Mechanisms controlling silicon nitride growth by plasma-enhanced atomic layer deposition

de Peuter, K.

Award date:
2014
Mechanisms controlling silicon nitride growth by plasma-enhanced atomic layer deposition

K. de Peuter

Feb. 2014

Under supervision of:

Dr.ir H.C.M. Knoops
Dr. C.K. Ande
Prof.Dr.ir W.M.M. Kessels

PMP 14-04
Summary

As the quest for increasing the number of transistors per chip continues, transistor designs are shifted to new materials. To fulfill the requirements for further performance increase and down-scaled, high-κ metal gate (HKMG) structures have been implemented as gate dielectrics. Within these transistors, sidewall spacers have the important task of protecting the HKMG structure and determining the spacing between the contacts in the transistor. Therefore, the spacer needs to have a high resistance against processing steps and a precise thickness. Silicon nitride is the most commonly used material for spacers, due to its high chemical inertness. The deposition of these films has to happen at low temperature, and the resulting films have to be conformal: the thickness of the film should not vary along the sidewall of the transistor and follow the topology of the structures very well. Conventional processes that deposit silicon nitride do not fulfill these requirements.

Plasma-Enhanced Atomic Layer Deposition is known to provide thin films with a low thermal budget, excellent conformality, and high quality. Although this technique has already been demonstrated for deposition of silicon nitride, an in-depth study of the mechanisms behind this process is necessary for further improvement. This thesis deals with several of these mechanisms involved in growth and material properties of SiN_x films grown with Bis(t-butylamino)silane (BTBAS) as precursor and a N_2 plasma. The growth mechanism was investigated by experiments using various reactants, complemented with simulations. As measurement for inertness, wet etch resistance in hydrogen fluoride was determined for a series of SiN_x films.

The growth experiments showed that a N_2 plasma is important in the growth of silicon nitride using BTBAS. The growth using a N_2 plasma was significantly higher than the growth using a NH_3 or N_2−H_2 plasma, which caused inhibition of the film surface. This inhibition could be reduced by exposing an inhibited surface to again a N_2 plasma. DFT simulations showed that this inhibition is occurring because the BTBAS precursor needs undercoordinated sites to adsorb, in contrast to other precursors that need N−H terminates surfaces. The NH_3 and H radicals that are present in H-containing plasmas can passivate these undercoordinated sites, disabling adsorption. Exposing a surface to a N_2 plasma creates undercoordinated sites, that enable adsorption.

By using the BTBAS/N_2 plasma ALD process at deposition temperatures of 300-500 °C, low etch rates of < 2 nm/min in buffered HF can be achieved. It was shown that the limiting factor in material quality is caused by redeposition of reaction products, which result in impurities in the film. This effect can be reduced by decreasing the residence time τ of particles in the plasma. A lower τ reduces the chance for products to dissociate and redeposit, achieving higher material purity and better etch resistance. Controlling the residence time allows to go to higher deposition pressures, which is favorable for film conformality. Excellent conformality was achieved in aspect ratios up to 3 whilst retaining low etch rates. The high inertness and conformality of these films are promising for usage of the BTBAS/N_2 process in spacer applications.
# Contents

1 **General introduction**  
1.1 Spacers in field-effect transistors ................................................. 1  
1.2 ALD of silicon nitride ................................................................. 4  
  1.2.1 Atomic layer deposition ....................................................... 4  
  1.2.2 Film conformality and quality ............................................... 5  
  1.2.3 Silicon nitride ALD: state of the art ....................................... 7  
1.3 Goal of the project ................................................................. 10  
1.4 Outline of this thesis ............................................................. 10

2 **Experimental setup and diagnostics** ............................................. 13  
2.1 Setup ......................................................................................... 13  
2.2 The silicon nitride ALD cycle ..................................................... 14  
2.3 Diagnostics ................................................................................ 16  
  2.3.1 Spectroscopic Ellipsometry ..................................................... 16  
  2.3.2 Scanning Electron Microscopy ................................................. 18

3 **Adsorption and inhibition mechanisms** .......................................... 19  
3.1 Growth experiments by multi-step cycles ....................................... 19  
  3.1.1 Conditions ............................................................................. 19  
  3.1.2 Results .................................................................................. 20  
3.2 Proposed inhibition mechanisms .................................................... 24  
3.3 Simulations .................................................................................. 25  
  3.3.1 Density Functional Theory ..................................................... 25  
  3.3.2 Crystal structure and properties ............................................. 26  
  3.3.3 Methodology ......................................................................... 28  
  3.3.4 Results .................................................................................. 29  
3.4 Summary ...................................................................................... 32

4 **Mechanisms regarding wet etch resistance and conformality** ............. 33  
4.1 Wet etching of silicon nitride ...................................................... 33  
  4.1.1 Steady-state versus apparent etch rate .................................... 34  
  4.1.2 Extraction of the steady-state etch rate ................................... 34  
  4.1.3 Etching mechanism in HF solutions ....................................... 36  
4.2 Etch experiments on planar wafers ............................................... 39  
  4.2.1 Etching procedure and conditions ........................................... 39  

v
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.2 Results</td>
<td>40</td>
</tr>
<tr>
<td>4.3 Identifying the governing mechanisms in material properties</td>
<td>44</td>
</tr>
<tr>
<td>4.4 Conformality and etch selectivity in HARS</td>
<td>48</td>
</tr>
<tr>
<td>4.5 Summary</td>
<td>51</td>
</tr>
<tr>
<td>5 General conclusions &amp; Outlook</td>
<td>53</td>
</tr>
<tr>
<td>5.1 General conclusions</td>
<td>53</td>
</tr>
<tr>
<td>5.2 Outlook</td>
<td>54</td>
</tr>
<tr>
<td>Bibliography</td>
<td>57</td>
</tr>
<tr>
<td>Appendix</td>
<td>65</td>
</tr>
<tr>
<td>A Relation between setpoint temperature and wafer temperature</td>
<td>65</td>
</tr>
<tr>
<td>B Matlab script for calculating etch rates</td>
<td>67</td>
</tr>
<tr>
<td>C Derivation of the residence time</td>
<td>69</td>
</tr>
</tbody>
</table>
1 General introduction

1.1 Spacers in field-effect transistors

Continuously increasing the number of transistors per chip is essential for the success of the semiconductor industry. The subsequent reduction of transistor size has come to a critical point as the conventional gate oxide has reached its scaling limit. Various problems are arising, which is causing the IC industry to shift its attention to new materials and structures. These new materials have their own challenges, for which new research is needed.

There are several different types of Field-Effect Transistors, FETs in short, but types that make use of a gate oxide are usually called metal-oxide-semiconductor FETs (MOSFETs). The MOSFET is the most-important building block for forefront high-density integrated circuits such as microprocessors and semiconductor memories [1]. The principle of the surface field-effect transistor was first proposed in the early 1930’s by Lilienfeld [2–4] and Heil [5]. The detailed early historical development of the MOSFET can be found in refs. [6, 7]. The technology, application, and device physics have been reviewed by many books [1, 8–10].

The basic structure of a MOSFET is illustrated in fig. 1.1a. Common MOSFETs are often divided into two types: NMOS and PMOS. An NMOS transistor consists of a $p$-type semiconductor substrate into which two $n^+$-regions, the source and drain, are formed, usually by ion implantation. The PMOS transistor consists of an $n$-type substrate where the source and drain are $p^+$ doped. The metal contact on the insulator is called the gate; so far heavily doped polysilicon or a combination of silicide and polysilicon are most commonly used as the gate electrode. The insulator is usually an oxide, and is called the gate dielectric.

The most important function of a MOSFET is a switch. This works in the following way. Usually, the Si substrate is insulating when there is no voltage applied across the gate dielectric. Now when a voltage $V$ is applied, a charge $q$ accumulates below the gate given by the relation $q = CV$, with $C$ the capacitance of the gate dielectric. This charge gives rise to a conducting channel through which current can flow from the source to the drain, making the device work as a switch. The capacitance $C$ is given by $C = \kappa \varepsilon_0 A/d$, where $\kappa$ is the dielectric constant of the gate dielectric (often also expressed as the relative permittivity $\varepsilon_r$), $\varepsilon_0$ the permittivity of vacuum, $A$ the surface of the dielectric, and $d$ its thickness.

This dielectric has been (nitrided) silicon oxide until the late 2000’s [11]. In the last few decades, tremendous improvements in device processing have produced high-quality SiON layers with the required thickness, very few electronic defects in amorphous structure, and excellent interface formation with Si. However, the nitried silicon oxide has reached a fundamental limit at which further downscaling gives rise to direct electron tunneling. This tunneling results in an increased leakage current and power consumption [12–14]. For gate oxide layers of $\sim$1 nm,
electron tunneling results in a leakage current level that is unacceptably high even for high-performance devices, and much too high for increasingly popular low-power applications [11, 15].

To overcome these problems, new high-κ gate oxides have been developed. A higher κ allows for thicker gate oxide layers whilst keeping the capacitance constant, therefore decreasing tunneling current but keeping the device physics the same. Amongst the most promising high-κ materials are hafnium oxide, HfO$_2$, or hafnium-based compounds such as HfSiO$_x$, with dielectric constants of about 16-26 (fig 1.1b) [11]. In comparison, the dielectric constant of SiO$_2$ is 3.9. These structures are already applied in industry: figure 1.2a shows Apple Inc.’s old A4 processor with a SiON gate dielectric introduced in September 2010, while figure 1.2b shows Apple Inc.’s new sub-30 nm A7 FET, introduced in September 2013, where the high-κ structure (together with a metal gate, MG, forming a so-called HKMG structure) has already been implemented [16].

The usage of the HKMG structures brings some challenges. HfO$_2$ and TiN (which is commonly used as the MG) are materials very sensitive to oxygen ingress and moisture from the surroundings [12, 17]. To protect these structures, a sidewall spacer is deposited around these layers (SWS, figs. 1.1, 1.2). A material commonly used for these kind of spacers is silicon nitride. Silicon nitride is a material that provides an excellent barrier against water molecules and oxygen. It is also chemically inert, which helps to protect the HKMG structure against further processing steps as source and drain formation by ion implantation. This inertness is an important requirement, because the spacer determines the positions of the source and drain. If the quality and thickness of the spacer are not uniform throughout the transistor sidewalls, undercutting of the spacer can occur, placing the source and drain too close together and thereby increasing leakage current or worsening transistor performance. As a measure for inertness (or quality) and barrier properties, often the resistance of the film against wet acid environments such as HF is determined.

Since these devices get smaller and closer packed, a technique that can create conformal layers
1.1. Spacers in field-effect transistors

of high-quality SiN, with a high step coverage is needed. As an indication for the closest distance between structures, often the term single- or narrow pitch is used. In 28/32 nm technology, single-pitch corresponds to an aspect ratio (the ratio of height of the FET and width between the FETs) of $\sim$1, and is expected to increase even more. A higher aspect ratio makes the deposition of conformal layers increasingly difficult. Furthermore, the deposition has to be done preferably at fairly low temperatures ($\leq 500 ^\circ$C) because the HKMG becomes unstable above that temperature [11, 15].

Until now, existing processes that deposit silicon nitride (LPCVD or PECVD) require deposition temperatures above the threshold of 500 °C, or do not result in conformal films for these aspect ratios [12]. In order to keep up with Moore’s Law, a new process is required that meets all above requirements. This project focusses on creating such a technique with Atomic Layer Deposition. In section 1.2 this technique is explained and an overview is given of different processes up to present.

![Figure 1.2:](image1.png)

**Figure 1.2:** (a) TEM image of Apple Inc.’s 45 nm technology transistor, as introduced in Sept. 2010, showing the polysilicon (poly) and the spacer (SWS). (b) TEM image of Apple Inc.’s new 28 nm (PMOS) transistor device introduced in Sept. 2013. The image shows the high-$\kappa$ material (HK), the metal gate (NMOS & PMOS MG), the sidewall spacer (SWS), the polysilicon (poly) and the source, drain and gate contacts. Adapted from ref. [16].
1.2 ALD of silicon nitride

1.2.1 Atomic layer deposition

ALD is a vapor phase thin film deposition technique that allows for the deposition of films with a good uniformity, high conformality, and accurate thickness control [18, 19]. It generally leads to films with a high material quality. ALD relies on self-limiting surface reactions that take place during the alternate pulsing of precursor and reactant gases, separated by purge or pump steps. To explain the details of these self-limiting reactions and how they translate to conformal and controlled deposition, the classical example of ALD of Al₂O₃ is discussed below [20]. See fig. 1.3 for a schematic representation of this ALD process.

![Figure 1.3: Schematic representation of thermal ALD of Al₂O₃ and plasma-assisted ALD. During the co-reactant step of the cycle (the 2nd half-cycle), the surface is exposed to a reactant gas or vapor such as NH₃ or H₂O, or to species generated by a plasma.](image)

The ALD process of Al₂O₃ consists of alternate pulses, referred to as half-reactions, of an Al(CH₃)₃ precursor and H₂O. Assume as a starting surface, an oxide substrate covered with hydroxyl surface groups. During the first step of the process, the substrate is exposed to the Al(CH₃)₃ precursor molecules that adsorb at the surface in a reaction with the surface hydroxyl groups. In a ligand-exchange reaction between the hydroxyl groups and the methyl ligands of
1.2. ALD of silicon nitride

the precursor, Al-O bonds are formed and CH$_4$ is released as a volatile reaction byproduct. The result is a surface terminated by $-CH_3$ groups, which limits the adsorption of more precursor molecules, and ensures the self-limiting nature of the first half-reaction. Subsequently, the excessive amount of precursor and the reaction products are purged from the chamber by an inert gas. The third step involves the exposure of the $-CH_3$ terminated surface to H$_2$O vapor. During this step, the H$_2$O molecules react with the $-CH_3$ surface groups, resulting in CH$_4$ reaction products and the formation of new hydroxyl groups at the surface. This half-reaction is also self-limiting, since it is completed when the H$_2$O has reacted with all $-CH_3$ groups. The last pulse of the process is again a purge step. After these four steps, together referred to as an ALD cycle, a submonolayer of $\sim$1 Å Al$_2$O$_3$ has been deposited. Since the surface is again terminated with hydroxyl groups at the end of a cycle, the process can be repeated until the desired thickness is reached.

The Al(CH$_3$)$_3$/H$_2$O ALD process is a so-called thermal ALD process. To widen the temperature window and increase the types of material that can be deposited, a plasma can be used as reactant as well (bottom figure in fig. 1.3). Besides the increased thermal budget and materials, using a plasma can also result in improved and controlled material properties, increased GPC and more process versatility [21]. These merits emerge from the fact that a plasma contains highly reactive species that need less activation energy to react with surface groups.

Although plasma-enhanced ALD can offer several benefits over thermal ALD, the method also faces some challenges when compared to its thermal counterpart. One limitation that is often highlighted is the higher difficulty to gain high conformality (or step coverage) in 3D structures. These can be substrates with surface structures of high aspect ratio, e.g. trenches or nano pillars, or substrate materials with high surface areas, e.g. porous materials. This reduced conformality has multiple causes, which can be influenced by plasma parameters as pressure, exposure time and gas flow rate. Besides influencing conformality, these parameters also influence material properties such as refractive index, stoichiometry and impurity content. Plasma-enhanced ALD brings an extra parameter to take into account: plasma power. This added parameter allows for extra optimization, but it makes the process more complex. Factors of influence in the conformality and quality of films deposited using plasma-enhanced ALD are discussed in the next section.

