Layout-Dependent Proximity Effects in Deep Nanoscale CMOS

John V. Faricelli
Advanced Micro Devices, Inc., 90 Central St, Boxborough MA 01719, USA
john.faricelli@ieee.org

***Abstract*—As CMOS scaling extends into the nanoscale regime, designers need to be aware that device behavior depends not only on traditional geometric parameters such as channel length and width, but also on layout implementation details of the device and its surrounding neighborhood. The advent of stress engineering, in which intentional mechanical stress is applied to improve device performance, also adds new geometric dependencies. This paper reviews the major process technology features that cause layout-dependent proximity effects and how to account for these effects in circuit and layout design.**

I. Introduction

In the submicron regime, dimensions were large enough that device behavior was largely independent of the neighborhood of the device. As device scaling continued into the nanoscale regime, the importance of modeling the device as well as its surroundings became apparent. This paper consists of three parts. First, it reviews the major features of modern CMOS process technology that give rise to layout dependencies on device behavior. Lithographic proximity effects have been purposely omitted. Second, it discusses some practical issues around implementation of layout-dependent models, and introduces the concept of a macro-model approach decoupled from the underlying compact model. Finally, it provides mitigation strategies for minimizing layout-dependent effects.

II. Well-Implant Proximity Effect

In modern CMOS processes, wells are formed using high-energy ion implantation that require a thick photoresist layer to mask the well implants. Ions implanted into the edge of the photoresist can scatter laterally into an adjacent device region [1]–[3] as shown schematically in Fig. 1.

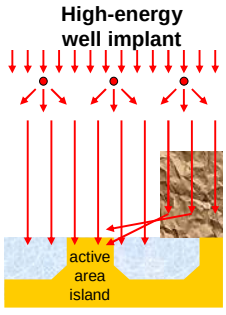

Fig. 1. Ions implanted into the edge of the photoresist can scatter out of the resist, increasing the implant dose of devices near the photoresist edge.

Since the well implant dopant (acceptor or donor) is the same type as the channel implant dopant, the additional doping increases the absolute value of the threshold voltage (V_T) of both NMOS and PMOS devices. The amount of additional dopant depends on the channel distance from the photoresist edge and is termed the well-implant proximity effect (WPE). The interaction distance can be on the order of 1 µm or more, with observed V_T variation of 50–100 mV [1]–[3]. Fig. 2 is an example of V_T variation vs. channel-to-well distance.

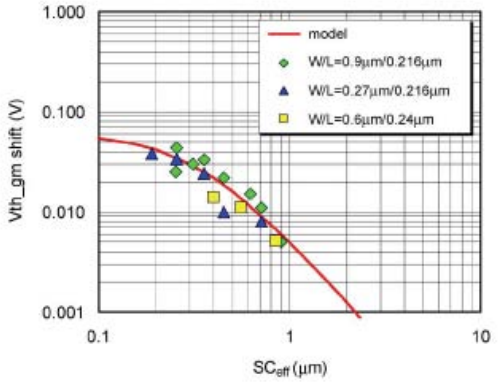

Fig. 2. Change in threshold voltage as a function of MOS channel-to-well distance, SC_{eff} [2].

If the process includes a deep triple well for noise isolation, the implanted well dopants will be opposite to channel dopant type and can cause V_T to become lower in magnitude or even invert the net doping type at the surface [1]. It should be noted that silicon-on-insulator (SOI) processes do not require deep well implants since the MOS device is formed on a thin silicon film that is naturally isolated by the underlying buried oxide. As such, they are largely immune to WPE.

The Compact Modeling Council (CMC) has standardized a WPE model applicable to MOS compact models for circuit simulation. Fig. 3 shows the layout measurements required for the WPE model. The model accounts for an irregular well shape by using a weighted average of the individual measurements in each direction. This WPE model has three MOSFET instance parameters (SCA, SCB, and SCC) which represent integrals of the first/second/third distribution functions for scattered dopants. Closed-form expressions for computing these parameters from the measurements in Fig. 3 are given in [6], with the primary WPE spatial dependence assumed to be $1/d^{\,2}$, where d is the distance from the channel to the well edge. Depending on process details, the SCA term may be sufficient since SCB and SCC represent the exponential tails of

978-1-4244-5759-5/10/$26.00 ©2010 IEEE

the dopant distribution. The SCA/B/C parameters are then used to adjust the device threshold, body coefficient, and low-field mobility in the MOS compact model [6].

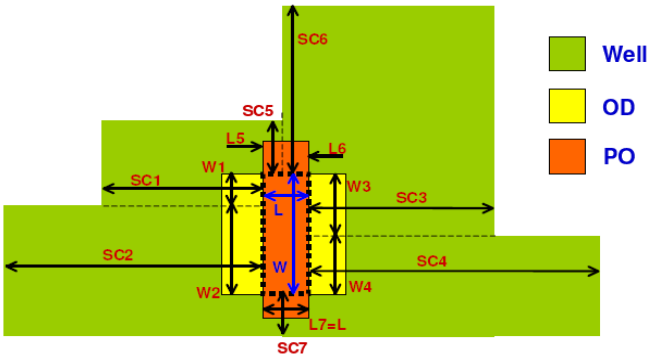

Fig. 3. WPE model layout measurements for the CMC WPE V2.2 model [2], [6]. OD is the oxide definition (active area). PO is the gate.

