

MODELS FOR INTERPRETING SCANNING CAPACITANCE MICROSCOPE MEASUREMENTS

J.F. Marchiando*, J.R. Lowney and J.J. Kopanski

Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899

(Received for publication May 12, 1996 and in revised form January 21, 1997)

Abstract

Theoretical high-frequency capacitance-versus-voltage curves have been calculated for silicon in order to correlate scanning capacitance microscope (SCM) measurements with semiconductor dopant profiles. For two-dimensional cases, the linear finite-element method is used to solve Poisson's equation in the semiconductor region and Laplace's equation in the oxide and the ambient regions. For three-dimensional cases, the collocation method is used in the semiconductor region, and the linear finite-element method is used outside this region. For a given oxide thickness, probe shape, and probe-tip size, the capacitance is calculated for a series of cases of uniform doping, and a few example solutions are found for a model graded doping profile. For the case of uniform doping, the theory can be used to form a database for rapid interpretation of SCM measurement data.

Key Words: Scanning capacitance microscopy, differential capacitance, dopant profiling, semiconductor.

Introduction

Profiling the dopant concentration along the surface of a processed semiconductor wafer with 20 nm spatial resolution and 10% accuracy is identified in the 1994 National Technology Roadmap for Semiconductors as a critical measurement need for the development of next generation integrated circuits [38]. This need is well documented and is the subject of a recent review [11]. One scanning probe method that holds great promise for two-dimensional (2D) and three-dimensional (3D) dopant profiling is scanning capacitance microscopy [7, 11, 14, 18, 19, 20, 21, 23, 24, 25, 29, 30, 40, 41, 42]. A scanning capacitance microscope (SCM) is based on an atomic force microscope (AFM) with a modified conducting tip and appropriate circuitry to measure the probe-to-sample capacitance variation as a function of both bias and probe position. A thin insulating oxide layer atop the sample separates the semiconductor from the conducting probe-tip, thus forming a metal-oxide-semiconductor (MOS) capacitor. The measured data are proportional to the change in the high-frequency capacitance caused by a modulation voltage and provide a measure of the field-induced changes in the semiconductor volume depleted of majority carriers. From these measurements, the dopant concentration is determined. Extracting dopant profiles from SCM data requires a model, and this ultimately establishes the accuracy of the method. While accuracy is important, there is also a need for a quick interpretation of the data. The first models [18, 21, 23] to quickly interpret SCM data used a number of simplifying assumptions, such as using the one-dimensional (1D) MOS capacitor model [16], but this tends to compromise the accuracy. These models are discussed further in the Appendix. To help correlate SCM data with dopant concentration, we have calculated theoretical high-frequency capacitance curves as a function of applied bias for a range of dopant densities in silicon for a given oxide thickness and probe-tip size. A set of calibration or conversion curves relating dopant density and derivative of the high-frequency capacitance is presented here that will provide the basis for a quick and accurate means to extract dopant densities from SCM data.

For 2D cases, the model samples are uniformly doped, and the probe is conically shaped and oriented normal to the surface of the sample, so that the system exhibits cylindrical

*Address for correspondence:

J.F. Marchiando
Semiconductor Electronics Division
Electronics and Electrical Engineering Laboratory
National Institute of Standards and Technology
Building 225, Room A305
Route 270 and Quince Orchard Road
Gaithersburg, MD 20899

Telephone number: (301) 975-2088
FAX number: (301) 948-4081
E.mail: jay.marchiando@nist.gov

symmetry. The linear finite-element method is used to solve Poisson's equation in the semi-conductor region and Laplace's equation in the oxide and the ambient regions.

Because there is need to understand SCM data near junctions, example solutions are also found for a model high/low like-dopant (p^+/p) graded profile junction. The net charge density distribution is found near the model junction for cases when a probe is absent, a V shaped probe is centered above the junction, and a conical-shaped probe is centered above the junction. The conical probe is tilted away from normal by a small angle as in a commercial SCM, and for this fully 3D case, the collocation method is used in the semiconductor region, while the linear finite-element method is used outside this region. These cases are intended to high-light some of the characteristics that need to be considered in the next generation of models that will be applied to more realistic but complicated 3D configurations.

Formulism

In modeling SCM data of a doped semiconductor wafer, it is useful to review some aspects of the measurement process, the MOS structure, and the approximations that are used to model them [6, 15, 16, 17, 18, 22, 28, 31, 36, 37, 39]. While there are a few different modes of operating an SCM, in each case, the measurement process involves placing a small highly conducting probe-tip near or on the surface of the thin (≈ 10 nm) insulating layer that covers the surface of the doped semiconductor substrate. A bias that contains both a steady-state and a small high-frequency alternating current component is applied between the probe and the semiconductor. The component, ΔV , displaces the electron and hole distributions in the semiconductor slightly away from their biased steady-state values. The resulting capacitance variation $\Delta Q/\Delta V$ is inversely proportional to the probe-to-sample circuit impedance. One mode of operating an SCM is based on the derivative of the high-frequency (HF) capacitance, where the measurement is proportional to dC_{HF}/dV_m , and V_m is the amplitude of a low-frequency modulation voltage. (The low-frequency component has a frequency between 1 kHz and 10 kHz with an amplitude between 0.1 V and 5 V, depending on the oxide thickness and the doping concentration. The high-frequency component has a frequency of 915 MHz with an amplitude of 0.1 V). Here, the bias includes both a low- and a high-frequency component, and the SCM measures the derivative of the HF depletion capacitance. The goal here is to model the derivative of the HF depletion capacitance [16].

In order to correlate SCM data with dopant profiles, $\Delta Q/\Delta V$ must be known as a function of both bias and dopant density [16, 31, 36]. This is a complicated problem, because a number of things may exist or occur in the sample or the measurement procedure that can affect the measurement, and

thus, the modeling. This includes the doping profile, the oxide thickness, interface states in the silicon band gap, light, vibration, etc., [31, 37]. Some of the interface states may be reduced by careful processing. To reduce the effects of any externally applied light and vibration, the SCM measurements are made in the dark in a closed and isolated chamber, and special care is given to the laser monitoring of the probe to prevent any illumination on the semiconductor. In order to make the problem more tractable, some simplification is needed. The model here uses idealized materials and conditions, no externally applied illumination on the semiconductor sample, a continuum model of band bending, no interface states, a uniform oxide thickness, and a uniform doping profile. The model is a first step toward interpreting SCM measurements. The electron and hole distributions in the semiconductor are determined by solving Poisson's equation,

$$\nabla \cdot (\epsilon_r \nabla \psi) = (q/\epsilon_0)(N_d - N_a + p - n), \quad (1)$$

where q refers to the elementary charge (1.602×10^{-19} C), ϵ_0 refers to the relative permittivity of free space (8.854×10^{-18} F/ μm), ϵ_r refers to the relative dielectric constant of the material (11.9 for Si, 3.9 for SiO_2 , and 1.0 for air), N_d refers to the number density of the ionized donor impurity distribution (μm^{-3}), N_a refers to the number density of the ionized acceptor impurity distribution (μm^{-3}), p refers to the number density of the mobile hole distribution (μm^{-3}), n refers to the number density of the mobile electron distribution (μm^{-3}), and ψ refers to the electric potential distribution (V). For the calculations, the zero of the potential is set by the conduction band minimum in the semiconductor substrate far away from the probed surface, $q\psi = E_F - E_C$, where E_F refers to the Fermi level that is constant for a system at equilibrium, and E_C refers to the bottom edge of the conduction band. The carrier number densities n and p are related to the potential ψ through the use of Fermi statistics and the band-bending approximation [6, 17, 31, 36]. (The energetic relations are comparable to that of Grove *et al.* [16], except that: (1) the electrostatic potential is measured from the conduction band minimum in the bulk, whereas Grove *et al.* [16] measured it from the intrinsic Fermi level in the bulk; (2) the surface state charge density is set to zero; and (3) Fermi statistics are used, whereas Grove used Boltzmann statistics.) At room temperature (300K) and concentrations used here, the dopants are fully ionized. Since the doping used here is p type and the capacitance measures displaced majority carriers, the minority carriers can be and are ignored, i.e., $N_d = 0 \approx n \ll p$. The high frequencies used in the measurements preclude the formation of an inversion layer. Since inversion is not allowed, n is negligible.

The electric potential in the insulator and the air is determined by solving Laplace's equation. The problem is then specified by the boundary conditions. Here, it is important to note that the equations must be solved on a domain region

that is sufficiently large so that further changes in the domain size will have little or no effect on the calculated derivative of the HF capacitance. Here, the domain region must contain the probe-tip and the neighborhood around the probe-tip, such as the probe shaft near the probe-tip, the air surrounding the probe, the oxide, and the doped semiconductor.

At the insulator-semiconductor boundary, the potential is continuous, and the discontinuity of the normal component of the electric-displacement vector depends on the trapped interfacial charge. For this work, the interfacial charge density is set to zero. Two Dirichlet boundary conditions are used; one grounds the backplane of the semiconductor, and the other one sets the bias along the probe boundary. The remaining outer boundaries of the domain satisfy Neumann boundary conditions where the normal derivative of the potential is set to zero. (This is the simplest approximation to impose on a supposedly sufficiently distant boundary. It is independent of bias and domain size, and its effect on the solution near the probe-tip ought to be within the “error” of the calculation). To remove the thermal equilibrium work-function difference between the probe and the sub-strate in the figure presentations, the Fermi levels of the probe and the sample are shifted with respect to each other at steady-state by the flat-band voltage, so that zero bias in the figures refers to the flat-band condition in the doped semiconductor sample beneath the probe-tip. This convention follows that of Grove *et al.* [16].

The net charge, Q , in the semiconductor is found by volume integration, i.e.,

$$Q = q \int d^3x (N_d - N_a + p - n). \quad (2)$$

The HF capacitance is determined by subtracting the results from two steady-state solutions with biases that differ by ΔV , and calculating $\Delta Q/\Delta V$. The HF capacitance is calculated for a range of biases and spline fitted. The derivative of the HF capacitance is found by differentiating the spline curve. These considerations guided the model calculations that are reported here.

Uniform Doping

Geometry

When the configuration of the SCM measurement is such that the conical-shaped probe exhibits cylindrical symmetry, the central axis of the probe is oriented in a direction normal to the surface of the sample, and the sample is uniformly doped, the geometry of the combined system exhibits cylindrical symmetry, and the system can be modeled as a 2D problem. Here, the probe shape is modeled after a commercially available probe. The probe is conically shaped with a rounded tip; the probe-tip radius of curvature is $0.01 \mu\text{m}$, and the cone apex half-angle is 10° . {For sake of easy modeling of the

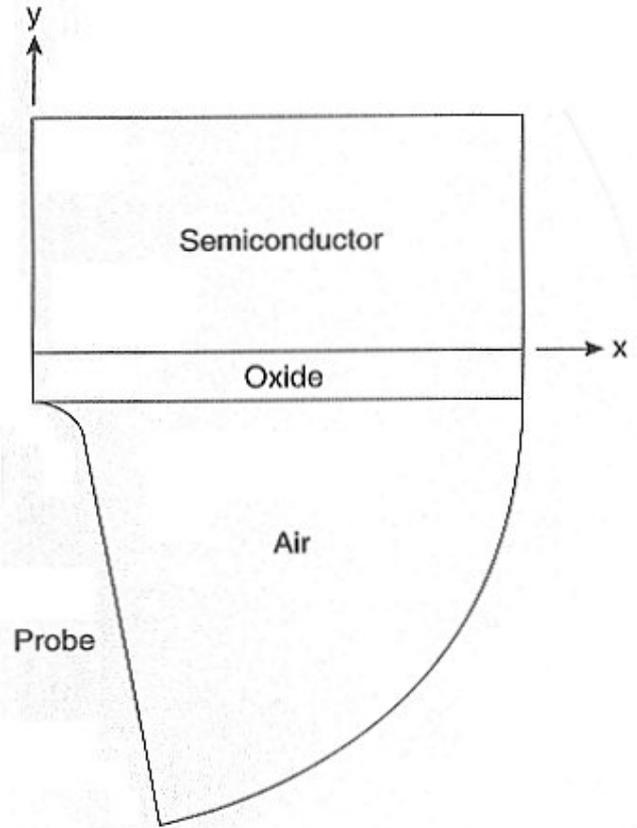


Figure 1. Geometry of the model structure. The three re-gions are the semiconductor region ($0 \leq y$), the oxide region ($-h_{\text{ox}} \leq y \leq 0$), and the ambient region ($y \leq -h_{\text{ox}}$), where h_{ox} refers to the oxide thickness. The oxide thickness is $0.01 \mu\text{m}$, and the probe-tip radius of curvature is $0.01 \mu\text{m}$. The radial cutoff distance is $0.1 \mu\text{m}$; the semiconductor substrate thickness is $0.05 \mu\text{m}$.

contact region between the probe-tip and the oxide boundaries, the probe-tip is blunted slightly by truncation (plane intersection), so that the angle between the two intersecting surfaces is 10° . The probe length is set by the radial cutoff distance, as explained later.

