Fiabilité des oxydes de grille ultra-minces sous décharges électrostatiques dans les technologies CMOS fortement sub-microniques

Adrien Ille

N° d'ORDRE : 2008AIX11040

THESE

présentée à

l'Université d'Aix-Marseille I

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE

Ecole doctorale **: Physique, Modélisation et Sciences pour L'Ingénieur**

Spécialité : **Micro et nano électronique**

par

Adrien ILLE

Ingénieur ISEN

FIABILITE DES OXYDES DE GRILLE ULTRA-MINCES SOUS DECHARGES ELECTROSTATIQUES (ESD) DANS LES TECHNOLOGIES CMOS FORTEMENT SUB-MICRONIQUES

Soutenue le 16 Juin 2008 devant la commission d'examen :

Président :	**Rachid BOUCHAKOUR**	**Directeur IM2NP**
Rapporteurs :	**Marise BAFLEUR**	**Dir. Recherche LAAS-CNRS**
	Pascal NOUET	**Pr., Université Montpellier II**
Directeur de Thèse :	**Alain BRAVAIX**	**ISEN-IM2NP**
Examinateur :	**Wolfgang STADLER**	**INFINEON**
	André TOUBOUL	**DRRT**

Im2np Institut Matériaux Microélectronique Nanosciences de Provence
UMR 6242 CNRS, Universités Paul Cézanne, Provence et Sud Toulon-Var

"Ein Prosit der Gemütlichkeit"

;-)

Acknowledgements

The work presented in this manuscript was obtained within the ESD laboratory of Infineon Technologies AG (IFX) located in Munich, Germany. This work has been funded by the European Medea+ project SIDRA (T-104) and realized with the collaboration of the Institut Matériaux Microélectronique Nanosciences de Provence (IM2NP), CNRS - UMR 6242, Institut Supérieur d'Electronique et du Numérique (ISEN) Toulon, France.

I would like to express my gratitude to my PhD thesis director Alain Bravaix (IM2NP & ISEN Toulon) for his interest in the crossed topics of ESD and reliability and I would like to acknowledge his high expertise in the advance CMOS technologies physics and reliability themes.

I would like to thanks particulary my PhD thesis advisor Dr. Wolfgang Stadler (IFX - Principal I/O & ESD Verification Methodology) for his guidance and his wide knowledge and experience in the field of ESD. I learned a lot from the many fruitful discussions all along the thesis. I am looking forward to be involved in common projects (or biking tours!) again in the future.

I want to thanks the jury, Rachid Bouchakour, Marise Bafleur, Pascal Nouet and André Touboul for their review and interest in this PhD work.

I am thankful to the Lib I/O laboratory team. First, a big thanks to Dr. Tilo Brodbeck (IFX - Interfaces, ESD & Latch-up) I appreciated the technical discussions and feedbacks as well as his assistance within this SIDRA project. I want to thank also Josef Niemesheim (IFX - IO & ESD testing Solutions) for his support in the measurements.

I want to cite the great link generated with the reliability community at IFX and their deep interest in the topic. I will not forget the strong motivation and very interesting discussions emanating from Dr. Andreas Kerber (IFX - dielectric reliability, currently at AMD). I would also like to address a special big thank to Dr. Thomas Pompl (IFX - Principal dielectric reliability), for his high expertise knowledge in the dielectric reliability and for his really good and continuous collaboration. I really acknowledge is high availability and enthusiasm to share his sharp knowledge. I have to recognize that his feedbacks and rigorism approach have always been very useful.

I would like to thanks Dr. Kai Esmark (IFX - ESD) which has with Dr. Wolfgang Stadler permit me to get familiar and interested in the ESD topics during my enclosure Master internship.

His support and advices along this thesis work was very profitable.

A big thanks to Dr. Harald Goßner (IFX - Head of ESD) whose interest in research works and support for internal training has been very beneficial. I especially appreciated his wide overview on ESD topics and his particular interest on practical case studies.

A BIG BIG thanks to the ESD community of IFX whose intense support and participation to PhD seminars is greatly acknowledged. I want to thanks Dr. Harald Goßner, Dr. Martin Wendel, Dr. Jens Schneider, Dr. Krzysztof Domanski, Dr. Gernot Langguth, Dr. Kai Esmark, Dr. Christian Russ, Mr. Franz Zängl. Thanks for the pleasant atmosphere and the collaboration.

I want to express my gratitude also to the LIB IO group and to my direct colleagues of box, Matthias knopf, Joachim Singer and Hafid Layachi who have been manifold solicited for any kind of topics!

I would like also to thanks the particular good knowledge sharing from I/O and ESD groups and the strong PhD students connection which have been developed within the group. Beyond the friendships, this have been very profitable from the seminars and the learning. Thank you Wolfgang Soldner, Enrico Franell, Moon-Jung Kim, Dr. Ulrich Glaser and Dr. Stephan Drueen. And god bless the Friday Burger event!

I would like to thanks Dr. Peter Baumgartner (IFX) for having given me the chance to access his modeling laboratory facilities. I would naturally thanks also Dr. Philipp Riess (IFX) for the discussions on the oxide degradation mechanisms and particulary on the SILC. Domagoj Siprak (IFX) and Antonius Koller are acknowledged for their support in the charge pumping measurements.

I would like also to thank the interest of the independent CMOS Bulk Alliance at IBM Hopwell Junction in this reliability topic. I want particulary to thanks Dr. David Alvarez (IFX), Dr. Robert Gauthier (IBM) and Dr. Kiran Chatty (IBM). Thank you for the good connection, silicon facilities and feedbacks!

I would like to express my gratitude to my friends, my family and the "french mafia" of Munich simply for their pleasant company, help and happiness to enjoy the living and relaxing the atmosphere. Last but not least I would like to thanks a lot my Céline for her patience and intensive support! Bises.

Résumé

I - Introduction

Ce mémoire traite un sujet de fiabilité spécifique aux composants électroniques modernes. Afin de mieux comprendre la problématique étudiée dans ce travail, une introduction sur le contexte industriel du secteur de l'électronique va tout d'abord être développée. Depuis le XX's l'impressionnante entrée de l'électronique dans nos modes de vie est très marquée. Cette omniprésence de l'électronique a pour origine un boom technologique basé sur les technologies semiconducteurs. Ce rapide avancement est dû tout d'abord au coût très faible de ces matériaux. Cependant le progrès technologique comporte deux principales forces motrices, la miniaturisation des composants [1] ainsi que la diversification des applications et fonctionnalités [2]. Jusqu'à présent le marché du semiconducteur est regi par la loi de Moore [1] qui annonce une complexité doublant tous les deux ans. La réduction des composants impliquant une diminution du prix des produits permet de financer la recherche ainsi que de dégager de larges bénéfices. Cependant, cette course à l'intégration des composants entraîne une forte complexité des solutions techniques à développer pour continuer le processus. Le revers du succès de cette poussée vers la miniaturisation est la sensibilité accrue des circuits intégrés (ICs) aux problèmes de fiabilités. Les décharges électrostatiques (ESD) constituent une des principales menaces nuisant à la fiabilité des composants [3, 4]. Depuis plus de 30 ans, la protection des ICs envers ces phénomènes naturels pose un réel défi pour les fabricants. Pour enrayer les défauts générés par les ESD sur les ICs, des éléments de protection sont implantés directement dans les puces. Sans ces éléments de protection, les ICs fabriqués dans les procédés technologiques récents de type Complementary Metal Oxyde Semiconductors (CMOS) pourraient difficilement éviter les décharges ESD, ce qui engendrerait des dommages irréversibles aux circuits ainsi qu'un rendement de production quasiment nul. De plus, la fiabilité des produits se doit d'être garantie hors des sites protégés pour les utilisateurs finaux et ce sur le long terme. Les cellules CMOS élémentaires (transistors à effets de champs nFET et pFET) comportent un film diélectrique isolant intervenant dans le contrôle des caractéristiques électriques du composant. Cette couche d'oxyde est principalement le dioxyde de silicium (SiO_2) ou un dérivé nitruré (SiO_xN) grâce à leurs facilités de fabrication et à leurs propriétés isolantes. La miniaturisation des ICs provoque la réduction des dimensions des transistors ainsi que la réduction de cette couche

d'oxyde. Dans les technologies avancées l'épaisseur de ce film d'oxyde (T_{ox}) approche maintenant la limite de 10 Angstrom (4 mono couches atomiques). Il est donc facile d'envisager la sensibilité extrême de ces couches d'oxyde minces vis à vis de décharges délivrant plusieurs Ampères et causant d'énormes surtensions. Dans ce contexte les conditions restrictives imposées par les procédés technologiques et par la complexité croissante des systèmes (mixité des blocs analogiques et numériques, multiplication des domaines d'alimentation, performances, applications haute fréquence, température,...) entraînent un défi considérablement accru pour le développement de produits robustes aux ESD.

Dans ce travail de recherche, le problème émergeant des défaillances des couches d'oxyde minces sous contraintes ESD dans les technologies CMOS avancées est étudié à travers une contribution à la compréhension des mécanismes de dégradation du diélectrique et des transistors MOSFETs soumis aux contraintes ESD. Cette étude a pour but d'améliorer les protocoles de conception de protections ESD ainsi que l'optimisation des designs afin de garantir la robustesse des produits dans une démarche de qualité approfondie tout en réduisant les coûts de sur-design.

II - Structure de la thèse

Une introduction sur le contexte industriel présentant les motivations et les attentes de ce travail de recherche amène le sujet de cette thèse articulée autour de six chapitres. Les deux premiers chapitres introduisent le contexte et l'état de l'art de la fiabilité et du développement des protections ESD dans les technologies CMOS avancées. Les conditions expérimentales sont décrites dans le chapitre 3. Les chapitres 4 et 5 constituent le corps expérimental de la thèse. Le chapitre 6 apporte l'application des résultats au design sécurisé des dispositifs MOSFETs pour mieux définir la protection ESD sur silicium.

Chapitre 1

Le premier chapitre constitue une introduction à la fiabilité du diélectrique utilisé dans les technologies MOS (Métal Oxyde Semiconducteur). Dans le but d'introduire les notions de base requises par la suite dans le manuscrit, un rappel de la physique intervenant dans la structure MOS est présenté jusqu'à l'exposition des paramètres caractéristiques des transistors MOS et des modes de courant de fuite à travers l'oxyde. La fiabilité du diélectrique des circuits intégrés sous contrainte électrique est développée dans ce chapitre en abordant les mécanismes et les aspects statistiques liés à la dégradation des oxydes. Finalement, les différents modèles d'extrapolation de la durée de vie des oxydes présents dans les circuits intégrés seront présentés.

