Introductory invited paper

Reliability of ultrathin gate oxides for ULSI devices

Chun-Yen Chang\textsuperscript{a,*}, Chi-Chun Chen\textsuperscript{a}, Horng-Chih Lin\textsuperscript{b}, Mong-Song Liang\textsuperscript{c}, Chao-Hsin Chien\textsuperscript{a}, Tiao-Yuan Huang\textsuperscript{a,b}

\textsuperscript{a}Institute of Electronics, National Chiao Tung University, Hsinchu, Taiwan
\textsuperscript{b}National Nano Device Laboratories, Hsinchu, Taiwan
\textsuperscript{c}Taiwan Semiconductor Manufacturing Co. Ltd., Hsinchu, Taiwan

Received 5 April 1999

Abstract

Ultrathin gate oxide, which is essential for low supply voltage and high driving capability, is indispensable for the continued scaling of ULSI technologies towards smaller and faster devices. Needless to say, the reliability of ultrathin oxide is of major concerns in the manufacturing of the state-of-the-art metal-oxide-semiconductor devices. This paper reviews the reliability issues regarding ultrathin gate oxide for present and future ULSI technologies. Issues including gate leakage current, time-dependent dielectric breakdown, poly-gate depletion, boron penetration, and plasma process-induced damage will be addressed. Several techniques such as nitrided oxide and alternative processes, which are proposed to improve gate oxide reliabilities, are also discussed. © 1999 Elsevier Science Ltd. All rights reserved.

1. Introduction

To ensure the continued shrinkage of CMOS technologies down to sub-quarter micron regime and beyond, ultrathin gate dielectric with low defect density and high reliability is indispensable. In general, high driving capability and well-controlled short channel characteristics require the use of ultrathin gate oxides. Since Momose et al. first proposed the use of 1.5 nm direct-tunneling (DT) gate oxide for \textit{n}-MOSFET with extremely high device performance in 1994 \cite{1}, many aggressive studies regarding direct tunneling gate oxide for sub-quarter micron devices have been carried out \cite{2-10}. The application of DT gate oxide not only achieves high speed for logic circuits, but also essential for high-performance RF applications \cite{8,9}. In addition, the use of thinner oxides is also critical to meet the demands of lower programming voltage for future nonvolatile memories \cite{10,11}.

However, the application of DT gate oxide to ULSI devices faces many challenges. Firstly, the presence of large quantum mechanical (QM) tunneling current is a serious scaling limitation in terms of standby power consumption \cite{12}. Secondly, breakdown characteristics for ultrathin oxides become even more critical due to the dramatic increase in electric field across the oxide during normal device operation. Whether oxide becomes inherently more robust or more vulnerable to electric stress as it thins down, therefore, plays a very crucial role for its applications to ULSI devices. Thirdly, poly-gate depletion effects (PDE) are known to get worse with oxide scaling. This is because the operating gate voltage normally does not scale proportionally to the oxide thickness, and therefore, the average surface field increases. The additional voltage drop at the poly depletion layer results in undesirable drive...
current degradation [13]. Since these effects worsen for thinner oxides, it could become a limiting factor for future device scaling.

Furthermore, as gate oxide thickness decreases, several process integration issues emerge as new challenges. Boron penetration from the $p^+$-poly-Si gate into the thin gate oxide and the channel region in $p$-MOSFETs is one of the major concerns for dual-gate CMOS technologies. The penetration of boron into and through gate oxides results in threshold voltage instability, low-field mobility reduction, and gate oxide degradation [14–18]. The undesirable boron penetration effects are aggravated as gate oxide is scaled down. Finally, plasma-process-induced charging damage can result in severe degradation in oxide integrity as gate oxide thickness is scaled down [19–27]. The frequent use of plasma steps with high-density plasma reactor in advanced circuit manufacturing, coupled with the restriction that no high temperature anneal is allowed in back-end processing, aggravates the situation. It is, therefore, very important to characterize the performance of plasma processes in production with scaled oxides.

In this work, we will review recent studies on issues relating to the reliabilities of ultrathin gate dielectrics for present and future ULSI technologies.