1.2.2 Film conformality and quality

As explained in the previous section, there are several factors to consider regarding the conformality and properties of PEALD films. A first factor is that reactive species from the plasma not only undergo ALD reactions (as in thermal ALD) but can also react on saturated surface sites. Radicals can recombine with other radicals and species residing at the surface, forming non-reactive molecules that desorb back into the plasma. In high aspect ratio structures the radicals typically have to undergo several surface collisions in order to reach deep into the surface features, which significantly reduces the local flux of radicals due to surface recombination [21, 22]. In many other reports, the impact of the surface recombination of plasma radicals is generalized and the poor conformality of plasma-assisted ALD films is often claimed. However, the conformality achieved by plasma-assisted ALD for certain conditions depends strongly on the value
Chapter 1. General introduction

of the recombination probability, \( r \), which itself depends on (a) the type of radicals responsible for film growth in a certain plasma-assisted ALD process and (b) the material being deposited. The influence of radical recombination in conformality can be compensated by increasing the exposure \( E = \rho t \), the product of radical density and exposure time. The longer one waits, the more particles can diffuse into the structure, eventually occupying all surface sites. Increasing the radical density increases the particle flux into the structure, leading to faster occupation of surface sites. The radical density can in general be increased by increasing the pressure of the plasma [21].

The radical recombination mechanism is often highlighted as being the biggest factor in conformality and is not so much of influence for material quality. However, there are other factors that are often overlooked but still need to be taken into account. A second factor that influences both conformality and quality is the directionality and energy of ions in the plasma. Ions have a trajectory which is mostly perpendicular to the surface due to acceleration in the plasma sheath, and therefore have anisotropic collisions with respect to horizontal and vertical surfaces. If an ALD process depends on these ions, e.g. it requires the presence of specific surface groups which are created or activated by ion impact, this anisotropy can induce growth differences and variation in film quality on sidewall and horizontal surfaces. The typical angle of dispersion within this ion angle distribution is given by \( \theta_{\text{IAD}} \) [23]:

\[
\theta_{\text{IAD}} \propto \arctan \left( \sqrt{\frac{k_B T_i}{E_i}} \right),
\]

where \( T_i \) is the temperature of the ionic species and \( E_i \) the kinetic energy of the ions upon collision with the surface. In order to have more isotropic collisions, the ion energy needs to be low. However, a positive effect of higher ion energy is that more energetic impacts by ions may break surface bonds, aiding in the removal of precursor remnants or other unwanted species and densifying the film, which generally results in a film of higher quality [24].

Ion bombardment is mostly influenced by the power coupled into the plasma and the plasma pressure. A higher power means more ions are created, increasing bombardment [24]. A higher pressure means that ions undergo more collisions before they reach the surface, their impact becoming less energetic. This in turn means that a higher pressure results in more dispersion of the ions.

A third factor in film conformality and quality can be redeposition of reaction products [25]. When reaction products, such as precursor ligands, are removed in the reactant half-cycle, there is a chance that these products can redeposit on the surface. For a half-cycle where a plasma is involved, this chance is bigger because reaction products can interact with the plasma, dissociate, form new products that may be more reactive, and redeposit on the surface more easily. These redeposited products can change the composition of the film, or inhibit precursor adsorption in the next cycle, poisoning the surface. Besides the fact that this effect can occur even on planar substrates, one can imagine that in a 3D structure, redeposited species at the bottom of the trench or pore are harder to remove and therefore can influence the composition or growth of the film along the sidewalls of the 3D structure. An important parameter for this effect is called the residence time \( \tau = V / q \), the average time a particle resides in a certain volume \( V \) when it is
1.2. ALD of silicon nitride

driven by a volume flow rate $q$ (in volume per time). A lower residence time means the reaction products have less time to become dissociated and diffuse to the surface, which is favorable for producing films with less impurities and inhibiting species.

To summarize this section: there are three main factors of influence in the conformality and quality of films. These are radical recombination, ion bombardment and residence time. Each of these factors can be influenced by plasma parameters as pressure, exposure time, or flow rate. For conformal films we would need: 1. A high exposure so sufficient particles can diffuse deep into the structure, 2. Low ion bombardment to reduce anisotropy, and 3. Low residence time so reaction products are removed more efficiently from the plasma, thus preventing redeposition.

1.2.3 Silicon nitride ALD: state of the art

The growth process of silicon nitride is similar to that of $\text{Al}_2\text{O}_3$ in the sense that it is a binary material and therefore usually needs a precursor that contains the first element and a reactant that contains the second element. In the case for silicon nitride, this is a silicon-containing precursor and a nitrogen-containing reactant. However, in some cases both the elements can come from only one reactant, and the second reactant is only necessary for activating the surface. Conventionally, the technique used for depositing silicon nitride is LPCVD, and this technique is still widely used in industry. For the case of ALD, to date, there are many precursors that can be considered to deposit silicon nitride, the main types being silicon hydride, silicon halide and organosilane precursors. Ammonia, nitrogen or mixtures thereof with hydrogen can be considered as reactants, either as a gas or as a plasma. A small summary of reported processes is found in table 1.1.

**Silicon halides**

Silicon halide precursors ($\text{SiCl}_4$, $\text{SiH}_2\text{Cl}_2$, or $\text{Si}_2\text{Cl}_6$) have mostly been reported in the literature and are mainly used in combination with NH$_3$ gas or plasmas and result in growth per cycle (GPC) between 0.9 Å and 2.7 Å for precursor dosing times of 15-90 s [26, 27], which is quite long compared to many other ALD cycles, with typical dosing times of 0.5-5 s [21]. The refractive index of these films ranges from 1.6 to 2.01 [27]. Typical temperature at which the depositions are conducted is 300-600 °C. Usage of these processes for deposition in high aspect ratio structures (HARS) has also been reported and resulted in a conformal film for so-called narrow or single pitch structures [12]. The wet etch resistance of these films was $< 1 \text{nm/min}$ in 1:100 diluted HF. A major disadvantage of using metal halide precursors is that they can lead to formation of ammonium chloride, which accumulates at/in several reactor parts and can have effects on further processing steps [28].

**Silicon hydrides**

An example of a silicon hydride is silane ($\text{SiH}_4$) of which an ALD process for silicon nitride was reported using a nitrogen plasma, resulting in GPCs between 0.25 Å and 2 Å at deposition temperatures of 250-400 °C [29]. Optimal material quality and GPC required 60-90 s plasma
exposure time, resulting in long cycle times which make the process basically not suitable for industrialization. In the same work also a reaction mechanism was proposed for the adsorption of precursor on the silicon nitride surface [29].

**Aminosilanes**

Recently, an ALD process for silicon nitride was reported using trisilylamine and a H$_2$-N$_2$ plasma [17]. This molecule contains three silicon atoms which surround a nitrogen. The precursor design seems promising, since it only contains elements that are present in silicon nitride, so there are no sources of impurities except hydrogen. With this process, GPCs of 1.2 Å to 2.1 Å were reported for deposition temperatures between 300 °C and 400 °C. Nitrogen content was typically within the range of 52-53 %. Wet etch rates in 1:100 HF were around 1 nm/min. Conformality of the layers was ∼ 84 % in single-pitch structures. Cycle times were not reported.

**Organosilanes**

A wide variety of PEALD using organosilane precursors was investigated in previous work [30]. Most of these processes did not result in stable films, partially due to the effect that with organosilane precursors some carbon or hydrogen contamination can remain in the film. GPCs of 0.2 Å to 0.4 Å were reported, together with refractive indices varying from 1.7 to 1.95, also partially due to oxidation. BTBAS was also used as precursor in combination with a NH$_3$ plasma, but that process did not lead to self-limiting growth [30].

**Recent work: BTBAS**

In the PMP research group at the TU/e, a new process using BTBAS has been developed for the deposition of Si$_N_x$ layers [31]. In contrast with the results from ref. [30], it was possible to grow silicon nitride by using an inductively coupled N$_2$ plasma instead of an NH$_3$ plasma. Using a H-containing plasma (NH$_3$ or N$_2$–H$_2$ in various compositions) did result in self-limiting growth as well, but the growth rates were extremely low. Several mechanisms for this inhibition were proposed, but additional research is needed to find the cause of this behaviour. This report will further investigate these observations.

By using only a N$_2$ plasma, growth rates for optimized processes were achieved between 0.33 Å and 0.15 Å for a dosing time of only 150 ms, for temperatures between 200 °C and 500 °C, respectively$^1$. The refractive index, N/Si ratio and mass density of films deposited in this temperature range, reached values close to those of stoichiometric Si$_3$N$_4$ and the carbon and hydrogen contents decreased with increasing $T$ to ∼ 5 at.% each (see fig. 1.4a). The uniformity of thickness and refractive index improved with increasing temperature as well. Also, the long-term stability of the films improved with increasing temperature, and for plasma exposure times $t \geq 7.5$ seconds, temperatures $T \geq 200$ °C, and pressures $p \leq 40$ mTorr the films showed no signs

$^1$Note that the temperatures mentioned here are the setpoint temperatures of the substrate table of the reactor; the actual wafer temperature will be approximately 50-100 °C lower depending on the setpoint temperature. This will be explained in section 2.1.


1.2. ALD of silicon nitride

of oxidation in the first two months after deposition. Room-temperature ALD was also investigated, but preliminary results showed low-density films with high impurity levels. A particularly strong dependence of film properties with the deposition pressure $p$ and the exposure time $t$ were observed, where low pressures and long exposure times yielded better results (fig. 1.4b). These trends were tentatively attributed to the contribution of ion bombardment to film growth.

Figure 1.4: (a) Concentration of C, H, and O in the SiN$_x$ as a function of setpoint temperature, as measured by Rutherford Backscattering Spectroscopy. The C and H content decreased with increasing temperature. The oxygen levels were low and did not vary significantly with temperature. (b) Concentration of N, Si, C, and O in the SiN$_x$ as a function of plasma exposure time for a deposition temperature of 200 °C. The C content decreased with increasing exposure time, while the Si content increased. Also here, the oxygen level did not vary significantly with plasma time. Graphs adapted from ref. [31].

Summary

Table 1.1 shows a summary of reported processes and results. Literature shows that in terms of etch resistance and conformality, the metal halide precursors yield the best results. The major downside of these precursors, as explained earlier in this section, is the formation of ammonium chloride that can accumulate in the reactor and block pump lines [28], together with the long exposure times that are necessary to reach saturation. These downsides are causing that the attention in IC industry is shifting towards organosilane precursors. There, the main challenges are to decrease and deplete the carbon and hydrogen contamination from ligand remnants in the silicon nitride film in order to increase the quality, and to increase the conformality of these films. In order to accomplish these goals, a thorough understanding is needed of the growth mechanism of PEALD SiN$_x$ and of the parameters that govern the etch resistance and conformality.

9
1.3 Goal of the project

The primary goal of this project is to study the underlying mechanics regarding the growth and properties of silicon nitride films grown using BTBAS and a \( \text{N}_2 \) plasma. This process was recently developed in this research group, and seemed promising for spacer applications. When the mechanisms of the BTBAS/\( \text{N}_2 \) plasma process are well understood, films with optimal properties can be produced and applied in spacer technology. This understanding can also help other PEALD processes as well. In order to find these mechanisms, the project is divided into two sub-goals:

**Adsorption and inhibition mechanisms**

Not yet much is known about the mechanism involved in the growth of Si\( \text{N}_x \) layers using BTBAS and a \( \text{N}_2 \) plasma. From recent research in this group (see section 1.2.3), it was shown that whereas using a \( \text{N}_2 \) plasma does show growth, using a \( \text{NH}_3 \) or \( \text{N}_2-\text{H}_2 \) plasma does not, or results in a very low growth rate. To investigate this behavior, computational and experimental results are combined to describe an adsorption mechanism of BTBAS to a silicon nitride surface and the role of N, H, and \( \text{NH}_3 \) therein.

**Mechanisms regarding wet etch resistance and conformality**

In order for silicon nitride films to be viable for spacer applications, they have to be of excellent quality along the transistor sidewalls, as explained in section 1.1. As a measurement of quality, often wet etch rates in HF are determined. Therefore, the wet etch resistance is determined by exposing a series of Si\( \text{N}_x \) films, deposited under several conditions, to various concentrations of solutions of HF. From those results, the limiting factors in material quality as mentioned in section 1.2.2 are identified. Thoroughly understanding these mechanics can be used to increase film wet etch resistance to an optimum.

The conformality of the process is investigated as well. Silicon nitride films are deposited in high aspect ratio structures (HARS) and the sidewall to top thickness ratio of the films in various aspect ratios is determined. Furthermore, qualitative information about etch uniformity is obtained by dipping these structures in HF and then again looking at the film thickness in these structures.

1.4 Outline of this thesis

This thesis will be divided into chapters devoted to the sub-goals of this project. Chapter 2 discusses the setup and ALD process used for conducting depositions and the \textit{in situ} measurements. It will also discuss the diagnostics used for process and material characterization.

Chapter 3 discusses the adsorption mechanism of BTBAS and how H-containing plasmas can inhibit this adsorption. It will show experimental results on the influence of different reactants on the growth per cycle, and suggest possible inhibition mechanisms. Subsequently, Density Functional Theory simulations are performed to computationally describe the adsorption and
1.4. Outline of this thesis

Inhibition processes. A summary of the experimental and computational results is found in section 3.4.

In chapter 4 the mechanisms regarding etch resistance and conformality are studied. First, a small introduction of wet etching is given, and the pitfalls in determining these etch rates are identified. Furthermore, an etching model is used to predict etch rates of silicon nitride and silicon oxides as a function of etchant concentration. These computed values are then compared with etch rates obtained from a series of experimentally grown films.

From the etch results, it is shown that redeposition of reaction products plays an important role in the ALD of silicon nitride, and that the residence time $\tau$ is the limiting factor in material quality. The dependency of material properties on this parameter is studied, and implications for deposition conditions are discussed. This parameter is then used to optimize the etch resistivity. Section 4.4 shows films deposited in high aspect ratio trenches for several optimal conditions.

This report will conclude with combining the results on adsorption, inhibition, etch resistance, and conformality in order to assess the applicability of the BTBAS/N$_2$ ALD process for spacers in new transistor designs.
## Table 1.1: Summary of the comparison between various reported silicon nitride ALD processes.