Chapitre 2

Dans le chapitre 2, une vue générale sur les phénomènes électrostatiques et leurs enjeux dans le milieu industriel est développée. Le contexte ainsi que les lourdes conséquences infligées par les ESD aux entreprises de composants semi-conducteurs sont détaillés depuis les menaces engendrées par les décharges électrostatiques sur les circuits intégrés jusqu'à la qualification de la robustesse des produits conformément aux modèles électriques de test standards. Ceux-ci sont développés jusqu'aux méthodes de caractérisation ESD nécessaires lors du développement des schémas de protection ESD. La solution de la protection ESD intégrée sur IC sera exposée à travers une introduction aux éléments de protection ainsi que sur les stratégies et les contraintes rendant délicate la protection des ICs. Ce sujet sera détaillé avec la présentation de l'évolution des contraintes résultant de l'intégration croissante des ICs. Nous montrerons que la diversité des applications provoque une forte augmentation du problème ESD avec l'intégration des circuits. Il est vital de contrôler ce problème croissant à l'aide de designs de circuits optimisés pour la réalisation de produits robustes vis à vis des modes de défaillance induits par les ESD. Plus précisément, nous soulignerons l'importance des limitations induites par la réduction de l'épaisseur d'oxyde ainsi que de son intégrité comme couche isolante pour la grille des transistors utilisée dans les technologies CMOS les plus récentes.

Chapitre 3

Les technologies CMOS étudiées dans cette thèse ainsi que les détails concernant le banc expérimental de mesure et les structures de tests sont exposés dans le chapitre 3. L'établissement d'un banc de mesure fiable et automatisé constitue la première base propice aux analyses approfondies des résultats expérimentaux.

Chapitre 4

Dans le chapitre 4, après un bref résumé bibliographique sur les travaux de modélisation du claquage des oxydes de grille dans le régime des nanosecondes, une nouvelle approche de caractérisation est développée pour évaluer le claquage des oxydes dans le domaine de temps des décharges électrostatiques. La vérification de la statistique intervenant dans la dégradation des oxydes de grille mince selon la théorie de percolation [5] est exposée. L'impact de la procédure de test ainsi que des structures sur l'évaluation de la dépendance temporelle du claquage est quantifié. Une méthodologie de stress plus précise a été développée pour la détermination du temps de claquage. Dans une deuxième partie, le claquage de l'oxyde de grille depuis les conditions de stress statiques (DC) jusqu'au régime temporel des ESD est caractérisé par la modélisation de l'accélération du temps au claquage en fonction de la tension de stress.

Chapitre 5

Dans le chapitre 5, l'impact des stress ESD non destructeurs sur les oxydes de grille est étudié en fonction de l'épaisseur de l'oxyde de grille. Une connaissance précise de la dégradation de l'oxyde et des dispositifs induite par des stress ESD modérés est requise pour la garantie d'une conception sûre et fiable. Les impacts de la dégradation des oxydes sous contraintes ESD ainsi que sur la fonctionnalité des dispositifs et leurs durée de vie sont reportés via des caractérisations DC et des cycles d'injections de porteurs chauds sur des transistors nFETs.

Chapitre 6

Le chapitre 6 présente la synthèse des résultats obtenus concernant la fiabilité du diélectrique (chapitre 4) et des dispositifs (chapitre 5) pour la conception de protections ESD fiables. Une sélection des critères de génération de défauts est discutée pour la détermination de la fenêtre temporelle de design ESD. Dans ce chapitre une procédure concernant les critères de conception vis à vis des critères de défaillance des oxydes est présentée. L'analyse des travaux de caractérisation réalisés est appliquée à la contribution d'un flow et d'un outil de développement utile pour les ingénieurs ESD dans leur travail. Pour la première fois, les défauts et la dégradation des couches minces d'oxyde sous contraintes de type ESD peuvent être inclus quantitativement et de manière systématique dans le développement des dispositifs associés à la protection ESD.

III - Synthèse des principaux résultats au cours de la thèse

III-A - Nouvelle méthodologie de caractérisation des oxydes de grille minces dans le régime ESD

La première motivation de notre étude concerne la dépendance temporelle du claquage du diélectrique (TDDB) en tentant de répondre à la question qui reste toujours ouverte sur l'origine des mécanismes de claquage du diélectrique mince sous contraintes électriques à tensions constantes (CVS) pour des conditions de forts champs et pour des temps ultra-courts. Ceci représente en effet une des caractéristiques des décharges électrostatiques.

La forte augmentation du phénomène de claquage dans les oxydes minces induit par les décharges électrostatiques dans les technologies actuelle [3, 4, 6, 7] génère le besoin de comprendre avec exactitude ces mécanismes et leurs facteurs d'accélération. La fiabilité des oxydes de grille est le sujet d'une intense recherche dans les régimes de basse tension à l'aide de stress à long terme [8, 9, 5, 10, 11, 12, 13]. Cependant ce thème est toujours ouvert dans le domaine des ESD et l'on peut se demander si les différents mécanismes de claquage ont une même origine. Un apport majeur de nos travaux est l'étude du claquage intrinsèque des oxydes minces sur une large gamme d'épaisseur d'oxyde, de 7 nm jusqu'à 1.1 nm, en fonction du temps de stress dans un intervalle qui varie du régime DC jusqu'aux nanosecondes. Contrairement aux rares études menées jusqu'à présent pour la modélisation du TDDB vers le régime temporel des ESD [14, 15, 16], une approche purement expérimentale a été effectuée de manière continue sur la totalité du domaine temporel et des épaisseurs d'oxyde de grille. Nous avons ainsi établi avec précision l'évaluation des paramètres d'accélération du modèle de claquage des oxyde minces basé sur 8 différentes épaisseurs d'oxyde et sur un vaste régime temporel couvrant 16 ordres de grandeur. Une partie essentielle de ce travail a donc été une pré-étude portant sur l'influence de la méthodologie de stress et des structures de test ainsi que sur l'extraction exhaustive des paramètres ayant un impact sur le claquage des oxydes.

III-A-1 - Banc expérimental et méthodologie de mesures

Le banc expérimental mis au point pour la caractérisation des oxydes de grille est basé sur un générateur de pulses Agilent (8114A) pouvant générer des stress rectangulaires de 20 ns jusqu'à 1 s, Figure 1. Le système de mesure utilise une méthodologie indépendante de stress et de mesure de type Kelvin. Une méthodologie de stress CVS (stress à tension constante) est appliquée pour la détermination du TDDB plutôt que la méthode traditionnelle VRS (rampe de tension à durée de stress constant) classiquement utilisée dans le cadre de caractérisation ESD. Les échantillons ont été stressés en configuration capacité avec la source, le drain et le substrat (bulk) connecté ensemble afin d'opérer un stress homogène sur l'oxyde de grille.

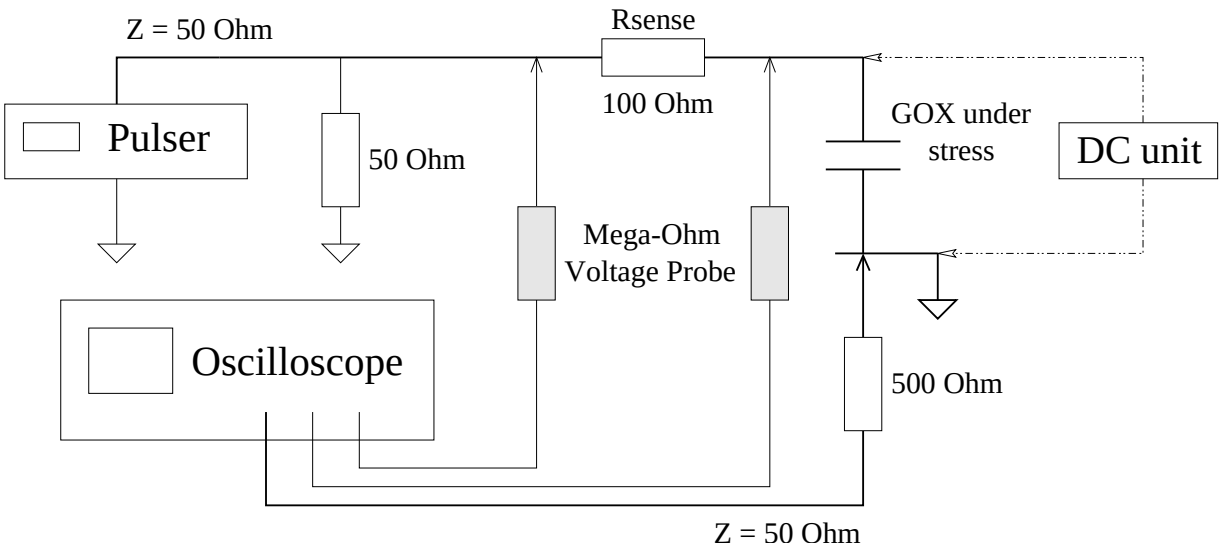

Figure 1: Configuration du banc de test utilisé pour la caractérisation de la dépendance temporelle du claquage du diélectrique.

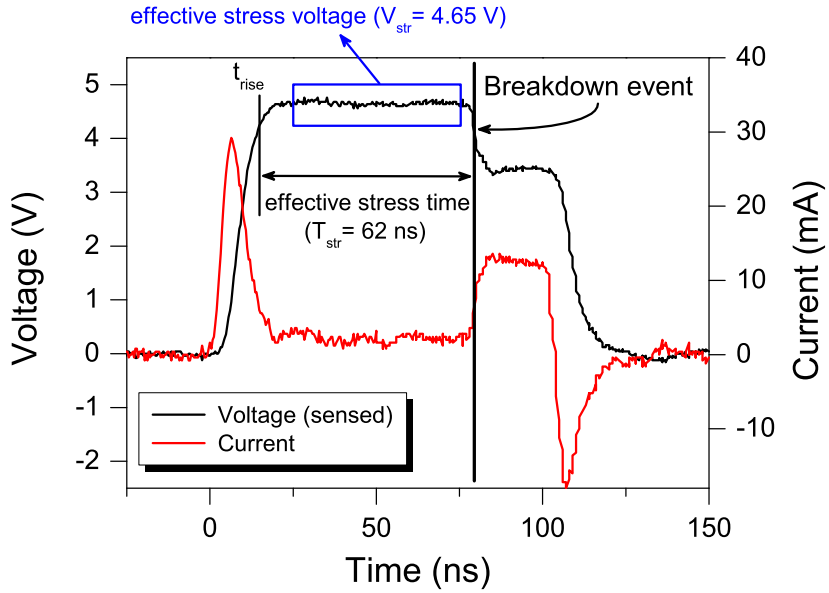

Figure 2: Forme d'onde mesurée à la grille pour un stress à 4.65 V. La détection de l'événement du claquage durant le pulse est présentée pour la détermination du temps effectif de claquage ($T_{\mathrm{BD,eff}} = 62$ ns). Les pics de courant observés au début et à la fin du pulse sont induits par la réflexion de la ligne de transmission.

