Modeling Oxide Regrowth During Selective Etching in Vertical 3D NAND Structures

T. Reiter[1]*, A. Toifl[2], A. Hössinger[2], and L. Filipovic[1]
[1]Christian Doppler Laboratory for Multi-Scale Process Modeling of Semiconductor Devices and Sensors, Institute for Microelectronics, TU Wien, 1040 Vienna, Austria
[2]Silvaco Europe Ltd., St Ives, Cambridgeshire, PE27 5JL, United Kingdom
*Email: reiter@iue.tuwien.ac.at

***Abstract*—One of the critical processing steps during the fabrication of modern 3D NAND flash memory structures is the selective etching of Si_3N_4 in the presence of SiO_2, which can cause the redeposition of byproducts on the SiO_2 layers. We present a physical process model for this phenomenon during etching and apply it to simulate oxide regrowth. The model describes the mass transfer of byproducts with a convection-diffusion equation which is solved on a cell-set volume representation of the etched solution. The simulated results, which combine Si_3N_4 etching, byproduct mass transfer, and the subsequent redeposition, show excellent agreement with experimental studies for the desired structures.**

***Index Terms*—Oxide regrowth, Selective etching, Process simulation, 3D NAND**

I. Introduction

Vertical three-dimensional (3D) NAND flash memory has established itself as one of the leading technologies for non-volatile storage devices. 3D NAND achieves higher storage capacities at lower power consumption and greater longevity by vertically stacking NAND cells, when compared to conventional 2D NAND [1], [2]. Increasing the number of stacked layers directly contributes to the increase in memory capacity, which is why having as many stacked layers as possible is desirable. However, this is often limited due to the complexities of the fabrication process, which requires a thorough physical understanding.

The initial structure of a 3D NAND stack consists of alternating silicon nitride (Si_3N_4) and silicon dioxide (SiO_2) thin films. After a slit is formed using a dry etching process, the patterned wafer is subject to a wet etching process to remove the sacrificial Si_3N_4 layer [3]. The wet etching process uses a hot aqueous solution of phosphoric acid (H_3PO_4) and must show high selectivity to Si_3N_4 over SiO_2 such that the SiO_2 layers are not damaged. For this reason, additives, such as silicic acid, are introduced to the H_3PO_4 solution [4]. However, recent experiments have shown an abnormal oxide regrowth behavior due to the addition of SiO_2 etching inhibitor and an increased Si_3N_4 etching byproduct generation [5]–[7] (c.f. Fig.1).

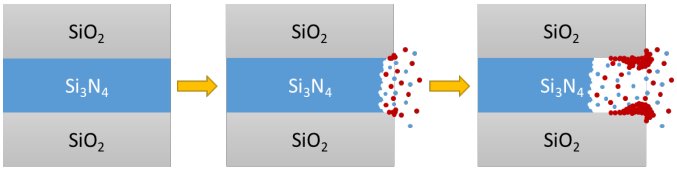

Fig. 1. Schematic diagram of oxide regrowth from generated byproducts (red circles) during Si_3N_4 (blue circles) etching.

In 3D NAND stacks with a large number of layers, the limited mass transfer of Si_3N_4 etching byproduct leads to a further increase in the oxide regrowth rate. SiO_2 layer thickening and etched Si_3N_4 trench clogging are defects which reduce the performance of NAND memories by preventing uniform metal deposition. Etching of the sacrificial nitride layer in a 3D NAND stack is part of the gate replacement process [8] and is followed by atomic layer deposition of tungsten to form the word line (WL). Increased scaling of the devices has been shown to lead to voids in the deposited tungsten mold, which can lead to serious chip failures such as electrode loss [9]. Additionally, a damaged dielectric layer can occur as a result of penetration of leftover fluorine atoms of the chemical ALD precursor (WF_6) during the tungsten deposition process [9]. To avoid these undesirable failures, voids in WL tungsten should be limited to a minimum. With increased vertical scaling, as well as XY-scaling [10], control over byproduct diffusion and oxide regrowth is expected to become more critical, as the width of the tungsten mold inlet continues to be optimized.

