Laterally Non-Uniform Doping Profiles in MOSFETs: Modeling and Analysis

by

Keith M. Jackson
B. S. E. Princeton University (1994)

Submitted to the
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 1996

Signature of Author_____
Department of Electrical Engineering and Computer Science, August, 1996

Certified by_____
Dimitri A. Antoniadis
Professor of Electrical Engineering
Thesis Supervisor

Accepted by_____
F. R. Morgenthaler
Professor of Electrical Engineering
Chair, Departmental Committee on Graduate Students

Laterally Non-Uniform Doping Profiles in MOSFETs: Modeling and Analysis

by

Keith M. Jackson

Submitted to the Department of Electrical Engineering and Computer Science
on August 23, 1996, in partial fulfillment of the
requirements for the degree of Master of Science in
Electrical Engineering

Abstract

As MOSFET dimensions continue to shrink, the promise of further optimizing device performance and the problem of the Reverse Short Channel Effect point to the need for an understanding of the effect of laterally non-uniform profiles on device output characteristics. A semi-analytical model is developed to calculate the current of a device with a laterally non-uniform channel profile, from which a threshold voltage is extracted. The model is tuned to match the results of a two-dimensional numerical simulator. Using this model and the same numerical simulator, the Reverse Short Channel Effect is shown to be caused by a potential barrier at the source. The model is used to explore the relationship between the shape of the lateral doping profile and the shape of the resulting threshold voltage versus channel length curve. By closely fitting data from real devices exhibiting a Reverse Short Channel Effect, the model is able to extract lateral doping profiles for these devices.

Thesis Supervisor: Dimitri A. Antoniadis
Title: Professor of Electrical Engineering

Acknowledgments

I would like to thank Professor Dimitri Antoniadis for his technical insight, patience, and support throughout this project. My office-mates Andy Wei and Melanie Sherony, I would like to thank for their discussions both fun and technical. I want to thank Mark Armstrong, our UNIX and MATLAB guru, for answering all of my questions and helping with technical issues. I would like to thank my fellow graduate students in the research group and beyond, who have provided help and good times along the way. Thanks also to Kevin Meyer-Golden who provided help in administrative matters and a refreshing view into life outside of MIT.

A special thanks goes out to Digital Semiconductor (D.S.), a division of Digital Equipment Corporation for providing the devices used in this thesis. Thank you to Ted Equi for arranging the use of the wafers and the data in this thesis. Thank you also to Candy Scarelli who helped set speed records in moving my thesis and proposal through the approval process at D.S. .

To my parents and sister, I thank you for your support and understanding these past two years. A very special and heartfelt thanks goes to my wife Stacy who provided the boundless love and support that kept me going in a stressful time.

Project support for this thesis was provided by ONR Grant # N00014-95-1-1297 and by a National Science Foundation Graduate Research Fellowship.

Table of Contents

List of Figures

List of Tables

Chapter One

Introduction

MOSFET design has historically centered around making device dimensions progressively smaller in order to increase the speed, the packing density, and the complexity of circuits. As device sizes have shrunk, more precise control of the vertical doping profile has been provided by such fabrication innovations as ion implantation and rapid thermal annealing. Beyond just vertical doping, angled implants are now used to control the source and drain depletion regions. At the same time, MOSFET dimensions have shrunk to the point where controlling the shapes of the ion implants, after diffusion, is critical to maximizing device performance. Because of the push to smaller lengths, an understanding of the effects of a lateral doping profile in the channel is necessary.

This thesis explores the effects of a laterally non-uniform doping distribution on the threshold voltage of a device by focusing on the Reverse Short Channel Effect (RSCE). This effect causes the threshold voltage (V_{th}) to rise as the gate length (L) shrinks because of the presence of laterally non-uniform channel doping or non-uniform oxide charge. This thesis explores the physical connection between a non-uniform lateral doping profile and an increasing V_{th} with length. Beyond this, a model is derived and used to identify trends in how different shapes of extra doping relate to different shapes of V_{th} vs. L curves. Finally, the model's usefulness in extracting a guess at the lateral doping profiles in real devices is shown by fitting it to threshold data from real devices exhibiting RSCE. The discussion and

modeling in this work considers only n-channel MOSFETs, although the results should also apply to p-channel MOSFETs.

Chapter Two reviews the threshold voltage of a MOSFET and the current understanding of processing causes of the Reverse Short Channel Effect. Chapter Three reviews the existing models for the RSCE (V_{th} vs. L), then designs a new model that aims to be more exact and provide physical insight. Chapter Four explores the connection between a non-lateral profile and a rise in threshold voltage, then looks at the relationship between the shape of the extra doping profile and the shape of the V_{th} vs. L curves. Chapter Five concludes the thesis with a summary and suggestions for future work.

Chapter Two

Background on the Reverse Short Channel Effect

Setting the stage for modeling the Reverse Short Channel Effect (RSCE), this chapter pulls together the background drawn upon in later chapters. Starting with a brief overview of the relevant basics of a MOSFET, this chapter also covers how the Reverse Short Channel Effect (RSCE) affects the output parameters of a MOSFET, and concludes with a discussion of probable lateral doping profiles for use in modeling the RSCE.

2.1 MOSFET Background

The Reverse Short Channel Effect (RSCE) is characterized by the threshold voltage of a MOSFET increasing as the length decreases. As such, it is important to examine the equation for the threshold voltage of a MOSFET[1]:

$$V_{th} = V_{fb} + 2\phi_f + \frac{Q_B}{C_{ox}} \qquad (2.1)$$

$$V_{fb} = -\frac{E_g}{2} - \phi_f - \frac{Q_{ox}}{C_{ox}}$$

$$\phi_f = \phi_t \ln(\frac{N_b}{n_i}) \quad ; \quad \phi_t = \frac{kT}{q}$$

$$Q_B = N_b y_{dep} = \sqrt{2q\varepsilon_s N_b (2\phi_f - V_{BS})}$$

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$$

In the above equations, the flatband voltage (V_{fb}) is written for devices with degenerately doped poly as their gate material, and N_b is the doping of the channel. This equation shows that as N_b increases, the threshold voltage will rise, roughly proportionally to the square root of the N_b, in the Q_B/C_{ox} term. For real devices, at short channels the threshold voltage is not independent of L and it decreases as L decreases further. (figure 2-1)

This Short Channel Effect (SCE) is often modeled via charge sharing, where the source and drain depletion regions control some portion of the charge under the gate, effectively reducing the doping in the area. [1, p. 181] In terms of equation 2.1, the SCE is added by substituting Q_B' for Q_B, where Q_B' is less than Q_B and decreases with L, thus causing V_{th} to decrease or "roll-off".

Making the connection between the theory and real data, there are two common methods for extracting the threshold voltage from an I_D vs. V_{GS} curve at a low V_{DS} (~50mV). One definition of threshold voltage, called an extrapolated V_{th}, is obtained by placing a line tangent to the I_D-V_{GS} curve at the peak g_m (dI_D/dV_{GS}) and taking V_{th} to be the voltage at which the line extrapolates to zero current. The other method, called a constant current V_{th}, yields similar results and defines a specific current (commonly $10^{-7}\ W/L$) at which V_{th} is defined. The constant current method will be used in this work, since a constant current V_{th} can be easily extracted from the subthreshold current, the region for which the model in this work is derived.

2.2 RSCE: What it is, and why it is a problem

The Reverse Short Channel Effect (RSCE) is the phenomenon in which the threshold voltage of a device increases as the length of the device decreases (Figure 2-1). This behavior is the opposite of what is expected from the Short Channel Effect (SCE) discussed above, thus the name *Reverse* SCE.

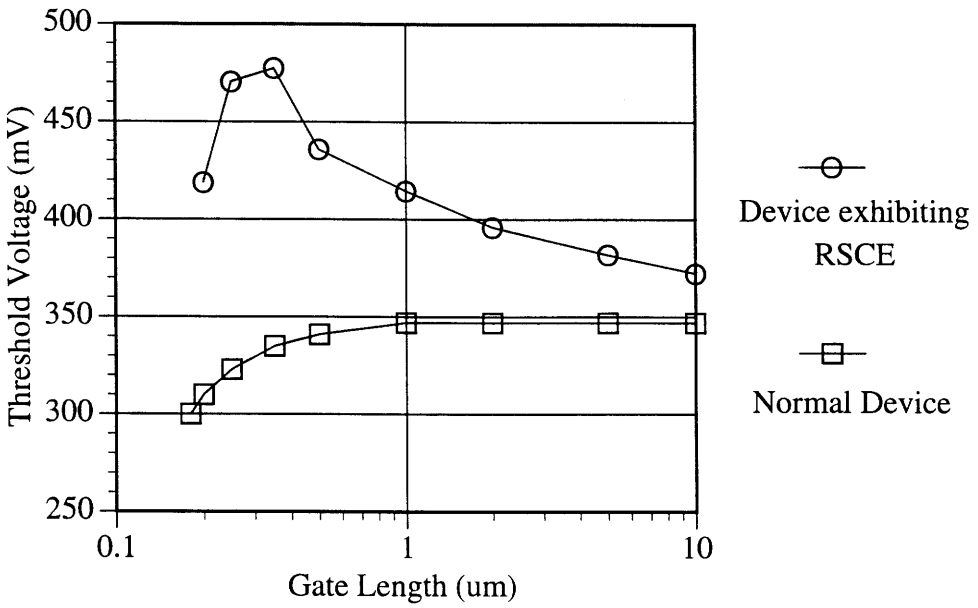

Figure 2-1: Example of the threshold voltage vs. length curves of a normal device and of one that exhibits RSCE.

This increase in V_{th} comes from extra doping or fixed oxide charge located near the source and drain. (figure 2-2). As the device's length shrinks, the V_{th} of the whole device rises. The reasons for this connection are examined in section 4-1. Note, however, that the V_{th} vs. L plot still decreases at some point as the SCE begins to dominate

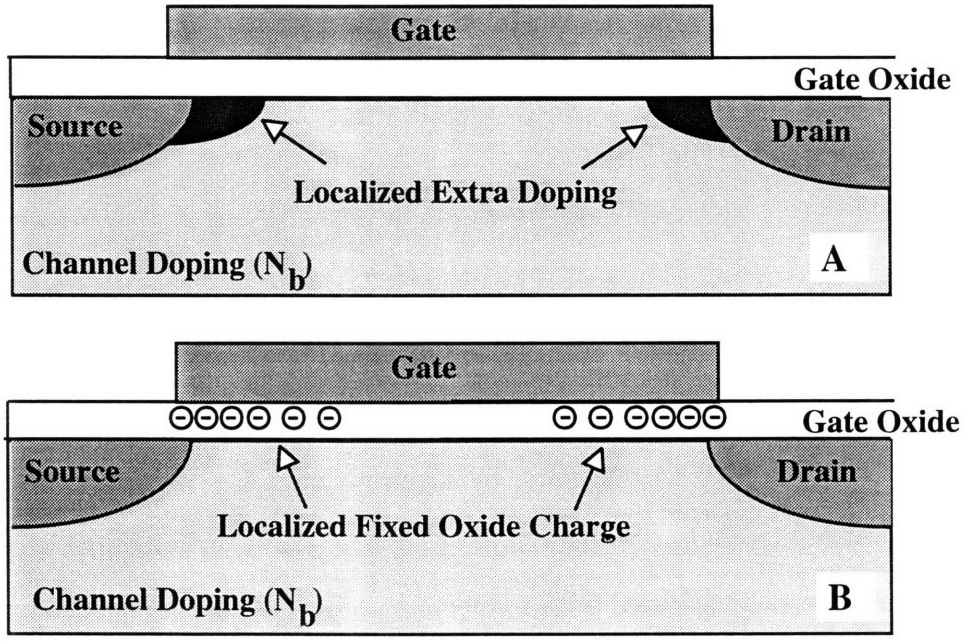

Figure 2-2: Illustration of the placement of extra doping (A) or extra fixed oxide charge (B) that causes the Reverse Short Channel Effect.

The rise in the threshold voltage caused by the RSCE adversely affects the performance of circuits designed for a specific channel length and V_{th}. Raising the V_{th} of a device lowers its drive current for a given power supply voltage. Since the output of a MOSFET almost always drives the gate of another MOSFET, this lower current results in a longer charging time of the load gate capacitor, so the second MOSFET will switch more slowly. At a more subtle level, a RSCE causes V_{th} to change much more versus length than anticipated. In any fabrication process, there will be a distribution of device lengths around the specified design length. If dV_{th}/dL is large around the specified/designed length, then this variation in length could translate into a large variation in V_{th} from device to device. This V_{th} variation could adversely affect the timing and switching levels of a circuit designed for devices of a specific threshold voltage and length.

2.3 Processing and Physical Causes

2.3.1 Enhanced Diffusion

From the standpoint of processing, the RSCE is caused by a high local concentration of point defects which affect the diffusion of dopants. Point defects are either vacancies (empty sites in the lattice) or interstitials (atoms that are not substitutional and that sit between the lattice atoms). Atoms move by acquiring enough energy (usually thermally) to move out of a lattice site to an interstitial position, move from an interstitial position or lattice site to a vacancy, or move from an interstitial position to a lattice site by pushing another atom into an interstitial position. In all three cases, increasing the concentration of point defects increases the rate of diffusion, and vice versa. [2]

At any temperature, there exists an equilibrium concentration of vacancies and interstitials. If an additional amount of interstitials are introduced, for instance, two things happen. One, the interstitials enhance the rate of diffusion of the atoms around them. Two, the interstitials decay back to equilibrium by recombining with vacancies, attaching to interfaces, and diffusing to areas of lower interstitial concentration. Thus, if during the processing of a device, a local concentration of interstitials were introduced, the diffusion of dopants would increase in that area, modifying the doping profile shape from that in other parts of the device. [2]

As a final suggestion that point defects are the root of the extra diffusion, it has been found that for devices fabricated on Silicon on Insulator (SOI), the enhanced diffusion for a given process is less than that for bulk silicon. Tsaimis et al. [3] monitored the movement of oxidation stacking faults caused by thermal oxidation and found they moved much less for SOI samples than for similarly prepared bulk silicon samples. This suggests that the extra oxide interface in the SOI acts as another sink for point defects, reducing the number available to cause the enhanced diffusion. Crowder et al. [4,5] found that similarly processed

SOI and Bulk wafers subjected to conditions that created enhanced diffusions and RSCE showed different effects. The SOI wafers consistently showed much less RSCE than the bulk wafers. Crowder attributes the lower amount of enhanced diffusion in SOI to the bottom oxide layer acting as a sink for the interstitials, lowering the number available to enhance the diffusion. [5]

2.3.2 Processing Steps

A variety of specific processes have been identified in the literature as introducing excess interstitials in the source and drain area and resulting in RSCE. These processes occur after the gate stack has been defined, at which point the source and drain areas are exposed while the channel is covered by the gate material. This sets up a situation in which the source and drain can be affected differently than the channel. In these cases, if excess interstitials are introduced into the source and drain, they can diffuse outward, causing localized enhanced diffusion in the area.

The first report of an anomalous increase in threshold voltage, as the RSCE is also called, is by Nishida and Onodera. [6] They observe an increase in threshold voltage with length for devices fabricated with heavy and deep boron channel implants. The magnitude of the roll-up increases as the dose of the implant is increased. Although they do not provide a physical explanation, their results are consistent with the slope of the retrograde doping profile providing a concentration gradient to move the channel dopants near the source and drain during a period of enhanced diffusion. Increasing the dose creates a steeper slope, which pushes more dopants to the surface near the source and drain.

Mazuré et al.[7-9] show that during the re-oxidation step after etching the gate stack, interstitials are injected into the channel from the oxidizing source and drain surfaces. This results in broadening of the retrograde channel profile, if such exists, near the source and drain edges, thus increasing the doping at the surface relative to the center of the channel.

They find that the amount of RSCE depends upon the sharpness of the retrograde channel profile - i.e., that the doping gradient is the driving force for the rate of movement of the impurities, while the interstitials facilitate the extra movement.

Rafferty et al. [10] examine the effects of damage caused by source/drain implants. In general, they find that in areas where the implants are not heavy enough to amorphize the silicon, the damage manifests itself in point defects. As before, these interstitials cause enhanced diffusion only near the source and drain, but in this case, also cause very peaked profiles at the oxide-silicon interface, as shown in Figure 2-3.

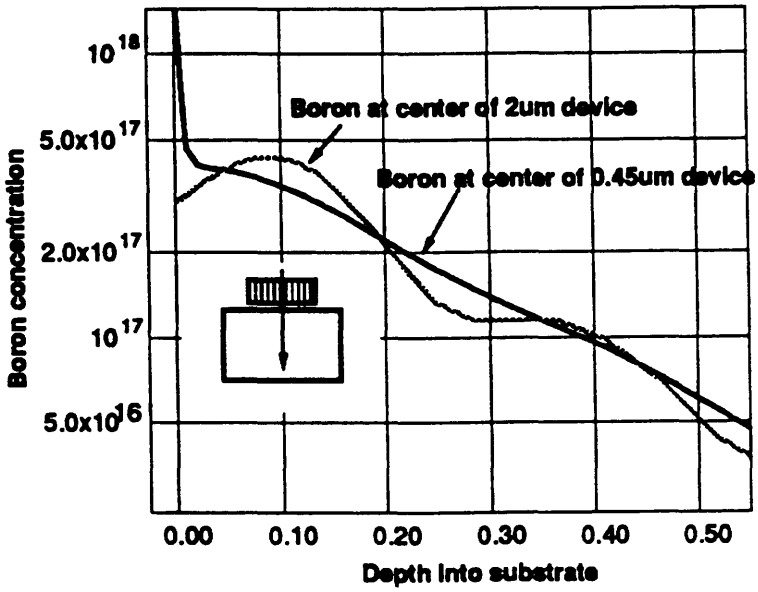

Figure 2-3 Simulated boron profiles at the center of short and long channel devices show the effects of enhanced diffusion. From Rafferty et al. [10]

The oxide-silicon interface acts as a sink for interstitials, so as the interstitials diffuse away from the source and drain, a concentration gradient of interstitials (high in implant areas and almost none at the oxide-silicon interface) is set up. Even if the channel doping is flat, this gradient of interstitials can push doping towards the oxide interface. Rafferty finds that the amount of RSCE depends on both the channel doping concentration (amount of doping available to move) and the dose and species of the source and drain implants. [10]

With heavier doses of source and drain implants or with higher channel doping, the amount of V_{th} roll-up increases.

Nishi et al. [11] creatively study the effect of the damage from source and drain implants on the magnitude of the Reverse Short Channel Effect. Using a set of devices of all the same gate length, they implant silicon into the source and drain at varying distances from the gate edge. They find (Figure 2-4) that the closer the extra silicon implant is to the gate edge, the further the threshold voltage rises, pointing to the cause of the enhanced diffusion as the point defects introduced from source and drain implantation.

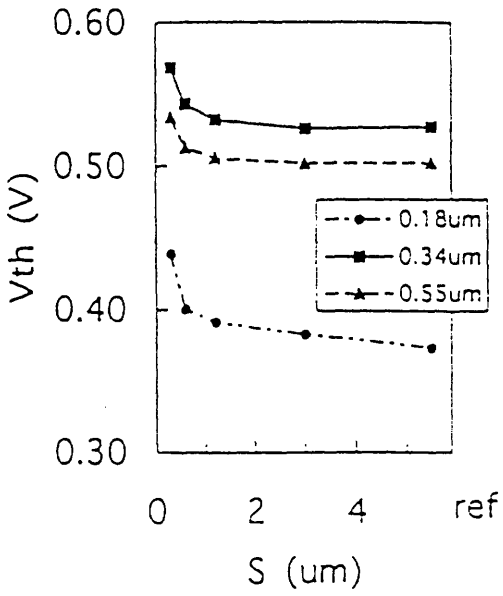

Figure 2-4 Decreasing S, the distance the silicon implant is from the source edge, results in the threshold voltage increasing for a given channel length. The reference device (ref) received no extra implant. From Nishi et al. [11].

From these experiments, they extract an enhanced diffusion coefficient that allows them to simulate the enhanced diffusion and match the V_{th} vs. L of a separate set of devices which exhibit RSCE without any additional silicon implants.

The use of a titanium silicide on the source and drain is found to cause a RSCE by Lu and Sung. [12] When the titanium and silicon react during the silicide formation, silicon moves into the titanium, leaving vacancies behind in the lattice. These vacancies, in a similar manner to interstitials, locally increase the diffusion of dopants. Lu and Sung vary the spacer width, changing the distance between the gate and the edge of the silicided area, and find that the amount of RSCE decreases as the spacer width is increased, pointing to a limited length over which these injected vacancies affect the diffusion.

Lutze and Venkatesan [13] expand upon the viewpoints above and identify the source and drain implant, among silicidation, re-oxidation, and source and drain implants, as the primary source of point defects that cause RSCE. As shown in figure 2-5 below, the amount of RSCE that occurs in similarly processed devices depends upon the channel profile. Even for a relatively flat profile at the surface (profile A), a RSCE occurs, confirming Rafferty's results.

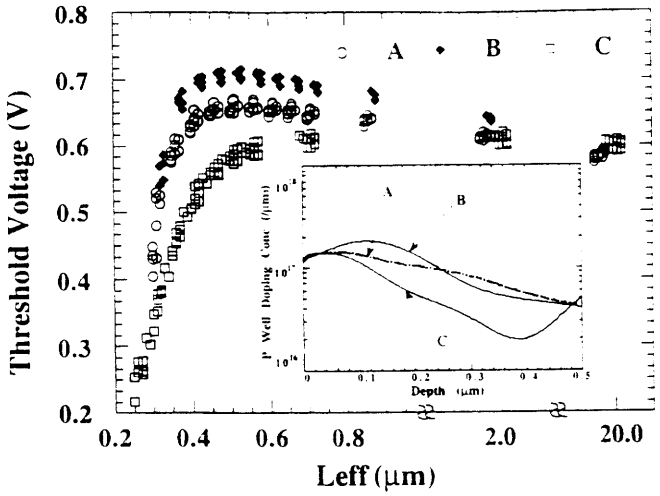

Figure 2-5 Threshold voltage vs. length for devices fabricated with different channel profiles. The inset shows the simulated channel profiles for long channel devices. From Lutze et al. [13]

Lutze also finds that an RF sputter clean before the titanium is deposited (for the silicide) significantly reduces the amount of RSCE. He suggests that vacancies, injected when the

surface is etched, recombine with the interstitials from the source and drain implants, thus reducing the amount of excess interstitials and consequently reducing the enhanced diffusion.

All of the above authors have taken the viewpoint that laterally non-uniform doping profiles are the cause of the observed RSCE. Jacobs et al. propose a model where the interstitials (originally near the source and drain) cause negative fixed oxide charge when they are captured at the gate oxide-silicon interface. [14] It is this laterally non-uniform distribution of oxide charge that causes the RSCE. Their main argument is that for the particular set of devices they examine, the back bias dependence of the threshold voltage is minimal. This is as expected for the initially flat channel profile, but counter to a redistribution of doping as the cause of the RSCE. Jacobs finds that with a process simulation model that allows the creation of negative fixed oxide charge when the excess interstitials are captured on the gate oxide, he is able to very closely match the V_{th} vs. L characteristics of his devices at different back biases.