La détermination du temps au claquage de l'échantillon est extraite directement de la tension mesurée à travers l'oxyde. Le claquage de l'oxyde est observé Figure 2, dans la chute brutale de cette dernière tension correspondante à la perte du caractère isolant du diélectrique pendant

l'application du stress. Toutefois l'évolution du courant de grille est mesurée après chaque stress pour la détection d'une éventuelle dégradation ainsi que pour la confirmation du claquage.

III-A-2 - Modélisation de la nature statistique du claquage des oxydes de grille minces

Des études statistiques sur un large nombre de structures ont permis de vérifier la continuité du modèle statistique de la dégradation de l'oxyde de grille proposé par le modèle de percolation [5] dans le domaine de temps de la nanoseconde. Ce modèle consiste à décrire la création d'un chemin de percolation à travers la couche d'oxyde basé sur la génération aléatoire de défauts dans l'oxyde. La probabilité cumulée de génération de défauts en fonction du temps de claquage est décrite par la loi statistique de Weibull [8], comme le montre la Figure 3. Cette loi s'exprime selon l'équation décrite ci-dessous,

$$F(t) = 1 - exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \text{ où } F(t = \tau = T_{63\%}) = 0.632 \tag{1}$$

Afin de tracer des fonctions linéaires l'équation (1) est ré-écrite sous la forme suivante:

$$ln(-ln(1-F)) = \beta\, ln(t) - \beta\, ln(\tau) \tag{2}$$

Ce nouveau résultat est primordial pour le design des éléments de protection ESD car la nature statistique de la dégradation des oxydes a jusqu'à maintenant complètement été négligée. Dans la même optique la dépendance du TDDB en fonction de la surface de l'oxyde stressé a été étudié. La statistique de Poisson modélise la dépendance en surface du temps au claquage dans le régime DC. Cette statistique s'exprime comme,

$$P(n) = \frac{(AD)^n}{n!}\ e^{-AD} \tag{3}$$

L'équation (3) combinée avec la statistique de Weibull (équation (2)) amène une puissante loi de normalisation en surface,

$$ln\left(\frac{A_{\mathrm{i}}}{A_{\mathrm{ref}}}\right) = -\beta \cdot (ln(T_{\mathrm{i}}) - ln(T_{\mathrm{ref}})) \tag{4}$$

Comme démontré dans la Figure 4, la dépendance surfacique au claquage est conforme à la statistique de Poisson dans le régime des stress de courte durée de type ESD. En effet la normalisation des distributions à une surface de référence d'oxyde de 0.024 μm^2 présentée dans la Figure 4 vérifie l'équation 4.

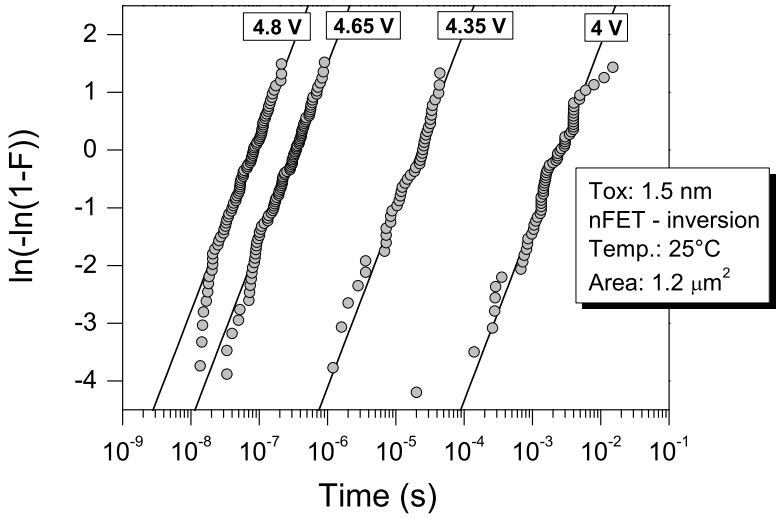

Figure 3: Distributions des claquages d'oxyde cumulés en fonction du temps de défaillance obtenues pour 4 différents niveaux de stress à tension constante sur des dispositifs nFET de 1.5 nm d'épaisseur d'oxyde de grille à température ambiante.

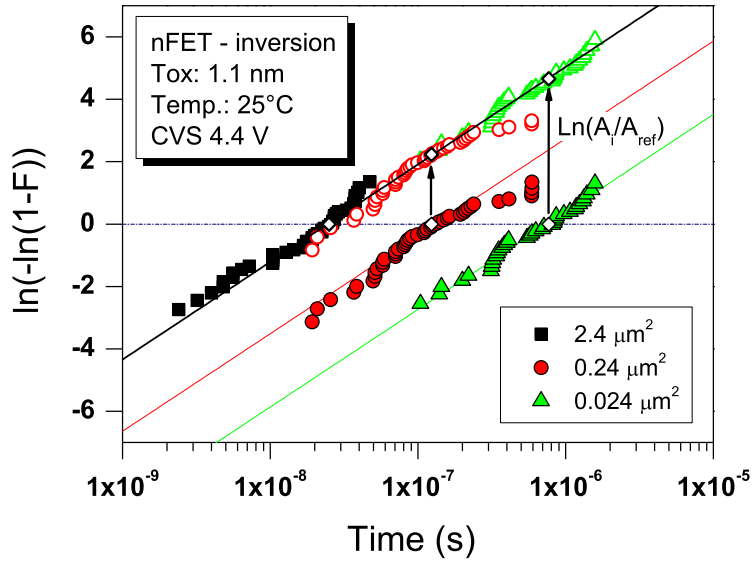

Figure 4: Les distributions de temps au claquage de dispositifs nFET de 1.1 nm d'épaisseur d'oxyde de différentes surfaces d'oxyde de grille sont tracées pour des stress à tension constante (CVS) à 4.4 V et 25 °C.

Contrairement aux idées reçues, pendant un stress ESD, le claquage intrinsèque des oxydes minces en configuration capacité ne résulte pas de phénomènes localisés dus à l'application de fort champs confinés ou à des effets de bords. Dans le cas de décharges électrostatiques, l'hypothèse de la génération d'un chemin de percolation dans l'oxyde de grille décrit comme un phénomène stochastique reste valable, impliquant une probabilité de défaillance aléatoire du film d'oxyde en fonction de la surface active de l'oxyde stressé.

Ceci démontre la sensibilité accrue des surfaces d'oxyde importante vis à vis des sur-tensions mises en cause lors d'événements ESD.

D'autre part une étude sur les différentes procédures de caractérisation a révélé directement l'aspect dangereux de la nature cumulative de la dégradation de l'oxyde à travers des stress répétitifs. Une comparaison entre les différentes méthodes de stress réalisée sur de larges statistiques d'échantillons a permis de distinguer la réduction du temps au claquage induite par ces effets de stress cumulés.

III-B - Modélisation du temps au claquage des oxydes de grille minces vers le régime ESD

Les effets induits par les diverses méthodologies de stress ainsi que par les effets parasites de layout sont présentés dans le chapitre 4. Pour la caractérisation du claquage de l'oxyde, les règles proposées par les laboratoires de recherche sur la fiabilité des procédés technologiques dans le long terme restent valables, à savoir l'utilisation de petites structures de test avec un layout optimisé concernant les résistances parasites [17, 18, 19]. Basé sur des stress CVS appliqués sur de nombreux échantillons (minimum 45 structures par niveau de stress) dans le domaine DC et ESD, une corrélation des temps aux claquages prélevés pour 63,2% de défaillance cumulé ($T_{63\%}$), a été établi avec les données obtenues pour des stress de longue durée comme le présente la Figure 5. La continuité du modèle d'accélération en tension du claquage des oxydes minces (< 7 nm) est observée sur 14 ordre de grandeurs en temps. Le claquage des oxydes minces est contrôlé par la tension selon une loi de puissance [13],

$$t = t_0 \left(\frac{V}{V_0} \right)^{-n} \tag{5}$$

Contrairement aux précédents travaux reportés sur les oxydes minces [14, 15, 16] considérant le champ (modèle en E) ou l'inverse du champ (modèle en 1/E) comme paramètres responsables de l'accélération du temps au claquage, la caractérisation approfondie et continue qui a été menée dans cette thèse permet de réfuter ces dernières hypothèses. Les résultats obtenus démontrent l'incapacité pour ces deux derniers modèles de décrire le TDDB des oxydes minces. Ceci est clairement exposé dans la Figure 6, où les trois modèles sont tracés pour la modélisation

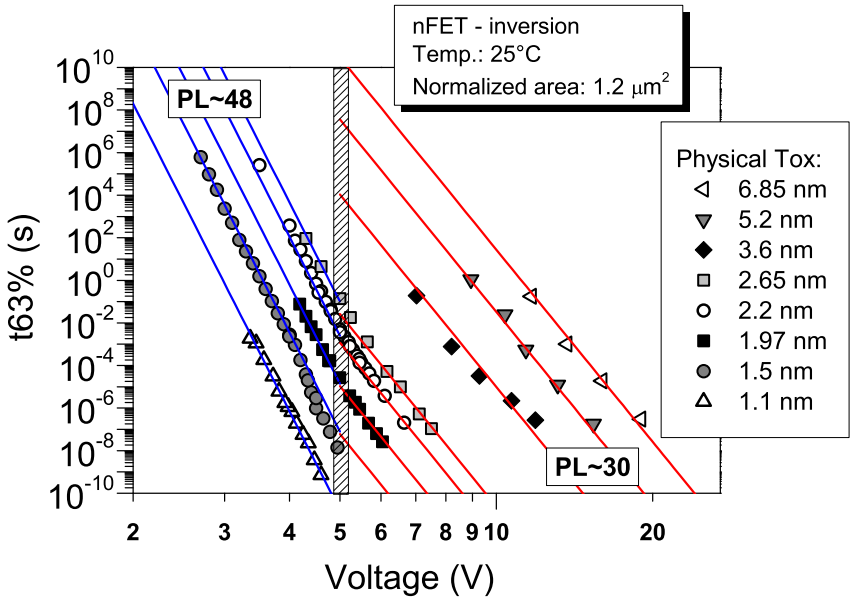

Figure 5: Accélération en tension des dispositifs nFET stressés en régime d'inversion à 25 °C normalisée à la même surface d'oxyde de grille de 1.2 µm^2 [20].