2. Gate leakage current

2.1. Quantum mechanical tunneling

Significant leakage current that flows through gate oxide due to QM tunneling may cause reliability problems and limits device application. It is well known that as the gate oxide thickness is scaled down to 3 nm and below, the gate leakage current increases dramatically due to the occurrence of DT. This happens when oxide voltage is reduced to lower than 3.2 V, the tunneling barrier for electron changes from triangular to trapezoidal. The current conduction mechanism is the so-called DT current [28]. Typical gate current characteristics of ultrathin oxides (2.5–4.2 nm) and the simulated Fowler–Nordheim (F–N) current fitting curves are as shown in Fig. 1, while the simulated DT current for thinner oxides down to 1.5 nm are shown in Fig. 2. The F–N $I$–$V$ fitting method for extracting ultrathin oxide thickness has been widely adopted because of its accuracy and convenience as an on-wafer measurement method [29,30]. By taking PDE into account, the method is well suited for ultrathin oxide measurements. This is especially so, considering the fact that PDE has made the extraction of oxide thickness from the traditional $C$–$F$ measurements difficult. However, this method is not suitable for measuring oxide thickness thinner than 2 nm because not enough F–N current is available for extraction before oxide breakdown. Therefore, to find an accurate and convenient method to determine the exact oxide thickness in DT conduction region is important and urgent.
2.2. Gate leakage current as a limitation of scaling

As a general rule, logic circuits can tolerate higher leakage current than memory circuits. It has been shown that the DT current associated with 1.5 nm oxide is comparably small to the large drain current of 0.1 \(\mu\)m MOSFET. However, leakage current of such magnitude could still result in lower data retention time for DRAM, higher chip standby power and accelerated device degradation. Thus, large DT current presents a scaling limitation for future CMOS technologies. To avoid CMOS operation at high electric field with large standby power, the power supply voltage \(V_{dd}\) has to scale more aggressively. This, however, defeats the purpose of scaling. To maintain device performance, it is projected that the oxide field during normal operation will stay at around 5 MV/cm [11], and that the minimum supply voltage will be around 1 V (0.9 V for worst case). If one assumes that the total active gate area per chip is of the order of 0.1 cm\(^2\) for future generation technologies, the maximum tolerable tunneling current will be about 1–10 A/cm\(^2\) [31]. As indicated in Fig. 2, the minimum thickness of silicon dioxide to meet this criterion as the gate dielectric will be around 1.85 nm.

2.3. Alternate gate dielectrics

To overcome the gate leakage current limitation for scaling of future CMOS technologies, many alternative materials with high dielectric constant \(\epsilon_r \gg \epsilon_{SiO_2}\) to serve as the gate dielectric were studied [32–44]. These ‘high-\(k\)’ materials can provide the same equivalent electrical thickness by a physically thicker layer which is, therefore, more controllable and hopefully more manufacturable. These films prepared with various advanced technologies have been shown to effectively reduce gate leakage current, suppress boron penetration and have higher time-dependent dielectric breakdown (TDDB) lifetime as well. These films fall into three main categories. The first group is nitride-related materials including oxynitride grown on nitrogen-implanted substrates [32], jet vapor deposited (JVD) nitride [33,34], and LPCVD nitride [35]. The second group belongs to metal oxides, such as Ta\(_2\)O\(_5\) [41,43], TiO\(_2\) [37,42,43], BaSrTiO\(_3\) (BST) [42], Al\(_2\)O\(_3\) [44] etc. The third group employs stack gate structures such as nitride/oxide [39,42], metal oxide/oxide [36,40] and metal oxide/nitride [39,42]. All three groups appear to be promising as potential replacement for the thermal-grown SiO\(_2\) dielectric. Nevertheless, process optimization and detailed reliability data of these novel materials need to be established before they can be confidently accepted into mass production.

