5

Intrinsic Parameter Fluctuations in Scaled MOSFETs

In chapter 2 we briefly discussed different sources of intrinsic parameter fluctuations (IPF) and reviewed some of the literature. The IPF are becoming one of the major factors limiting the scaling of conventional MOSFETs and their integration into billion transistor count chips. In this chapter we study IPF due to random dopants and Line Edge Roughness (LER) in well calibrated 35 nm MOSFETs and in the scaled transistors corresponding to the next generations of technology nodes according to the ITRS. Recall from the discussions in chapter 2 that, the main sources of IPF are: random discrete dopants, gate LER and Si/SiO$_2$ interface roughness leading to oxide thickness fluctuations. However, only the simulation results with the first two sources of fluctuations are presented in this thesis.

Following introductory discussion on discrete random dopants, the results from the statistical simulator on the IPF in scaled conventional MOSFETs due to random discrete dopants are presented in section 5.1. Section 5.2 deals with the simulation of IPF introduced by LER in 35 nm and 25 nm transistors. In the case of the 35 nm device the impact of different LER RMS amplitudes ($\Delta = 3$ nm, 2 nm, 1 nm) is also studied.

In addition to the statistical LER simulations, which are based on randomly generated gate edges, a study of “deterministic” LER patterns is presented in section 5.3. In this case the gate edges are intentionally designed to have specific shapes representing specific roughness configurations. A full process simulation with Taurus has been performed, in the case of LER simulation, for both the 35 nm and 25 nm n-MOSFETs.
5.1 Random discrete dopants

The number of dopants in long and wide channel MOSFETs is large and the IPF due to the variations in the number of dopants and their position in the active region are not so significant. However as MOSFETs scale down to sub-100 nanometer dimensions the statistical variation in the number of dopants and their physical location in the active region of devices results in significant IPF. The atomistic nature of these sub-100 nm MOSFETs is illustrated in figure 5:1a [5.1]. The devices, with sub-50 nm dimensions shown in figure 1b and c, can no longer be described, modelled or simulated based on the traditional assumptions of continuous dopant distribution, smooth interfaces and straight gate edges, illustrated in figure 5:1a. The granularity of charge and the atomicity of matter result in structural variations from device to device which is the cause of IPF [5.1] [5.]

![Figure 5:1](image)

**Figure 5:1** The evolution of atomistic devices: (a) MOSFET with continuous ionised dopant charge and smooth boundaries and Si/SiO<sub>2</sub> interfaces; (b) Sketch of 22 nm channel length MOSFET required for 45 nm technology node with rough interface, LER and random discrete dopants; (c) Impression of 5 nm channel length MOSFET with the silicon crystal lattice superimposed. Courtesy of Asenov et al.[5:1]

MOSFETs with 7 nm gate lengths are expected to be in mass production around 2018. Such devices, similar to the one in figure 5:1c, will have approximately 10-15 silicon atoms along the channel length and the position of each silicon, dopant or insulator atom is likely to have significant macroscopic impact on the device characteristics. Conventional MOSFETs at such dimensions require a high doping concentration in the channel in order
to suppress short channel effects. Aside from its adverse effect on device performance the high doping concentration will introduce intolerable IPF.

Due to the increasing impact of discrete random dopants on critical device parameters, it becomes very important to consider this phenomenon in the process of nano-scale device modelling and simulation. This means that the introduction of statistical analysis into process and device simulation is vital. The Glasgow University atomistic device simulation software (discussed in chapter 3) is a specifically developed simulation tool to investigate this problem, which is one the major stumbling blocks of CMOS device scaling.

Most simulation studies of intrinsic parameter fluctuations have been restricted to devices corresponding to one particular technology node, or to idealized devices. In this work, we have extended the study of intrinsic parameter fluctuations to a family of realistic MOSFETs corresponding to all technology nodes until the end of the current edition of the ITRS. This allows us to investigate the magnitude of parameter fluctuations in future technology generations and will enable us to understand the trend of fluctuations in decanano-meter MOSFETs.

5.1.1 Simulation Approach

The simulation of IPF shifts the paradigm of traditional device simulation. In the presence of microscopic variations from device to devices it is inadequate to simulate one device in order to characterize the operation and parameters of all macroscopically similar but microscopically different devices. It becomes necessary to simulate a statistically significant sample of devices. This will allow a meaningful statistical analysis including estimates of mean values, variances and higher moments of the statistical distributions.

In this section, using statistical 3D simulations, we study the random dopant induced IPF in conventional MOSFETs scaled to 25, 18, 13, and 9 nm gate lengths according to the requirements of the ITRS for high performance devices in the 65, 45, 32 and 22 nm technology nodes respectively.

In the scaling procedures, described in more details in chapter 2, the generalized scaling rules, in parallel with ITRS predictions’ have been adopted. The doping concentration was broadly scaled by the factors, \( \alpha \) and \( \kappa (N_a = \alpha \kappa N_a) \), where \( \kappa = \sqrt{2} \) and \( \alpha = \kappa (V_{sd}/V_{dd}) \) where, \( V_{sd} \) and \( N_a \) are supply voltage and channel doping concentration.
of the scaled devices respectively. Super-retrograde channels have been achieved using ion implantation. The doping in the middle of the channel is reduced by progressive use of pocket implants (see chapters 3 and 4).

The assignment of discrete dopants is based on rejection technique [5:2]. Given the expected number of dopants estimated from the continuous doping distribution obtained from the Taurus process simulator [5.3], the probability that there is a dopant in a mesh ‘brick’ (a cell of 3-D mesh) is calculated. Then, using rejection, the dopants are placed randomly according to the initial continuous-doping distribution.

During the atomistic device simulation mobility is handled as follows. Based on the results of calibration and scaling obtained from full process simulation steps, a continuously doped device is simulated for all $I-V$ points and the corresponding mobility profile is stored. This mobility profile is then used in the course of the atomistic device simulation.

A typical simulation domain used in the simulation of the 35 nm Toshiba MOSFET is illustrated in figure 5:2. The position of individual discrete dopants is clearly identifiable from the potential landscape. The fluctuation in the surface potential and depletion layer edge inflicted by individual dopants are clearly seen.