The model domain contains three subregions: the semiconductor substrate region, the insulating oxide layer region, and the air or ambient region surrounding the probe. A cross section of the model geometry for the probe-sample structure is shown in Figure 1. The coordinate system is chosen such that the central axis of the cylindrical coordinate system is the y axis. The $y=0$ plane forms the SiO_2 -Si interface boundary, the positive y axis is directed into the substrate region, and the negative y axis passes through the probe axis. The radial direction is along the x axis. This is upside down

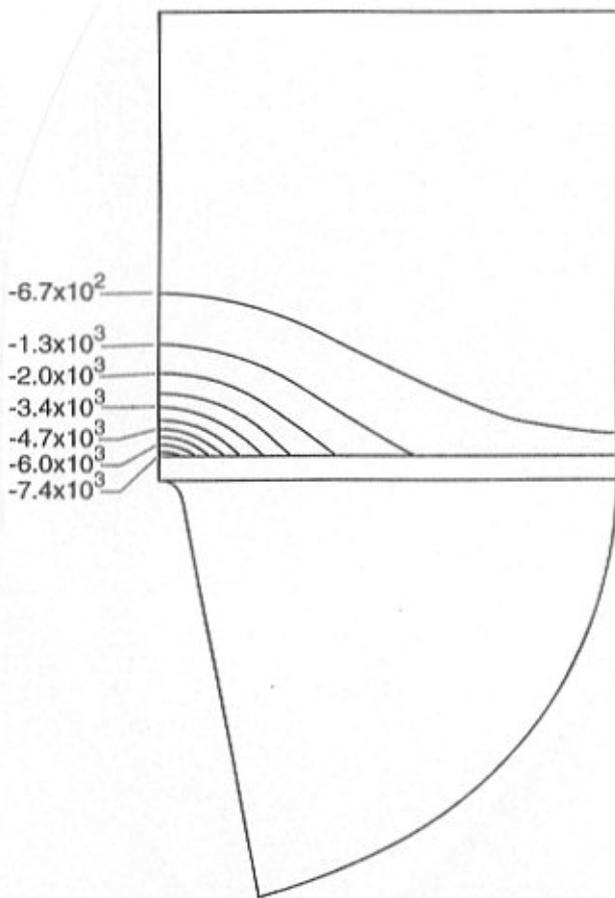


Figure 2. Contour plot of the net charge density distribution, where the dopant density is $1 \times 10^4 \mu\text{m}^{-3}$, and the bias is 0.2 V. The radial cutoff distance is 0.2 μm ; the semiconductor substrate thickness is 0.2 μm . Contour levels are expressed in units of μm^{-3} .

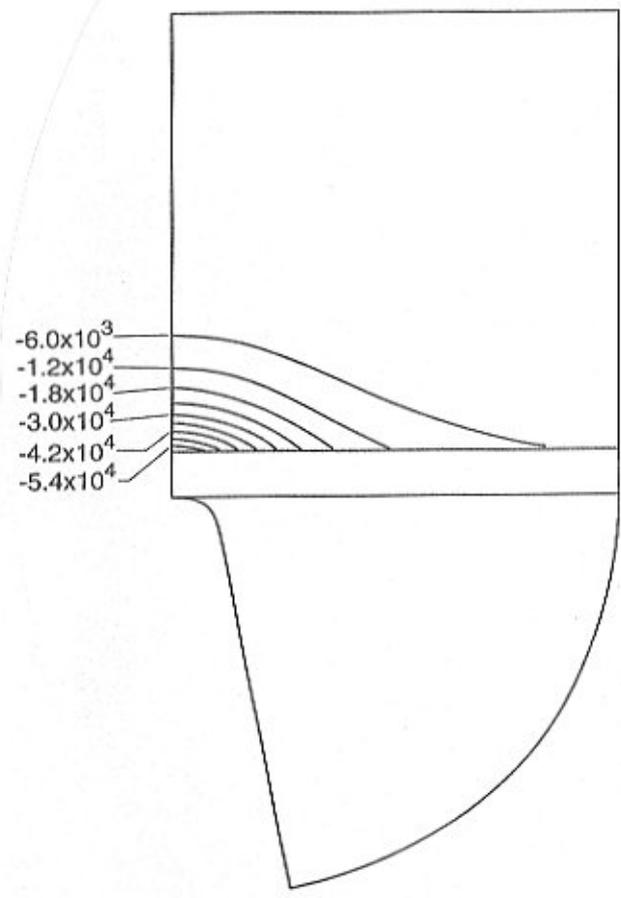


Figure 3. Contour plot of the net charge density distribution where the dopant density is $1 \times 10^5 \mu\text{m}^{-3}$, and the bias is 0.2 V. The radial cutoff distance is 0.1 μm ; the semiconductor substrate thickness is 0.1 μm . Contour levels are expressed in units of μm^{-3} .

from the usual SCM configuration. The semiconductor region is where $y \geq 0$, the oxide region is where $-h_{\text{ox}} \leq y \leq 0$, and the ambient region is where $y \leq -h_{\text{ox}}$, where h_{ox} is the oxide layer thickness. The unit of length is expressed in μm , and here, $h_{\text{ox}} = 0.01 \mu\text{m}$.

The size of the substrate region is set in part by the radial cutoff distance or length of the x axis. Since the interest here is to maintain the spatial resolution of the measurement near that of the probe-tip radius of curvature, the radial cutoff distance was set to usually 10 times the probe-tip radius of curvature. Here, the radial cutoff distance was set to 0.2 μm when $1 \times 10^4 \mu\text{m}^{-3} \leq N_a \leq 9 \times 10^4 \mu\text{m}^{-3}$; 0.1 μm when $1 \times 10^5 \mu\text{m}^{-3} \leq N_a \leq 3 \times 10^7 \mu\text{m}^{-3}$; and 0.05 μm when $4 \times 10^7 \mu\text{m}^{-3} \leq N_a \leq 1 \times 10^8 \mu\text{m}^{-3}$.

The substrate depth (y) cutoff was determined so that the substrate region could contain the depletion region, and the charge neutrality condition could be maintained deep inside the substrate for the given maximum bias. The maximum bias was roughly set by determining when the contour value of $N_a/10$ reached the radial cutoff distance.

The length of the probe is determined by using a circular arc to form the outer boundary of the ambient region and requiring the arc to intersect perpendicularly with the boundaries of the probe and the oxide. Therefore, the probe length is set by the radial cutoff distance. Changing the radial cutoff distance changes the probe length, and for a nonzero bias, this changes the charge on the probe, the net charge in the semiconductor, and the capacitance. However, when the

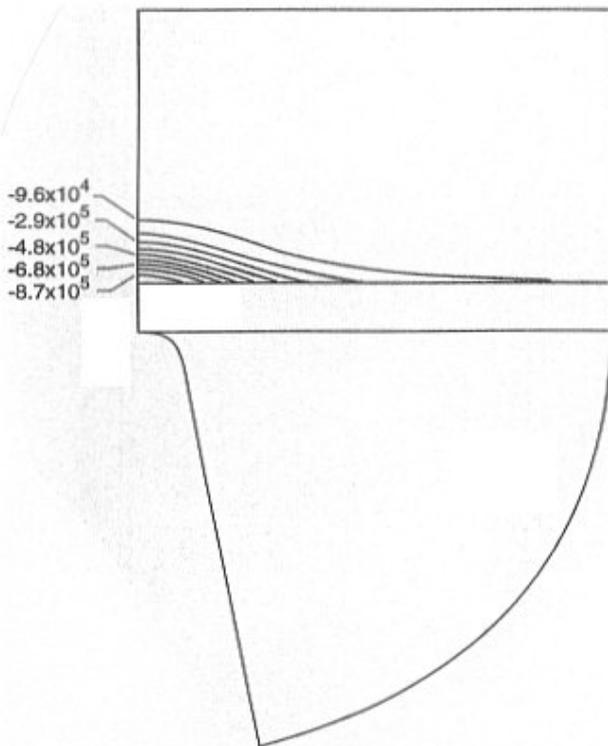


Figure 4. Contour plot of the net charge density distribution where the dopant density is $1 \times 10^6 \mu\text{m}^{-3}$, and the bias is 1.2 V. The radial cutoff distance is $0.1 \mu\text{m}$; the semiconductor substrate thickness is $0.06 \mu\text{m}$. Contour levels are expressed in units of μm^{-3} .

radial cutoff distance is sufficiently large, the change in the derivative of the high-frequency capacitance is found to be small and is within the estimated error of the calculation.

This condition, where the calculated derivative of the high-frequency capacitance becomes insensitive to changes in the size of the domain, is both important and necessary for modeling an SCM measurement where the objective is to determine a meaningful absolute measurement and not just a relative measurement. Conversely, for an SCM measurement to be practical, the derivative of the high-frequency capacitance must be insensitive to and separable from the stray capacitances in the system.

Method of solution

To solve both the nonlinear Poisson equation in the semiconductor region and the Laplace equation in the oxide and the ambient regions, we used PLTMG (Piece-wise Linear Triangular finite-element MultiGrid) [5], a software package for solving elliptic partial differential equations for scalar

problems in two dimensions. The package provides support for adaptively refining the mesh and for plotting the contours or the surface profile of the solution or a function of the solution.

To calculate the capacitance as a function of applied bias V_B , the depletion region must be suitably meshed to have an accurate volume integration to find the net charge in the semiconductor. To help reduce the grid dependence in the charge calculations, the procedure used here was to find one grid that would suitably mesh the depletion region at the largest allowed bias setting and then use that mesh for solutions at other smaller biases, starting from deep depletion ($V_B > 0$) and moving to accumulation ($V_B < 0$). This is done for each dopant density considered in the work.

Finding a suitable mesh over the maximal depletion region was difficult. The default adaptive meshing algorithm used in PLTMG was found to mesh the region around the probe-tip in the ambient region quite well, but only at the expense of the oxide and the substrate regions; they were meshed too coarsely. Because PLTMG provides little direct control of the mesh step size, and the user options for the adaptive meshing algorithm are limited, the only way to cause PLTMG without modification to form a different mesh is to perturb the equation that PLTMG is trying to solve. One method for improving the mesh was found by: (1) equalizing the media by setting the relative dielectric constants to one; (2) perturbing the doping profile near the oxide-semiconductor surface to force the meshing algorithm to sense the nonuniformity; and (3) solving this perturbed problem at a bias of 5 V or 10 V beyond that for the maximal depletion region used in the capacitance calculations. Here, the perturbed dopant density is allowed to vary quadratically in the x direction and decay normally (Gaussian) in the y direction; the perturbation is localized near the SiO_2 -Si boundary. The perturbation is used only during the mesh refinement procedure; after finding a suitable or much improved mesh, the correct equations are solved to find the capacitance.

The mesh is undoubtedly one source of error in the calculation. Some error may be expected at larger biases when the depletion region penetrates into regions that are less densely meshed. Estimating this kind of error with PLTMG is difficult. One method would involve halving the stepsize of the finest mesh, but this is not possible, because PLTMG provides no option for uniformly refining the finest mesh, unlike for the initial coarse mesh. The only option available then is to add points to the mesh. For meshes containing 1×10^4 and 2×10^4 points, the derivative high-frequency capacitance curves were compared and were found to differ by less than 0.5% and 3% for low and high dopant densities of $1 \times 10^5 \mu\text{m}^{-3}$ and $1 \times 10^8 \mu\text{m}^{-3}$, respectively. While the number of mesh points is important, this approach is somewhat open to question, because there is little user control over the

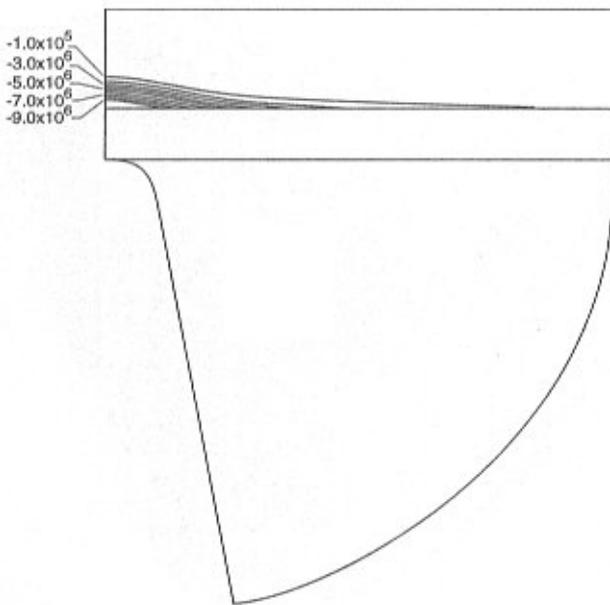


Figure 5. Contour plot of the net charge density distribution where the dopant density is $1 \times 10^7 \mu\text{m}^{-3}$, and the bias is 4.4 V. The radial cutoff distance is $0.1 \mu\text{m}$ the semiconductor substrate thickness is $0.02 \mu\text{m}$. Contour levels are expressed in units of μm^{-3} .