III. Mechanical Stress Variation

A. Physics of Mechanical Stress on Silicon MOSFETs

CMOS processing uses a variety of materials processed at a wide range of temperatures. Differences in thermal expansion coefficient, lattice mismatch, etc., impose mechanical stress (pressure per unit area), which induces strain (change in length per unit length) on the silicon lattice. Since silicon is piezoelectric, perturbing the equilibrium lattice positions of silicon atoms changes the silicon band structure, causing band splitting, changes in band shape (and effective mass), and redistribution of carriers into different sub-bands [7]. These effects are illustrated in Fig. 4. At the device level, these changes in the band structure modify carrier transport quantities like mobility, which in turn affect device current. Stress can also alter the device threshold by either modulating dopant diffusivity [8] or altering the band structure [9].

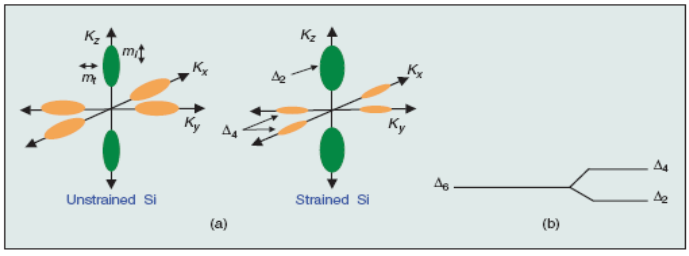

Fig. 4. Effect of strain on silicon band structure [7].

Recent generations of CMOS device architectures have incorporated intentional, or engineered, uniaxial stress along the channel length to enhance device performance [10]–[14]. Depending on the carrier type, one may want to introduce tensile or compressive stress to improve carrier mobility and thus drain current. The desired stress orientations for improved device performance are illustrated in Fig. 5. Note that NMOS and PMOS transistors are not symmetric in their response to applied stress. In the channel-length direction, NMOS transistors favor tensile stress while PMOS transistors favor compressive stress. However, both types favor tensile stress in the channel-width direction.

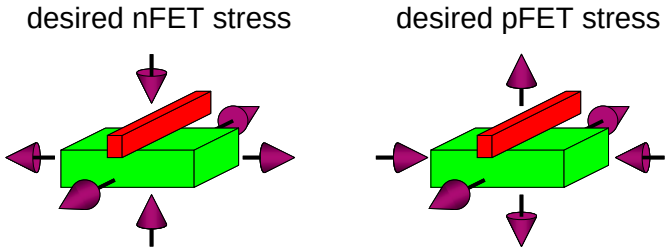

Fig. 5. Desired stress orientations for improving device performance.

B. Shallow Trench Isolation Stress

One of the first stress-related effects to be investigated and modeled was the effect of shallow trench isolation (STI) proximity on device performance [15]–[17]. STI formation induces compressive stress on the active device region (oxide definition or OD as noted in Fig. 3). Fig. 6 shows an idealized stress profile compared against channel position for two different lengths of OD (LOD). The stress seen by the device depends on the distance from the device channel to both STI edges. When both edges are close, the stress is additive.

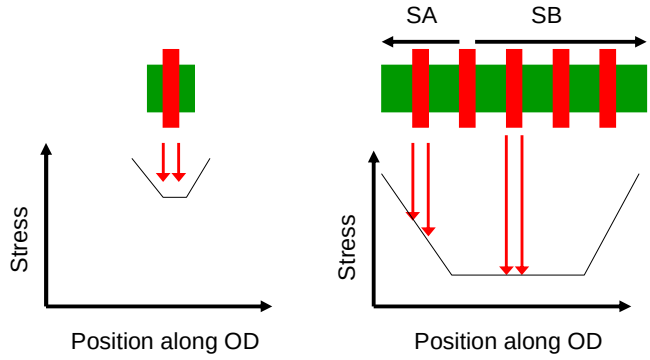

Fig. 6. Stress vs. position for two different lengths of OD (after [9]).

The stress profiles are typically approximated by a 1/LOD dependence [17]. Typically, this is implemented in the compact model using the SA and SB parameters which measure the distance from the gate to the left and right STI edges respectively.

Since STI introduces compressive stress, proximity to the STI edge degrades NMOS performance while enhancing PMOS performance. To minimize variation among “identical” gate fingers, the middle fingers of a long OD stripe should be used, or the OD stripe should be extended and dummy devices introduced on the left and right.

C. Stress Liners

Stress liners (also known as contact etch stop layers) are a common method for introducing intentional uniaxial stress to enhance device performance [10]–[14]. These layers are typically silicon nitride films deposited by plasma-enhanced chemical vapor deposition over the fabricated transistor gates. Depending on the deposition conditions, either tensile or compressive layers may be formed [18]. Since NMOS devices prefer tensile stress in the channel direction while PMOS devices prefer compressive stress, a dual-stress liner (DSL) approach is used. TEM cross-sections of such a process are shown in Fig. 7.

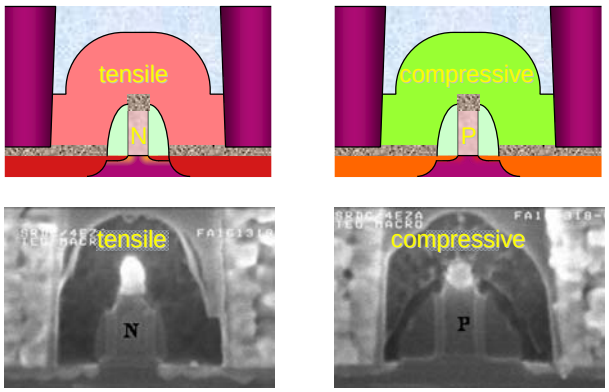

Fig 7. DSLs with tensile layer over NMOS and compressive layer over PMOS [12].

