

Variation in Transistor Performance and Leakage in Nanometer-Scale Technologies

Sharad Saxena, *Senior Member, IEEE*, Christopher Hess, *Senior Member, IEEE*, Hossein Karbasi, *Member, IEEE*, Angelo Rossoni, Stefano Tonello, Patrick McNamara, Silvia Lucherini, *Member, IEEE*, Seán Minehane, *Member, IEEE*, Christoph Dolainsky, and Michele Quarantelli, *Member, IEEE*

(Invited Paper)

Abstract—Variation in transistor characteristics is increasing as CMOS transistors are scaled to nanometer feature sizes. This increase in transistor variability poses a serious challenge to the cost-effective utilization of scaled technologies. Meeting this challenge requires comprehensive and efficient approaches for variability characterization, minimization, and mitigation. This paper describes an efficient infrastructure for characterizing the various types of variation in transistor characteristics. A sample of results obtained from applying this infrastructure to a number of technologies at the 90-, 65-, and 45-nm nodes is presented. This paper then illustrates the impact of the observed variability on SRAM, analog and digital circuit blocks used in system-on-chip designs. Different approaches for minimizing transistor variation and mitigating its impact on product performance and yield are also described.

Index Terms—Design for manufacturability (DFM), semiconductor device variation, tolerance analysis, yield estimation, yield optimization.

I. INTRODUCTION

VARIABILITY in transistor performance and leakage is a critical challenge to the continued scaling and effective utilization of CMOS technologies with nanometer-scale feature sizes. Minimizing transistor variation through technology optimization and ensuring robust product functionality and performance have always been the important concerns for integrated circuits [1]–[4]. However, the various technology enablers employed to obtain the benefits of device scaling at the 65- and 45-nm technology nodes make the problem of variability particularly severe.

Some of the factors contributing to the variability increase are fundamental to the planar CMOS transistor architecture. Random dopant fluctuations (RDFs) and line-edge roughness (LER) are two examples of such intrinsic sources of variation [5]. Other reasons for the variability increase are the advanced resolution-enhancement techniques (RETs) required to print patterns with feature sizes smaller than the wavelength of

lithography [6]. Additional phenomena, such as stress from shallow-trench isolation (STI), where the impact on device characteristics was negligible in previous technologies, have a noticeable effect on the transistor characteristics in advanced CMOS technologies [7]. New materials and performance-enhancement techniques add additional sources of variation; an example is the use of various types of local-stress techniques for mobility enhancement [8]. Supply-voltage reduction which often accompanies device scaling decreases the margins available to designers, making it difficult to provide sufficient guard bands to ensure functionality and performance for the expected amount of variation [9], [10]. Finally, the transition to 300-mm wafers increases the impact of across-wafer nonuniformity.

Optimizing the die-size and performance entitlement of scaled technologies requires a comprehensive approach for variability characterization, minimization, and mitigation. This paper begins in Section II with a classification of the sources of variation in nanometer-scale CMOS. Statistically accurate characterization of all these different sources of variation requires a large number of measurements on a wide variety of devices, layout styles, and environments. Efficient acquisition and analysis of this large volume of data measurements require new test structures, test hardware, test methods, and data analysis. Section III describes a transistor variability-characterization infrastructure. Section IV describes the results we have obtained from utilizing this infrastructure to characterize the different sources of systematic variation, in particular the performance variation that arises from transistor layout and other layout environment differences. Section V provides some of the results we have obtained from the characterization of random variation, which is the variation in transistors with identical dimensions, layout, and local layout neighborhood.

Transistor variation affects many aspects of IC manufacturing and design. Section VI describes some of these impacts. We first illustrate some new difficulties and challenges for technology development arising from increased variability. Then, we illustrate the impact of increased transistor variability on product performance and yield. The yield of high-density SRAM and the performance and power consumption of analog and digital blocks are all affected by the variation in transistor characteristics. This can result in lost revenue and expensive design respins. We show examples of the impact of variability

Manuscript received June 1, 2007; revised October 9, 2007. The review of this paper was arranged by Editor B. Zhao.

The authors are with PDF Solutions, Inc., San Jose CA 95110, USA (email: sharad.saxena@pdf.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TED.2007.911351

on each of these circuit types found in a typical system on chip (SoC).

Section VII discusses the problem of minimizing transistor variation and mitigating its impact. We begin with process and technology solutions for minimizing variation. Approaches such as RET for lithography, robust device designs and integration schemes, and manufacturing line control are illustrated. Next, we discuss a variety of techniques that can be termed as design for manufacturability (DFM) approaches. These approaches include design rules for variability reduction and the use of recommended design rules, compact modeling solutions to enable the consideration of random and systematic sources of variation during transistor-level circuit simulation and block-level timing simulation, design automation algorithms and tools for variation-aware design, and design methodologies to minimize the impact of variation.

II. SOURCES OF VARIATION

Variation in transistor characteristics can be characterized in many different ways, each of which being suitable for different objectives. A critical distinction is whether variations are present instantaneously during testing (time-zero variation) or whether they change the device behavior over time as the device is stressed under a bias. Such temporal variation affects product reliability, and it is due to the effects like channel hot carriers and negative-bias temperature instability. These reliability concerns are outside the scope of this paper; here, we focus on instantaneous variation.

One categorization of instantaneous variations divides it into intrinsic and extrinsic variations [4], [5]. Intrinsic variation is caused by fundamental phenomena associated with today's CMOS transistor architectures and manufacturing technologies. Examples of such variation are RDF and gate LER. Extrinsic variation arises from the lack of manufacturing control or the lack of fidelity in reproducing mask patterns onto the wafer. Common examples include transistor dimension and gate-oxide-thickness variation across chip, die to die, and wafer to wafer.

Another categorization, motivated by compact modeling, distinguishes among three types of variation: predictable, correlated, and uncorrelated [11]. Predictable variation is modeled by adding additional layout parameters to compact models to enable the calculation of the impact of layout effects like stress, orientation, and well proximity [11]–[13]. Uncorrelated variation refers to the component of variation where the information about the characteristics of one transistor does not provide any additional information about the characteristics of other transistors. The intrinsic sources of variation mentioned earlier result in uncorrelated variation. Uncorrelated variation is modeled by independent random variables with distributions that do not depend on layout or spatial parameters [14]. Correlated variation occurs when the characteristics of one transistor in a group restrict the distribution of other transistors in the group. Across-chip variation that depends on the distance between transistors is an example of this type of variation. Correlated variation is also modeled by random variables, but the distribution depends on the layout and other spatial parameters [11], [15].

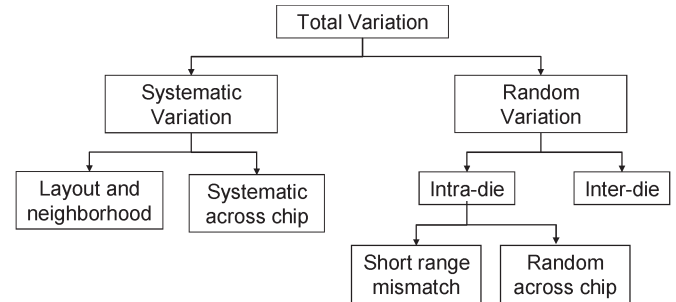


Fig. 1. Classification of the different types of variation in transistor characteristics.

Fig. 1 shows the classification of variation employed in this paper. We divide variation into two main categories: systematic and random variations. This classification is motivated by the differences in the root causes of the different types of variation.

Systematic variation is the difference in the electrical characteristics of two transistors with identical width and length (W/L), where there is a clearly identifiable difference in either the device layout or the layout neighborhood. Examples include the impact of gate poly pitch on gate length due to optical proximity effects, stress effects, orientation effects, etc. Random variation is the difference in the electrical characteristics of devices with identical geometry (W/L), layout, and neighborhood within the interaction distance of the known sources of variation. For example, the impact of gate poly-silicon pitch differences becomes negligible for pitches greater than $3\lambda/NA$ [16], where λ is the wavelength, and NA is the numerical aperture of the exposure system. For $\lambda = 193$ nm and $NA = 0.75$, these give an interaction distance of approximately 775 nm. This forms a lower bound on the distance over which the local neighborhood is required to be identical for the variation to be considered random. Recently, a new source of variation with millimeter-scale interaction distance has been observed: rapid thermal anneal (RTA) temperature variation due to the density of STI regions not covered by gate poly [11], [17]. Such extreme long interaction-distance effects are considered to be random variation in this paper.

Each of the main categories can be further divided. Classification of random variation into intradie (within-die) and interdie variations (between different dies) helps in identifying the root causes and possible improvement actions. For example, intradie variation can be addressed by lithography improvements, such as off-axis illumination [6] and within-field-exposure dose compensation [18]. Across-wafer nonuniformity's contribution to interdie variation can be improved by techniques like exposure dose compensation across the wafer [18], ion-implantation uniformity, RTA uniformity, etc. Systematic variation can be divided into variation arising from layout or neighborhood differences within an interaction distance.

III. VARIABILITY-CHARACTERIZATION INFRASTRUCTURE

Accurate and efficient characterization of the different types of variation places new requirements on the device characterization infrastructure. Statistically significant variability characterization requires large measurement samples, covering all

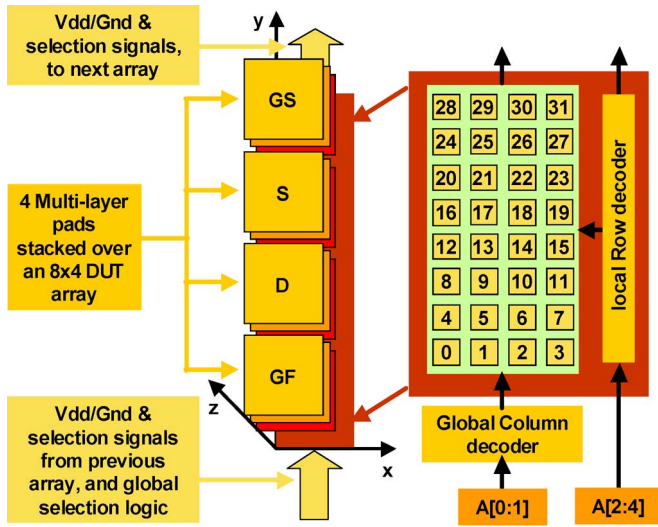


Fig. 2. Concept of arranging 32 devices (right) under probe pads (left).

major sources of variation. Test structures, test chips, test hardware and methodology, and the analysis of test results all need to be optimized to enable accurate and fast measurements and analysis. This section describes our infrastructure for transistor variability characterization.