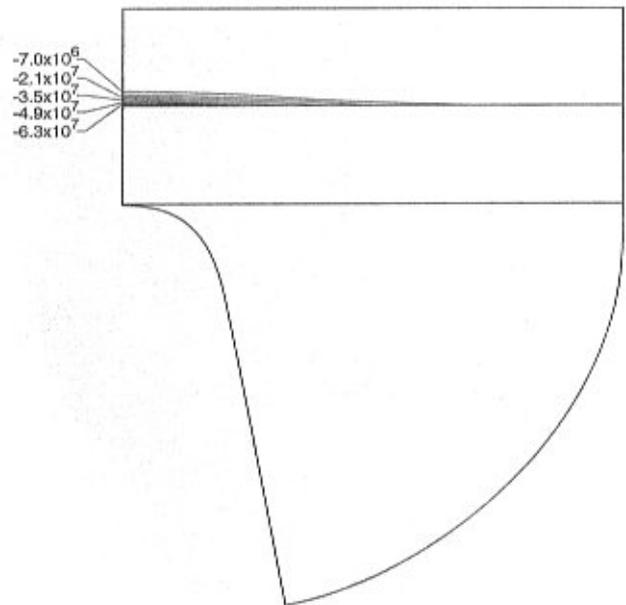


Figure 6. Contour plot of the net charge density distribution where the dopant density is $1 \times 10^8 \mu\text{m}^{-3}$, and the bias is 5.0 V, the breakdown voltage of the oxide. The radial cutoff distance is $0.05 \mu\text{m}$ the semiconductor substrate thickness is $0.01 \mu\text{m}$. Contour levels are expressed in units of μm^{-3} .

placement of the points. However, the results are encouraging. For this work, the meshes used 2×10^4 points.

Another contribution to the error of the calculation involves satisfying the boundary conditions. As the bias is increased above zero, the depletion region forms and expands along the SiO_2 -Si boundary with a finite skin-depth. At large biases, the depletion region can expand beyond the radial cutoff distance and move beyond the boundaries of the domain, and of course, any net charge outside the domain is neglected in the capacitance calculations. The maximum bias used in a capacitance calculation was determined by observing the depletion region expand along the oxide boundary as a function of bias until the net number density at the radial cutoff distance became nearly 10% of the dopant density. The HF capacitance and the derivative of the HF capacitance may be expected to exhibit a larger percentage of this kind of error at the larger biases. It follows then that the greatest accuracy and resolution occur when the depletion region is small, when the bias is small, i.e., near the flat-band condition.

Results of calculations

In order to help interpret capacitance data and the model accuracy, we show, in Figures 2 through 6, contour

plots of the net charge density distribution in the semiconductor calculated as a function of dopant density and applied bias. The dopant density and the applied bias are: $1 \times 10^4 \mu\text{m}^{-3}$ and 0.2 V in Figure 2; $1 \times 10^5 \mu\text{m}^{-3}$ and 0.2 V in Figure 3; $1 \times 10^6 \mu\text{m}^{-3}$ and 1.2 V in Figure 4; $1 \times 10^7 \mu\text{m}^{-3}$ and 4.4 V in Figure 5; and $1 \times 10^8 \mu\text{m}^{-3}$ and 5.0 V in Figure 6. The latter bias is the approximate breakdown voltage of the oxide.

The shape of the depletion region depends on the bias and the dopant density via the characteristic screening length in the semiconductor. The shape of the depletion region is more hemispherical at lower dopant densities and flatter at higher dopant densities due to screening. The size and shape of the depletion region sets the resolution of the measurement, and this sets the maximum bias of the measurement. In calculations with dopant densities below $1 \times 10^5 \mu\text{m}^{-3}$, it was found that the depletion expanded with bias into regions less densely meshed in the bulk than that near the surface, so that the volume integral of the calculated charge was less accurate.

Figures 7 to 10 present the results of calculations of the capacitance and the derivative of the capacitance for a range of dopant densities from $1 \times 10^5 \mu\text{m}^{-3}$ to $1 \times 10^8 \mu\text{m}^{-3}$ in

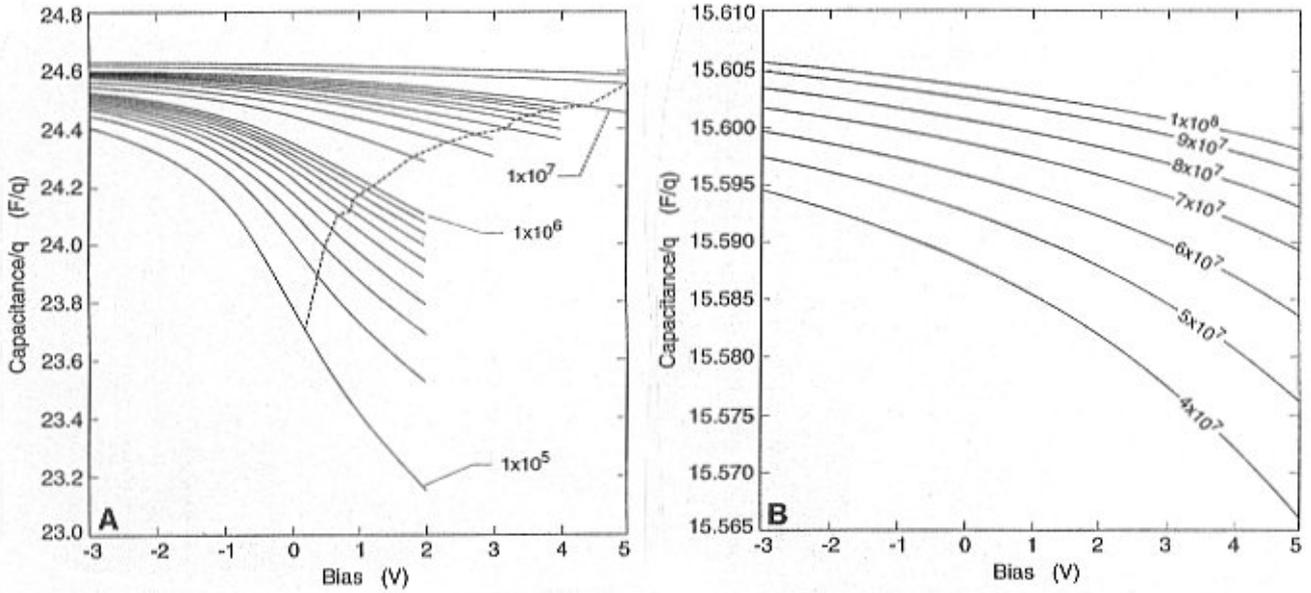


Figure 7. The high-frequency capacitance (solid line) calculated for a range of dopant densities N_a . The dopant density range is: (A) 1×10^5 to $3 \times 10^7 \mu\text{m}^{-3}$, where the radial cutoff distance is $0.1 \mu\text{m}$ and (B) 4×10^7 to $1 \times 10^8 \mu\text{m}^{-3}$, where the radial cutoff distance is $0.05 \mu\text{m}$. The dotted line in (A) refers to the approximate maximum bias for which the calculation is valid. The range is stepped in unit increments per decade. The capacitance is scaled by the elementary charge, $q = 1.602 \times 10^{-19} \text{ C}$.

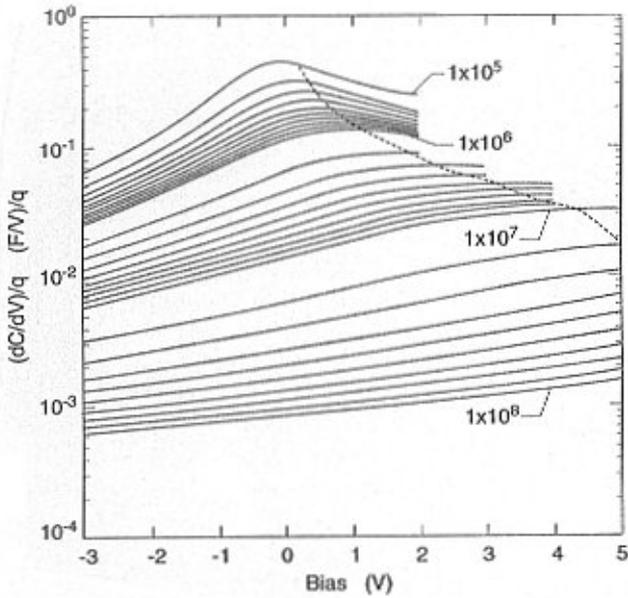


Figure 8. The derivative high-frequency capacitance (solid line) and the approximate maximum bias (dotted line) calculated for a range of dopant densities N_a . The dopant density range is 1×10^5 to $1 \times 10^8 \mu\text{m}^{-3}$. The range is stepped in unit increments per decade.

cutoff distance were used in the calculations for this large range of dopant densities. (The domain size is set by the radial cutoff distance, the radial cutoff distance sets the probe length, and the probe length controls the stray capacitance). Figures 7A and 7B present the capacitance calculations where the radial cutoff distances are $0.1 \mu\text{m}$ and $0.05 \mu\text{m}$, respectively. Figure 8 presents the calculated derivative HF capacitance. The dotted line in Figures 7A and 8 shows the approximate maximum bias for which the calculations are believed to be valid; the depletion region expands outside the domain at larger biases. Two noticeable features of Figure 8 are that the curves tend to peak at values of positive bias (depletion), except for the lowest dopant density ($1 \times 10^5 \mu\text{m}^{-3}$), and that the peaks move with dopant density toward higher bias. The smooth spacing between the curves in Figure 8 suggest that the domain size and the adaptive meshing density were suitably determined so that the calculated derivative capacitance curves appear to be insensitive to the changes in the domain size. This is a necessary condition, otherwise the modeling and the measurement would be impractical.

Figure 9 presents derivative capacitance calculations showing sensitivity to the dopant density variations, the domain size, and the mesh. Curves A, C, and D involve different meshing parameters used during the adaptive mesh refinement procedure, but with the same dopant density. Curves A and E involve similar meshing parameters, but with different dopant densities. Curves A and B involve the same meshing parameters, same dopant density, but with different domain

unit increments per decade. Two different values of the radial

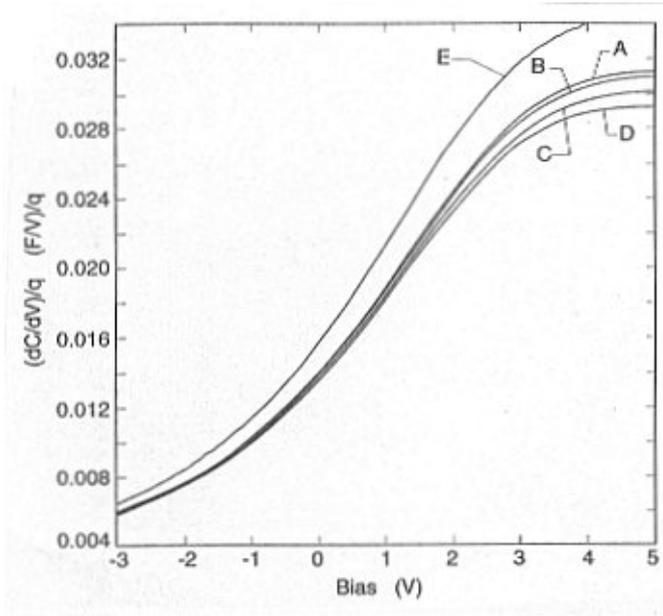


Figure 9. Sensitivity of the calculated derivative high frequency capacitance. The dopant density is $1 \times 10^7 \mu\text{m}^{-3}$ for curves (A, B, C and D) and is $9 \times 10^6 \mu\text{m}^{-3}$ for curve (E). The radial cutoff distance is $0.1 \mu\text{m}$ for curves (A, C, D, and E) and is $0.05 \mu\text{m}$ for curve (B). Curves (A, C, and D) involve different meshing parameters, but the same dopant density; curves (A and E) involve similar meshing parameters, but different dopant densities; and curves (A and B) involve different radial cutoff values.