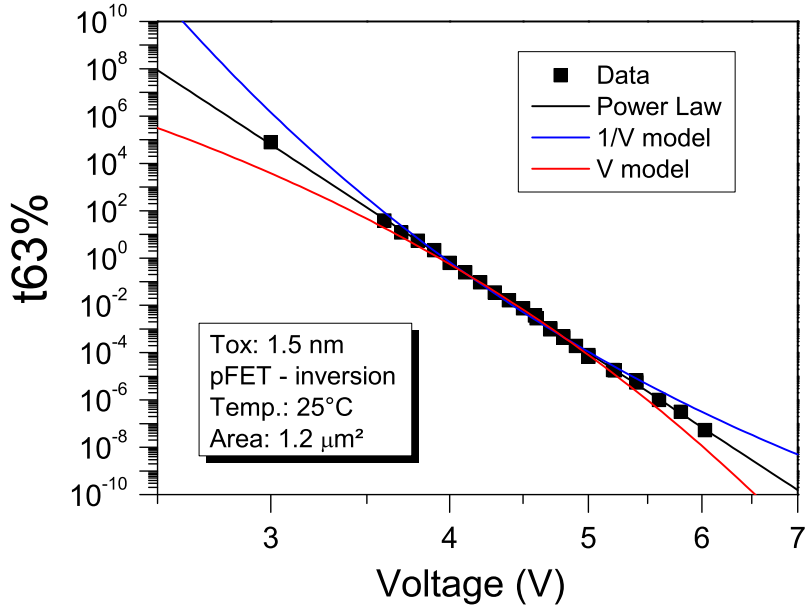

Figure 6: TDDB des dispositifs de type pFET pour une épaisseur d'oxyde de 1.5 nm stressé en inversion à 25 °C [21].

du TDDB de dispositifs pFET ayant un diélectrique T_{ox} = 1.5 nm. Les modèles en 1/E et E sont tracés dans leur forme d'accélération en tension, respectivement $t \cdot exp(G/V)$ et $t \cdot exp(-\alpha \cdot V)$.

Une caractérisation complète du TDDB en fonction du type du dispositif et en fonction de la polarité permet d'établir la généralisation de la description du claquage de l'oxyde à travers une loi de puissance possédant deux différents domaines d'accélération. Dans les résultats, un coude est très clairement aperçu aux alentours de 5-6 V comme le montre la Figure 5. Une explication physique sur la valeur de cette tension de transition est discutée en fonction des différents modes de dégradation de l'oxyde basé sur la libération de l'énergie maximale des électrons à l'anode de l'isolant stressé. En dessous de cette valeur, le facteur d'accélération (exposant n de la loi de puissance) est élevé impliquant une sensibilité encore plus accrue des oxydes minces vis à vis des stress ESD. La nature cumulative de la dégradation des oxydes est quantifiable grâce à cette modélisation du claquage par la loi de puissance en tension. L'intensité de l'impact d'un stress sur l'oxyde est dépendant du facteur d'accélération. Nous avons montré que le niveau d'un stress ESD au delà de 90 % de la tension de claquage induit une réduction significative de la durée de vie de l'oxyde de grille.

Une hiérarchisation de la robustesse en fonction du type du dispositif et de la polarité est observée comme le montre la Figure 7. Le cas des dispositifs nFETs stressés en régime d'inversion constitue le pire cas de dégradation. Ceci mène à des directives ESD directes concernant la définition des cellules standards dans les technologies CMOS sub-microniques. Par example ceci affecte tout particulièrement le choix des capacités de type buffer basées sur les oxydes minces ainsi que le choix de leurs polarisation [21]. Cette généralisation du modèle de claquage permet l'élaboration d'un outil d'extrapolation puissant et intéressant pour les développeurs de protections ESD.

Une alternative à la caractérisation dans le régime des nanosecondes peut être basée sur l'extrapolation des données acquises par les laboratoires de fiabilité qualifiant les durées de vie des procédés technologiques. Cependant les conditions de stress sont très différentes et demandent des extrapolations extrêmes en température et en tension. En effet les conditions de stress sont effectués à 140 °C pour les tests de fiabilité. Une forte dépendance de l'énergie d'activation de l'accélération en température (E_a) est observée en fonction de la tension comme l'indique la Figure 8. Le facteur d'accélération en tension du TDDB augmente en direction des faibles températures et l'activation thermique devient négligeable à fort champs. L'accélération en température du TDDB peut être modélisée par une loi d'Arrhenius avec une dépendance de l'activation en tension. Cette accélération qui est modélisée par l'équation 6 est exposée dans la Figure 9.

$$t_{BD} = t_0 \cdot \exp\left[\frac{E_a(V)}{k_B} \cdot \left(\frac{1}{T} - \frac{1}{T_0}\right)\right] \tag{6}$$

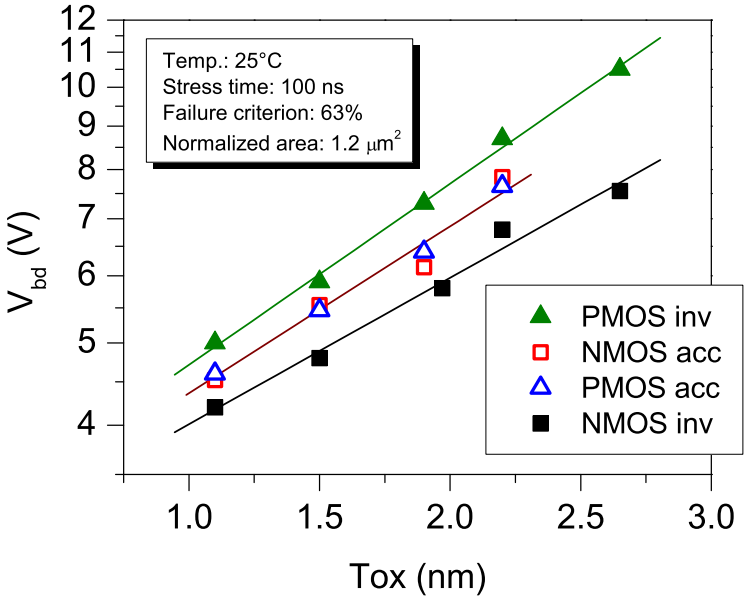

Figure 7: Accélération en tension du claquage en fonction de l'épaisseur physique de l'oxyde pour des dispositifs nFET et pFET stressés à 25 °C dans le régime HBM (∼ 100 ns pulse) pour une surface de grille normalisée à 1.2 μm^2 [20].

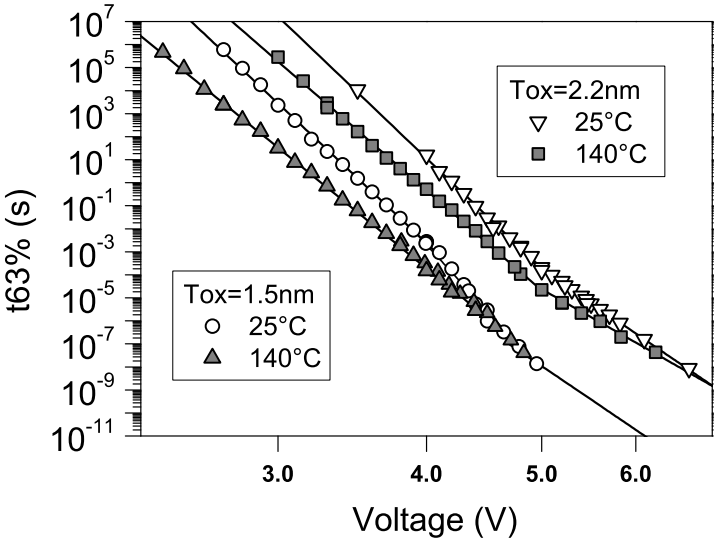

Figure 8: Accélération en tension des dispositifs de type nFET d'épaisseur d'oxyde de grille T_{ox} = 1.5 nm et 2.2 nm à 25 °C et à 140 °C [22].

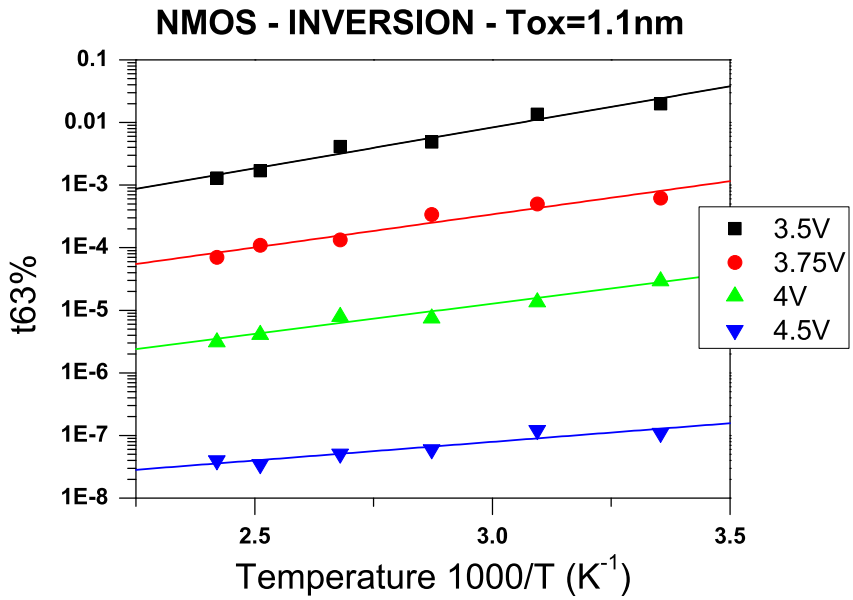

Figure 9: Temps au claquage extraits des distributions des dispositifs nFET avec T_{ox} = 1.5 nm stressés en inversion forte en fonction de la température pour 4 différents niveaux de CVS.

III-C - Dégradations de l'oxyde de grille des dispositifs sous stress ESD

Un important sujet pour la fiabilité des composants électroniques est la garantie de l'intégrité de l'oxyde de grille et du bon fonctionnement des circuits après un stress ESD. Le chapitre 5 est dédié à l'étude de l'impact de stress ESD modérés sur les caractéristiques du dispositif et à la dégradation potentielle de sa durée de vie pour des échantillons nFET stressés en inversion (*i.e.* $V_g > 0$). Nos études ont distingué les dispositifs selon leurs épaisseurs d'oxyde (T_{ox}): les oxydes épais ($T_{ox} > 7$nm), les oxydes intermédiaires, c'est à dire compris entre 3 nm et 7 nm puis les oxydes minces ($T_{ox} < 3$nm).