3. Breakdown characteristics

3.1. Occurrence of soft-breakdown

Breakdown characteristics of ultrathin gate oxides are discussed in this section. It was observed that when the oxide is thinner than 6 nm, an anomalous failure mode may be induced during high electric field stressing [45,46]. Specifically, thin gate oxide sometimes exhibits a significant leakage current increase accompanying a characteristic ‘noisy’ or fluctuating leakage current [47,48]. In contrast to the conventional ‘hard breakdown’ (HBD), this new breakdown phenomenon is named soft-breakdown (SBD), quasi-breakdown or partial-breakdown. Its occurrence complicates oxide reliability evaluation. Fig. 3 shows typical \(V-t\) curves during charge-to-breakdown \((Q_{bd})\) measurements using constant current stress, which is a well-known and widely accepted method for evaluating oxide reliability, for various gate oxide thickness. The abrupt drop in applied voltage after a certain period indicates the occurrence of oxide breakdown. It can be seen that the magnitude of post-breakdown voltage for SBD and HBD events is quite different. While the post-breakdown voltage after HBD is around 1 V or less, the voltage after SBD can be more than 1 V of magnitude. Moreover, the SBD events depict the characteristic ‘noisy’ behavior with fluctuating voltage, accompany by an increase in non-switching 1/f noise and random telegraph noise after soft-breakdown [49]. Such
behaviors could be ascribed to on/off switching events of one or more local conduction spots [48].

3.2. Oxide thickness dependence on $Q_{bd}$ characterization

The oxide breakdown characteristics show strong thickness dependence, as is evidenced from Fig. 3. For thicker oxides (e.g., 8.6 nm), only HBD is observed, while both SBD and HBD events are induced in oxide with intermediate thickness (e.g., 4 nm). For ultrathin oxide (e.g., 2.5 nm), SBD is found to dominate exclusively the breakdown events [50]. Moreover, for even thinner oxide (<2 nm), due to the extremely large DT current, the voltage drop after SBD is very small, and may not be noticeable. Therefore, it becomes very difficult to detect oxide breakdown from the $V-t$ curves [4]. Fig. 4 shows the 50% $Q_{bd}$ values, measured at 25°C and 180°C, as a function of oxide thickness. At room temperature, it is found that $Q_{bd}$ of 2.6 nm-thick oxide is about three orders of magnitude higher than that of thicker oxides. Such ‘turn-around’ behavior indicates that oxides show higher tolerance to DT current since negligible energy is deposited inside the oxide layer, comparing to the F–N tunneling [47]. However, when temperature is raised to 180°C, $Q_{bd}$ of 2.6 nm-thick oxide is only about one order of magnitude higher than that of thicker oxides. This implies a very significant temperature acceleration effect for ultrathin oxide under DT stressing. This suggests that more attention should be paid to burn-in tests as well other wafer processing steps that encounter elevated temperature in order not to jeopardize the gate oxide integrity (GOI) of ULSI devices employing ultrathin gate oxide [51].

The polarity dependence of $Q_{bd}$, which is the $Q_{bd}$ difference between gate injection ($V_g < 0$) and substrate injection ($V_g > 0$) stressing, is shown in Fig. 5. The stressing current density is fixed at 1 A/cm$^2$. It has been well documented in previous reports that the polarity dependence increases with decreasing oxide thickness ($T_{ox}$) for oxides thicker than 4 nm. This is ascribed to the different properties between the poly-Si/oxide and oxide/Si interfaces [52,53]. As shown in Fig. 5, as oxide is further scaled down, the polarity dependence becomes even more dramatic. This is mainly due to the rapid rise in $Q_{bd}$ under substrate injection polarity as oxide is thinned down. $Q_{bd}$ under gate injection polarity, however, remains relatively unchanged during F–N stressing.

3.3. Evaluation of ultrathin oxide reliability

As described above, the evaluation of ultrathin oxide reliability becomes much more complicated as oxide is scaled down. Recently, several studies have pointed out that oxide breakdown is a strong function of device geometry [54–56]. As such, the conventional use and its interpretation of $Q_{bd}$ for comparing different MOS processes may lead to erroneous conclusion. For the same token, traditional use of large-area sample (i.e., capacitor) for evaluating oxide reliability
may also lead to erroneous conclusion for device level applications [56]. It is also reported that the oxide degradation by SBD may not significantly affect transistor's drain current characteristics but increase gate current and noise [4, 47]. However, in more recent studies, it is reported that device can still suffer HBD failure if the channel length of the device is sufficiently small (i.e., 0.2 μm) [55]. Therefore, evaluation of ultrathin oxide reliability remains a critical concern for scaled oxides.