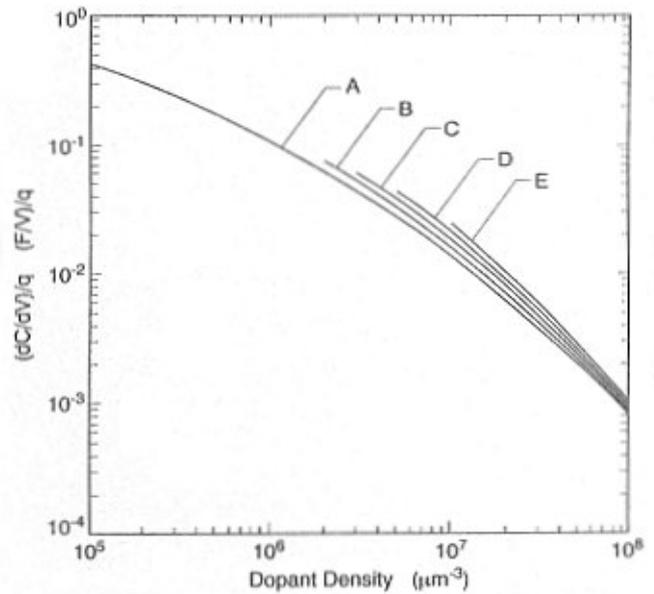


Figure 10. Conversion curves relating dopant density with SCM derivative high-frequency capacitance data. Curves A, B, C, D, and E refer to bias settings of (0.0, 0.5, 1.0, 1.5, and 2.0) V, respectively.

sizes. The dopant density is $1 \times 10^7 \mu\text{m}^{-3}$ for curves A, B, C and D, and is $9 \times 10^6 \mu\text{m}^{-3}$ for curve E. The radial cut-off distance is $0.1 \mu\text{m}$ for curves A, C, D, and E, and is $0.05 \mu\text{m}$ for curve B. The closeness of curves A and B ($< 2\%$) re-veals that the domain size is sufficiently large and the probe-tip is suitably represented so that the calculation is relatively insensitive to small changes in the domain size. The distance between curves A and D ($\approx 8\%$) reveals the sensitivity of the calculation on the domain mesh. A better mesh would reduce this value. The distance between curves A and E reveals the sensitivity of the method in determining dopant density at this concentration. Curves A and E are used in Figure 8.

Figure 10 presents conversion curves relating dopant density with derivative HF capacitance data for a representative set of biases. Curves A, B, C, D, and E refer to bias settings of 0.0, 0.5, 1.0, 1.5, and 2.0 V, respectively. The curves differ in length as a function of positive bias. This is due to limiting the maximum domain size in the calculations. For a given domain size and positive bias, there is a critical value of dopant density below which the depletion region will

expand beyond the bounds of the domain, and the capacitance calculations would not be valid. (Recall that the dotted line in Figure 8 indicates the approximate maximum bias for which the calculations are believed to be valid). The curvature and the bias dependence of the conversion curves reveal the importance of reproducing identical measurement parameters in order to extract accurate values of the dopant density from SCM measurements.

Database

Since SCM data are a smooth function of geometrical parameters, such as the probe-tip radius of curvature, the probe cone apex angle, and the oxide thickness, it is then possible to form a database set of calculations with which one may use linear interpolation to quickly analyze a set of SCM measurements [23]. Such a database is planned for future work.

Graded Doping

One dimension

The SCM technique can image a semiconductor with contrast to variations in dopant density and spatial resolution of the order of the probe-tip diameter. The previous section sought to interpret SCM measurements of materials where the doping is uniform, but there is also a need for interpreting

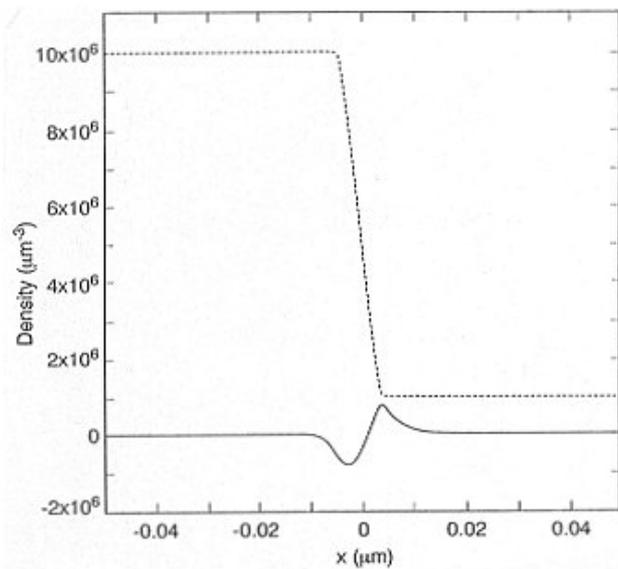


Figure 11. Dopant density calculations for the 1D model high/low like-dopant graded profile junction. The net charge density distribution is given by the solid line, and the dopant density distribution is given by the dotted line. The transition region is centered at $x = 0$ and its full width is $0.01 \mu\text{m}$.

SCM measurements of materials where the doping is not uniform, and the interest is in scanning across graded dopant junction profiles. While the pn-junction may be one of special interest, this paper restricts its consideration to one example of a model high/low like-dopant graded profile junction. The goal here is to initiate a preliminary investigation of the charge distribution near a junction and its redistribution when probed during an SCM measurement.

The model high/low like-dopant graded profile junction, which is defined here and used in the remainder of the paper, is presented as the dotted line profile in Figure 11. The dopant profile is characterized by variation in one direction, the x direction, with different left- and right-side values of the dopant density, and a graded dopant profile in the transition region. The transition region is centered near the origin, $x = 0$, and its diameter is set to the value of the probe-tip radius of curvature of the previous sections, 10 nm . The acceptor dopant density is $1 \times 10^7 \mu\text{m}^{-3}$ for $x \leq 0.005 \mu\text{m}$, and is $1 \times 10^6 \mu\text{m}^{-3}$ for $x \geq 0.005 \mu\text{m}$. The graded dopant profile is approximated by a cubic-spline interpolating polynomial that matches both the function and the first derivative at the endpoints of one interval for the transition region, i.e., Hermite's formula [35]. The dopant profile function and its first derivative are continuous.

The software package COLSYS [2, 3, 4] was used to solve Poisson's equation in one-dimension for the equilibrium

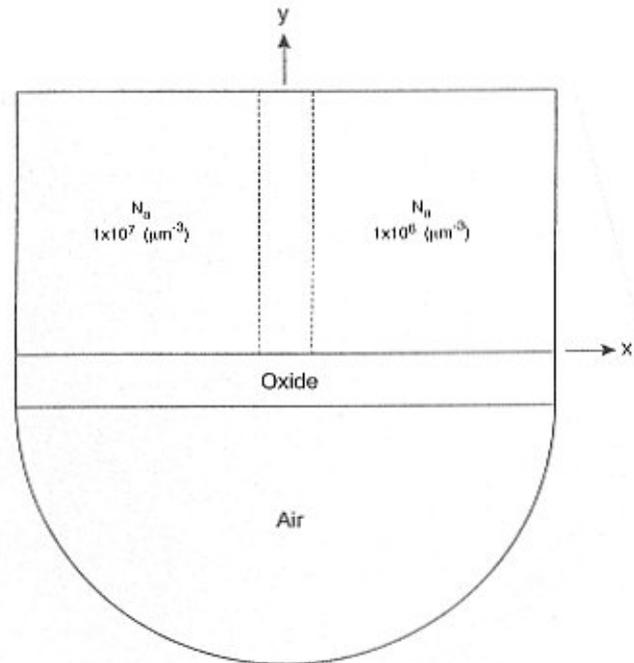


Figure 12. Geometry of the 2D model high/low like-dopant graded profile junction with no probe near the junction. The semiconductor region is given by $-0.05 \leq x \leq 0.05$ and $0.0 \leq y \leq 0.05$; the oxide region is given by $-0.01 \leq y \leq 0$; and the air or ambient region is the semicircular bounded region where $y \leq -0.01$. The two dotted lines bound the transition region, $-0.005 \leq x \leq 0.005$. There is translational invariance in the direction normal to the plane of the figure.

solution in the semiconductor. Figure 11 presents the results of this calculation for the net charge density distribution (solid line); negative values indicate depletion and positive values indicate accumulation. Holes near the junction move from regions of high density to regions of low density, as is well known. The calculated electric potential distribution is the equilibrium solution deep in the bulk of the semiconductor, and this is used to specify the Dirichlet boundary condition along the grounded backplane of the semiconductor for the 2D and 3D problems that are discussed in the following sections.

Two dimensions, no probe

This section considers the charge distribution near a high/low like-dopant graded profile junction of a model 2D structure without the presence of an SCM probe. The model geometry is shown in Figure 12. The semiconductor region is given by $-0.05 \leq x \leq 0.05$ and $0.0 \leq y \leq 0.05$; the oxide region is given by $-0.05 \leq x \leq 0.05$ and $-0.01 \leq y \leq 0$; and the air or

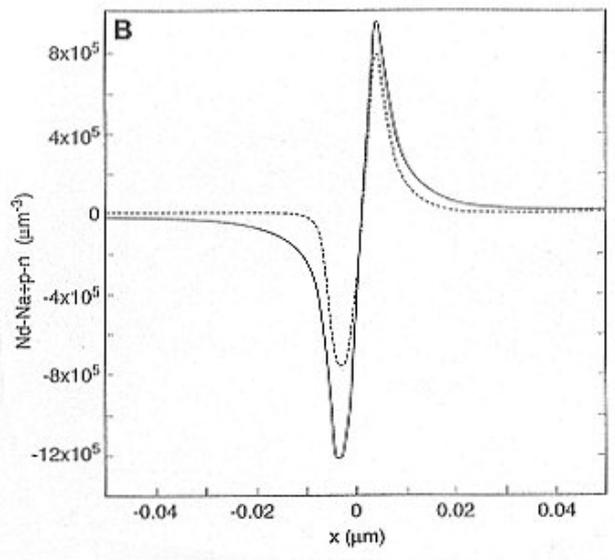
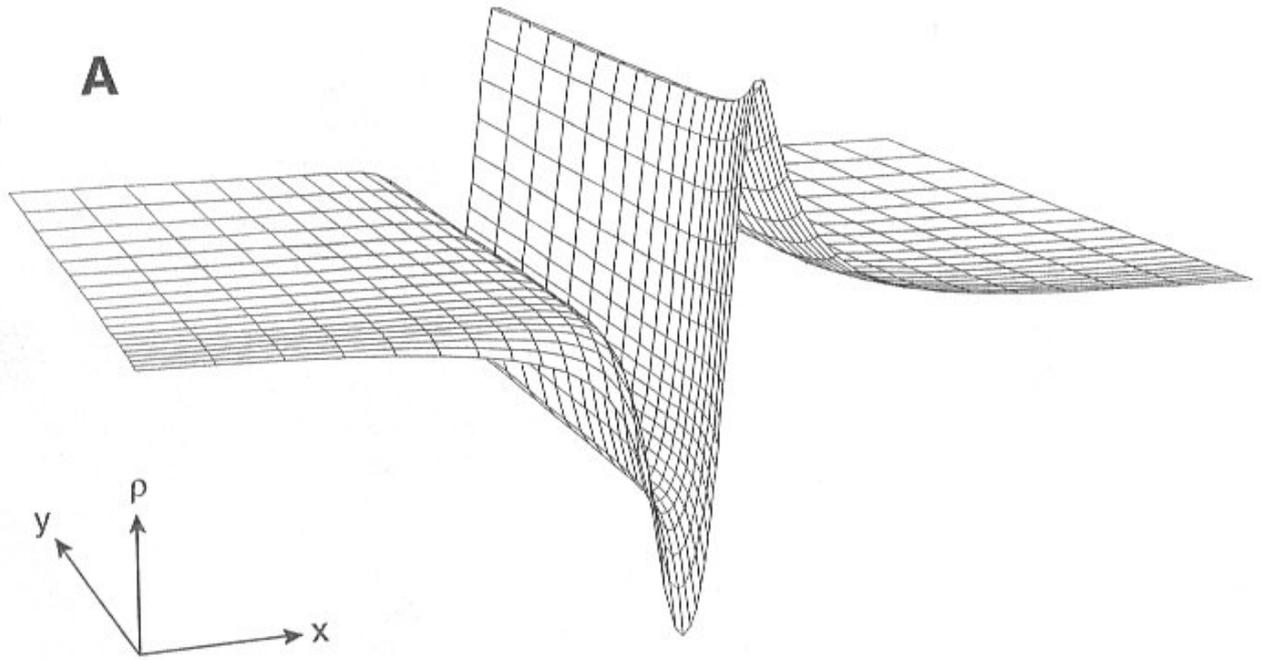


Figure 13. Results of calculations for the 2D model high/low like-dopant graded profile junction with no probe near the junction. **(A, across at top)** Surface profile of the net charge density distribution, and **(B)** line profiles of the net charge density distribution at $y = 0$ (solid line) and at $y = 0.05$ (dotted line). Net charge density values less than zero indicate depletion and values greater than zero indicate accumulation.

ambient region is the semicircular bounded region where $y \leq -0.01$.