![Figure 5:2](image)

**Figure 5:2** Image of 3D potential distribution in one of the simulated 35 nm MOSFETs illustrating the position of discrete random dopants in the channel (indium) and source/drain regions (arsenic). The continuous yellow line depicts the approximate position of $p$-$n$ junction.
The spread of the current voltage characteristics of a sample of 200 microscopically different devices is illustrated in figure 5:3. The open-dotted line represents an experimental data from [5.4]. Results of the statistical analysis of such devices are reported in the next subsections.

![Figure 5:3 Spread $I_d$-$V_g$ characteristics of 200 transistors with 35 nm gate length. The open circle shows the $I-V$ curve from the experimental 35nm Toshiba device.](image)

Three device parameters namely, $V_T$, $I_{inv}$ and $I_{off}$ are statistically studied and presented in the following sections. The values of off-current ($I_{off}$) and strong inversion current ($I_{inv}$) are extracted at the points where $I_d(V_g=0V)$ and $I_d(V_g=V_{dd})$ respectively. $\sigma V_T$ was calculated using a standard statistical procedure once the $V_T$ is automatically extracted from each of the simulated $I$-$V$ curves. For extraction purposes, a current threshold voltage criterion $I_g(V_g=V_T)=I_s(W/L)$ has been adopted. An integrated device simulation methodology, described in [5.5], has been implemented. In the following sub-sections the results of statistical simulation of IPF due to the number of dopants and their position in the depletion region of the scaled MOSFETs is presented.


5.1.2 Threshold Voltage variation

Threshold voltage is an important parameter in the MOSFET design. In concert with the subthreshold slope it determines the off-state leakage current [5.6]. The gate voltage overdrive, which is the difference between the supply voltage and the threshold voltage determines the drive current. One of the main roles of the channel profile design is for $V_T$ adjustments [5.7] in conjunction with the short channel control. Well defined, steady and stable threshold voltage a crucially important for analogue and digital circuits, i.e., less variation of $V_T$ is highly desirable.

Therefore it is, important to keep $\sigma V_T$ within an acceptable degree of tolerance in order to deliver a stable integrated circuit and properly working system. However in real nano-scale MOSFETs and integrated circuits the discrete random dopants in the channel region introduce $V_T$ fluctuations. Moreover, the fluctuations increase significantly as the gate length decreases.

![Figure 5.4](image.png)

Figure 5.4 Threshold voltage variation as a function of scaled device gate length.
Figure 5.4 shows the standard deviation of the threshold voltage, $\sigma V_T$, as a function of the gate length shrinks. $\sigma V_T$ rapidly increases when the transistors are scaled below 20 nm gate length. For example a standard deviation of approximately 60mV for the 18 nm transistor gives a spread in $\pm 3\sigma V_T$ of $\pm 360$ mV, which is very large, resulting in some of the devices in the integrated circuit not turning on. As the channel length is further reduced the fluctuations become excessively high with a spread for $\pm 3\sigma V_T$ of $\pm 510$mV for the 9 nm transistor. This is in excess of the supply voltage expected to be in the range of $(V_{dd} \sim 500-600$mV).

Although most studies focus on conventional MOSFETs, the $V_T$ variation due to random discrete dopants is not a unique phenomenon for these devices. A recent simulation study in [5.8] showed that the random discrete dopants in source and drain of double gate MOSFETs also induce $V_T$ fluctuations. For example the double gate MOSFET with effective channel lengths of 4, 6, 8, 10 nm exhibits a $\pm 3\sigma$ variation of 204, 90, 45, 30mV respectively. Although the degree of $\sigma V_T$ in the conventional MOSFETs presented in this thesis is much higher, the fluctuations in non-conventional MOSFETs are not negligible [5.8].

The histogram distributions of the threshold voltage in the prototype 35 nm and scaled 25, 18 nm transistors are illustrated in figure 5.5. All the results are on 3D atomistic device simulation of samples of 200 devices each. For all these sets of devices the distribution of the threshold voltage is close to the normal distribution, with a standard deviation increasing with reduction of the channel length.

Another important information that can be drawn from frequency distribution of threshold voltage in the all three devices is the increase in range of $V_T$ distribution as the transistor channel length decreases. The ranges of distribution are approximately 162, 249, and 330mV for 35, 25 and 18 nm gate lengths respectively (see figure 5.5). Comparing between the range of $V_T$ distribution of the 25 nm transistor and the 35 nm prototype transistor, there is a 52% increase and in the case of 18 nm device the range increases more than 200%. This reflects an increase in $V_T$ fluctuation in the smaller channel length transistors.
Figure 5:5 Histograms of $V_T$ distributions for the prototype 35 nm and the scaled 25 and 18 nm MOSFETs
5.1.3 Off-current variation

The current flow in a MOSFET has two major components: the diffusion current resulting from the gradient carrier concentration and drift associated with the applied electric field [5.7]. While the strong inversion regime is dominated by the drift component of the overall drain current, the subthreshold current is dominated by diffusion [5.9] [5.10].

The off-current is another important parameter in MOSFETs, which to great extent determines the standby leakage current in integrated circuits. The impact of discrete random dopants on $I_{\text{off}}$ is well illustrated in figure 5:3. The off-current fluctuations are in the range of $1.2 \times 10^{-8} – 1.5 \times 10^{-6}$ A/µm. There are nearly two orders of magnitude differences between the maximum and minimum subthreshold currents in the sample. Such a level of fluctuations is highly unacceptable for particular circuit applications, like 6T SRAM cells (see section 5.1.5).

The mean and the standard deviation of $I_{\text{off}}$ is depicted in figure 5:6. The presence of random discrete dopants results in a substantial increase in the average subthreshold leakage current with the reduction of a channel length. For example $\sigma I_{\text{off}}$ for 18 nm is 220nA/µm, which is more than 180% greater than the mean off current for a 35 nm MOSFET. This is highly undesirable from the static power dissipation point of view.