A. Test Structures and Test Chips

We have developed two types of test chips for variability characterization. The first is included as part of a full reticle characterization vehicle (CV) test chip for characterizing the impact of front end of the line processing on yield and transistor performance and variability. This test chip contains test structures for statistically characterizing the impact of the main types of variation described in Section II. The second test chip, called the device scribe (dScribe CV), contains a much smaller set of test structures suitable for monitoring transistor variation and its impact in mass production [19]. These test structures are placed in the scribe line of product wafers.

A novel addressable array test structure was developed to improve the packing and pad efficiency of the large number of test structures required for variability characterization. In addition, fast parallel parametric test hardware (pdFasTestII) was developed and optimized to make use of the device arrays and to minimize test time. This parametric fast test hardware is described in the next section. The device array and the fast parametric test hardware together address the challenge of efficiently measuring the data required for statistically significant variability characterization.

1) *Device-Array Test Structure*: Addressable arrays have been employed in the past for improving area and pad efficiency for transistor characterization [20]–[23]. Our array architecture is optimized for area and fast parametric testing. The device array packs 32 devices within an array, where all routing requires only two metal layers. This improves silicon efficiency by allowing arrays to be placed underneath the pads, which are implemented in the third and higher levels of metal. The left side of Fig. 2 shows this approach.

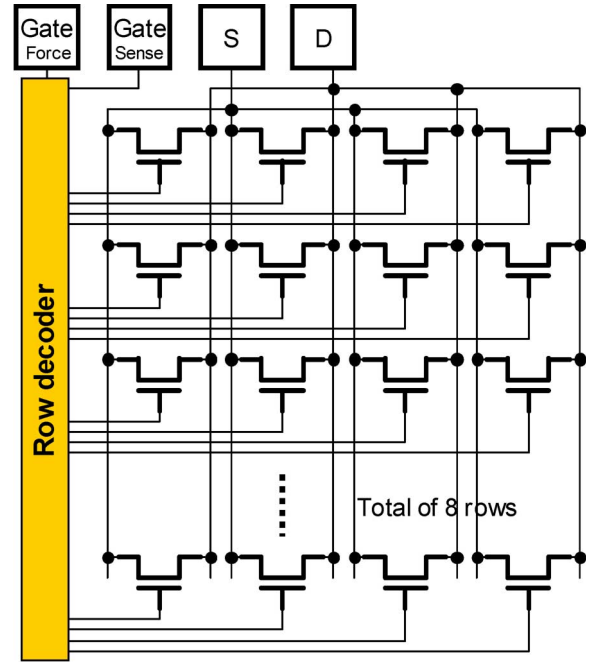


Fig. 3. Schematic of one device array.

All 32 devices share the source and drain pins. The drain and source paths do not have any selection devices, which improves measurement accuracy. It also eliminates the risk of back-biasing the wells, which can happen if a selection device is implemented along the pad-to-source path. Balanced routing limits the resistance between each pad and the source or drain pins of each device under test (DUT) to less than 5Ω . The only selection device is in the gate path in order to turn on a specific DUT. For the gate, a force and sense path is used to monitor and adjust the gate voltage if required. The selection logic, shown in Fig. 3, is purely combinatorial, and it is shared among several arrays. The absence of timing critical circuitry reduces design time, simplifies testing, and ensures a robust design across a variety of technologies in different stages of maturity.

The sensitivity of array test structures to distinguish systematic variation from random variation is increased by placing multiple replicates of identical transistor layouts in an array and, whenever necessary, placing multiple arrays close to each other on a test chip. The numbers of replicates and arrays are determined by the resolution required to detect the systematic effect of interest. For example, say that we want to detect with 95% confidence a difference in drive current of $\Delta\mu$ between two layouts. Further assume that the random variation is the same for both the layouts and that it is normally distributed with variance σ^2 . From the equation for 95% confidence interval for the mean of a normal distribution with variance σ^2 , we have $\Delta\mu = 1.96 \times (\sigma/\sqrt{n})$, where n is the number of replicates [24]. Fig. 4 shows the I_{drive} difference that can be detected with different number of replicates for $\sigma = 10\%$. As seen in this figure, a single array (32 replicates) is sufficient for detecting systematic differences in I_{drive} of less in 5% in presence of random variation with $\sigma = 10\%$.

2) *Test-Chip Contents*: The test structure set contains four categories of structures: structures for characterizing random

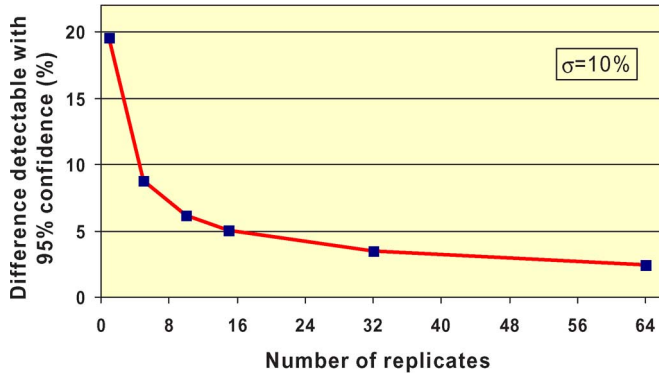


Fig. 4. Difference due to systematic effects detectable with different numbers of replicates.

TABLE I
LAYOUT AND SYSTEMATIC SOURCES OF VARIATION COVERED BY SPECIAL TEST STRUCTURES (PARTIAL LIST)

Systematic variation	Root causes
Gate poly orientation	Poly lithography, stress layers
Gate poly pitches	Poly OPC, lithography & stress
Poly corner rounding	Poly OPC and lithography
Transistor location in multi-gate transistor	Poly patterning impacted by local neighborhood differences
Active corner rounding	Active OPC and lithography
Un-modeled narrow width effects	Stress effects, poly step height
Gate to active edge	STI stress, e-SiGe stress
Well proximity	Implant scattering
Nwell-Pwell separation	Gate counterdoping and misalignment
Contact density and placement	Silicide sheet resistance and stress

and systematic variations, across-chip variability (ACV) characterization, leakage characterization, and support structures for decomposing the sources of variations.

Table I shows the main layout and environment effects covered in the systematic-variation component of the test chip. This set of structures covers most of the sources of layout-driven variation discussed in the literature, such as proximity effects impacting lithography [25], stress effects [7], [8], well-proximity effects [26], and poly-step-height effects [27], [28]. It also includes additional effects that have not been discussed much in the literature, such as counterdoping due to N+/P+ spacing and active corner rounding. Test structures are designed to be sensitive to each of these effects, with the layout being optimized to minimize the impact of other layout effects. Fig. 5 shows the examples of some of these structures. Fig. 5(a) is for characterizing poly-pitch effects and the impact of poly end cap on transistor performance and leakage, Fig. 5(b) is for characterizing STI stress effects, and Fig. 5(c) shows a test structure for characterizing the impact of poly-silicon corner rounding and the effectiveness of optical proximity correction (OPC) in overcoming such corner rounding. Similar test patterns were designed for each of the layout effects. A design of experiments (DOE), varying the different layout parameters of interest, is performed to cover the range of layout variation found in typical designs. A subset of the systematic variation patterns, particu-

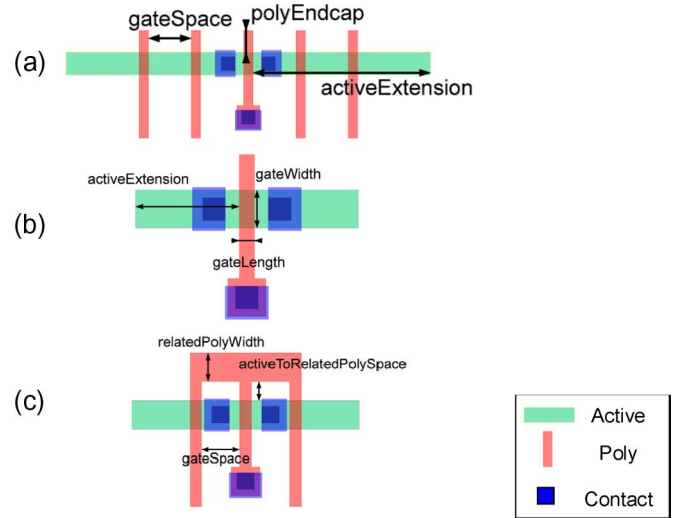


Fig. 5. Example layout patterns for characterizing layout effects: (a) Poly-pitch-induced variation due to the impact of lithography and stress, (b) active-extension-induced variation due to STI stress, and (c) gate-poly-shape-induced variation due to corner rounding.

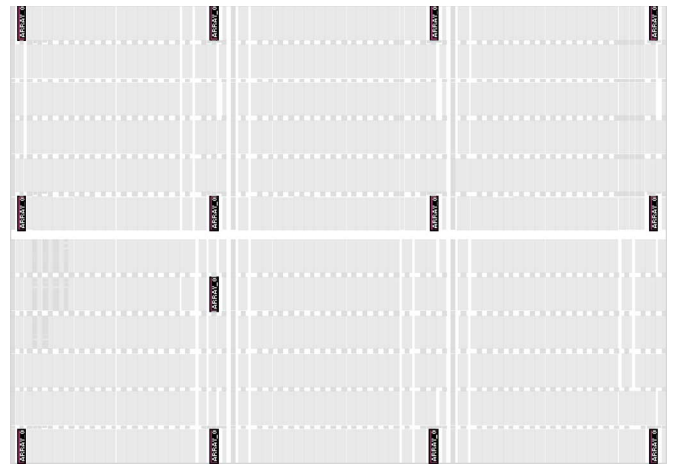


Fig. 6. Typical placement of ACV structures.

larly those patterns sensitive to patterning variation, is selected for ACV evaluation. This is accomplished by placing them in device arrays and replicating these arrays multiple times across the reticle. In addition to transistor structures, a set of support structures is also included in the ACV structure set. These support structures help quantify the root causes of variation. For example, the test structures for electrical measurement of gate-length variation are included in the across-chip characterization module. The actual number of across-chip placements depends on the size of the reticle, the available area, and the tradeoff with other structures.