<table>
<thead>
<tr>
<th>Precursor</th>
<th>Organosilane</th>
<th>Ammonosilane</th>
<th>Ammonia</th>
<th>Mass density (g/cm³)</th>
<th>N/Si</th>
<th>Cl (at%)</th>
<th>H (at%)</th>
<th>O (at%)</th>
<th>GPC (˚A)</th>
<th>Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTBAS</td>
<td>2.6</td>
<td>0.4</td>
<td>13</td>
<td>6.3 / 0.5 &gt; 7.8</td>
<td>100</td>
<td>1.5</td>
<td></td>
<td></td>
<td>2.2-2.9</td>
<td></td>
</tr>
<tr>
<td>SiH₄</td>
<td>1.35</td>
<td>0.9</td>
<td>1.2</td>
<td>1.7-1.8</td>
<td>180</td>
<td>1.2</td>
<td></td>
<td></td>
<td>2.6-3.1</td>
<td></td>
</tr>
<tr>
<td>SiCl₄</td>
<td>2.0</td>
<td>1.9</td>
<td>2.2</td>
<td>1.7-2.1</td>
<td>150</td>
<td>1.0</td>
<td></td>
<td></td>
<td>2.4-2.7</td>
<td></td>
</tr>
<tr>
<td>SiCl₂H₂</td>
<td>2.5</td>
<td>2.4</td>
<td>2.4</td>
<td>2.3-2.5</td>
<td>150</td>
<td>1.0</td>
<td></td>
<td></td>
<td>2.5-2.9</td>
<td></td>
</tr>
<tr>
<td>SiCl₂(2H)</td>
<td>2.8</td>
<td>2.5</td>
<td>2.8</td>
<td>2.3-2.5</td>
<td>150</td>
<td>1.0</td>
<td></td>
<td></td>
<td>2.5-2.9</td>
<td></td>
</tr>
</tbody>
</table>

When there are no values mentioned, the corresponding properties have not been reported.

**Reference**

[33] [26] [27] [32] [17] [31] [29] [33] [32] [27] [26] [33] [12]
2 Experimental setup and diagnostics

In this chapter, a short overview of the experimental setup and diagnostics will be given. The ALD reactor which is used for the deposition of silicon nitride will be explained in section 2.1. To provide a better understanding of the silicon nitride ALD process, section 2.2 will explain the various steps in the ALD cycle used in this work, and describe the BTBAS precursor. Section 2.3 will explain the techniques used for diagnostics and the measurement procedures.

2.1 Setup

The depositions of the silicon nitride layers were done in a remote plasma ALD reactor developed by Oxford Instruments, called FlexAL 2. Figure 2.1 shows a schematic of this reactor. It can be used to deposit on regular wafers varying from 2 to 8 inch directly, and on other samples (e.g., pieces of wafers, double-side polished wafers, etc.) via a carrier wafer. These samples can be loaded into the reactor through a loadlock. The reactor’s effective volume is 55 L.

![Figure 2.1: A schematic overview of the FlexAL 2, showing the main components of the reactor. The valves are depicted in red and the heated substrate table is shown in blue. The ICP source is positioned on top of the reaction chamber. The in situ diagnostics SE, QMS and OES are depicted as well. The loadlock, precursor bubblers and precursor delivery lines are not shown.](image-url)
Chapter 2. Experimental setup and diagnostics

The substrate table of the FlexAL 2 can be heated to a maximum of 500 °C. The deposition wafer is placed directly on the substrate table. The wafer is not clamped to the table, and contacts the table only at several points. These factors cause the temperature of the deposition wafer $T_{\text{wafer}}$ to be 50-100 °C lower than the substrate table setpoint temperature $T_{\text{set}}$ at higher temperatures, and to depend on the pressure in the reactor, since the pressure influences thermal conductivity [31]. As a guideline for the difference between the setpoint temperature and actual wafer $T_{\text{wafer}}$ temperature, a relation between those two temperatures can be found in Appendix A.

The walls and precursor supply lines are heated to 150 °C and 70 °C, respectively, which are necessary to avoid precursor condensation. The precursor bubbler is heated to 50 °C. During the ALD process the precursor is vapor drawn into the reactor through a supply line which has an injection point below the plasma source. This supply line can also be used to introduce gases other than the precursor in the reactor.

The plasma source is positioned on top of the reactor. Its inductively coupled plasma is controlled by an automatic matching network for an efficient power coupling from the RF source, operating at 13.56 MHz, into the plasma. The input power is set to a maximum of 600 W. The plasma gases are introduced through a supply line above the plasma source.

The pressure in the reactor is controlled by a butterfly valve located above the turbo pump. The angle of this valve determines the effective pumping of the reactor. The control software offers options to actively regulate the pressure by varying the valve angle, or setting a fixed angle on the butterfly valve.

Furthermore, the reactor has three in situ diagnostic tools: spectroscopic ellipsometry, quadrupole mass spectrometry and optical emission spectroscopy. Of these three tools, only SE is used to monitor the film growth during the deposition process. A short overview of this diagnostic is found in section 2.3. The ellipsometer is attached to the reactor under an angle of 70° with respect to the vertical plane by quartz windows. During deposition, these windows are closed by valves to avoid depositions on the windows that might influence the measurements.

2.2 The silicon nitride ALD cycle

The ALD process for silicon nitride from the BTBAS precursor and a $\text{N}_2$ plasma consists of the repetition of the following cycle:

1. Pre-dosing: An argon flow is used to purge the precursor delivery lines from any contaminants or left reaction products from previous ALD cycles;

2. Precursor dosing: The bubbler valve is opened and BTBAS is vapor drawn into the reactor. During this step (150 ms) the argon flow is diverted to the exhaust;
3. Reaction step: Because the precursor dosing time is very short, a reaction step of 3 seconds is included to allow the precursor to adsorb on the surface. The precursor lines are purged with argon to ensure no precursor is left in the lines. To avoid the pressure building up too high, the butterfly valve is nearly closed by setting the angle to 10°;

4. Precursor purge: The butterfly valve is opened completely to remove all reaction products and pump down the chamber. Argon is still flowing to purge the precursor delivery lines;

5. Pre-plasma: The plasma gases are switched on to ensure that the flows are stabilized when the plasma is ignited;

6. Plasma: The plasma is ignited. Typical plasma exposure times are 8-15 s;

7. Purge plasma: The reaction chamber is purged using argon and pumped down to remove the remaining plasma species and reaction products from the reactor.

All purge steps have a duration of 1 s, this was found to be enough to obtain self-limiting behavior. The ALD cycle, with a total duration of about 19 s, is repeated and after a certain number of cycles the process can be halted to perform an SE measurement. The desired film thickness can be achieved by repeating this sequence of cycles. For experiments where other gases are involved, more steps can be inserted in the recipe.

The BTBAS (molecular formula: SiH$_2$(NHC$_4$H$_9$)$_2$) is purchased from Air Products Inc. with a purity of ≥98.5% [34]. The BTBAS molecule consists of a silicon atom coordinated with two hydrogen atoms and two t-butylamine ligands, as shown in figure 2.2. Since the molecule itself already contains N, a part of the N in the deposited SiN$_x$ film might come from the precursor instead of the N$_2$ plasma. When the precursor is heated to 50 °C (the bubbler temperature), it has a vapor pressure of 8.4 Torr [34]. The nitrogen gas that is used in the ICP is industrial grade with 99.999% purity. In several experiments ammonia gas is used, of commercial grade with 99.6% purity. Also hydrogen is used during other experiments, with a purity of 99.999%.

![Figure 2.2](image.png)

*Figure 2.2:* A schematic view of the bis(t-butylamino)silane (BTBAS) precursor, showing the Si atom (blue), the amino groups (gray) and the butyl ligands (brown). The pink spheres represent H.
Chapter 2. Experimental setup and diagnostics

2.3 Diagnostics

2.3.1 Spectroscopic Ellipsometry

Ellipsometry is an optical measurement technique that is based on light reflection (or transmission) from samples [35] and is generally used for determining film thickness, refractive index and other optical constants. The key feature of ellipsometry is that it measures the change in polarization of light upon reflection on a sample (or transmission by a sample in transmission ellipsometry). The name 'ellipsometry' comes from the fact that linearly polarized light often becomes elliptically polarized upon light reflection. Ellipsometry measures two values, $\psi$ and $\Delta$.

These represent the amplitude ratio $\psi$ and phase difference $\Delta$ between light waves known as $p$- and $s$-polarized light waves (see fig. 2.3). In spectroscopic ellipsometry, $(\psi, \Delta)$ spectra are measured by using a broadband light source. In general, the spectroscopic ellipsometry measurement is carried out in the ultraviolet/visible region, but measurements in the infrared region have also been performed widely.

![Figure 2.3: Measurement principle of spectroscopic ellipsometry.](image)

Ellipsometry measures the complex reflectance ratio, $\hat{\rho}$, of a system, which can be parameterized by the amplitude component $\Psi$ and the phase difference $\Delta$. The polarization state of the incident light on the sample can be decomposed into an $s$ and a $p$ component (the $s$ component is oscillating perpendicular to the plane of incidence and parallel to the sample surface, and the $p$ component is oscillating parallel to the plane of incidence). The amplitudes of the $s$ and $p$ components, after reflection and normalized to their initial value, are denoted by $\hat{r}_s$ and $\hat{r}_p$, respectively. The angle of incidence is chosen close to the Brewster angle of the sample to ensure a maximal difference in $\hat{r}_p$ and $\hat{r}_s$. The measured complex reflectance ratio, $\hat{\rho}$ (a complex quantity), is now given by the ratio of $\hat{r}_p$ over $\hat{r}_s$:

$$\hat{\rho} = \frac{\hat{r}_p}{\hat{r}_s} = \tan(\Psi)e^{i\Delta}.$$  \hspace{1cm} (2.1)

Thus, $\tan(\Psi)$ is the amplitude ratio upon reflection, and $\Delta$ is the phase shift (difference). (Note that the right hand side of the equation is simply another way to represent a complex number.)
2.3. Diagnostics

Since ellipsometry is measuring the ratio (or difference) of two values (rather than the absolute value of either), it is very robust, accurate, and reproducible. For instance, it is relatively insensitive to scattering and fluctuations, and requires no standard sample or reference beam. The downside is that information cannot be extracted from \( \Psi \) and \( \Delta \) in a straightforward way; they have to be fitted with an optical model specified for the material that is measured. After that information can be obtained about the thicknesses of the layers, the refractive indices and absorption coefficients. The optical constants can be expressed in terms of the complex dielectric function \( \varepsilon = \varepsilon_1 + i\varepsilon_2 \) [35].

The used ellipsometers were a J. A. Woollam Variable Angle Spectroscopic Ellipsometer with a VB-400 Control Module and an HS-190 monochromator (1.2 eV - 6.5 eV) for ex situ measurements, with angles varying from 65°-85°, and a J. A. Woollam Spectroscopic Ellipsometer M-2000F (1.2 eV - 5.0 eV) for in situ measurements under an angle of 70°. The model used for layer analysis is discussed below.

SE model

The data was analyzed using CompleteEASE SE software. The model used for extracting the data consisted of different model layers, that could be stacked onto each other in order to get the desired model that would represent the experimentally grown film best. In our case, this model contained a silicon substrate with a native oxide incorporated, and then a layer model for Si\(_{\text{n}}\) available in the software library, which uses a Tauc-Lorentz oscillator. In the Tauc-Lorentz model, the dielectric function is modeled from the product of a unique bandgap of amorphous materials (Tauc gap) and the Lorentz model [35].

The \( \varepsilon_2 \) of the Tauc-Lorentz model is expressed as

\[
\varepsilon_2 = \begin{cases} 
\frac{AE_{p,0}C(E_p-E_g)^2}{(E_p^2-(E_{p,0})^2)^2+C^2E_p^2} & (E_p > E_g) \\
0 & (E_p \leq E_g)
\end{cases}
\]  

Here, \( E_p \) is the photon energy of the used light. \( E_{p,0} \) is the \( \varepsilon_2 \) peak position and \( E_g \) the Tauc gap of the material. The \( A \) and \( C \) represent the amplitude and half width of the \( \varepsilon_2 \) peak, respectively. The expression for the real value \( \varepsilon_1 \) can be found in [35].

In situ SE has the advantage that it can be used to determine the growth per cycle (GPC) for several conditions on one substrate, while ex situ measurements require the samples to be exposed to atmosphere, which makes the samples unusable for further deposition steps while obtaining only a single data point. In situ SE therefore can easily be used to investigate the growth properties as a function of a certain parameter (e.g., plasma exposure time, pressure) and determining the GPC per condition by linear regression from a graph that contains the film thickness versus the number of ALD cycles. For the in situ GPC results reported, a thickness vs. number of cycles graph was analyzed consisting of 50 ALD cycles, where an SE measurement was made every 5 cycles.
2.3.2 Scanning Electron Microscopy

A scanning electron microscope (SEM) is a type of electron microscope that images a sample by scanning it with a high-energy beam of electrons in a raster scan pattern [36]. The electrons interact with the atoms that make up the sample producing signals that contain information about the sample’s surface topography, composition, and other properties such as electrical conductivity.

The types of signals produced by an SEM include secondary electrons, back-scattered electrons (BSE), characteristic X-rays, light, specimen current and transmitted electrons. Secondary electron detectors are common in all SEMs, but it is rare that a single machine would have detectors for all possible signals. The signals result from interactions of the electron beam with atoms at or near the surface of the sample. In the most common or standard detection mode, secondary electron imaging or SEI, the SEM can produce very high-resolution images of a sample surface, revealing details less than 5 nm in size. Due to the very narrow electron beam, SEM micrographs have a large depth of field yielding a characteristic three-dimensional appearance useful for understanding the surface structure of a sample. A wide range of magnifications is possible, from about 10 times to more than 500,000 times. Back-scattered electrons (BSE) are beam electrons that are reflected from the sample by elastic scattering. BSE are often used in analytical SEM along with the spectra made from the characteristic X-rays. Because the intensity of the BSE signal is strongly related to the atomic number (Z) of the specimen, BSE images can provide information about the distribution of different elements in the sample. Characteristic X-rays are emitted when the electron beam removes an inner shell electron from the sample, causing a higher energy electron to fill the shell and release energy. These characteristic X-rays can be used to identify the composition and measure the abundance of elements in the sample.

The used SEM is a JEOL High Resolution 7500FA SEM.
3 Adsorption and inhibition mechanisms

The influence of the reactant in growth of silicon nitride using BTBAS is not well understood. Other SiN$_x$ processes mostly use a NH$_3$ plasma instead of a N$_2$ plasma as reactant (see section 1.2.3), but initial experiments using BTBAS and a NH$_3$ plasma showed significantly reduced growth when compared to growth with a N$_2$ plasma. Several scenarios for possible growth inhibition have been given, but more information is needed to get a good understanding of the growth mechanism. Therefore, in this chapter, insights into the growth of silicon nitride layers using the BTBAS/N$_2$ ALD process are obtained by the interpretation of growth experiments using modified ALD cycles, where a silicon nitride surface is exposed to different reactants that inhibit film growth. These experiments are complemented by Density Functional Theory simulations, where the adsorption of a BTBAS precursor molecule to a silicon nitride surface is simulated for various surface compositions. The goal of these simulations and experiments is to find under which conditions growth is occurring, i.e., what surface compositions allow precursor adsorption to the surface, and what influence do different reactants have on the surface composition.

The layout of this chapter is as follows. First, growth experiments will show how the growth of silicon nitride films is reduced when the surface is exposed to the inhibitors. Furthermore, it is investigated if this inhibition can be recovered to the uninhibited state. Possible inhibition mechanisms explaining the growth experiment results will be suggested in section 3.2. These suggestions are subsequently validated with the DFT simulations in section 3.3, where also the principles of DFT and the simulation conditions will be discussed. Section 3.4 gives a summary of the results.