Since $Q_{bd}$ tests is no longer a valid tool to evaluate ultrathin oxide reliability [56], the constant voltage stressing (CVS) method which represents a more realistic situation for practical applications may be more suitable for evaluating ultrathin oxide reliability [55]. Fig. 6 illustrates the TDDB characteristics for various oxide thicknesses (e.g., 4.2, 3.0 and 2.5 nm). Time-to-breakdown ($T_{bd}$) tests under different oxide voltages at high field stressing are often employed for predicting the oxide lifetime under normal operating field (e.g. 5 MV/cm). As shown in Fig. 6, TDDB improves with decreasing oxide thickness. This can be ascribed to reduced trapped charges as well as smaller interface states generation after electrical stress as oxide is scaled to direct-tunneling regime. By extrapolating the data from Fig. 6, an electric field of over 9 MV/cm at room temperature is projected for 10-year lifetime. It is worthy to note that oxide lifetime is strongly dependent on gate area [57]. While the total gate area on a chip is of the order of 0.1 A/cm² for future generation technologies, the projected lifetime from Fig. 6, therefore, may be overestimated. Nonetheless, one can still conclude that TDDB performance is significantly improved with scaled oxides.

Finally, hot-carrier induced oxide degradation has also received much attention for scaled oxides since...
hot-carrier generation becomes more significant [58]. Figs. 7 and 8 depict hot carrier degradation results performed on 2.5 and 4.2 nm oxides. The hot carrier stress was performed with constant drain voltage of 3 V for 2.5 nm oxide and 4 V for 4.2 nm oxide, with appropriate gate bias to ensure maximum substrate current injection. These results confirm that the degradation in transconductance (Fig. 7) and threshold voltage (Fig. 8) are much smaller for the thinner oxide, consistent with data proposed in [58] ($T_{ox}$ range: 1.5–3.8 nm). Thus ultrathin gate oxide do exhibit higher hot-carrier resistance, despite the fact that it is biased under a higher electric field stress due to a not proportionally scaled $V_{dd}$.

4. Poly-gate depletion effects

Poly depletion effects, caused by insufficient active dopant concentration near the poly/SiO$_2$ interface, become non-negligible as the miniaturization of devices proceeds into deep sub-micrometer regime. A lower effective surface electric field, due to additional voltage drop across the depletion layer, was reported to cause a significant reduction in drive current [59], difficulties in determining accurate oxide thickness [60], and overestimation of dielectric lifetime [61].

Significant degradation of inversion capacitance caused by poly-gate depletion was found to degrade device driving capability. The effect was more severe for devices with lower rapid thermal annealing (RTA) cycles and/or implant dose [59]. For n-MOS transistors with n$^+$ poly-Si gate, the effect can be explained by the physical model that As or P atoms tend to segregate to grain boundaries and become inactive [61]. Since the depletion width from the poly/SiO$_2$ interface increases if the oxide electric field increases, the PDE worsens for ultrathin oxide, as the supply voltage usually does not scaled proportionally. Thus, improvement in device performance by oxide scaling can not be achieved without resolving this issue. Fig. 9 shows the calculated gate capacitance to oxide capacitance ratios for devices under inversion polarity as a function of effective ploy gate doping, without taking the QM effect into consideration. The gate oxide thickness ranges from 1.5 to 3.5 nm. We can see that for typical effective doping concentration of $1 \times 10^{20}$ cm$^{-3}$, gate capacitance plunges to only 80% of oxide capacitance for 3.5 nm oxide (i.e., 20% degradation). The degradation increases to almost 35% for 1.5 nm oxide. This trend indicates that drain current gain from thinner oxide is partly compensated by poly depletion effect. A heuristic solution to this problem is to enhance the thermal budget and/or implant dose. This, however, will cause adverse side effects. Since for deep submicron technologies, a very steep channel dopant profile and ultra-shallow S/D junction are required to suppress the so-called short-channel effects. Increasing thermal budget certainly jeopardizes the formation of steep doping profile. Furthermore, prolonged activation cycle and higher dopant implantation bring about disastrous penetration effect for pMOS transistors with p$^+$ gate. These factors reduce the available process windows for optimal process integration as the technology is further scaled down.

