PREDICTIVE SIMULATION OF VOID FORMATION DURING THE DEPOSITION OF SILICON NITRIDE AND SILICON DIOXIDE FILMS

C. Heitzinger[1,3], A. Sheikholeslami[1], H. Puchner[2], and S. Selberherr[1]

[1] Institute for Microelectronics, TU Vienna,
Gußhausstraße 27–29/E360, A–1040 Vienna, Austria
[2] Cypress Semiconductor,
3901 N. First Street, San Jose, CA 95134, USA
[3] Corresponding author, e-mail address: Heitzinger@iue.tuwien.ac.at

The simulation of deposition processes enables to predict the shape of deposited films and voids for different trench shapes and process conditions. The position and shape of the voids determine the electrical characteristics of the final devices, e.g., memory cells. For these topography simulations a simulator was developed which uses an enhanced level set algorithm for describing the moving boundaries and handles a variety of deposition and etching processes. In experiments silicon nitride and silicon dioxide films were deposited into trenches of test structures of varying shapes. The processes simulated are backend processes for memory cells, where ILD (interlevel dielectric) materials are deposited and voids are formed in all layers. SEM images of the test structures were compared to simulation results. Very good agreement between measurements and simulations was found.

INTRODUCTION

Deposition in and etching of silicon trenches are crucial processes in today's semiconductor manufacturing, e.g., for state of the art memory cells and power MOSFETs. Understanding and simulating the transport of gas species and surface evolution enable to predict the resulting profiles and voids, to achieve desired shapes, and thus to optimize process parameters with respect to manufacturing throughput and the electrical characteristics of the devices. In backend processes for memory cells ILD (interlevel dielectric) materials and processes result in void formation during gap fill. This approach lowers the overall k-value of a given metal layer and is economically advantageous. The impact of the voids on the total capacitive load is tremendous.

In order to provide predictive simulations of the overall capacitance, the shapes and positions of the voids must be simulated accurately. These topography simulations serve

as the input to the following extraction of the capacitance. Based on these extracted capacitances the properties of the whole device can be simulated and hence the mask shapes and process parameters adjusted to yield the desired device characteristics.

In the following sections the simulation method is described. A general purpose topography simulator was developed based on the level set method. It consists of a level set module, a radiosity module, a diffusion module, and a surface reaction module. The level set method is an interesting alternative to conventional methods like the cell based method [2, 9] . One of the main advantages of the cell based method is its robustness [5]. It can handle critical structures such as high aspect ratio trenches well [3, 6]. On the other hand, the level set method provides higher spatial resolution on grids of the same size and calculating surface normals – crucial for radiosity simulations – is more precise than when using a cellular format, but the level set method involves more complicated algorithms.

An advanced level set algorithm which combines dynamic narrow banding and extending the speed function is presented. These seemingly unconnected concepts were combined yielding a fast and accurate level set algorithm. Furthermore an algorithm which coarsens the surface reduces the computational demands of the simulations significantly. This allows to achieve the high resolutions which are indispensable for accurate simulation of the surface evolution at the trench opening, the trench bottom, and during void formation.

After a description of the modeling approach to particle transport and surface reactions, the test structures are discussed and simulation results are presented. The simulations reproduced the shapes of the trenches very well and good quantitative agreement was achieved as well. The effects of the surface coarsening algorithm on the accuracy of the simulations and their computational effort were found to be very satisfactory.

THE ELSA SIMULATOR

We divide the feature scale simulation into three main steps. (a) The transport of particles in the boundary layer above the wafer must be simulated. This can happen in the radiosity or diffusion regime. (b) The fluxes of particles at the wafer surface found in the previous step determine the chemical reactions that take place at the wafer surface. (c) The surface of the wafer changes according to the previous steps. This is tracked using the level set method. The corresponding simulation flow of a typical topography simulation is shown in Figure 1.