3.1 Growth experiments by multi-step cycles

To recap the observations on growth inhibition by several reactants, this section shows results of additional experiments where a silicon nitride surface is exposed to inhibitors and the effect on growth is studied. Furthermore, attempts to reduce this inhibition are studied.

3.1.1 Conditions

The BTBAS/N$_2$ cycle was extended with an extra step that involved exposing the surface to several inhibitors: NH$_3$ gas, NH$_3$ plasma, and H$_2$ plasma. Subsequently, an attempt to recover the surface to the uninhibited state was studied by two processes: thermal desorption under gas purging and plasma activation. Therefore, after the inhibitor exposure, the inhibited surface was again exposed to a N$_2$ gas to study thermal desorption, or a N$_2$ plasma to study the plasma activation.
This fourth step was varied in time to investigate time-dependency. Table 3.1 shows a summary of these multi-step cycles. In comparison, the normal silicon nitride ALD cycle consists only of the first two steps. Experiments were performed at 500 °C and 200 °C. These temperatures were chosen because 500 °C would give the material with properties closest to stoichiometric Si$_3$N$_4$ (in terms of stoichiometry and density), and would have almost no impurities [31], while at 200 °C the material would have more impurities while still being of decent quality. The growth per cycle (GPC) was measured as an indication of the level of inhibition.

Table 3.1: A summary of the multi-step cycles used for investigating the influence of H$_2$ and NH$_3$ in the BTBAS adsorption process. Step 1 and 2 contain the regular ALD cycle to get a surface with properties as close to a normal SiN$_x$ film as possible. In step 3 the surface is exposed to an inhibitor: NH$_3$ gas, NH$_3$ plasma, or H$_2$ plasma. In step 4 recovery of the original surface is investigated by either thermal desorption under a gas purge or by N$_2$ plasma exposure.

<table>
<thead>
<tr>
<th>Step</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Precursor</td>
<td>BTBAS 150 ms, 3 s reaction time</td>
</tr>
<tr>
<td>2</td>
<td>Reactant</td>
<td>N$_2$ plasma 10 s, 40 mTorr</td>
</tr>
<tr>
<td>3</td>
<td>Inhibitor</td>
<td>NH$_3$ gas 10 s, 40 mTorr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NH$_3$ plasma 10 s, 40 mTorr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>H$_2$ plasma 10 s, 40 mTorr</td>
</tr>
<tr>
<td>4</td>
<td>Recovery</td>
<td>N$_2$ gas 8 mTorr (maximum pumping)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N$_2$ plasma 40 mTorr</td>
</tr>
</tbody>
</table>

3.1.2 Results

Inhibition at 500 °C

First, the influence of inhibitors at a deposition temperature of 500 °C was investigated. Figure 3.1 shows the GPC as function of recovery time, either by thermal desorption using a N$_2$ gas purge or N$_2$ plasma exposure (step 4 in table 3.1). The dotted line represents the growth per cycle of a "standard" process that only involves BTBAS and a N$_2$ plasma. It can be seen that for each of the three inhibitors, the GPC is reduced when no attempt of recovery is made (recovery time 0 s). This inhibition was observed in earlier research, and it seems that it does not depend on the type of inhibitor: for all three inhibitors, the reduced GPC is ~40% of the standard GPC. Inhibition also occurred when using a N$_2$–H$_2$ mixture plasma, for 5-95% H$_2$/N$_2$ flow ratio. Causes for this inhibition will be suggested in section 3.2.

Using a N$_2$ gas to recover the original GPC seemed to have no effect on the NH$_3$ and H$_2$ plasma inhibition, and a very slight effect on the inhibition by NH$_3$ gas. Differences in recovery are not thoroughly studied. The main point here is that inhibited films were not affected by thermal desorption at 500 °C.

When trying to recover the growth rate using N$_2$ plasma exposure, we do see an effect. In the case of NH$_3$ gas inhibition, the growth is recovered to ~75%. For the case of NH$_3$ plasma and
3.1. Growth experiments by multi-step cycles

H₂ plasma, growth rates close to that of the standard cycle are obtained. The results show that using a N₂ plasma is important for reactivating the surface. An explanation could be that reactive plasma species (such as N* or N₂*) or ions can remove inhibiting species from the surface.

Figure 3.1 also shows that after a recovery time of 30 seconds, the GPC of the process that used a H₂ plasma is slightly higher than the GPC of the standard cycle. Although this difference is within the error of the GPC measurement, the increase may come from the fact that the standard cycle itself is also slightly inhibited. As will be shown in chapter 4, the GPC of the standard cycle at 500 °C can be increased when the plasma exposure time is prolonged. The mechanisms responsible for this effect will be discussed there.

![Figure 3.1: The growth per cycle at 500 °C as a function of recovery time (step 4 in table 3.1), for the inhibition with NH₃ gas, NH₃ plasma, and H₂ plasma. The solid lines represent recovery using a N₂ plasma. The dashed lines represent recovery using by thermal desorption using a N₂ gas purge. In the legend it is shown which inhibitor was used. In the graph the GPC of the standard cycle at 500 °C is also shown as a dotted line.](image)

**Inhibition at 200 °C**

To investigate the inhibition at 200 °C, experiments were performed with NH₃ gas and H₂ plasma as inhibitors. Of the latter, recovery was investigated only with a N₂ plasma. Experiments were not performed with N₂ gas in the recovery step after H₂ plasma inhibition, since at 500 °C there was no recovery observed, and therefore recovery at lower temperature was unlikely to occur.

At 200 °C similar trends are visible, as shown in figure 3.2. For comparison, both the growth rates for standard cycles at 200 °C and 500 °C are shown. Also at this temperature we see reduction in GPC when the films are exposed to the inhibitors. The inhibition by NH₃ gas seems less pronounced than the inhibition by H₂ plasma. In the case of NH₃ gas inhibition, we see no
recovery by thermal desorption. Recovery using a N\textsubscript{2} plasma also seems to have no observable effect on NH\textsubscript{3}-inhibited surfaces. This result is in contrast with the results obtained at 500 °C, where a recovery was observed. An explanation is that, for the case of inhibition by NH\textsubscript{3} gas, the N\textsubscript{2} plasma does reduce inhibition, but also reduces the growth. It was shown in earlier research that at 200 °C, still relatively much carbon is present in the film, as can be seen in figure 1.4. A prolonged plasma exposure reduces this carbon content, but with that the total growth of the film reduces as well. It might be that in this case, the recovery from inhibition and growth reduction from removal of carbon cancel out. Films grown at 500 °C do not have a high carbon content, so there the reduction in GPC due to carbon removal is far less significant.

The recovery curve of the H\textsubscript{2} plasma at 200 °C seems very similar to that at 500 °C. There is a strong inhibition at first, but by again exposing the surface to a N\textsubscript{2} plasma the growth can be recovered to values close to the standard GPC at 500 °C. The similarities of the recovery curve at these two temperatures might suggest that the properties of films grown using an additional H\textsubscript{2} plasma at 200 °C are also similar to properties of films grown at 500 °C. This can be a big advantage because, as will be shown further in the report, the film grown at 500 °C is of much higher quality. Being able to grow films with similar properties at lower temperatures, although using additional steps in the ALD cycle, is a big advantage for industrial viability.

Extra evidence for the hypothesis that using an additional H\textsubscript{2} plasma results in a higher-quality film, was found because an additional effect when using an extra hydrogen plasma was observed. We see an initial reduction of film thickness when starting a deposition that involves a H\textsubscript{2} plasma
3.1. Growth experiments by multi-step cycles

on a SiN$_x$ film that was grown without a H$_2$ plasma. Figure 3.3 shows the film thickness as a function of the number of cycles of the ALD process with an additional H$_2$ plasma. Before the start of this process, a film with a thickness of 10.5 nm was grown, using the standard recipe: at 200 °C, with a plasma exposure time of 10 s and a pressure of 40 mTorr, without using a H$_2$ plasma. The subsequent deposition involving an additional H$_2$ plasma did not have a recovery step (table 3.1) in the cycle, so the cycle consisted of 1. precursor exposure, 2. N$_2$ plasma exposure, 3. H$_2$ plasma exposure.

What can be seen from figure 3.3 is that the film thickness, upon start of the ALD process including the extra H$_2$ plasma, initially decreases before reaching a steady growth per cycle of 0.05 Å. This effect is not observed at a higher deposition temperature of 500 °C. One explanation for these observations can be that hydrogen from the additional H$_2$ plasma step can diffuse into the silicon nitride, removing the carbon that was incorporated during the growth of the original film. At higher temperatures the standard cycle produces films with an already low carbon content, so there the H-containing plasma may be less effective. These observations could again be an indication that using a H-containing plasma could increase film quality at lower temperatures. Preliminary X-ray Photoelectron Spectroscopy elemental analysis showed no traces of carbon left in such a film, but oxygen levels were elevated. No further analysis on films grown using an additional H-containing plasma was performed so far.

![Figure 3.3](image-url): Film thickness at 200 °C as a function of the number of cycles of ALD cycle involving an additional H$_2$ plasma. The multi-step cycle did not contain a recovery step. Before start of this process, a film of 10.5 nm was grown using the standard process at 200 °C.
3.2 Proposed inhibition mechanisms

Several inhibition scenarios can be proposed to explain the observations from section 3.1. Figure 3.4 shows a schematic of these scenarios. One proposed mechanism is the blocking of adsorption sites by NH$_3$. This scenario assumes a reactive surface to be N–H terminated, so that the precursor adsorbs to such a surface by replacing a hydrogen that resides on the surface. The hydrogen could be coming from precursor ligands that dissociate in the plasma half-cycle and adsorb to the surface. Note that silicon nitride ALD processes that use halide precursors have N–H terminated surfaces that are reactive towards adsorption [26, 27]. Ammonia and N–H groups contain dipole moments, so if the nitride surface contains N–H groups, hydrogen bonding could occur, causing physisorption of NH$_3$ to the surface. If the bond energy is higher than the thermal energy at 500 °C, this would explain the stable surfaces at that temperature. NH$_3$ is present in the gas and plasma of course, but not in pure H$_2$ plasma. Since inhibition was also shown for that case, the ammonia should come from elsewhere. It could be that the hydrogen plasma also abstracts some nitrogen from the surface, creating ammonia that can block surface sites.

![Figure 3.4: Schematic of different inhibition scenarios as proposed in [31], showing the main reactive surface groups in an ALD cycle after a pure N$_2$ plasma and the surface after a hydrogen-containing plasma. Scenario 1 assumes the formation and physisorption of NH$_3$, blocking adsorption sites. Scenario 2 assumes that the precursor can only adsorb on undercoordinated sites (dangling bonds), and that a H-containing plasma can passivate these sites.](image)

A second scenario was also proposed by King [29] for the case of silicon nitride growth with a silane precursor (scenario 2 in fig. 3.4). This scenario assumes that the precursor adsorbs to dangling bonds or undercoordinated sites at the surface. When such a surface is exposed to a NH$_3$ or N$_2$–H$_2$ plasma, the H radicals that are present in these plasmas can passivate these dangling bonds, thereby disabling precursor adsorption. This would explain the inhibition by plasma exposure, but not by ammonia gas, since a gas does not contain free radicals. The ammonia could adsorb dissociatively, leaving also a H-terminated surface. The nitrogen plasma then can remove hydrogen from the surface, creating new reactive sites and recovering the growth. A combination of scenario 1 and 2 could also be possible.
3.3 Simulations

In order to discern between the proposed inhibition mechanisms, calculations were performed where a silicon nitride model surface is exposed to a precursor molecule and different reactants, and subsequently the adsorption of these species is studied. Since the type of reactive surface on which BTBAS can adsorb would already discern between the scenarios in section 3.2, these simulations are an important part of the results. The calculations are based on Density Functional Theory. In this section first a short overview of this theory is given. Secondly, since silicon nitride can occur in different orientations and for the simulations a good model surface needs to be chosen, the crystal structures are discussed. In section 3.3.3 the choice for model surface is justified and simulation conditions are shown. The simulation results are shown in section 3.3.4. The DFT calculations were performed by dr. C.K. Ande; this chapter merely provides an interpretation of these simulations.

3.3.1 Density Functional Theory

Density functional theory is used to calculate the configuration of a collection of atoms with the lowest energy. This configuration can be, for instance, the geometry of an isolated molecule or the crystal structure of a mineral. In order to find this configuration, one needs to know what the energy of a specific configuration is and how it changes when the atoms are moved around. In DFT the electronic structure of the system is considered, since electronic structure and energy are closely related. A helpful observation regarding atoms is that the nuclei of the atoms are much heavier than the electrons, and therefore the electrons respond more rapidly to changes in the environment. Therefore the problem of finding configurations can be split into two separate problems for the nuclei and the electrons (Born-Oppenheimer approximation), where the nuclei are fixed and generate a potential \( V \) for the electrons. For a system of \( N \) electrons, their lowest energy (ground state) can be calculated with the time-independent Schrödinger equation:

\[
\left[ -\frac{\hbar^2}{2m} \nabla_i^2 + \sum_{i=1}^{N} V(r_i) + \sum_{i=1}^{N} \sum_{j<i} U(r_i, r_j) \right] \psi = E \psi.
\]  

Here, \( \hbar \) is the reduced Planck constant, \( m \) is the electron mass, \( V \) the potential landscape created by the nuclei, \( U \) the electron-electron interaction, \( E \) the total energy of the system, and \( \psi \) the many-particle wavefunction that describes the position of all the electrons in space. Because this equation couples all the electron interactions to one another, solving this equation for large \( N \) takes a very long time.

Density functional theory works around this problem by not looking at the energy and position of every electron separately, but by looking at the electron density as a function of position in space. This density can be written in terms of the individual electron wave functions \( \psi_i \) as

\[
n(r) = \sum_i \psi_i^*(r) \psi_i(r).
\]
Using this quantity, the energy of a system can now be given by the Kohn-Sham equation [37]:

\[
\left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) + V_H(r) + V_{XC}(r)\right] \psi_i(r) = \varepsilon_i \psi_i(r).
\] (3.3)

This equation looks a lot like the Schrödinger equation (eq. (3.1)), except that instead for solving \( N \) coupled equations, one only has to solve \( N \) equations that involve a single electron, greatly reducing computing costs. Another difference with eq. (3.1) is the appearance of \( V_H \) and \( V_{XC} \). \( V_H \) is called the Hartree potential and is defined by

\[
V_H = e^2 \int \frac{n(r')}{|r - r'|} d^3 r'.
\] (3.4)

This potential describes the Coulomb repulsion between the electron being considered in one of the Kohn-Sham equations and the total electron density defined by all electrons in the problem. The Hartree potential includes a so-called self-interaction contribution because the electron that is described is also part of the electron density, so a part of \( V_H \) involves a Coulomb interaction between the electron and itself. This is not a physical effect, and therefore a correction should be made. This is done, amongst other effects, in the exchange-correlation potential \( V_{XC} \), which defines exchange and correlation contributions to the single-electron equations. There are lots of different types of exchange-correlation functionals, some created \( ab \ initio \), some empirical. Depending on the type of problem where DFT calculations have to be done, a suitable functional is used. For more information about exchange-correlation functionals, the reader is referred to [37, 38].