Recently, several alternative process technologies were proposed to reduce or eliminate altogether poly depletion effects [62–69]. Poly-Si$_{1-x}$ Ge$_x$ has been proposed as a very promising alternative gate material, since it suffers from less poly depletion problem and is also less susceptible to boron penetration [62,63]. The suggested Ge content for optimum performance is about 20% [62]. In addition, refractory metals such as TiN or WN$_x$ have also been proposed as alternate gate materials due to its low gate resistance, and midgap work function [64–69]. Furthermore, being metallic, it is inherently immune to gate depletion problem. Nevertheless, more studies are needed regarding their process compatibility and impacts on device reliability before their acceptance by the ULSI industry.

5. Boron penetration

The effects of boron penetration into and through the gate oxides of p-channel devices are discussed in this section. P$^+$-polysilicon gate for p-channel device is indispensable for deep-submicron CMOS technologies since it offers better short-channel behavior than conventional
buried-channel structure employing $n^+$-polysilicon gate. However, boron diffusion from the $p^+$-poly-Si gate into the thin gate oxide and the underlying channel region causes $P$-channel devices to depict flatband voltage shift, threshold voltage instability, low field mobility, drive current reduction, and gate oxide degradation [15,16], etc. The boron penetration effects may be modeled as the creation of a very thin sheet of negative charge at the Si/SiO$_2$ interface, which can be regarded as either a fixed negative charge in the gate oxide or a shallow $p$-type dopant layer in the silicon substrate [14]. Previous studies have shown that boron penetration becomes more severe with increasing anneal temperature [15]. As oxide thickness scales to direct tunneling regime, the situation also worsens. Since the thermal budget must be sufficient to achieve adequate gate dopant activation in order to avoid performance loss caused by gate depletion. These conflicting requirements make the process window between boron penetration and poly-gate depletion very stringent [70,71]. To resolve these issues, boron penetration effects and methods to minimize them have thus been widely studied.

Although it has been reported that fluorine can be used to improve gate oxide integrity for $n$MOS device, incorporation of fluorine from BF$_2$ implantation is known to exacerbate boron penetration [72,73]. It is also known that annealing in the presence of hydrogen will enhance boron diffusion [74]. Thus it is crucial to keep the content of fluorine and hydrogen in the process flow as low as possible to minimize boron penetration. Several techniques have also been proposed for suppressing boron penetration, and they could be divided into two categories: the first category is to lower the boron diffusivity in the poly-silicon gate. A typical example is to employ gate structure such as $\alpha$-Si or stacked poly-Si gates [75–77]. Another example is the use of poly-Si$_{1-x}$ Ge$_x$, which is shown to depict reduced boron penetration [62,63]. The second category is to improve the immunity of gate dielectric to boron penetration. A typical example is to introduce nitrogen into gate oxide by various processes (e.g., nitridation or nitrogen implantation), therefore, effectively suppress boron penetration and improve gate oxide integrity [78].

Fig. 10 depicts threshold voltage ($V_{th}$) as a function of post-implantation annealing condition for $p$-MOSFETs with pure oxides and N$_2$O-nitrided oxides. The oxide thickness is 4.2 nm. Both boron and BF$_2$ implants were activated for 20 s.