The ELSA (Enhanced Level Set Applications) simulator is a general tool for topography simulation in the realm of semiconductor manufacturing processes. It consists of four modules, namely a level set module, a surface reaction module, and modules for particle transport by radiosity or diffusion. It can be used for simulating all common deposition and etching processes. Its main features are an efficient and precise level set algorithm including narrow banding and extending the speed function and a surface coarsening algorithm which reduces the computational effort significantly while ensuring high resolution in critical areas. The initial surface is given by an arbitrary number of points and a coarsening algorithm can be applied recursively.

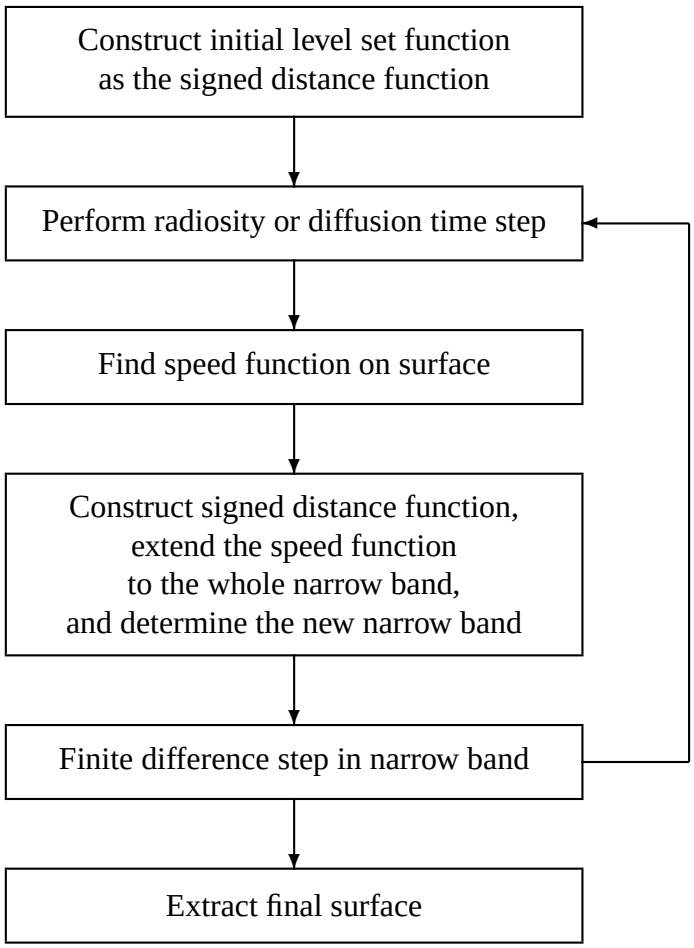


Figure 1: Overview of the simulation flow combining transport by physical models and surface evolution using the level set method. The simulation stops when a prescribed time is reached or when a layer of prescribed thickness has been deposited.

AN ENHANCED LEVEL SET ALGORITHM

The idea leading to fast level set algorithms stems from observing that only the values of the level set function near its zero level set are essential, and thus only the values at the grid points in a narrow band around the zero level set have to be calculated. As the zero level set moves, the signed distance function in the narrow band must be maintained. The usual approach to determining the current active narrow band points is to reinitialize whenever the zero level set reaches the boundary of the narrow band. Reinitialization, however, introduces errors and hence must be avoided. The level set algorithm used in our simulator traces the active narrow band points dynamically in each time step at no additional computational cost.

Narrow banding and extending the speed function are performed by one algorithm, which provides several benefits and is described below. First, the speed function is retained as the signed distance function throughout the simulation, which assures good accuracy till the end of the simulation. Second, narrow banding reduces the number of active points that have to be updated from $O(n^2)$ to $O(n)$. By retaining the signed distance function the

width of the narrow band can be kept down to two points on each side without decreasing accuracy. Third, time consuming calculations (cf. [1]) are reduced to a minimum by intertwining the computations necessary for narrow banding and extending the speed function. Finally the width of the narrow band can be adjusted if desired.

In detail the steps of this algorithm are as follows, where several auxiliary functions are used. Concerning the data structures, the information about the level set grid, the distance function, the extended speed function, and the tags for the fast marching method are stored in arrays. The trial points are stored in a heap.