Now we can find the ground-state energy of the configuration, by finding an appropriate electron density. The ground-state energy of a specific configuration of atoms can be found by solving equation (3.3) iteratively, so guessing an initial \( n(r) \), from that defining \( V \), \( V_H \), and \( V_{XC} \), solving the KS equations, calculating the electron density again from the KS single electron wave functions with eq. (3.2), and checking the result with the initial \( n(r) \). If they are the same, then this is the electron density belonging to this specific configuration of atoms one has started with, and the ground-state energy of this configuration is found. If not, the initial guess for \( n(r) \) is changed until this criterion is met.

When the ground-state energy of a configuration is found, one can now change this configuration, compute the new potentials, and calculate again the ground-state energy using the method as explained above. In this way the lowest-energy configuration of atoms (which is what one would like to find) can be found using an optimization algorithm. A flowchart for this process is shown in figure 3.5 [38].

### 3.3.2 Crystal structure and properties

In order to perform simulations that can be translated to processes that happen in reality, one first needs to know what model would be useful to perform simulations with. This section will discuss the various structures that silicon nitride can occur in, and compare the properties of these structures. In the next section, the choice for the model surface is justified.
3.3. Simulations

Silicon nitride can occur in both crystalline and amorphous structures. Stoichiometric silicon nitride (Si$_3$N$_4$) occurs in three crystallographic structures, designated as $\alpha$-, $\beta$-, and $\gamma$-phases [39]. The $\alpha$ and $\beta$ phases are the most common forms of Si$_3$N$_4$, and can be produced under normal pressure conditions ($p \leq 1$ atm), but since the $\alpha$-state is chemically unstable compared with the $\beta$-phase, $\beta$-Si$_3$N$_4$ is the most common form [40].

The $\alpha$- and $\beta$-Si$_3$N$_4$ have trigonal and hexagonal structures, respectively, which are built up by corner-sharing SiN$_4$ tetrahedra, as shown in figure 3.6. They can be regarded as layers of silicon and nitrogen atoms in the sequence ABAB... in $\beta$-Si$_3$N$_4$ or ABCDABCD... in $\alpha$-Si$_3$N$_4$. The AB layer is the same in the $\alpha$ and $\beta$ phases, and the CD layer in the $\alpha$ phase is related to AB by a c-glide plane. The SiN$_4$ tetrahedra in $\beta$-Si$_3$N$_4$ are interconnected in such a way that tunnels are formed, running parallel with the c axis of the unit cell. Due to the c-glide plane that relates AB to CD, the $\alpha$ structure contains cavities instead of tunnels.

Films grown with ALD or CVD will often be amorphous, as they are in this work. Amorphous silicon nitride (SiN$_x$) can have a wide range of silicon/nitrogen compositions, as already shown in section 1.2.3. It can retain local Si–N bonding but lacks long-term translational order. The density of amorphous SiN$_x$ films is generally slightly less than the density of $\beta$-Si$_3$N$_4$, but due to the disorder in crystal structure do not contain the tunnels, making SiN$_x$ more suitable as moisture barrier, while still having chemical inertness [41].
3.3.3 Methodology

As explained above, the experimentally grown silicon nitride films are amorphous, and will have varying stoichiometry. Because this reduces simulation reproducibility and elongates simulation time, the calculations consider mostly a bare stoichiometric $\beta$-Si$_3$N$_4$ surface as model system, with the (0001) direction parallel to the c-direction of the unit cell. The bond lengths in $\beta$-Si$_3$N$_4$ are very similar to that of amorphous silicon nitride, making this surface a good approximation to the experimental surface. An extra advantage is that the $\beta$-state has less atoms in its primitive cell, which speeds up the simulations.

The simulations made use of a cell with periodic boundaries, a so-called slab. The size of the slab is 1 unit cell in the a- and b- directions, and 2 unit cells in the c- direction to mimic bulk effects. A schematic of the slab is depicted in figure 3.7. Since the c-boundaries are also periodic, extra space in that direction has to be incorporated to exclude non-physical effects. The (0001) unit cell has 3 undercoordinated Si atoms and 3 undercoordinated N atoms per side. The undercoordination would make it assumable that this surface is more reactive towards adsorption than a fully coordinated surface, since having fully coordinated/filled shells is in general energetically more favorable.

Adsorption is simulated on one (0001) side of the surface. To exclude effects from dangling bonds on the other side of the surface, all the undercoordinated sites from that side are passivated with H, as visible in figure 3.7. The a- and b- sides of the slab do not need to be passivated, since these sides are periodic.

In order to simulate the interactions of the precursor and different plasmas with the Si$_3$N$_4$ surface, the simulations involve adsorption of H, N, NH$_3$, and BTBAS. H atoms are the main created species in a H$_2$ plasma. N atoms are present in a N$_2$ plasma, and are typically believed to be the main reactive species [29]. NH$_3$ is likely to be present in a NH$_3$ or N$_2$−H$_2$ plasma. H, N, and NH$_3$ adsorption is simulated on a bare, undercoordinated surface, while BTBAS adsorption is simulated on both a bare surface and a surface that is totally passivated with H. The latter is done because we would like to find out whether a H-terminated surface is reactive towards
3.3. Simulations

adsorption of BTBAS, since this would already discern between the two proposed scenarios in section 3.2. It is very likely that an undercoordinated silicon nitride surface is H-terminated after exposing it to a H₂ plasma [42].

The simulations are performed using the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [43] and consider a temperature of 0 K, so no atomic vibrations are taken into account. After finding a stable geometry, Molecular Dynamics simulations can be used to check stability at higher temperatures. When this is done, it is mentioned in the results. The DFT software that was used for simulations is VASP\(^1\), the software used for visualising the data is VESTA\(^2\).

![Figure 3.7: The slab model used for simulations. Blue spheres represent Si, grey spheres represent N and pink spheres represent H. Unit cell vectors are also shown.](image)

3.3.4 Results

H adsorption

To study the passivation of silicon nitride by a hydrogen plasma, the $\beta$-$\text{Si}_3\text{N}_4$ was first exposed to H atoms. The calculations showed that a fully passivated surface is energetically more favorable than an undercoordinated surface, with an average gained energy per adsorbed hydrogen of

\(^1\)https://www.vasp.at/
\(^2\)http://jp-minerals.org/VESTA/en/
−0.78 eV. These results confirm the assumption that an undercoordinated surface is more reactive towards adsorption, and show that a silicon nitride surface can be passivated by exposing it to a H\textsubscript{2} plasma.

**N adsorption**

To simulate a "normal" BTBAS/N\textsubscript{2} ALD process, the surface was exposed to N atoms which are likely to be present in the N\textsubscript{2} plasma. During these simulations, the surface was allowed to relax (positions of surface atoms were not fixed). Afterwards, molecular dynamics simulations were performed to check stability at 300 K. The resulting surface is shown in figure 3.8a. The three N atoms were adsorbed and formed interlinking N chains. Adsorption of N leaves undercoordinated sites at the surface. Such a surface as shown in figure 3.8a probably resembles an actual surface during silicon nitride growth with BTBAS and a N\textsubscript{2} plasma.

**NH\textsubscript{3} adsorption**

For the case of NH\textsubscript{3} adsorption, three molecules were placed near the surface and again the geometry with lowest energy was calculated, with an MD stability check at 300 K afterwards. This geometry is found in figure 3.8b. The simulation shows that all NH\textsubscript{3} molecules have been dissociated and chemisorbed to the surface. The ammonia dissociated in NH\textsubscript{2} and H. The NH\textsubscript{2} group attached to an undercoordinated silicon; the H attached to an undercoordinated N. Adsorption of NH\textsubscript{3} eliminates all dangling bonds of the β-Si\textsubscript{3}N\textsubscript{4} surface, while this is not the case for N adsorption.

![Figure 3.8](image-url)

**Figure 3.8:** (a) Calculated surface after N adsorption on a bare surface, where the bridging N structures are visible. The bridging N atoms are undercoordinated. (b) Calculated surface after NH\textsubscript{3} adsorption on a bare surface. Here, all the atoms are fully coordinated.
3.3. Simulations

BTBAS adsorption

Another simulation involves the adsorption of a BTBAS molecule onto a bare Si$_3$N$_4$ surface. The result is shown in figure 3.9a. It is visible that the precursor can adsorb to an undercoordinated surface. Time-resolved analysis showed that initially a N atom from the molecule binds to a silicon atom residing at the surface. Subsequently, the silicon binds to a neighboring surface N atom, donating a H atom to another Si atom. Finally, the second N atom from the molecule binds to a second surface Si atom. This creates a stable geometry.

When BTBAS is placed near a passivated surface, it does not adsorb, as shown in figure 3.9b. To compare the adsorption stability: the adsorption energy of BTBAS to a bare surface is $-1.9$ eV, while to a passivated surface it is only $-0.05$ eV. The very low binding energy of BTBAS to a passivated surface means that it can be easily removed by thermal desorption or purging. These results are consistent with scenario 2 from section 3.2 as also suggested by King [29].

![Figure 3.9: (a) BTBAS adsorption to an undercoordinated Si$_3$N$_4$ surface, with an adsorption energy of $-1.9$ eV. (b) BTBAS adsorption to a passivated Si$_3$N$_4$ surface, with an adsorption energy of $-0.05$ eV.](image)

From earlier research, Quadrupole Mass Spectrometry data showed that in the precursor adsorption process, when binding to a surface site, one N-t-butyl group is released [31]. The simulation results do not yet contradict these observations, since the simulations are done at 0 K. Preliminary MD simulations at 1000 K show the removal of a butyl ligand [44], which is fairly consistent with the QMS data, with the exception of the additional N attached to the butyl. Since the actual surface during growth will look more the one depicted in 3.8a, which has additional N adsorbed to the surface with respect to the bare Si$_3$N$_4$ surface, those N atoms might be responsible for the extra N attached to the removed N-t-butyl ligand.
3.4 Summary

In this chapter we have studied the adsorption mechanism of BTBAS to a silicon nitride surface and the role of NH$_3$ gas, NH$_3$ plasma and H$_2$ plasma in this mechanism. Furthermore, DFT simulations were performed where the adsorption process of atomic N, NH$_3$, and BTBAS to a silicon nitride surface was visualized.

The growth experiments showed that indeed NH$_3$ and H-containing plasmas do inhibit growth: when the BTBAS/N$_2$ cycle was extended with an additional step where the surface was exposed to NH$_3$ gas, NH$_3$ plasma, or H$_2$ plasma, the growth-per-cycle was significantly lower. This inhibition could (partially) be made undone by exposing a passivated surface to a N$_2$ plasma. A N$_2$ plasma thus is important in creating reactive sites for precursor adsorption. This can either happen by hydrogen removal from the surface by N radicals, or by ion impact which might break surface bonds.

There are several indications that exposing a surface to H-containing plasmas might have a beneficial effect on the quality of films deposited at lower temperatures. Hydrogen might be more favorable in removing impurities from the film than nitrogen.

The DFT simulations showed that the inhibition is occurring because the BTBAS precursor needs undercoordinated sites to adsorb to. When the surface is terminated by hydrogen, BTBAS only has an adsorption energy of $-0.05$ eV, which is very low compared to the adsorption energy of $-1.9$ eV when BTBAS adsorbs to an undercoordinated surface. The simulations also showed that a surface can become fully H-terminated when the surface is exposed to NH$_3$ or a H$_2$ plasma. Exposing a surface to N$_2$ plasma leaves dangling bonds. The difference in surface termination between a H-containing plasma and N$_2$ plasma explains the inhibition observed with a H-containing plasma or NH$_3$ gas. The DFT simulations give theoretical evidence that NH$_3$ and H-containing plasmas can inhibit growth, while a N$_2$ plasma would still enable precursor adsorption by leaving or even creating dangling bonds on the surface.
4 Mechanisms regarding wet etch resistance and conformality

To identify the mechanisms that dictate the properties that are more essential for spacer applications, namely the inertness and barrier properties of the films, and conformality (section 1.1), several experiments were conducted. As a measurement for the chemical inertness and barrier properties, often the wet etch rate of silicon nitride in a HF solution is used. Therefore, etch rates of SiNx films were determined. However, extracting the etch rates is not a straightforward task. In order to achieve satisfying quantification of trends, accurate etch rate data of the involved materials are needed.

Besides having a low etch rate, is having a good conformality essential for the application of silicon nitride as spacer in high aspect ratio structures. Therefore, silicon nitride films were deposited in structures with a high aspect ratio, and the difference in thickness between horizontal and vertical surfaces was determined. Furthermore, since the quality of the material has to be uniform throughout the structure, the etch uniformity of the films in these structures was quantitatively determined.

This chapter will first discuss the pitfalls in determining the wet etch rate of silicon nitride, and provides a model for predicting the etch rates as a function of etchant concentration. Secondly, it will discuss the procedure used for determining the etch rates and show obtained wet etch rates of films that were deposited using various parameters. In section 4.3 the underlying mechanisms of these etch rates will be discussed, and it is shown that these mechanisms have important consequences for choosing optimal parameters when depositing high quality silicon nitride. In section 4.4 the conformality is shown for two optimized conditions.

4.1 Wet etching of silicon nitride

Many methods for determining the etch rate of materials exist. In situ etch rate experiments, like using a quartz-crystal balance, acoustic microscopy and controlled sample immersion into an etchant, determine the amount of etched material in real time. Hence, the etch rate can directly be measured as a function of time. Ex situ measurements on the other hand do not show what is occurring during the etch process but only give an average picture. These type of measurements, performed for a specific etch time, are easier or cheaper to do since only a method of measuring thickness is needed. Not addressing a time-varying etch rate in a proper way can result in either an over- or underestimation of the determined etch rate.

Van Barel et al. [45, 46] presented a detailed analysis of the time evolution of the etch rate applicable to most etching methods. In this section a small part of their work will be shown.
to get an idea of the time evolution of a general wet etch process. The onset, steady-state and offset periods in a typical etch rate experiment are identified. These periods are associated with the initial increase of the etch rate after immersion into the etchant (incubation), the constant etch rate regime until the removal from the etchant (steady-state or propagation), and the final gradual etch rate reduction due to the removal of the etchant from the sample surface (rinsing or termination).

The most commonly used etchants for silicon nitride are hot phosphoric acid [47] and diluted hydrofluoric acid [48]. The hot phosphoric acid has problems such as particle contamination of the wafers and low etch rates. An additional technical problem is the high operational temperature, usually between 140 and 180 °C, which raises safety issues. Therefore it is more favorable to perform etch experiments with hydrofluoric acid (even though HF should also be handled with great care). Some models describing the etch process of silicon nitride in HF based solutions exist [49, 50]. These models are used to predict etch rates as a function of etchant concentration. A similar model is used for the etching of silicon dioxide to understand the etch selectivity between silicon nitride and silicon oxide [51].