The DSL method enhances drive current significantly for both NMOS and PMOS devices but has a drawback in that it introduces an additional source of variation at the boundary between the two liner types of opposite stress, as shown in Fig. 8. At the boundary between the two stress liners, the opposite stresses naturally relax. NMOS devices near this boundary are degraded compared to an NMOS far from the boundary because they experience less tensile stress. PMOS transistors, on the other hand, may see either degradation or enhancement, depending on whether the DSL boundary is parallel to the channel-width or -length direction. See Fig. 5.

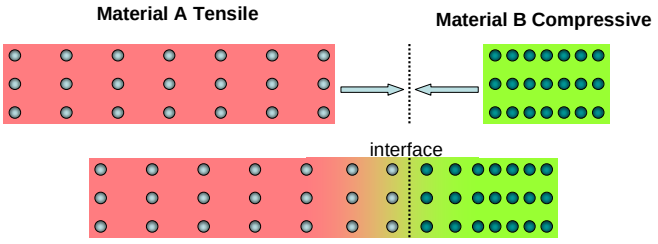

Fig. 8. Stress relaxation at the boundary between tensile and compressive stress liners (plan view).

Although detailed simulation studies of the stress profiles on individual transistors have been reported, e.g., [19], there is no commonly accepted compact device modeling strategy to capture the effect of proximity to the DSL boundary. An example of a DSL boundary proximity model is described in [20]. This model decomposes the two-dimensional stress field into one-dimensional components parallel and perpendicular to the direction of current flow. Using test structures with varying channel-to-DSL boundary distances, a phenomenological model of stress vs. distance is built. Applied stress is then related to device mobility by an additional relationship.

In addition to proximity to the DSL boundary, the effect of stress liners can be modulated by the distance to the neighboring poly as well as contacts that cut through the liner [21]. As the poly-to-poly spacing decreases, the stress liner loses its effectiveness in transferring stress to the channel of the device [14]–[22]. This has implications for the continued used of DSL stressors at nanoscale dimensions. The width of the poly spacer also plays a role – larger spacers provide less area for the stress liner to "grab" and transfer stress to the channel.

D. Contact-Induced Stress

In some processes, the source/drain contact itself can be used to introduce intentional stress into the device [14]. Fig. 9 shows the benefit of introducing a tensile trench contact to NMOS transistors in a 45-nm technology. The common tensile stress in the contacts should degrade PMOS performance, but the authors claim this is mitigated by the raised compressive source/drain regions used in their embedded SiGe process. If the trench contact extends over the entire width of the device, there should be minimal layout sensitivity to this scheme.

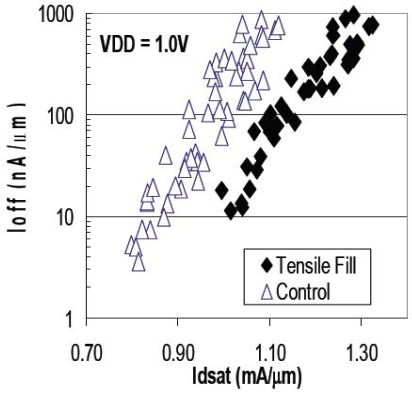

Fig. 9. I_{dsat}-I_{off} benefit of tensile contact fill on NMOS performance [14].

E. Stress Memorization and Gate-Induced Stressors

Another source of engineered stress is the stress memorization technique (SMT) [23], [24]. A temporary tensile stress layer is placed over the NMOS transistors. The stress is transferred into the underlying source/drain regions during thermal anneal after which the disposable nitride layer is removed. Drain current enhancement of more than 27% was reported at the 65-nm node [24]. One might expect SMT to

impose similar layout-dependent proximity effect as DSL since the disposable tensile liner is placed only over the NMOS transistor and stress should be relaxed near PMOS transistors. But this effect is yet to be discussed in the literature.

In processes employing a replacement gate or "gate last" integration flow to circumvent high-*K*/metal-gate thermal stability issues, channel stress can be induced by appropriate choice of gate material. The use of compressive gate material on NMOS transistors increased I_{dsat} by 6% at the 45-nm node [14]. Since the stressor is applied directly to the channel, there is likely no proximity effect from neighboring devices, but there may be some length dependence. Any such effect would likely be captured by the base compact model during fitting for different channel lengths or by binning.

F. Source/Drain Stressors

Many CMOS processes induce channel stress by growing regions of different materials in the source/drain through selective epitaxy [13], [14]. Embedded SiGe (eSiGe) is most commonly used to introduce compressive stress on the PMOS channel. In this process, the source/drain regions are etched and SiGe (with a larger lattice constant than Si) is selectively grown. Depending on the process details, the source/drain regions may become raised. Fig. 10 shows a TEM cross-section of a 45-nm PMOS device with eSiGe. The angular shape of the eSiGe is related to the etch process used to remove part of the Si source/drain prior to epitaxy [14].