Device arrays are particularly suited for across-chip variability evaluation because they allow the measurement of many devices in a single test touchdown. Fig. 6 shows a typical placement employed for ACV evaluation. The efficiency of device arrays and fast parametric testing becomes even more significant when denser placements of across-field test structures are employed, such as the 5 × 5 placements reported in the literature [29], [30].

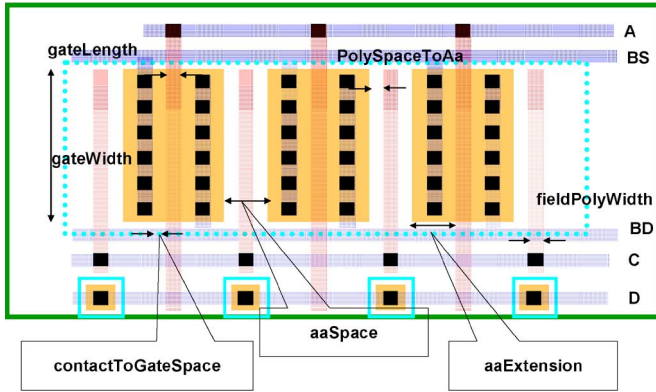


Fig. 7. Schematic layout and some design attributes of the test structure used for leakage characterization.

Accurate characterization of short-distance mismatch is essential for SRAM, analog, and RF design [10], [31]. Device arrays with identical devices are a commonly employed test structure for mismatch characterization [32]. We have employed both the device-array architecture described previously and the larger arrays in order to increase the sample size and, therefore, the statistical significance of mismatch measurements [22], [23].

Leakage and leakage-variability characterization have additional requirements. The measurement infrastructure should be sensitive to the various sources of leakage. This includes both parametric sources like gate-length variation, which causes variation in OFF-state current (I_{off}), and defect sources like silicide stringers and pipe defects, which result from the lateral diffusion of silicide under the spacer and into the channel. Low current levels expected in leakage measurement require that the test structures and other parts of the test methodology do not introduce measurement noise.

We make use of a large nonaddressable array of transistors connected in parallel for leakage characterization. As with the test structures used for transistor-performance characterization, a DOE is performed on the layout of the transistors in these arrays to evaluate the leakage variability introduced by the layout. Fig. 7 shows a schematic of this test structure. Different biasing and measurement schemes make this structure sensitive to different leakage mechanisms. For example, in an NMOS transistor, by biasing the drain to V_{dd} and all other nodes to ground, the current on the gate terminal is the sum of the current from silicide stringers and silicide encroachment of the spacer and the gate current in the overlap region (I_{govl}). I_{govl} contribution is minimized by making a subset of the DOE on thick gate oxides, typically available in most modern CMOS processes. Similarly, NMOS subthreshold current ($I_{subthes}$) is measured by biasing drain to V_{dd} , grounding all other terminals, and measuring current at the source. Changing the bias node, bias voltage, and measurement node of this structure allows measurement of different components of leakage.

Leakage arrays are effective at characterizing the contribution of various sources of leakage to average leakage and leakage variability. The large number of parallel structures provides high enough current to avoid measurement-induced noise.

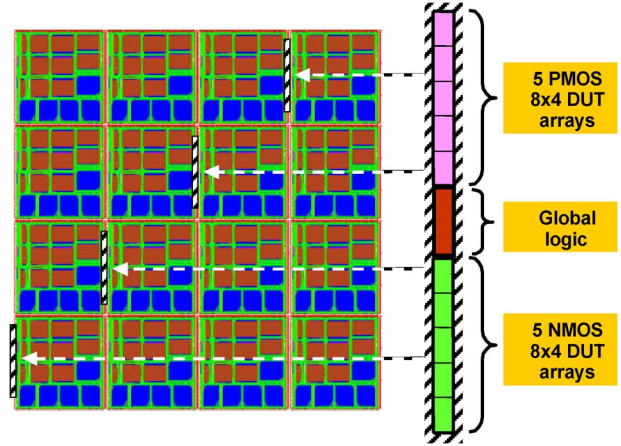


Fig. 8. Packing of arrays into one Scribe CV test chip (right) and potential placement on a multiple product reticle (left).

In addition, they provide enough critical area for sufficient observability of leakage events that are triggered by defectlike mechanisms mentioned earlier.

3) *Device Scribe*: Lot-to-lot variation is an important component of total variation. Traditional approaches for monitoring lot-to-lot variation make use of process-control monitor (PCM) structures. Test-time constraints limit the number of PCM measurements to a few sites on a wafer; five to nine sites measured on a few wafers in a lot is common. These measurements are also used to evaluate whether a lot is within process-control limits and are sometimes referred to as wafer acceptance tests. Wafer averages evaluated from a limited number of sites do not provide variability as a metric for acceptance. For 300-mm wafers, across-wafer nonuniformity is a significant contributor to the total variation, as will be illustrated in Section V. This makes sparse sampling of PCM ineffective as variability monitoring in mass production. A special configuration of device arrays, used in conjunction with fast parametric test, solves this problem by allowing measurement of transistor characteristics on a large number of sites. Potentially even die on every wafer can be measured in production [19]. This combination of device arrays and fast parametric test has been applied so far to measure up to 150 sites per wafer in production.

Fig. 8 shows one implementation of the Device Scribe CV test chip. Ten device arrays are packed in one pad group to be tested in parallel. The Device Scribe CV is then placed in one or more locations next to product chips, as shown on the left side of Fig. 8. Each array is placed under four pads.

B. Fast Parametric Test

Test time is a critical bottleneck for variability characterization in nanometer-scale technologies. Statistically accurate and complete characterization requires measurement of a large sample of devices for a variety of device types, layouts, and supporting structures. We have developed a fast parametric test system to address this problem.

The parallel pdFasTestII parametric test system provides up to 72 parallel analog testing channels. The ten arrays of the dScribe CV are tested in parallel. Within each array, the

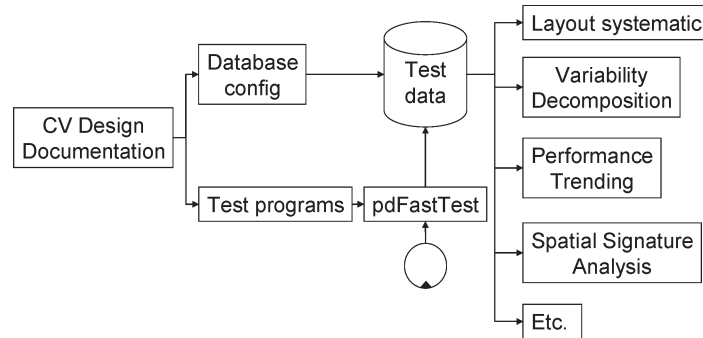


Fig. 9. Architecture of the tools developed for the rapid analysis of large quantities of measurements.

32 devices are tested in sequence. The pdFasTestII measures two I/V curves for each of the 320 devices: one for the linear region, and one for the saturation region. From these curves, the tester derives key device parameters like V_{tlin} , I_{dlin} , V_{tsat} , I_{dsat} , G_{mlin} , and G_{msat} . This configuration of pdFasTestII test hardware has provided 60 times speed gain over a regular parametric tester to perform the same task. Despite the large speed advantage, the accuracy in measuring the aforementioned key indicators of transistor performance is not compromised. Device-array measurements with pdFasTestII were found to be in good agreement with measurements of individually padded PCM transistors tested with standard parametric test [19].

C. Analysis Tools

The test chips and fast parametric test infrastructure described in the previous sections typically produce multiple gigabytes of measurements. Efficient software tools are needed to rapidly summarize and analyze these data. Fig. 9 shows the architecture of analysis tools we have developed to perform such analysis utilizing the capabilities of dataPOWER yield management software [33].

Configuration files, database schema, and test programs are automatically generated from the test-chip design documentation. The test programs drive the fast test hardware, and the measurements are loaded into a database according to the schema and analysis configuration specifications. A number of software tools access this database and perform various analysis functions: analysis of layout systematics, variance decomposition, trending, spatial signature analysis, lot equipment history, etc. Detailed description of these applications is outside the scope of this paper. The next few sections illustrate the output from these applications, showing the application of the variability-characterization infrastructure.

IV. CHARACTERIZING SYSTEMATIC VARIATION

Unmodeled layout and environment effects change device characteristics and appear as variation in device performance. While the exact amount of variation differs between technologies, we have found that the key phenomena are similar across many different technologies, and often, the amount of variation is similar.

TABLE II
TYPICAL LAYOUT EFFECTS CAUSING VARIATION IN
NANOMETER-SCALE DEVICES

Layout effect	Typical impact for 65 nm technologies
Poly pitch: printability	3-5% change between weak pitches and reference pitch
Poly pitch: stress	5-10% change in I_{drive}
Poly orientation	2%-7% change in I_{drive} depending on poly pitch. Larger impact for weak pitches
Poly local neighborhood; e.g. center vs. edge gate in a multi-gate layout	Depending on OPC and dummy poly schemes, 1-10% difference between center and edge gates
Poly corner rounding	2-7% decrease in I_{drive} for worst case spacing
STI Stress: Variation between minimum active extension to maximum active extension	PMOS I_{drive} : 5-8% NMOS I_{drive} : 12-18%
Active corner rounding	1-5% I_{drive} increase for worst case spacing
Gate counter-doping	6-10% decrease in PMOS I_{drive}
Contact density	3-5% I_{drive} decrease between dense and spare contacts

Table II shows the impact of some of the main layout effects that change the device characteristics. The term weak pitch in this table refers to the poly pitch with the smallest lithography process window. The impact shown in this table is typical over a number of 65-nm technologies and is not specific to any particular technology. The measurements for this table were taken on a number of 65-nm technologies in early manufacturing. All the measurements were made on design-rule clean test structures of the type shown in Fig. 5. Therefore, these systematic sources of variation can be expected to cause variation in product designs if they are not modeled adequately.