4.1.1 Steady-state versus apparent etch rate

A wet, dry or vapor etch process can be divided into three time intervals as depicted in figure 4.1. The first domain is called the incubation period. During this period the etching starts from an etch rate of 0 nm s\(^{-1}\) and increases until a steady-state etch rate is reached. The time to reach this constant etch rate is defined as the incubation time \(T_i\). During this incubation time a certain amount of material has been etched which is defined as the incubation distance \(D_i\). In the subsequent time domain the etching proceeds at a constant etch rate \(R_c\) until the rinsing period is started at time \(t_m\). The measurement time \(t_m\) is the moment at which the time monitoring of the etch process is actually stopped and the rinsing process starts. It is practically very difficult to determine the time \(t_e > t_m\) at which the etching itself actually stops when performing \textit{ex situ} experiments. In the rinsing process the etch rate decreases to 0 nm s\(^{-1}\); the time it takes to reach this etch rate and the distance etched during that time are called the rinsing time \(T_r\) and the rinsing distance \(D_r\), respectively.

A common pitfall is to determine the etch rate directly from the ratio of the total etch distance \(d_e\) and the monitored etch time \(t_m\). When using these numbers, the calculated etch rate changes to an apparent etch rate \(R_m \neq R_c\) (fig. 4.1). The difference between the apparent etch rate and the actual steady-state etch rate depends on the relative importance of the incubation and rinsing period, which in turn also depends on the measurement time. Using \(R_m\) instead of \(R_c\) can lead to large deviations in etch data.

4.1.2 Extraction of the steady-state etch rate

A single \textit{ex situ} experiment involving an incubation and a rinse period is not sufficient to determine the steady-state etch rate. However, repeating the same etch rate experiment for consecutive measurement times \(t_m\) allows the calculation of the steady-state etch rate through linear regression (see figure 4.2).
4.1. Wet etching of silicon nitride

Figure 4.1: Theoretical representation of a typical etch process consisting of an incubation period, a steady-state or constant etch rate period and a rinsing period. The direct calculation of the etch rate out of the measured etched distance $d_e$ and the measured time $t_m$ (leading to the apparent etch rate $R_m$) can lead to a large deviation from the steady-state etch rate $R_c$. Adapted from ref. [45].

Figure 4.2: Graphical representation of four experiments that are stopped at different measurement times $t_{mi}$ ($i = 1, 2, 3, 4$). These etch experiments assume the same incubation time and rinse period. Under these conditions the steady-state etch rate can be extracted from the slope of the weighted linear fit through the four data points representing the etch distance $d_e$ versus measured etch time $t_m$. Adapted from ref. [45].
In order to do this correctly, three conditions need to be met. At first, all experiments need to fulfill the condition of equal incubation and rinsing period. This can be done by placing all the test material into an etch recipient under the same timing. Also the rinsing has to be similar for the different samples.

The second condition is that the etch experiment must continue at least until the incubation period is finalized. If this condition is not met, no information about the steady-state etch rate can be extracted.

Likewise, it is required that the rinsing starts during the steady-state etching period. This means that the experiment must not be self-limiting and continues until the samples are removed form the etchant.

4.1.3 Etching mechanism in HF solutions

Knotter and Denteneer [50] proposed a reaction mechanism for the etching of silicon nitride in hydrofluoric acid. The etch rate \( R \) of silicon nitride can be described with a semi-empirical equation depending on the concentration of active species in the hydrofluoric acid solution, such as \( F^- \), \( HF_2^- \), HF and \( H_2F_2^- \), together with their "rate constants" (weight factors) \( k_i \) that depend on the type of silicon nitride that is being etched:

\[
R_{SiN_x} = \frac{k_0[HF_2^-] + (k_1[HF] + k_2[F^-]) K_3[H^+]}{1 + K_3[H^+]} \quad \text{(Å s}^{-1}). \tag{4.1}
\]

In this equation, \( K_3 \) is the equilibrium constant of the reaction that occurs when a surface group \( SiNH_2 \) is protonized by \( H^+ \): \( SiNH_2 + H^+ \leftrightarrow SiNH_3^+ \). The \( NH_3 \) group then can release from the silicon, leaving an open space on the coordination sphere of Si with which HF, \( F^- \) or \( HF_2^- \) can react, forming an Si-F bond. The electronegativity of the fluoride reduces the bond strength of the remaining Si-N bonds, allowing other \( F^- \) ions to replace the remaining N and thus creating volatile \( SiF_4 \).

We also would like to be able to compare the performance of the deposited silicon nitride with a different type of material, silicon oxide for instance. Knotter also described a reaction mechanism and semi-empirical model for the etching of silicon dioxide [51]. Here, the etch rate also depends on the different concentrations of active species, but since the mechanism differs from that from silicon nitride, the model that describes the etch rate \( R \) is also different:

\[
R_{SiO_2} = \left( k_0[H_2F_2] + k_1[HF_2^-] \right) \frac{K_3[H^+]}{1 + K_3[H^+] + 1/(K_4[H^+] + \alpha)} + k_2[HF_2^-] \frac{K_4[H^+]}{1 + \alpha + \alpha K_3^2[H^+] + K_4[H^+] + K_3 K_4[H^+]} \quad \text{(Å s}^{-1}). \tag{4.2}
\]

Again, \( k_i \) are the rate constants that depend on the type of material that is being etched. The capital \( K_i \) are equilibrium constants of the surface reactions. Since the full understanding of the etch mechanism of silicon oxide is outside the scope of this project, the reader is referred to [51]...
4.1. Wet etching of silicon nitride

for more information. What we will do in this section is use these equations to predict the etch rates and selectivity with changing etchant concentrations.

The concentration of each species in an etch solution can be calculated from the equilibrium solution reactions together with the conservation formulas of mass and charge. These equations are stated below [50]:

\[
\begin{align*}
[H^+] + [F^-] & \leftrightarrow [HF] \\
\frac{[H^+][F^-]}{[HF]} & = 7.2 \times 10^{-4} \text{mol/L;} \\
[HF_2^-] & \leftrightarrow [HF] + [F^-] \\
\frac{[HF]][F^-]}{[HF_2^-]} & = 3.963 \text{L/mol;} \\
[H_2F_2] & \leftrightarrow 2[HF] \\
\frac{[H_2F_2]}{[HF]^2} & = 2.7 \text{L/mol;} \\
[H^+] + [NH_4^+] & = [F^-] + [HF_2^-]; \\
C_F & = [F^-] + [HF] + 2[HF_2^-] + 2[H_2F_2].
\end{align*}
\]

In equation (4.7), the constant \(C_F\) stands for the analytical concentration of fluoride species in the solution. This analytical concentration can be tuned by the number of species that is put in the solution when creating it. The \(NH_4^+\) term in eq. (4.6) emerges when the solution is buffered using \(NH_4F\). Buffering keeps the pH of the solution constant, and reduces fluctuations in etch rate due to changing concentrations. When \(NH_4F\) is added, it is assumed that it dissociates into \(NH_4^+\) and \(F^-\) completely.

Using equations (4.1) and (4.2), rate constants were defined for various types of materials. Table 4.1 shows these constants for LPCVD SiN, PECVD SiN:H and SiO\(_2\). To get a feeling of the typical etch rates for such materials, the etch rates were calculated as a function of the percentage HF in the etch solution by plugging the above equations into Matlab software. The Matlab script used for calculations can be found in appendix B. The trends are shown in figure 4.3a.

Table 4.1: Calculated weights for LPCVD SiN, PECVD SiN:H, and SiO\(_2\).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SiN</th>
<th>SiN:H</th>
<th>SiO(_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_0) (Å L mol(^{-1}) s(^{-1}))</td>
<td>0.29 ± 0.02</td>
<td>10 ± 1</td>
<td>7 ± 1</td>
</tr>
<tr>
<td>(k_1) (Å L mol(^{-1}) s(^{-1}))</td>
<td>0.21 ± 0.01</td>
<td>9.1 ± 0.5</td>
<td>18 000 ± 4000</td>
</tr>
<tr>
<td>(k_2) (Å L mol(^{-1}) s(^{-1}))</td>
<td>28 ± 7</td>
<td>820 ± 30</td>
<td>12.6 ± 0.5</td>
</tr>
<tr>
<td>(K_3) (L mol(^{-1}))</td>
<td>25 ± 7</td>
<td>25 ± 7</td>
<td>0.28 ± 0.06</td>
</tr>
<tr>
<td>(K_4) (L mol(^{-1}))</td>
<td>na</td>
<td>na</td>
<td>13 000 ± 2000</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>na</td>
<td>na</td>
<td>0.85 ± 0.06</td>
</tr>
<tr>
<td>Reference</td>
<td>[50]</td>
<td>[50]</td>
<td>[51]</td>
</tr>
</tbody>
</table>

We can see from figure 4.3a that the etch rate goes up with increasing HF concentration. Furthermore, it varies greatly per type of material. When the silicon nitride is hydrogenated, the
Chapter 4. Mechanisms regarding wet etch resistance and conformality

Figure 4.3: (a) Calculated etch rates as a function of mass percentage HF in water for LPCVD SiN, PECVD SiN:H and SiO$_2$ using equation (4.1) and weights as shown in table 4.1. (b) Calculated etch rates for SiN as a function of the percentage of fluorine coming from the NH$_4$F. $C_F=0.25$ M is equivalent to the concentration of 0.5 mass% HF; $C_F=19.13$ M is equivalent to the concentration of 40 mass% HF.

The etch rate is a factor of 30 higher than for normal SiN, but shows the same trend. This could mean that the amount of hydrogen present in the sample is essential for the etch resistance properties. This result was also shown experimentally in several studies [28, 52–55]. The etch rate of silicon dioxide varies greatly with concentration, as also visible in figure 4.3a, from an etch rate of 0.6 nm/min at 0.5% HF to 100 nm/min at 40% HF. When comparing this increase with SiN, the selectivity of the etching between these two materials can vary from 1:2 to 1:50, which can have great implications on the type of application the etching is used for. Several studies confirmed this selectivity [50, 56–59].

Furthermore, to see how the buffer influences the etch rate, etch rates were calculated for SiN as function of the percentage of fluorine coming from the buffer. This is shown in figure 4.3b for two analytical concentrations chosen to match the concentrations of two commonly used etchants: 0.5 mass% HF in water and 40 mass% HF in water. From figure 4.3b we see that for various fluorine concentrations the amount of fluoride coming from NH$_4$F influences the etch rate significantly. For low analytical concentrations ($C_F=0.25$ M, equivalent to a 0.5% HF solution), the etch rate decreases when NH$_4$F is added. This decrease is because at low concentrations the dominant term in equation (4.1) is the $[F^-]$ term, which is influenced by $[H^+]$. $[H^+]$ drops faster than $[F^-]$ rises, so the total etch rate drops. At higher analytical concentrations ($C_F=19.1$ M, equivalent to 40% HF), the $[HF_2^-]$ term is dominant in the etching process, and this concentration rises initially with increasing percentage of $F^-$ coming from NH$_4$F. Figure 4.3 shows that a small variation in concentration can give large deviations in etch rates. Therefore, if we want to compare the etch rates of multiple samples, the etching concentrations should be as much the same as possible.
4.2 Etch experiments on planar wafers

To investigate the etch rate of silicon nitride films as a function of deposition parameters as temperature, pressure and plasma exposure time, etch tests were performed on planar wafers with a SiN_x film grown on top. This section will discuss the etching procedure and the results from the etch experiments.

4.2.1 Etching procedure and conditions

For determining the wet etch rates of various silicon nitride films, we used two types of HF solutions: a buffered HF (BHF) solution (using NH_4F) and a non-buffered 0.5% (0.25 mol/L) HF solution, from now on called diluted HF or dHF. For the buffered etch a stock solution was used of a 7:1 volume ratio of 40% NH_4F in water (18.0 mol/L) and 49% HF in water (48.01 mol/L). Of these solutions, the concentrations of various species were determined using equilibrium equations and mass and charge balance equations as shown in section 4.1.3. These concentrations are shown in Table 4.2. If we plug in these concentrations in the formula for the etch rate (4.1) and (4.2) with the material constants shown in Table 4.1, we find an etch rate of 10.4 nm/min and 214 nm/min for LPCVD SiN and SiO_2, respectively, for the BHF, and 0.34 nm/min and 0.6 nm/min for the dHF.

Table 4.2: Concentrations calculated using equations (4.3)-(4.7) for both the buffered and the standard 0.5% solution.

<table>
<thead>
<tr>
<th>Species</th>
<th>BHF</th>
<th>dHF</th>
</tr>
</thead>
<tbody>
<tr>
<td>[H^+]</td>
<td>1.04 × 10^{-5}</td>
<td>0.0127</td>
</tr>
<tr>
<td>[F^-]</td>
<td>10.08</td>
<td>0.008101</td>
</tr>
<tr>
<td>[HF_2]</td>
<td>5.742</td>
<td>0.00458</td>
</tr>
<tr>
<td>[HF]</td>
<td>0.1448</td>
<td>0.1427</td>
</tr>
<tr>
<td>[H_2F_2]</td>
<td>0.0566</td>
<td>0.055</td>
</tr>
<tr>
<td>C_F</td>
<td>21.75</td>
<td>0.25</td>
</tr>
</tbody>
</table>

The procedure for etching the films was as follows. First the samples were cleaned using nitrogen gas. Then the samples were placed in a holder that could process multiple samples at once so no influence of different incubation or rinse times had to be taken into account. Eventually the holder was immersed in the HF solution for a set time and subsequently rinsed in DI water with resistance R > 10 MΩ cm. After rinsing, the samples were blown dry with nitrogen gas and then taken for characterization by SE or SEM. For every etch rate shown in the sections below, the steady-state etch rate was determined from linear regression of at least 3 thickness measurements per condition.
Chapter 4. Mechanisms regarding wet etch resistance and conformality

4.2.2 Results

Following the etch procedure from section 4.2.1, wet etch rates in 0.5 % HF and 7:1 buffered HF solutions were determined of SiN\textsubscript{x} layers deposited at various pressures, temperatures and plasma exposure times. Figure 4.4 shows the thickness of the deposited layers and the etch distance defined in section 4.1.1 as a function of etch time for samples deposited at 40 mTorr, 10 s plasma exposure and substrate table temperatures as shown in the figure. Thermally grown silicon dioxide (SiO\textsubscript{2}) is added as a reference.

![Figure 4.4](image)

Figure 4.4: (a) Layer thickness as a function of etch time in 0.5% HF for samples deposited at 40 mTorr, 10 s plasma exposure and different temperatures. (b) Etch distance as a function of etch time for samples as described in (a). Thermally grown SiO\textsubscript{2} is added as a reference.

From figure 4.4 it can be seen that within the first 15 seconds of etching the etch rate is higher than after this time. This might be attributable to lower quality deposition of the top layer of the samples, or ingress of polluting species since the samples had been exposed to atmosphere for longer than a month after synthesis. According to figure 4.4a this part is about 1 nm thick for the SiN\textsubscript{x} layers and about 3 nm thick for the SiO\textsubscript{2} sample. After 15 seconds the layer thickness decrease seems linear, indicating a constant etch rate. Figure 4.4b shows that the etch rate of the sample deposited at 200 °C is fairly high compared to the other temperatures, and that as the temperature increases the etch rate drops.

To get a better comparison of those etch rates, they are shown in figure 4.5, together with etch rates of films deposited with other parameters. The etch rates vary from 70 nm/min for samples deposited at 200 °C, 40 mTorr and 7.5 s plasma exposure, to etch rates below 0.5 nm/min for samples deposited at 400 °C and higher. Table 4.3 shows the etch rates in nm/min for various films, together with the etch rate for thermally grown SiO\textsubscript{2}. Under these etching conditions, the samples deposited at \( T \geq 400 \) °C perform better than thermally grown SiO\textsubscript{2}, which has an etch rate of 0.56 nm/min.