In addition to the above papers which look at primary causes of RSCE, two others look at additional situations that lead to RSCE. Chang et al. [15] study the effect of the shape of the etched gate profile (caused by different etching gases in the dry etching of the gate stack) on threshold voltages of n-channel and p-channel devices shown in figures 2-6 and 2-7 below:

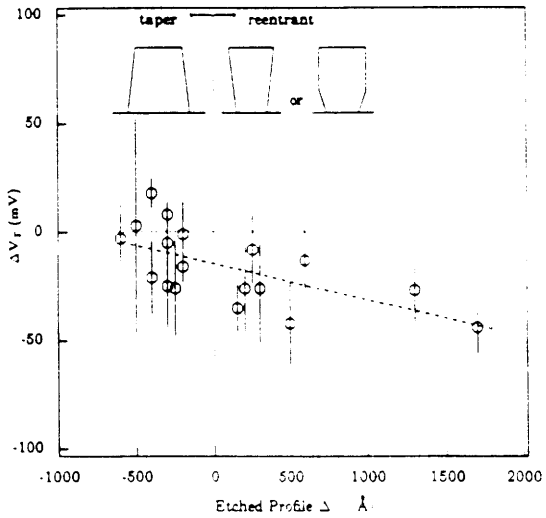

Figure 2-6 ΔV_{th} (V_{th} at 20μm - V_{th} at 1μm) vs. Δgate length due to etching profile for N-MOS devices. From Chang et al. [15]

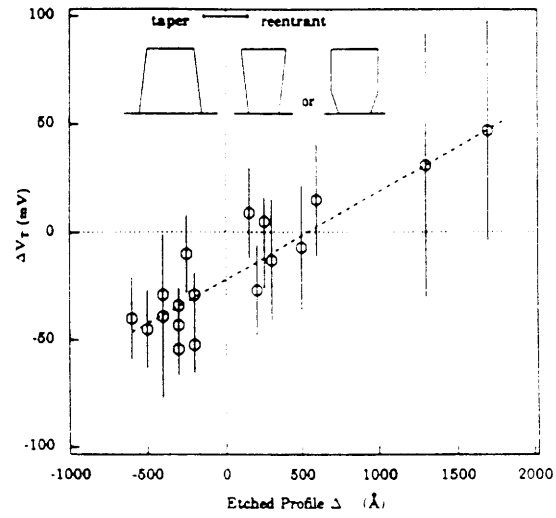

Figure 2-7 ΔV_{th} (V_{th} at 20μm - V_{th} at 1μm) vs. Δgate length due to the etching profile for P-MOS devices. From Chang et al. [15]

They find that the results for n-channel devices follow standard short channel effect theory (Figure 2-5). That is, tapered gate profiles create longer L_{eff} and less V_{th} roll-off while re-entrant profiles have smaller L_{eff} and greater V_{th} roll-off. For p-channel devices, they find the opposite (Figure 2-6). As the etched gate becomes re-entrant, the Reverse Short Channel Effect appears. The p-channel devices are buried channel type with a boron implant at the surface. Chang proposes that for far enough re-entrant profiles, a small area of the boron in the channel beyond the source and drain implant experiences oxidation-enhanced diffusion during the post-etching oxidation step. Thus, the primary cause is oxidation-enhanced diffusion of the boron in the channel, but its magnitude is controlled by how far the re-entrant profile reaches beyond the source and drain implants.

Kalnitsky et al. look at the effect of poly-depletion on the Reverse Short Channel Effect. [16] Studying a set of devices with the Reverse Short Channel Effect, they vary the poly-doping separately from the source and drain doping. They achieve this by doping the poly before etching the gate stack. They find that the lower the doping of the poly, the lower the inversion capacitance measured, due to more poly-depletion. At the same time, for devices with the same processing, the lower the poly-doping, the more the Reverse Short Channel Effect. This is summarized in Figure 2-8 below: From the perspective of extra doping as the cause of the RSCE, this effect makes sense. If C_{ox} decreases due to poly-depletion, then any increase in Q_b from the extra doping is magnified because it is divided by a reduced C_{ox}, due to poly-depletion, in the threshold voltage equation (2.1).

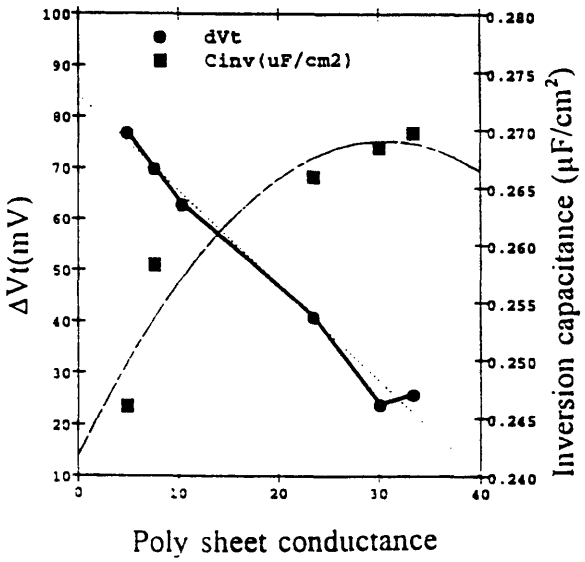

Figure 2-8 Measured V_{th} roll-up (solid dots, left axis) and corresponding inversion capacitance (solid squares, right axis) as a function of poly sheet conductance (1000/ρ). From Kalnitsky et al. [16]

The model that is used in this work assumes that the RSCE is caused by changes in the doping distribution, an assumption supported by the large majority of papers in the literature. Lutze's plot of the effects of different channel profiles on the amount of RSCE strongly points to doping as the primary cause for RSCE. (Figure 2-5) This thesis focuses on modeling the effects of the final doping distributions and does not consider how the particular doping distribution formed. As such, Crowder's, Harafi et al.'s, and Chung et al.'s work provide a nice transition from processing causes of the RSCE to the general shapes of the resulting lateral doping profiles.

2.4 Motivation for the Doping Distribution Shape

In his Ph.D. thesis, Crowder models the effects of transient enhanced diffusion via modifying the SUPREM code and the model parameters to match a set of devices designed to exhibit a RSCE. [5] Included in his study are a set of process simulations for a given set

of processing conditions that generate the RSCE. Lateral doping profiles 100 Å below the oxide surface are shown for a series of lengths (Figure 2-9) along with the corresponding V_{th} vs. L curves. (Figure 2-10) Figure 2-9 shows doping profiles for two different orders of thermal processes which give different amount of enhanced diffusion, but similar shapes. The shape of the profiles shown in figure 2-9 looks like an exponential drop-off of extra doping down to the base doping level, which is the doping profile selected for the model derived in this work. Note also that the doping in the center of the channel increases steadily as the channel length shrinks, but that it never exceeds the level of extra doping at the edges. It only becomes a flat profile at that level. This is also included in the behavior of the extra doping model used in this thesis.

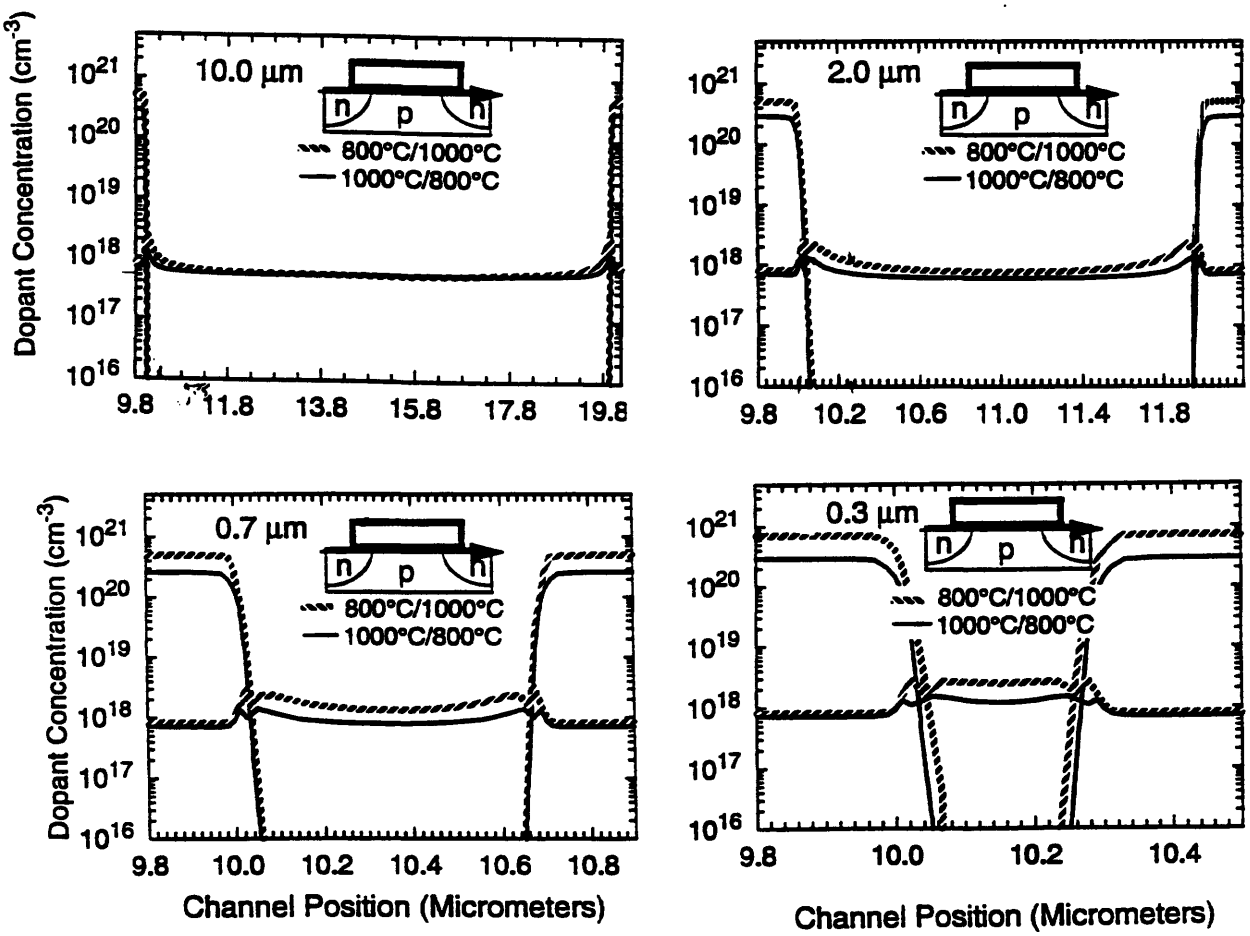

Figure 2-9 Simulated boron and arsenic lateral profiles at a depth of 100Å after furnace annealing at 800°C for 28 minutes and rapid thermal annealing at 1000°C for 10 seconds or the reverse anneal. From Crowder [5]

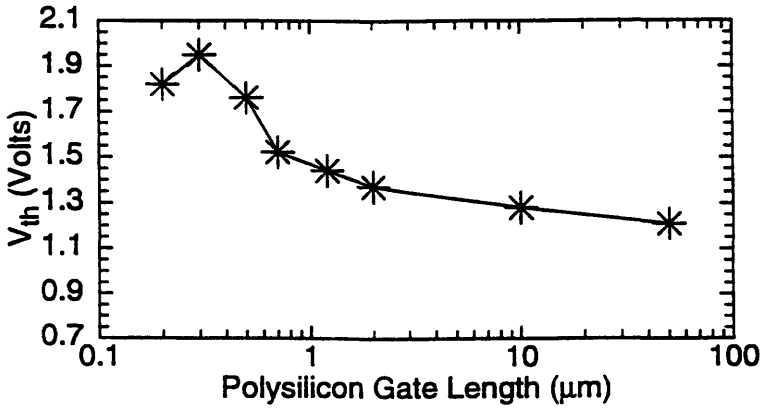

Figure 2-10 Resulting simulated threshold voltages of the devices that received the 800°C then 1000°C annealing from Figure 2-8. From Crowder [5]

Another suggestion for a doping profile shape comes from Hanafi et al. who use a two-dimensional numerical simulator to look at possible shapes of doping distributions causing RSCE. [17] They use a block model for the doping and assume that the source of the extra doping is the center of the channel and that some of the extra doping is sucked into the source and drain. Their lateral channel distribution looks like: (Figure 2-11)

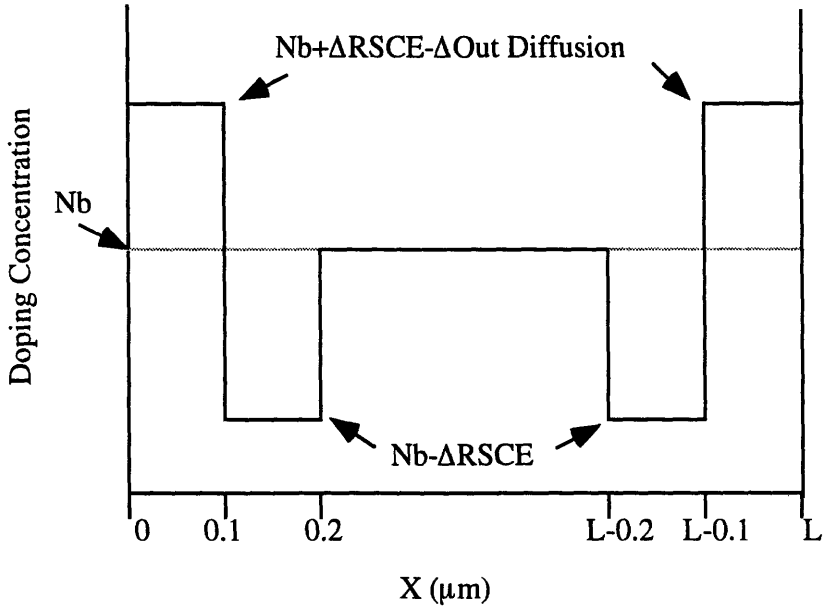

Figure 2-11 Lateral doping profile used by Hanafi et al. [17] in their device simulations.

Although such a profile works well in simplifying parameters for 2-D numerical simulation, the model in this work needs a continuous, integrable function, unlike these profiles. Also, the exponentially decaying profiles provide a better match in shape to Crowder's simulations and the basic form of a diffused profile.

Finally, Chung et al. have extracted lateral non-uniform profiles using charge pumping measurements. [18,19] Making the assumption that the RSCE was caused by doping, Chung converts the extracted number of interface traps to doping and comes up with

profiles as shown in Figure 2-11, which give a good match to the V_{th} vs. L curves for the same device when used in device simulations.

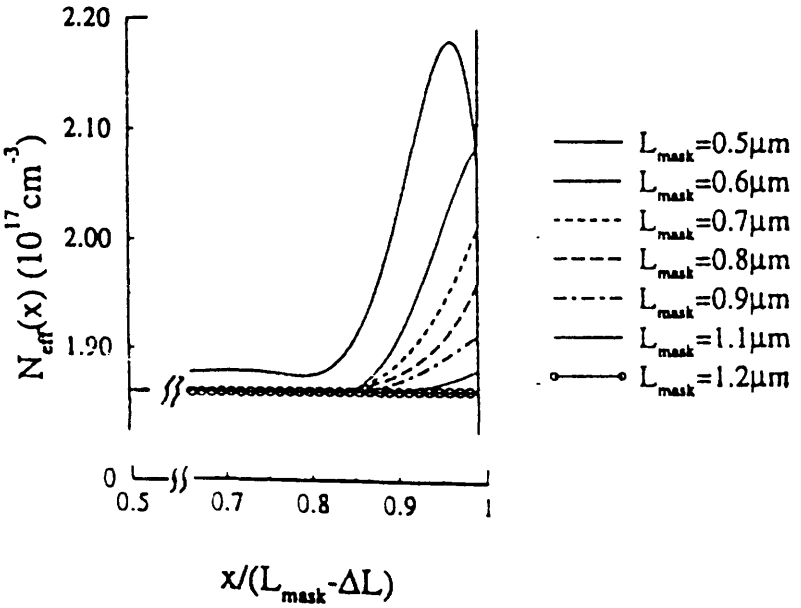

Figure 2-11 Extracted effective lateral non-uniform doping profile along the channel. From Chung et al. [18]

Although these measurements cannot determine the difference between interface traps and doping, the method does provide a way to gain a view of the possible lateral doping profile in the device. The distribution they extract again looks like an exponentially decaying function. However, they find that the magnitude and decay length change with channel length. The devices they examine are near the part of the distribution where the source and drain profiles overlap, which could cause such an effect.

Based upon this data and Crowder's simulations, the model developed in Chapter Three will assume an exponential tail-type extra doping distribution.

Chapter Three

Device Modeling

Creating a model for a device creates a framework which allows one to examine a device's reaction to different inputs without actually testing a real device. Additionally, such a framework will often provide greater insight into the device's operation. Any modeling usually starts from a basic theoretical description, focuses on the dominant phenomena, and often translates this into a set of mathematical equations. Given the typical tradeoffs between generality, accuracy, and speed, the types of models for the MOS transistors span a large range. Drawing upon a number of analytic models of the MOSFET, this chapter derives a model in order to examine the effect on device parameters of laterally non-uniform channel doping. This model aims to provide a level of detail in its calculation, beyond that of current models, that would be useful for designers or for process engineers trying to evaluate the lateral doping distribution in a MOSFET exhibiting a RSCE.

3.1 Introduction

MOSFET models of different levels of complexity can be roughly divided by their intended use. The most general and complicated models are used at the level of individual device design or analysis. Using basic relationships such as Poisson's equation and the

continuity equations for current, these simulators allow one to model the current outputs, internal electric fields, carrier distributions, etc. for an arbitrary device structure. Experimental knowledge is added via models that mimic such effects as mobility degradation with vertical fields and band gap narrowing from heavy doping. Such simulators take the basic equations and models, discretize them across a grid of points on the structure, and then numerically solve them. These types of simulations can take extensive computation time, but allow one to look at the effects of large changes in the device structure and doping profiles.

Analytic or semi-analytic models can be derived for a device structure that has only a few parameters that vary within a certain range. These types of models are used to design a device with defined output characteristics. These models consist of one or two equations, often numerically solvable, that govern the device's behavior. In this case, only the key parameters are focused on and included in the model. Often these parameters are identified within a certain region of device operation, or for certain types of geometries. Although a number of simplifying assumptions are made to reach a solution, the models are often quite accurate with a minimum of fitting parameters. This type of solution, because it can give specific equations for such things as surface potential, inversion charge, or current, allows one to gain physical intuition into the effects of different variables on the output characteristics of the device. Often a model ends up with transcendental equations for current or potential. Actually solving these equations is avoided by looking at specific device parameters (like V_{th}) that can be defined at a specific potential or current. These types of equations are often used in finding the V_{th} or subthreshold slope of devices at the design stage.

Another major use of device models is found in circuit design. Here the key issues are fast computational speed and accurate output from a single model for many device lengths and widths. Such models take purely analytically solvable equations that have been arrived at via major simplifications (again based upon theoretical and experimental knowledge of the device), then add empirical dependencies to make the equations better able

to mimic the real behavior of the device. For example, many models use the concept of charge sharing to add short channel behavior to long channel device equations. This approach often gives a final set of equations that have fitting parameters that must be extracted from real data. Such models can then be used to show how a complex circuit will operate when such devices are used as building blocks. Simplified versions of these same models are often the ones used in textbooks on device operation.

The models with the most relevance to the perspective taken in this thesis are the kind that allow device design given a specific device structure. Given that the model aims to predict both threshold voltage and subthreshold slope, it will focus on the equations for current in the subthreshold area of operation. The model will focus on solutions for the 2-D potential in the channel and then calculate a uniform current flow given such potential distributions.

Deriving the potential distribution in the channel is the starting point for any of the MOSFET models for threshold voltage. The most simplistic approach is to approximate the device as a one-dimensional vertical slice of the real device. This simplification can be justified using the gradual channel approximation which states that the vertical (y) term is much larger than the horizontal (x) term in a 2-D Poisson's equation. [20, p.168] Improving upon this, a form for the potential in y can be plugged into the 2-D Poisson's equation resulting in a second-order differential equation in the surface potential along the channel.

Ko examines this problem in the process of modeling electric fields near the drain of a MOSFET. [21] He starts by using the integral form of Gauss's law over a box next to the drain. Approximating the integral of E(x,y) with depth by multiplying an average E field by an average depletion depth, he derives a differential equation purely in terms of the surface potential and horizontal electric fields, which can be solved for the surface potential. Although effective for the hot carrier modeling for which Ko uses the model, the model does not suggest a particular form for the potential versus depth, which is necessary to calculate an accurate current.

Toyabe and Asai start from the idea that a uniformly doped channel gives a quadratic vertical potential and point out that this shape does not hold near the source and drain. [22] Since there are four natural boundary conditions in y (the value of the potential and its first derivative at the gate oxide-silicon interface and at the depletion edge), they use a cubic equation for $\phi(y)$. This equation for y is plugged back into the 2-D Poisson's equation and evaluated at y=0, yielding a second-order differential equation in $\phi_S(x)$, this time without any fitting parameters.

The model used in this thesis is based on Toyabe's approach with the modification of letting the channel doping vary along the channel.

3.2 Present RSCE Models

Three analytic models currently exist in the literature specifically for calculating the threshold voltage of devices exhibiting a RSCE. Of these, two are aimed at providing models that allow a length-independent description of such devices for compact modeling. The other paper examines the effects of a RSCE on the drain current and output conductance, modeling it as a potential barrier at the source and drain. All of the models adapt simple first-order equations for the threshold voltage and drain current by using an additional resistance, adjusted flat band voltage, or increased channel doping to model the RSCE.

Brut et al. picture the RSCE as caused by extra doping of a Gaussian-like shape at the edge of the channel. [23] Brut then finds an average doping of the channel ($N_{b,eff}$), integrating its doping from source to drain and dividing by the channel length. This gives an effective channel doping dependent upon channel length. Using the definition of threshold voltage shown in equation (2.1), he replaces the N_b with this $N_{b,eff}$ and adds short channel effects (SCE) via a charge sharing description, which reduces Q_B by a multiplicative factor:

$$V_{th} = V_{fb} + 2\phi_f + \frac{Q_B[1 - y_{dep}(\frac{\tan(\alpha)}{L})]}{C_{ox}} \qquad (3.1)$$

$$N_{b,eff}(L) = N_b + 2\frac{N_0 x_0}{L}\{1 - \exp[-(\frac{L_e}{y_0})^{\eta}]\}$$

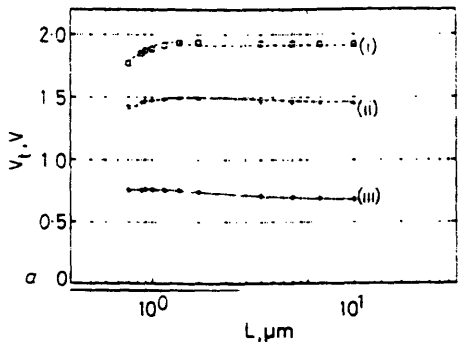

Figure 3-1 Experimental (symbols) and theoretical (lines) variations of threshold voltage with channel length for V_{BS} = 0 (iii), -2.5 (ii), -5.0 (i). From Brut et al. [23]

His model tracks actual V_{th} data closely versus both length and back bias. (Figure 3-1) Brut uses five fitting parameters in his model. Three of them, the magnitude (N_0), the exponent of the Gaussian (η), and the characteristic decay length (x_0) of the Gaussian of extra doping, are physically based. The other two are purely experimental parameters that allow him to fit the roll-off of the threshold voltage due to short channel effects (α and n).

Arora et al. take a similar approach but use negative fixed oxide charge as the cause of the RSCE. [20,24] Assuming an exponentially decreasing distribution of fixed oxide charge near the source and drain, Arora also takes an average, by integrating the oxide charge from source to drain, and then dividing by L. He then uses this effective Q_{ox} to adjust the flat band voltage (equation 2.1).

$$Q_{ox,eff} = \frac{2Q_{ox}x_0}{C_{ox}L}[1 - \exp(-\frac{L}{x_0})] \qquad (3.2)$$

Implementing the SCE differently, the final model has four fitting parameters. Similarly to Brut, the two parameters that are physically based are the magnitude (Q_{ox}) and characteristic decay length (x_o) of the extra Q_{ox}. The other two parameters allow Arora to fit the roll-off of the threshold voltage due to short channel effects.