We propose a physical model which is capable of capturing the generation of etching byproduct concentration

and physical mass transfer, leading to an oxide regrowth behavior as previously demonstrated in experiments. The model is implemented in the in-house ViennaPS [11] tool, which provides combined level-set (LS) and cell-set (CS) geometry descriptions and detailed physical process models, as well as efficient process emulation tools.

II. Model

A. Combined Surface and Volume Description

For an accurate and efficient description of the structure throughout the simulation, we use a combination of the LS method for the surface topography and a CS for the underlying volume. The surface is described implicitly by a level set function $\phi(\vec{x})$ which is defined at every point $\vec{x}$ in space. This function is obtained using signed distance transforms, describing the surface S as the zero level set:

$$S = \{\vec{x}\colon\ \phi(\vec{x}) = 0\}. \tag{1}$$

In order to propagate the surface, the level-set equation

$$\frac{\partial \phi(\vec{x},t)}{\partial t} + v(\vec{x})|\nabla\phi(\vec{x},t)| = 0 \tag{2}$$

is solved in time t, given the scalar velocity field $v(\vec{x})$ which describes the surface normal velocity at each point. This is achieved by discretizing the level set function on a regular grid and applying a finite difference scheme to solve Eq. (2).

In our simulations, we employ the sparse field approach which is a memory-efficient implementation of the LS method [12]. In the sparse field implementation, only a few grid points close to the surface are stored enabling efficient memory usage. This is shown in Fig. 2, where the surface (white line) is described only by grid points with an LS value smaller than a given threshold value. In order to describe multi-material regions, multiple LS functions are stored in a prescribed order, such that each LS represents a material interface, while only the last LS describes the geometry surface.

The CS is stored over the whole simulation domain, above and below the surface described by the LS. It uses the same grid as the LS, such that each LS grid point acts as the corner of the cells around the point. To determine which material region a cell lies in, the LS values at the cell corner of the material interface are inspected. If the sum of the values is negative the cell belongs to the underlying material, otherwise it represents the material on the other side of the interface. The CS volume description is also shown in Fig. 2 by the differently colored voxels in the domain. In Fig. 2 only the surface LS is shown and therefore, all cells in gray represent the etching solution, while the green voxels show the top material layer of the geometry. Additionally, a different material region inside the geometry is shown by the yellow voxels.

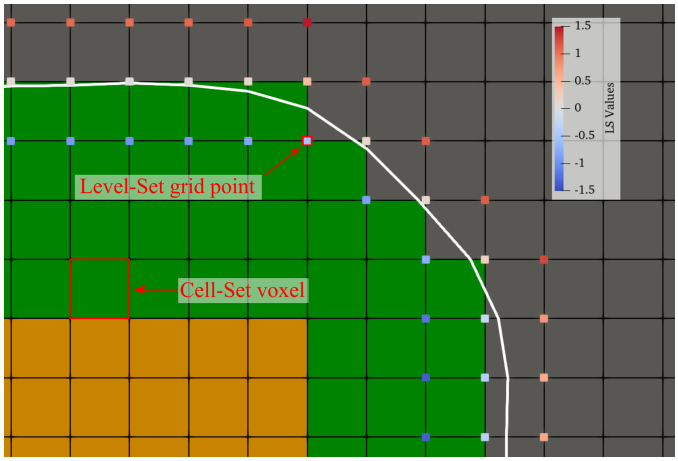

Fig. 2. Depiction of the surface of the structure (white line) described by the LS method on top of the CS volume description (voxels). While the LS stores the distance to the interface between material and etching solution, the CS can represent multiple material regions, shown by differently colored voxels, inside the material's volume.