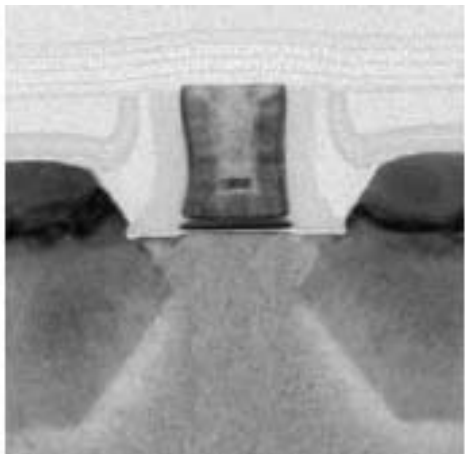

Fig. 10. TEM cross-section of PMOS transistor with eSiGe source/drain [14].

For NMOS transistors, the use of embedded SiC to introduce tensile stress has been reported [25], [26]. Since SiC has a smaller lattice than the underlying silicon, it pulls on the silicon under the source/drain, which in turn induces a tensile stress in the channel region. A 9% improvement in I_{on} due to SiC source/drain was reported in [26] at the 45-nm node, compared to a "best practices" device employing SMT and tensile liners.

In either case, the stress profile on the device may depend on both the channel length and the cumulative amount of source/drain region (volume effect) in the OD stripe containing the device [27]. The trend is for the induced stress to drop off sharply at narrow OD lengths and saturate at wide OD lengths. This implies that PMOS devices with eSiGe will exhibit I_{dsat} degradation at narrow OD lengths, unlike the non-eSiGe case, in which the STI effect would dominate and cause enhancement.

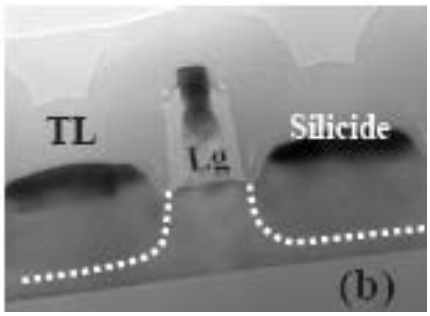

Fig. 11. Cross-sectional TEM image of eSiC poly-gate NMOS device [26]. The eSiC region is denoted by the dotted line.

Wang proposed the following model for the stress σ_p on each device finger due to eSiGe [28]:

$$\sigma_p = \left(1 + \frac{m}{L} + \frac{m}{L + L_{sd}} + \frac{m}{2L + L_{sd}}\right) \frac{L_{SD}}{1 + L_{SD}} \sigma_m \qquad (1)$$

where L is the channel length, L_{sd} is the length of the source/drain diffusion regions between gate stripes, and m and σ_m are model parameters. The second two terms in parentheses account for the effect of the two nearest neighbor devices. The model assumes each source/drain diffusion region has the same size although the model equation can be modified for differently sized diffusion regions. A comparison of the model vs. measured data is shown in Fig. 12. Measured data in [29] also confirms that the presence of adjacent poly gates on the same OD stripe will reduce the amount of PMOS enhancement due to reduced eSiGe volume.

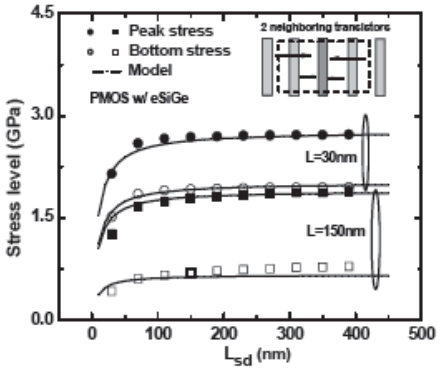

Fig. 12. Stress variation of eSiGe on both channel length and source/drain lengths [28].

G. Effect of Neighboring Active Regions on STI Stress

The amount of STI stress seen by a MOSFET depends not only on the width of the stripe, but also on its proximity to other OD stripes [30]–[32], as shown in Fig. 13. This can be inferred from simple physical reasoning: as the width of the STI goes to zero (OD regions merge), STI stress is eliminated. At large STI widths, the effect of the neighboring OD region becomes negligible. The stress at the edge of the OD region can be modeled as

$$\sigma_{P\_STI} = \left(1 + \frac{m}{LOD}\right) \frac{W_{STI}}{(A_{STI} + W_{STI})} \sigma_{m\_STI} \tag{2}$$

where W_{STI} is the distance between OD regions, LOD is the length of the OD region, σ_{m_STI} is the STI saturation stress, and m and A_{STI} are modeling parameters [32]. Fig. 13 presents measured data on the OD-OD stress effect. The effect of neighboring OD regions is most pronounced when the gate to STI edge distance is small.

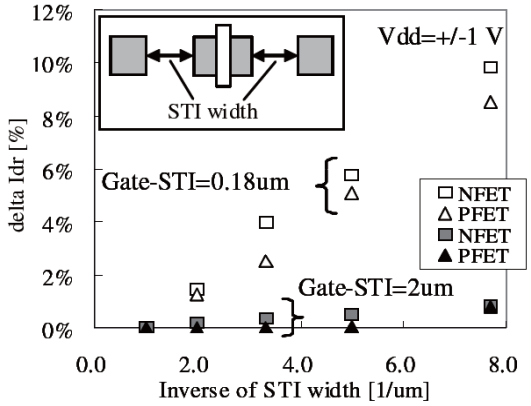

Fig. 13. Change in drain current due to proximity of neighboring OD regions [30].