Looking in general at the effect of a potential barrier at the source or drain, Hsu et al. use RSCE as an example to create their model. [25] Taking the standard equation for current over a barrier (eqn. 3.3), they linearize its dependence on the voltage applied across it, giving an effective conductance modulated by the gate voltage (eqn. 3.4):

$$I_B = I_\infty \exp(-\beta V_B)[\exp(\beta V_{APP}) - 1] \tag{3.3}$$

$$G = \beta I_\infty \exp(-\beta V_B) \tag{3.4}$$

Summing up voltage drops across the source, drain, channel, and barriers:

$$V_{DS} \approx \frac{2I}{\beta I_\infty \exp(-\beta V_B)} + \frac{I}{K(V_{GS} - V_{th})} + IR_{SD} \qquad (V_{DS} << V_{GS} - V_{th}) \tag{3.5}$$

$$I_{Linear} = \frac{KV_{DS}(V_{GS} - V_{th})}{1 + \frac{2K(V_{GS} - V_{th})}{\beta I_\infty \exp(-\beta V_B)} + KR_{SD}(V_{GS} - V_{th})} \tag{3.6}$$

where $K=\mu Cox(^W/_L)$ and R_{SD} is the total resistance of the source and drain regions. The V_{th} in the above equation is defined for the uniform part of the channel doping (equation 2.1), but is no longer the V_{th} of the device. This equation shows that the output current is reduced by the source/drain resistance and a resistance due to the barrier. Interestingly, the feature that differentiates the barrier resistance from the source/drain resistance is that it depends upon gate voltage. Empirically coming up with an equation for the relationship between the gate voltage and barrier height, Hsu shows a nice fit to an experimental current-voltage curve with

the model. They conclude that a potential barrier at the source and drain can severely degrade a MOSFET's performance.

Each of these models accurately describes the Reverse Short Channel Effects in terms of threshold voltage vs. channel length characteristics of a MOSFET. The model derived in the next section aims to be more exact in modeling the effects of a non-uniform lateral channel doping profile by avoiding taking an average of the channel doping. The model's results will be compared to Brut's method to gauge any improvements. Hsu's approach of viewing a potential barrier as the connection between the extra doping and a rising threshold voltage will be shown in the next chapter to be a very useful description.

3.3 Motivation for a New Model

Understanding the effects of laterally non-uniform channel doping profiles on MOSFET output characteristics, especially the threshold voltage, is important both from the perspective of designing a smaller device and inverse modeling devices that show abnormal threshold voltage versus length characteristics.

From the perspective of trying to debug a process exhibiting unintentional RSCE, gaining a sense of the doping profiles inside the channel can give a starting point for detailed inverse modeling of the device. Knowing the final doping distribution inside the device is a starting point for figuring out what part of the process could have caused the extra diffusion that resulted in a RSCE. At the same time, having a sense of how much of a laterally non-uniform doping is needed to significantly affect device output parameters is useful in designing a process. Given a sense of the acceptable limits, specific thermal budgets or cycles could be pinpointed and carefully simulated to see if they could contribute to RSCE.

From a designer's perspective, purposefully creating a non-uniform channel with more doping near the source and drain can help alleviate Short Channel Effects. In this case, it would be helpful to have a sense of the design space, in terms of shape and amount, for this

extra doping to reduce the SCE but not cause a RSCE. Both of these applications require a model which accurately correlates different lateral doping profiles with different V_{th} vs. L curves.

The model derived in the next section aims to allow a more exact calculation of the effects of a lateral doping profile on the subthreshold current and threshold voltage than current models provide. Additionally, by taking the least number of approximations possible, the model aims to help provide a better understanding of the connection between the extra doping and the resulting V_{th} vs. L curve that exhibits a RSCE.

3.4 Model Derivation

From reverse engineers' and designers' perspectives, the key criteria of any model are:

1) Realistic doping shapes
2) Reasonably accurate V_{th} vs. L curves
3) Flexibility
4) Rapid calculation
5) Provides insight into connections between input and output parameters

These are the basic guidelines for the derivation of the model.

The first part of any derivation is defining the intended focus and range of the model. The focus of this model is to explore the effects of changes in the shape of laterally non-uniform extra doping on MOSFET device output characteristics. As such, simplifications have been made to the device to allow this extra doping to be the focus. To create simple potential border conditions, the source and drain have been simplified to be abrupt and very deep. This will affect the SCE, but should have limited effect on the rest of the V_{th} vs. L curve. The channel has been simplified to be uniform although because of the use of ion implantation, channels are often modeled as a step function which is higher at the surface, then abruptly lower at some depth. This seemed to add an unnecessary complication since the back bias dependency (the device characteristic that depends most heavily on doping vs. depth) was not being considered. Additionally, the model assumes that the extra doping is just added - it is not removed from somewhere else in the channel. The general form of the extra doping is assumed to be constant with changes in channel length. From a processing standpoint, this assumes that the extra doping distribution originating from the source is unaffected by the drain and vice-versa.

In terms of actual device operation, the model will focus on the subthreshold region of operation. In this region, the doping level is much greater than the inversion charge, so the potential distribution is determined by the doping only. In this region, no lateral electric fields exist in the channel, so only current flowing by diffusion needs to be considered. The

gradual channel approximation (GCA), that vertical electric fields vary more quickly than horizontal electric fields, holds in this region and the solution to the vertical and horizontal potentials can be separated. Also, the potential distribution and current flow is assumed to be uniform with width so the problem is reduced to a two-dimensional one. Finally, the depletion approximation is used to give a simple boundary condition in y in the bulk of the devices.

The range of the model encompasses a variety of shapes of the extra doping at the edges. Taking the lead from the general form of diffused profiles (Gaussian or exponential) and from simulations by Crowder described in section 2.4, the profile is assumed to have the shape of an exponential tail. The equation used to describe this shape is:

$$N_{(x)} = N_b\{1 + \frac{A}{1+\exp[\frac{-L}{L_d}]}(\exp[\frac{-x}{L_d}] + \exp[\frac{x-L}{L_d}])\} \tag{3.7}$$

which looks like:

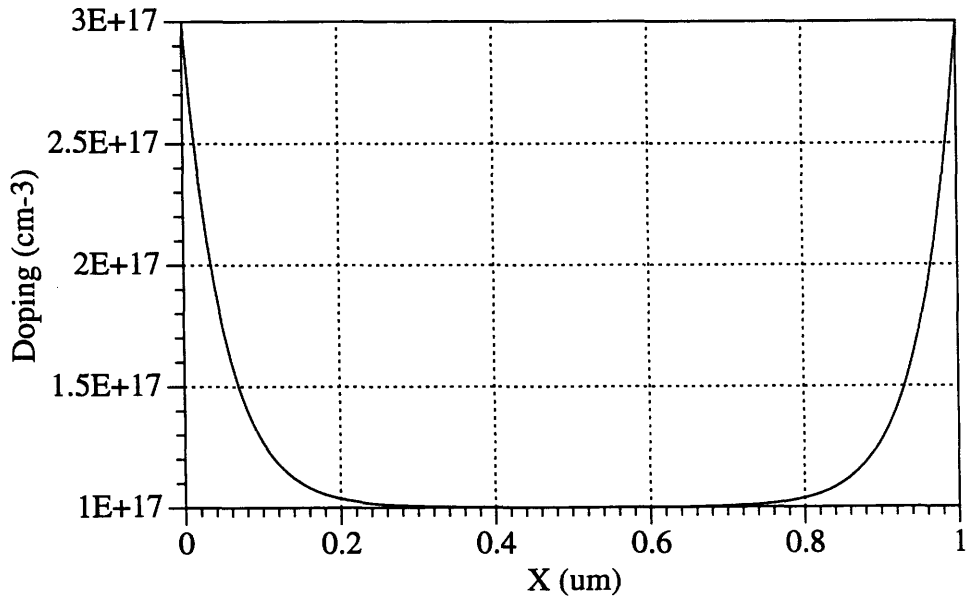

Figure 3-2 Plot of the doping profile used by the model in this thesis for a 1μm device with A=2, L_d=0.05, and N_b=1x10^{17}.

where N_b is the uniform base doping, AN_b is the magnitude of the extra doping at the source and drain edges {$(A+1)N_b$ is total doping at the source and drain edges}, and L_d is the exponential decay length of the extra doping towards N_b. The term that divides A keeps the maximum extra doping equal to $(A+1)N_b$ even as the distributions overlap as observed in Crowder's simulations [5].

The parameters here are: (1) the magnitude of the extra doping, and (2) the characteristic length of the exponential drop-off. In terms of the range of these parameters, the model will be used in this work to examine V_{th} in the 0.2-0.5 V range (corresponding to N_b in the $8x10^{16}$ to $3x10^{17}$ range), L_d in the 0-0.5 μm range, and A in the 0-4 range. This form (eqn 3.1) of the doping is easily integrable and is a reasonable representation for a heavily diffused profile. Crowder's 2-D results suggest a similar shape and show an A of ~2 and L_d in the 0.2-0.3 μm range. [5]

The model consists of two parts. First, a solution for the two dimensional (2-D) potential in the device, and then the calculation of a current from this 2-D potential. The solution for a potential draws upon the approaches of both Toyabe and Asai [22] and Ahmed et al. [26] Although both Ahmed and Toyabe make some large assumptions in the use of their models, their initial equations are carefully derived and useful for the following model.

The starting point for any electrostatic problem is setting up Poisson's equation and establishing boundary conditions. Poisson's equation in two dimensions is:

$$\frac{\partial^2\phi(x,y)}{\partial x^2}+\frac{\partial^2\phi(x,y)}{\partial y^2}=\frac{qN(x)}{\varepsilon_s} \tag{3.8}$$

Drawing a picture of the simplified device and labeling the boundary conditions:

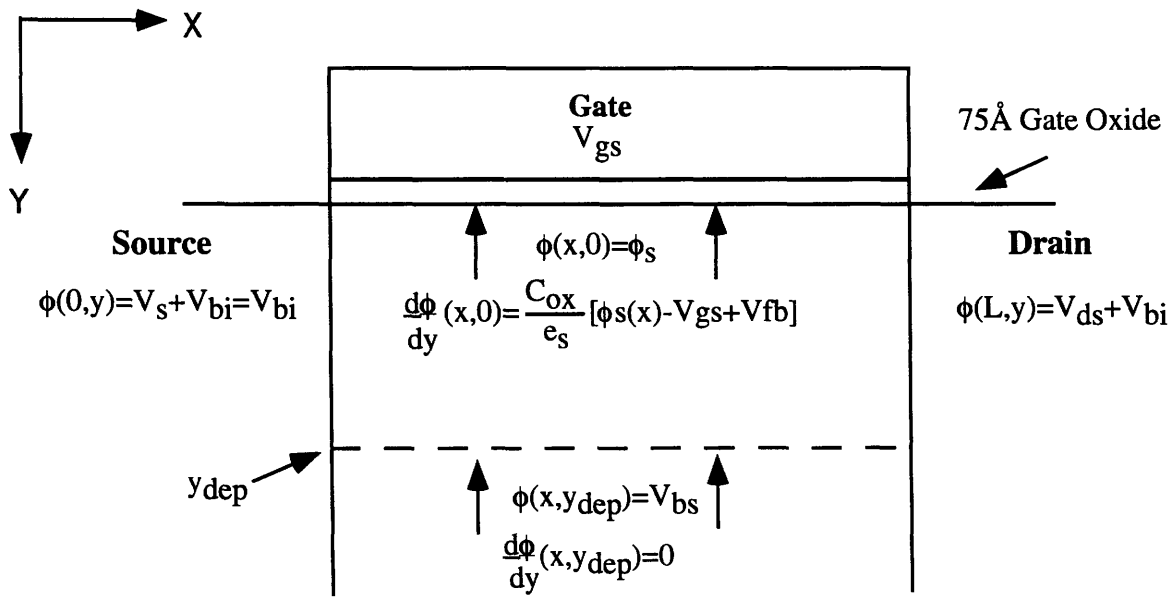

Figure 3-3 Simplified device structure used in deriving the model with the boundary conditions included.

The derivation and subsequent use of the model revolves around the surface potential ($\phi_s(x)$). As shown in Figure 3-4, $\phi_s(x)$ is defined as the band bending at the gate oxide silicon interface from the bulk potential level. As noted in the boundary conditions above, the potential is defined to be V_B at the edge of the depletion region.

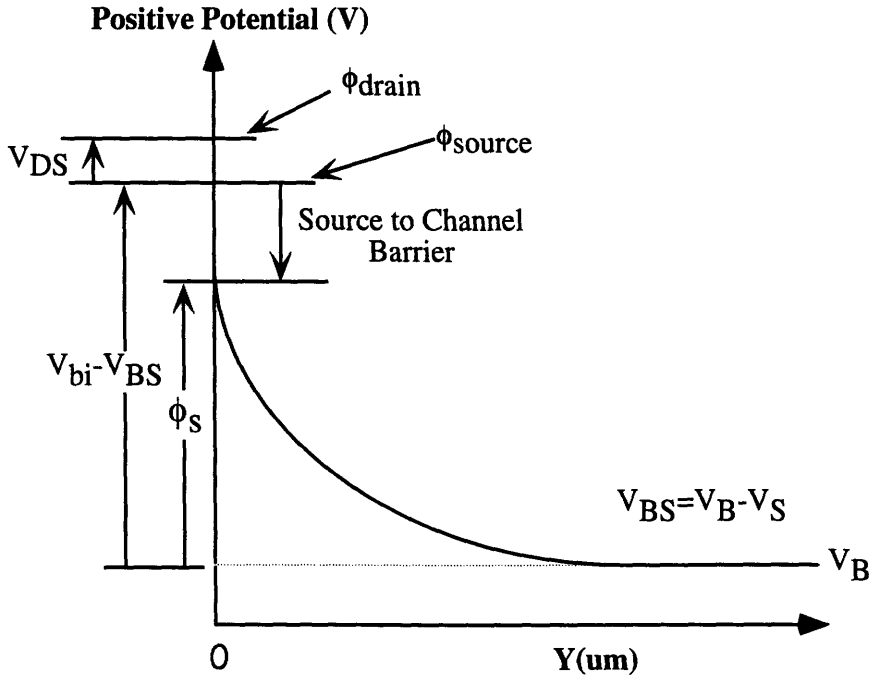

Figure 3-4 Schematic defining the reference point and magnitude of the surface potential in the channel (ϕ_s) and the source, drain, and back potentials.

The source is at (V_{bi} - V_{BS}) above the V_B reference level. The drain is at (V_{bi} - V_{BS} + V_{DS}) above the V_B reference level. In reference to the source and drain, ϕ_s is at a lower potential. As the gate voltage increases, ϕ_s increases; as the doping is increased, ϕ_s decreases.

Next, since the model will focus only on the subthreshold region of operation, the gate voltage is low enough that the channel is in weak inversion, where the number of free carriers is still much less than the doping level. Thus the charge on the right-hand side of Poisson's equation is only the doping. This same view can also be applied in the horizontal direction, so the built-in and applied potentials of the source and drain are dropped fully across depletion regions. Consequently, no potential is dropped across the entire channel and the current will flow only via the diffusion of carriers from the source to the drain. Using the gradual channel approximation (GCA):

$$\frac{\partial^2\phi(x,y)}{\partial x^2} << \frac{\partial^2\phi(x,y)}{\partial y^2} \tag{3.9}$$

Poisson's equation can be approximated as:

$$\frac{\partial^2\phi(x,y)}{\partial y^2} = \frac{qN(x)}{\varepsilon_s} \tag{3.10}$$

It is worth noting that the GCA breaks down near the source and drain in the depletion region where electric fields exist and change rapidly, especially with non-uniform extra doping added. However, except at very short channels where the depletion region reaches across the whole channel, the approximation has been found to be quite accurate.

The solution to equation 3.10 is a quadratic, but as Toyabe points out, this is a poor fit near the drain, so a cubic equation will be used. Also, since there are four boundary conditions, a cubic will work well. The solution then is:

$$\phi(y) = a_0 + a_1 y + a_2 y^2 + a_3 y^3 \tag{3.11}$$

$$a_0 = \phi_s(x)$$

$$a_1 = \frac{C_{ox}}{\varepsilon_s}[\phi_s - V_{GS} + V_{fb}]$$

$$a_2 = \frac{1}{y_{dep}^2(x)}\{-3(\phi_s(x) + V_{BS}) + y_{dep}(x)(2a_1)\}$$

$$a_3 = \frac{1}{y_{dep}^3(x)}\{2(\phi_s(x) + V_{BS}) + a_2 y_{dep}(x)\}$$

However, it is the variation of the potential in x that is the most interesting, so plugging y back into equation 3.8 and evaluating the resulting equation at y=0 (the surface) gives:

$$\frac{\partial^2 \phi_s(x)}{\partial x^2} - P^2 \phi_s(x) = \frac{qN(x)}{\varepsilon_s} + R \tag{3.12}$$

$$P^2 = \frac{1}{y_{dep}^2(x)}\left(\frac{4C_{ox}}{C_D} + 6\right)$$

$$R = \frac{1}{y_{dep}^2(x)}\left(6V_{BS} - \frac{4C_{ox}(V_{GS} - V_{fb})}{C_D}\right)$$

$$C_D = \frac{\varepsilon_s}{y_{dep}(x)}$$

Notice that P depends on y_{dep}, which will depend on x as the doping varies with x. This makes the differential equation unsolvable analytically, because the expression for $y_{dep}(x)$ depends on the square root of $\phi_S(x)$. Both Toyabe and Ahmed treat P as a constant with x, allowing them to solve the differential equation. Ahmed then lets P vary with L when he uses the final equation.

Proceeding with solving the differential equation, treating P as a constant, the homogeneous solution is of the form:

$$\phi_s(x) = C_1 \sinh(Px) + C_2 \cosh(Px) \tag{3.13}$$

The particular solution, following Ahmed's derivation [26] is:

$$\phi_s(x) = -\frac{R}{P^2} + e^{Px}\left[\int e^{-2Px}\left(\int N(x) e^{Px} dx\right) dx\right] \tag{3.14}$$

Using the doping distribution described by equation 3.1 and applying the boundary conditions at the source and drain from figure 3-3, the total solution becomes:

$$\phi_{s(x)} = -\frac{R}{P^2} - Q_{E(x)} + (V_{bi} + \frac{R}{P^2} + Q_{E(x=0)})\frac{\{\sinh(Px) + \sinh(P[L-x])\}}{\sinh(PL)} + (V_{DS})\frac{\cosh(Px)}{\sinh(PL)}$$

$$Q_{E(x)} = \frac{q}{\varepsilon_s} N_b \{\frac{1}{P^2} + \frac{A}{1+\exp[\frac{-L}{L_d}]} * [\frac{1}{P^2 - (\frac{1}{L_d^2})}](\exp[\frac{-x}{L_d}] + \exp[\frac{x-L}{L_d}])\} \quad (3.15)$$

This equation is an implicit one in $\phi_S(x)$ since y_{dep} also depends upon $\phi_S(x)$. The equation for y_{dep} for a uniform profile is:

$$y_{dep(x)} = \sqrt{\frac{2\varepsilon_s(\phi_{s(x)} - V_{BS})}{qN_{(x)}}} \quad (3.16)$$

These two equations (3.15,3.16) are solved numerically in MATLAB [28] given a good starting guess. From here, ϕ_S at each x point can be plugged back into the equation 3.11 for $\phi(y)$ to generate the 2-D potential. From this potential distribution, the diffusion current is calculated using a method developed by Greenfield and Dutton [27].

Greenfield and Dutton's approach [27] makes a number of reasonable assumptions:

1) Current flows only along the channel in the x direction
2) I_{DS} is constant in the channel:

   There are no sources or sinks of current from surface states, gate leakage, or impact ionization
3) Mobility is constant in all regions of the channel

Greenfield defines an inversion charge as:

$$n_{(x,y)} = \frac{n_i^2}{N_{(x)}} \exp\{\frac{q}{kT}[\phi_{(x,y)} - \phi_{Fn(x)}]\} \quad (3.17)$$

where ϕ_{Fn} is the quasi-fermi level for electrons. He writes the equation for current flow as:

$$J(x,y) = -q\mu_n n_{(x,y)} \frac{\partial \phi_{Fn}(x)}{\partial x} \hat{x} \tag{3.18}$$

Integrating this equation with depth and multiplying by the width gives I_{DS}.

$$I_{DS} = -q\mu_n W \frac{n_i^2}{N(x)} \exp[\frac{-\phi_{Fn}(x)}{\phi_t}](\frac{\partial \phi_{Fn}(x)}{\partial x}) \int_{y_{dep}}^{0} \exp[\frac{\phi(x,y)}{\phi_t}] dy \tag{3.19}$$

Since I_{DS} is constant in the channel (see assumption 2 above), multiplying both sides by dx and integrating from source to drain (0 to L for dx, V_S to V_D for $d\phi_{Fn}$) gives:

$$I_{DS} \int_0^L N(x) \{ \int_{y_{dep}}^{0} \exp[\frac{\phi(x,y)}{\phi_t}] dy \}^{-1} dx = -q\mu_n W n_i^2 \int_{V_S}^{V_D} \exp[\frac{-\phi_{Fn}(x)}{\phi_t}] \partial \phi_{Fn}(x) \tag{3.20}$$

$$I_{DS} = \frac{-q\phi_t \mu_n W n_i^2 [\exp(\frac{-Vs}{\phi_t})][1 - \exp(\frac{-V_{DS}}{\phi_t})]}{\{ \int_0^L N(x) \{ \int_{y_{dep}}^{0} \exp[\frac{\phi(x,y)}{\phi_t}] dy \}^{-1} dx \}} \tag{3.21}$$

The resulting equation gives current from a 2-D potential via two integrations. In the model, this is implemented in MATLAB as two numeric integrations, which take the majority of the evaluation time.

From the current vs. voltage characteristics (equation 3.21), one can derive device output parameters. In order to speed up the execution of the model, V_{th} and the subthreshold slope (SS) are extracted using a method similar to one suggested by Biesemans et al. [29,30] S, defined as $^1/_{SS}$, is:

$$S = \frac{50mV - 0mV}{\log_{10}[I_{(50mV)}] - \log_{10}[I_{(0mV)}]} \tag{3.22}$$

I_{off} is defined as I at V_{GS}=0. Finally, V_{th} is defined at a constant current of 10^{-7} W/L and extrapolated from I_{off}:

$$V_{th} = S\log_{10}\left(\frac{10^{-7}\,W/L}{I_{(0mV)}}\right) \tag{3.23}$$

3.5 Verification

Before using the model, its output needs to be compared to the more exact output of MEDICI,[31] a two-dimensional numerical device simulator. In order to compare the model and simulation directly, the same assumptions about the device structure used in the model are used in the simulated structure. The structure (like figure 3-3) has a very deep (0.5μm) and abrupt source and drain (rectangular-shaped) and the lateral doping profile is again uniform with depth. The source, drain, and gate doping, normally degenerately doped, are picked to give E_{Fn}=E_c with the Maxwell-Boltzman statistics the simulator employs. These simulations will allow an evaluation of the model's ability to calculate the output characteristics of the simplified device it should exactly match. Later in section 4.3, the model will be applied to fit the output characteristics of real devices with RSCE, evaluating its ability to model much more complex structures.

Using the simulator to model the insides of a device, the assumptions made in the model's derivation can be considered. Ignoring the mobile carriers when solving Poisson's equation required that the doping level be much greater than the electron concentration. Figure 3-5 shows the electron concentration at $V_{GS} = V_{th}$ for a base doping level of $1x10^{17}$. The amount of carriers is about 3 orders of magnitude lower, justifying the assumption. Also, in order to use the gradual channel approximation, the E field must change more rapidly with depth than along the channel. Figure 3-6 shows dE_x/dx and dE_y/dy at the

surface (y=0). Except near the source and drain, the approximation holds, as expected. Additionally, E_x in the channel was found to be zero, reinforcing the concept that there is no drift current, only diffusion current.