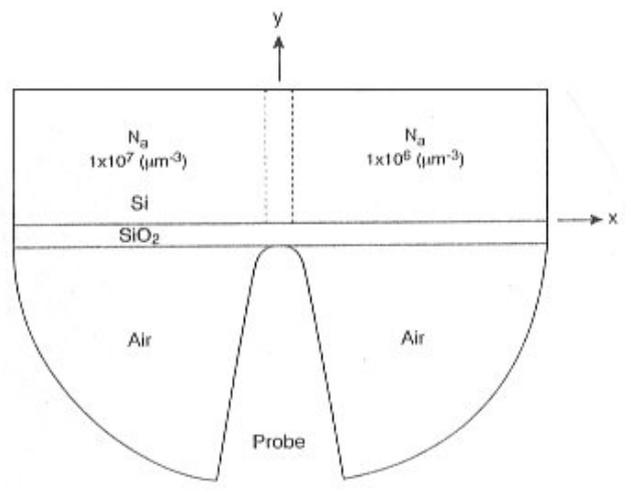


Figure 14. Geometry of the 2D model high/low like-dopant graded profile junction with a V shaped probe centered above and along the junction. The semiconductor region is given by $-0.1 \leq x \leq 0.1$ and $0.0 \leq y \leq 0.05$; the oxide region is given by $-0.01 \leq y \leq 0.0$; and the air or ambient regions border the probe region where $y \leq -0.01$. The two dotted lines bound the transition region, $-0.005 \leq x \leq 0.005$. There is translational invariance in the direction normal to the plane of the figure.

The software package PLTMG was used to solve Poisson's equation. The electric potential distribution of the previous section was used to specify a Dirichlet boundary

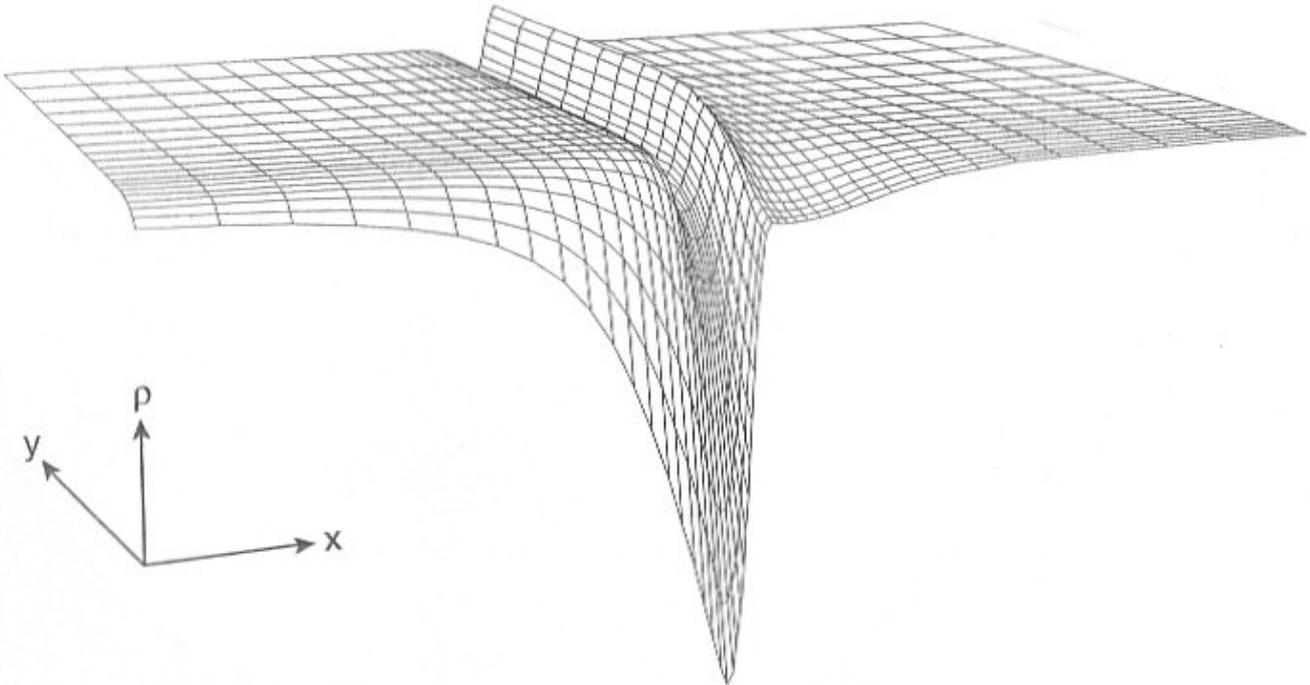


Figure 15. Surface profile of the net charge density distribution calculated for 1.0 V bias for the model shown in Figure 14. Net charge density values less than zero indicate depletion and values greater than zero indicate accumulation.

condition along the grounded backplane ($y = 0.05$) of the semiconductor region. Neumann boundary conditions were used along the remaining outer boundaries. The dopant density distribution in the semiconductor region is that of the previous section.

Calculated results are shown in Figure 13. Figure 13A presents a surface profile of the net charge density distribution, and Figure 13B presents line profiles of the net charge density distribution across the junction along the SiO_2 -Si interface at $y = 0$ (solid line) and along the Si backplane at $y = 0.05$ (dotted line). Net charge density values less than zero indicate depletion, and values greater than zero indicate accumulation. Some charge redistribution occurs near the junction near the SiO_2 -Si boundary in accord with the spreading of the electric field distribution; the depletion peak appears to be more enhanced than that for accumulation.

Two dimensions, 2D probe

This section considers the charge distribution near a high/low like-dopant graded profile junction with a biased V shaped probe centered above and along the junction. This is a 2D problem. The geometry of the model is shown in Figure 14. The semiconductor region is given by $-0.1 \leq x \leq 0.1$ and $0.0 \leq y \leq 0.05$; the oxide region is given by $-0.1 \leq x \leq 0.1$ and $-0.01 \leq y \leq 0.0$; and two air or ambient regions border the probe region where $y \leq -0.01$.

The software package PLTMG was used to solve Poisson's equation. Dirichlet boundary conditions are used to specify the electric potential along the grounded backplane ($y = 0.05$) of the semiconductor region deep in the bulk as in the previous section and along the boundary of the probe to specify the applied bias. Neumann boundary conditions are used along the remaining outer boundaries. The dopant density distribution in the semiconductor region is that of the previous section.

The results of the calculation are presented in Figure 15; the surface profile of the net charge density distribution is shown for a 1.0 V bias. Net charge density values less than zero indicate depletion, and values greater than zero indicate accumulation. For increasing positive bias, the depletion region expands near the surface toward the higher dopant density region; see Figure 13 for comparison. The accumulation region located near the surface and the lower doped side of the junction is depressed. The presence of the junction lessens the expansion of the depletion region into the lower doped region. The geometry is similar to that of parallel capacitance.

Two dimensions, 3D probe

This section considers the charge distribution near a high/low like-dopant graded profile junction with a biased conical-shaped probe centered above the junction, but

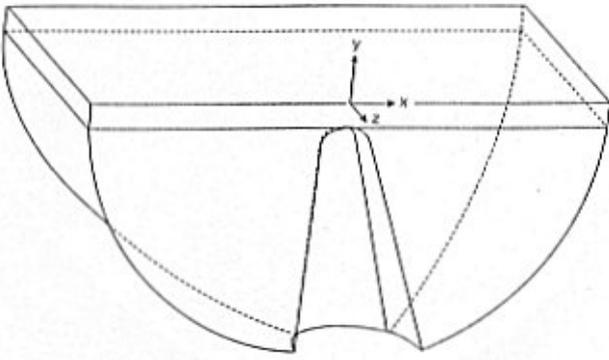


Figure 16. Geometry of the model 3D structure for the oxide and the air or ambient regions with a conical-shaped probe centered above the junction. The oxide region is given by $-0.1 \leq x \leq 0.1$, $-0.01 \leq y \leq 0.0$, and $-0.1 \leq z \leq 0.0$. The ambient region surrounds the probe region and is the bounded region where $y \leq -0.01$. The probe-tip radius of curvature is $0.01 \mu\text{m}$, the probe cone apex half-angle is 10° ; and the central axis of the probe is tilted 10° away from normal toward the x axis. The semiconductor region (not presented) is adjacent to the oxide region and is given by $-0.1 \leq x \leq 0.1$, $0.0 \leq y \leq 0.05$, and $-0.1 \leq z \leq 0.0$; the dopant junction is centered at $x=0$, and the dopant distribution is translational invariant in both y and z directions.

oriented 10° off the normal to the oxide surface. This is a 3D problem. Because software was not available that could both solve the nonlinear Poisson equation in the semiconductor and model the geometry of the air or ambient region surrounding the probe, it was necessary to break the problem into two parts.

In order to model the geometry of both the oxide region and the ambient region and solve Laplace's equation, the finite-element software package ANSYS [1] was used. In order to model the semiconductor region and solve Poisson's equation, the collocation software package ESPDESC (Elliptic Systems of Partial Differential Equations Solved by Collocation) [26] was used. The solution is found by relaxing and matching the boundary conditions along the oxide-semiconductor boundary, and iterating to self-consistency [26, 27, 32, 33, 34]. (Relaxing refers to an iterative procedure for determining the boundary conditions between two bordering domain regions; the solutions from two successive iterations are averaged before proceeding with the next iteration). From this, the capacitance may be found [8, 9, 26]. The model geometry of the oxide and the ambient regions is presented in Figure 16. The oxide region is given by $-0.1 \leq x \leq 0.1$, $-0.01 \leq y \leq 0.0$, and $0.1 \leq z \leq 0.0$; and the ambient region is the bound region where $y \leq -0.01$. The probe-tip radius of curvature is $0.01 \mu\text{m}$; the

probe cone apex half-angle is 10° and the central axis of the probe is tilted 10° away from normal toward the x axis as in a commercial instrument.

The dopant distribution is that of the previous section. The dopant junction is centered at $x=0$, and the dopant distribution is translational invariant in both y and z directions. The semiconductor region is given by $-0.1 \leq x \leq 0.1$, $0.0 \leq y \leq 0.05$, and $-0.1 \leq z \leq 0.0$.

The boundary conditions are set as in the previous section. Dirichlet boundary conditions are used to specify the electric potential along the grounded backplane ($y=0.05$) of the semiconductor region deep in the bulk and along the boundary of the probe for the applied bias voltage. Neumann boundary conditions are used along the remaining outer boundaries.

The results of the calculation are shown in Figure 17; the surface profile of the net charge density distribution is shown for the $z=0$ plane for 1.0 V bias. Net charge density values less than zero indicate depletion, and values greater than zero indicate accumulation.

Again, as in the previous section, for increasing positive bias, the depletion region expands near the surface toward the higher dopant density region; see Figures 13 and 15 for comparison. The depletion region is more localized near the probe-tip in Figure 17 than in Figure 15, as may be expected to occur when comparing the results of a 2D and a 3D probe. The accumulation region located near the surface and the lower doped side of the junction is depressed. The presence of the junction lessens the expansion of the depletion region into the lower doped region.

Summary

In order to help correlate SCM data with semiconductor dopant concentrations, model capacitance curves have been calculated for silicon. For 2D cases, the linear finite-element method is used to solve Poisson's equation in the semiconductor region and Laplace's equation in the oxide and the ambient regions. For 3D cases, the collocation method is used in the semiconductor region, and the linear finite-element method is used outside this region. For a given oxide thickness, probe shape, and probe-tip size, the high-frequency capacitance is calculated for cases of uniform doping for a range of dopant concentrations. The derivative high-frequency capacitance is calculated and is used to form a set of conversion curves relating dopant density with derivative high-frequency capacitance data. For uniform doping, the theory can be used to form a data-base for rapid interpretation of SCM measurement data.