For example, Fig. 10 shows the normalized change in I_{drive} for different gate pitches in a 65-nm technology. Fig. 11 shows a different impact of poly pitch in technologies that employ dual stress layers for performance improvement of both NMOS and PMOS transistors. The plot shows the shift in $I_{dsat}-V_{tsat}$ characteristics between the tightest contacted pitch and one of the loose pitches available in this technology. This shift is an indication of the mobility difference. This figure shows that the stress layer and the poly pitch interact to provide the mobility enhancement obtained for a particular transistor. Tight pitches

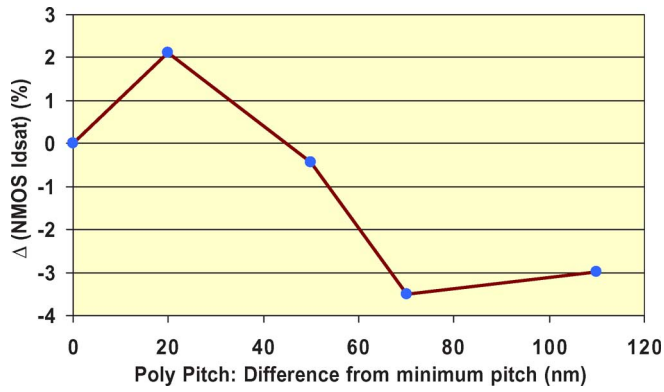


Fig. 10. Poly-pitch dependence of Idrive for minimum gate-length NMOS transistors.

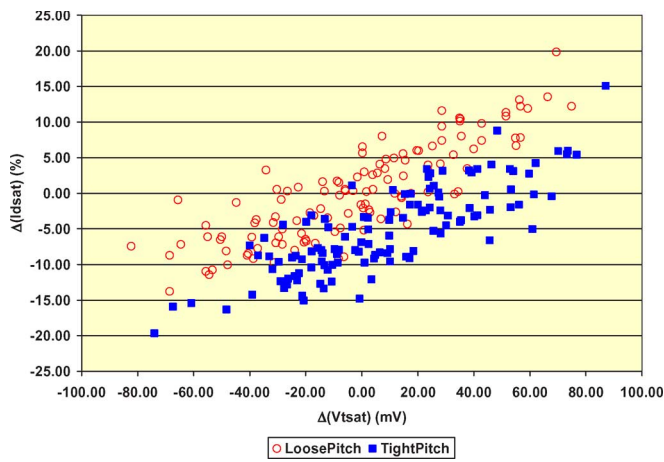


Fig. 11. Idrive variation for minimum gate-length PMOS transistors caused by the interaction of poly pitch and stress layer.

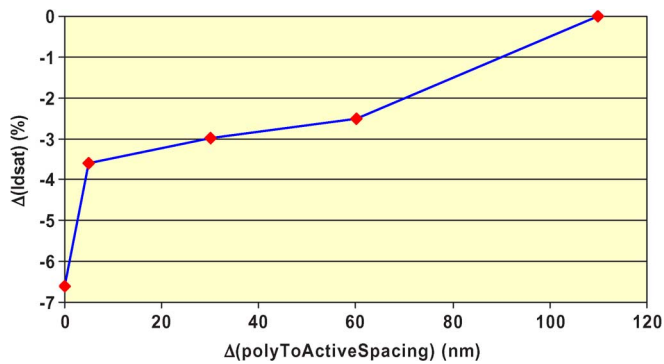


Fig. 12. Impact of poly corner rounding on W_{min}/L_{min} NMOS transistor Idrive. Corner rounding depends on poly to active spacing [Fig. 5(c)]. Idrive difference is calculated with respect to the largest spacing in the DOE.

can result in a reduction of mobility improvement obtained from stress layers because a smaller volume of stressor film is present in the proximity of the transistor [34]. Fig. 12 shows the impact of poly corner rounding on NMOS Idrive. Fig. 13 shows the impact of active corner rounding on the transistor OFF-state current (I_{off}). In addition to the effective width increase due to corner rounding, we have observed increased likelihood of high-leakage transistors due to active corners.

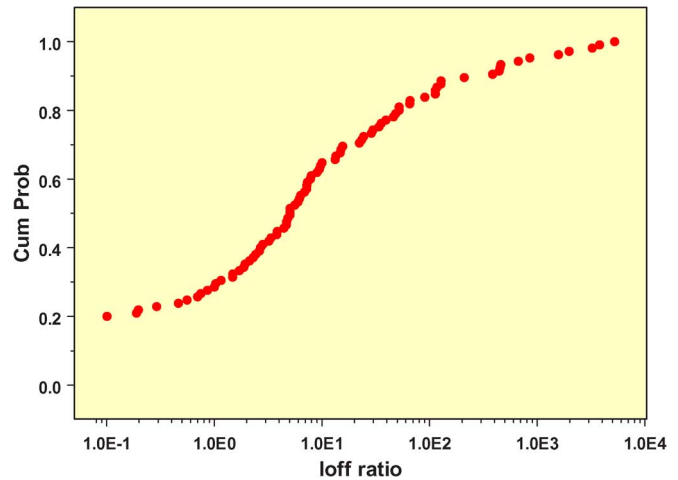


Fig. 13. Cumulative probability distribution of Ioff ratio for W_{min}/L_{min} NMOS transistor with active corner rounding. Ioff ratio is the ratio of Ioff for a transistor layout without the possibility of corner rounding to a transistor with a layout that has potential for corner rounding.

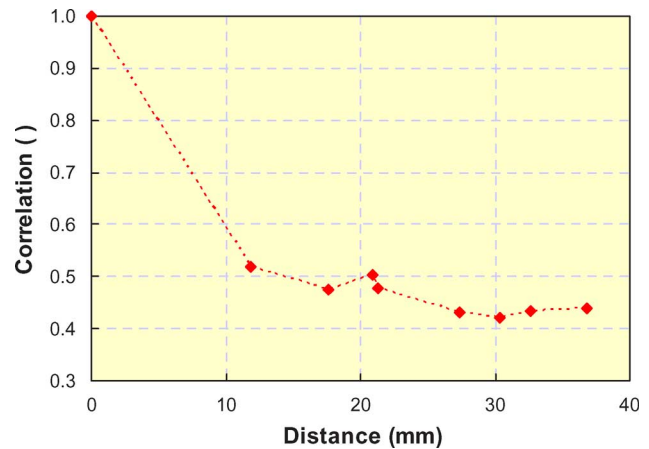


Fig. 14. Vtsat correlation coefficient as a function of distance between the transistors for a W_{min}/L_{min} NMOS transistor in a 65-nm technology.

Another type of systematic variation is the distance-dependent component of ACV. It is considered systematic because the amount of variability is a function of the distance between the transistors. Fig. 14 shows a typical dependence of the linear correlation coefficient of Vtsat of identical transistors on the distance between the transistors. Over a range of distances, the correlation coefficient is found to monotonically decrease with distance, followed by an approximately constant correlation coefficient. This behavior allows modeling of the distance-dependent component of ACV using the concept of interaction distance. This is the distance over which correlation is a linear or a quadratic function of distance, followed by constant correlation. Similar behavior has also been reported in [15] for across-chip gate-length variation with similar interaction distances, as shown in Fig. 14, and for threshold-voltage variation in [35] for large geometry devices but not for small geometry devices, suggesting that the exact nature of the distance-dependent component depends on the process and its maturity.

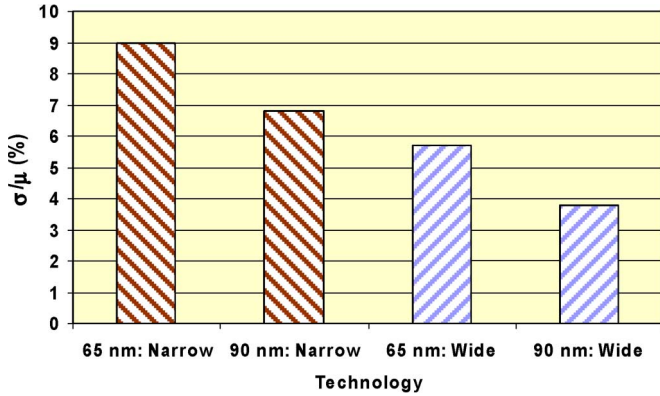


Fig. 15. Relative standard deviation of I_{drive} for L_{min} NMOS transistors over multiple generations. Wide transistors have $W = 1 \mu m$; narrow transistors are the minimum width logic transistors allowed in the technology.

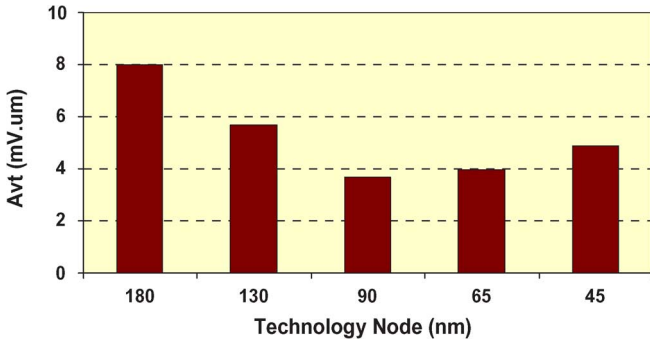


Fig. 16. NMOS local mismatch trend over multiple generations.

V. CHARACTERIZING RANDOM VARIATION

This paper defines random variation as the variation in the characteristics of a population of transistors with identical layout and environment. Random variation arises from extrinsic causes like manufacturing control and from intrinsic causes like RDF and LER. This section describes some of the random-variation trends we have observed by utilizing our variability-characterization infrastructure. As was the case for systematic variation in the previous section, the variability trends are typical for a number of technologies and not specific to any particular technology. The measurements corresponding to 65-nm variability were made from technologies in early manufacturing.

Fig. 15 shows a typical trend in relative standard deviation of I_{drive} with technology scaling that we have observed. The variability of both narrow and wide devices increases with scaling, suggesting an increase in both extrinsic and intrinsic components of random variation.