When comparing these found values with the calculated values from the etching model in section 4.2.1, 0.34 nm/min for LPCVD SiN and 0.6 nm/min for SiO\textsubscript{2}, the SiO\textsubscript{2} value is consistent, but the etch rate of our grown SiN\textsubscript{x} at higher temperatures seems to be less. This is not strange,
4.2. Etch experiments on planar wafers

since the deposition technique we used differs from the LPCVD. However, the lower etch rate is an indication that the quality of the silicon nitride deposited at ≥400 °C is already better than the existing LPCVD technique that was used for depositing the SiN.

The relatively large errors for samples deposited at 400 °C and 500 °C are probably arising from the fact that the changes in thickness used for calculating the etch rate were so small that they were susceptible to small concentration changes and errors in dip times. The absolute thickness errors are in the same order of magnitude of each other.

Figure 4.5: Comparison of the etch rates in dHF for samples deposited at various plasma exposure times, pressures and temperatures. The conditions under which a film was deposited are indicated. The etch rate of SiO$_2$ under the same conditions is added for comparison.

From figure 4.5 and table 4.3 three trends can be distinguished. As the plasma exposure time increases, the etch rate drops significantly. Furthermore, as the plasma pressure is lowered, there also is a drop in etch rate. Another important trend is the decrease in etch rate with increasing temperature.

The lowering of the etch rates with increasing temperature, increasing plasma exposure time and lower pressures can be attributed to the change in composition of the film with these parameters. In our previous research on this process it was shown that, following the same trend as the etch rates with increasing temperature, pressure, and plasma exposure time, the deposited material shifts towards more stoichiometric Si$_3$N$_4$ and more dense films with less H and C impurities, as shown in figure 1.4 [31]. Either of these properties could have influence on the etch rate, as also shown by other studies [28, 52–55, 60] and for the the case of hydrogen content by the simulations in section 4.1.3. An explanation is that when there are already N–H and Si–H bonds present in the SiN$_x$ film, the acid needs less protonization of the bonds before NH$_3$
can be released, thus speeding up the etch process. A different reason could be that films with higher impurity levels also have a lower density, which allows the etchant to ingress into the film more easily and therefore etch the film faster. A third explanation is that the impurities could be present in the form of CH$_x$ species, which has lower binding energies to Si than an NH$_x$ bond and could be therefore more easily removed [42].

The etch experiments with the buffered solution showed the same trends as with the unbuffered solution, as shown in figure 4.6 and table 4.4. The etch rate of the film deposited at 40 mTorr and 7.5 seconds was not shown because the entire film was etched after one dip. Also in this environment, a higher deposition temperature, lower pressure and longer plasma exposure time led to a lower etch rate. Furthermore, table 4.4 shows that the etch selectivity between SiO$_2$ and the SiN$_x$ can change when the concentration of species in the etch solution is also changed. When comparing the nitride film deposited at 500 °C and SiO$_2$, this selectivity increased from a factor of fifty in the dHF to a factor of hundred in the BHF. This difference in selectivity is consistent with the calculations shown in section 4.1.3.

When comparing the experimentally determined values, resp. 0.8 nm/min and 82.8 nm/min for best-quality SiN$_x$ and SiO$_2$, with the values calculated for LPCVD SiN and SiO$_2$ in the BHF in section 4.2.1, resp. 10.4 and 214 nm/min, similar differences occur as with the the etch rates in dHF. The difference in etch rate for the SiN$_x$ is more pronounced, probably due to the higher etchant concentration. Now also for the SiO$_2$ the calculated etch rate is higher than the experimentally obtained value. The SiO$_2$ that Knotter used for determining the weight constants was sputtered on a quartz crystal balance, and therefore it might be of less quality than our used material as well. It could be that this difference was negligible in the dHF, due to its lower concentration.

Other ALD processes for silicon nitride show etch rates of 1-2 nm/min for films deposited at 400 − 500°C when dipped in 0.1-0.5% HF [12, 26]. When comparing these values to ours, it shows that the etch resistance of the films deposited in our research is significantly better, even at lower deposition temperatures and without using halide-based precursors. Since our PEALD process is not yet fully optimized, the quality of our films could even be more improved by using lower deposition pressures and longer exposure times at high temperature.

### Table 4.3: Etch rates in dHF for films deposited at various temperatures, pressures and plasma exposure times. Unless stated otherwise, the used pressure was 40 mTorr. Thermally grown SiO$_2$ is added as a reference.

<table>
<thead>
<tr>
<th>Plasma exposure time (s)</th>
<th>200 °C</th>
<th>300 °C</th>
<th>400 °C</th>
<th>500 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>70</td>
<td>1.3 ± 0.6</td>
<td>0.2 ± 0.4</td>
<td>0.01 ± 0.05</td>
</tr>
<tr>
<td>10</td>
<td>28</td>
<td>(2.8 ± 0.6 @ 10 mTorr)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>11 ± 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal SiO$_2$</td>
<td>0.56 ± 0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.2. Etch experiments on planar wafers

Figure 4.6: Etch rates of deposited films in BHF. Unless stated otherwise, the used pressure and plasma exposure time were 40 mTorr and 10 s, respectively. The etch rate of SiO$_2$ under the same conditions is added for comparison.

Table 4.4: Etch rates in a buffered oxide etch for films deposited at various temperatures, pressures and plasma exposure times. Unless stated otherwise, the used pressure was 40 mTorr. Thermally grown SiO$_2$ is added as a reference.

<table>
<thead>
<tr>
<th>Plasma exposure time (s)</th>
<th>200 °C</th>
<th>300 °C</th>
<th>400 °C</th>
<th>500 °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>&gt; 45</td>
<td>5.6 ± 0.2</td>
<td>1.1 ± 0.3</td>
<td>0.8 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>(14 ± 2 @ 10 mTorr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>38 ± 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal SiO$_2$</td>
<td></td>
<td></td>
<td></td>
<td>82.8 ± 0.7</td>
</tr>
</tbody>
</table>
4.3 Identifying the governing mechanisms in material properties

The results shown in section 4.2.2 show a strong dependency not only on temperature, but as well on pressure and plasma exposure time. These results coincide with the trends in C, H, and O content in the films as shown in fig. 1.4. We would like to understand what is the relevant parameter in these trends, in order to improve material quality and, later on, the conformality of these films.

In section 1.2.2 we mentioned that there are three major plasma-related mechanisms that have influence on material quality and conformality. The trends in the previous section could indicate that redeposition plays a significant role in PEALD of SiN, and that therefore, from these three parameters, residence time is an important parameter in improving the process. A lower pressure means a lower residence time, so reaction products have less time to form new species in the plasma and therefore less products are redeposited, leaving a cleaner surface after a certain time. If there is much redeposition, it takes longer time to reach a clean surface, explaining the improved etch resistance with prolonged plasma exposure. The redeposition process is visualized in figure 4.7. One also might think that ion bombardment plays a significant role in material quality, but we will see later on that this is not the limiting factor in this case.

![Figure 4.7](image-url)  
**Figure 4.7:** Schematic of how redeposition takes place. During the plasma step, ligands are removed and form reaction products. In the plasma, these products can undergo dissociation and ionization by collisions with electrons and form new products, which can redeposit on the surface. A short residence time of particles in the plasma means that particles are flushed away before they can redeposit.

Residence times in the reactor are calculated by adapting the definition (see section 1.2.2) to a steady-state flow controlled by a Mass Flow Controller in a reactor:

\[
\tau = \frac{pV}{p_0q_{MFC}} \text{ (min),} \tag{4.8}
\]
4.3. Identifying the governing mechanisms in material properties

where $p$ is the pressure in the reactor in Torr, $V$ is the volume of the reactor in cm$^3$, $p_0$ is the standard pressure (760 Torr), and $q_{\text{MFC}}$ the flow rate through the MFC in sccm. The derivation of this equation from the definition of $\tau$ is found in Appendix C.

The maximum flow of $N_2$ into the reactor is limited to 100 sccm. To increase the range of residence times we can get with the reactor, we add Ar to the plasma, of which the flow rate can be varied from 0 sccm to 200 sccm. In this way a higher total flow rate can be achieved, and with that lower residence times can be reached. This way we can also isolate the effect of pressure-dependent mechanisms as ion bombardment and radical recombination effects, since now we can go to higher pressures whilst keeping the residence time constant by also increasing flow rate. Higher pressure reduces ion bombardment and increases exposure, making these effects of less influence.

Table 4.5 shows residence times for various pressures and flow rates. In all experiments, the $N_2$ flow rate is kept constant at 100 sccm. From literature it is known that adding an inert gas to a $N_2$ plasma may promote formation of active species by a factor of $\sim 1.5$ through different ionization processes [61, 62]. However, since the partial pressure of $N_2$ is lowered in a $N_2-$Ar mixture by a similar or larger factor, we expect this to make little difference for the surface interactions.

<table>
<thead>
<tr>
<th>$p$ (mTorr)</th>
<th>100 sccm</th>
<th>200 sccm</th>
<th>300 sccm</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>3.5</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td>40</td>
<td>1.7</td>
<td>0.87</td>
<td>0.58</td>
</tr>
<tr>
<td>24</td>
<td>1.0</td>
<td>0.52</td>
<td>0.35 (pump limit)</td>
</tr>
<tr>
<td>20</td>
<td>0.87</td>
<td>0.43</td>
<td>Above pump limit</td>
</tr>
<tr>
<td>13</td>
<td>0.56 (pump limit)</td>
<td>Above pump limit</td>
<td>Above pump limit</td>
</tr>
</tbody>
</table>

To first see the influence of $\tau$ on the growth per cycle, we plot the GPC against the residence time at 200 °C and 500 °C with a plasma exposure time of 10 s, as shown in figure 4.8. From these graphs, two distinct trends can be distinguished. At 200 °C, the GPC increases with increasing $\tau$, whereas at 500 °C the GPC seems to drop with increasing $\tau$. At first these opposing trends may seem confusing, since $\tau$ itself is not so much dependent on temperature. An explanation can be derived from looking at the difference in material properties at these temperatures. At 200 °C it was shown that a relatively high impurity level remains in the film upon deposition, compared to films deposited at 500 °C (see section 1.2.3). The reduction in GPC at 200 °C with decreasing $\tau$ thus can be explained by the reduced redeposition of reaction products. If more products are flushed away, less products are redeposited on the surface, decreasing the GPC. At 500 °C there are already few impurities present, therefore the trend is less clear. The opposite slope then can originate from better flushing of growth-inhibiting products. From chapter 3
we know that adsorbed hydrogen can inhibit growth. A lower residence time also reduces the redeposition of inhibiting species, thereby increasing the growth per cycle.

A more detailed explanation can be given using figure 4.9. There, the GPC is plotted versus the plasma exposure time for processes with various residence times at 200 °C. The blue line with a residence time of 1.7 s is the original process, here the GPC rises initially and then decreases slowly towards an equilibrium value. This value is reached when all reaction products are flushed away and there are no more products left to be removed. The black and red lines with a much lower residence time of resp. 0.58 s and 0.35 s show a GPC that decreases much faster towards equilibrium. This behaviour can again be traced to redeposition of reaction products. At high \( \tau \), much products are redeposited, therefore it takes longer to remove all ligands and leave finally a clean surface. For processes with lower \( \tau \), much less is redeposited and a clean surface is obtained faster. Note that the timescales with which the GPC decreases towards saturation scales linearly with the residence time (inset fig. 4.9). These timescales were extracted from fits of the function

\[
y = 0.6 \left(1 - e^{-t_1/\tau}\right) - B \left(1 - e^{-t_2/\tau}\right)
\]

where \( t_2 \) is the timescale dictating the GPC decrease. When a small amount of \( \text{H}_2 \) (10 sccm, olive line) is added to the plasma, again inhibiting behaviour can be seen as described in chapter 3. The results from figure 4.9 could indicate that by using a lower residence time, a higher quality film can be reached much faster with respect to the original process.

This hypothesis is even more strengthened when we look at the etch rates of films that are deposited with a specific residence time. Figure 4.10 shows the etch rates for several of those films, deposited at 400 °C. Here it is visible that shorter residence times indeed showed a factor 4 improvement in the etch resistivity of deposited films. Now, etch rates of 0.6 nm/min in BHF can be reached at less-than-optimal temperatures, which are already lower that the values originally obtained using the standard process (0.8 nm/min at a deposition temperature of 500 °C). Using a
4.3. Identifying the governing mechanisms in material properties

**Figure 4.9:** The GPC at 200 °C versus plasma exposure time for processes with various residence times. The blue line is the original process \((p = 40\text{mTorr, } 100 \text{ sccm } N_2)\), the red and black lines are processes where the pressure is lowered and Ar is added. The olive line represents a process where a small amount (10 sccm) of \(H_2\) is added to the plasma. The inset shows the timescales of the decrease of the GPC versus the residence time.
low residence time process at 500 °C will probably result in an even lower etch rate.

Preliminary X-ray Photoelectron Spectroscopy data showed similar trends, where the C and O content also decreased linearly with decreasing residence times. Figure 4.10 also shows that ion bombardment is not the limiting factor in material quality, because now even at higher pressures (black dot, reduced ion bombardment) we get similar or lower etch rates when compared to a lower pressure process (magenta dot, more ion bombardment).

The ability to now go to higher pressures while still getting good material quality can reduce the anisotropy caused by ions and increase the diffusion of particles in high aspect ratio structures, which in turn can improve the conformality of deposited films.

![Figure 4.10: Etch rates in BHF for films deposited with various residence times. The deposition temperature was 400 °C, the plasma exposure time was 10 s. The legend shows what other parameters led to a specific \( \tau \). The dashed line serves as a guide to the eye.](image)

### 4.4 Conformality and etch selectivity in HARS

Using the extra freedom in flow rates and pressures that the residence time provided, optimal conditions for depositing conformal layers in high aspect ratio structures were determined. Again referring to section 1.2.2, it is desired to have a high pressure to get reduced anisotropy from ions and a high exposure, and low residence times to reduce the redeposition of reaction products. Therefore, two depositions were done in HARS using the following conditions: 1. 24 mTorr, 10 s exposure (lowest residence time possible, \( \tau = 0.35 \) s); 2. 80 mTorr, 10 s exposure (to have high pressure and decent material quality, \( \tau = 1.2 \) s). In both experiments, the \( \text{N}_2 \) and Ar flow
rates were maximized to 100 sccm and 200 sccm, respectively. The deposition temperature was 400 °C.

Figure 4.11 shows several HARS with a silicon nitride film deposited under the conditions as mentioned above. To investigate the conformality of the films, the sidewall to top (S/T) and bottom to top (B/T) thickness ratios were measured by SEM. Because of the skewing of the SEM images, the top thickness was not measured from the horizontal surface but from the film directly below the entrance of the trench. The results are shown in table 4.6. Condition 1 yields a >80% conformality for aspect ratios ≤2, with little variation. Above AR 2 the conformality decreases rapidly. This was as expected: when structures are placed closer together, radical recombination becomes of influence, and this effect results in worse conformality at lower pressures. Condition 2 gives excellent conformality of >90% for aspect ratios up to 3, with virtually no variation. The higher pressure causes reduced anisotropy from ions and a higher exposure, making sure all surface sites in the trench are saturated. At aspect ratio 4 the conformality of condition 2 also decreases significantly.