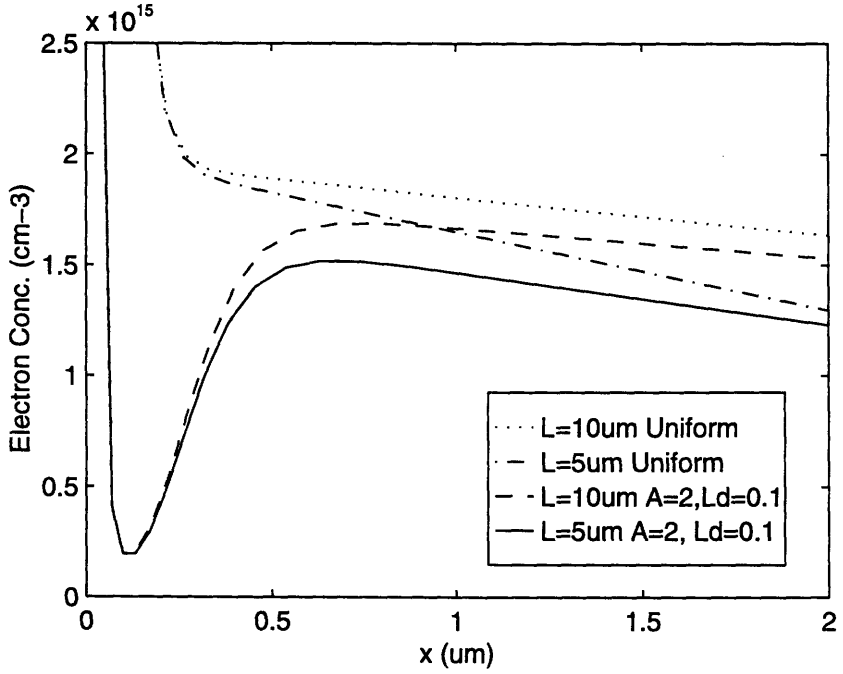

Figure 3-5 Electron concentration at the surface (y=0) in one half of a 10μm & 5μm device with A=0 & A=2, L_d=0.1, and N_b=1x10^{17}.

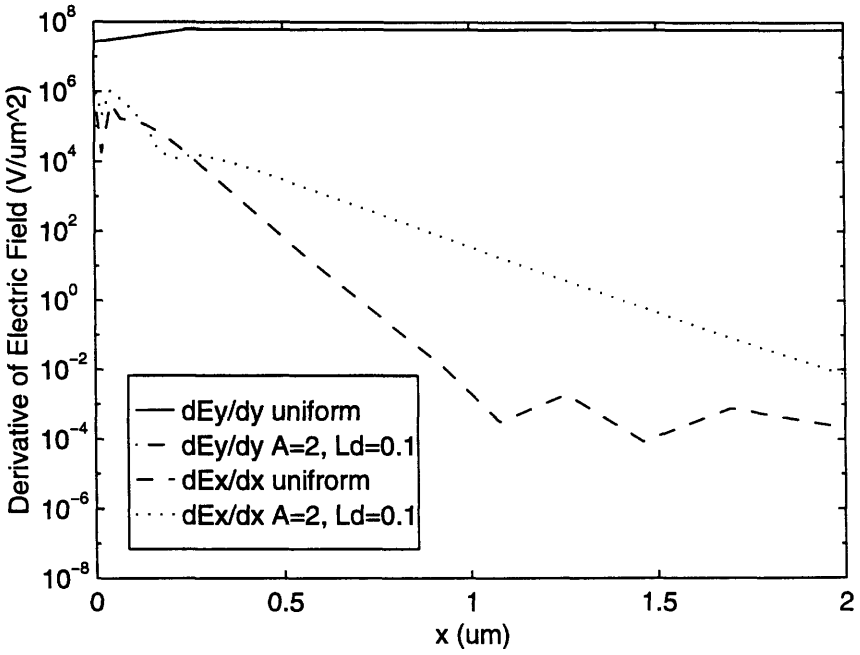

Figure 3-6 Verifying the gradient channel approximation, $dE_x/dx << dE_y/dy$.

The model is able to come quite close in shape and magnitude to the surface and vertical potentials that MEDICI predicts for the device. In order to make the potentials to match, two values had to be adjusted. First, the flat band voltage had to be increased from its expected value by 45 mV to get the long channel potentials to match at each V_{GS}. Second, the P used in the sinh and cosh terms (see equation 3.15) was multiplied by 0.4. Since $1/P$ is equivalent to a characteristic length at which the potential falls off from the source and drain, 0.4P lengthens this decay. The P calculated by the model without this factor gives too steep a decay of the potential at the source and drain. This separate adjustment of parameters for the constants of the sinh and cosh terms was also done by Toyabe and Asai [22], who first show that the model has the right shape of potential and then determine the coefficients for the equation by comparison to 2-D numerical simulations. Although for a perfect fit, this multiplier of P must vary with N_b and A, the value of 0.4 gives reasonably good results across the range of parameters considered here and will be fixed at this value.

For the rest of the modeling in this thesis, the multiplier of P and adjustment to the flat band voltage is not changed. Although better fits could be made if these parameters were varied, leaving them constant leaves the model with only physically-based adjustable parameters (N_b, A, and L_d) with which to fit a real device. Even with these parameters constant, quite good fits are found

In terms of V_{th} versus L curves, the match between the model and MEDICI is remarkably good. Figure 3-7b shows the comparison for $N_b=1x10^{17}$, A=2, L_d=0,0.1,0.3. In figure 3-7 the triangles represent the threshold voltages from MEDICI determined by equation 3.13 where the constant current V_{th} is found from an I_{off} and SS around V_{GS}=0. These come consistently close to the model's points and are also close to an actual constant current V_{th}. A similarly close fit is found for $N_b=8x10^{16}$, L_d=0,0.1,0.3 in Figure 3-7a. A larger difference is visible in the $3x10^{17}$ comparison (figure 3-7c). The values of the $3x10^{17}$ curves don't overlap very well because the V_{fb} was not adjusted correctly, but the ΔV_{th} vs. L, visible in the parallelism of the two sets of curves, is again very close. Increasing the V_{fb} an additional 25 mV gave curves that perfectly matched the MEDICI output, but the parameters were not changed, as explained above. The model seems to work well in the range set out for it, but becomes slightly less accurate as the extra doping amount increases. Overall, the model matches the MEDICI output very well, especially in terms of changes in V_{th} with L.

The derived model is used in Chapter 4 to calculate a threshold voltage for a variety of different shapes of extra doping, correlating them with the resulting V_{th} vs. L curves. Additionally, the model is used to help solidify a viewpoint on the connection between the extra doping and the RSCE V_{th} vs. L characteristic.

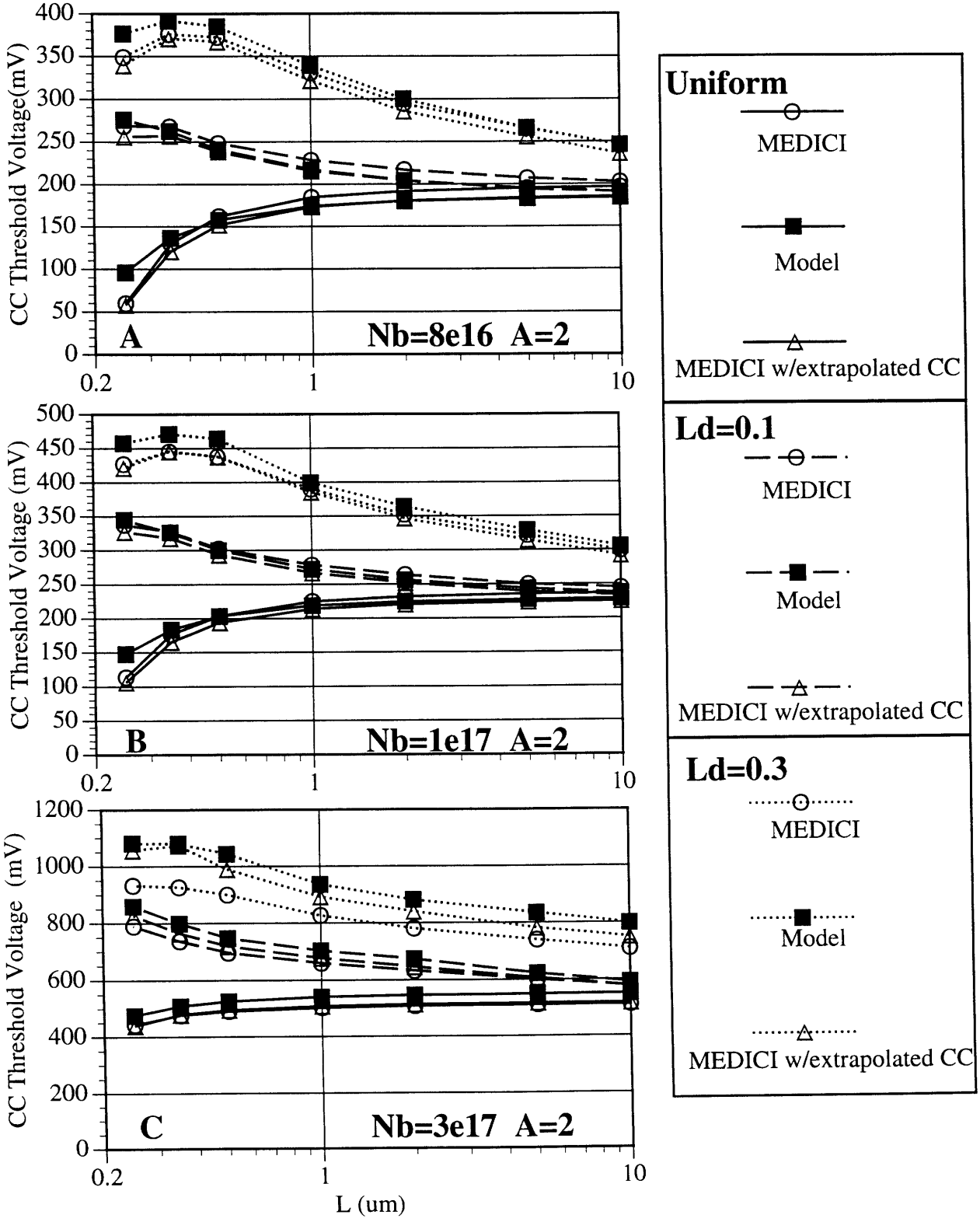

Figure 3-7 The match between the V_{th} vs. L curves calculated by the model and those calculated by MEDICI.

Chapter Four

Discussion and Analysis

Using both the model derived in the last chapter and the two-dimensional numerical device simulator MEDICI [31], this chapter explores the connection between a laterally non-uniform doping profile and the resulting rise in threshold voltages as the channel length (L) decreases. The model is then used to illuminate the relationship between the parameters of the doping profile and the resulting shape of the V_{th} vs. L curve. Using these dependencies, threshold voltage data from real devices exhibiting the Reverse Short Channel Effect (RSCE) is matched. These real devices, fabricated by Digital Semiconductor (D.S.), a division of Digital Equipment Corporation, are also used to provide real data for examples throughout the chapter. Finally, the model is compared to the other RSCE modeling approaches detailed in Chapter 3.

4.1 Connecting the Extra Doping to the Increased Threshold Voltage

Evidence from work by Rafferty [10], Lutze [13], and Crowder [5] points strongly to the existence of laterally non-uniform extra doping in the channel causing the Reverse Short Channel Effect (RSCE). At short channel lengths (<1 μm) where the source and drain doping distributions overlap and the doping is increasing everywhere in the channel, it makes

sense that the threshold voltage rises, since $V_{th} \propto \sqrt{N_b}$ for a uniformly doped channel (eqn 2.1). Yet, at long channels (10 μm) where the extra doping distributions only exist at the edges of the channel and a majority of the channel has no extra doping, the threshold voltage still rises as the channel length (L) decreases. This suggests that the connection between the long channel threshold voltage and the extra doping at the edges is different than the normal relationship (equation 2.1).

The extra doping at the edges causes the surface potential to be smaller in this area than in the rest of the channel. Since the doping at an edge does not change versus L, at an L long enough that the source and drain extra doping distributions do not overlap, the potential distribution around this area of extra doping (potential well area in figure 4-1) should be constant as L changes for V_{GS} and V_{DS} constant. Figure 4-1, which shows the surface potential as simulated by MEDICI, shows this to be the case.

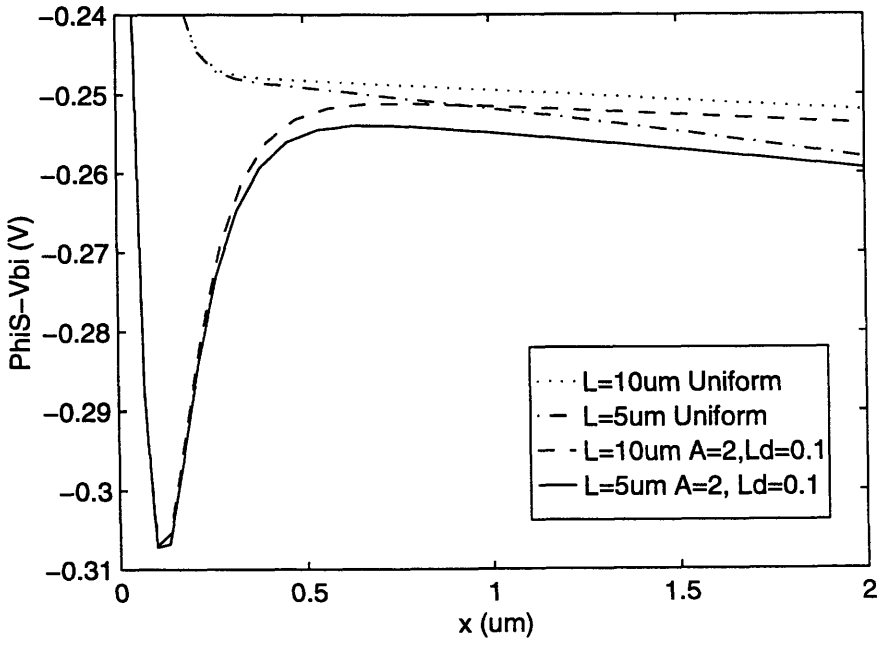

Figure 4-1 Simulated ϕ_s-V_{bi} near the source of a 5 μm and 10 μm $1x10^{17}$ base doping device with uniform profile or a non-uniform A=2, L_d=0.1 profile at V_{GS}=0.1.

The dip in potential at the edge of the channel, due to the extra doping, can be viewed as a barrier between the source and the rest of the channel. The equation for the current over a barrier is: [32, p.258]

$$I = WC\exp[\frac{-\phi_B}{\phi_t}](\exp[\frac{V_{AP}}{\phi_t}]-1) \tag{4.1}$$

where ϕ_B is the barrier height, V_{AP} is the voltage dropped across the barrier, W is the width of the device, and C is a constant. In the subthreshold region of operation, the number of carriers is far less than the doping, so the built-in potential of the p-n junctions and the source and drain voltages are completely dropped across depletion regions near the source and drain. The size of these depletion regions will not vary with channel length until short L. Thus, the voltage applied across the potential barrier will be assumed to be approximately constant with respect to channel length L and V_{DS}, set only by how the potential of the source and drain decay into the channel. In this case, the current that flows over the barrier is simplified to:

$$I = WD\exp[\frac{-\phi_B}{\phi_t}] \equiv I_{\max} \qquad D = C(\exp[\frac{V_{AP}}{\phi_t}]-1) \tag{4.2}$$

This is the maximum current that can flow over the barrier and for long L, it does not change with L since the potential does not change (figure 4-1).

Taking the barrier height to be related to the potential difference between the source potential and the surface potential in the region of extra doping gives:

$$\phi_B \propto V_{SB} + V_{bi} - \phi_s - \delta \tag{4.3}$$

Since for a regular device the barrier would be $V_{SB}+V_{bi}-\phi_s$ (with a smaller ϕ_s) and this barrier does not cause any anomalous effects, ϕ_B must be less than the full potential difference from source to channel, thus the subtraction of δ. In subthreshold, when $\phi_s(x)$ is

between ϕ_f and $2\phi_f$, $\phi_s(x)$ can be approximately solved analytically: [20,p.260]

$$\phi_s = V_{GB} - V_{fb} + \frac{\gamma^2}{2}[1 - \sqrt{1 + \frac{4}{\gamma^2}(V_{GB} - V_{fb})}] \approx V_{GB} - V_{fb} \tag{4.4}$$

thus giving an equation relating ϕ_B and V_{GB} (with all of the constants lumped into θ):

$$\phi_B \propto -V_{GB} + \theta \tag{4.5}$$

The existence of a barrier at the source and drain creates a bottleneck for the number of carriers able to get from the source, over the barrier, and into the channel. For a normal device, the current is proportional to $^1/_L$ and increases as L decreases. However, the barrier will act to limit current to be less than or equal to I_{max} , which is set by the gate voltage. Thus, at some L, the current will saturate at I_{max} as L decreases further. Extracting a constant current threshold voltage (V_{GS}=V_{th}) for lengths shorter than this saturation point gives:

$$I_{CC} = 10^{-7}\frac{W}{L} = WD\exp[\frac{Vth - \theta}{\phi_t}] = I_{\max}$$
$$V_{th} \propto -\phi_t \ln(L) + \ln(\frac{10^{-7}W}{D}) - \phi_t \tag{4.6}$$

So, as L drops, the threshold voltage rises. Looking at equation 4.6, it says that V_{th} vs. log(L) should be a straight line. Looking at data from devices fabricated by Digital Semiconductor (D.S.) (figure 4-21 and 4-22) and data from the model in this thesis (figure 4-10) shows that, for a decent amount of extra doping, constant current V_{th} vs. log L is a straight line. (See section 4.4 for a description of the D.S. devices.)

For a normal MOSFET with a laterally uniform channel, the equation for diffusion current can be written as:

$$I_{diffusion} = Wq\mu_n\phi_t[\frac{n_{source} - n_{drain}}{L}] \approx Wq\mu_n\phi_t[\frac{n_{source}}{L}] \tag{4.7}$$

The source is considered to be an infinite source of electrons, so the number of electrons at the source end of the channel is set solely by the maximum number the inversion layer can support at the source end at a given gate potential. In the same vein, the drain is then considered to be an infinite sink of carriers, setting n_{drain} to zero for a finite drain voltage (equation 4.7).

Thus, for the laterally uniform device, the bottleneck that determines the current in subthreshold is the number of carriers the inversion layer can support at the source end of the channel. The current for a uniform device increases proportionately to $^1/_L$ (eqn. 4.7). This makes sense since the gradient in electron concentration between the source and drain, which drives the diffusion current, is increasing proportionately to $^1/_L$, while the number of carriers at the source and drain are staying constant. Thus, when a constant current threshold is extracted at $10^{-7}W/_L$, the $^W/_L$ factor cancels with the $^W/_L$ factor in the current equation (4.7), so V_{th} stays constant with L.

For the non-uniform channel case, even though the current is limited to I_{max}, the current still flows via diffusion. Thus, using the equation for diffusion current (4.7):

$$n_{source} = L\frac{I_{MAX}}{Wq\mu_n\phi_t} + n_{drain} \tag{4.8}$$

If I_{max} is constant, then n_{source} must be dropping as L decreases. Looking at data from a MEDICI simulation for a uniform profile and a non-uniform profile with A=2, L_d=0.1, and N_b=1x10^{17} (figure 4-2), it is evident that the n_{source} (@ x=.25μm) stays roughly constant for the laterally uniform channel, but drops (@x=0.6μm) as L decreases for the non-uniform case . So, even though the inversion layer could support more carriers, the source cannot supply them. Although the number of carriers has changed, the potential distribution, and thus the barrier, has not (figure 4-1) because in the subthreshold region the potential is determined only by the doping. The bottleneck has shifted from the number of carriers the

inversion layer can support for a uniform channel MOSFET to the number of carriers that can get into the channel over the barrier for a non-uniform channel.

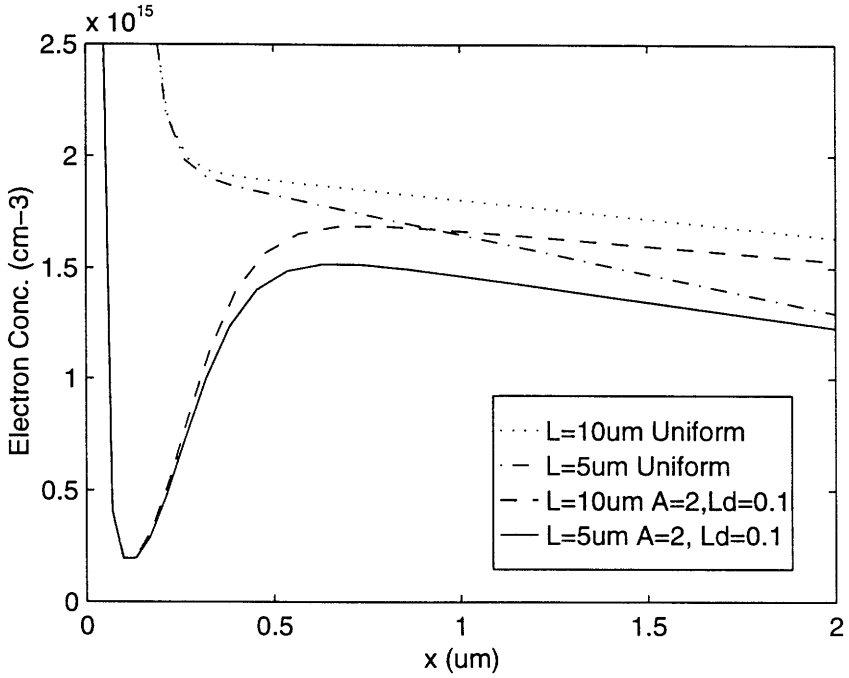

Figure 4-2: Simulated electron concentration for a 5 μm and 10 μm $1x10^{17}$ base doping device with uniform profile or a non-uniform A=2, L_d=0.1 profile at V_{GS}=0.1. (same device as figure 4-1.)

Looking at a series of I_D vs. V_{GS} plots from the devices fabricated by D.S., it is visible that the current at a given V_{GS} does not increase as $^1/_L$ (figure 4-3). Nishida et al. [6] show the same phenomenon for their devices. Another way to view this is, if the SS is reasonably constant (true for L >0.5 μm in figure 4-4 below), then the current at V_{GS} = 0 should increase as $^1/_L$ if the V_{th} is to stay constant with L. In Figure 4-4, results of the model show that for a uniform doping that is indeed true, but as the amount of extra doping increases, the I_{off} vs. L curves flatten out at I_{max}@(V_{GS}=0). As can be seen in figure 4-4, as the amount of extra doping increases (for this figure, an increase in L_d), I_{max} decreases. Thus, the current is fixed starting at longer and longer lengths, so the RSCE becomes

apparent at longer and longer channel lengths. For the L_d=0.1 curve, the I_{off} is still increasing as L decreases at 10 μm, but by 2 μm the I_{off} has become constant. For higher L_d, I_{off} has already become constant at 10 μm.

This constant current with L results in a $V_{th} \propto \log(L)$ as discussed above. Because the current becomes limited at longer lengths with an increase in the barrier height (decrease in I_{max}), then the amount of total V_{th} rise that can occur before the peak region increases.