B. Selective Etching and Byproduct Generation

The selective etching of Si_3N_4 is modeled as an isotropic etching process with constant rate $R_{Si_3N_4}$ and ideal selectivity ($R_{SiO_2} = 0$) due to the presence of SiO_2 etching inhibitors. The byproduct concentration in the etching solution is increased at the surface points while etching the Si_3N_4 layers. For each surface point which is moved during a time step Δt, a byproduct concentration proportional to $(\Delta t/\Delta x)R_{Si_3N_4}$, where Δx is the CS grid spacing, is added to the underlying voxel.

C. Mass Transfer

The mass transfer of the etching byproduct is modeled as a combination of convection and diffusion. Therefore, we use the transient convection-diffusion equation

$$\frac{\partial c(t,\vec{x})}{\partial t} = D\nabla^2 c(t,\vec{x}) - \vec{v}\cdot\nabla c(t,\vec{x}) + S \tag{3}$$

to describe the evolution of the byproduct concentration $c(t,\vec{x})$ in the etching solution over time. Here, D is a constant diffusion coefficient, $\vec{v}$ is a convection stream velocity field, and S describes a source or sink of the byproduct concentration $c(t,\vec{x})$. While the diffusion coefficient is chosen to be constant over the entire structure, the convection velocity field is set as follows: in the

rectangular vacancies after etching, a convection stream towards the central cavity with velocity v_t is assumed. In the previously etched trench in the middle of the structure, it is assumed that a convection stream with velocity v_c flows towards the bottom. The convection stream is assumed as a result of stirring the etching solution during the process. Additionally, a sink with constant strength S is placed on top to decrease the concentration outside the structure and model the removal of the etching byproduct.

The convection-diffusion equation is then solved numerically on the CS. We consider a Neumann boundary condition at the material interface, resulting in no concentration flux inside the material layers

$$\left.\frac{\partial c(t,\vec{x})}{\partial \vec{n}}\right|_{\vec{x}=\mathcal{I}} = 0, \tag{4}$$

where $\mathcal{I}$ describes the interface between the etching solution and the Si_3N_4/SiO_2 stack geometry. The initial byproduct concentration in the solution is assumed to be zero everywhere. After each etching step, the transient convection-diffusion equation is solved with the updated concentration as described in Section II-B until the previously performed etching time step Δt is reached.

D. Oxide Regrowth

To model the redeposition of byproducts, we accumulate byproduct concentration over every time step, expressed in the sum

$$c_{\text{acc}}(\vec{x},t) = \frac{1}{t}\sum_{i=0}^{N} \Delta t_i c(t_i,\vec{x}), \tag{5}$$

where Δt_i are the discrete etching time steps from time t_{i-1} to t_i. This distribution around each surface point, multiplied by a factor R_D, is used in the redeposition step to determine the growth rate of the oxide. Since the redeposition of oxide happens simultaneously while etching the surface, it is executed in regular time intervals during the selective etching process. This way, the clogging of the etched vacancy is intuitively considered in the model.

III. Results

The model uses four parameters to describe the mass transfer of byproduct concentration (D, v_c, v_t, and S) and one parameter to control the amount of oxide redeposition (R_D). These parameters were manually calibrated to fit the results of experimental studies by Kim *et al.* [5] and their values are shown in Table I.

TABLE I
CALIBRATED PARAMETERS OF THE OXIDE REGROWTH MODEL. THE PARAMETERS WERE CALIBRATED TO MATCH THE FINAL SURFACE PROFILES AS REPORTED IN REF. [2].

Parameter		Value	Unit
Diffusion coefficient	D	50	nm^2/s
Cavity stream velocity	v_c	5	nm/s
Trench stream velocity	v_t	7.5	nm/s
Sink	S	0.1	concentration/s
Redeposition factor	R_D	0.01	-

The simulated structure consists of a 64 SiO_2/Si_3N_4 pair-layered stack with a layer thickness of 30 nm and a trench opening with a width of 150 nm in the middle. The trench is generated using geometric process emulation, where a rectangular shape is removed from the surface and interface LS functions by performing Boolean operations. Subsequently, the selective etching model, in combination with the oxide redeposition model, is applied until the target etch depth of 200 nm is reached. The domain is discretized with a grid spacing of 2.5 nm resulting in approximately 475 000 cells. Since surface propagation, as well as byproduct diffusion, can be computed individually for each cell, the complete workflow can be parallelized with high efficiency. All simulations were carried out on a desktop workstation, using an 8-core AMD Ryzen 7 (3.80 GHz) CPU and a single run took around 5 minutes to complete. The resulting stack, with the final etching byproduct concentration, is shown in Fig. 3.