1. First find the grid points whose speed function values are initially known. These values are computed in the physical simulation step and translated to the grid. Then compute the distance for the initial points, and tag them as known, and initialize the corresponding grid points of the speed function. Find the trial points which are the neighbors of the initially known points, and compute their tentative distance values. All other points are far points.

2. While there are trial points, do the following:

   (a) Remove the first point from the heap and call it a. It has the smallest distance from the zero level set of all points in the heap.

   (b) If narrow banding is used and the maximum width of the narrow band is smaller than the distance of a, return from the loop.

   (c) Mark a as known.

   (d) For all neighbors b of a, do the following:

      i. If b is a far point, recompute its distance and speed function values and mark it as a trial point.

      ii. If b is a trial point, recompute its tentative distance and speed function values, unless it was computed in the previous step.

3. If narrow banding shall be used, set the distance values of all points which are not marked as known to the width of the narrow band. Set the sign of the distance function.

4. Finally return two objects, namely the grid containing the new signed distance function and the grid containing the extended speed function.

A second order space convex finite difference scheme [9] is used for updating the level set grid. The main level set function must perform the bookkeeping for the narrow band considering the old narrow band from the previous iteration and the new narrow band. The points which are outside the old narrow band but inside the current narrow band are initialized to the signed distance function just computed.

SURFACE COARSENING

When using radiosity models for simulating the transport of particles above the wafer, two operations consume the most part of the computation time. The first operation is determining the visibility between all surface elements, which is an $O(n^2)$ operation, where n denotes the number of surface elements extracted from the level set grid. The second operation is solving a certain system of linear equations, which leads to calculating the inverse of a matrix with n^2 elements, which is an $O(n^3)$ operation.

Obviously increasing the number of surface elements is not a remedy in cases where high resolution is required. High resolution is needed, e.g., near the trench opening and the bottom of the trench. One approach is to devise a refinement and coarsening strategy for unstructured grids at the level of the level set implementation and the algorithms working on it. This, however, complicates the fast marching algorithm necessary for extending the speed function. In this work a different approach was taken by coarsening the surfaces after they have been extracted from the level set grid.

The algorithm works by walking down the list of surface elements extracted as the zero level set and calculating the angle α between two neighboring surface elements. Whenever $|\pi - \alpha|$ is below a certain threshold value of a few degrees, the neighboring elements are coalesced into one. After one sweep through the list, the algorithm can be reapplied for further coarsening. After k coarsening sweeps, at most 2^k surface elements are coalesced into one. The resulting longer surface elements are used for the radiosity calculation, after which the fluxes are translated back from the coarsened elements to the original ones.

TRANSPORT BY RADIOSITY AND LUMINESCENT REFLECTION

A formulation of the radiosity method for the case of luminescent reflection is described in the following [9]. It pertains to low energy particles and is used in the following simulations. The flux reaching the (straight) surface elements obtained in the surface extraction step may be written as a vector, i.e.,

$$\begin{aligned}\text{Flux} &= \beta_0 I_S + \beta \Psi L I_R \\ &= \frac{\beta - \beta_0}{1-\beta} I_S + \frac{\beta(1-\beta_0)}{1-\beta} \underbrace{L^{-1}(L^{-1} - (1-\beta)\Psi)^{-1}}_{T:=} I_S.\end{aligned}$$

Here I_S is the vector of fluxes coming from the sources to the surface elements, I_R is the vector of fluxes that arrive because of reflections, β_0 the sticking coefficient for particles coming directly from the source, β the one for secondary bounces, L the diagonal matrix containing the lengths of the surface elements, and

$$\Psi_{ij} = \frac{n_i \cdot (t_j - t_i) n_j \cdot (t_i - t_j)}{\pi |t_j - t_i|^3} [i \text{ visible } j],$$

where t_i are the centroids of the surface elements, n_i their unit normal vectors, and $[i \text{ visible } j]$ is 1 or 0 if the surface element j is visible from i or not. The second line in the equation above is obtained from the first line and the relationship $I_R = (1-\beta_0)I_S + (1-\beta)\Psi L I_R$ after some straightforward algebraic manipulations.