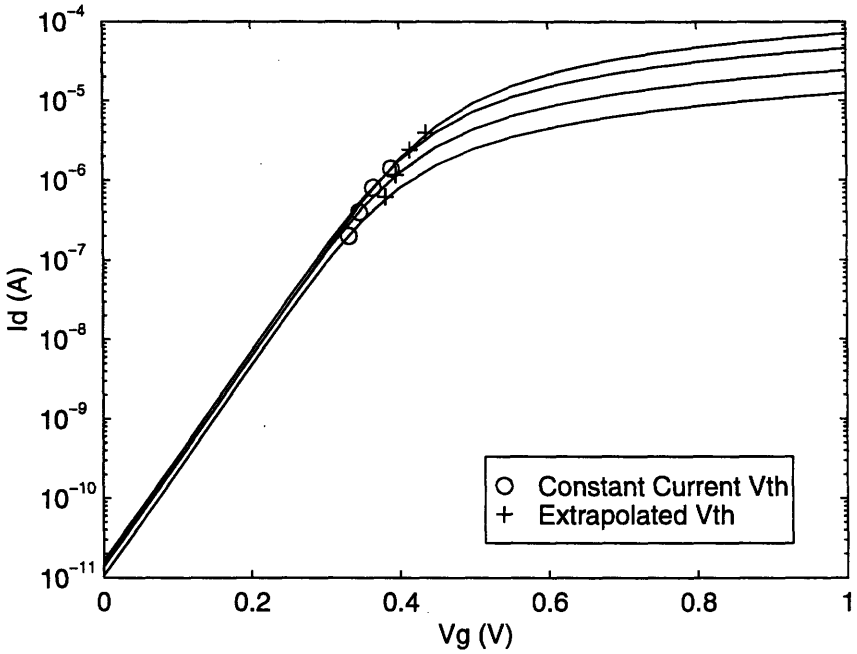

Figure 4-3 I_D-V_{GS} characteristics of devices on D.S. wafer #9 that show a Reverse Short Channel Effect and a resulting constant current in subthreshold as L changes. L=10,7,3.5,1.75 μm, W=14 μm for these devices.

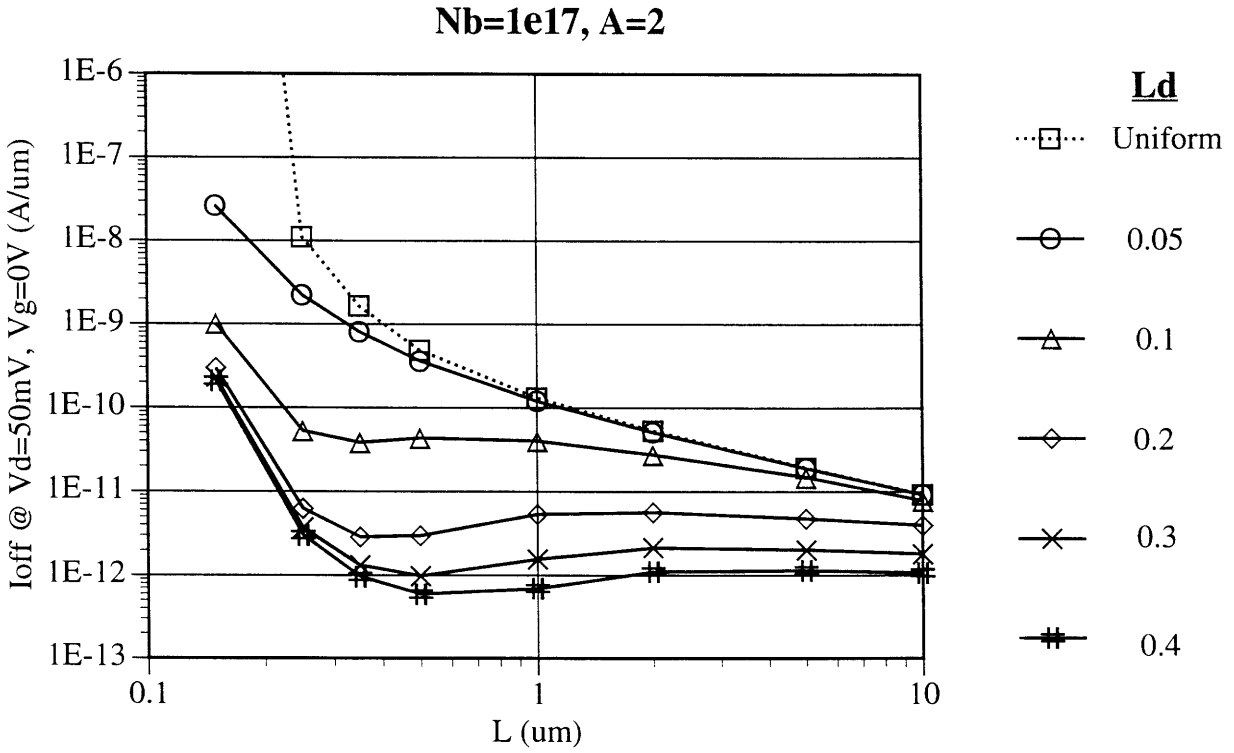

Figure 4-4: Off current vs. L for different amounts of extra doping showing the 1/L dependence of the uniform device and the more constant off current of the devices with extra doping.

The definition of V_{th} at a specific current is an artificial one, yet the extrapolated method gives a similar V_{th} vs. L distribution for real devices. The extrapolated V_{th} for the D.S. devices (+ in figure 4-3) correspond to a constant current V_{th} of $3\times10^{-7}\ W/L$ for $L\geq1\mu m$. Nishida et al. [6] also show that their extrapolated V_{th} all correspond to a single constant current definition of V_{th}, suggesting that the two definitions of V_{th} are similar.

From the definition of an extrapolated V_{th} in Chapter 3, V_{th} can be written as:

$$V_{th} = V_0 - \frac{I_0}{g_{m(V_0)}} \tag{4.9}$$

where I_0, V_0 is the point at which the tangent line of slope g_m was placed. Using Hsu's equation for I_0 (eqn 3.6) and differentiating it to get g_m (dI_D/dV_{GS}), the above equation can be written as:

$$V_{th,device} = V_0 - (V_0 - V_{thc})\left[\frac{WI_{MAX} + 2\frac{\mu C_{ox} W}{L}(V_0 - V_{thc})}{WI_{MAX} + 2\frac{\mu C_{ox} W}{L}(V_0 - V_{thc})\beta\frac{\partial V_B}{\partial V_{GS}}}\right] \qquad (4.10)$$

V_{thc} is the threshold voltage of the center of the channel (doping N_b), given by eqn. 2.1. At a given $V_0 > V_{th}$ this can be simplified to:

$$V_{th,device} \approx V_{thc}\left[\frac{L+\Delta}{L+\theta}\right] \qquad \Delta > \theta \qquad (4.11)$$

$V_{th,device}$ changes with length, for L large, as:

$$\frac{\partial V_{th,device}}{\partial L} \propto -\frac{1}{L^2} \qquad (4.12)$$

$$V_{th,device} \propto \frac{1}{L} \qquad (4.13)$$

A plot of the extrapolated V_{th} vs. $^1/_L$ for the D.S. devices, in figure 4-5, is approximately a straight line for long L (small $^1/_L$), pointing out this dependence.

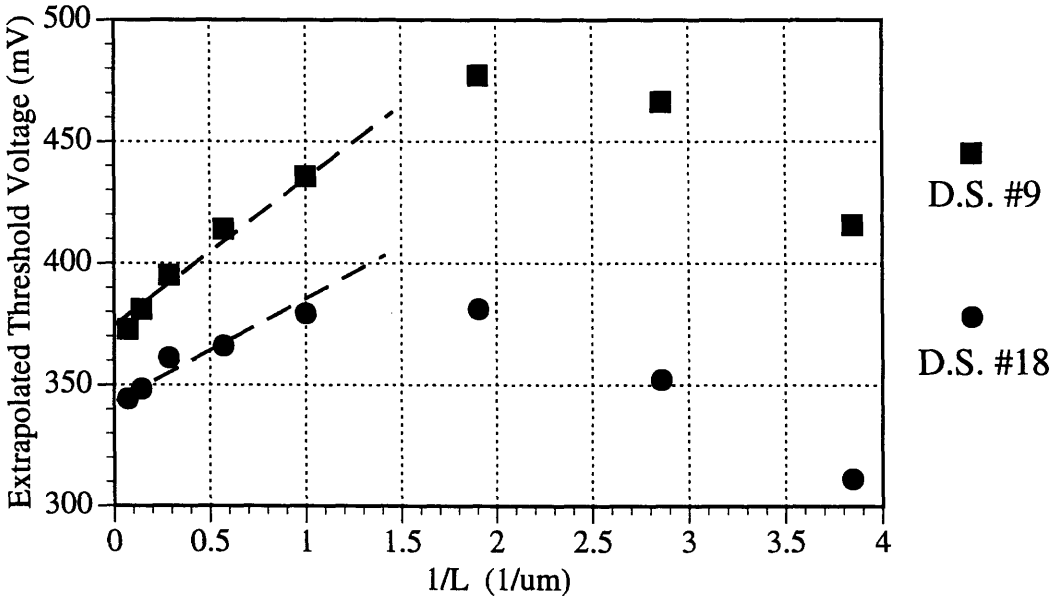

Figure 4-5 Plot of the extrapolated threshold voltage vs. L for two different devices fabricated by D.S., showing the 1/L variation of the threshold voltage at long L.

Outside of the region where the constant current $V_{th} \propto \log(L)$ and extrapolated $V_{th} \propto 1/L$, the V_{th} vs. L increases more slowly at very long channels and more quickly at short channels. As the channel length of a normal uniform channel MOSFET increases, the amount of current that flows drops for the same V_{GS} and V_{DS} bias point. Thus, at much longer channel lengths, for the non-uniform channel case, I_D predicted by the normal subthreshold current equation will be less than I_{max}. In this case, the barrier will no longer limit the current flow, and V_{th} will not depend on L.

At much shorter channel lengths, the source and drain doping distributions begin to overlap and the height of the barriers increase since more doping has been added. Thus I_{max} decreases and V_{th} rises faster than before.

The effect of I_{off} being approximately constant with L (Figure 4-4) and the potential around the extra doping not changing with L is predicted correctly by the model used in this thesis. As such, it is instructive to look at how the model predicts this. Greenfield and Dutton's [27] equation for current, used in this thesis, calculates the current from a potential distribution and the source and drain voltages:

$$I_D = \frac{-q\phi_t\mu_n Wn_i^2[\exp(\frac{-Vs}{\phi_t})][1-\exp(\frac{-V_{DS}}{\phi_t})]}{\{\int_0^L N_{(x)}\{\int_{y_{dep}}^0 \exp[\frac{\phi_{(x,y)}}{\phi_t}]dy\}^{-1}dx\}} \tag{4.14}$$

At a given V_S and V_{DS}, the numerator is a constant. Also, for L large enough that the extra doping distributions do not overlap, the integral of $\phi(x,y)$ with y is the same function of x for different L. Thus I_D can be rewritten as:

$$I_D = \frac{C}{\int_0^L \frac{N_{(x)}}{\delta_{(x)}}dx} \quad ; \qquad \delta_{(x)} = \int_{y_{dep}}^0 \exp[\frac{\phi_{(x,y)}}{\phi_t}]dy \tag{4.15}$$

Right at the source and drain edge, the depletion depth (y_{dep}) is very large and ϕ_s very large, resulting in a large $\delta(x)$ and thus very small $1/\delta(x)$. In the center, $\delta(x)$ is finite, so $1/\delta(x)$ has a finite value. At the edges where the extra doping is, $\delta(x)$ is smaller than in the center of the channel, since ϕ_s and y_{dep} are small, so $1/\delta(x)$ is larger. (Figure 4-6) The integral with L is dominated by the region of extra doping where $N_b(x)$ is large and $1/\delta(x)$ is also large. The integral over these extra doping regions has no L dependence, so I_D becomes constant with L. For a uniformly doped channel, $\delta(x)$ is constant over the channel and zero near the source and drain, so I_D is fully dependent on L in this case.

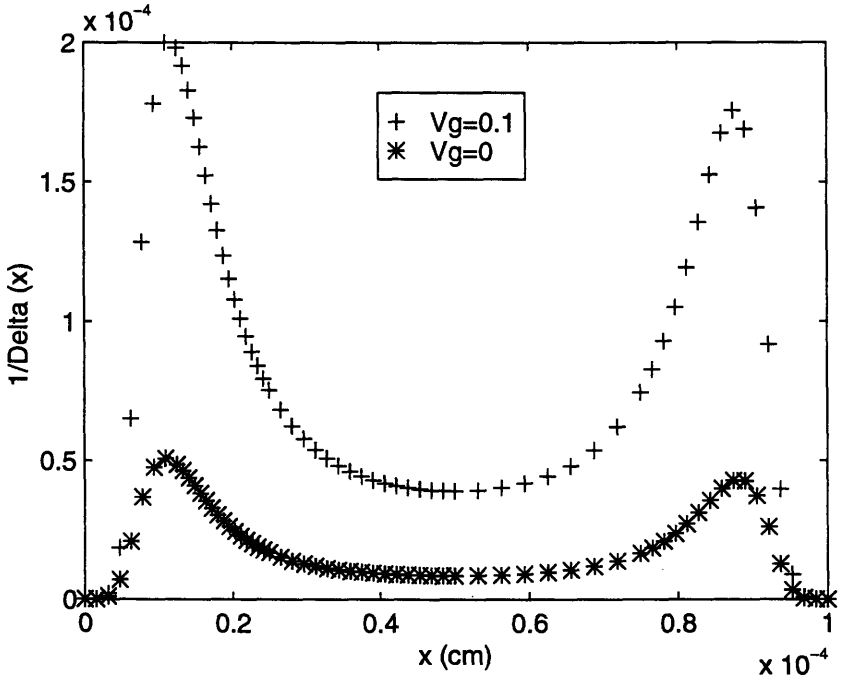

Figure 4-6 Plot of $1/\delta(x)$ (defined in equation 4.15) as calculated by the model.

Although all of this discussion has revolved around doping, fixed oxide charge has also been postulated to be a cause of RSCE. Looking back at equation 4.3 and 4.5, the V_{GB} is always constant across the length of a device, so it is the changes in the V_{fb} which primarily causes the potential barriers. Such changes in V_{fb} can happen either by doping or

charge. Since Q_{ox} is contained in V_{FB} as $-Q_{ox}/C_{ox}$, if Q_{ox} is varied with x, it too could cause a potential barrier to form and cause the same effects. However, a difference between doping and fixed oxide charge would be the effect of a back bias on the threshold voltage. Looking at the equation for $\phi_s(x)$, the change with back bias is found in the V_{GB} term.

$$\phi_s = V_{GB} - V_{fb} + \frac{\gamma^2}{2}[1 - \sqrt{1 + \frac{4}{\gamma^2}(V_{GB} - V_{fb})}] \qquad (4.16)$$

Since $\gamma \propto \sqrt{N_b}$ the magnitude of the back bias effect is greater for more highly doped regions. As a negative back bias is applied, V_{GB} increases, and due to differences in body coefficient (γ), the ϕ_s at the top of the barrier increases faster than the ϕ_s in the center of the channel. At large enough back biases, the barrier should disappear and the threshold voltage look flat with L. Figure 4-7 shows this effect for the D.S. devices. Nishida et al.[6] also show similar results for their devices.

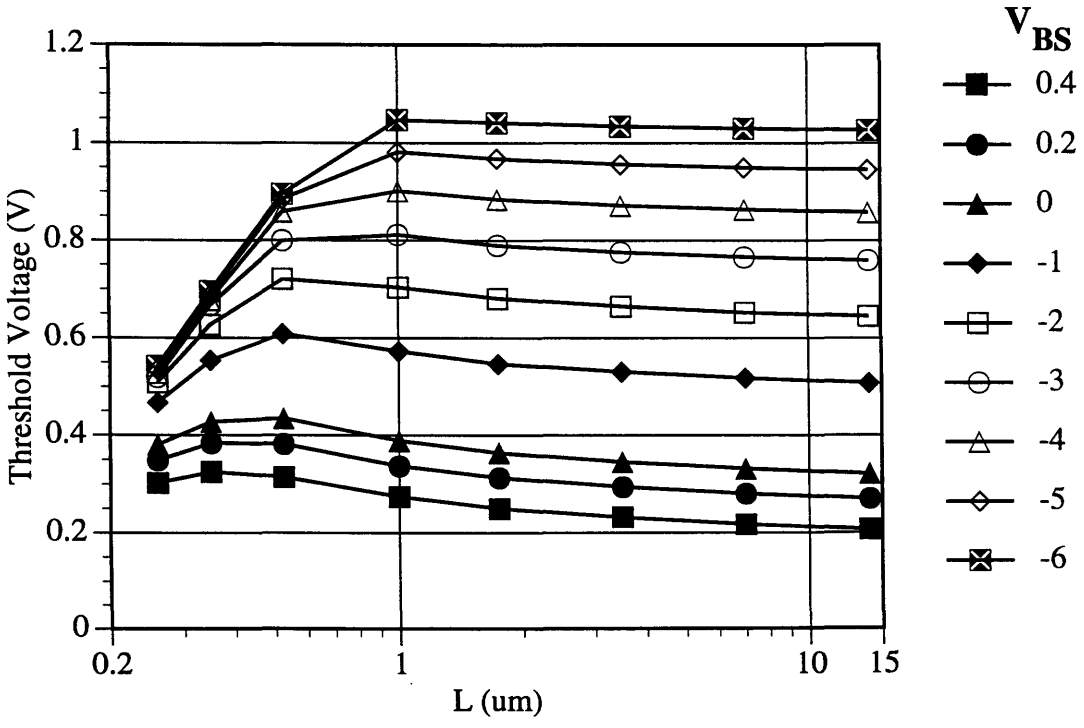

Figure 4-7 V_{th} vs. L at different back biases for D.S. wafer #9, showing the RSCE disappearance at high back biases.

For fixed oxide charge, the body coefficient is the same across the whole channel, so the whole potential would be shifted up and down at the same rate by a back bias. Thus, the shape of the V_{th} vs. L curve would be insensitive to back bias. For the D.S. devices shown in Figure 4-7, the fact that the V_{th} vs. L curve flattens out at large back biases points to extra doping as the cause of the RSCE they exhibit.

Rafferty [10] and Jacobs [14] argue their viewpoints based on this back bias effect. The devices Rafferty studies show a flattening of the V_{th} vs. L curve with back bias (figure 4-8). More short channel effects are also seen at short L since the depletion regions around the source and drain have been expanded by the back bias. Jacobs does not see any change in V_{th} vs. L shape with his devices (figure 4-9). Jacobs' data has been corrected for short channel effect changes in the body coefficient, using simulations, so even for short channels the ΔV_{th} vs. L coincide. From the above arguments, each author has analyzed their devices correctly, which points to both oxide charge and doping as causes of the RSCE.

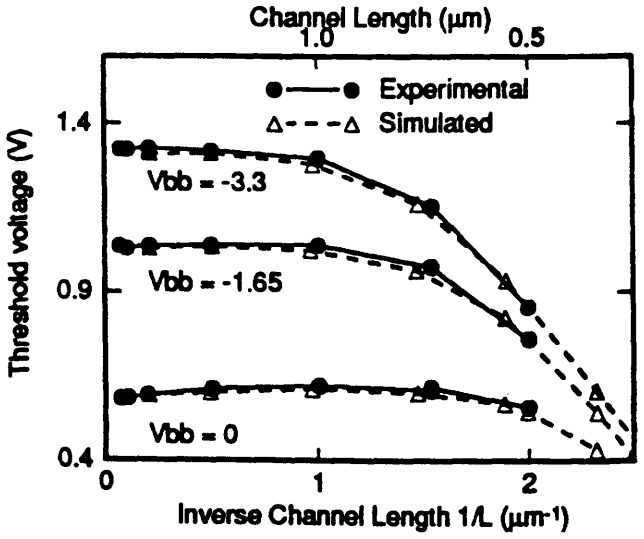

Figure 4-8 V_{th} vs. 1/L for different back biases, showing the reduced threshold voltage for higher back biases. From Rafferty et al. [10]

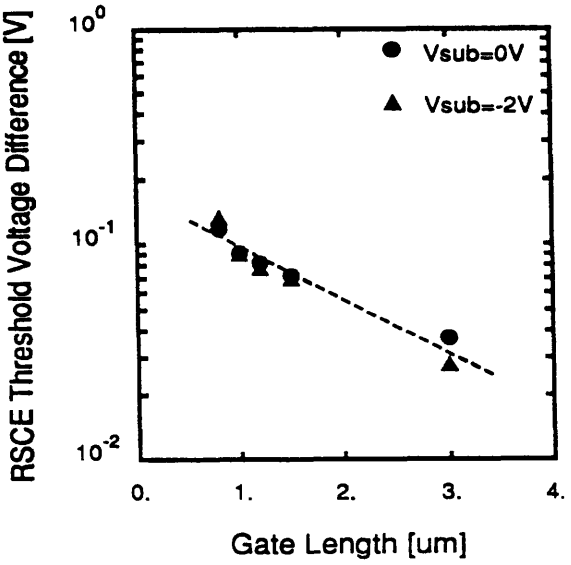

Figure 4-9 Threshold voltage increase from the predicted V_{th} vs. L (which accounted for the Short Channel Effect); ΔV_{th} is independent of back bias. From Jacobs et al. [14]

4.2 Qualitative Relationships

In the last section, the model was used to solidify a theory of the connection between a laterally non-uniform channel doping and the resulting threshold voltage roll-up versus L. This section will use the model to examine how different doping distributions affect the resulting V_{th} vs. L curve. In this section, the V_{th} will be defined as a constant current V_{th} at $10^{-7W}/_{L}$. The three parameters of the extra channel doping which will be varied are the magnitude of the extra doping (A), the characteristic length of the exponential decay (L_d), and the base uniform channel doping (N_b). (Figure 3-2)

4.2.1 Barrier Relationships

From the perspective that it is a potential barrier at the source side which is causing the Reverse Short Channel Effect, then varying the lateral doping profile should vary the size of the barrier and thus the shape of the resulting V_{th} vs. L curve. Increasing the amount of extra doping reduces ϕ_s (V_{GS} constant) and creates a higher barrier . Since the total amount of the extra doping is AL_d per side, then any increase in A or L_d should cause the threshold voltage to rise. More specifically, changing the L_d will modulate the extent of the extra doping, while changes in A will change the magnitude, and to a small extent, the extent of the doping. Looking at the series of plots in Figure 4-10, V_{th} does indeed increase with any increase in L_d and A. The rest of this section examines the relationships between A and L_d, the resulting barrier height, I_{off}, and V_{th}. Throughout, $N_b=1x10^{17}$ and $t_{ox}=75$ Å.