![Figure 5:6 Mean and standard deviation of $I_{\text{off}}$ in the 35, 25 and 18 nm MOSFETs](#)
In order to visualize and compare the distributions the off-current for different channel lengths, the histograms of these distributions are presented for all three devices in figures 5:7. The column one represent plot of linear scale and two represent $\log(I_{\text{off}})$.

<table>
<thead>
<tr>
<th>Off-current: linear scale plot</th>
<th>$\log(I_{\text{off}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_g = 35\text{ nm}$</td>
<td>$L_g = 35\text{ nm}$</td>
</tr>
<tr>
<td><img src="image" alt="Linear scale plot" /></td>
<td><img src="image" alt="Logarithmic scale" /></td>
</tr>
<tr>
<td>$L_g = 25\text{ nm}$</td>
<td>$L_g = 25\text{ nm}$</td>
</tr>
<tr>
<td><img src="image" alt="Linear scale plot" /></td>
<td><img src="image" alt="Logarithmic scale" /></td>
</tr>
<tr>
<td>$L_g = 18\text{ nm}$</td>
<td>$L_g = 18\text{ nm}$</td>
</tr>
<tr>
<td><img src="image" alt="Linear scale plot" /></td>
<td><img src="image" alt="Logarithmic scale" /></td>
</tr>
</tbody>
</table>

**Figure 5:7** $I_{\text{off}}$ distribution of the calibrated and scaled (25 and 18nm) devices both in linear scale and logarithmic scale.
Unlike the distributions of the threshold voltage, which are close to normal distributions, the $I_{\text{off}}$ distribution is a log-normal distribution (see figure 5:7 above). This is due to the exponential dependence of the subthreshold current on gate voltage. It is also important to note that there is substantial increase in the average subthreshold leakage current with the reduction of channel length. For example the average $I_{\text{off}}$ in 25 and 18 nm device samples are 95 and 152\text{nA}/\mu\text{m} respectively. This corresponds to a percentage increase of 11.7% and 72.7% respectively compared to the off-current in the 35 nm transistor.

Figure 5:8 illustrates the channel length dependence of the extracted skew and kurtosis of the distributions of the off-current which in this case differ substantially from the expected value for the normal distribution. The other important information that can be drawn from the results plotted on figure 5:8 is that the smaller the gate length of the devices the more the distribution is skewed to the right (positive skew) with reference to the prototype device of 35 nm gate length.

![Figure 5:8](image)

Figure 5:8  3\textsuperscript{rd} and 4\textsuperscript{th} moments of the $I_{\text{off}}$ distribution in 18, 25, and 35nm MOSFETs
5.1.4 Drain current variation in the strong inversion region

For simplicity and computational efficiency, we have studied the drain current in strong inversion at low drain voltage as an indication of the fluctuations in the drive current ($I_{on}$). Although there will be a quantitative difference in the $I$-$V$ simulation results under high drain voltage, the qualitative difference estimated in previous simulation studies [6.2] is rather small. Hence the results presented in this section are good indications to the fluctuation of “$I_{on}$” despite the fact that the data are obtained from the atomistic device simulation at low drain voltage of 50mV.

The corresponding data are extracted in strong inversion at gate voltage of $V_g = 850$mV which is equivalent to the supply voltage of $V_{dd} = 850$mV. The drain current fluctuations in strong inversion are not as strong as in the subthreshold current. At high gate voltage the inversion layer charge screens the Coulomb potential of the individual discrete dopants reducing the potential fluctuations in the channel and the corresponding current fluctuations.

![Figure 5:9 Standard deviation and mean of drain current in the strong inversion regime.](image)

**Figure 5:9** Standard deviation and mean of drain current in the strong inversion regime.
As expected there is a trend of increasing $I_{on}$ fluctuation magnitude with the reduction of the channel length illustrated in figure 5:9. For example, the ±3σ variation of the strong inversion current for an 18nm transistor is about 288µA/µm, which is a significant current variation in such a small device. As with the distribution of the threshold voltage, the distribution of the on-current is also close to normal in form (see figure 5:10).

The inversion drain current is crucial to the performance of MOSFETs. One of the motivations of scaling down transistors size is to achieve high relative drive current in order to boost circuit performance. However variation in the drive current may result in mismatch in analogue applications and variation in the signal propagation times of digital circuits.

The mean and the standard deviation of $I_{inv}$ and $I_{off}$ for the different channel lengths are summarised in table 5:1. The skew and kurtosis of the off-current distribution is also added to the table. The $I_{inv}$ current variations are relatively smaller comparing it with the corresponding $I_{off}$ variations. On the other hand, the increases in the fluctuation of $I_{inv}$ in the scaled devices are not negligible. There is a 35% increase of $\sigma_{I_{ev}}$ in the 25 nm transistor compared to the prototype 35 nm transistor. This percentage of fluctuation in $I_{inv}$ increases to 176% in the 18 nm transistor.

<table>
<thead>
<tr>
<th>$L_g$ [nm]</th>
<th>$\langle I_{inv} \rangle$ [µA/µm]</th>
<th>$\langle I_{off} \rangle$ [nA/µm]</th>
<th>$\sigma_{I_{inv}}$ [µA/µm]</th>
<th>$\sigma_{I_{off}}$ [nA/µm]</th>
<th>$\sigma_{V_T}$ [mV]</th>
<th>Skew</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>282</td>
<td>35</td>
<td>18</td>
<td>78</td>
<td>33</td>
<td>2.4</td>
<td>8.1</td>
</tr>
<tr>
<td>25</td>
<td>300</td>
<td>95</td>
<td>23</td>
<td>112</td>
<td>46</td>
<td>2.5</td>
<td>8.5</td>
</tr>
<tr>
<td>18</td>
<td>350</td>
<td>152</td>
<td>48</td>
<td>220</td>
<td>61</td>
<td>3.0</td>
<td>12.0</td>
</tr>
</tbody>
</table>

Table 5:1 Summary of selected statistical parameters
Figure 5:10 Normalized histogram of drain current distribution in the strong inversion region ($I_{inv}$) of the 35 nm; the 25 nm and 18 nm n-channel transistors.
5.1.5 Impact of random discrete dopant fluctuation on circuits

The IPF resulting from random dopants in individual devices discussed in the previous sections contribute significantly to mismatch [5.11] between neighboring devices in integrated circuits. Recent simulation studies show that random doping fluctuations have already adversely affected the yield and the stability of SRAM illustrated in figure 5:11 [5.12][5.13].