In the case of multiple, low energy species the calculation of the visibility matrix and the inverse T only depends on topographic information and thus does not have to be repeated for each species.

DEPOSITION MECHANISMS

When modeling topography processes it is in general possible to write down quite complicated reaction paths. This route is taken, e.g., in the investigations of surface reactions in [4, 11–13], where dozens of reactions are listed. In the conclusion of [13] it is remarked that it is not straightforward to determine the vital reactions and their constants. Therefore it is mandatory to reduce the possible reaction paths to an essential minimum which can reproduce the observed phenomena.

The silicon nitride films were deposited by PECVD from silane and NH_3 and they were not doped. The reaction thus is $SiH_4 + NH_3 \rightarrow SiNH + 3H_2$ [8]. For simulation purposes this was considered the essential reaction, where a detailed model of triaminosilane condensation can be found in [10].

The silicon dioxide films were deposited by pyrolytic decomposition of TEOS in a LPCVD process. In order to calculate the thickness Δd of the film deposited during a time interval of length Δt, we observe that Δd is proportional to Δt, to an Arrhenius term, and to the deposition rate R corresponding to the deposition model chosen. This implies $\Delta d = \Delta t \cdot k_e \mathrm{e}^{-E/kT} \cdot R$. Here $k_e \mathrm{e}^{-E/kT}$ is the Arrhenius term with activation energy E, absolute temperature T, and pre-exponential constant k_e, and R is the deposition rate [7].

SIMULATION RESULTS

Test structures of interconnect lines of memory cells were fabricated and several SEM images thereof were used to validate the corresponding simulations. For metal lines 1 and 2 the deposition of silicon dioxide films from TEOS was considered and for metal line 3 the deposition of silicon nitride was simulated. The detail in Figure 2 shows the metal layers 2 and 3.

The path of the species above the wafer surface is tracked in radiosity simulations where reflection happens in a luminescent manner. The computational effort of the level set algorithm with narrow banding as described above is negligible compared to the computation time of the physical models. This, however, is not the case when narrow banding is not employed. The simulation results with coarsening are nearly identical to those when no