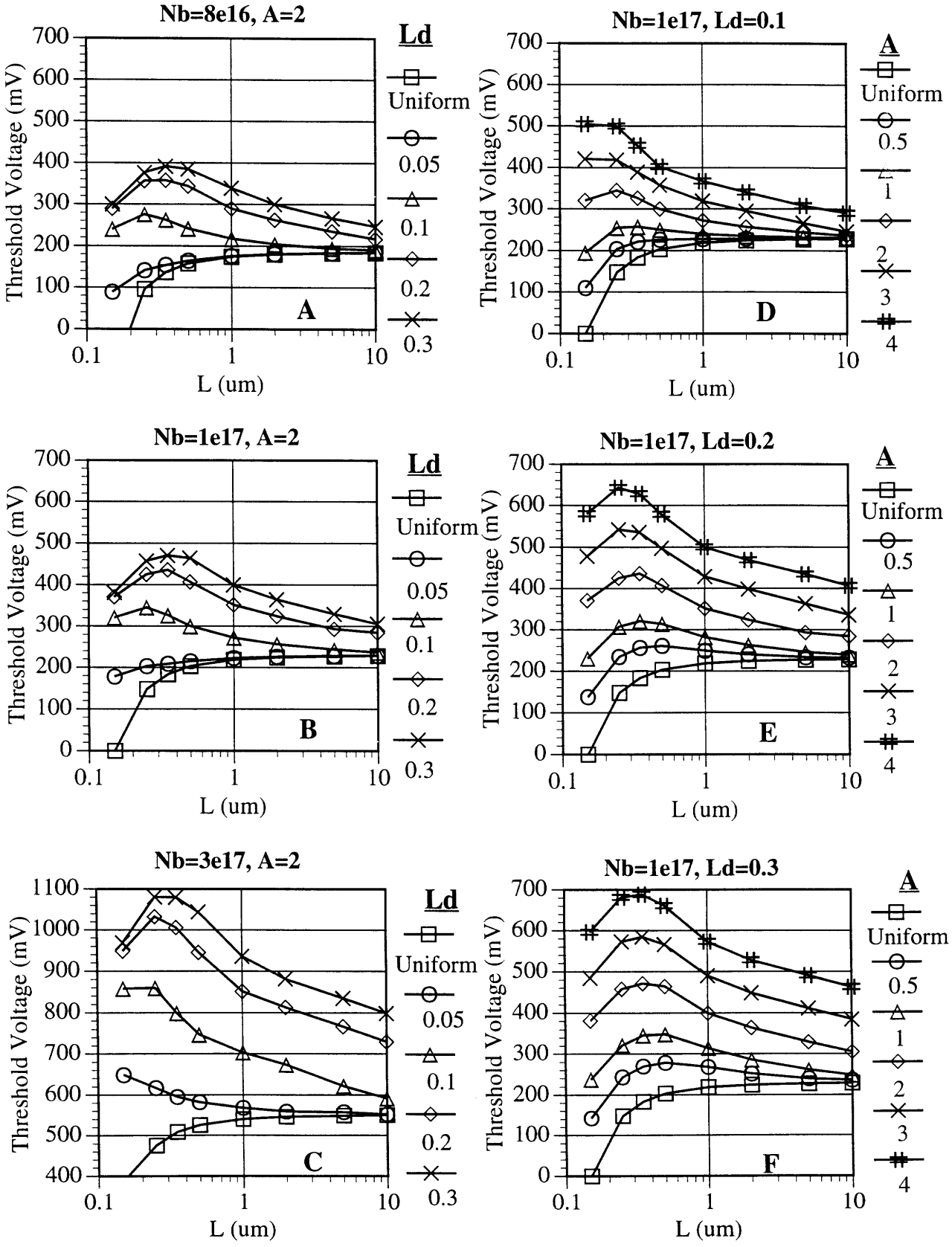

Figure 4-10 V_{th} vs. L curves for a variety of L_d, A, and N_b as predicted by the model

Changing the doping shape changes the potential barrier height, which directly affects the maximum current that can flow into the channel. Examining the current at one V_{GS} as L_d and A change should provide a sense of how L_d and A affect the barrier height. The threshold voltage can be directly related to I_{off} (I at V_g=0) as shown in chapter 3 (eqn 3.23), so I_{off} will be used in this case. Examining a long channel (L=10 μm) device, Figure 4-11A shows that for conditions that cause some threshold roll-up, I_{off} is proportional to $L_d^{-\alpha}$ for a variety of different As. Figure 4-12A shows that I_{off} is proportional to the exp(-A) for a variety of different L_ds.

Since the extra doping distribution is exponentially shaped near the source and drain, it has a slope of $-1/L_d$ for L_d>>x. As L_d gets larger, the slope gets smaller and the doping concentration near the source and drain rises, raising the barrier height. However, the rate of increase of the doping level at a given point in the channel decreases ($\propto 1/L_d^2$) as L_d gets larger, so the rate of increase of V_{th} slows, thus the $V_{th} \propto \log(L_d)$ type dependence seen in figure 4-11B.

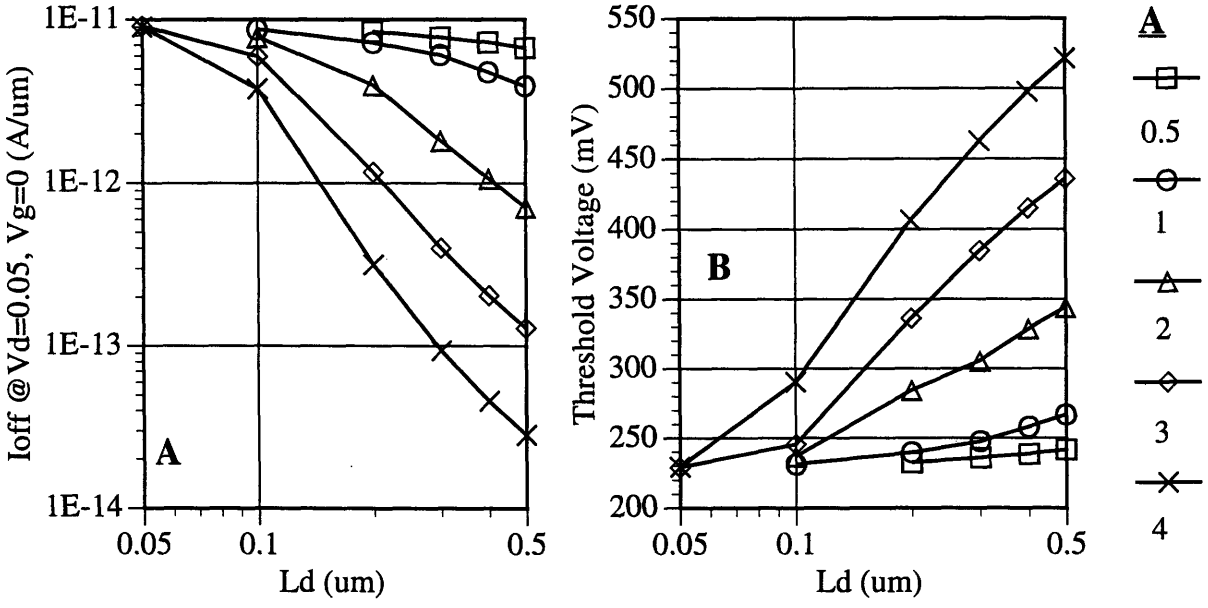

Figure 4-11 Plot showing log $I_{off} \propto L_d$ which translates into $V_{th} \propto \log(L_d)$ for a variety of values of A at N_b=1x10^{17}. Only points for curves that show a RSCE are plotted.

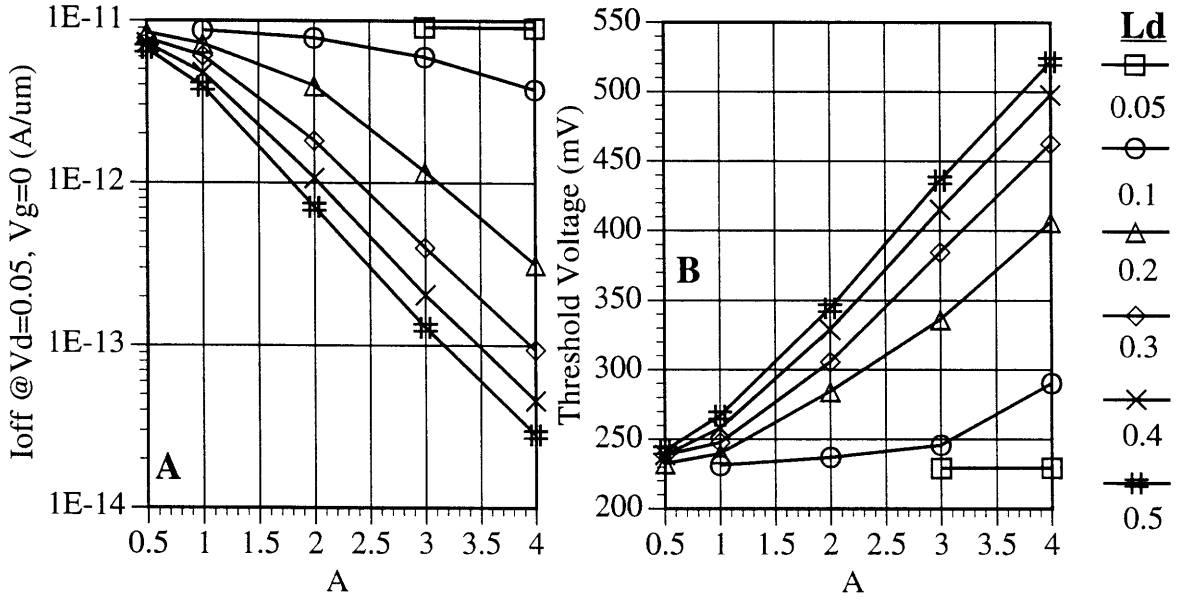

Figure 4-12 Plot showing log $I_{off} \propto exp(-A)$ which translates into $V_{th} \propto A$ for a variety of values of L_d at $N_b=1x10^{17}$. Only points for curves that show a RSCE are plotted.

From equation 3-23, the I_{off} relationships can be translated into V_{th} relationships, assuming the subthreshold slope is relatively constant with V_{GS}:

$$V_{th,long} \propto \log(L_d) \tag{4.17a}$$

$$V_{th,long} \propto A \tag{4.17b}$$

Figures 4-11b and 4-12b show that these relationships do hold for a long channel device. Since $I_{off} \propto \exp(-\phi_B/\phi_t)$ (equation 4.2), this suggests that $\phi_B \propto (A - \alpha\ln(L_d))$ at long channels.

For a 0.35 μm device, Figure 4-13a shows that for conditions that cause some V_{th} roll-up, I_{off} is proportional to $\exp(\beta\text{-}1/L_d)$, with β a constant, and Figure 4-14a shows that I_{off} is proportional to exp(-A).

Translating these dependencies into V_{th} vs. L dependencies gives:

$$V_{th,long} \propto \beta - \frac{1}{L_d} \quad (4.18a)$$

$$V_{th,long} \propto A \quad (4.18b)$$

Figures 4-13b and 4-14b show that these relationships do hold for a short channel device. For these I_{off} relationships, $\phi_B \propto (A\text{-}1/L_d)$ at short channels. The differences in the V_{th} relationships for long and short channels is probably related to the overlap of the extra doping that occurs only at shorter channels.

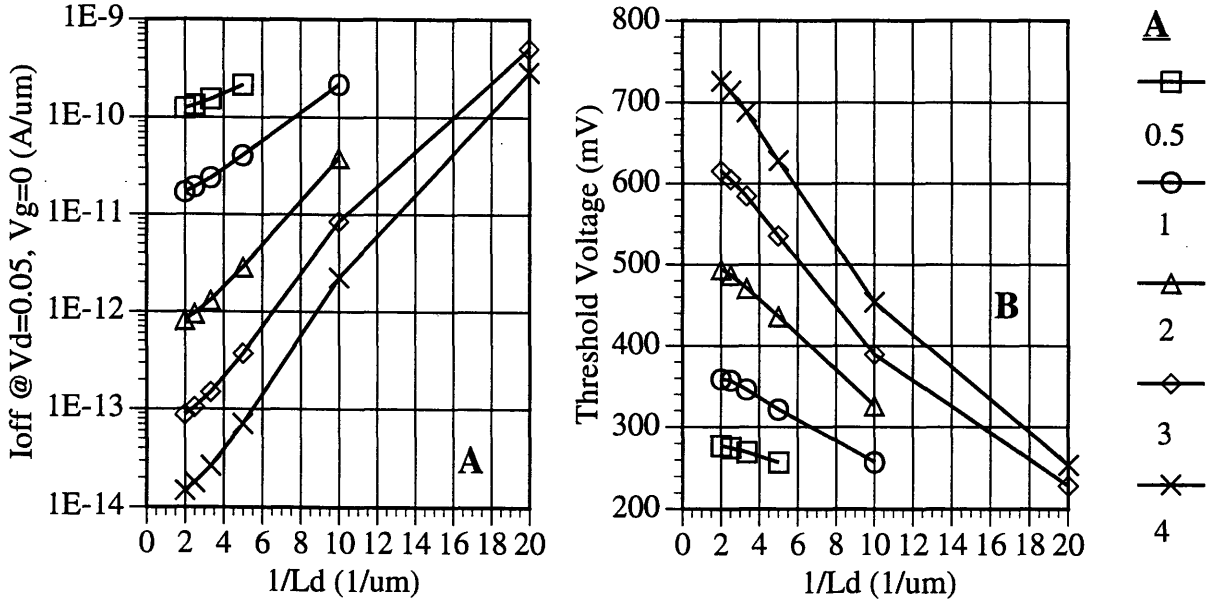

Figure 4-13 Plot showing log $I_{off} \propto \beta\text{-}1/L_d$ which translates into $V_{th} \propto \beta\text{-}1/L_d$ for a variety of values of A at $N_b=1x10^{17}$. Only points for curves that show a RSCE are plotted.

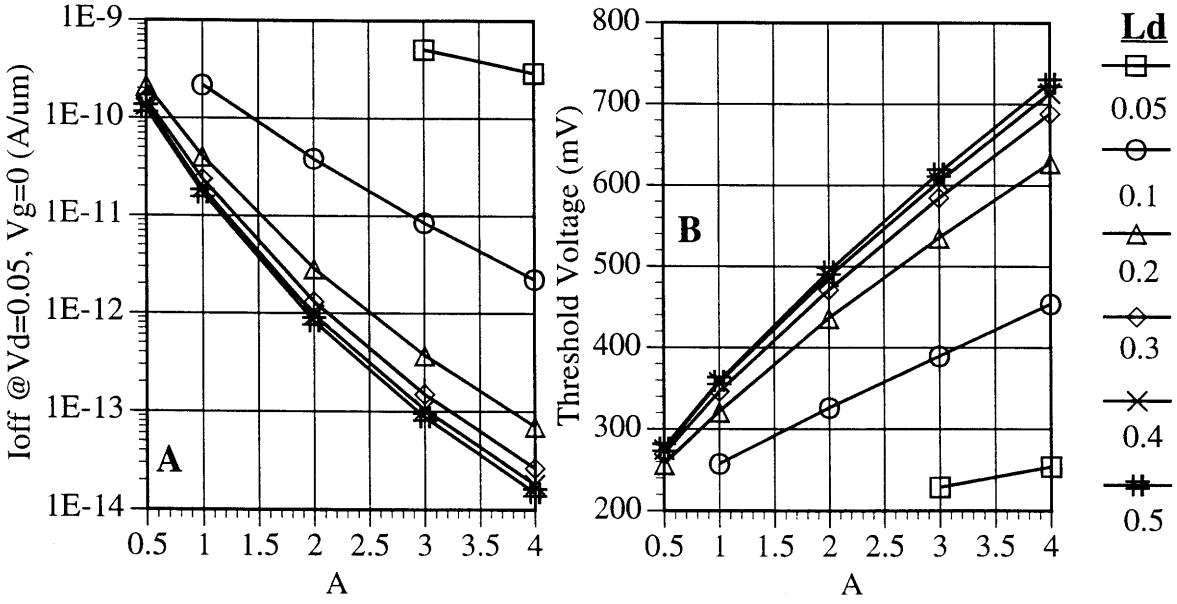

Figure 4-14 Plot showing log $I_{off} \propto A$ which translates into $V_{th} \propto A$ for a variety of values of L_d at $N_b=1x10^{17}$. Only points for curves that show a RSCE are plotted.

Although numerical dependencies could be extracted from the above figures, they would probably be dependent on the specific equation for the doping. Instead, the form of the relationships, assumed to be independent of the type of doping shape, will be used to help explain the shape of a single V_{th} vs. L curve. From an understanding of the shape, an approach to extracting the lateral doping profile that causes the Vth vs. L shape will emerge.

4.2.2 Threshold Voltage vs. L Shapes

Looking more closely at figure 4-10, four different shape descriptions suggest themselves :

1) Change in long channel (10μm in this study) V_{th}:
   $\Delta V_{th,long}(A,L_d) = V_{th,long}(A, L_d) - V_{th,long}$ (uniform)
2) Change in threshold voltage from long channel to peak V_{th}:
   $\Delta V_{PL}(A,L_d) = V_{th,peak}(A, L_d) - V_{th,long}(A,L_d)$
3) Length at which the V_{th} vs. L curve peaks (L_{peak})
4) Sharpness of the peak of the V_{th} vs. L curve

Concentrating first on the dependence of $\Delta V_{th,long}$, the relationships between V_{th} and L_d and A developed in the last section (eqn 4.17a &b) suggest a starting point. Looking at the definition for $\Delta V_{th,long}$ above, it should look like V_{th} vs. L shifted by a constant value. Given that $V_{th} \propto \log(L_d)$, $V_{th,long}$ should constantly increase as L_d increases, but the rate of increase should become less as L_d increases. Given the linear dependence of V_{th} on A, whose slope does not seem to depend on L_d (for larger L_d), the shape of the curves should be the same for any A. As A increases, the magnitude of $\Delta V_{th,long}$ will increase as well. Examining figure 4-15, $\Delta V_{th,long}$ follows these trends. This long channel threshold increase occurs because, as detailed above, the maximum current that can flow is reduced as the doping, and thus the barrier height, rises. Reducing the current at a given V_{GS} results in a higher threshold voltage.

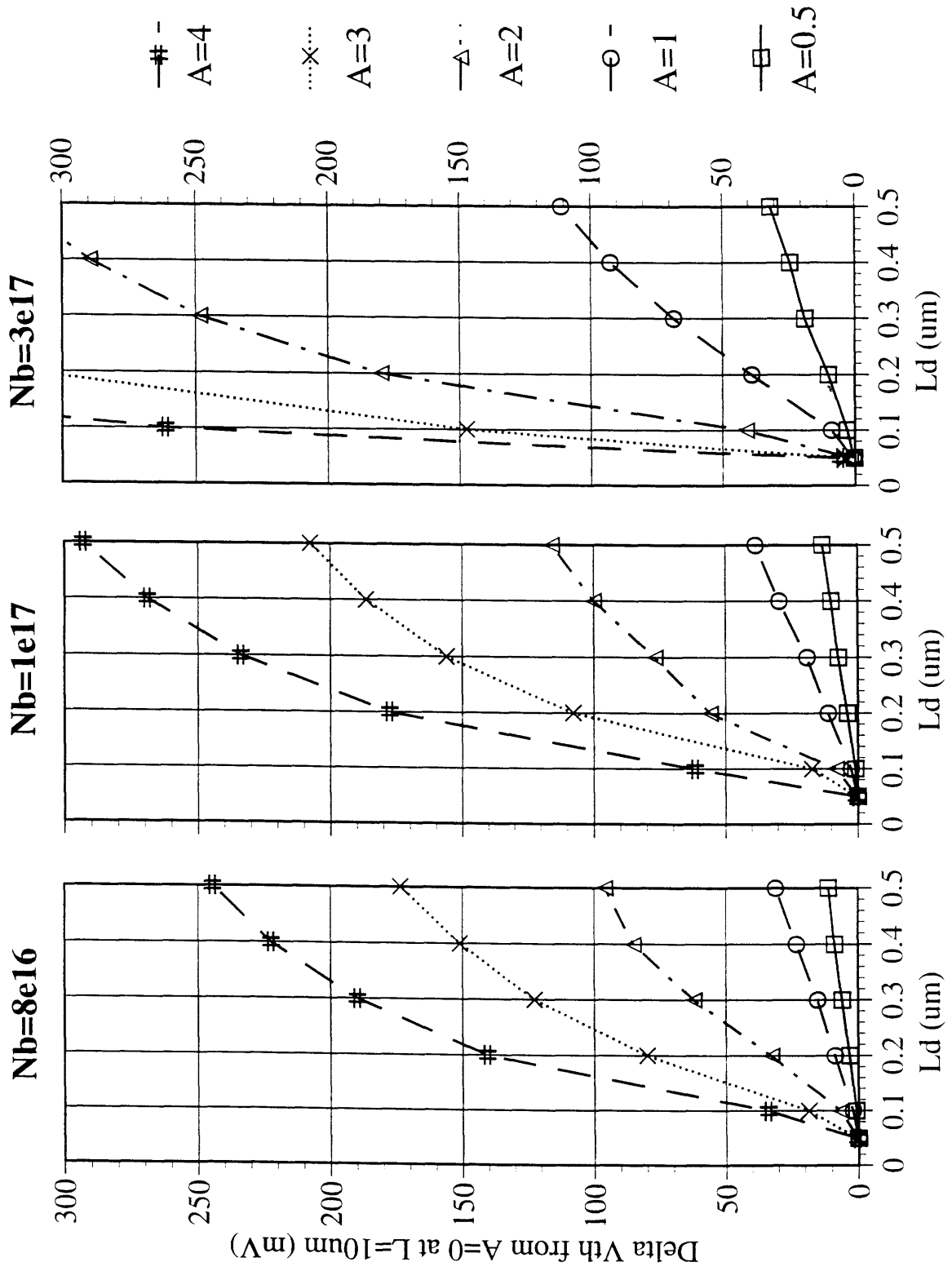

Figure 4-15 Plot of $\Delta V_{th,long} = V_{th} - V_{th,10\mu m}$ for a variety of N_b, A, and L_d as calculated by the model.

Using the relationships extracted above for the L=0.35μm device (approximately L_{peak}) in combination with those for the L=10μm device, the dependence of the difference between the peak V_{th} and long channel V_{th} (ΔV_{PL}) on L_d and A can be examined. Since $V_{th,long} \propto \log(L_d)$ and $V_{th,0.35\mu m} \propto \beta - {}^1/L_d$ then for small L_d, $V_{th,peak}$ will increase faster than $V_{th,long}$ and ΔV_{PL} will increase. For longer L_d, $V_{th,long}$ will increase more quickly than $V_{th,peak}$ and ΔV_{PL} should turn over and start decreasing. Since both $V_{th,long}$ and $V_{th,peak}$ are linearly dependent on A, increasing A always increases ΔV_{PL}. This behavior is visible in figure 4-16 for 3 different dopings. The ${}^1/L_d$ dependence of $V_{th,peak}$ reflects the fact that at some long L_d, the doping saturates at a uniform doping profile. This is visible in figure 4-17 where an A=2, $N_b=1x10^{17}$ curve saturates into a uniform $3x10^{17}$ curve as L_d increases.

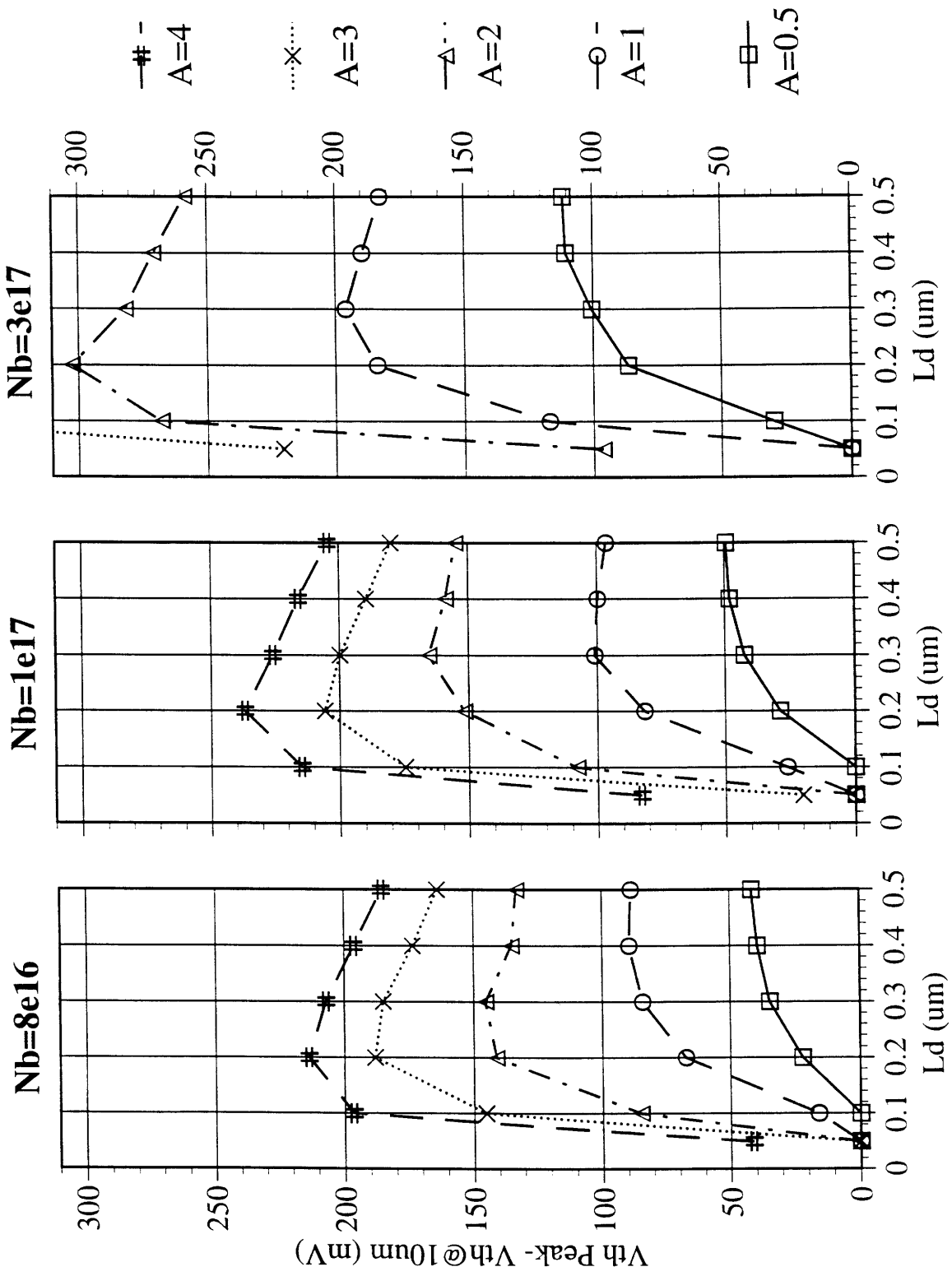

Figure 4-16 Plot of $\Delta V_{PL} = V_{th,peak} - V_{th,10\mu m}$ for a variety of N_b, A, and L_d as calculated by the model.