Figure 5:11 Circuit schematics of SRAM cell (a); Distribution of SNM; adapted from [5.14]

Figure 5:11 illustrates the random dopant induced distribution of static noise margin (SNM) in an ensemble of SRAM cells of various cell ratios at the transition characteristic for the advanced stages of the 90nm technology node. To obtain acceptable yield, in the presence of random dopants fluctuation, the cell ratio, \( r = (W_d / L_d) / (W_a / L_a) \) (where \( W \) and \( L \) are width and length of the driver and the access transistors in SRAM cell) must be increased from its nominal value of \( r = 1 \) to \( r = 3 \). However, this increase in cell ratio leads to a larger cell area and thus compromises SRAM scaling and hence the scaling of the majority of silicon based digital systems. Moreover, the simulation study in [5:14] also shows that fluctuation of SNM due to random dopants in the 35 nm device is comparable with the fluctuation caused by instabilities in the power supply of \( \pm 25\% \).
5.2 Line edge roughness in conventional MOSFETs

Ideal MOSFETs are considered to have a straight gate edges as illustrated in figure 5:10a. However in real devices, the gate is prone to line edge roughness (LER) inherited from various imperfect process steps. In the past, LER had little impact on device operation because the gate length was much larger compared to the roughness of the gate edges. However, with transistors gate length scaled to sub-50nm the contribution of LER to overall IPF is becoming proportionally significant. Therefore, in this work, we have investigated on the effect of LER on the variation in the electrical characteristics of advanced decananano MOSFETs.

![Figure 5:12](image)

**Figure 5:12** Comparative illustrations of (a) an ideal transistor with straight gate edges and uniform channel width and (b) a top view of hypothetical gate edges which are constructed from a pair of the randomly generated rough lines used in this work for to investigate LER

In the following section, brief discussions of the main sources of LER are followed by results from process and device simulations that take in to account the effect of LER on the transistor’s parameters. The simulation results are for gate lengths 35 and 25 nm.
5.2.1 Contributing factors to LER

LER is a major process related problem, which affects MOSFET parameters. As a result of gate LER, gate geometry will vary from transistor to transistor. Since the drain current is related to the gate geometry the overall current may vary introducing also $V_T$ variation from transistor to transistor [5.15]. At the same time, the doping distribution near the p-n junctions will follow the gate shape introducing variations in the channel gate length. It is therefore important to control the LER in order to minimize the corresponding IPF.

The main contributing factors for LER are: molecular structure (polymer aggregate) of resist,\(^*\) which will generate a line width fluctuation in resist pattern that will affect the final shape of the transistor line edge during the etching process [5.16]; the degree of mixing of resist materials (photosensitive compound, base resin, and organic solvent); the fluctuation of photo chemical events, like absorption of radiation and reflected photons and the pattern transfer process (etching) where the acid diffusion length can degrade the line edge and space pattern [5.17].

5.2.2 Characterisation of LER

LER can quantitatively be characterized by an autocorrelation function with two parameters; a correlation length ($\Lambda$) and the root mean value ($\Delta$). From a statistical point of view, the autocorrelation function (ACF) describes the dependence of values of random data at one interval on the values at another one [5.18]. It can directly be estimated from the sampled data values as the product of a distance $dx$-shifted with itself [5.19] [5.20]. Assuming that $\langle f(x) \rangle = 0$ over a large sample of discrete data, the ACF can be expressed as:

$$ACF(r\Delta x) = \frac{1}{N-r} \sum_{n=1}^{N-r} (f(x_n))(f(x_{n+r})),$$

$$r=1,2,...m: \ (m < N) \quad (6.1)$$

\(^*\) The role of resist materials in the device manufacturing process is to perform the function of transforming a pattern placed on the mask onto a desired location.
in the case of LER $f(x_n)$ is an estimation to the displacement from the average line $\langle f(x) \rangle$ at point $x_n$ (see figure 5:12b), $N$ is the number of samples regularly spaced by $\Delta x$ over the width of the transistor, $r$ is often called the lag number and $m$ is its maximum value [5.21]. Usually LER is assumed to have a Gaussian or exponential ACF($x$) given by equations (5.2) and (5.3) respectively.

$$ACF_{G}(x) = \Delta^2 \exp\left(-\frac{x^2}{\Lambda^2}\right) \quad (5.2)$$

$$ACF_{E}(x) = \Delta^2 \exp\left(-\frac{x^2}{\Lambda}\right) \quad (5.3)$$

There is an additional parameter used for LER characterization known as a roughness exponent ($\alpha$) discussed in the literature [5.20], [5.22], [5.23], [5.24] that takes into account the relative contribution of high frequency fluctuation to the spatial roughness. It is usually associated with LER in photo-resists and 2D-rough surfaces (for example Si/SiO$_2$ interface roughness). However, in this work we concentrate on the two parameters, $\Delta$ and $\Lambda$ of the correlation function, which we believe are sufficiently adequate to characterize the gate LER.