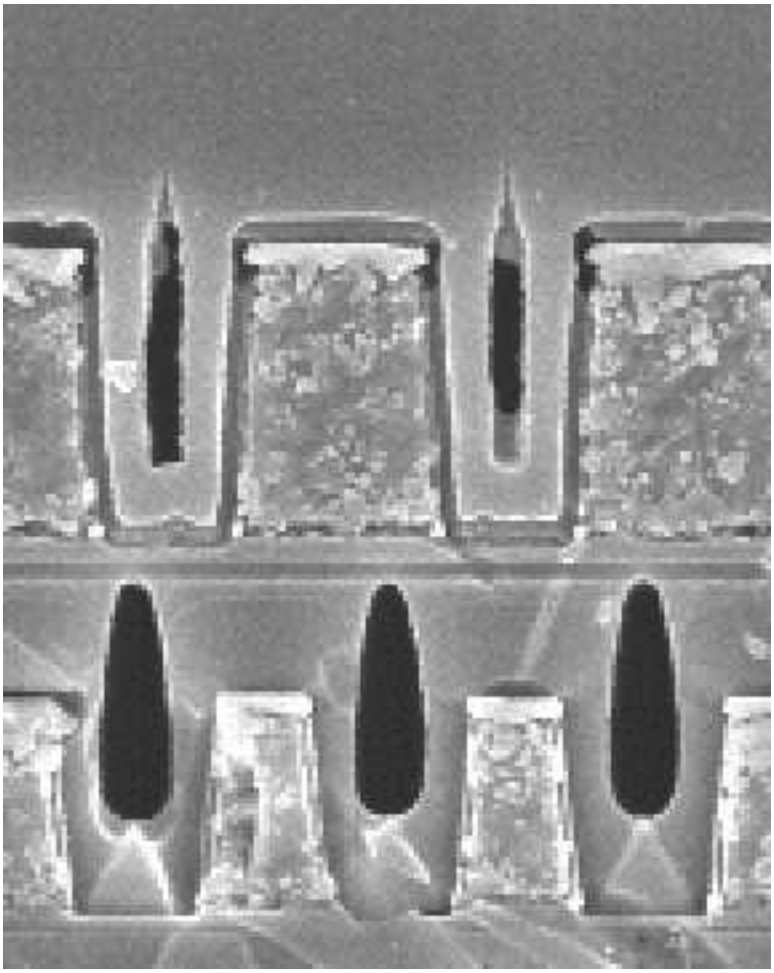


Figure 2: SEM image of a test structure with trenches of different widths. This detail shows metal layers 2 and 3 of the interconnect structure of a memory cell. In the upper row (M3 layer) the trenches are about $0.45\,\mu\mathrm{m}$ wide and a nitride film was deposited. In the lower row (M2 layer) silicon dioxide was deposited from TEOS.

coarsening was applied. Accuracy is hardly affected, but the simulation time significantly decreased.

An example of void formation after silicon nitride deposition is shown in Figure 3. Figure 4 shows the corresponding level set function. The shape and position of the void are reproduced correctly in the simulation. The test structures contain trenches of different widths and the influence of the width was reproduced in the simulations.

CONCLUSION

State of the art algorithms for surface evolution processes like deposition and etching processes have been devised and implemented. Based on these, a general simulator called ELSA was developed, which can be used for simulating all common deposition and etching processes.