Fig. 16 shows our measurements of NMOS mismatch over multiple technology generations. In this plot, mismatch is quantified as the slope of the area-dependent term in $\sigma(\Delta(V_t)) = (A_{vt}/\sqrt{WL}) + k$ [31], where L and W are the drawn length and width of the gate, respectively. This figure shows that, while A_{vt} has the expected reduction with T_{ox} scaling until about 90-nm technologies [36], in recent technology generations, A_{vt} has not been decreasing. This could be due to the reduction in T_{ox}

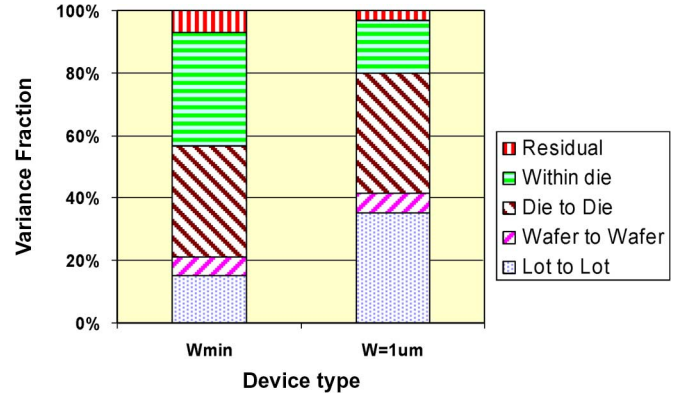


Fig. 17. Variance decomposition of I_{drive} for L_{min} NMOS transistors in a typical 65-nm technology. The Y-axis shows the fraction of the total variance that can be attributed to a particular source. Residual denotes the variance fraction not explained by the variance decomposition model.

scaling, the increase in channel doping required to reduce short-channel effects, and the contribution of additional process steps like gate predoping to V_t variation.

Fig. 17 shows a decomposition of total I_{drive} variation of NMOS transistors into lot-to-lot, wafer-to-wafer, and across-wafer components using the standard variance-decomposition methods [37]. We have found across-wafer nonuniformity, including across-chip variation, to be the largest source of variation over multiple nanometer-scale technologies. The next largest component tends to be lot-to-lot variation. Wafer-to-wafer variation within a lot is often small, indicating the good wafer-to-wafer control obtained by single-wafer processing tools employed in nanometer-scale fabrication.

VI. IMPACT OF VARIATION

The increase in transistor variability affects many aspects of technology development, manufacturing, and design. This section describes some of these impacts. Section VI-A provides an example of the impact of increased variability on a common technology-development methodology: the design and use of split lots for technology optimization. Sections VI-B–VI-E illustrate the impact of increased variability on each of the major circuit blocks in state-of-the-art SoCs.

A. Impact on Technology Development

Variability increase poses many difficult challenges for technology development. The most critical challenge is the development and adoption of variability-reduction methods. These range from techniques for robust transistor architecture and process integration options to the evaluation of process-control improvements [38]. In addition to these technological challenges, the increase in variability necessitates modifications and improvements in the common technology-development methodologies.

One such methodology is the use of split lots for technology optimization. In this approach, the technology is optimized for the target application through a series of DOE, where the goal of each experiment is to obtain incremental improvement in the

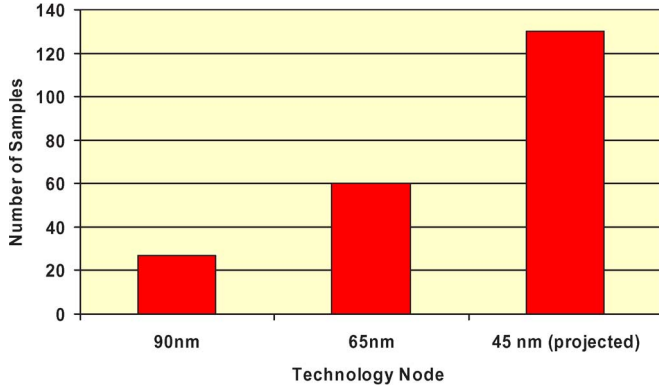


Fig. 18. Sample size required to detect 2.5% Idrive difference with 95% confidence. A 45-nm variability increase is projected based on trends from 90 and 65 nm.

technology characteristic of interest. For example, suppose that the purpose of the experiment is to improve transistor Idrive at a given Ioff by 2.5%. The Idrive measurements for each of the splits will typically be distributed normally around a mean value μ . For normally distributed measurements, the number of samples required to identify with 95% confidence that the two splits have a difference in the mean of $\Delta\mu$ is given by [24]

$$\mu \pm \Delta\mu = \mu \pm 1.96 \times \frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{1.96 \times \sigma}{\Delta\mu} \right)^2.$$

Fig. 18 shows the increase in sample size required to detect 2.5% change in Idrive as technology scales, assuming that the mean Idrive remains the same for each generation. This increase in sample size, resulting in an increase in test budgets and test cycle times, increases the difficulty of technology development. Additional challenges arise from spatial distributions and lot-to-lot variation, which result in a different population of devices in identical splits.

B. Impact on SRAM

High-density SRAMs are an essential component of most complex ICs, and they are a key technical and business differentiating factor between competing technologies. This competition for the highest density SRAM has resulted in an industry-wide challenge at the 45-nm technology node to produce highly manufacturable SRAM using a core bit cell with less than $0.3\text{-}\mu\text{m}^2$ area. The small feature sizes of SRAM devices, the tight timing constraints, and the large size of the memory blocks make them particularly sensitive to variation in transistor performance and leakage. All key SRAM performance metrics: read fails, write fails, retention fails, static-noise margins, etc., are sensitive to transistor variation. This section illustrates the impact of variation on one such metric: the probability of read fails.

A read failure happens when sufficiently large voltage separation does not develop between bit line (BIT) and bit line complement (/BIT) lines of a SRAM in the time interval when the read enable (REN) is active. As a result, the sense

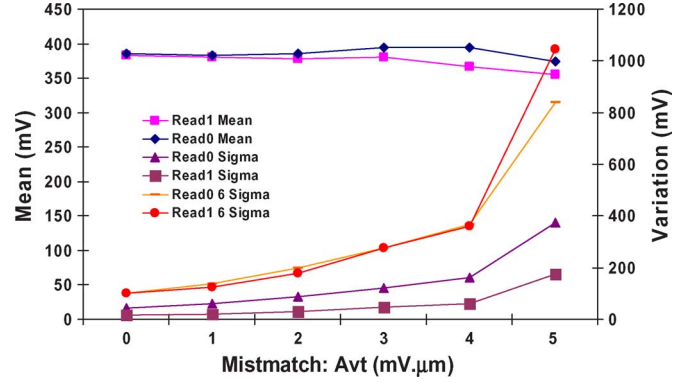


Fig. 19. Simulation of the impact of V_t mismatch on the variation in BIT /BIT voltage separation in high-density SRAM.

amp is unable to latch the state of the bit. The need for fast memory access reduces the time allowed for REN to remain active, which in turn requires that a sufficient voltage difference between BIT and /BIT develops quickly. Mismatch between the transistors in the SRAM bit cell causes variation in the BIT /BIT voltage separation in a fixed REN time period. A read failure occurs when the voltage separation is not sufficient.

Fig. 19 shows the simulated variation in BIT to /BIT voltage separation as a function of mismatch for a 45-nm SRAM bit cell based on typical sizing found in the literature [39], [40]. Mismatch is quantified by A_{vt} for the technology. This figure shows that the mean value of BIT /BIT separation is approximately 350–400 mV. However, as mismatch increases to values observed in 45-nm technologies (Fig. 16), the standard deviation (σ) of the BIT /BIT separation approaches 100 mV. Multimegabit SRAMs require robustness greater than $\pm 6\text{-}\sigma$ variation. Thus, we find that, with $6\text{-}\sigma$ variation, the BIT /BIT voltage separation approaches the range of voltage supply (Vdd) of 1.0–1.2 V, which increases the probability of read failures and the sensitivity to supply voltage during minimum Vdd tests. This has resulted in the search for design solutions, such as the assist circuits to overcome the single-bit fails associated with increased transistor variability [41].

C. Impact on Power Consumption

Fig. 19 shows another consequence of increasing variability; it limits the allowable Vdd reduction. As $6\text{-}\sigma$ variation approaches Vdd, the option of Vdd reduction for power consumption becomes unavailable. Fig. 20 shows the International Technology Roadmap for Semiconductors values for Vdd for low standby-current technologies from 180 to 45 nm [42]. In addition to the need to maintain high threshold voltage for leakage minimization, the increase in variability also contributes to the reduced pace of Vdd reduction.

An increase in transistor variability also increases leakage at a fixed Vdd. Typically, Ioff has a log-normal distribution. That is, $\log(\text{Ioff}) \sim N(\mu, \sigma^2)$, where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 . The log-normal distribution for Ioff implies that $E(\text{Ioff}) = \exp(\mu + (\sigma^2/2))$, where $\exp(\cdot)$ denotes the exponential function, and $E(\cdot)$ denotes the expected value. The increase in $E(\text{Ioff})$ due to Ioff variability increases the leakage power of an IC.

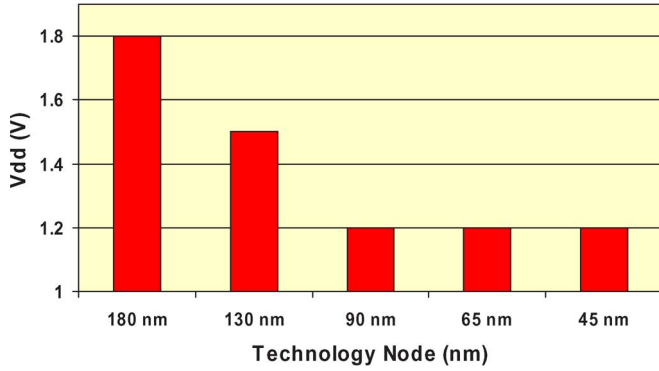


Fig. 20. Supply voltage (Vdd) trend from 180- to 45-nm technologies.

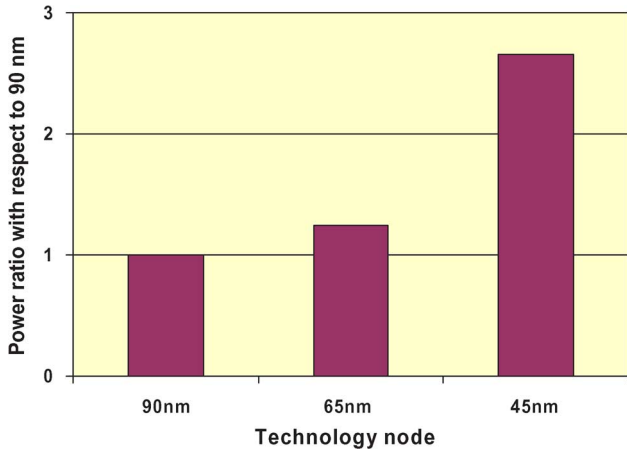


Fig. 21. Estimated power increase to maintain same bandwidth and accuracy product as 90 nm.