IV. Model Implementation

A. Reference Device Concept and Model Parameter Extraction

To extract layout-dependent parameters for a compact model efficiently, a common strategy is to partition the modeling into two parts: a base model and a set of layout-dependent models that modulate the base model. The base model is typically fit to a set of measured devices that spans the *L* and *W* space for that model bin. Care must be taken in the design of these test device structures to minimize any layout variation that will subsequently be modeled separately, although some differences, such as channel length or width, are unavoidable.

The layout specification, such as distance to well edge, OD length, number of adjacent device gates, etc., constitutes the reference device layout. It simplifies the base model parameter extraction if the reference layout minimizes the layout-dependent effects. For example, one could measure the middle device fingers of a multi-stripe OD region, with neighboring well, OD, and stress liner edges "far" away. The designer should not assume, however, that the base device is always one in which layout-dependent effects are minimized. After performing parameter extraction, it is possible that the model is re-normalized so that the "nominal" device resembles more realistic layout, such as a ring oscillator used for model validation.

Given the complex nature of layout-dependent modeling with so many effects and parameters, the wise circuit designer should experiment with the model by creating test layouts to understand how important each effect is and to look for anomalous behavior.

B. Extraction of Modeling Dimensions from Layout

The job of extracting the modeling dimensions from layout is usually done as part of a layout-versus-schematic (LVS) run performed before layout parasitic *RC* extraction. LVS tools provide a rich set of geometric manipulation functions that can measure the distances between the MOS device gate region and well(s), OD edges, stress liner edges, etc. Designers should be aware, however, that these measurements do not come without a penalty in LVS runtime. Depending on the size of the layout, extracting distances for layout-dependent models can increase LVS runtime from minutes to many hours. Often, runtime can be improved dramatically by carefully limiting the search distance the LVS tool will use to look for neighboring geometry. Checks on OD geometry are particularly sensitive because of the large number of OD shapes compared to well shapes.

C. Micro- vs. Macro-Modeling

Many of the models used to capture layout-dependent effects modify the low-field mobility of the compact device model. Unfortunately, device current is not necessarily linearly dependent on low-field mobility, even at micron-scale dimensions. In addition, the choice of fitting parameters used in the compact model can exacerbate the non-linear response of modeled current compared to mobility. This can lead to a vicious cycle of adjusting other model parameters, such as saturation velocity, to compensate.

One way out of this dilemma is to take a macro-modeling approach, as was done for a high-performance 32-nm SOI process [33]. Rather than indirectly through compact modeling parameters, a device current scaling factor and an offset in account for the layout-dependent effects. This simplifies the fitting procedure and model regression against data, and led to a more accurate reflection of the model against measurement. It is transparent, because the amount of deviation from the reference device is obvious and easily quantified. It decouples the layout-dependent modeling from the choice of compact model used for the base device, so supporting a different compact model does not require re-extraction and qualification

of the layout-dependent modeling. Finally, the layout-dependent model could be implemented as a separate software package and used in a variety of other tools, as will be described in Section V-D.

The macro-modeling approach does have some limitations. By using a per-instance current scale factor, we assume the layout-proximity effects have the same quantitative effect in the linear and saturation regions of the device. The use of a threshold voltage shift implies that the change to the body effect coefficient due to WPE can be ignored. For digital circuits and many analog applications, however, this level of accuracy has proven sufficient.

V. Mitigation of Layout-Dependent Effects

A. Layout Guidelines to Reduce Variability

One strategy to reduce layout-induced device mismatch or minimize variation between pre- and post-layout simulations is to employ layout guidelines [34]. It should be noted that guidelines for device mismatch are not the same as those for minimizing differences between pre- and post-layout. In the first case, geometric consistency is sufficient but absolute matching with schematic-only simulation is not required. Minimizing differences between pre- and post-layout simulation is more difficult to achieve. One way to accomplish this is to specify all the layout-dependent parameters at the schematic level and have the layout generated from that specification list. This approach is achievable for a single device but becomes complex for realistic layout because of interactions between devices.

While the specifics of layout guidelines are dependent on the details of the manufacturing process being used, general guidelines to reduce layout-induced device mismatch are:

- Keep well edges at consistent distances in all four directions from MOS devices to minimize variation due to WPE effects.
- Maintain consistent distances to any DSL boundaries in all four directions. Since this layer is usually derived from other drawn layers, this may be difficult to do unless the foundry provides a way to visualize the stress layers in the layout editing tool.
- Use an active region (OD) and gate definition of the same size, shape, and orientation.
- Add dummy devices and/or dummy poly over STI so device fingers at the edge of a shared OD behave similarly to inner fingers.
- Add dummy OD shapes to minimize OD-OD effects. Wide OD stripes are less susceptible to neighboring OD effects.

Fig. 14 gives an example of applying these guidelines to reduce layout-induced mismatch. It should be noted that these guidelines will force an increase in the size of layout, and schematics will need to be altered when dummy transistors are added to pass LVS.

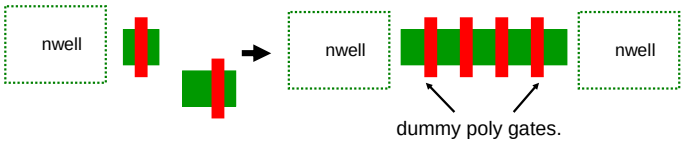

Fig. 14. Reducing layout-induced variability by following layout guidelines.