To investigate the etch uniformity in the structures, the as-deposited samples were dipped in BHF for 3 minutes. Afterwards, they were again analyzed with SEM. Because low thicknesses are difficult to measure with SEM, a qualitative analysis will be given.

Figure 4.12 shows SEM images of several HARS with a silicon nitride film after 3 min in BHF for condition 1. We can see that for the higher aspect ratios (AR 2, fig. 4.12a) the film is etched non-uniformly; the etch rate at the bottom of the trench is higher than in the top of the trench. This is unwanted when applying these layers as spacers, then a uniform etching throughout the trench is wanted. However, for the lower aspect ratios (AR 1.2, fig. 4.12b), the film is still conformal and virtually no film seems to have been etched away. For applications using this aspect ratio, condition 1 seems to be an excellent solution.
Table 4.6: Sidewall to top (S/T) and bottom to top (B/T) thickness ratios as measured with SEM in trenches with various aspect ratios, for the conditions as mentioned above. All the ratios have a 5 percent point error margin.

<table>
<thead>
<tr>
<th>AR</th>
<th>Condition 1 S/T (%)</th>
<th>Condition 1 B/T (%)</th>
<th>Condition 2 S/T (%)</th>
<th>Condition 2 B/T (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>93</td>
<td>95</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>1.7</td>
<td>88</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>80</td>
<td>94</td>
<td>93</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>93</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>93</td>
<td>75</td>
<td>90</td>
</tr>
</tbody>
</table>

Similar tests were performed for films deposited under condition 2, as shown in figure 4.13. Here, the etching is more uniform throughout the trench, even at higher aspect ratios (AR 3, fig. 4.13a). The overall etch rate seems to be slightly worse than for condition 1, which is as expected. Also for condition 2 films in low aspect ratio trenches were etched uniform and have good resistivity (AR 1-1.5, fig. 4.13b). Whereas overall the etch rate is slightly worse than for condition 1, condition 2 etches uniform and therefore can also be used for creating good protection barriers.

Figure 4.12: SiN$_x$ films deposited under condition 1 in HARS after 3 min BHF dip. (a) in HARS of aspect ratio $\sim$2; (b) in HARS of aspect ratio $\sim$1.

If we compare the conformality of our SiN$_x$ process with those reported reported in literature, we find that our results are very encouraging. Whereas the conformality of our process is still of less quality compared to existing SiN$_x$ processes that use a halide precursor at 500 °C [12], the films grown with BTBAS perform as good as or better than existing low-temperature halide-free
4.5. Summary

Figure 4.13: SiN$_x$ films deposited under condition 2 in HARS after 3 min BHF dip. (a) in HARS of aspect ratio $\sim$3; (b) in HARS of aspect ratio 1-1.5.

processes [17]. In single pitch structures (AR 1), our process results in $> 90\%$ conformal films and very low etch rates throughout the structure, with overall etch rates lower than for existing low-$T$ halide ALD processes [12].

4.5 Summary

In this chapter, the underlying mechanisms that have influence on material properties as etch resistance and conformality were determined. It was shown that the residence time of particles is an important parameter in ALD of silicon nitride, where a low residence time leads to a better etch resistance and faster saturation of growth. The low residence time prevents redeposition of reaction products, creating a denser film with lower impurity content.

In general, wet etch rates of $< 2 \text{nm/min}$ in BHF and dHF were achieved for deposition temperatures of 300-500 °C, performing similar to or better than existing processes up to now in that temperature range. These wet etch rates follow trends similar to impurity content in the films and density, making it plausible that the wet etch rate is determined by these properties.

The conformality of this process was investigated for two conditions optimized for residence time, exposure, and ion bombardment. Both of these conditions resulted in excellent conformality of $> 90\%$ in single-pitch structures, whereas for higher aspect ratios the condition utilizing a higher pressure obtained better conformality. The etch uniformity in these structures was also determined and was high in single-pitch structures for both conditions. For higher aspect ratios, the film deposited with the condition utilizing higher pressure showed uniform etching, but the overall etch rate seemed to be slightly worse than condition 1. These observations could be explained by that the condition with higher pressure resulted in a more isotropic deposition, but since the residence time in this process was higher, the film was overall of less quality. This
apparent tradeoff can be used to specify conditions depending on the application of the films.
5 General conclusions & Outlook

5.1 General conclusions

This thesis deals with several aspects regarding the underlying mechanisms involved in synthesis and growth of silicon nitride using a plasma-enhanced atomic layer deposition process. This process made use of Bis(t-butyramino)silane (BTBAS) as precursor and a N₂ plasma (and N₂−Ar plasma) as reactant. The following conclusions can be drawn from this work:

- Hydrogen-containing plasmas or NH₃ gas can inhibit the growth of silicon nitride using BTBAS. A N₂ plasma can reduce this inhibition. The N₂ plasma is able to abstract hydrogen from the surface, presumably through reactions with N radicals or bond weakening by ion bombardment. By using a sequential cycle of a H-containing plasma and a N₂ plasma, growth rates of 0.16 Å/cycle can be achieved at deposition temperatures of 200 °C and 500 °C. There were also indications that using an additional H-containing plasma in the silicon nitride ALD cycle may result in higher-quality films at lower temperatures.

- It was shown by Density Functional Theory simulations on β-Si₃N₄ that BTBAS needs undercoordinated surface sites to adsorb. If there are none of such sites available, the precursor will not adsorb, reducing film growth. DFT also showed that H-containing plasmas or NH₃ gas can passivate a Si₃N₄ surface, and thus effectively disabling BTBAS adsorption. When a N₂ plasma reacts with a silicon nitride surface, undercoordinated sites remain, explaining higher growth when using a N₂ plasma. The DFT gave theoretical evidence for the inhibition observed when using H-containing plasmas or NH₃ gas.

- Using the standard BTBAS/N₂ plasma ALD process, low wet etch rates of < 2 nm/min can be achieved at deposition temperatures of 300-500 °C, which are comparable to or lower than etch rates obtained with existing ALD processes within this temperature range, and can be further optimized.

- The limiting factor in material quality is related to redeposition of reaction products. This effect can be reduced by decreasing the residence time of particles in the plasma. When doing so (by increasing effective pumping), even lower etch rates can be achieved with respect to the original process, of 0.5 nm/min in BHF at a substrate temperature of 400 °C. Furthermore, this higher quality film can be reached in shorter cycle times. The residence time also allows to use higher pressures, which is favorable for deposition of conformal films. This effect is often overlooked regarding PEALD, but here we have shown that it is an important factor to consider.
Chapter 5. General conclusions & Outlook

- By using a low residence time process, conformal films were achieved up to aspect ratio 1.7 with non-uniform etching in higher aspect ratio structures. In single-pitch (AR 1) trenches, the films were conformal and showed excellent uniformity in etching. By using a high-pressure, higher residence time process (decent blanket etch resistance), films of excellent conformality were achieved in aspect ratios up to 3, and maintained high conformality after etching.

In short: the fundamental mechanisms involved in growth and material properties have been studied. These mechanisms are found to have important technological implications, and can be extended to other ALD processes as well. By controlling these mechanisms, an industrially viable process for silicon nitride that fulfills the requirements of high etch resistance and good conformality can be developed. Therefore, the BTBAS/N$_2$–Ar plasma ALD process is an excellent choice as new spacer deposition process for the new generations of FETs.

5.2 Outlook

Eventually, it is preferred to have a deposition temperature that is as low as possible whilst maintaining good film properties and etch resistivity. At present, films deposited at temperatures below 300°C do not meet these requirements, even with a low residence time process. Observations regarding inhibition by hydrogen led us to presume that while hydrogen does inhibit the surface, it is very effective in removing the precursor ligands. Therefore, using a H-containing plasma might be more favorable for low temperature depositions. To overcome the inhibition, an ALD process with multiple steps could be devised, where after the precursor step the surface is first exposed to a H-containing plasma and afterwards a low residence time N$_2$–Ar plasma, to have both good removal of ligands and a fast recovery of the growth rate. Such an ALD cycle is depicted in figure 5.1.

The inhibition could also be used to advantage. It might be used for selective ALD, if intentional low growth is desired. The selectivity could also be used for patterning of structures. When the hydrogen of certain parts of the surface is removed by either e-beam lithography or UV exposure, the growth rate on those parts would be higher. The advantages of such area-selectivity are discussed in ref. [63].

A different way to overcome the redeposition of reaction products is by using other precursors. Since QMS data from earlier research showed that the BTBAS-covered surface still has butyl ligands when entering the plasma half-cycle, this probably causes the redeposition effect [31]. If the precursor would only have one butylamine ligand, this ligand would be removed during the precursor half-cycle, leaving a C-free surface when entering the plasma half-cycle. Without any C, there would not be any redeposition effects, creating a better-quality surface and faster saturation. The DSBAS precursor [64] only has one di-butylamine ligand, so this could be a promising precursor for low temperature silicon nitride deposition. The adsorption of DSBAS to a silicon oxide surface has been studied in ref. [64], where the removal of the dibutyl ligand was observed. Of course, the adsorption process to a silicon nitride surface might be different, but it may be worth it to investigate this process. The theoretical adsorption process of DSBAS to a silicon nitride surface is shown in figure 5.2.
5.2. Outlook

**Figure 5.1:** Illustration of a multi-step ALD cycle where the precursor-saturated surface is first exposed to a H-containing plasma in order to effectively remove all precursor ligands. Afterwards, the surface is exposed to a N\textsubscript{2} plasma which removes the adsorbed hydrogen, allowing new precursor adsorption.

**Figure 5.2:** Schematic of how the DSBAS precursor would adsorb to a silicon nitride surface. The butylamine ligand could be removed during precursor adsorption, leaving a carbon-free surface when entering the plasma step, preventing slow saturation by redeposition.
Bibliography


Appendix
A Relation between setpoint temperature and wafer temperature

Table A.1: A guideline for the relation between the actual wafer temperature $T_{wafer}$ and the setpoint temperature $T_{set}$ for a pressure of 0 mTorr. When processing at higher pressures (20-80 mTorr) the wafer temperature will be approx. 25-60 °C higher [31].

<table>
<thead>
<tr>
<th>$T_{set}$ (°C)</th>
<th>$T_{wafer}$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>85</td>
</tr>
<tr>
<td>200</td>
<td>155</td>
</tr>
<tr>
<td>300</td>
<td>210</td>
</tr>
<tr>
<td>400</td>
<td>275</td>
</tr>
<tr>
<td>500</td>
<td>355</td>
</tr>
</tbody>
</table>
B  Matlab script for calculating etch rates

clear all;
format short;
warning off;
tot = 20; % in how much intervals you want to split the range

%Declare variables
syms x1 positive; % [H+]  
syms x2 positive;%( [ F- ] *)  
syms x3 positive;%( [HF] *)  
syms x4 positive;%( [HF2-] *)  
syms x5 positive;%( [H2F2] *)  
R=zeros(tot+1,1); 
sx1=zeros(tot+1,1);  
sx2=zeros(tot+1,1);  
sx3=zeros(tot+1,1);  
sx4=zeros(tot+1,1);  
sx5=zeros(tot+1,1);  

for i=0:tot

% Etch rate constants for material taken from Knotter et al. (2001) *)
k0 = 0.29;  
k1 = 0.21;  
k2 = 28;  
k3 = 25;  

%(*Analytical concentrations *)  
CF = 33.32; %(*all fluoride containing species *)  
NH = CF/100.*(100/tot*i);  
HF = CF/100.*(100-i); %concentration of NH4+ for buffered solution  

%HF concentration calculations *)
eqn = [ 
x1.*x2./x3 == 0.00072, %(*from HF <-+ H+ + F- *)  
x4./(x3.*x2) == 3.963, %(*from HF2- <-+ HF + F- *)
]
x5./(x3.^2) == 2.7, %(*from 2 HF <-> H2F2 *)
x1 + NH == x2 + x4, %law of neutrality [H+] +[NH4+] = [F-] + [HF2-]
CF == x2 + x3 + 2*x4 + 2*x5
];

%(*gives concentrations of all species*)
[sx1(i+1), sx2(i+1), sx3(i+1), sx4(i+1), sx5(i+1)] =
solve(eqn,x1,x2,x3,x4,x5,'PrincipalValue',true);

R(i+1) = (k0.*sx4(i+1)./(1 + K3.*sx1(i+1)) + (k1.*sx3(i+1)
  + k2.*sx2(i+1)).*K3.*sx1(i+1)./(1 + K3.*sx1(i+1))).*6;
end
x=0:100/(tot):100;

plot(x,R);
C Derivation of the residence time

The definition of the residence time $\tau$ is as follows:

\[ \tau = \frac{V}{q}, \quad (C.1) \]

where $V$ is a certain volume and $q$ is the flow rate through that volume, in volume per time. We want to express the residence time in parameters we can control experimentally, which are the pressure $p$ inside the volume (the reactor) and the flow rate of the gases entering the reactor $q_{\text{MFC}}$. For an ideal gas ($pV = NkT$), the flow rate is given by

\[ q = \frac{\phi kT}{p}, \quad (C.2) \]

with $\phi$ the particle flux through the volume, $k$ Boltzmann’s constant, $T$ the temperature of the gas, and $p$ the pressure inside the volume $V$. During an experiment the gas flow and pressure are stabilized, so we have a steady-state flow. This means that the particle density does not change over time. This in turn means that the particle flux coming in the chamber ($\phi_0$) and going out of the chamber ($\phi$) are equal, $\phi_0 = \phi$. The particle flux coming into the chamber travels through an MFC. The MFC measures the flow rate in sccm, standard cubic centimeter per minute. A standard cubic centimeter is the amount of particles that takes up 1 cm$^3$ under standard pressure $p_0$ (760 Torr) and standard temperature $T_0$ (293.15 K). The particle flux $\phi_0$ coming through the MFC is thus given by (using again the ideal gas relation and $q_{\text{MFC}}$ the flow rate as measured by the MFC):

\[ \phi_0 = \frac{p_0q_{\text{MFC}}}{kT_0}. \quad (C.3) \]

Note that the units of the variables are not yet taken into account, this will be done at the end of this section.

When we now combine equations (C.2) and (C.3), together with the requirement that $\phi_0 = \phi$, and also with the assumption that the gas temperature does not change when the gas is entering the reactor ($T_0 = T$), we get as a result for the flow rate in the reactor:

\[ q = \frac{q_{\text{MFC}}p_0}{p}. \quad (C.4) \]

This we can fill in the definition of the residence time (C.1), yielding as result for the residence time $\tau$ in the reactor:

\[ \tau = \frac{pV}{p_0q_{\text{MFC}}}. \quad (C.5) \]

Of course, one has to take into account the different units of the variables. If $q_{\text{MFC}}$ is in cm$^3$/min, then $\tau$ is in minutes, $p$ and $p_0$ have to have the same units (Torr or Pa or any other arbitrary unit), and $V$ is in cm$^3$.  

69