D. Impact on Analog and RF Circuits

The impact of variability, particularly mismatch, on analog and RF circuits is well known and extensively studied [10], [31], [43]. High-resolution data converters are particularly sensitive to variation and mismatch because they resolve small changes in either input or output voltages. For example, Kinget [43] derives the following relationship for the tradeoff between transistor mismatch, accuracy, and power consumption for a class of circuits that includes flash analog-to-digital converters (ADC)

$$\frac{BW_{Acc_{DC}}}{P} \leq \frac{K}{C_{ox}A_{vt}^2}$$

where BW denotes the bandwidth, which determines the maximum input/output frequency or speed, Acc_{DC} denotes the accuracy of the ADC, P denotes the power consumption, K is a constant depending on the circuit, and C_{ox} is the gate capacitance.

The increase in power consumption required to maintain the same bandwidth-accuracy product with scaling can be estimated from this tradeoff relationship, typical oxide thickness for the various technology nodes, and the threshold-voltage mismatch obtained from the scaling trend shown in Fig. 16. This increase in power consumption is shown in Fig. 21. This figure shows that maintaining the same accuracy and speed at a

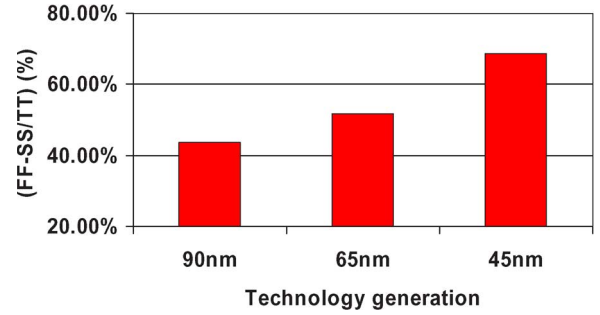


Fig. 22. Spread in Idrive between SS and FF corners for multiple generations of low standby-power technologies.

45-nm technology node will consume approximately 2.5 times the power as a 90-nm node. Conversely, to maintain or reduce the power consumption will require compromising either the speed or the accuracy.

One straightforward method of reducing the impact of mismatch is by increasing the size of the critical transistors. The optimal sizing can be determined by statistical circuit simulation [44]. However, in addition to the increased power consumption mentioned previously, this approach results in a less competitive design because the area of the analog and RF components of a SoC does not scale as fast as the digital components with technology generation. This limits the number of dies per wafer that can be obtained for products with significant analog content, reducing the economic benefits of scaling.

E. Impact on Digital Timing

The impact of the increase in transistor variation on the delay, maximum clock frequency (FMAX), and leakage of digital designs has been reported in the literature [4], [45]. Here, we illustrate the impact of variation on design margins. Digital designers cope with the variation in transistor performance through the use of design margins or guard banding obtained from worst case corner models. However, as the contribution of the additional sources of variation described in Section II increases, the margins required for worst case design become very large. Ensuring a desired performance with these increased margins limits the area and leakage scaling that can be obtained from technology scaling, again limiting the economic benefit of scaling.

Fig. 22 shows a typical trend of Idrive spread between slow-slow (SS) and fast-fast (FF) corners that we have observed for the minimum gate-length and gate-width transistors for multiple technology generations. The spread appears to be increasing every technology generation, requiring an increase in the design margins. This increase in spread is consistent with the variability increase shown in Fig. 15 if we take the corners to represent approximately $3\text{-}\sigma$ variation. Inclusion of systematic effects in the worst case corners would increase the spread even more.

VII. MINIMIZING THE IMPACT OF VARIATION

A wide range of process and design solutions are being explored to address the challenges posed by the increase

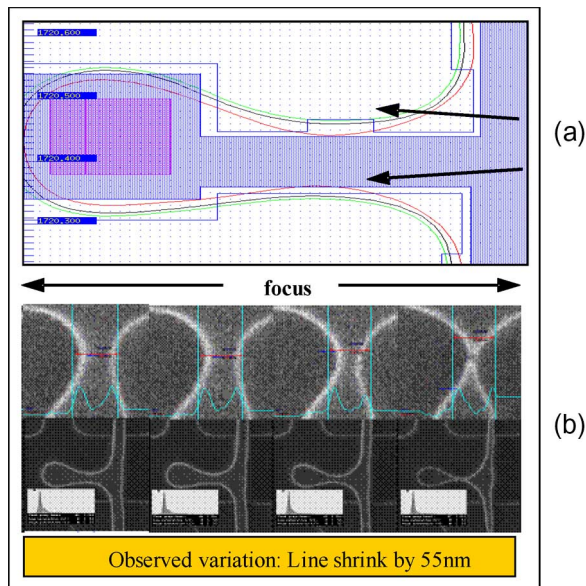


Fig. 23. (a) Identification of sensitive layout patterns using lithography simulation including manufacturing variation. The different lines indicate simulated pattern under nominal lithography conditions, the worst case grow due to focus variation, and the worst case shrink due to focus variation. (b) SEM from an experiment to verify the sensitivity of the pattern to focus variation.

in transistor variability to the cost-effective utilization of nanometer-scale CMOS technology. This section illustrates some of the main approaches.

A. Process Solutions

Statistical process control has long been the main approach for minimizing manufacturing variation. New process-control techniques like model-based, run-to-run, and feedforward controls are being developed and deployed to achieve the tight tolerances required for state-of-the-art processes [38]. For example, gate critical-dimension control by exposure dose compensation in scanners for improving across-wafer and across-field nonuniformity is being evaluated and employed in mass production [18]. Additional techniques to reduce variability due to lithography include immersion lithography, polarized illumination, and double exposure [46]. Laser annealing also holds the promise of reducing some systematic sources of variation, e.g., variation in transistor characteristics due to local changes in active density [47].

OPC robustness is an important avenue for reducing the variability induced by patterning. One approach is to perform calibrated lithography simulations that include the impact of main sources of variation in lithography, such as focus, exposure, mask error factor, and misalignment [48], [49]. These simulations can be performed at a full-chip level to identify critical spots in the layout, where the deviation from target dimensions over the expected range of variation is large, or on test chips to identify sensitive layout patterns where the OPC can be further optimized. Fig. 23 shows the application of this approach to find a layout pattern sensitive to focus variation, which causes large gate-length variation. The ability to identify patterns sensitive to the worst case manufacturing variation makes it possible to find and correct OPC weaknesses before product manufacturing.

SEM images from a focus-variation experiment confirm the sensitivity of this pattern to focus variation.

Another area of active research is the adoption of device designs and process integration options aimed at reducing transistor variability. Attempts range from process integration options for standard CMOS device architectures to the use of new device architectures. The use of amorphous silicon as the gate electrode to reduce variability arising from the interaction of pocket and extension implants with the gate edge [50] and the formation of ultrashallow junctions using coimplantation of pocket implant with carbon or fluorine and laser anneals [51] are examples of the first category of techniques. Proposals for device architectures that employ low- and delta-doped channels [52], thin-body silicon-on-insulator [53], and FinFETs are examples of new device architectures with the potential for reducing device variability [54].

B. Design for Manufacturability Solutions

A variety of design solutions ranging from physical design approaches to new architectures for common design blocks are being attempted and adopted. These techniques form a part of DFM approaches.

The use of advanced OPC techniques like subresolution assist features is common at 65- and 45-nm nodes [6]. Another approach is the use of recommended design rules. These rules are relaxed in comparison to minimum design rules but are recommended as a means of reducing variability. They are to be used in addition to good design practices like the use of nonminimum area transistors for mismatch control and having redundant contacts on wide transistors to reduce the impact of IR drop. Recommended design rules typically target systematic sources of variation. For example, each systematic source of variation listed in Table II is a candidate for a recommended design rule. Common examples of recommended design rules are single poly orientation for critical transistors, recommended gate-poly pitches for gate-length control, recommended poly to active spacing for poly flaring control, etc.

Improvement in device models to include random and systematic sources of variation is useful for estimating and, whenever possible, reducing variation during circuit design [11], [12]. Some systematic effects like STI stress and well proximity are now included in many circuit simulators using industry standard device models [13]. Other effects like orientation are handled by the use of special switches in model files [11]. Random variation is modeled by the use of statistical SPICE models which enable Monte Carlo simulation and other more efficient statistical-analysis techniques for block-level statistical design [14], [55]. Additional techniques include reducing the pessimism of corner models by deriving corner models from statistical compact models that include the impact of random variation on device characteristics [55]. At the full-chip level, the use of statistical static timing analysis is an active area of research to avoid the pessimism of worst case corners and also to include additional sources of variation [56].

Another design approach is to map logic functions to a small set of physical primitives that have been optimized for manufacturability [57], [58]. The goal of these approaches is to select

a small set of logic blocks and robust layouts for these logic blocks in order to reduce the area and performance overhead compared to the traditional approaches like gate arrays.

The use of design architectures that are robust to process variation is also being pursued. For example, read and write assist circuits have been proposed as design solutions improving the manufacturability of high-density SRAM [41]. Digital averaging and compensation of offset errors introduced by mismatch is another approach being explored for reducing the sensitivity of analog circuits to mismatch [59]. Finally, there are also attempts to implement common design functions using variation-tolerant algorithms. Bernstein *et al.* [60] show that different adder architectures and logic implementations have different sensitivities to variation.

VIII. CONCLUSION

The scaling of CMOS transistors to nanometer-scale feature sizes is increasing the variability in transistor characteristics. This variability increase poses a challenge to the cost-effective utilization of scaled technologies. Meeting this challenge requires an efficient and comprehensive infrastructure for accurate characterization of the various types of variation. Accurate characterization forms the basis for variability minimization and circuit design and layout techniques to reduce the impact of variation.

This paper presented our infrastructure for variability characterization along with a sample of results obtained from the application of this infrastructure to a number of 90-, 65-, and 45-nm technologies. Some of the impacts of the increased transistor variability on SRAM and analog and digital circuit blocks were described. The range of methods being explored and adopted for reducing variability and minimizing its impact was also outlined.