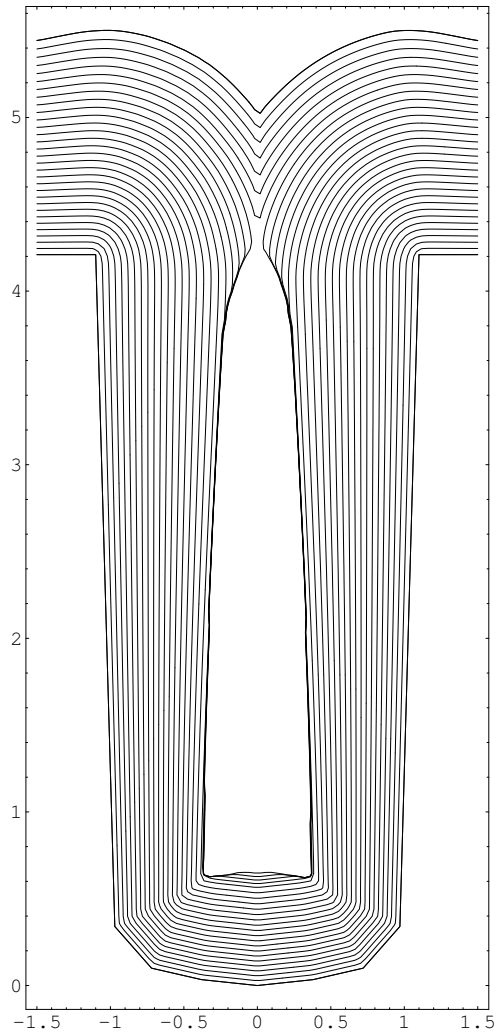
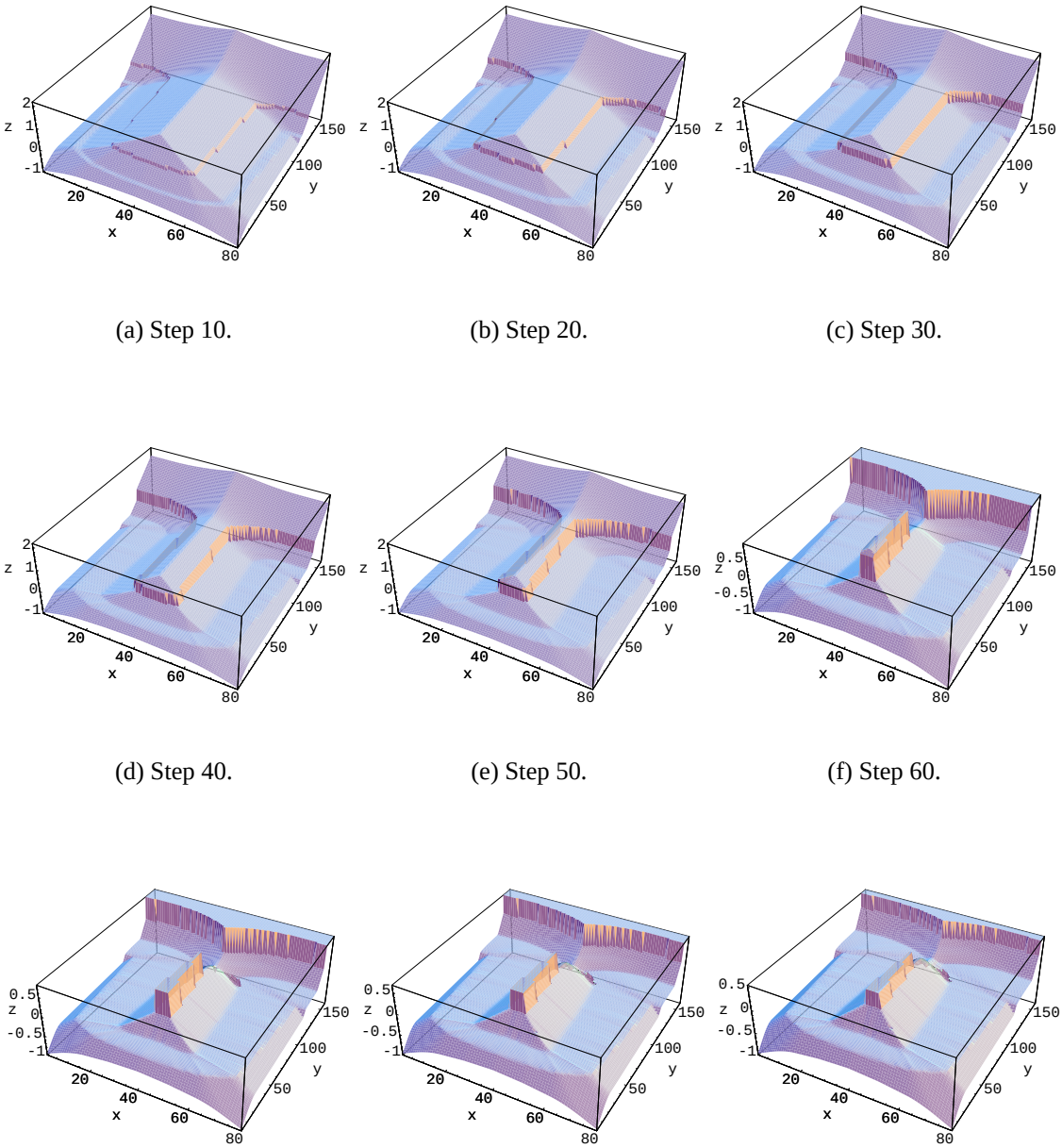


Figure 3: Simulation of void formation as shown in Figure 2. Here a level set grid of $80 \cdot 160$ points was used, 90 iterations were performed, and the coarsening algorithm was applied twice, substituting at most four neighboring surface elements by a new one.

The accuracy and speed of radiosity simulations was improved by two methods. The first method is an algorithm which performs three level set computations in parallel: calculating the signed distance function via a fast marching algorithm, extending the speed function, and dynamically moving the narrow band according to the new zero level set. This ensures a fast and accurate algorithm. The second method is a coarsening algorithm which guarantees fine resolution of the surface in parts of the boundary with relatively high curvature, i.e., where it is needed most. These parts are typically the opening of the trench and its bottom. At the same time the resolution is lowered where possible which reduces the demand on computational resources significantly. Typically radiosity simulations run five to ten times faster when this algorithm is employed.

The shape and position of voids in silicon dioxide and nitride films occurring in backend manufacturing processes were simulated accurately. The voids determine the capacitance of the interconnect lines, which is crucial for the performance of the final memory cell. All observed effects match in the SEM pictures and in the simulation results.