A more formal approach to layout guidelines is to incorporate them into the design rule check (DRC) deck. Transistors are tagged with a special ID layer that triggers special rules to reduce layout variability. While this method is satisfying because it is "correct by construction," it can induce even more layout "bloat" than guidelines.

B. Post-Layout Optimization

If we have a good set of layout guidelines coded as DRC rules, it becomes possible to identify a problematic layout and apply fixes. A simple example of this is shown in Fig. 15. Two PMOS transistors are adjacent to an empty region. Since each PMOS has an n-well boundary in close proximity, both devices will be subject to WPE. A simple "fix up" rule would close those gaps with n-well by merging the two n-wells.

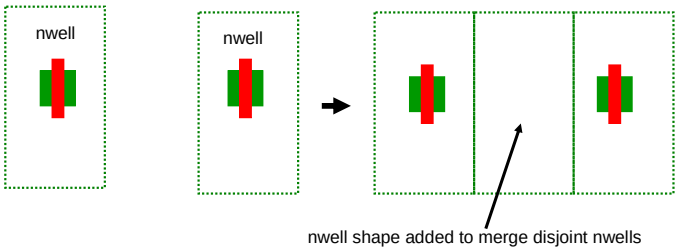

Fig. 15. Merging n-well shapes alleviates WPE.

C. Layout Analysis Tools

An alternate methodology for mitigating layout-dependent proximity effects is to provide a tool to the layout engineer that analyzes the layout and provides guidance on how to minimize them. The SEISMOS-LX tool [35], for example, can perform detailed stress analysis of layouts and present the results to the layout designer as 2D stress maps that indicate where stress variation is greatest.

The tool can also work in a more conventional macro-modeling mode by extracting measurements and converting them via a compact model to a mobility scale factor and threshold voltage offset for use in circuit simulation [36]. A similar macro-modeling approach using Mentor tools was described in [37], in which measurements performed by an LVS tool were fed into a model for layout-dependent effects and then back-annotated onto the layout. Both of these analyses require co-development between the CAD tool vendor and the foundry, since the models to convert layout dimensions to device performance need to be tuned to each process.

A layout analysis flow based on information available from the foundry design kit suitable for small to medium-sized layouts has been implemented at AMD. The flow performs an LVS run to extract each transistor finger in the layout, along with the layout-dependent parameters used by the compact model. The netlist is read into a script that connects each transistor terminal to a DC source, and a DC operating point is run to compute I_{eff} [38] for each device. Because correct net connectivity from the LVS run is not required, the analysis can be performed right after transistor placement while the layout is still LVS "dirty."

The flow was later enhanced to give the layout designer a more detailed breakdown of the effects causing variability. A series of circuit simulations is run, with each run turning on the set of parameter responsible for a particular physical effect. The results of the analysis are displayed on an error-layer overlay in the layout design system, so the layout designer can simply select a transistor and get a summary of the layout-dependent variation broken down by physical effect. A sample screenshot is shown in Fig. 16.

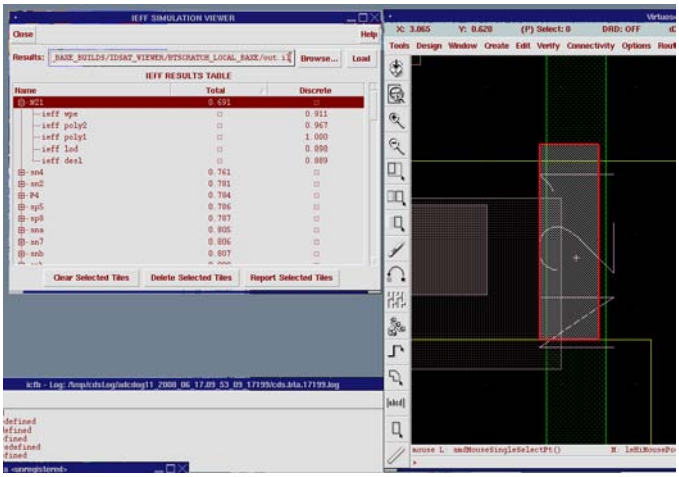

Fig. 16. Screenshot of variation viewer.

D. Standard Cell Characterization

Including layout-dependent proximity effects in the analysis of standard cells presents an interesting challenge, since the neighborhood around the standard cell is not known until after placement. Fortunately, a fair approximation of the surrounding layout can be made by enclosing the cell with a "collar" that includes a typical well and OD placement [33]. An example is shown in Fig. 17. This collar assumes the standard cells are placed in rows of N/P/P/N, with n-wells shared between rows of adjacent cells. Rows and columns of dummy cells can be added during placement to mimic the layout conditions used for cell characterization.

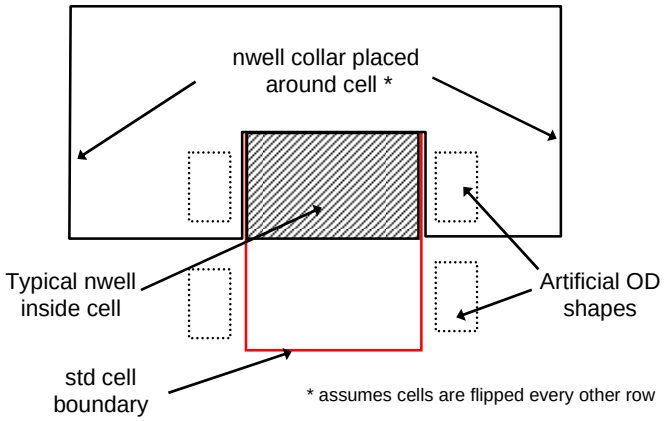

Fig. 17. Use of "collar" layout to mimic surroundings of a standard cell in a typical placement scenario.