The parameter $\Delta$ characterizes the deviation of roughness magnitude with respect to the stationary length $\langle f(x) \rangle$ (see figure 5:12b). The equation for $\Delta$ has a similar definition as the standard deviation in statistics and is given by:

$$\Delta = \sqrt{\frac{\sum_{i=1}^{N} (f(x_i) - \langle f(x) \rangle)^2}{N}} \quad (5.4)$$

Despite the simplicity of equation 5.4 to measure the vertical roughness ($\Delta$) it does not capture the property of roughness along the gate width [5.22]. Hence additional parameter to characterize LER is essential.
The ITRS (2001) defines LER as $3\Delta$. However in the latest ITRS 2003 edition a new definition has been introduced. The new lithography guideline for LER is defined in terms of line width roughness (LWR) as $LER = \frac{LWR}{\sigma}$, and the required figures are given in table 6:2. The work on LER in this thesis has been performed prior to the guidance line; hence the values used in this work do not exactly parallel the roadmap. However the result obtained still highlights the problem that the presence of LER creates for devices.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology nodes</td>
<td>hp90</td>
<td>hp65</td>
<td>hp45</td>
<td>hp32</td>
<td>hp22</td>
</tr>
<tr>
<td>Dram ½ pitch</td>
<td>90</td>
<td>65</td>
<td>45</td>
<td>32</td>
<td>22</td>
</tr>
<tr>
<td>Printed gate length [nm]</td>
<td>53</td>
<td>35</td>
<td>25</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Printed gate CD* control (3σ) [nm]</td>
<td>3.3</td>
<td>2.2</td>
<td>1.6</td>
<td>1.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Line Width roughness (3σ) [nm]</td>
<td>3.0</td>
<td>2.0</td>
<td>1.4</td>
<td>1.0</td>
<td>0.7</td>
</tr>
</tbody>
</table>

* Critical dimensions (CD) are referred to dimensions of the smallest geometrical features like printed gate length, width of interconnect line, contacts, trenches, etc., which can be formed during device manufacturing process. [5.27]

Table 5:2 ITRS guidelines for LER, LWR, and printed gate critical dimension (CD) variation

The other parameter needed to more completely characterize the gate LER is $\Lambda$, which is a measure of the damping distance or the width of the autocorrelation function. $\Lambda$ signifies the narrowness of the power spectrum [5.25] of the roughness along the gate width. The value of $\Lambda$ often used is in the range of 20-50 nm. For most work presented in this thesis, the correlation length is taken to be 30 nm.

5.2.3 LER Simulation approach

In this work the investigation of LER and its effect on transistor parameters is mainly based on full device process simulation, in which the randomly generated rough line pattern is introduce onto the nominally straight gate in order to produce an LER effect during device simulation. The random lines are generated using a Fourier technique explained in detail in [5.26].
The RMS height and the correlation length are used as inputs to the random rough line generator along with the corresponding choice of autocorrelation function, either the Gaussian or exponential ACF. Various values of RMS amplitude ($\Delta = 1$ nm, $2$ nm, and $3$ nm) and a correlation length of $\Lambda = 30$ nm were used for each set of process simulated devices. In addition to this further values of $\Lambda$ ($5$ nm and $15$ nm) were considered to study how the gate shape and the metallurgic p-n junctions are affected by the autocorrelation length.

The randomly generated rough lines of both gate edges are then used as a mask to transfer the roughness pattern to the poly gate during the etching process simulation step. The device structure obtained from the full process simulation is used as input to device simulation. In order to identify the contribution of LER to IPF, all the device process conditions including doping concentration and annealing time are identical to the originally calibrated and scaled devices with the exception of the introduction of LER to the poly gate.

### 5.2.4 IPF due to LER

In this section the effects of LER on the electrical parameters of MOSFETs including $I_{\text{off}}$, $I_{\text{on}}$ and $V_T$ are presented for different $\Delta$ and $\Lambda$.

#### 5.2.4.1 The effect of LER on device parameters with various value of $\Delta$

Figure 5:13 shows the simulated $I_d-V_g$ characteristics of 50 n-MOSFETs, with nominal gate length of 35 nm, gate width 100 nm and a randomly generated gate edges. Values of $\Delta$ and $\Lambda$ are 1 nm and 30 nm respectively. The inset images show the 2-D net doping profile of the Si/SiO$_2$ interface of two extreme cases from the sample corresponding to $I_{\text{max}}$ and $I_{\text{min}}$. The two devices are very different in terms of effective channel length. Compared to the effective channel of the 35nm MOSFET, which is 33 nm, ‘device A’ and ‘device B’ have 22% smaller and 12% larger channel length respectively.
The approximate position of the polysilicon gate is also indicated in the figure 5:13. The p-n junction follows the shape of the gate. In order to study in more detail the effect of LER on impurity doping distribution and carrier mobility in those two extreme cases, the doping profile at the cross-section C-C is shown in figures 5:14 and 5:15.

![Figure 5:13](image)

**Figure 5:13** $I_d$-$V_g$ characteristics for a sample of 50 n-channel MOSFETs, including LER effect with an RMS height of 1 nm. Values at high drain voltage of 850mV are plotted in logarithmic scale and those at low drain voltage of 50mV are plotted in linear scale. The image to the right is the net doping profile of the two samples with the minimum and maximum off currents.

The main reason for the low performance of device B and the high performance of device A are the large and the small effective channel length respectively. At the same time according to figure 5:14, device A has lower doping concentration in the channel and higher average channel mobility, $\langle \mu_{ch} \rangle = 60.7 \text{ cm}^2/\text{Vs}$, compared to device B. In addition to larger effective channel length, device B has higher doping concentration in the channel giving average channel mobility, $\langle \mu_{ch} \rangle = 51.5 \text{ cm}^2/\text{Vs}$. Hence device B delivers lower performance.
In order to be able to compare how different values of $\Delta$ will affect the $I_d$-$V_g$ characteristics and the doping profiles of the 35 nm MOSFET, simulation results for $\Delta=2$ nm and $\Delta=3$ nm are presented in figures 5:16 and 5:17 respectively. The variation of the distance between the metallurgical p-n junctions is shown for extreme devices in figures 5:16 and 5:17 respectively. As expected, roughness increases with the increasing $\Delta$. As a consequence of gate width variation, the irregularity of the effective channel length defined by the position of the metallurgical $p$-$n$ junction of the transistors increases.

![Figure 5:14](image1.png)  
**Figure 5:14** 1-D net-doping profile in the interface and mobility of device A at the cross section A-A

![Figure 5:15](image2.png)  
**Figure 5:15** 1-D net-doping profile in the interface and mobility of device B at the cross section A-A

![Figure 5:16](image3.png)  
**Figure 5:16** $I_d$-$V_g$ characteristics of 50 $n$-channel MOSFET. RMS height of 2 nm.
Figure 5:17 $I_d$-$V_g$ characteristics of 50 n-channel MOSFETs, including LER effects with an RMS height of 3 nm. The image to the right depicts the two extreme cases occurring as a result of LER in devices E and F among the simulated 50 devices.