Profitable utilization of scaled CMOS technologies requires techniques for variability minimization and robust circuit design and layout methods. We believe that accurate and efficient characterization and modeling of transistor variability are indispensable for the development of these capabilities.

ACKNOWLEDGMENT

PDF Solutions is a registered trademark, and CV, dScribe CV, Device Scribe CV, pdFasTest-II, and dataPOWER are trademarks of PDF Solutions, Inc.

REFERENCES

- [1] A. R. Alvarez, B. L. Abdi, D. L. Young, H. D. Weed, J. Teplik, and E. R. Herald, "Application of statistical design and response surface methods to computer-aided VLSI device design," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 7, no. 2, pp. 272–288, Feb. 1988.
- [2] S. W. Director, P. Feldmann, and K. Krishna, "Statistical integrated circuit design," *IEEE J. Solid-State Circuits*, vol. 28, no. 3, pp. 193–202, Mar. 1993.
- [3] S. R. Nassif, "Within-chip variability analysis," in *IEDM Tech. Dig.*, 1998, pp. 283–286.
- [4] K. A. Bowman, X. Tang, J. C. Eble, and J. D. Meindl, "Impact of extrinsic and intrinsic parameter fluctuations on CMOS circuit performance," *IEEE J. Solid-State Circuits*, vol. 35, no. 8, pp. 1186–1193, Aug. 2000.
- [5] A. Asenov, A. R. Brown, J. H. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decanometer and nanometer-scale MOSFETs," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1837–1852, Sep. 2003.
- [6] L. W. Liebmann, S. M. Mansfield, A. K. Wong, M. A. Levin, W. C. Leipold *et al.*, "TCAD development for lithography resolution enhancement," *IBM J. Res. Develop.*, vol. 45, no. 5, pp. 651–665, Sep. 2001.
- [7] G. Scott, J. Lutze, M. Rubin, F. Nouri, and M. Manley, "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," in *IEDM Tech. Dig.*, 1999, pp. 827–830.
- [8] G. Eneman, P. Verheyen, R. Rooyackers, F. Nouri, L. Washington, and R. Degraeve *et al.*, "Layout impact on the performance of a locally strained PMOSFET," in *VLSI Symp. Tech. Dig.*, 2005, pp. 22–23.
- [9] R. Gonzalez, B. Gordon, and M. A. Horowitz, "Supply and threshold voltage scaling for low power CMOS," *IEEE J. Solid-State Circuits*, vol. 32, no. 8, pp. 1210–1216, Aug. 1997.
- [10] M. J. M. Pelgrom, H. P. Tuinhout, and M. Vertregt, "Transistor matching in analog CMOS applications," in *IEDM Tech. Dig.*, 1998, pp. 915–918.
- [11] S. K. Springer, S. Lee, N. Lu, E. J. Nowak, J.-O. Plouchart, and J. S. Watts *et al.*, "Modeling of variation in submicrometer CMOS ULSI technologies," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 2168–2178, Sep. 2006.
- [12] R. A. Bianchi, G. Bouche, and O. Roux-dit-Buisson, "Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance," in *IEDM Tech. Dig.*, San Francisco, CA, 2002, pp. 117–120.
- [13] *BSIM4.60 Users Manual*. [Online]. Available: <http://www-device.eecs.berkeley.edu/bsim3/~bism4.html>
- [14] C. C. McAndrew, "Statistical modeling for circuit simulation," in *Proc. IEEE Int. Symp. Quality Electron. Des.*, 2003, pp. 357–362.
- [15] P. Friedberg, Y. Cao, J. Cain, R. Wang, and C. Spanos, "Modeling within-die spatial correlation effects for process-design co-optimization," in *Proc. IEEE Int. Symp. Quality Electron. Des.*, 2005, pp. 516–521.
- [16] A. K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. Bellingham, WA: SPIE Press, 2001.
- [17] I. Ashan, N. Zamdmer, O. Glushchenkov, R. Logan, E. J. Nowak, and H. Kimura *et al.*, "RTA-driven intra-die variations in stage delay and parametric sensitivities for 65 nm technology," in *VLSI Symp. Tech. Dig.*, Honolulu, HI, 2006, pp. 170–171.
- [18] H. van der Lann, R. Carpajj, J. Krist, O. Noordman, Y. van Dommelen, and J. van Schoot *et al.*, "Etch, reticle, and track CD fingerprint corrections with local dose compensation," *Proc. SPIE*, vol. 5755, pp. 107–118, 2005.
- [19] C. Hess, S. Saxena, H. Karbasi, S. Subramanian, M. Quarantelli, and A. Rossoni *et al.*, "Device array scribe characterization vehicle test chip for ultra fast product wafer variability monitoring," in *Proc. Int. Conf. Microelectron. Test Struct.*, Tokyo, Japan, 2007, pp. 145–149.
- [20] R. Lefferts and C. Jakubiec, "An integrated test chip for the complete characterization and monitoring of a 0.25 μm CMOS technology that fits into five scribe line structures 150 μm by 5000 μm ," in *Proc. IEEE Int. Conf. Microelectron. Test Struct.*, Mar. 2003, pp. 59–63.
- [21] U. Schaper and J. Einfeld, "Parameter variation on chip-level," in *Proc. IEEE Int. Conf. Microelectron. Test Struct.*, Apr. 2005, pp. 155–158.
- [22] M. Quarantelli, S. Saxena, N. Dragone, J. A. Babcock, C. Hess, and S. Minehane *et al.*, "Characterization and modeling of MOSFET mismatch of a deep submicron technology," in *Proc. Int. Conf. Microelectron. Test Struct.*, Monterey, CA, 2003, pp. 238–243.
- [23] S. Saxena, S. Minehane, J. Cheng, M. Sengupta, C. Hess, and M. Quarantelli *et al.*, "Test structures and analysis techniques for estimation of the impact of layout on MOSFET performance and variability," in *Proc. Int. Conf. Microelectron. Test Struct.*, Awaji, Japan, 2004, pp. 263–266.
- [24] E. J. Dudewicz and S. N. Mishra, *Modern Mathematical Statistics*. Hoboken, NJ: Wiley, 1988.
- [25] M. Choi and L. Milor, "Impact on circuit performance of deterministic within-die variation in nanoscale semiconductor manufacturing," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 7, pp. 1350–1367, Jul. 2006.
- [26] T. B. Hook, J. Brown, P. Cottrell, E. Alder, D. Hoyniak, and J. Johnson *et al.*, "Lateral ion implant straggle and mask proximity effect," *IEEE Trans. Electron Devices*, vol. 50, no. 9, pp. 1946–1951, Sep. 2003.
- [27] Y. Kim, S. Sridhar, and A. Chatterjee, "Trench isolation step-induced (TRISI) narrow width effect on MOSFET," *IEEE Electron Device Lett.*, vol. 23, no. 10, pp. 600–602, Oct. 2002.
- [28] Y. Kim, S. Sridhar, and A. Chatterjee, "Correction to "Trench isolation step-induced (TRISI) narrow width effect on MOSFET,"," *IEEE Electron Device Lett.*, vol. 23, no. 11, p. 676, Nov. 2002.
- [29] M. Orshansky, L. Milor, P. Chen, K. Keutzer, and C. Hu, "Impact of spatial intrachip gate length variability on the performance of high-speed