VI. Conclusion

A review of the sources of layout-dependent device variation has been presented: well-edge proximity, STI (LOD) effect, stress liner boundary effect, contact-induced stress variation, source/drain stressor volume effect, and adjacent OD effect. Methodologies for accounting for these effects at the compact modeling level, given layout, have been described.

Being able to account for layout-dependent variation is not enough, however, to ensure an efficient and timely design process. Circuit and layout designers must work together closely to account for these effects up front. This can be done through layout guidelines, correct-by-construction techniques, or by layout analysis tools that identify variation issues in the early stages of layout. Failure to do so may result in unpleasant surprises and many cycles of revision late in the design.

Acknowledgments

The author acknowledges the team of AMD and GLOBALFOUNDRIES colleagues who worked on modeling layout-dependent stress effects, in particular Akif Sultan (GF), Sushant Suryagandh (GF), Alvin Loke (AMD), Tom Daum (AMD), and Greg Constant (AMD).

REFERENCES

[1] T. B. Hook, J. Brown, P. Cottrell, E. Adler, D. Hoyniak, J. Johnson, and R. Mann, "Lateral ion implant straggle and mask proximity effect," *IEEE Trans. Electron Devices*, vol. 50, no. 9, Sep. 2003, pp. 1946–1951.

[2] Y.-M. Sheu, K.-W. Su, S. Tian, S.-J. Yang, C.-C. Wang, M.-J. Chen, and S. Liu, "Modeling the well-edge proximity effect in highly scaled MOSFETs," *IEEE Trans. Electron Devices*, vol. 53, no. 11, Nov. 2006, pp. 2792–2798.

[3] J. Watts, K.-W. Su, and M. Basel, "Netlisting and modeling well-proximity effects," *IEEE Trans. Electron Devices*, vol. 53, no. 9, Sep. 2007, pp. 2179–2186.

[4] T. H. Morshed *et al*., "BSIM4.6.4 MOSFET model user's manual," University of California, Berkeley, Apr. 2009.

[5] X. Li, W. Wu, G. Gildenblat, G. D. J. Smit, A. J. Scholten, D. B. M. Klaassen, and R. van Langevelde, "PSP 103.1 technical note," Arizona State University, Apr. 2009.

[6] http://www.eigroup.org/cmc/proximity_effect_modeling/wpe_guidelines_v2.2.pdf

[7] N. Mohta and S. E. Thompson, "Mobility enhancement," *IEEE Circuits and Devices Mag.*, vol. 21, no. 5, Sep./Oct. 2005, pp. 18–23.

[8] Y.-M. Sheu, S.-J. Yang, C.-C. Wang, et.al. "Modeling mechanical stress effect on dopant diffusion in scaled MOSFETs," *IEEE Trans. Electron Devices*, vol. 52, no. 1, Jan. 2005, pp. 30–37.

[9] C.-C. Wang, W. Zhao, F. Lin, M. Chen, and W. Cao, "Modeling of layout-dependent stress effect in CMOS design," in *Proc.* 2009 *Int. Conf. on Computer Aided Design* (San Jose, CA), Nov. 2009, pp. 513–520.

[10] A. Shimizu *et al*., "Local mechanical stress control (LMC): A new technique for CMOS-performance enhancement," in *IEEE Int. Electron Devices Meeting Tech. Dig.* (Washington, DC), Dec. 2001, pp. 433–436.

[11] V. Chan *et al*., "High speed 45nm gate length CMOSFETs integrated into a 90nm bulk technology incorporating strain engineering," in *IEEE Int. Electron Devices Meeting Tech. Dig*. (Washington, DC), Dec. 2003, pp. 77–80.

[12] H. S. Yang *et al*., "Dual stress liner for high performance sub-45nm gate length SOI CMOS manufacturing," in *IEEE Int. Electron Devices Meeting Tech. Dig.* (San Francisco, CA), Dec. 2004, pp. 1075–1077.

[13] M. Horstmann *et al*., "Integration and optimization of embedded-SiGe, compressive and tensile stressed liner films, and stress memorization in advanced SOI CMOS technologies," in *Int. Electron Devices Meeting Tech. Dig.* (Washington, DC), Dec. 2005, pp. 233–236.

[14] C. Auth *et al*., "45nm high-k + metal gate strain-enhanced transistors," in *IEEE Symp. on VLSI Technology Tech. Dig.* (Honolulu, HI), Jun. 2008, pp. 128–129.

[15] 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 *IEEE Int. Electron Device Meeting Tech. Dig.* (Washington, DC), Dec. 1999, pp. 827–830.

[16] R.A. Bianchi, G. Bouche, amd O. Roux-dit-Buisson, "Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance," in *IEEE Intl. Electron Devices Meeting Tech. Dig.* (San Francisco, CA), Dec. 2002, pp. 117–120.

[17] K.-W. Su et al., "A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics," in *Proc. Custom Integrated Circuits Conf.* (San Jose, CA), Sep. 2003, pp. 245–248.

[18] M. Maeda and K. Ikeda, "Stress evaluation of radio-frequency-biased plasma-enhanced chemical vapor deposited silicon nitride films," *J. Applied Physics*, vol. 83, no. 7, Apr. 1998, pp. 3865–3870.