For a model high/low like-dopant graded profile junction, the net charge density distribution is found for cases with no probe, a V shaped probe, and a conical-shaped probe centered near the junction. The presence of the junction

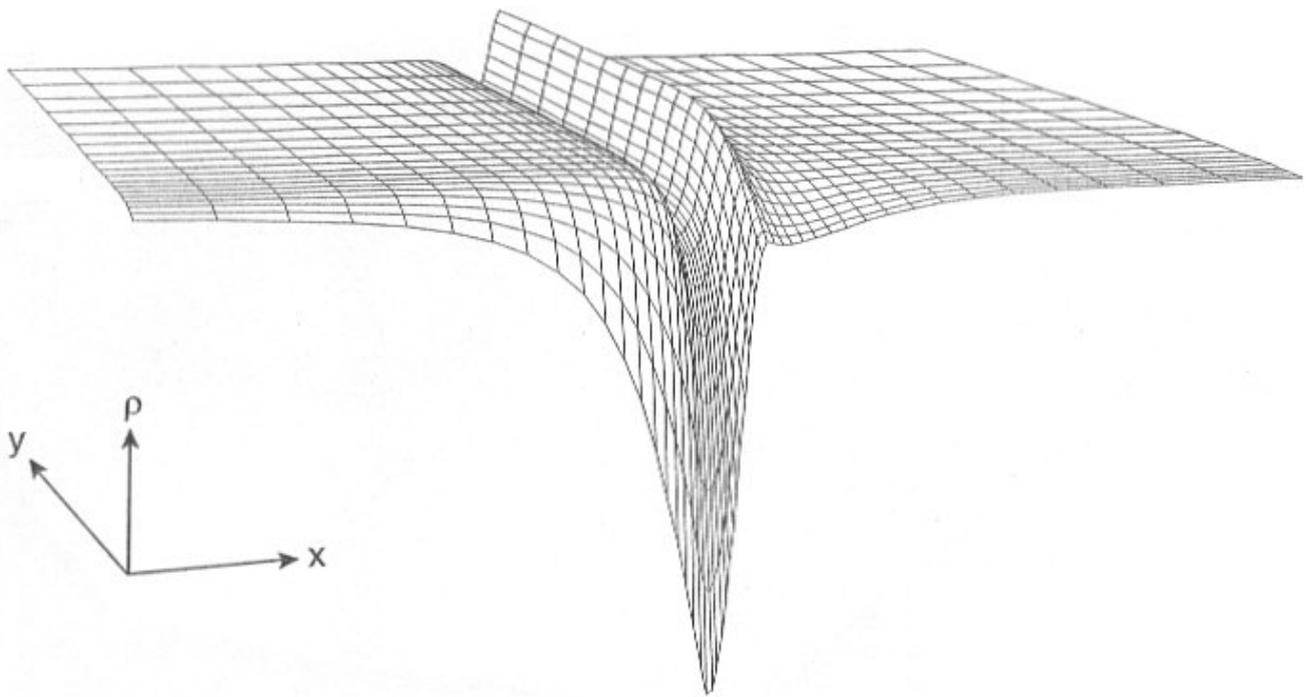


Figure 17. Surface profile of the net charge density distribution in the $z = 0$ plane calculated for 1.0 V bias for the model involving Figure 16. Net charge density values less than zero indicate depletion and values greater than zero indicate accumulation.

lessens the expansion of the depletion region into the lower doped region. The geometry is similar to that of parallel capacitance, and understanding the role of the equilibrium charge distribution near the junction is important in determining the doping profile around the junction.

Appendix

Overview of models

A number of models have been used in the literature to interpret SCM data, and they differ by their degree of approximation. The fast models tend to use analytic models such as the 1D metal-oxide-semiconductor capacitor (MOSC) model [16], and the slower models tend to use 2D or 3D finite-element method (FEM) calculations. This section considers five models; two are slow and three are fast. The two slow FEM-based models compare effects due to probe shape and provide a reference with which to compare the results of the three fast 1D MOSC-based models. For a given geometric structure, the models are explained, and the curves of the derivative of the high-frequency capacitance are calculated and plotted together for comparison.

The domain geometry is shown in Figure 1 and is the same as that presented for the case of uniform doping. The

probe-sample system exhibits cylindrical symmetry with dependence only along the radial and cylindrical axes. The probe shape is conical with a rounded tip. The probe-tip rests on the oxide surface, and the probe-axis is oriented normal to the oxide surface. The probe-tip radius of curvature is $0.01 \mu\text{m}$, and the probe-cone apex half-angle is 10° . The oxide thickness is $0.01 \mu\text{m}$. The radial axis cutoff distance is $0.1 \mu\text{m}$. Dirichlet boundary conditions set the potential for the biased probe and the grounded semiconductor. Neumann boundary conditions are used on the remaining outer boundaries. Calculations are done for two dopant density (N_d) values, one low ($1 \times 10^5 \mu\text{m}^{-3}$) and one high ($1 \times 10^8 \mu\text{m}^{-3}$).

FEM: Conical probe (model A)

To provide a frame of reference for comparing models, the first slow model or Model A is the same numerical approach that is used in much of this paper. The linear finite-element software package PLTMG is used to model the conical probe and to solve Poisson's equation in the semiconductor region and Laplace's equation in the oxide and ambient regions.

FEM: Spherical probe (model B)

To compare effects due to probe shape, the second slow model or Model B replaces the conical probe of Model A

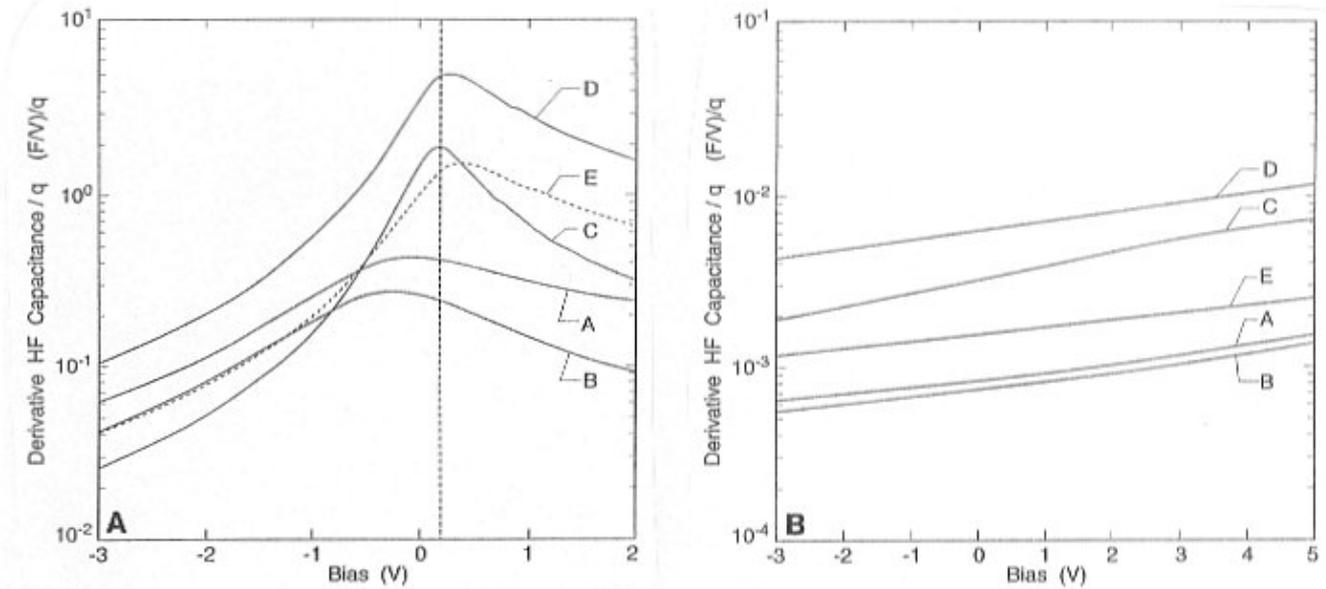


Figure 18. Derivative of the high-frequency capacitance for the five models. The dopant density (N_a) is (A) $1 \times 10^5 \mu\text{m}^{-3}$ and (B) $1 \times 10^8 \mu\text{m}^{-3}$. The vertical dotted line at 0.2 V bias in (A) is the approximate maximum bias for which the calculations are valid.

with a spherical probe. The probe-tip radius of curvature is the same for both models, and PLTMG is used to solve the problem numerically.

One way to speed the use of slow FEM-based models is to note that the capacitance curves are smooth functions of the model parameters. Thus, it is possible to form a database of calculations whereby linear interpolation may be used to extract dopant densities. But this requires further work.

1D MOSC: Conical probe in air on oxide (model C)

Contrasting the two slow FEM-based models are three fast 1D MOSC-based models. One fast model [23] or Model C assumes that the probe-air-oxide-semiconductor capacitance (per unit area) can be approximated to lowest order by first partitioning the conical probe into a set of concentric rings and then assuming that the capacitance (per unit area) between a ring and the semiconductor may be found by using the 1D MOSC model. With the air-gap or probe-to-oxide distance known, the net air-oxide capacitance (per unit area) can be found. {The oxide capacitance (per unit area) is equal to $\epsilon_0 \epsilon_{ox} / t_{ox}$, t_{ox} refers to the oxide thickness. The capacitances (per unit area) of the air and oxide, being in series, can be added as in a parallel circuit to form a net oxide capacitance (per unit area) that is needed by the 1D MOSC model}. With this and the 1D MOSC model, the radial distribution of the net air-oxide-semiconductor capacitance (per unit area) can be found. This radial distribution is similar to having capacitances in parallel, and the net capacitance is found by summing the

capacitances in series, i.e., by integration in the radial direction.

1D MOSC: Spherical probe in oxide (model D)

Another fast model [18, 20, 21] or Model D is more robust in finding the net oxide capacitance (per unit area) that is needed by the 1D MOSC model. The probe is modeled to lowest order by a sphere embedded in oxide. The net oxide capacitance (per unit area) is determined by finding the capacitance (per unit area) of a sphere-oxide-metal system where the semiconductor is treated as being metallic. The probe is biased at 1 V, the metal is grounded at 0 V, and the method of images is used to solve the electrostatics problem. The normal component of the electric field along the oxide-metal boundary gives the surface charge density distribution, which in turn gives the capacitance (per unit area) of the system. This capacitance (per unit area) is then equated to the net oxide capacitance (per unit area). With this and the 1D MOSC model, the radial distribution of the net probe-oxide-semiconductor capacitance (per unit area) is determined. The net capacitance is then found by integration as in Model C.

Spherical probe in air on oxide (model E)

An interesting improvement to Model D is Model E. The probe is again spherical, but now it is surrounded by air and is set on the uniformly thick oxide layer. The net oxide capacitance (per unit area) is determined as before by finding the capacitance of a model electro-statics problem; i.e., the semiconductor is treated as being metallic and is grounded at

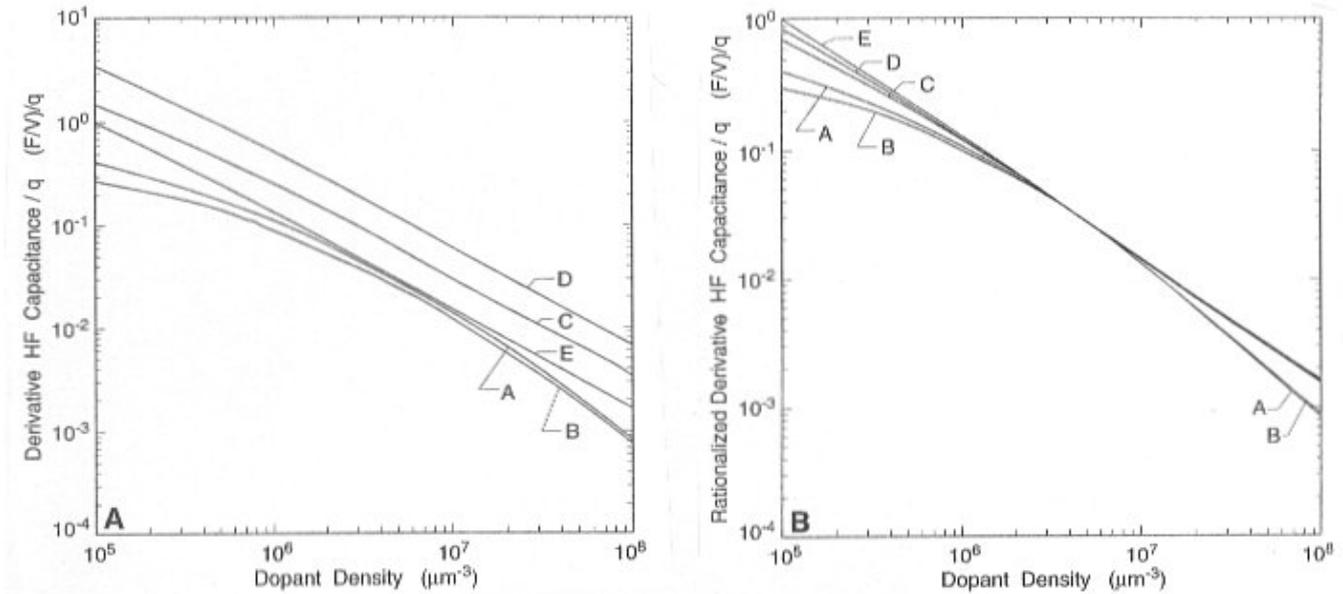


Figure 19. Conversion curves relating dopant density with derivative high-frequency capacitance data at zero bias for the five models as (A) calculated and (B) scaled to the value of model A at dopant density value $7 \times 10^6 \mu\text{m}^{-3}$.