(a) Step 10. (b) Step 20. (c) Step 30.

(d) Step 40. (e) Step 50. (f) Step 60.

(g) Step 70. (h) Step 80. (i) Step 90.

Figure 4: The figures show the intermediate level set grids at steps 10, 20, 30, 40, 50, 60, 70, 80, and 90 during the simulation in Figure 3. In the narrow band the signed distance function is retained until the end of the simulation, whereas other points have not been touched after the initialization. The flat area remains after being touched by the moving narrow band.

ACKNOWLEDGMENTS

The first and last authors acknowledge support from the “Christian Doppler Forschungsgesellschaft,” Vienna, Austria.

REFERENCES

[1] D. Adalsteinsson and J.A. Sethian. The Fast Construction of Extension Velocities in Level Set Methods. *J. Comput. Phys.*, 148(1):2–22, 1999.

[2] C. Heitzinger, J. Fugger, O. Häberlen, and S. Selberherr. On Increasing the Accuracy of Simulations of Deposition and Etching Processes Using Radiosity and the Level Set Method. In G. Baccarani, E. Gnani, and M. Rudan, editors, *Proc. 32th European Solid-State Device Research Conference (ESSDERC 2002)*, pages 347–350, Florence, Italy, September 2002. University of Bologna.

[3] C. Heitzinger, W. Pyka, N. Tamaoki, T. Takase, T. Ohmine, and S. Selberherr. Simulation of Arsenic In-Situ Doping with Poly-Silicon CVD and its Application to High Aspect Ratio Trenches. *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, 2002. (In print).

[4] R.J. Hoekstra, M.J. Kushner, Ph. Schoenborn, and V. Sukharev. Microtrenching Resulting from Specular Reflection during Chlorine Etching of Silicon. *J. Vac. Sci. Technol. B*, 16(4):2102–2104, 1998.

[5] W. Pyka. *Feature Scale Modeling for Etching and Deposition Processes in Semiconductor Manufacturing*. Dissertation, Technische Universität Wien, 2000. http://www.iue.tuwien.ac.at/phd/pyka.

[6] W. Pyka, C. Heitzinger, N. Tamaoki, T. Takase, T. Ohmine, and S. Selberherr. Monitoring Arsenic In-Situ Doping with Advanced Models for Poly-Silicon CVD. In D. Tsoukalas and C. Tsamis, editors, *Proc. Simulation of Semiconductor Processes and Devices (SISPAD 2001)*, pages 124–127, Athens, Greece, September 2001. Springer, Wien, New York.

[7] G.B. Raupp, F.A. Shemansky, and T.S. Cale. Kinetics and Mechanism of Silicon Dioxide Deposition Through Thermal Pyrolysis of Tetraethoxysilane. *J. Vac. Sci. Technol. B*, 10(6):2422–2430, November 1992.

[8] G. Schumicki and P. Seegebrecht. *Prozeßtechnologie*. Springer, 1991.

[9] J.A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge, 1999.

[10] D.L. Smith, B. Wacker, S.E. Ready, C.C. Chen, and A.S. Alimonda. Mechanism of SiN_xH_y Deposition from NH_3–SiH_4 Plasma. *J. Electrochem. Soc.*, 137:614–623, February 1990.

[11] D. Zhang and M.J. Kushner. Mechanisms for CF_2 Radical Generation and Loss on Surfaces in Fluorocarbon Plasmas. *J. Vac. Sci. Technol. A*, 18(6):2661–2668, 2000.

[12] D. Zhang and M.J. Kushner. Surface Kinetics and Plasma Equipment Model for Si Etching by Fluorocarbon Plasmas. *J. Appl. Phys.*, 87(3):1060–1069, February 2000.

[13] D. Zhang and M.J. Kushner. Investigations of Surface Reactions during C_2F_6 Plasma Etching of SiO_2 with Equipment and Feature Scale Models. *J. Vac. Sci. Technol. A*, 19(2):524–538, 2001.