In the upper region of the stack, the etch byproduct can escape more quickly to the bulk etching solution outside the structure, while at the bottom, it tends to accumulate. As a result, the oxide regrowth rate is slightly increased at the bottom of the stack. This behavior matches the experimental findings by Kim *et al.* [5]. The redeposition factor R_D, on the other hand, can be directly connected to the amount of SiO_2 etch inhibitor in the solution. As reported by Kim *et al.* a higher amount of etch inhibitor leads to significantly more redeposition on the SiO_2 layers.

In Fig. 4, we compare the thickness profile of the redeposited layer to experimental observations. While using the same mass transfer parameters as in Fig. 3, the redeposition factor is adjusted to the experimental data. The width profile of the SiO_2 layer matches very well with the experimental results in Ref. [5].

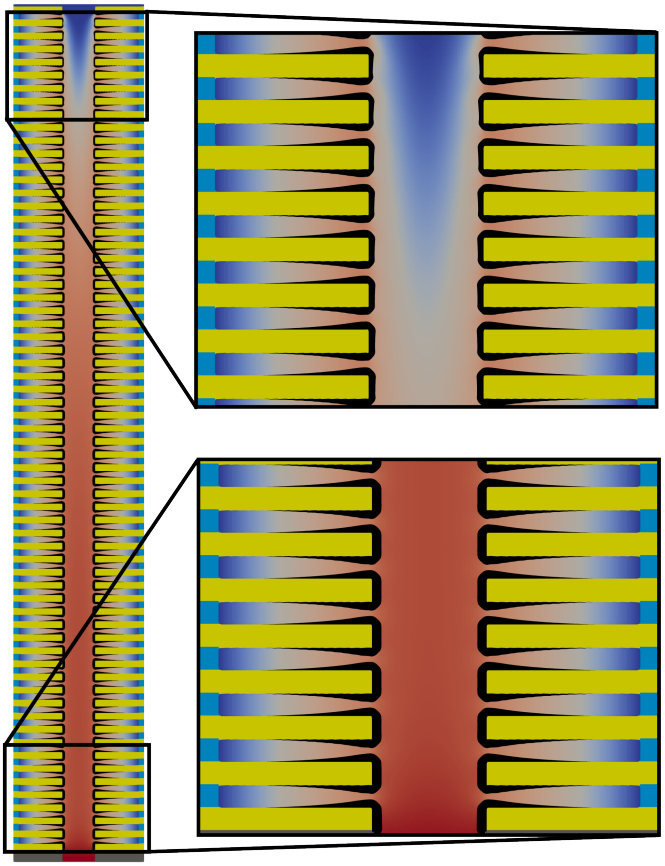

Fig. 3. Results of the selective etching and redeposition simulation in a 3D NAND stack with 64 SiO_2 (yellow)/Si_3N_4 (blue) pair-layers. The redeposited oxide (black) is shown on the SiO_2 layers, while in between the materials, the final byproduct concentration is shown. To qualitatively match the profiles of experimental results, the parameters were calibrated to the values presented in Table I.

IV. Conclusion

A physical process model is applied to simulate the oxide redeposition during selective etching of Si_3N_4 in a 3D NAND stack. The model can capture the formation and mass transfer of byproducts to identify areas of increased oxide regrowth. The proposed framework enables the analysis of the etch process' physical behavior, with the ultimate aim to improve the fabrication environment. Because of the physical nature of the presented model, it serves as a basis for understanding the abnormal oxide regrowth mechanism and its impact on the final 3D NAND structure.