0 V, and the sphere is biased at 1 V. While a method of images solution would be needed here to qualify it as a fast method, we used PLTMG to find the normal component of the electric field along the oxide-metal boundary. This gives the surface charge density distribution, which in turn gives the radial distribution of the net oxide capacitance (per unit area). The 1D MOSC is used as before in Model D to find the net capacitance.

Results and comparisons of models

Figure 18 shows the derivative of the high-frequency capacitance of the five models, where the dopant densities N_a are $1 \times 10^5 \mu\text{m}^{-3}$ (Fig. 18A) and $1 \times 10^8 \mu\text{m}^{-3}$ (Fig. 18B). The vertical dotted line at bias 0.2 V in Figure 18A is the approximate maximum bias for which the calculations are valid. At larger bias, the depletion region expands beyond the set boundaries of the domain region, and the net charge that is outside the set boundaries is not included in the calculation of the capacitance. (The domain size is set, in part, by the radial cutoff distance, and this is determined subjectively as explained in the section entitled **Geometry** under the main heading **Uniform Doping**. Here, the radial cutoff distances are $0.1 \mu\text{m}$ and $0.05 \mu\text{m}$ for dopant densities $1 \times 10^5 \mu\text{m}^{-3}$ and $1 \times 10^8 \mu\text{m}^{-3}$, respectively).

The degree of agreement (comparative scaling) of the results of the fast 1D MOSC-based models and the slow FEM-based models may be expected to depend, in part, on the degree to which the assumptions of the 1D MOSC model are

satisfied. The best agreement may be expected at high dopant densities where the screening length is small compared to the probe-tip radius of curvature and the oxide thickness. The worst agreement may be expected at low dopant densities where the screening length is large compared to the probe-tip radius of curvature and the oxide thickness.

Figure 19 shows the calibration or conversion curves relating dopant density with derivative high-frequency capacitance data at zero bias (flat-band) for the five models as calculated (Fig. 19A) and scaled (Fig. 19B) to the value of model A at the dopant density value $7 \times 10^6 \mu\text{m}^{-3}$. The difference between curves A and B shows that probe-shape effects become more pronounced at lower dopant densities, as may be expected. The fast 1D MOSC-based model curves C, D, and E deviate significantly from the slow FEM-based model curves A and B, i.e., curves A and B show more curvature than curves C, D, and E.

While the fast 1D-MOSC-based models can provide qualitative information in the matter of seconds compared to hours for a slow FEM calculation, these figures show the importance of relating the 1D-MOSC-based results to the more accurate FEM based results. This observation is the basis for proposing the use of fast models that interpolate over a database of calculations of slow FEM-based calculations in order to have both speed and accuracy.

Acknowledgments

It is a pleasure to acknowledge those at NIST who helped us to accomplish this work. They include: Dr. David G. Seiler for his concern regarding semiconductor characterization modeling that led to this project, Dr. Ronald F. Boisvert for his assistance with the software libraries [10, 12, 13], Dr. William F. Mitchell for his assistance with the PLTMG software package and the many helpful discussions, Drs. Hai C. Tang, Jeffrey T. Fong and Mrs. P. Darcy Barnett for their assistance with the ANSYS software package, Mr. N. Alan Heckert for his assistance with the NCAR graphics package, Dr. James S. Sims for his assistance with providing machine time for calculations, Mr. Denis A. Lehane for computer operating system support, and Mr. Raymond J. Mele for his graphics art work.

This work was supported in part by the National Semiconductor Metrology Program at the National Institute of Standards and Technology (NIST); not subject to copyright. Certain software are identified in this paper to adequately specify the modeling procedures. Such identification does not imply recommendation or endorsement by the NIST, nor does it imply that the methods identified are necessarily the best available for the purpose.

References

- [1] ANSYS (1994) ANSYS user's manual for revision 5.1. Vol. 1: Procedures. ANSYS, Inc., 201 Johnson Road, Houston, PA. Chapter 4, 9-15.
- [2] Ascher U, Christiansen J, Russell RD (1981) Collocation software for boundary-value ODEs. *ACM Trans Math Softw* **7**: 209-222.
- [3] Ascher U, Christiansen J, Russell RD (1981) Algorithm 569, COLSYS: Collocation software for boundary-value ODEs. *ACM Trans Math Softw* **7**: 223-229.
- [4] Bader G, Ascher U (1987) A new basis implementation for a mixed order boundary value ODE solver. *SIAM J Sci Stat Comput* **8**: 483-500.
- [5] Bank RE (1994) PLTMG: A software package for solving elliptic partial differential equations, users' guide 7.0. Society for Industrial and Applied Mathematics. Philadelphia, PA. Software is available at URL: <http://www.netlib.org/>.
- [6] Barber HD (1967) Effective mass and intrinsic concentration in silicon. *Solid-State Electron* **10**: 1039-1051.
- [7] Barrett RC, Quate CF (1991) Charge storage in a nitride-oxide-silicon medium by scanning capacitance microscopy. *J Appl Phys* **70**: 2725-2733.
- [8] Berntsen J, Espelid TO, Genz A (1991) An adaptive algorithm for the approximate calculation of multiple integrals. *ACM Trans Math Softw* **17**: 437-451.
- [9] Berntsen J, Espelid TO, Genz A (1991) Algorithm 698, DCUHRE: An adaptive multidimensional integration routine for a vector of integrals. *ACM Trans Math Softw* **17**: 452-456.
- [10] Boisvert RF, Howe SE, Kahaner DK, Springmann JL (1990) Guide to Available Mathematical Software, NISTIR 90-4237. Center for Computing and Applied Mathematics, National Institute of Standards and Technology (NIST), Gaithersburg, MD (U.S. Department of Commerce). Available from: Mathematical and Computational Sciences Division within the Information Technology Laboratory of the NIST. Also available at URL: <http://math.nist.gov/gams/>.
- [11] Diebold AC, Kump MR, Kopanski JJ, Seiler DG (1996) Characterization of two-dimensional dopant profiles: Status and review. *J Vac Sci Technol B* **14**: 196-201.
- [12] Dongarra JJ, Grosse E (1987) Distribution of mathematical software via electronic mail. *Comm ACM* **30**: 403-407.
- [13] Dongarra J, Rowan T, Wade R (1995) Software distribution using Xnetlib. *ACM Trans Math Softw* **21**: 79-88.
- [14] Dreyer M, Wiesendanger R (1995) Scanning capacitance microscopy and spectroscopy applied to local charge modifications and characterizations of nitride-oxide-silicon heterostructures. *Appl Phys A* **61**: 357-362.
- [15] Gaitan M, Mayergoyz ID (1989) A numerical analysis for the small-signal response of the MOS capacitor. *Solid-State Electron* **32**: 207-213.
- [16] Grove AS, Deal BE, Snow EH, Sah CT (1965) Investigation of thermally oxidised silicon surfaces using metal-oxide-semiconductor structures. *Solid-State Electron* **8**: 145-163.
- [17] Halen PV, Pulfrey DL (1985) Accurate, short series approximations to Fermi-Dirac integrals of order $1/2$, $1/2$, 1 , $3/2$, $5/2$, 3 , and $7/2$. *J Appl Phys* **57**: 5271-5274. Erratum. *J Appl Phys* **59**: 2264-2265.
- [18] Huang Y (1995) Quantitative two-dimensional dopant profile measurement on semiconductors by scanning probe microscopy. Doctoral Dissertation. Department of Physics, University of Utah, Salt Lake City, UT. pp. 42-60.
- [19] Huang Y, Williams CC (1994) Capacitance-voltage measurement and modeling on a nanometer scale by scanning C V microscopy. *J Vac Sci Technol B* **12**: 369-372.
- [20] Huang Y, Williams CC, Slinkman J (1995) Quantitative two-dimensional dopant profile measurement and inverse modeling by scanning capacitance microscopy. *Appl Phys Lett* **66**: 344-346.
- [21] Huang Y, Williams CC, Smith H (1996) Direct comparison of cross-sectional scanning capacitance microscope dopant profile and vertical secondary ion-mass spectroscopy profile. *J Vac Sci Technol B* **14**: 433-436.
- [22] Johnson WC, Panousis PT (1971) The influence of Debye length on the C-V measurement of doping profiles. *IEEE Trans Electron Devices* **18**: 965-973.
- [23] Kopanski JJ, Marchiando JF, Lowney JR (1997) Scanning capacitance microscopy applied to two-dimensional

dopant profiling of semiconductors. *Mater Sci Eng B* **44**:46-51.

[24] Kopanski JJ, Marchiando JF, Lowney JR (1996) Scanning capacitance microscopy measurements and modeling for dopant profiling in silicon. In: *Semiconductor Characterization: Present Status and Future Needs*. Bullis WM, Seiler DG, Diebold AC (eds.). AIP Press, Woodbury, NY. pp. 308-312.

[25] Kopanski JJ, Marchiando JF, Lowney JR (1996) Scanning capacitance microscopy measurements and modeling: Progress towards dopant profiling of silicon. *J Vac Sci Technol B* **14**: 242-247.

[26] Marchiando JF (1996) Application of the collocation method in three dimensions to a model semiconductor problem. *Int J Num Meth Eng* **39**: 1029-1040. {The long version appears in (1995) On using collocation in three dimensions and solving a model semiconductor problem. *J Res Natl Inst Stand Technol* **100**: 661-675.}

[27] McFaddin HS, Rice JR (1992) Collaborating PDE solvers. *Applied Numerical Math* **10**: 279-295.

[28] Morgan SP, Smits FM (1960) Potential distribution and capacitance of a graded p-n junction. *Bell Sys Tech J* **39**: 1573-1602.

[29] Neubauer G, Erickson A, Williams CC, Kopanski JJ, Rodgers M, Adderton D (1996) 2D-Scanning capacitance microscopy measurements of cross-sectioned very large scale integration test structures. In: *Semiconductor Characterization: Present Status and Future Needs*. Bullis WM, Seiler DG, Diebold AC (eds.). AIP Press, Woodbury, NY. pp. 318-321.

[30] Neubauer G, Erickson A, Williams CC, Kopanski JJ, Rodgers M, Adderton D (1996) Two-dimensional scanning capacitance microscopy measurements of cross-sectioned very large scale integration test structures. *J Vac Sci Technol B* **14**: 426-432.

[31] Nicollian EH, Brews JR (1982) *MOS (Metal Oxide Semiconductor) Physics and Technology*. John Wiley & Sons, New York. pp. 371-422.

[32] Renka RJ (1988) Multivariate interpolation of large sets of scattered data. *ACM Trans Math Softw* **14**: 139-148.

[33] Renka RJ (1988) Algorithm 660, QSHEP2D: Quadratic Shepard method for bivariate interpolation of scattered data. *ACM Trans Math Softw* **14**: 149-150.

[34] Renka RJ (1988) Algorithm 661, QSHEP3D: Quadratic Shepard method for trivariate interpolation of scattered data. *ACM Trans Math Softw* **14**: 151-152.

[35] Scheid F (1968) *Schaum's Outline of Theory and Problems of Numerical Analysis*. McGraw-Hill, New York. pp. 65-67.

[36] Schroder DK (1990) *Advanced MOS Devices*. Addison-Wesley Publishing Co., Reading, MA. pp. 91-176.

[37] Schulz M (1986) Space-charge layers at semi-conductor interfaces. In: *Crystalline Semiconducting*

Materials and Devices. Butcher PN, March NH, Tosi MP (eds.). Plenum Press, New York. pp. 425-481.

[38] Semiconductor Industry Association (1994) *The National Technology Roadmap for Semiconductors*. Available from: Semiconductor Industry Association, 4300 Stevens Creek Boulevard, Suite 271, San Jose, CA. pp. 19-21, 56.

[39] Seiwatz R, Green M (1958) Space charge calculations for semiconductors. *J Appl Phys* **29**: 1034-1040.