III-C-1 - Etude de l'intégrité fonctionnelle des dispositifs nFETs sous contraintes ESD modérées

Pour les oxydes épais à intermédiaires, un net effet des stress non destructif sur les caractéristiques des dispositifs est notable. Il se caractérise dans ce cas par une dérive de la tension de seuil (V_{th}) des transistors, l'augmentation des courants de fuite de l'état "I_{off}" correspondant à l'état 0 en entrée du transistor ($V_{gs} = 0$, $V_{ds} = V_{dd}$). Pour les oxydes minces dont les épaisseurs sont en

dessous de la distance tunnel, le piégeage de charges n'opère plus dû au direct dé-piégeage des charges par l'effet tunnel direct. Pour les oxydes de grille minces aucune dérive significative des paramètres principaux des dispositifs n'est observée. Pour les oxydes intermédiaires et épais, la dérive des caractéristiques peut se révéler fatale pour le fonctionnement de circuits analogiques. La dérive des tensions de seuil des transistors nFET en régime saturé post stress ESD (de 100 ns de durée) est montrée dans la Figure 10 pour cinq différentes épaisseurs d'oxyde. Si l'on considère qu'une modification de 10 % d'un des paramètres du dispositif représente un critère de défaillance, alors pour un niveau de stress de l'ordre de 80 % de la tension de claquage de l'oxyde (V_{bd}), le dispositif ne respecte plus le critère de fiabilité. Cette marge doit être considérée dans le design ESD.

Avec l'apparition des dérives des caractéristiques DC, une augmentation graduelle du courant de grille est observée pour les oxydes intermédiaires (T_{ox} = 3 - 7 nm). Ce phénomène, particulièrement critique pour les mémoires non volatiles, connu sous le nom de SILC (Stress Induced Leakage Current) [23, 24] est observable aussi sous condition de stress ESD. Ce courant de fuite a pour origine un mode de conduction assisté par les défauts présents dans l'oxyde et reflète donc la dégradation du volume de l'oxyde. L'accroissement du SILC peut être corrélé avec le taux de génération des défauts. Cette cinétique de génération des défauts mesurée via le SILC en fonction de la charge injectée (Q_{inj}) est décrite Figure 11 par une loi de puissance similaire à celle observé dans le régime DC [23, 25, 26]. Ceci marque encore une preuve de la continuité des mécanismes de dégradation du régime DC jusqu'au domaine des nanosecondes.

III-C-2 - Etude de la fiabilité à long terme des dispositifs nFETs sous contraintes ESD modérées

En terme de fiabilité ESD, l'éventualité de défauts latents générés par des décharges électrostatiques ne menant pas a une défaillance franche constitue un sujet critique pour la fiabilité des circuits intégrés à long terme [27, 28, 29, 30]. Afin d'évaluer l'impact de stress ESD non destructif sur la durée de vie des dispositifs, des études croisées de stress ESD et d'injections de porteurs chauds (HCI) ont été menées. Les stress HCI permettent une estimation accélérée de l'intégrité des dispositifs pour un critère de durée de vie, défini par exemple comme 10 % de réduction du courant saturé, grâce à une extrapolation de la durée de vie à partir des stress vieillissant les dispositifs [31]. Deux épaisseurs d'oxydes ont été utilisées pour ces expériences: 5.2 nm et 2.2 nm. L'influence du niveau de stress ESD sur la cinétique de dégradation en mode HCI est montrée dans la Figure 12 pour un oxyde de 5.2 nm. On observe un effet purement cumulatif du stress ESD sur la dégradation des paramètres sous stress porteurs chauds. Le grand nombre de défauts générés par le stress ESD induit initialement un décalage de la cinétique HCI de réduction en courant. Ce décalage est dépendant du niveau de stress ESD et sur le long terme, des charges injectées lors de l'événement ESD jusqu'à une phase de convergence de la dégradation relative au stress HCI vers le même niveau (20 %). Pour des stress porteurs chauds réalisé

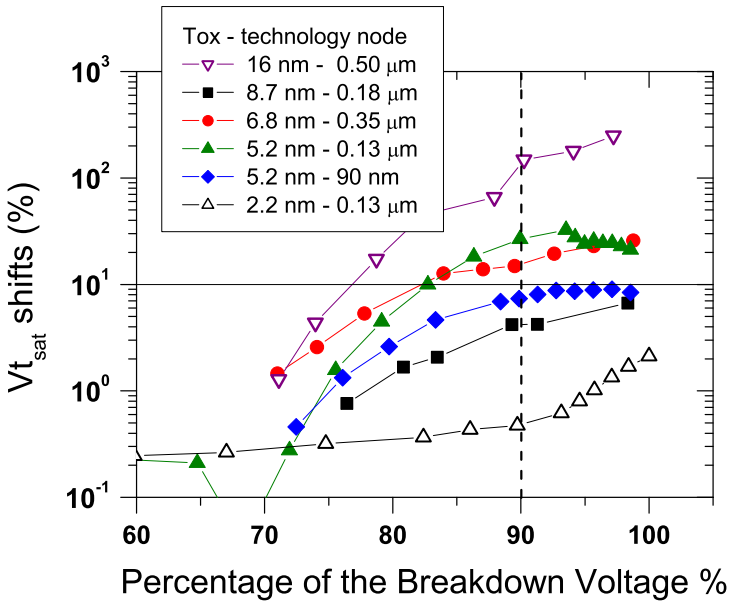

Figure 10: Dérive de la tension de seuil mesurée en mode saturé ($V_{t,sat}$) induite par une rampe de stress en tension (VRS) de 100 ns tracée en fonction du pourcentage de la tension de claquage (V_{bd}) [20].

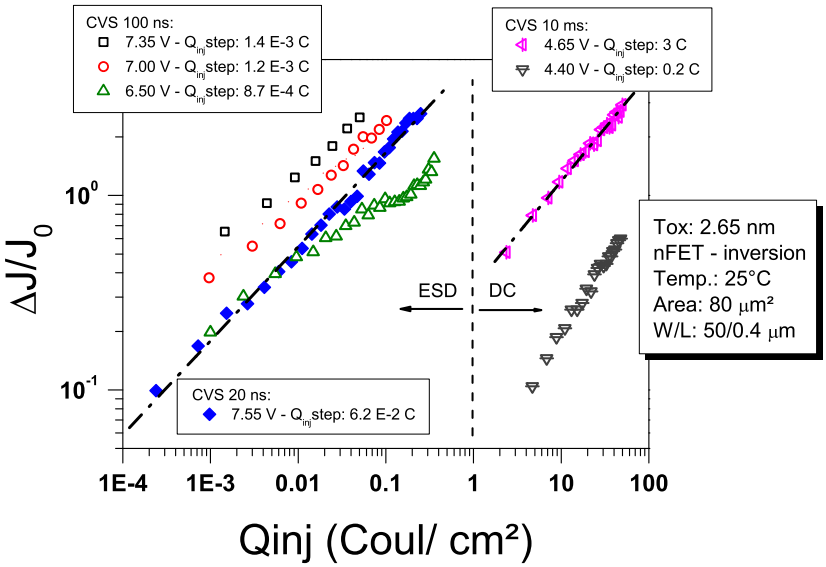

Figure 11: Augmentation des courants de fuite ($\Delta J/J_0$) de type Stress Induced Leakage Current (SILC) tracée en fonction de la charge injectée (Q_{inj}) pour des stress à tension constante de durées variables. Depuis le régime DC jusqu'aux pulses ESD courts, la cinétique de génération du SILC reste similaire selon une loi de puissance [20].

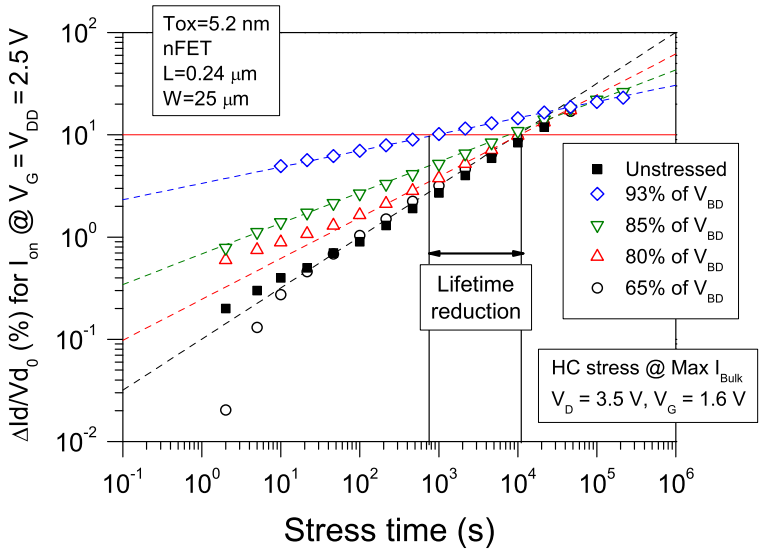

Figure 12: Impact du niveau de pré-stress ESD sur la cinétique de dégradation porteurs chauds. Un effet purement cumulatif engendré par la charge induite lors du stress ESD est observé. Pas d'effet sur le long terme est noté pour des niveaux de stress inférieure à 85 % de la tension de claquage [20].