The final shape of the metallurgical p-n junction is determined by all processing steps. Results of the process simulation at 4 stages of the simulation flow are presented in figure 5:18. The values of $\Delta$ are: 1 nm, 2 nm and 3 nm for the 35 nm transistor, and 3 nm for the 25 nm transistor. The correlation length is 30 nm in all cases. The width is 100 nm, which is more than three times the correlation length, adequate enough to capture the roughness effect along the device width that gives 2-D characterization of the LER.

Images in the first row represent the shape of the gate after the etching step. Row two shows the deposited nitride offset spacer, which controls the effective channel length during the extension and contact area implantation. The images in row three show 2-D arsenic doping profile after the ion implantation procedure for the formation of source and drain extensions.

At this stage only a very short spike anneal is performed at a relatively low temperature and for a very short period of time. Still some smoothing of the metallurgical junction compared to the gate edge shape is observable. The strongest effect on the doping
distribution is the rapid thermal annealing (RTA) process after the source/drain contact implantation. After the full process steps the $p$-$n$ junctions in all three cases of different RMS height are smoothed by the RTA step. In addition to the smoothing the RTA advances the junctions reducing the effective channel length.

<table>
<thead>
<tr>
<th>$L_G$ [nm]</th>
<th>35</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta$ [nm]</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gate shape (Top view)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 nm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Off-set spacer</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Junction formation</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Full process</th>
</tr>
</thead>
</table>

**Figure 5:18** LER with different values of $\Delta$. The red colour in pictures of row 1 and 2 represent the etched polysilicon gate material and yellow represents silicon dioxide insulator. The green colour in row 2 represents the silicon oxy-nitride spacer. The red colour in pictures of row 3 (after the formation of $p$-$n$ junction where the polysilicon gate is removed for visualisation purpose) represent the arsenic (As) doping concentration in source and drain region and the blue colour represent minimal As concentration in the channel near the interface. The lines in pictures of row 3 and 4 show the position of $p$-$n$ junctions. The width of the device is 100 nm
Figure 5:19 depicts samples of 1-D net doping profiles of the devices presented in figure 5:18. The net doping profile shown in closed and open symbols corresponds to devices in row three and four respectively. After full device process steps the increase in the net channel doping (near the interface of Si/SiO$_2$) is shown in figure 5:19 for all cases of RMS heights.

![Figure 5:19 1-D channel doping profiles near the interface of devices shown in figure 5:18. Images in row three represent 2D arsenic doping profile just after the implantation and the corresponding 1-D profile is shown in filled symbols. Images in row 4 represent after full process simulation and the corresponding 1-D profile is shown in open symbols.](image)

The overall effects of LER on IPF using various values of RMS are illustrated in figure 5:20. The third row depicts the histogram for the $I_{on}$ distributions. The $I_{on}$ fluctuation in the 35 nm MOSFET doesn’t show a significant difference for different values of $\Delta$. On the other hand there is a steady increase in average derive current from 670 $\mu$A/$\mu$m to 741 $\mu$A/$\mu$m as the RMS value increases from 1 nm to 3 nm. This is due to shorter channel length as the RMS value increases from 1 nm to 3 nm.

The corresponding off-current distribution is shown in row 1 and 2 both in linear and logarithm scales in figure 5:20. The off-current distribution often can be described as a lognormal. The fluctuation in off-current significantly increase with the increase of $\Delta$. When $\Delta = 1$, 2 and 3 nm the standard deviation of the off-current is 55, 79, and 157nA/$\mu$m respectively (table 5:4). Since the roughness increases with increasing RMS values, the off current fluctuation is increasing as expected.
Figure 5.20: Histograms showing distributions of intrinsic device parameters in the 35nm n-MOSFET. The columns and rows represent RMS values and the corresponding distribution of electrical parameters.
respectively. The correlation length is 30nm in all cases. Row 1 and 2 show linear scale and logarithmic scale of the \( I_{\text{off}} \) current distributions respectively.

The negative impact of LER on the threshold voltage variations and subthreshold slope is also shown on the histogram distributions of \( V_T \) and \( S \) illustrated in figure 5:20. The two important features of electrostatic integrity (threshold voltage and subthreshold slope) in the 35 nm transistor are compared in row four and five. For example, the \( \pm 3\sigma V_T \) variation for corresponding RMS values of 1, 2 and 3nm are \( \pm (22, 30 \text{ and } 80 \text{mV}) \) respectively. The other aspect of \( V_T \) degradation is the short channel effect, which is the decrease of threshold voltage as the effective channel length is reduced [5.7]. As previously discussed, the length between the metallurgical \( p-n \) junction decreases for the

The simulated values of important MOSFET parameters for the gate length of 35 nm and 25 nm with the effect of LER included are summarised Tables 5:3 and 5:4.

<table>
<thead>
<tr>
<th>( L_G ) [nm]</th>
<th>( \Delta ) [nm]</th>
<th>35</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_{\text{on-MAX}} ) [( \mu A/\mu m )]</td>
<td>761</td>
<td>797</td>
<td>817</td>
</tr>
<tr>
<td>( I_{\text{on-MIN}} ) [( \mu A/\mu m )]</td>
<td>574</td>
<td>575</td>
<td>627</td>
</tr>
<tr>
<td>( I_{\text{off-MAX}} ) [nA/( \mu m )]</td>
<td>285</td>
<td>411</td>
<td>710</td>
</tr>
<tr>
<td>( I_{\text{off-MIN}} ) [nA/( \mu m )]</td>
<td>41</td>
<td>14</td>
<td>46</td>
</tr>
<tr>
<td>( I_{\text{d-max}} ) @ 50mV [( \mu A/\mu m )]</td>
<td>137</td>
<td>143</td>
<td>146</td>
</tr>
<tr>
<td>( I_{\text{d-min}} ) @ 50mV [( \mu A/\mu m )]</td>
<td>97</td>
<td>104</td>
<td>110</td>
</tr>
</tbody>
</table>