[19] F. Payet, F. Boeuf, C. Ortolland, and T. Skotnicki, "Nonuniform mobility-enhancement techniques and their impact on device performance," *IEEE Trans. Electron Devices*, vol. 55, no. 4, Apr. 2008, pp. 1050–1057.

[20] R. Q. Williams, D. Chidambarrao, J. H. McCullen, S. Narasimha, T. G. Mitchell, and D. Onsongo, "Compact model methodology for dual-stress nitride liner films in a 90nm SOI ULSI technology," in *NSTI Nanotech*, (Boston, MA), May 2006, pp. 858–859.

[21] X. Zhang, W. Wu, G. Du, R. Han, and X. Liu, "Influence of contact positioning on the performance of stress induced MOSFETs," in *Proc. IEEE Int. Conf. on Electron Devices and Solid State Circuits* (Hong Kong, China), Dec. 2008.

[22] G. Eneman, P Verheyen, A DeKeersgieter, M. Jurczak, and K. DeMeyer, "Scalability of stress induced by contact-etch-stop layers: A simulation study," *IEEE Trans. Electron Devices*, vol. 54, no. 6, Jun. 2007, pp. 1446–1453.

[23] C.-H. Chen *et al*., "Stress memorization technique (SMT) by selectively strained-nitride capping for sub-65nm high-performance strained-Si device application," in *IEEE Symp. on VLSI Technology Tech. Dig.* (Honolulu, HI), Jun. 2004, pp. 56–57.

[24] A. Wei *et al*., "Multiple stress memorization in advanced SOI CMOS technologies," in *IEEE Symp. on VLSI Technology Tech. Dig.* (Kyoto, Japan), Jun. 2007, pp. 216–217.

[25] K.-W. Ang, "Enhanced performance in 50nm N-MOSFETs with silicon-carbon source/drain regions," in *IEEE Int. Electron Devices Meeting Tech. Dig.* (San Francisco, CA), Dec. 2004, pp. 1069–1071.

[26] B. Yang *et al*., "High-performance nMOSFET with *in-situ* phosphorus-doped embedded (ISPD eSi:C) source-drain stressor, " in *Proc. IEEE Int. Electron Devices Meeting* (San Francisco, CA), Dec. 2008, pp. 51–54.

[27] G. Eneman *et al.*, "Scalability of the $Si_{1-x}Ge_x$ source/drain technology for the 45-nm technology node and beyond," *IEEE Trans. Electron Devices*, vol. 53, no. 7, Jul. 2006, pp. 1647–1656.

[28] C.-C. Wang, W. Zhao, F. Liu, M Chen, and Y. Cao, "Compact modeling of stress effects in scaled CMOS," in *Proc. IEEE Simulation of Semiconductor Processes and Devices* (San Diego, CA), Sep. 2009, pp. 131–134.

[29] C.-Y. Cheng, Y.-K. Fang, J.-C. Hsieh, S.-J. Yang, Y.-M. Sheu, and H. Hsia, "New observations in LOD effect of 45-nm P-MOSFETs with strained SiGe source/drain and dummy gate," *IEEE Trans. Electron Devices*, vol. 56, no. 8, Aug. 2009, pp. 1618–1623.

[30] H. Tsuno *et al*., "Advanced analysis and modeling of MOSFET characteristics fluctuation caused by layout variation," in *IEEE Symp. on VLSI Technology Tech. Dig.* (Kyoto, Japan), Jun. 2007, pp 204–205.

[31] A. B. Kahng, P. Sharma, R. O. Topaloglu, "Exploiting STI stress for performance," in *Proc. IEEE Int. Conference on Computer-Aided Design* (San Jose, CA), Nov. 2007, pp. 83–90.

[32] C.-C. Wang, W. Zhao, F. Liu, M. Chen, and Y. Chao, "Modeling of layout-dependent stress effect in CMOS design," in *Proc. IEEE Int. Conf. on Computer-Aided Design* (San Jose, CA), Nov. 2009, pp. 513–520.

[33] J. Faricelli, "Layout-dependent proximity effects in deep nanoscale CMOS," in *IEEE Solid-State Circuits Society Seminar* (Fort Collins, CO), Apr. 2009. http://ewh.ieee.org/r5/denver/sscs/Presentations/2009_04_Faricelli.pdf

[34] A. Sultan *et al*., "CAD utilities to comprehend layout-dependent stress effects in 45nm high-performance SOI custom macro design," in *Proc. Int. Symp. on Quality Electronic Design* (San Jose, CA), Mar. 2009.

[35] Seismos-LX, http://www.synopsys.com

[36] X.-W. Lin, "Layout proximity effects and device extraction in circuit designs," in *Proc. Int. Conf. on Solid-State and Integrated Circuit Technology* (Beijing, China), Oct. 2008, pp. 2228–2231.

[37] H. Eissa and A. E. El Mously, "Physical aware design methodology for analog and mixed signal integrated circuits," in *Proc. Int. Workshop for Design and Test* (Riyadh, Saudi Arabia), Nov. 2009, pp. 1–5.

[38] M. Na, E. Nowak, W. Haensch, and J Cai., "The effective drive current in CMOS inverters," in *IEEE Int. Electron Devices Meeting Tech. Dig.* (San Francisco, CA), Dec. 2002, pp. 121–124.