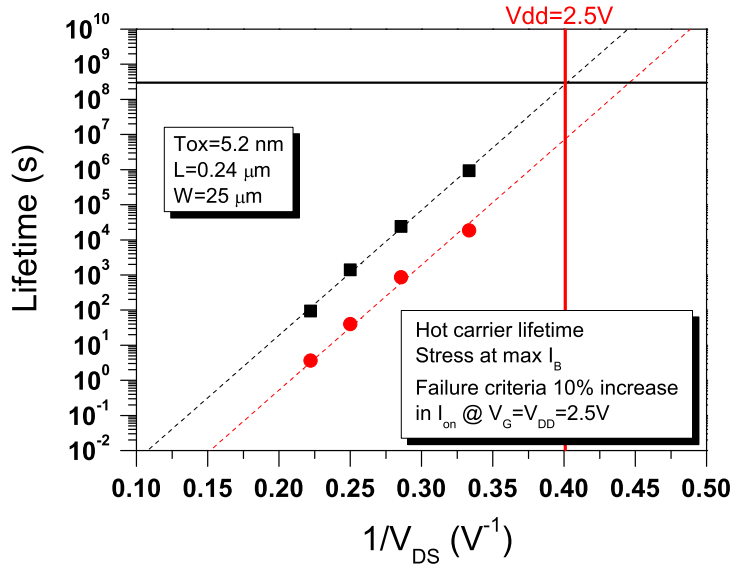

Figure 13: Tracé de durée de vie HCI pour des niveaux de pré-stress ESD à 90% de la tension de claquage. La durée de vie du dispositif est réduite de 1.5 décades en temps [20].

en DC, le critère de défaillance standard est fixé pour une réduction des paramètres de 10 %. En considérant ce critère, on observe que la durée de vie du dispositif n'est pas fortement impactée jusqu'à un niveau de stress ESD de l'ordre de 90 % de la tension de claquage. Cependant pour un stress à 90 % de V_{bd} une réduction de la durée de vie par un facteur 15 dans le long terme est observé comme l'indique la Figure 13. Dans le cas d'un oxyde mince, pas de réduction de durée de vie est notée sur le long terme comme le présente la Figure 14. La différence de comportement entre les oxydes minces et les oxydes de grille intermédiaires et épais s'explique par le rôle plus important du piégeage de charges dans l'oxyde épais à intermédiaire par rapport aux états d'interface et aux états lents (border traps) dans les oxydes minces. Ces pièges ont un impact direct sur les paramètres électriques du dispositif comme nous l'avons exposé dans la partie précédente. Une étude plus poussée doit aussi considérer les possibles phénomènes de relaxation des charges piégées sur les pièges proches de l'interface en tenant compte de leurs cinétiques de rétablissement (recovery).

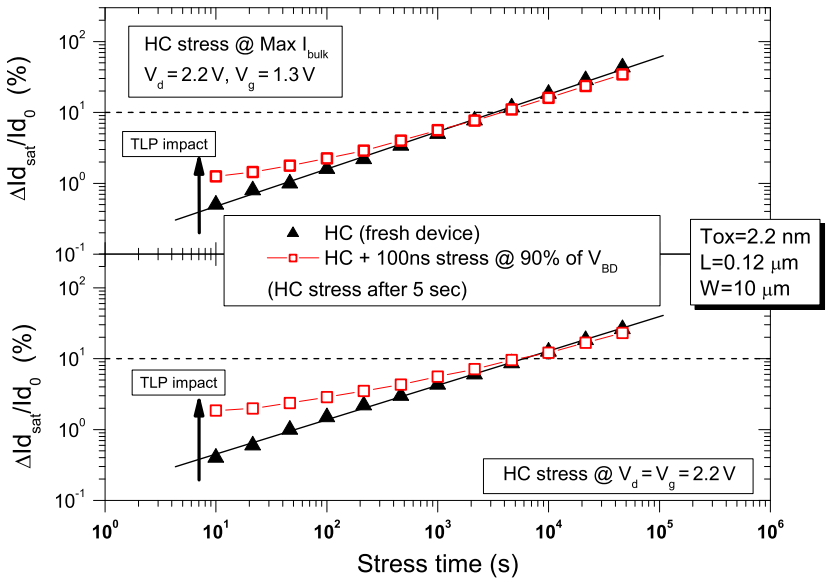

Figure 14: Impact d'un stress TLP ESD de 100 ns sur la cinétique de dégradation HCI pour un oxyde de 2.2 nm d'épaisseur de type nFET [20].

III-D - Établissement d'une fenêtre de design ESD fiable

III-D-1 - Loi de puissance du TDDB du claquage franc

Pour le développement de produits CMOS robustes et fiables, la procédure qui consiste à établir une fenêtre de design ESD adaptée constitue une étape primordiale pour la conception du schéma de protection [3, 4, 6]. Le principe de la limitation imposée par la fenêtre de design ESD est basée sur le critère de claquage franc des oxydes de grille sous stress TDDB dans le régime temporel ESD.

La détermination des lois de puissance temporelle en fonction des tensions de stress (TDDB) peuvent être déduites de deux manières différentes, par une caractérisation approfondie ou par des extrapolations. Il y a deux possibilités pour l'obtention des valeurs de claquage de l'oxyde basé sur les méthodes d'extrapolation. Pour les fabricants, une extrapolation directe des données peut être effectuée à l'aide des qualifications de durée de vie du diélectrique liées à la qualité des procédés technologiques pour la réalisation de la structure de grille. Cette dernière procédure est, par exemple, décrite dans la Figure 15.

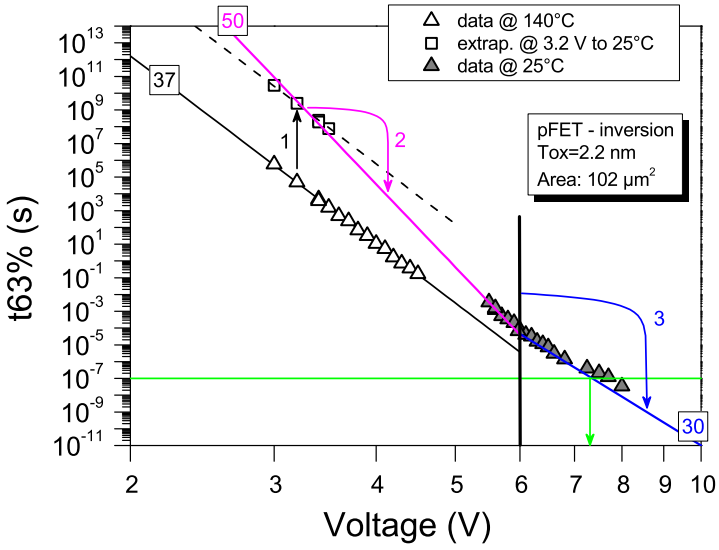

Figure 15: Extraction de la tension de claquage des oxydes minces depuis l'extrapolation des données de fiabilité à long terme. 1) extrapolation des données de qualification des procédés technologiques de 140 °C jusqu'à 25 °C effectuées pour une faible tension en utilisant les modèles d'extrapolation DC en température. 2) modification du facteur d'accélération en tension de 140 °C à 25 °C. 3) prise en compte de la plus petite valeur d'accélération en tension au delà du coude pour 6 V (valeur du pire cas du transistor pFET stressé en inversion). Finalement, la valeur de la tension au claquage de l'oxyde dans le domaine de l'ESD peut être extraite.

Dans le cas où aucun accès aux données de fiabilité à long terme n'est possible, les extrapolations peuvent être basées sur les travaux de caractérisation publiés en utilisant les règles théoriques ou empiriques d'extrapolation (accélération du TDDB en fonction de l'épaisseur d'oxyde par exemple). Une approche sécuritaire peut se baser sur le pire cas constitué par le nFET stressé en inversion reporté pour de nombreuses épaisseurs d'oxyde dans [20].

III-D-2 - Critères de défaillance et design ESD

La détermination d'une fenêtre de design ESD sûre concernant la limitation liée aux dégradations des oxydes de grille ne se limite pas à l'établissement de la loi au claquage de l'oxyde sous stress TDDB [20]. La détermination de la condition limite de design doit inclure les effets statistiques de la dégradation du diélectrique ainsi que les dégradations liées au dispositif. Un choix judicieux des critères de défaillance ESD ainsi que pour son niveau d'évaluation pour établir la résistance au claquage constitue une tâche délicate mais strictement nécessaire à l'aboutissement de produits fiables. Les conditions restreignant le design ESD, exposées dans la Figure 16, ne sont pas fixées et sont fortement dépendantes de l'épaisseur d'oxyde.

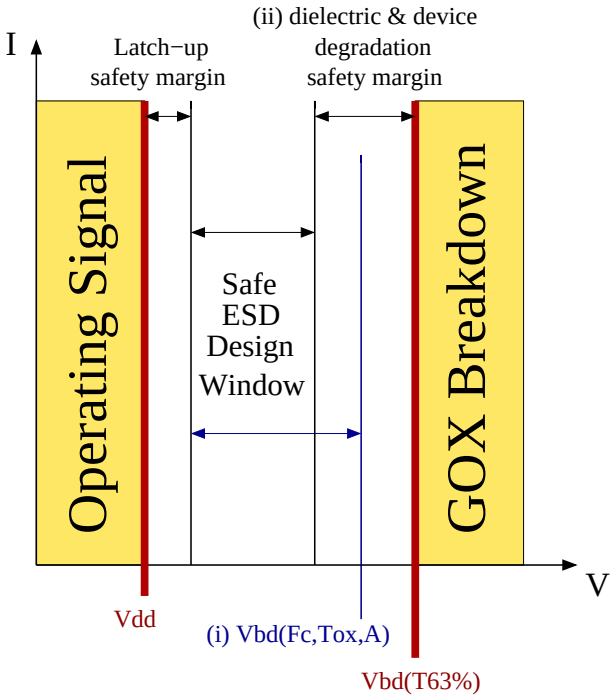

Figure 16: Fenêtre de design ESD considérant deux différentes conditions limites pour la dégradation des oxydes de grille: (i) le claquage du diélectrique en fonction d'un niveau statistique acceptable de défaillance, (ii) la limite résultante de la réduction de la durée de vie du diélectrique et du dispositif (donné en pourcentage de V_{bd}). L'ordre des deux critères mentionnés est arbitraire.

(i) En ce qui concerne le claquage franc des oxydes, la valeur des pentes des distributions des claquages cumulés en fonction du temps de défaillance (pentes de Weibull décrites par le paramètre β) est dépendante de l'épaisseur de l'oxyde de grille. Ces pentes diminuent pour des oxydes plus minces comme le montre la Figure 17. D'après l'idée de la théorie de percolation depuis ses origines [32, 33, 34] jusqu'à ses multiples adaptations focalisées sur les oxydes ultra-minces incluant des effets de saturation [35, 36, 37], la taille critique du défaut dans l'oxyde (de diamètre a_0) impose une dépendance en T_{ox} que l'on peut écrire génériquement comme,

$$\beta \sim \frac{T_{ox}}{a_0} + Cste \tag{7}$$

Le claquage de l'oxyde dépend de l'épaisseur d'oxyde, du niveau du critère de défaillance cumulé, de la surface d'oxyde, du type du dispositif, de la polarité du stress et de la température.

(ii) En ce qui concerne les dégradations du diélectrique et du dispositif

1- L'impact de l'effet cumulatif de la dégradation du diélectrique dépend de l'accélération en tension du TDDB qui est fonction du type du dispositif et du domaine de tension. Une dépendance indirecte avec l'épaisseur du diélectrique peut être observée comme le claquage des oxydes minces qui est localisé dans un autre domaine d'accélération en tension que ceux des oxydes épais dans le domaine des ESD. Ceci est dû à la transition intervenant dans le facteur d'acceleration en tension vers 5 - 6 V.