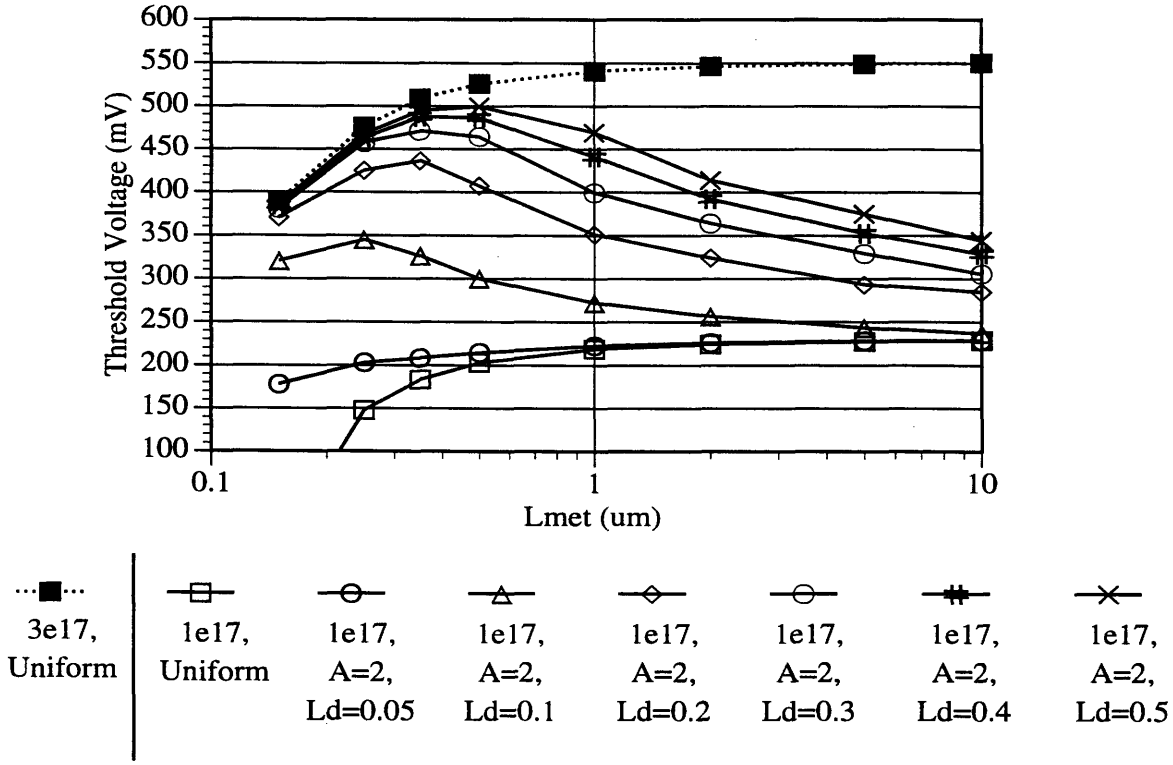

Figure 4-17 Plot of V_{th} vs. L showing that the $1x10^{17}$, A=2 curves saturate into the $3x10^{17}$ curves as L_d increases.

The V_{th} vs. L curve peaks at the point where short channel effects start to dominate. From the perspective of charge-sharing, the depletion region of the source and drain reduce the amount of doping in the channel that the gate images, reducing the Q_b in equation (2.1) and lowering the threshold voltage. At the point at which the source and drain depletion regions control a significant percentage of the charge, the threshold voltage rolls off. Increasing the doping near the source and drain shrinks the width of the depletion regions and decreases the length at which the SCE dominates. Thus, for higher A or N_b, the channel length at which the distribution peaks is shorter, which means that the V_{th} can increase further before rolling off. As L_d increases, the doping in the channel increases more slowly as the channel length decreases. Thus the threshold voltage peak saturates earlier, becoming broader and peaking at longer L with further increases in L_d. Figure 4-18 shows these relationships with L_{peak} decreasing as N_b and A increase, but increasing as L_d is increased.

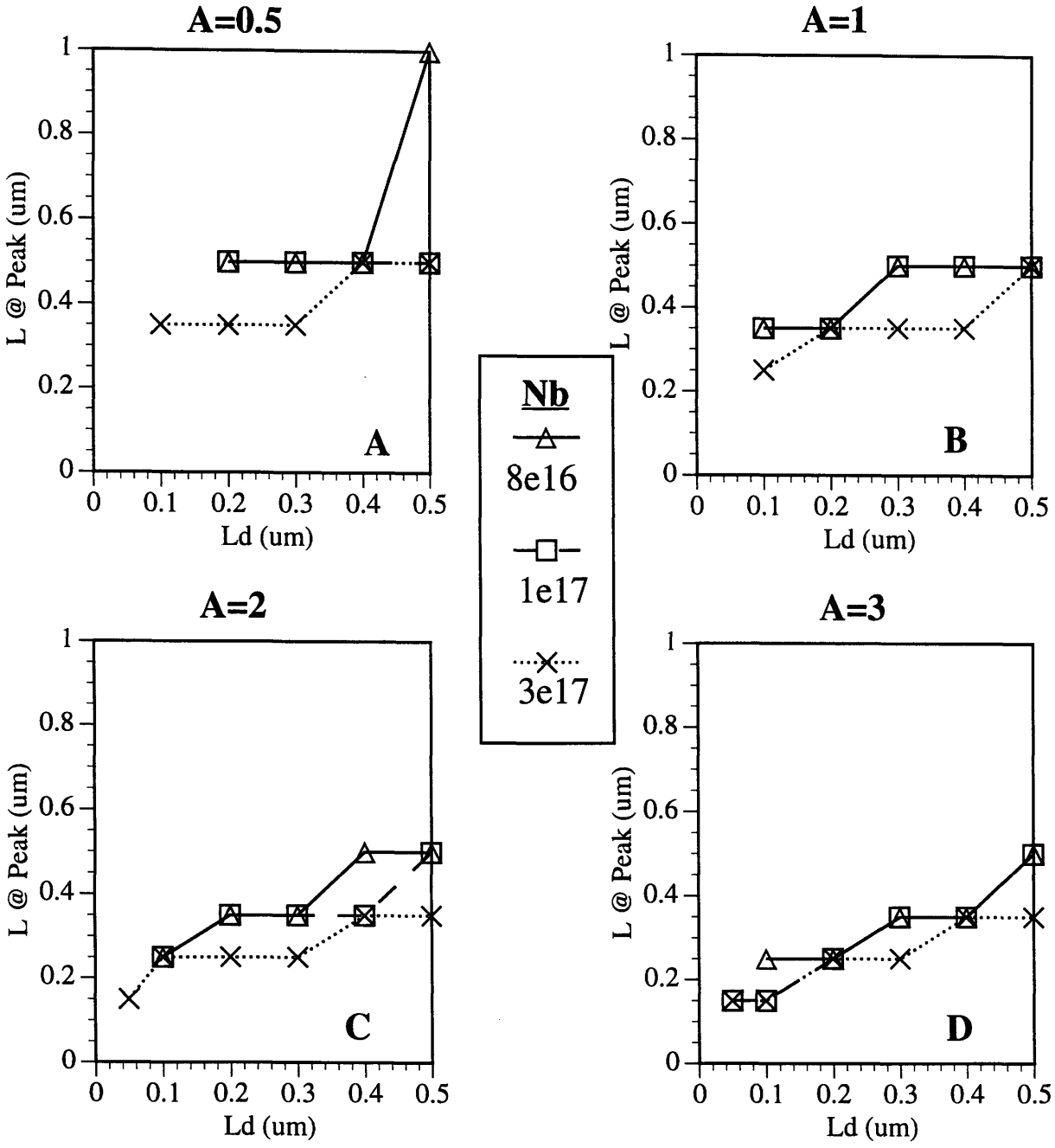

Figure 4-18 Plots of the length at which the V_{th} vs. L curves peak (L_{peak}) for a variety of N_b, A, and L_d as calculated by the model.

Finally, the peak width of the V_{th} vs. L curves broaden as L_d increases, reflecting the larger range of L over which the lateral doping profile is becoming uniform. Representing

the peak sharpness by looking at the instantaneous slope off the longer L side of the peak normalized to the average slope from peak to long L:

$$Peak \quad Sharpness = \frac{\frac{dVth}{dL}\Big|_{Peak}}{\frac{\Delta V_{PL}}{L_{peak} - L_{long}}} \tag{4.19}$$

If this ratio were 1, the curve would be a straight line from long L to the peak. A higher ratio indicates that the curve is more sharply peaked and the V_{th} increases much more quickly at short L than it does at long L. Although the discrete number of channel lengths causes some noise in the plot, figure 4-19 shows that as L_d increases, the peak width broadens, independent of the A or N_b.

These relationships of the shape descriptors to A and L_d (table 4-1) suggest an approach to extracting a profile from a set of real devices exhibiting a RSCE. An initial guess at N_b can be either the expected N_b from process simulations or one that gives a V_{th} in the range of the measured long channel V_{th} (eqn 2.1). Since the model assumes an uniform N_b with depth, the equivalent N_b for a vertical doping profile can be calculated via a doping transformation, such as the one described by Arora [33]. The sharpness of the peak (figure 4-19) suggests whether the L_d is long or short. The amount of ΔV_{PL} (figure 4-16) can suggest the appropriate A with the sense of the size of L_d gained from the peak sharpness. With this A the L_{peak} (figure 4-18) can give a smaller range for L_d. At this point ,a small enough range of parameters has been chosen that the optimization of a fit should be a relatively easy task.

	L_{peak}	ΔV_{PL} ($V_{th,peak}$ - $V_{th,long}$)	$\Delta V_{th,long}$	Peak sharpness
$N_b \uparrow$	$\downarrow$	$\uparrow$	$\uparrow$	–
A $\uparrow$	$\downarrow$	$\uparrow$	$\uparrow$	–
$L_d \uparrow$	$\uparrow$	$\uparrow$, levels off	$\uparrow$	$\downarrow$

Table 4-1 Summary of the dependencies described in section 4.2.2

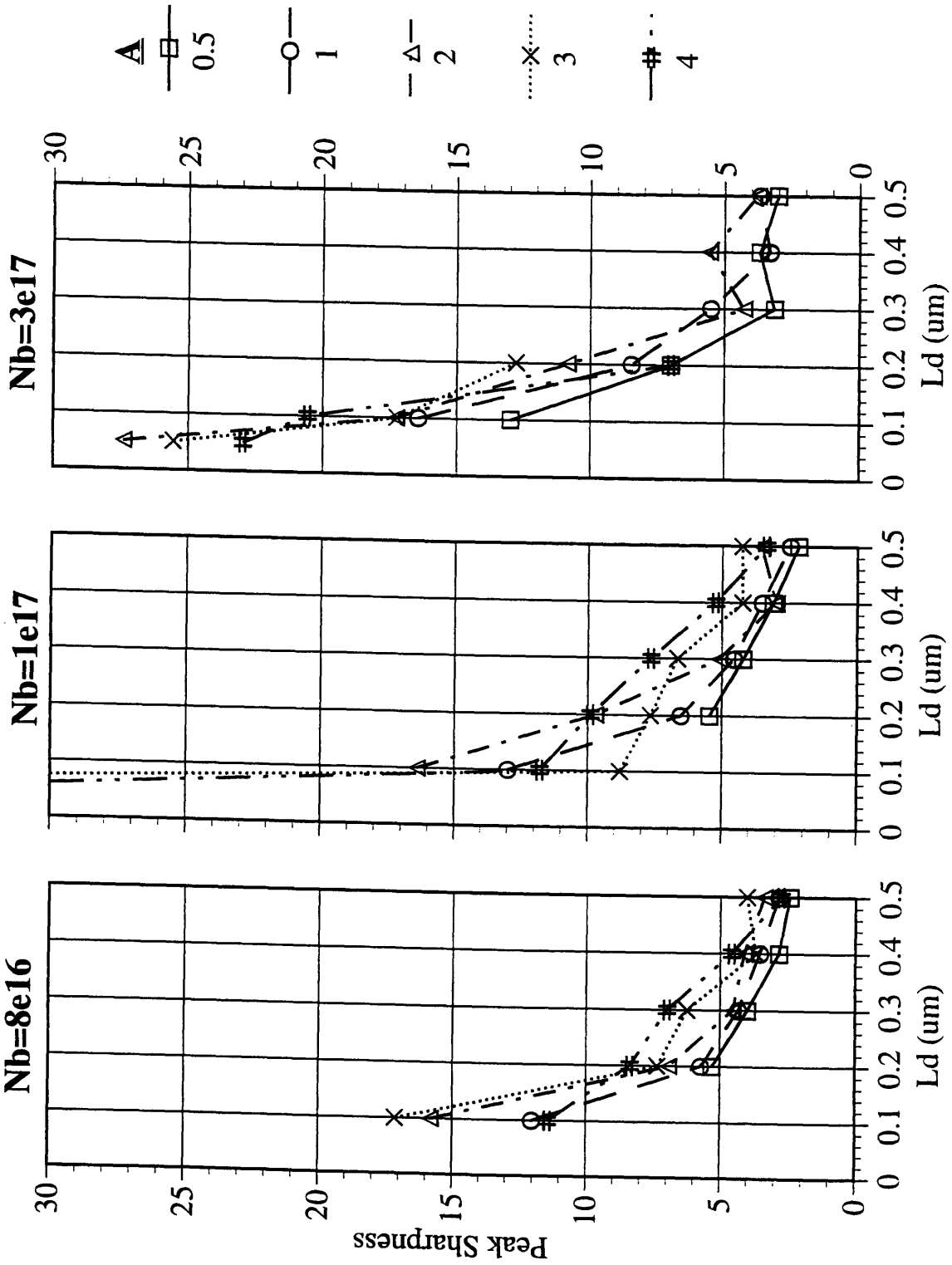

Figure 4-19 Plots of the sharpness of the V_{th} vs. L_{peak} for a variety of N_b, A, and L_d.

4.3 Matching to Actual Devices

Beyond helping to illuminate the correlation between the shape of the extra doping and the V_{th} vs. L characteristics, the model can be used in reverse to extract a lateral channel profile from a set of V_{th} vs. L data from a real device. The data in this section are from devices on two silicon wafers fabricated by Digital Semiconductor (D.S.), a division of Digital Equipment Corporation in Hudson, MA. The processing of the two wafers was the same except for the process steps and thermal cycles specific to the fabrication of the gate stack. As such, although the two wafers show very different V_{th} vs. L curves (figures 4-21 and 4-22 below) , their long channel base doping should be the same.

Motivated by Kalnitsky et al.'s work [16] on the influence of poly-depletion on the amount of RSCE, high-frequency capacitance vs. voltage measurements were taken on these devices. (Figure 4-20) Very little poly-depletion was found, thus the different amount of roll-up of the two devices has to do with interior doping. Additionally, from this plot, an electrical oxide thickness in inversion of about 75Å was extracted.

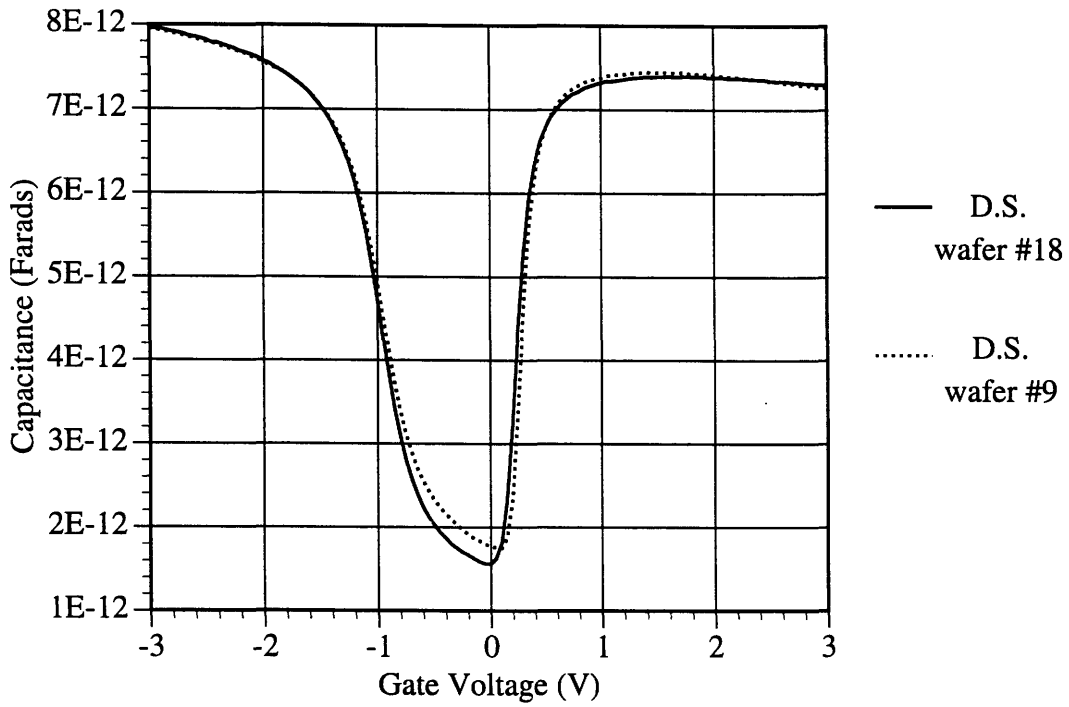

Figure 4-20 High frequency (100KHz) gate to source, drain, & back capacitance-voltage trace of a 40μmx40μm MOSFET on both D.S. wafer #9 and D.S. wafer #18.

D.S. wafer #18 has a long channel V_{th} of 280mV, corresponding to uniform doping in the range of $1.3x10^{17}$. Examining the V_{th} vs. L curve, the peak is not sharp, the ΔV_{PL} (10μm to peak) is small (~40mV), and the L_{peak} is at about 0.6μm. Looking back to figures 4-18 and 4-19 ($1x10^{17}$ chart), the flat peak and large L_{peak} suggest a high L_d. From figure 4-16, the small ΔV_{PL} with a large L_d suggests a small A. An initial guess is that L_d should be in the range of 0.3-0.5μm, A around 0.5, and $N_b=1.3x10^{17}$ for this device. Plugging these values into the model gives the results show in figure 4-21.

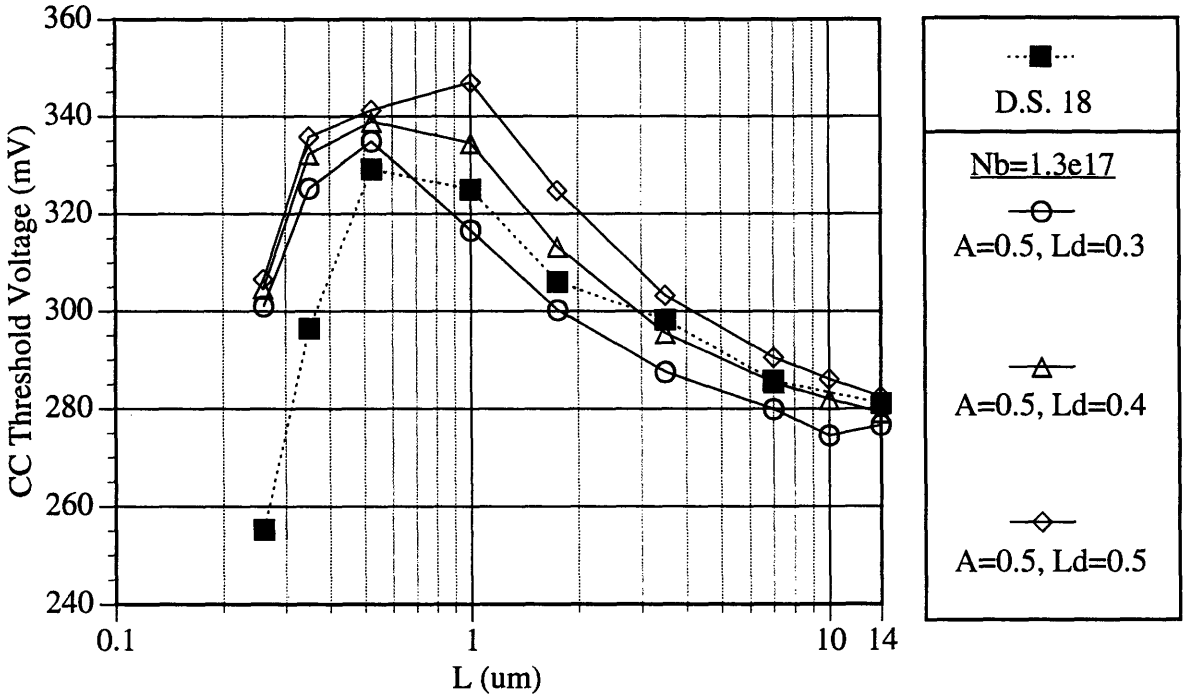

Figure 4-21 Threshold voltage vs. L for devices from D.S. wafer #18 and the approximately matching curves generated by the model.

The best match found was around A=0.5, L_d=0.4, N_b=1.3x10^{17}. These fits to the D.S. data were done by manual iteration, which was stopped when a reasonable, but perhaps not optimal, fit was found. The shape of the curves suggest that a lower L_d might give a better fit to the peak shape, which would then require a higher A to achieve the correct L_{peak}, and a lower N_b to achieve the correct V_{th} magnitude.

For D.S. wafer #9, the RSCE is much more significant. In this case, the V_{th} vs. L curve shows a much sharper peak, an L_{peak} of ~ 0.5μm, and a ΔV_{PL} of about 90mV. The long channel threshold of 320mV suggests a base doping of about 1.5x10^{17}. Although the base channel doping from the two wafers might be expected to be the same since they received the same channel implant, the best fit for wafer #9 was found at a higher N_b calculated for the long channel V_{th}. The L_{peak} again suggests an L_d of about 0.5μm. (Figure 4-18) Looking at the 1x10^{17} chart in Figure 4-15, the magnitude of ΔV_{PL} of ~90mV at an L_d

of 0.5, suggests an A of 1. An initial guess is that L_d should be in the 0.4-0.5 range, A in the 0.75-1.0 range and N_b about $1.5x10^{17}$.