Fig. 11. Gate leakage current as a function of post-implantation annealing condition for $p$-MOSFETs with pure oxides and N$_2$O-nitrided oxides. The oxide thickness is 4.2 nm. Both boron and BF$_2$ implants were activated for 20 s. Gate leakage current was measured at a gate voltage $V_g = -2$ V under inversion polarity and a low drain bias $V_d = -0.1$ V.
boron penetration on gate leakage current are shown in Fig. 11. Consistent with Fig. 10, severe boron penetration results in large leakage current. Samples with N₂O-nitrided oxide and boron implant depict the smallest leakage current, even under the worst annealing condition. By comparing Figs. 10 and 11, it appears that device performance (i.e., threshold voltage) is more sensitive to boron penetration than oxide leakage current for ultrathin gate oxides. To sum up, process optimization and the use of alternative materials for gate dielectrics or gate electrodes may be necessary for solving boron penetration problem encountered in future ultrathin gate oxide devices.

6. Plasma charging damage

Plasma charging effect, which may lead to severe oxide degradation and profile distortion during processing, has become one of major reliability concerns in ULSI manufacturing since the late 80s [19–27]. This can be attributed to several reasons: (1) oxides become more susceptible to charging damage as \( \text{T}_{\text{ox}} \) is scaled below 10 nm. (2) The number of plasma steps employed in wafer fabrication increases dramatically as the chip functionality and complexity advance. (3) In order to improve the throughput or to meet the critical requirements of deep-submicron manufacturing, process tools with high-density plasma (HDP) reactors for etching and deposition steps are widely used. These process steps may potentially aggravate the extent of charging.

6.1. Causes for plasma charging damage

It is well known that plasma exposure may cause degradation of device performance and decrease device yield and reliability. In general, many processing steps including polysilicon etch [79–81], ion implantation [82], dielectric deposition, oxide contact and via etch [83], metal interconnect etch [84], and resist ashing [85] may each contribute to device damage. Damage could happen by the synergistic effect of ion bombardment [86], charging due to plasma non-uniformity [87–89], contamination [90], and ultraviolet radiation [91]. Here, we focus on charging damage. Charge imbalance is the main cause responsible for plasma charging damage. The local ion and electron fluxes are out of balance due to charge imbalance, and result in charge built-up. When voltage due to charge built-up is sufficiently large, F–N tunneling occurs. Current collected by the antenna structure is channeled through the thin gate oxide by F–N tunneling. Under such situations, electron current injected into the oxide may deposit energy in the oxide and lead to trap creation and interface state generation. Consequently, gate oxide integrity is degraded. However, the damage depends on many factors from different process conditions and device applications.

Fig. 12 shows the \( Q_{\text{bd}} \) results as a function of device location and antenna area ratio for oxides with thickness ranging from 8 to 2.6 nm. Charging damage is induced by a photoresist ashing step after wet metal pad definition. For devices with large antenna area ratio (AAR, e.g., 10K), significant damage begins to appear at the wafer center as oxide thickness is scaled below 6 nm. For oxide thinner than 4 nm, oxide breakdown is induced at the wafer center. These results suggest that plasma charging damage is strongly dependent on thickness and device location.
6.2. Evaluation of plasma charging damage

When entering sub-quarter micron era (i.e., $T_{ox} < 6$ nm), the dependence of plasma charging damage on oxide thickness presents an important and controversial topic. Park and Hu studied the damage induced in oxides ($2.2 \text{ nm} < T_{ox} < 7 \text{ nm}$) during metal and contact etching processes, and concluded that thinner oxide has superior immunity [21]. Alavi et al. showed that, as oxide is thinned down, the damage increases up to a point ($\sim 4 \text{ nm}$), and then decreases due to directing tunneling [22]. Similar trends were also reported by Noguchi et al. in investigating the electron shading effects [23]. On the other hand, the results of Bayomi et al. [(24), $T_{ox}$ range: 8–4 nm], Krishnan et al. [(25), $T_{ox}$ range: 6–3.5 nm] and Chien et al. [(92), $T_{ox}$ range: 8–4 nm] showed that the susceptibility to damage increases with decreasing oxide thickness. More recently, Krishnan et al. further indicated that severe damage could be induced in gate oxide as thin as 2.1 nm under certain inductively coupled plasma (ICP) metal etching conditions [93]. These different findings are understandable since process conditions and equipment configurations can be very different from one study to another. Besides, oxide degradation characteristics under high field stressing may change significantly as $T_{ox}$ is thinned down. Thus different indicators (e.g., $Q_{bd}$, breakdown field, threshold voltage ($V_{th}$), etc.) that are used to characterize the damage by different investigators may lead to different outcomes.