**Table 5:3** Comparison of electrical parameters 35 and 25nm gate length of n-channel MOSFETs with LER effect.

<table>
<thead>
<tr>
<th>( L_G ) [nm]</th>
<th>RMS [nm]</th>
<th>( \sigma(I_{\text{off}}) ) [nA/( \mu m )]</th>
<th>( \sigma(I_{\text{on}}) ) [( \mu A/\mu m )]</th>
<th>( \sigma(V_T) ) [mV]</th>
<th>( \sigma S ) [mV/dec]</th>
<th>( \langle I_{\text{on}} \rangle ) [( \mu A/\mu m )]</th>
<th>( \langle I_{\text{off}} \rangle ) [nA/( \mu m )]</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>1</td>
<td>55</td>
<td>46</td>
<td>7.4</td>
<td>1.7</td>
<td>670</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>79</td>
<td>45</td>
<td>10.03</td>
<td>1.5</td>
<td>690</td>
<td>131</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>157</td>
<td>41</td>
<td>26.5</td>
<td>2.7</td>
<td>741</td>
<td>258</td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>393</td>
<td>95</td>
<td>49.5</td>
<td>10.8</td>
<td>640</td>
<td>168</td>
</tr>
</tbody>
</table>

**Table 5:4** Summaries of device parameters with the effect of LER
The $\pm 3\sigma_{I_{on}}$ for the 25 nm MOSFET is about $\pm 285 \ \mu A/\mu m$ (see table 5:4) for the RMS value of 3 nm. Furthermore, comparing the 35 and 25 nm gate lengths MOSFETs, $\pm 3\sigma_{I_{off}}$ variation in 25 nm transistor is approximately $\pm 1100 \ nA/\mu m$ which reflects a 60% higher off-current variation than in the 35 nm MOSFET.

**Figure 5:21** Linear scale of $I_{off}$ distribution in the 25 nm transistor with RMS=3 nm

**Figure 5:22** Logarithmic scale of $I_{off}$ distribution in the 25 nm transistor with RMS=3 nm.

**Figure 5:23** Histogram of drive current distribution for 25 nm transistor with RMS =3 nm.
5.2.4.2 The effect of correlation length ($\Lambda$)

As discussed earlier, one of the parameters used to characterise LER is the correlation length. In this section the effect of different values of $\Lambda$ (5 nm, 15 nm and 30 nm) on metallurgical and $p$-$n$ junctions and the device characteristics are presented. In all simulations here, the RMS value is kept 3 nm in order to separate the contribution of correlation length in characterizing the LER. The simulated devices have a 35 nm gate length. With the exception of different correlation lengths, the process corresponds to the calibrated 35 nm MOSFET.

![Gate shape and arsenic doping profile](image)

Figure 5:24 Gate shape and arsenic doping profile of three randomly generated MOSFETs using correlation lengths 30, 15, and 5nm. Row 1 shows top view of gate shape before As-implantation (the grey colour represents nitride spacer and red represents polysilicon gate material) and rows 2-3 depict the mapping of arsenic concentration (orange colour), 5 seconds and then after 15 seconds annealing respectively and row 4 shows distribution of arsenic and $p$-$n$ junctions after full process.
The gate shapes of three randomly generated MOSFETs using $\Lambda$ of 5, 15 and 30 nm are illustrated at the top of figure 5:24 (first row). The arsenic doping profile and the metallurgical $p$-$n$ junction immediately after the arsenic implantation, and at 5 seconds and after 15 seconds activation, are illustrated in rows 2 and 3 respectively.

The $p$-$n$ junction position follows almost precisely the shape of the gate edge in the case of large correlation lengths. The 3D implantation and the diffusion associated with the activation step smear almost completely short wavelength features at small correlation lengths.

![Image](image_url)

**Figure 5:25** $I_d$-$V_g$ Characteristics at high drain voltage of 850mV for the 35nm $n$-channel MOSFET, (a) log scale and (b) linear scale, with the effect of LER simulated for various correlation lengths

Figure 5:25a and b compare the $I_d$-$V_g$ characteristics among three devices with LER and the reference 35 nm transistor without LER. Exactly the same doping conditions and process flow simulation have been applied to all four devices. The degradation of subthreshold slope is noticeable in figure 5:25a while the drive current variation can be deduced from figure 5:25b. When the correlation length is 5 nm $I_{off}$ is almost an order of magnitude larger than the calibrated reference 35 nm device without LER.

Although $I_{on}$ increases with decreasing correlation length, the more rapid increase in $I_{off}$ results in an overall deterioration of device performance. The dependence of $I_{off}$ and
$I_{on}$ on $\Lambda$ are summarized in table 5:5. The simulation results presented in this sub-section do not represent statistical measures of ensembles of devices (like the mean value and the standard deviation). Curves shown in figures 5:25a and 5:25b represent four individual devices. However, the results and trends of variations reasonably reflect the impact of different values of correlation length ($\Lambda$) on intrinsic device parameters.

<table>
<thead>
<tr>
<th>$\Lambda$ [nm]</th>
<th>$I_{on}$ [$\mu$A/$\mu$m]</th>
<th>$I_{off}$ [nA/$\mu$m]</th>
<th>$\frac{I_{on}}{I_{off}}$</th>
<th>$\Delta I_{on}$ [$\mu$A/$\mu$m]</th>
<th>$\Delta I_{off}$ [nA/$\mu$m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>836</td>
<td>1830</td>
<td>456</td>
<td>110</td>
<td>1730</td>
</tr>
<tr>
<td>20</td>
<td>794</td>
<td>1450</td>
<td>547</td>
<td>68</td>
<td>1350</td>
</tr>
<tr>
<td>30</td>
<td>751</td>
<td>330</td>
<td>2275</td>
<td>25</td>
<td>230</td>
</tr>
<tr>
<td>No LER</td>
<td>726</td>
<td>100</td>
<td>7255</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5:5 Off-current and on-current in 35 nm n-MOSFET with LER of different correlation lengths