2- La dégradation de l'oxyde de grille qui entraîne la défaillance partielle (quasi-claquage) ou complète (claquage franc) du dispositif dépend surtout de l'épaisseur de l'oxyde puis du type de dispositif canal N ou canal P, du circuit et du type d'application.

III-D-3 - Sélection du critère de défaillance pour le design ESD

La sélection du critère de défaillance pour la fenêtre ESD est fonction de l'épaisseur d'oxyde.

- Pour les oxydes épais: la dégradation de l'oxyde de grille est gouvernée par les charges piégées dans l'oxyde induites au cours du stress ESD qui ne perturbent pas significativement la durée de vie du diélectrique mais qui entraînent des dérives importantes des caractéristiques électriques des dispositifs (V_{th}, $I_{d,lin}$, $I_{d,sat}$, G_m,...). Le critère de défaillance pour une condition de stress ESD peut être exprimé en fonction d'un pourcentage de la tension au claquage franc V_{bd} pour lequel les dérives des paramètres du dispositif restent dans les spécifications tolérées.

- Pour les oxydes minces: la nature statistique de la croissance des défauts liée à la dégradation des diélectriques minces contrôle la restriction du design ESD. Ceci est basé principalement sur deux phénomènes, *(1)* sur la réduction des pentes des distributions des claquages cumulés en fonction du temps de défaillance (Figure 17) et *(2)* sur le facteur important d'accélération de la loi de puissance du TDDB en dessous du coude dans lequel les oxyde minces sont situés. Le critère de défaillance est établi dans ce cas par le choix d'un niveau de défaillance cumulé satisfaisant les normes qualité (%, ppm).

La corrélation entre la marge de sécurité nécessaire pour un design sûr avec le niveau cumulé de défaillance est présentée Figure 18 pour un stress de type HBM (Human Body Model [3]) pour plusieurs épaisseurs d'oxyde. Cette corrélation est basée sur une large base de résultats expérimentaux. La nette discontinuité observée sur le graphe à trois dimensions résulte du coude intervenant dans le TDDB qui est observable dans le Figure 5.

III-D-4 - Démarche générale pour un design ESD cohérent vis à vis des défaillances d'oxyde de grille

La démarche à suivre pour l'ingénieur ESD est donnée par les spécifications des applications ainsi que par la fenêtre de design ESD quantifié pour un design pertinent. Concernant la limite liée aux défauts d'oxydes, le domaine accessible pour le design ESD est défini par la condition limite pire-cas applicable aux dispositifs en fonction de son épaisseur d'oxyde de grille. Pour un stress de type HBM, l'évolution de la fenêtre de design ESD en fonction de l'épaisseur d'oxyde est exposée en Figure 19. Dans cet exemple, la pire condition limite pour les oxydes épais a été choisie pour une marge de sécurité de 25% de V_{bd}. Pour les oxydes minces un niveau cumulé de défaillance de 1 ppm est appliqué. La transition entre oxydes épais et oxydes minces est définie pour 2.8 nm. Cette tendance est ici exposée plus spécifiquement pour les étages d'entrée considérant une surface d'oxyde de l'ordre de 1 μm^2. Le défi imposé à la protection des oxydes de grille minces est clairement démontré avec une marge d'action fortement réduite pour la conception ($\sim$ 2 V) en comparaison avec les oxydes de grille épais.

En ce qui concerne l'évaluation de la pertinence du concept ESD vis à vis de la protection des oxydes, la loi de puissance en tension du TDDB permet d'évaluer avec précision l'impacte des surtensions intervenant pendant les états transitoires ou lors de stress de type CDM (Charged Device Model, soit pour une durée de stress de 1 ns). Ces types de stress font actuellement l'objet d'une attention croissante dus à leurs sévérités prédominantes dans les technologies CMOS avancées [3, 38, 39] ainsi qu'aux efforts requis nécessaires pour surmonter les difficultés expérimentales rencontrées à la maîtrise de la durée des stress inférieurs à 5 ns. L'alternative possible obtenue par une méthode d'extrapolation en temps précise se révèle donc très utile et permet l'identification des problèmes potentiels dans la phase de développement d'une technologie avant d'aborder sa qualification.

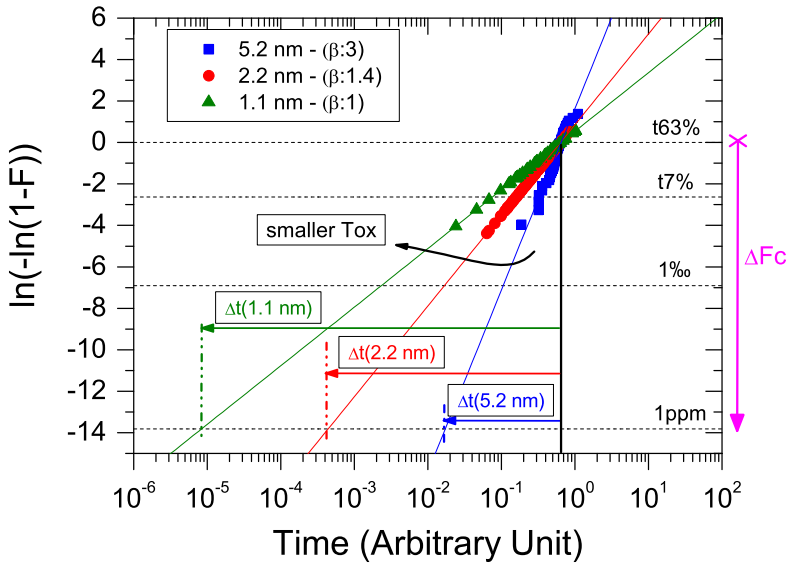

Figure 17: Impact des pentes de Weibull β (relatives à l'épaisseur d'oxyde) sur le temps au claquage en fonction du niveau de défaillance cumulé [20].

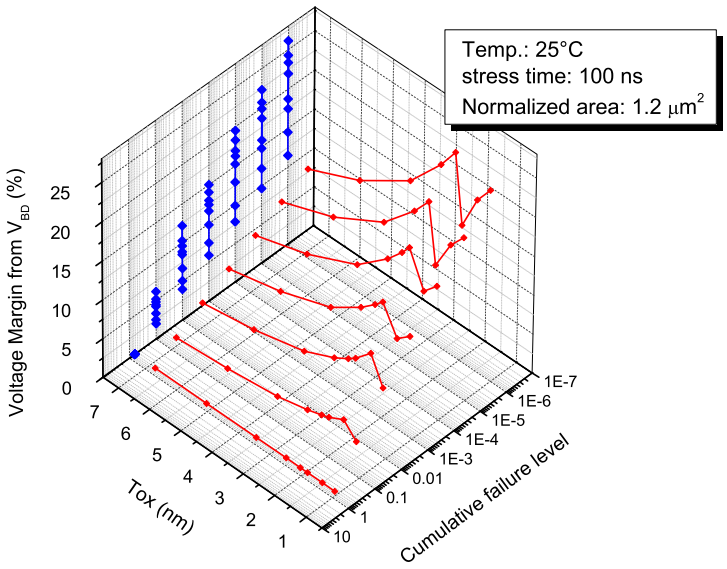

Figure 18: Marge de sécurité pour le design ESD établie depuis la tension de claquage franc exprimé en pourcentage dans le domaine en temps correspondant au stress HBM en fonction de l'épaisseur d'oxyde et du niveau de défaillance cumulé.

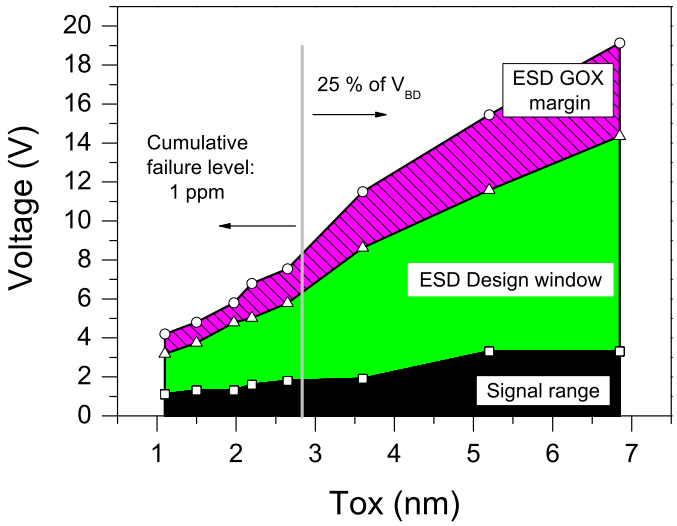

Figure 19: Tendance de la fenêtre de design ESD considérant la protection de l'oxyde de grille d'étages d'entrée vis à vis de stress HBM.

IV - Conclusion

Les travaux présentés dans ce manuscrit ont abordé le problème spécifique des décharges électrostatiques de façon innovante à travers la caractérisation de la fiabilité du diélectrique de grille au claquage, utilisé dans les dernières technologies CMOS, en fonction des intervalles en temps spécifiques aux stress ESD.

Cette contribution a permis d'acquérir une meilleure compréhension des différents mécanismes de dégradation des oxydes minces sous contraintes de type ESD et d'en dégager les conséquences pour la conception de dispositifs résistants aux ESD. Nos travaux ont notamment participés à une avancée significative concernant la modélisation des mécanismes de claquage des oxydes épais à ultra-minces dans le régime des nanosecondes soumis à de très fort champs électriques.

La méthodologie et le modèle compact d'extrapolation proposés ont été développés grâce à une caractérisation approfondie des différents mécanismes dans le but de fournir un nouvel outil puissant pour les développeurs de protections ESD au niveau design. Ceci contribuera à réaliser des produits fiables et robustes concernant l'ESD qui reste un problème majeur dans les technologies submicroniques CMOS à l'origine des défaillances irréversibles ou progressives des oxydes minces. L'optimisation de la conception des circuits a aussi été traitée en considérant ces aspects de fiabilité couvrant l'intégrité des fonctionnalités ainsi que la durée de vie des dispositifs soumis à des stress ESD de niveau modéré. Cette thèse ouvre aussi de nouvelles perspectives pour l'évaluation de la robustesse ESD des futurs matériaux diélectriques à haute permittivité qui commencent à faire leur apparition dans les derniers produits 45nm en remplacement des Oxydes minces SiO_xN pour les prochains nœuds technologiques.

Bibliography

[1] Moore G.E. Progress in digital integrated electronics. In *International Electron Device Meeting Technical Digest*, pages 11–13, 1975.