Many test vehicles have been traditionally used in evaluating plasma-induced antenna effects, including $Q_{bd}$ [94], initial-electron-trapping-rate (IETR) [95], relative linear transconductance reduction [96], hot-carrier-injection (HCI) [97–100], stress-induced leakage current (SILC) [101,102], and charge pumping (CP) [103,104], etc. However, as oxide is scaled into direct tunneling regime, some traditional measurements may become insensitive in detecting the charging damage. Fig. 13 depicts threshold voltage, subthreshold swing and transconductance of transistors as a function of cell location. We can see that, even with a large AAR value, these parameters vary only slightly across the wafer. However, significant damage is actually identified by the $Q_{bd}$ measurements, as illustrated in Fig. 14. The insensitivity of transistor parameters to charging damage could be explained by the negligible surface state generation and bulk trapping after plasma process, which is intrinsic to ultrathin oxides. Therefore, one would have to conclude that some traditional methods are no longer sensitive for detecting charging damage for ultrathin oxides. Other methods, such as charge-to-breakdown, gate leakage current and noise measurement, are more sensitive and, therefore, more suitable for characterizing the plasma charging effects. In fact, for deep submicron devices, the situation may probably become more complicated. Thus, evaluation of plasma charging damage of ultrathin oxide device should be carefully examined.
6.3. Methods to improve plasma charging immunity

Since plasma charging damage could cause serious device degradation, several approaches are employed to alleviate it. These approaches include using protection diode [105, 106], setting tight design rules for circuit, and introducing robust gate dielectrics [107]. Protection diode may be the most effective method to eliminate charging damage. However, additional layout/process and loss of integrity are its main drawbacks. The resort to design rule, on the other hand, limits the circuit layout flexibility. Finally, N$_2$O-nitrided oxide can be employed to effectively suppress the charging damage. In contrast with pure oxide, the leakage current characteristics of antenna devices with nitrided oxide are significantly improved, as shown in Fig. 15. Only slight increase in gate leakage current is observed on antenna devices with nitrided oxide. These improvements can be attributed to the formation of strong Si–N bonds replacing strained Si–O bonds and weak Si–H bonds [107]. The enhanced interface hardness of nitrided oxide results in improved gate oxide integrity. Therefore, N$_2$O-nitrided oxide is an extremely effective approach to improve the immunity to plasma damage in ultrathin oxides.

7. Conclusion

The continued scaling of ULSI technologies has demanded immediate attention to reliability issues of ultrathin oxides. In this work, we have reviewed recent status on the issues relating specially to the reliabilities of ultrathin gate dielectrics for present and future ULSI technologies. Several challenges facing the application of ultrathin gate oxides are discussed and can be summarized as followed:

1. The quantum mechanical tunneling current will eventually become a scaling limitation for the continued brute-force scaling of thermal silicon dioxide as gate dielectric. Alternative gate dielectrics with lower leakage current and comparable performance have to be introduced before the thermal oxide eventually becomes the show-stopper.

2. Ultrathin gate oxides exhibit superior TDDB characteristics. However, its increased dependence on temperature and polarity should be carefully taken into consideration in process integration, especially for processing steps that require high temperature.

3. Better understanding of breakdown mechanism of ultrathin oxide is essential for device lifetime prediction under normal operation condition, and for choosing proper reliability evaluation indicators.

4. The conflicting requirements in thermal cycle for poly depletion effect and boron penetration will severely restrict the process window. Alternative gate dielectrics and electrode materials may be inevitable in order to resolve this issue.

5. Plasma charging damage remains a major concern for ultrathin oxide reliability. More efforts need to be made to understand and minimize the charging damage.

References


tor with TiN gate technology for 0.1 μm MOSFETs. Symp VLSI Technol Dig Tech 1997:135.


[62] Lee WC, King TJ, Hu C. Optimized poly-Si₁₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋₋˓