### 5.3 “Deterministic” LER simulation

In order to understand better the effect of LER/LWR on transistor performance and IPF we have simulated different specific patterns reflecting generic roughness configurations, which are referred in this work as “deterministic” LER. Although these specific geometric gate patterns do not represent the overall property of randomly generated rough edges, they can assist in investigating the effect of asymmetric gate edges on the $I$-$V$ characteristics and the local distribution of dopants near the p-n junctions. It is also possible to determine which of the two irregular edges, source or drain, has the larger impact on device properties. In this section, the investigations of eight different combinations of geometric gate patterns are presented.
5.3.1 Local doping variation

The variation of gate length along the width of the device will affect the local doping distribution, in particular the pattern of the $p$-$n$ junctions. Ideally the doping concentration will have the same profile everywhere along the width of the channel. This will result in a uniform effective channel length along the channel width. The uniformity results in uniform $V_T$, $I_{on}$ and $I_{off}$. However, the presence of LER or LWR changes this idealistic picture of devices.

To highlight the impact of the gate pattern on the doping distribution, we examine the doping profiles of different positions along the channel width. One of the 3D device structures with a particular gate obtained from the process simulation was partitioned into four parts as illustrated in figure 5:26. The positions are selected where the doping profile along the channel was changed significantly as a result of the gate shapes changes. A vertical slice in those positions gives a corresponding 2-D image on $x$-$y$ plane, from which the 1-D doping profiles near the interface are extracted.

Images of the 2-D doping profiles for arsenic in source and drain regions and the net doping profile in the device for all four different positions are shown in column one and two of figure 5:26. The cut plane positions are marked on the image that shows the top view of the gate. The profiles for both arsenic and net doping are not symmetric with respect to the middle of the channel.

For example, the image for the 2-D profile in row one of figure 5:26 depicts fewer dopants near the source junction than the drain junction which indicates that the doping distribution depends strongly on the 2D shape of the gate edge. The roughness of the gate edge, shown in figure 5:26b is passed onto the $p$-$n$ junction during the implantation process but also smoothed by the high temperature annealing steps.
Table 5.26: 2-D doping profiles of arsenic dopants (first column, which is represented by red colour) and net doping profiles (second column where the blue colour represents all p-type red represents all n-type dopants) in different positions along the device width. The figures in rows correspond to the doping profiles in the marked cut planes position (a), p-n junction (b), and the 3-D doping profile of arsenic is shown in (c). The net doping is usually defined as the difference between the acceptors ($N_a$, like boron and indium) and donors ($N_d$, like arsenic) dopants in the device.
The 1-D profiles of all dopants in the 4 marked positions on figure 5:26a are shown in figure (5:27). Although the net doping concentration in the middle of the channel is almost the same for all four positions, the distance between the metallic $p$-$n$ junctions varies significantly from position to position along the gate width. In addition to the effective channel width variation, the doping profile also becomes asymmetric in the neighbourhood of the source/drain junctions.

Figure 5:27 1-D doping profiles in the Si/SiO$_2$ interface along the channel of the same device but at different position along the gate width as shown in figure 5:26a.
5.3.2 The effect Deterministic LER on I-V characteristics

The eight specific geometric gate patterns which are employed to study the deterministic LER are shown in figure 5:28. The first row depicts the top view of the gate and the second row depicts the p-n junction of the corresponding geometric patterns after full process simulations.

Figure 5:28 Basic geometric elements of LER

![Basic geometric elements of LER](image)

Figure 5:29 Id-Vg characteristics of the basic geometric gate patterns shown in figure 5:23

![Id-Vg characteristics of the basic geometric gate patterns](image)
Device D7 has the worst $I_{off} = 1160 \text{nA/\mu m}$, which is more than eleven times the off-current of the original prototype 35 nm MOSFET (100 nA/\mu m). This is because the channel length in the middle of the device is greatly reduced, which results in an overall reduction of the effective channel length and as a consequence there is an increase in current. Another interesting observation is that it does not make a significant difference whether the defect is at the source or at the drain edge of the gate. For example the defects in devices D1 and D3 are mirror images, and their impact on the drain current is almost the same. The same result holds for the other two pairs of mirror devices, (D2, D4) and (D5, D6).

5.4 Chapter summary

Intrinsic parameter fluctuations in decanano MOSFETs induced by the number and position of discrete random dopants in the channel are statistically investigated, together with the impact of line edge roughness. Conventional transistors with gate lengths of 35 nm, 25 nm, 18 nm, 13 nm, and 9 nm, which correspond to all technology node generations in the 2001 edition of the ITRS, are used in the investigation.

In the case of random discrete dopants, the fluctuation of all three important MOSFET parameters namely $V_T, I_{on}$ and $I_{off}$ increase with decreasing transistor gate length. For example, the $\pm 3\sigma V_T$ variation for 18 nm n-channel MOSFET is $\pm 360\text{mV}$. In the 9nm device the $\pm 3\sigma V_T \sim \pm 510\text{mV}$, which is unacceptably high bearing in mind that the supply voltage for the corresponding technology node is expected to be $V_{dd} \sim 500$-600mV. Similarly, the fluctuations in $I_{off}$ and $I_{on}$ are increasing with the decreasing gate length. $\sigma I_{off}$ for 18 nm n-channel MOSFET reaches 220 nA/\mu m, which is 180% more than the standard deviation in the prototype 35 nm transistor. $\sigma I_{on}$ for the same device is $\sim 45\mu A/\mu m$.

LER which is rapidly becoming one of the major sources of IPF has also been investigated. Different values of RMS and correlation lengths have been used to investigate their impact on the MOSFET parameters. The threshold voltage fluctuation in the 35 nm MOSFET as a result of LER is $\sigma V_T = 26.5 \text{mV}$. This increases significantly to $\sigma V_T = 45.5 \text{mV}$, in the 25 nm MOSFET, an increase of 72%. The corresponding $\sigma I_{off}$ for the same devices are 157nA/\mu m and 393nA/\mu m respectively. However, in the 35 nm or 25 nm MOSFETs, random discrete dopants remain the dominant source of IPF.


5.5 Chapter references