- digital circuits," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 21, no. 5, pp. 544–553, May 2002.
- [30] M. Orshansky, L. Milor, and C. Hu, "Characterization of spatial intrafield gate CD variability, its impact on circuit performance and spatial mask-level correction," *IEEE Trans. Semicond. Manuf.*, vol. 17, no. 1, pp. 2–11, Feb. 2004.
- [31] H. P. Tuinhout, "Impact of parametric mismatch and fluctuations on performance and yield of deep-submicron CMOS technologies," in *Proc. ESSDERC*, Firenze, Italy, 2002, pp. 95–101.
- [32] T. Mizuno, J. Okamura, and A. Toriumi, "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFET's," *IEEE Trans. Electron Devices*, vol. 41, no. 11, pp. 2216–2221, Nov. 1994.
- [33] *dataPower User's Manuals*, PDF Solutions, Inc. [Online]. Available: <http://dp.pdf.com/site/index.html>
- [34] G. Eneman, P. Verheyen, A. De Keersgieter, M. Jurczak, and K. De Meyer, "Scalability of stress induced by contact-etch-stop layers: A simulation study," *IEEE Trans. Electron Devices*, vol. 54, no. 6, pp. 1446–1453, Jun. 2007.
- [35] A. Cathingnol, S. Bordez, K. Rochereau, and G. Ghibaudo, "From MOS-FET matching test structures to matching data utilization: Not an ordinary task," in *Proc. Int. Conf. Microelectron. Test Struct.*, Tokyo, Japan, 2007, pp. 230–233.
- [36] P. A. Stolk, F. P. Widdershoven, and D. B. M. Klasesn, "Modeling statistical dopant fluctuations in MOS transistors," *IEEE Trans. Electron Devices*, vol. 45, no. 9, pp. 1960–1971, Sep. 1998.
- [37] P. S. R. S. Rao, *Variance Components Estimation: Mixed Models, Methodologies and Applications*. London, U.K.: Chapman & Hall, 1999.
- [38] G. S. May and C. J. Spanos, *Fundamentals of Semiconductor Manufacturing and Process Control*. New York: Wiley-Interscience, 2006.
- [39] F. Boeuf, F. Arnauad, C. Boccaccio, F. Salvetti, J. Todeschini, and L. Pain *et al.*, "0.248 μm^2 and 0.334 μm^2 conventional bulk 6T-SRAM bit-cells for 45 nm node low cost—General purpose applications," in *VLSI Symp. Tech. Dig.*, 2005, pp. 130–131.
- [40] M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Ohbayashi, S. Imaoka, and H. Makino *et al.*, "A 45 nm low-standby-power embedded SRAM with improved immunity against process and temperature variations," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2007, pp. 326–328.
- [41] H. Pilo, C. Bardwin, G. Braceras, C. Browning, S. Lamphier, and F. Towler, "An SRAM design in 65-nm technology node featuring read and write-assist circuits to expand operating voltage," *IEEE J. Solid State Circuits*, vol. 42, no. 4, pp. 813–819, Apr. 2007.
- [42] *International Technology Roadmap for Semiconductors*. [Online]. Available: <http://www.itrs.net>
- [43] P. R. Kinget, "Device mismatch and tradeoffs in the design of analog circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 6, pp. 1212–1224, Jun. 2005.
- [44] P. McNamara, S. Saxena, C. Guardiani, H. Taguchi, E. Yoshida, and N. Takahashi *et al.*, "Design for manufacturability characterization and optimization of mixed-signal IP," in *Proc. IEEE CICC*, 2001, pp. 271–274.
- [45] K. A. Bowman, S. G. Duvall, and J. D. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 183–190, Feb. 2002.
- [46] B. J. Lin, "Lithography for manufacturing of sub-65 nm nodes and beyond," in *IEDM Tech. Dig.*, 2005, pp. 48–51.
- [47] K. Adachi, K. Ouchi, N. Aoki, H. Tsujii, T. Ito, and K. Matsuo *et al.*, "Issues and optimization of millisecond anneal process for 45 nm node and beyond," in *VLSI Symp. Tech. Dig.*, 2005, pp. 142–143.
- [48] H. Eisenmann, K. Peter, and A. J. Strojwas, "Generalization of the photo process window and its application to OPC test pattern design," *Proc. SPIE*, vol. 5042, pp. 42–50, 2003.
- [49] S. Jank, V. Temchenko, P. G. Karakatsanis, R. Veerasingam, R. Vallishayee, and C. Dolanski *et al.*, "Method for fast and accurate calibration of litho simulator for hot spot analysis," in *Proc. ISSM*, 2006, pp. 453–456.
- [50] H. Fukutome, Y. Momiyama, T. Kubo, E. Yoshida, H. Morioka, and M. Tajima *et al.*, "Suppression of poly-gate-induced fluctuations in carrier profiles of sub-50 nm MOSFETs," in *IEDM Tech. Dig.*, San Francisco, CA, 2006, pp. 1–4.
- [51] E. Josse, S. Parihar, O. Callen, P. Ferreira, C. Monget, and A. Farcy *et al.*, "A cost-effective low power platform for the 45-nm technology node," in *IEDM Tech. Dig.*, 2006, pp. 693–696.
- [52] A. Asenov and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1- μm MOSFETs with epitaxial and δ -doped channels," *IEEE Trans. Electron Devices*, vol. 46, no. 8, pp. 1718–1724, Aug. 1999.
- [53] T. Mizuno, "New channel engineering for sub-100 nm MOS devices considering both carrier velocity overshoot and statistical performance fluctuations," *IEEE Trans. Electron Devices*, vol. 47, no. 4, pp. 756–761, Apr. 2000.
- [54] A. V.-Y. Thean, Z.-H. Shi, L. Mathew, T. Stephens, H. Desjardin, and C. Parker *et al.*, "Performance and variability comparisons between multi-gate FETs and planar SOI transistors," in *IEDM Tech. Dig.*, San Francisco, CA, 2006, pp. 1–4.
- [55] M. Sengupta, S. Saxena, L. Daldoss, G. Kramer, S. Minehane, and J. Cheng, "Application-specific worst case corners using response surfaces and statistical models," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 24, no. 9, pp. 1372–1380, Sep. 2005.
- [56] C. Visweswariah, K. Ravindran, K. Kalafala, S. G. Walker, S. Narayan, and D. K. Beece *et al.*, "First-order incremental block-based statistical timing analysis," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 10, pp. 2170–2180, Oct. 2006.
- [57] L. Pileggi, H. Schmit, A. J. Strojwas, G. Gopalakrishnan, V. Kheterpal, and A. Koorapathy *et al.*, "Exploring regular fabrics to optimize the performance–cost trade-off," in *Proc. ACM/IEEE Des. Autom. Conf.*, 2003, pp. 782–787.
- [58] A. J. Strojwas, "Conquering process variability: A key enabler for profitable manufacturing in advanced technology nodes," in *Proc. ISSM*, 2006, pp. xxv–xxxii.
- [59] C. Paulus, H.-M. Bluthgen, M. Low, E. Sicheneder, N. Bruls, and A. Courtois *et al.*, "A 4GS/s 6b flash ADC in 0.13 μm CMOS," in *Proc. VLSI Circuits Symp. Dig.*, 2004, pp. 420–423.
- [60] K. Bernstein, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, and S. Nassif *et al.*, "High-performance CMOS variability in the 65-nm regime and beyond," *IBM J. Res. Develop.*, vol. 50, no. 4/5, pp. 433–449, Sep. 2006.



Sharad Saxena (S'87–M'90–S'91–SM'97) received the Ph.D. degree in computer science from the University of Massachusetts, Amherst.

He was a Senior Member of Technical Staff with the R&D Division, Texas Instruments, Dallas. He is a Fellow with PDF Solutions, Richardson, TX, where his responsibilities include developing methods and techniques for characterizing and modeling transistor performance and variation in advanced technologies. He has published 30 conference proceedings and journal papers. He was an Associate Editor of

the IEEE TRANSACTIONS ON COMPUTER AIDED DESIGN OF ELECTRONIC CIRCUITS AND SYSTEMS from September 2002 to December 2005. He is the holder of 15 patents.

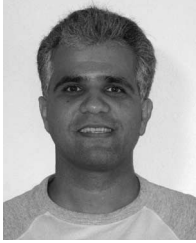
Dr. Saxena has also served on the Program Committee of the International Symposium on Quality in Electronic Design.



Christopher Hess (S'94–A'98–M'00–SM'01) received the diploma degree in electrical engineering and the Dr.-Ing. (Ph.D.) degree in computer science from the University of Karlsruhe, Baden-Württemberg, Germany, in 1992 and 1998, respectively.

He was a Founding Member of the Defect Diagnosis Group, University of Karlsruhe in 1992. He joined PDF Solutions, San Jose, CA, in 1998 as a Fellow, and he is currently responsible for yield and performance characterization. He has published about 40 conference proceedings and journal papers. He is the holder of 10 patents. Since 1992, he has been involved in the design of more than 100 test chips for the microelectronics industry over numerous technologies. For years, his main R&D focus is on the efficient merging of experiments by minimizing chip area usage and manufacturing cycle time. His activities also include developing new architectures of testing equipment, providing several orders of magnitude more data from high-density test chips compared to testing traditional stand-alone experiments. Most recently, he is working on high-density test chips for scribe line applications.

Dr. Hess is a Member of the Electron Devices Society. He is a Technical Committee Member of several semiconductor manufacturing related conferences. He has served as the Technical Chairman of the 2000 International Conference on Microelectronic Test Structures and General Chairman of ICMTS in 2003.



Hossein Karbasi (M'95) received the B.S. degree in electrical engineering from Tehran Polytechnic University, Tehran, Iran, in 1988 and the M.S. and Ph.D. degrees in electrical engineering from the University of Missouri, Columbia, in 1994 and 1998, respectively.

He joined Cadence Design System, San Diego, CA, in 1998, where he worked on various projects related to CMOS design, characterization and reliability. Since 2001, he has been with PDF Solutions, San Diego, working on device characterization (DC/AC/RF) and development of fast DC characterization methods.



Angelo Rossoni received the Dr.Ing. degree in electronic engineering from the University of Brescia, Brescia, Italy, in 2000.

He joined PDF Solutions Inc., Desenzano Del Garda, Italy, in 2001, where he worked on various projects related to VLSI technology design of CMOS process characterization structures.



Stefano Tonello received the Laurea degree in electrical engineering from the Università degli Studi di Padova, Padova, Italy, in 1994.

From September 1995 to June 2000, he was a NOR FLASH and Embedded SRAM Designer with STMicroelectronics, Agrate B.za, Milano, Italy. Since 2001, he has been a Designer of CMOS process characterization structures with PDF Solutions, Desenzano del Garda, Italy.



Patrick McNamara received the MS degree in electrical engineering from the University of Southern California, Los Angeles, CA.

He was with TRW, implementing wideband communication systems. He is currently a Product Marketing Director with PDF Solutions, San Diego, CA, working on analog/RF and digital DFM.



Silvia Lucherini (M'07) received the M.Sc. degree in physics from the Massachusetts Institute of Technology, Cambridge, MA.

From 1980 to 1988, she was with Bell-Northern Research (now Nortel) and Italtel, working on semiconductor lasers R&D. From 1988 to 2005, as an STMicroelectronics Senior Device/Modeling Engineer, she developed flash memory processes and statistical models for analogue applications. Since 2006, she has been a Senior Consulting Engineer with PDF Solutions, Richardson, TX, where she has worked on 65nm CMOS variability studies and improvement.



Seán Minehane (S'95–M'00) received the B.E. degree (Hons.) in electrical and electronic engineering from the National University of Ireland, Cork, Ireland, in 1995 and the M.Eng.Sc. and Ph.D. degrees in microelectronics, also from the National University of Ireland, in 1997 and 2000, respectively.

He was with the National Microelectronics Research Centre (now, the Tyndall Institute), Cork, Ireland, in 1995. From 2001 to 2007, he was with PDF Solutions, Inc., working on design for manufacturability, PDK development, and the simulation of advanced CMOS technologies for digital and mixed-signal/RF applications. He joined the Advanced Technology Group, CSR plc, Cambridge, England, in 2007.



Christoph Dolainsky received the Ph.D. degree in physics from Technische Universität München, Munich, Germany, in 1995.

Since 1996, he has been with PDF Solutions, Munich, Germany, working on OPC tools, litho simulation and applications in verification, and DFM area.



Michele Quarantelli (M'94) was born in Parma, Italy, in 1965. He received the Laurea degree in electronics engineering at the University of Bologna, Italy, in 1990 and the M.S. and Ph.D. degree in electrical and computer engineering at Carnegie Mellon University, Pittsburgh, PA, in 1997 and 2003 respectively, in the field of mismatch modeling and characterization. From 1996 to 1999, he has been Assistant Professor and since year 2000 he is holding a position as Adjunct Professor in the field of Analog Integrated Circuits Design at the University of Brescia Brescia, Italy.

He is now with PDF Solutions, Desenzano del Garda, Italy where he leads the design team. His main research interests are in the field of process characterization, test structures design, and in the design of integrated circuits for applications such as wireless systems, analog signal processing, and basic analog building blocks.