Acknowledgment

Financial support by the Federal Ministry of Labour and Economy, the National Foundation for Research, Technology and Development and the Christian Doppler Research Association is gratefully acknowledged.

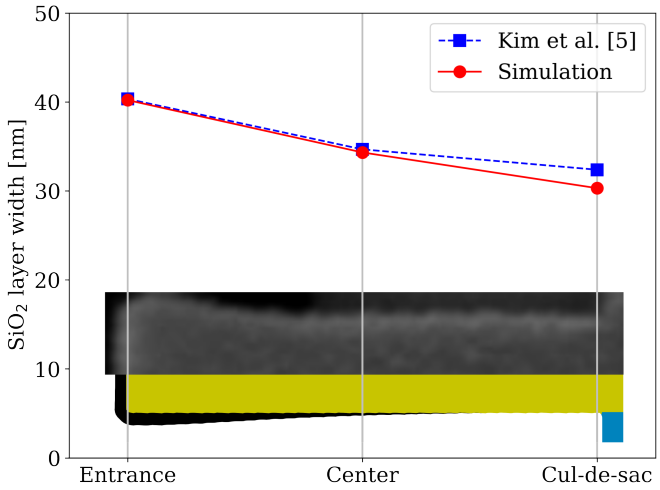

Fig. 4. The SiO_2 layer width, measured at the entrance, center, and back of the etched vacancy. The measured SiO_2 layer lies in the top region of the etched stack and is simulated with the same mass transfer parameters as in Fig. 3. The redeposition factor is calibrated to $R_D = 0.005$ to match the experimental results by Kim *et al.* [5].

References

[1] Y. Li, "3D NAND memory and its application in solid-state drives: Architecture, reliability, flash management techniques, and current trends," *IEEE Solid-State Circuits Magazine*, vol. 12, pp. 56–65, 2020.

[2] H. Kim *et al.*, "Evolution of NAND flash memory: From 2D to 3D as a storage market leader," in *IEEE International Memory Workshop (IMW)*, pp. 1–4, 2017.

[3] D. Bassett *et al.*, "Etching of silicon nitride in 3D NAND structures," *ECS Transactions*, vol. 69, pp. 159–167, 2015.

[4] D. Seo *et al.*, "Selective wet etching of Si_3N_4/SiO_2 in phosphoric acid with the addition of fluoride and silicic compounds," *Microelectronic Engineering*, vol. 118., pp. 66–71, 2014.

[5] T. Kim *et al.*, "Oxide regrowth mechanism during silicon nitride etching in vertical 3D NAND structures," *Microelectronic Engineering*, vol. 221, pp. 111191-1–111191-7, 2020.

[6] K.-W. Teng *et al.*, "Abnormal redeposition of silicate from Si_3N_4 etching onto SiO_2 surfaces in flash memory manufacturing," *Journal of Materials Science*, vol. 55, pp. 1126—1135, 2020.

[7] Z. Zhou *et al.*, "Redeposition mechanism on silicon oxide layers during selective etching process in 3D NAND manufacture," *Journal of Industrial and Engineering Chemistry*, vol. 119, pp. 218—225, 2023.

[8] J. Jang *et al.*, "Vertical cell array using TCAT(Terabit Cell Array Transistor) technology for ultra high density NAND flash memory," in *Proc. IEEE Symposium on VLSI Technology*, pp. 192–193, 2009.

[9] S. Chung *et al.*, "Process improvements for 7th generation 1Tb quad-level cell 3D NAND flash memory in mass production," in *IEEE International Memory Workshop (IMW)*, pp. 1–4, 2023.

[10] L. Heineck and J. Liu, "3D NAND flash status and trends," in *IEEE International Memory Workshop (IMW)*, pp. 1–4, 2022.

[11] T.Reiter *et al.*, ViennaPS-1.2.0 https://github.com/ViennaTools/ViennaPS.

[12] S. Ohser and R. Fedkiw, "Level Set Methods and Dynamic Implicit Surfaces," in *Applied Mathematical Series*, vol. 153., 2003.