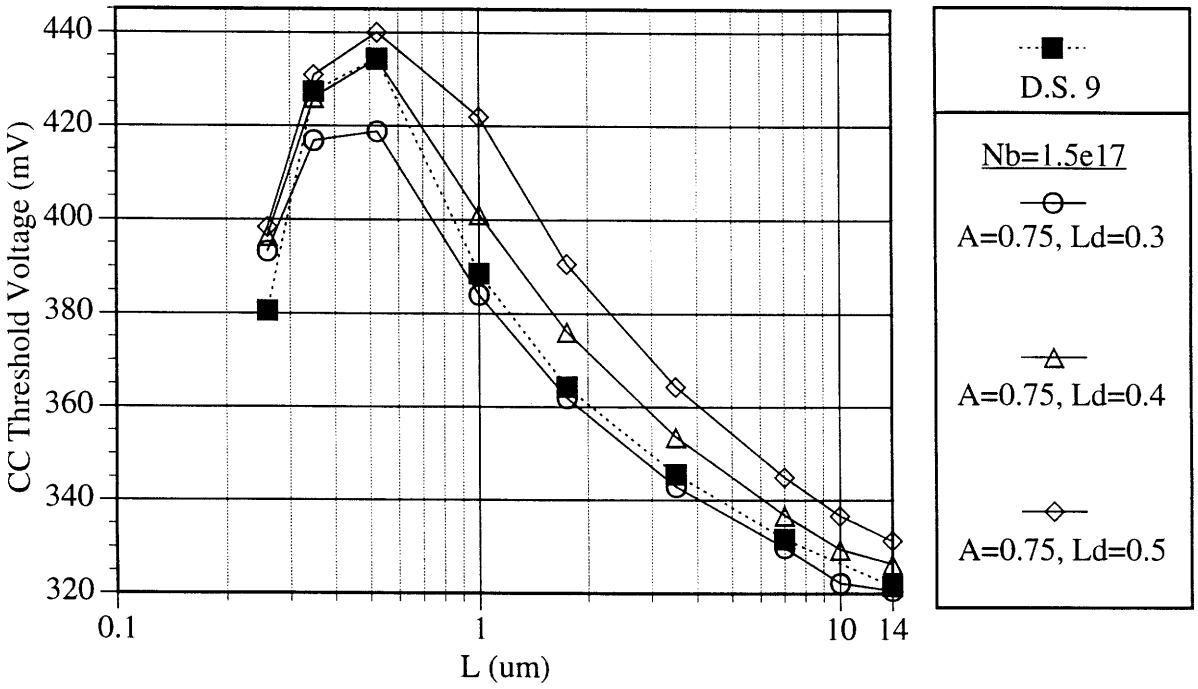

Figure 4-22 Threshold voltage vs. L for devices from D.S. wafer #9 and the approximately matching curves generated by the model.

The parameters of the doping profile that produce a good fit are $N_b=1.5x10^{17}$, A=0.75, and $L_d \approx 0.4$. These parameters and those for D.S. #18 correspond to profiles at the source and drain that are shown in figure 4-23.

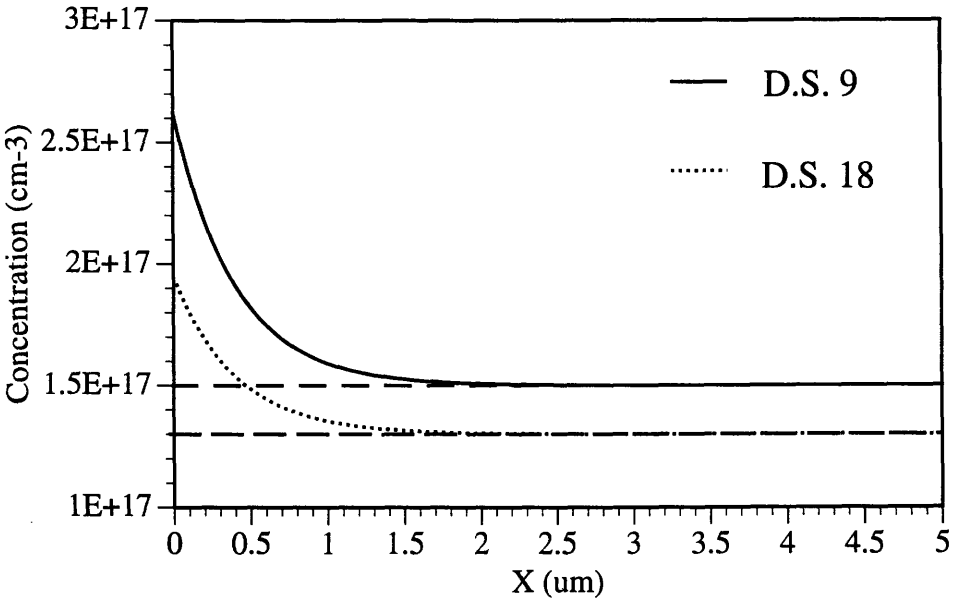

Figure 4-23 Extracted profiles from the D.S. devices corresponding to the V_{th} vs. L matches in Figures 4-21 and 4-22. D.S. #9: N_b=1.5x10^{17}, A=0.75, L_d=0.4. D.S. #18: N_b=1.3x10^{17}, A=0.5, L_d=0.4.

The differences in the base doping that gives the best fit for each wafer suggest that the doping vs. depth shapes of these profiles are different. Since the model considers them to be uniform, different shapes in y show up as a different levels of doping. This difference in doping versus depth could either be a different change in shape of the original doping versus depth profile, or a different pile-up of extra doping at the surface. The model's ability to fit these V_{th} vs. L curves demonstrates its ability to model the behavior of real devices with shallow sources and drains and non-uniform vertical profiles. This ability also shows that the model has met its goal of extracting an initial guess at the lateral profile of a device exhibiting a RSCE.

4.4 Comparison to Other RSCE Models

The approach Brut et al. [23] and Arora et al. [24] take in their models is to find an average channel doping level or oxide charge level in the channel. Since this average increases as length decreases, when it is used in standard equations for the threshold voltage a Reverse Short Channel Effect V_{th} vs. L curve is calculated. Following this example, this section compares the full model in this work with Brut's integrated average of a Gaussian doping distribution and with an integrated average of the doping distribution used in this thesis (eqn 4.20), which will be referred to as the "average model" for the rest of the thesis. Besides differing in the shape of the distribution, the average model differs from Brut's in that it keeps the maximum doping fixed at $(A+1)N_b$, rather than letting the two distributions sum up to $(2A+1)N_b=N_b+2N_0$ as Brut does.

$$N_{b,eff(L)} = N_b + AN_b \frac{2L_d}{L} \tanh(\tfrac{L}{2L_d}) \qquad (4.20)$$

Figure 4-24 compares the V_{th} vs. L curves calculated by the full model presented in this thesis and for the average model (substituted as N_{eff} into Brut's V_{th} model, eqn 3.1) at $N_b=1x10^{17}$, A=2, L_d=0.1. The curves are remarkably similar given that the only fitting parameter in the average model relates to Short Channel Effects.

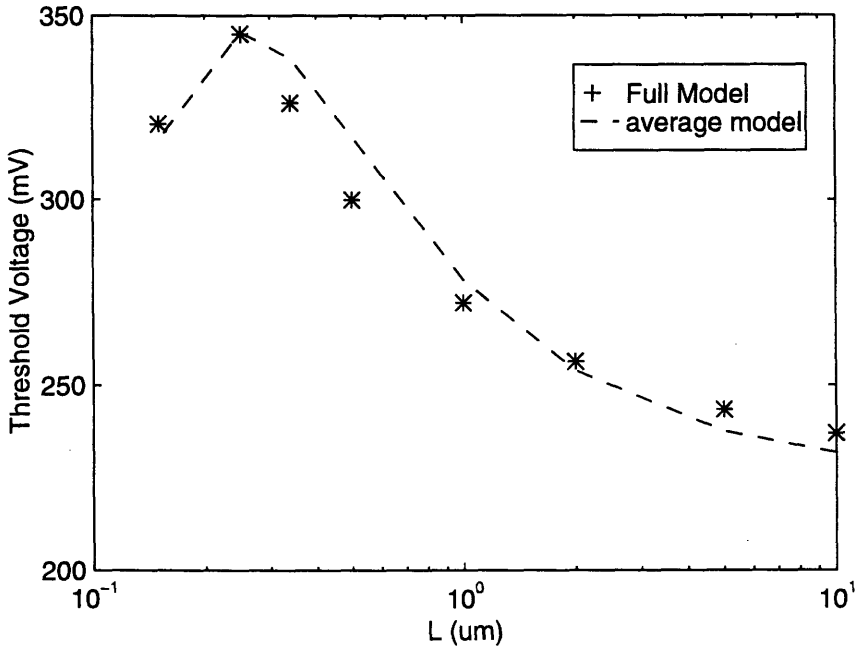

Figure 4-24 Comparison between the full model developed in this thesis and a model employing an average of the same doping profile for N_b=1x10^{17}, A=1, and L_d=0.1.

Fitting the full model, the average model, and Brut's model to the V_{th} vs. L data for D.S. #9, all three models fit the curve quite closely, as shown in figure 4-25. Table 4-2 summarizes the parameters used in each model.

	N_b	A	L_d	α	n
Full model	1.5	0.75	0.4	-	-
average model	1.45	1.4	0.2	46	-
Brut's model	1.45	0.9	0.3	45	1.2

Table 4-2 Summary of the parameters used by the 3 models to get the best fit of the V_{th} vs. L characteristics of a device from D.S. #9.

The differences in the parameters for the full model and the average model may indicate that a better fit would be found with the parameters used by the simplified model, since the fit for the full model was not completely optimized . Also, V_{th} as defined by equation (2.1) is closer to an extrapolated V_{th} [20], so a more appropriate match might be found using extrapolated V_{th} data. The difference in parameters between Brut's model and the other two models probably stems from the different doping shapes assumed by each model.

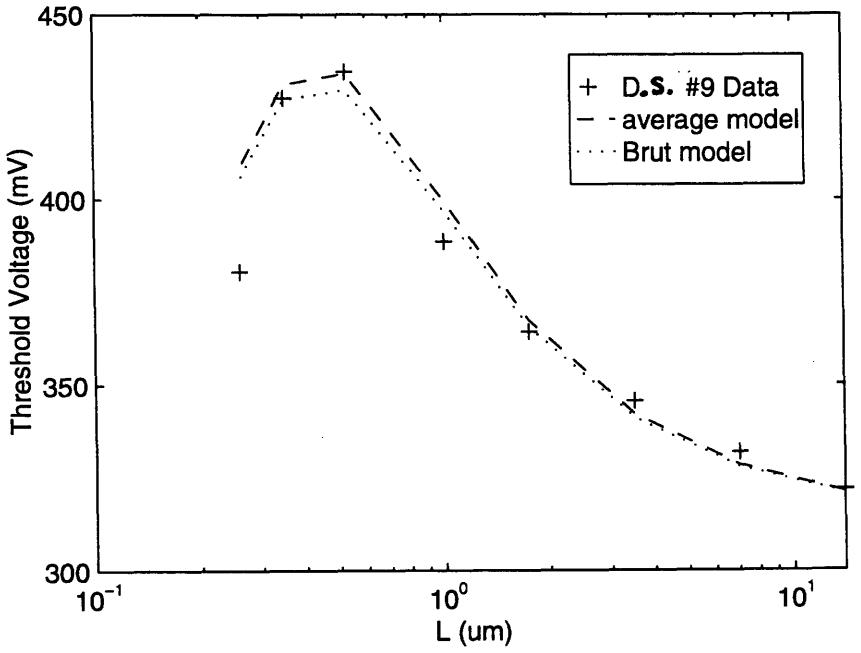

Figure 4-25 Comparison of the fit of Brut's model and the average doping model to the V_{th} vs. L of a device from D.S. wafer #9.

Examining the dependence of V_{th} on L for Brut's model (eqn 3.1) and the integrated average of this thesis' doping model (eqn 4.20):

$$N_b \propto 1+\frac{1}{L} \tag{4.21}$$

Using equation (2.1):

$$V_{th} \propto \log(1+\tfrac{1}{L})+\sqrt{\tfrac{1}{L}\log(1+\tfrac{1}{L})} \tag{4.22}$$

So, for L>1:

$$\log(1+\tfrac{1}{L}) \approx \frac{1}{L} \Rightarrow V_{th} \propto \frac{1}{L} \tag{4.23}$$

which is the dependence the extrapolated V_{th} shows versus L. Since the D.S. data is constant current V_{th} vs. L, which gives $V_{th} \propto \log(L)$, the Brut model is of a slightly different shape as visible in Figure 4-25, suggesting that a better comparison might be with extrapolated V_{th} data.

The average doping models, although empirical, give quite a good fit to actual V_{th} vs. L data. To test whether the profile that gives a good fit with the model is a good guess at the realistic profile, the profile would need to be used in a 2-D numerical simulator. If the results from this simulation gave a similar V_{th} vs. L curve, then the empirical models could provide a very fast way of extracting profiles. In contrast, the model derived in this thesis has been shown to match MEDICI simulations and to be sensitive to small changes in profile shape. Additionally, from a perspective of understanding the reason that extra doping at the edges causes the threshold voltage to rise, the model in this thesis works far better.

Chapter Five

Conclusions

5.1 Summary

In this thesis, the Reverse Short Channel Effect, as exhibited by V_{th} rising as the channel length (L) decreases, has been shown to be caused by a potential barrier at the source originating from additional doping in that area. In the subthreshold region of operation, this barrier limits the current that can flow in the channel causing the current at one gate voltage to be roughly constant as the length decreases. This constant current as L decreases results in an increasing constant current threshold voltage. This viewpoint nicely explains why the threshold voltage increases at long channel lengths and also why the V_{th} vs. L curves flatten out with a large enough back bias.

A semi-analytical model was derived for a laterally non-uniformly doped channel. Comparing the output to 2-D numerical simulations, the model was found to accurately model the simplified device structure for which it was derived. The model was also able to closely match the V_{th} vs. L curves of real devices exhibiting a RSCE. This close fit suggests that the doping distribution used, which was motivated by process simulation, was of a correct shape. In terms of speed of evaluation, the model calculated V_{th} at about 20% faster than MEDICI with no models turned on. The numerical integration routine in MATLAB

used by the model should be able to be sped up with better placement of points and a routine designed to integrate exponentials.

The model provided insight both into how the extra doping causes the V_{th} to rise and into the relationship between changes in A and L_d and the shape of the V_{th} vs. L curve. In the model, the smaller surface potential in the region of the barrier dominates the current equation, giving an approximately constant current as L changes, as seen in real devices.

Examining the V_{th} at long and short channel lengths, relationships between V_{th}, L_d, and A were expressed in equation form, which suggested how the barrier height depended on the doping profile parameters. The V_{th} vs. L curves for a range of extra doping shapes (different magnitudes and exponential decay lengths) were examined and relationships between the doping profile parameters and the amount of V_{th} roll-up, the long channel threshold voltage, the shape of V_{th} peak, and the L at which the V_{th} vs. L curve peaks were identified. Beyond the specific doping distribution used, these relationships give a sense of the effect on V_{th} vs. L curves that changing the lateral extent or magnitude of extra doping at the channel edges would have. This set of relationships was used to help fit V_{th} vs. L curves for real devices that exhibited RSCE. The initial guess made from these relationships gave results very close to the actual data.

5.2 Future Work

Both process simulation and the fits to the D.S. devices suggest that the change of doping with depth is an important parameter ignored by this model. By including depth, the subthreshold slope versus L could be modeled accurately, providing another piece of data to help extract the channel profile.

In its current formulation, the model could be used inside of an optimization loop to gain a sense of the lateral channel profile by fitting V_{th} vs. L curves. These fits could be used as an initial guess to speed optimization by a full 2-D numerical simulator to extract more

exact profiles. In this case, the numeric integration routine would need to be sped up to make this optimization occur more quickly. The average doping models in the literature should also be carefully evaluated, given their comparable fits to actual data and their faster evaluation time.

Besides inverse modeling devices in order to debug a process, the model could be used to understand the design space for counter-doping, where doping is intentionally added at the source and drain to reduce the short channel effect. Figure 4-10 points to the fact that a delicate balance exists for such extra doping between reducing the SCE and causing a RSCE, which is just as detrimental to device operation. In this case, since the extra doping would be implanted into the area, rather than diffused, a differently shaped doping distribution might need to be used in the model.

Whether in detrimental effects like the RSCE or in the use of counter-doping to reduce the SCE, laterally non-uniform doping profiles will play a key role as device dimensions are reduced beyond the 0.1 μm range and as low-power considerations push the threshold voltage lower.

Bibliography

[1] Y. Tsividis, *Operation and Modeling of the MOS Transistor.* New York: McGraw-Hill, 1987.

[2] S. Sze, ed. *VLSI Technology.* New York: McGraw-Hill, 1988.

[3] C. Tsamis, D. Tsoukalas, and P. Normand. "Decrease of the Lateral Distribution of Interstitials in Silicon-on-Insulator Structures." *Microelectronic Engineering*, 28, 1995, pp. 463-66.

[4] S. Crowder, P. Rousseau, J. Snyder, J. Scott, P. Griffin, and J. Plummer. "The Effect of Source/Drain Processing on the Reverse Short Channel Effect of Deep Sub-Micron Bulk and SOI nMOSFETs." *IEDM95*, pp. 427-30.

[5] S. Crowder. *Processing Physics Silicon-on-Insulator Material.* Ph.D. Thesis, Stanford University: Technical Report ICL95-047.

[6] M. Nishida and H. Onodera. "An Anomalous Increase of Threshold Voltages with Shortening the Channel Lengths for Deeply Boron-Implanted N-Channel MOSFET's." *IEEE Transactions on Electron Devices*, ED-28(9), September 1981), pp. 1101-03.

[7] C. Mazuré and M. Orlowski. "Oxidation Induced Local Channel Dopant Accumulation." *Solid State Devices*, 1988, pp. 439-42.

[8] C. Mazuré and M. Orlowski. "Guidelines for Reverse Short-Channel Behavior." *IEEE Electron Device Letters*, 10(12), December 1989, pp. 556-58.

[9] M. Orlowski, C. Mazuré, and F. Lau. "Submicron Short Channel Effects Due to Gate Reoxidation Induced Lateral Interstitial Diffusion." *IEDM 1987*, pp. 632-35.

[10] C. Rafferty, H.-H. Vuong, S. Eshraghi, M. Giles, M. Pinto, and S. Hillenius. "Explanation of Reverse Short Channel Effect by Defect Gradients." *IEDM 93*, pp. 311-14.

[11] K. Nishi, H. Matsuhashi, T. Ochiai, M. Kasai, and T. Nishikawa. "Evidence of Channel Profile Modification Due to Implantation Damage Studied by a New Method, and Its Implication to Reverse Short Channel Effects of NMOSFETs." *IEDM 95*, pp. 993-95.

[12] C.-Y. Lu and J. Sung. "Reverse Short-Channel Effects on Threshold Voltage in Submicrometer Salicide Devices." *IEEE Electron Device Letters*, 10(10), October 1989, pp. 446-48.

[13] J. Lutze and S. Venkatesan. "Techniques for Reducing the Reverse Short Channel Effect in Sub-0.5 μm CMOS." *IEEE Electron Device Letters*, 16(9), September 1995, pp. 373-75.

[14] H. Jacobs, A. v. Schwerin, D. Scharfetter, and F. Lau. "MOSFET Reverse Short Channel Effect Due to Silicon Interstitial Capture in Gate Oxide." *IEDM 93*, pp. 307-10.

[15] C.-P. Chang, K. Ng, W. Lindenberger, T. Kook, F. Preuninger, and A. Kornblit. "Etching-related Reverse Short Channel Effect in Buried Channel P-MOSFET." *Microelectronics Manufacturing and Reliability*, 1802, 1992, pp. 46-57.

[16] A. Kalnitsky, R. Frjns, C. Mallardeau, E. Daemen, M. Bonis, M. Varrot, M.-T. Basso, R. Penning de Vries, and M. Brillouet. "Suppression of the V_t Roll-up Effect in Sub-Micron NMOST." *ESSDERC 94*, pp. 377-80.

[17] H. Hanafi, W. Noble, R. Bass, K. Varahramyan, Y. Lii, and A. Dally. "A Model for Anomalous Short-Channel Behavior in Submicron MOSFET's." *IEEE Electron Device Letters*, 14(12), December 1993, pp. 575-77.

[18] S. Chung, S. Cheng, G. Lee, and J. Guo. "Direct Observation of the Lateral Nonuniform Channel Doping Profile in Submicron MOSFET's from an Anomalous Charge Pumping Measurement Results." *1995 Symposium on VLSI Technology Digest of Technical Papers*, pp. 103-4.

[19] J-C Guo, C. Hsu, and S. Chung. "Direct Observation of Channel-Doping-Dependent Reverse Short Channel Effect Using Decoupled C-V Technique." *Japanese Journal of Applied Physics*, 33(1,1B), January 1994, pp. 630-34.

[20] N. Arora, *MOSFET Models for VLSI Circuit Simulation: Theory and Practice*. New York: Springen-Verlag, 1993.

[21] P. Ko. *Hot-Electron Effects in MOSFET's*. Ph.D. Dissertation, Dept. of Electrical Engineering and Computer Science, Univ. of CA, Berkeley, June 1982.

[22] T. Toyabe and S. Asai, "Analytical Models of Threshold Voltage and Breakdown Voltage of Short-Channel MOSFET's Derived from Two-Dimensional Analysis." *IEEE Journal of Solid-State Circuits*, SC-14(2), April 1979, pp. 375-82, .

[23] H. Brut, A. Juge, and G. Ghibaudo. "Physical Model of Threshold Voltage in Silicon MOS Transistors Including Reverse Short Channel Effect." *Electronics Letters*, 31(5), 2 March 1995, pp. 410-12.

[24] N. Arora and M. Sharma. "Modeling the Anomalous Threshold Voltage Behavior of Submicrometer MOSFET's." *IEEE Electron Device Letters,* 13(2), February 1992, pp. 92-94.

[25] S. Hsu, I. Kalish, K. Suzuki, R. Kawabata, and H. Shibayama. "Physical Mechanism of the 'Reverse Short-Channel Effect' in MOS Transistors." *Solid-State Electonics*, 34(6), 1991, pp. 605-8.

[26] A. Ahmed, R. Carter, and S. Mohammad. "The Effect of Focused Ion-Beam Implantation on the Threshold Voltage of Short-Channel Silicon Metal-Oxide-Semiconductor Field-Effect Transistors." *Journal of Applied Physics*, 78(12), 15 December 1995, pp. 7007-17.

[27] J. Greenfield and R. Dutton, "Nonplanar VLSI Device Analysis Using the Solution of Poisson's Equation." *IEEE Transactions on Electron Devices*, ED-27(8), August 1980, pp. 1520-32.

[28] MATLAB™, The MathWorks, Inc Program.

[29] S. Biesemans and K. DeMeyer. "Analytical Calculation of Subthreshold Slope Increase in Short-Channel MOSFET's by Taking Drift Component into Account." *Japanese Journal of Applied Physics*, 34(1,2B), February 1995, pp. 917-20.

[30] S. Biesemans, S. Kubicek, and K. DeMeyer. "New Current-Defined Threshold Voltage Model from 2D Potential distribution Calculations in MOSFETs." *Solid-State Electronics*, 39(1), 1996, pp. 43-48.

[31] MEDICI™, Technology Modeling Associates Program.

[32] S. Sze, *Physics of Semiconductor Devices*. New York: Wiley-Interscience, 1981.

[33] N. Arora. "Semi-Empirical Model for the Threshold Voltage of a Double Implanted MOSFET and its Temperature Dependence." *Solid-State Electronics*, 30(5), 1987, pp. 559-69.