[40] Tomiye H, Kawami H, Izawa M, Yoshimura M, Yao T (1995) Scanning capacitance microscope/ atomic force microscope/scanning tunneling microscope study of ion-implanted silicon surfaces. *Jpn J Appl Phys, Part 1*, **34**: 3376-3379.

[41] Williams CC, Hough WP, Rishton SA (1989) Scanning capacitance microscopy on a 25 nm scale. *Appl Phys Lett* **55**: 203-205.

[42] Williams CC, Slinkman J, Hough WP, Wickrama-singhe HK (1989) Lateral dopant profiling with 200 nm resolution by scanning capacitance microscopy. *Appl Phys Lett* **55**: 1662-1664.

Discussion with Reviewers

D.J. Thomson: Could the authors comment about the "resolution" of SCM given their calculations. For example, if the objective was to measure doping concentrations to 10% accuracy, what do the authors' calculations predict about the limits of SCM for the particular cases they have simulated?

Authors: Determining the "resolution" is rather complicated at this time. Uncertainties in the measurement and the model parameters must be accounted for in order to ascribe an uncertainty to a value of doping. This has yet to be done, but some sense of the "resolution" (within the constraints of the model) may be inferred from the conversion curves presented in Figure 10. The curves are far from horizontal suggesting that the resolution could be within the 10% criterion for very well controlled experimental situations.

D.J. Thomson: Could the authors also comment on the applicability of their techniques to the interpretation of SCM data from "real" devices where there are multiple doped regions?

Authors: Interpreting SCM data from "real" devices is complicated by changes in things that are both included and not included in the model. There are doping gradients, changes in electrical type, charging, surface contamination, changes in oxide thickness, probe-tip, etc. The database was calculated for uniformly doped material under idealized conditions. Therefore, the database ought to apply to profiles that are near these idealized conditions (i.e., the doping gradient is not too steep), where the quasi-neutrality condition is satisfied near the measurement point. It remains to be determined how best to implement or augment the database for profiles where

quasi-neutrality is less satisfied. Using an iterative procedure with some self-consistency criterion has been suggested by Clayton Williams at the University of Utah, but we have not studied that yet. Interpreting SCM data from “real” devices is under development.

P. Koschinski: Can the authors explain why the macroscopic Poisson equation, which is a continuum equation, is applicable to nanoscopic problems? For example, at a doping level of $1.0 \times 10^4 \mu\text{m}^{-3}$ ($1.0 \times 10^{16} \text{cm}^{-3}$), only one doping atom per μm depth is present underneath the tip (doping concentration/tip area = one-dimensional concentration under the tip). For calculated depletion regions, which extend below one μm , less than one atom is forming the space charge region underneath the tip. Can you comment on this rough estimation?

Authors: When the doping of the semiconductor becomes sufficiently low, so that the screening length becomes sufficiently greater than the probe-tip radius of curvature and the oxide thickness, the spatial discreteness of the impurities will modulate the macroscopic charge density and be detectable by the probe. The uniformity assumption imposed on the charge density by the form of Poisson equation used here will become less valid and break down. Apart from Figure 2 that is included for demonstration purposes, the lowest dopant density for which the capacitance is calculated and shown in Figures 7A and 8 is not $1.0 \times 10^4 \mu\text{m}^{-3}$, but rather is $1.0 \times 10^5 \mu\text{m}^{-3}$, where the screening length is near 13 nm, which is a little larger than the 10 nm used here for the probe-tip radius of curvature and the oxide thickness. True, this is near the edge.

P. Koschinski: The extent of depletion or accumulation layers in semiconductors does not only depend on tip bias and doping level of the semiconductor but on localized surface and interface charges, caused by surface states, too. Since it is known that, e.g., silicon exhibits various surface states with different charges, can the authors explain why they believe that their calibration curves calibrated for charge free surfaces/interfaces are valid for real materials? The same statement is also true for charged deep traps in the bulk semiconductor. Can the authors comment on this problem?

Authors: The usefulness of the SiO_2 -Si interface in devices stems from the fact that the interface state density can be made very low by proper processing. The model calculations here are a reasonable first step toward interpreting SCM measurements under idealized conditions, i.e., using a zero interface-state charge density as a zeroth order approximation. We agree that there is room for improvement. The calculated results need to be compared to measurements of real materials, and this is the next step. While real materials involve physics beyond that modeled here, it is important that the measurements be made in a regime where the fewest complicating physical mechanisms dominate, so that model

parameter extraction is both quick and meaningful. We intend to include a measure of charge trapping at some level of approximation and see the effects. As long as the surface charge is constant under the probe, its effects can be compensated with the alternating current (AC) bias voltage.

P. Koschinski: All calculations are performed by solving Poisson’s equation for a time independent non-equilibrium case, i.e., a biased tip located above a semiconductor surface. Can the authors explain why it is justified to neglect non-equilibrium phenomena like recombination or generation processes, by simply solving Poisson’s equation, since any non-equilibrium state of a semiconductor is accompanied by these phenomena? For example, in the case of accumulation one would expect an enhanced recombination of charge carriers in the accumulation layer influencing the carrier concentrations.

Authors: The technique takes advantage of the fact that the minority carrier recombination/generation times are much larger than the majority carrier dielectric relaxation time. The low-frequency component of the applied bias modulation must be sufficiently fast compared to that of minority carrier generation, so that an inversion layer does not form. As the low-frequency component cycles through accumulation, the number density of the minority carriers are duly reduced by the enhanced number of majority carriers. As the low-frequency component cycles through depletion, the minority carriers are unable to respond in time, and due to their small number density compared to that of the majority carriers, the minority carriers are essentially frozen out and ignorable. The high-frequency component of the bias modulation involves times much longer than the majority carrier response time, so that the majority carriers can be modeled to lowest order of approximation as responding nearly instantaneously to the applied bias field.

P. Koschinski: In real capacitance measurements an AC component of the applied bias with a specific frequency is necessary. Since many phenomena in semiconductors are depending on the frequency used for the measurement, like charging of surface or bulk states, how can the calibration curves obtained by time independent simulations be correlated with measurements with time dependent signals?

Authors: Again, to the lowest order of approximation, the effects due to surface states are assumed to be negligible, and the objective is to use frequencies at which all time dependent factors other than those with which we are concerned are frozen out and ignorable.

P. Koschinski: The authors explained that the simulation mesh was once generated with a perturbation approach for the special case of maximum depletion and then used for all other simulations. What is the criterion that the authors believe

that the resulting mesh is most appropriate for further simulations? Why did the authors not use the established criterion for mesh generation of equidistributing of the local discretization error of the whole simulation area?

Authors: The approximate error in the solution in PLTMG is an a posteriori local error estimate based on the jump discontinuity of the normal direction of the vector function of the product of the solution gradient and the dielectric constant. The values of the dielectric constants of the three spatial regions were found to significantly weight the refinement procedure ($\epsilon_{\text{Si}} = 11.9$, $\epsilon_{\text{oxide}} = 3.9$, and $\epsilon_{\text{air}} = 1$). Equidistributing the local discretization error of the whole simulation area caused the bulk of the refinement to be in the air around the probe, less in the oxide, and even less in the semiconductor. The mesh in the semiconductor region had comparatively large triangles over areas where the net charge density had significant variation. This led to a coarse evaluation of the volume integral of the net charge density to find the displaced charge, and the capacitance. Remeshing at different values of bias would introduce additional error into the calculation. See Figure 9; compare lines A, C, D.

D.P. Kilcrease and D.C. Cartwright: Can you quantify the error of taking the interfacial charge density to be zero at the insulator-semiconductor boundary?

Authors: Not at this time. This is something that needs to be done. A uniform distribution of trapped charge shifts the capacitance curves along the direct current (DC) bias voltage axis [7, 14, 16]. Some of this problem is removed by the measurement procedure that is used to determine the bias voltage near flat-band in the low doped region of a doping profile. A nonuniform distribution further complicates things by allowing a spatial variation as well, and of course, this will modulate the estimates of the doping profile. Unless variations in the interface charge density can be ignored, the usefulness of the SCM technique is questionable.

H. Edwards: This work uses a semiclassical model for the carriers. Would a quantum treatment change the results significantly?

Authors: A quantum treatment would be much more complicated than that done here. One thing that it would do is make the carrier charge density at the SiO_2/Si boundary go to zero. This becomes significant when inversion or accumulation occurs. The measurements are to be done near flat-band and where inversion is not allowed to occur. The difference in accumulation should be small, since the field is then mostly across the oxide. Further, our spatial sizes are still rather large (≥ 10 nm), so that the effects due to a quantum treatment ought to be relatively small. However, it is something that needs to be considered in the future.

H. Edwards: One big question in SCM is how to set the DC

sample bias so that the true position of a pn junction may be measured in a cross-sectional experiment. Based on their model, can the authors suggest the proper choice of DC bias for such a measurement, as well as how to verify that the correct bias is being applied?

Authors: Generally, the DC bias should be chosen so that the surface is held near flat-band. Otherwise, the probe bias will greatly alter the charge balance under the probe. The bias can be determined from the peak in the $\Delta C_{\text{HF}}/\Delta V$ in a relatively lightly doped region. We have not yet modeled a pn junction with this method. The calculations have been for uniform doping. To interpret measurements by using a conversion curve and a data-base, the doping profile needs to be slowly varying, so that it varies relatively little over the region perturbed by the probe. The phase of the SCM signal changes when crossing a pn junction, so this may be used to estimate the junction boundary.

H. Edwards: Spatial resolution is so far the weakest point of SCM as applied to shallow-junction profiling. Does this work illuminate whether true nm-scale SCM imaging will ever be possible?

Authors: We have not yet analyzed a profile with this method. Spatial resolution depends, in part, on the probe-tip radius of curvature and the signal-to-noise ratio of the instrument. There is a balance; smaller tips give less signal. This is related to the preceding question about the bias. If the bias voltages can keep the size of the depletion region near that of the probe-tip size, there is hope.

H. Edwards: The present work seems to target an accuracy of a few percent in the numerical solutions. However, variations in oxide thickness, dielectric constant, surface charge, interface-state density, trapped charge, and surface contamination could change the signal intensity by orders of magnitude. These variations also could effectively change the voltage scale and shift the DC offset by several volts. How do the authors plan to account for these important but uncontrolled factors? Is the model robust enough to extract these parameters from real data?

Authors: All these factors and effects are important, and they complicate things. There is probably insufficient functional dependence in the capacitance measurements by themselves to independently resolve all the model parameters. For the measurements to be meaningfully interpreted, the measurements must be done in a manner where the fewest factors dominate and are calculable. Consequently, some control must be exercised over the sample preparation and the measurement process. In general, the inverse problem does not have a unique solution, and other measurements need to be made on the samples to determine some of the input parameters. All these effects need to be considered, but they require work beyond that presented here.

C.C. Williams: In the sections entitled **Two dimensions, 2D probe** and **Two dimensions, 3D probe**, it is noted that the depletion region expands into the highly doped side. I believe that this is due to the fact that the figure represents the net charge distribution and is not normalized against the background concentration. It would be more instructive to define a condition for depletion related to a percent modulation of the local carrier density. It would be interesting to see whether the depletion really occurs toward the high side.

Authors: The depletion expansion toward the highly doped side is due to the dopant density dependent work-function difference. The finite-sized probe-tip is an equipotential surface that spans across the doping gradient, and its presence is felt across the gradient region. The higher doped side sees a relatively larger bias (depletion); the lower doped side sees a relatively lower bias (accumulation).

Normalizing against the background concentration is comparable to scaling that of the low doped region with a large value and to scaling that of the high doped region with a low value, and this, of course, will shift the distribution to the lower doped region. A plot of ρ/N_a amplifies the accumulation in the junction. A plot of $d\rho/dV$ is similar to that of ρ , except that: (1) the variation associated with the static built-in field across the junction is suppressed; and (2) a small variation is introduced at the surface near the probe that seems unphysical and may possibly be due to the error of the calculation. Plots of $(1/N_a)(d\rho/dV)$ and $(1/p)(d\rho/dV)$ show distributions that are deeper and broader in the lower doped region than in the higher doped region. Rationalizing with N_a magnifies small variations in ρ at the surface near the probe-tip in a way that seems unphysical. Rationalizing with p gives a smooth distribution at 1 V bias, but this has limits, because at large bias part of the lower doped region becomes fully depleted and p becomes negligible. Then ρ/p becomes singular. So yes, it is possible to form a distribution with ρ that is shifted more towards the lower doped region than the higher doped region, as motivated by that expected when considering uniformly doped regions separately. But then too, the capacitance is a measure of charge, a volume integral of the displaced charge density, that is a quantity that is not rationalized, and here, the main contribution comes from the higher doped region.