[2] Ziebart Wolfgang (Infineon Technologies President & CEO). Technical and economical trends in micoelectronics. In *European Solid-State Device Research Conference / European Solid-State Circuits Conference*, pages 1–10, 2007.

[3] Amerasekera Ajith and Duvvury Charvaka. *ESD in Silicon Integrated Circuits*. John Wiley & Sons, LTD, Chichester, England, second edition, 2002.

[4] Esmark Kai, Gossner Harald, and Stadler Wolfgang. *Advance Simulation Methods for ESD Protection Development*. ELSEVIER Science Ltd, Oxford, England, 2003.

[5] Degraeve R., Groeseneken G., Bellens R., Depas M., and Maes H.E. A consitent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides. *Annual International Reliability Physics Symposium*, pages 863–866, 1995.

[6] Boselli G., Rodriguez J., Duvvury C, and Smith J. Analysis of ESD protection components in 65nm CMOS technology: Scaling perspective and impact on ESD design window. In *EOS/ESD Symposium*, 2005.

[7] Smedes T. and Guitard N. Harmful voltage overshoots due to turn-on behaviour of ESD protections during fast transients. In *EOS/ESD Symposium*, pages 366–375, 2007.

[8] Wolters D.R. Breakdown and wearout phenomena in Si/SiO2. *Insulating films on semiconductors*, pages 180–194, 1981.

[9] DiMaria D.J, Cartier E., and Arnold D. Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon. *Journal of Applied Physics*, 73(7):3367–3384, 1993.

[10] Stathis J.H. and DiMaria D.J. Reliability projection for ultra-thin oxides at low voltage. *International Electron Device Meeting*, pages 167–170, 1998.

[11] Chen I.C., Holland S., and Hu C. A quantitative physical model for time-dependent breakdown in Si/SiO2. In *International Reliability Physics Symposium*, pages 24–31, 1985.

[12] Kimura M. and Ohmi T. Time-dependent dielectric degradation (TDDB) influenced by ultrathin film oxidation process. In *Journal of applied physics*, volume 35, pages 1478–1483, 1996.

[13] Wu Ernest Y. and Suñé Jordi. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45:1809–1834, 2005.

[14] Wu J., Juliano P., and Rausenbaum E. Breakdown and latent damage of ultra-thin gate oxides under ESD stress conditions. In *EOS/ESD Symposium*, 2000.

[15] Weir B.E., Leung C.C, Silverman P.J., and Alam M.A. Gate dielectric breakdown: a focus on ESD protection. In *International Reliability Physics Symposium*, 2004.

[16] Matsuzawa K., Satake H., C. Sutou, and Kawashima H. Gate oxide reliability under ESD-like pulse stress. In *Simulation of Semiconductor Processes and Devices*, pages 129 – 132, 2003.

[17] Martin A., O'Sullivan P., and Mathewson A. Dielectric reliability measurement methods: a review. *Microelectronics Reliability*, 38(1):37–72, 1998.

[18] Pio F. Sheet resistance and layout effects in accelerated tests for dielectric reliability evaluation. In *Microelectronics Journal*, volume 27, pages 675–685, 1996.

[19] Degraeve R., Kaczer B., and Groeseneken G. Degradation and breakdown in thin oxides layers: mechanisms, models and reliability prediction. *Microelectronics Reliability*, 39:1445–1460, 1999.

[20] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. In *EOS/ESD Symposium Proceedings*, pages 328–337, 2007.

[21] Ille A., Stadler W., Gossner H., Brodbeck T., Pompl T., and Bravaix A. Thin gate oxides time-to-breakdown in the ESD time domain and consequences for applications. In *International Electrostatic Discharge Workshop*, pages 176–186, 2007.

[22] Ille A., Stadler W., Kerber A., Pompl T., Brodbeck T., Esmark K., and Bravaix A. Ultra-thin gate oxide reliability in the ESD time domain. In *EOS/ESD Symposium Proceedings*, pages 285–294, 2006.

[23] DiMaria D.J, Buchanan, Sthatis J.H, and Stahlbush R.E. Interface states induced by the presence of trapped holes near the silicon-silicon-dioxide interface. *Journal of Applied Physics*, 77(5):2032–2040, 1995.

[24] Riess Philip. *Etude de la fiabilité des oxydes minces: analyse des mécanismes de transport et de génération du SILC*. Physique des composants a semiconducteurs, Institut National Polytechnique de Grenoble, 1999.

[25] Esseni D. and Bude J.D. On interface and oxide degradation in VLSI MOSFETs - part II: Folwer-nordheim stress regime. *Transaction on Electron Devices*, 49:254–263, 2002.

[26] Wu E.Y., Sune J., Nowak E., Lai W., and McKenna J. Weibull slopes, critical defect density, and the validity of stress-induced-leakage current (SILC) measurements. *International Electron Device Meeting*, pages 125–128, 2001.

[27] Aur S., Chattejee A., and Polgreen T. Hot-electron reliability and ESD latent damage. *Transaction on Devices*, 35(12):2189–2193, 1988.

[28] Huh Y., Lee M.G., Jung H.C., Li T., Song D.H., Lee Y.J., Hwang J.M., Sung Y.K., and Kang S.M. A study of ESD-induced latent damage in CMOS integrated circuits. In *International Reliability Physics Symposium*, 1998.

[29] Song M., Eng D.C, and MacWilliams K.P. Quantifying ESD/ESD latent damage and integrated cirecuit leakage currents. In *EOS/ESD Symposium*, pages 304–310, 1995.

[30] Reiner J.C., Keller T., Jäggi H., and Mira S. Impact of ESD-induced soft drain junction damage on CMOS product lifetime. *Microelectronics Reliability*, 40:1619–1628, 2000.

[31] Takeda E. and Suzuki N. An empirical model for device degradation due to hot carrier injection. *Electron Device Letters*, 4:111, 1983.

[32] Suñé J., Placencia I., Barniol N., Farrés E., Martin F., and Aymerich X. On the breakdown statistics of very thin Si/SiO2 films. *Thin SiO2 Films*, 185:347–362, 1990.

[33] Dumin D.J., Mopuri S.K., Vanchinathan S., Scott R.S., Subramoniam R., and Lewis T.G. High field related thin oxide wearout and breakdown. *Transactions on Electron Devices*, 42(4):760–772, 1995.

[34] Stathis J. H. Percolation models for gate oxide breakdown. *Journal of applied physics*, 86(10):5757–5765, 1999.

[35] Suñe J., Jimenez D., and Miranda E. Breakdown modes and breakdown statistics of ultrathin Si/SiO2 gate oxides. *International Journal of High Speed Electronics and Systems*, 11(3):789–849, 2001.

[36] Nicollian P.E, Krishnan T., Chancellor C.A, Khamankar R.B., Chakravarthi S., Bowen C., and Reddy V. The current understanding of trap generation mechanisms that leads to the power law model for gate dielectric breakdown. In *International Electron Device Meeting*, pages 197–206, 2007.

[37] Krishnan A.T. and Nicollian P.E. Analytic extension of the cell-based oxide breakdown model to full percolation and its implications. *Annual International Reliability Physics Symposium*, pages 232–239, 2007.

[38] Etherton Melanie. *Charged Device Model (CDM) ESD in ICs : physics, modeling, and circuit simulation*. Technische wissenschaften, Eidgenössische Technische Hochschule ETH Zürich, 2006.

[39] Ito C. and Loh W. A new mechanism for core device failure during CDM ESD events. In *EOS/ESD Symposium*, 2006.

Contents

Symbols

Symbols	Description	Units
C_{ox}	Oxide effective capacity	F
C'_{ox}	Oxide capacity per area unit	F/cm^2
D_{it}	Interface states density per unit area and energy	eV^{-1}cm^{-2}
E_a	Thermal activation energy of the TDDB from the Arrhenius model	eV
E_c	Energy level of the conduction band	eV
E_F	Fermi energy level	eV
E_g	Energetic forbidden band of the silicon (1.12 eV)	
E_i	Intrinsic energy level	eV
E_{ox}	Field across the oxide	V/cm
E_v	Energy level of the valence band	eV
F_c	Statistical failure criterion defined as the cumulative failure level	% (or ppm)
G_m	Transistor transconductance	S
$G_{m,max}$	Maximum transistor transconductance	S
h	Planck constant, $6.626 \cdot 10^{-34}$	J/s
I_{cp}	Pumped substrate current from a CP measurement	A
$I_{d,lin}$	Drain current in the linear regime	A
I_{ds}	Drain current	A
$I_{d,sat}$	Drain current in the saturated regime	A
I_g	Gate current	A
I_{off}	"Off" drain current defined for $V_{gs} = 0V$ and $V_{ds} = V_{dd}$	A
I_{on}	"On" drain current of a MOS transistor defined for $V_{gs} = V_{ds} = V_{dd}$	A
I_{silc} (J_{silc})	Current (current density) of the gate leakage at low field	A (A/cm^2)
I_{sub} (J_{sub})	Substrate current (substrate current density)	A (A/cm^2)
J_g	Gate current density	A/cm^2
k_B	Boltzmann constant, $1.38 \cdot 10^{-23}$ J/K $= 8.617 \cdot 10^{-5}$ eV/K	
L_g	Gate length of a MOS transistor	µm
m_e	Electron rest mass, $9.11 \cdot 10^{-31}$ kg $= 5.69 \cdot 10^{-16}$ eV s^{2}cm^{-2}	
m^*_{ox}	Effective electron mass in the oxide	kg
n_{BD}	Critical trap density at the oxide breakdown per unit volume	cm^{-3}
N_A	Acceptor doping concentration per unit volume	cm^{-3}
N_D	Donor doping concentration per unit volume	cm^{-3}
N_{it}	Charge number trapped at the interface per unit area	cm^{-2}
N_{poly}	poly-silicon doping concentration per unit volume	cm^{-2}
q	Elementary charge, $1.60218 \cdot 10^{-19}$	C

Continued on next page

Symbols	Description	Units
Q_{BD}	Charge to oxide breakdown	$\mathrm{C/cm^2}$
Q_{i}	Inversion charge in the semiconductor per unit area	$\mathrm{C/cm^2}$
Q_{inj}	Injected charge through the gate	$\mathrm{C/cm^2}$
Q_{inv}	Mobile charge present in the inversion channel layer	$\mathrm{C/cm^2}$
Q_{it}	Charge trapped in the interface states	$\mathrm{C/cm^2}$
Q_{ot}	Charge trapped in the oxide volume	$\mathrm{C/cm^2}$
T	Temperature	K
T_{BD}	Time to oxide breakdown	s
T_{ox}	Oxide thickness	nm
$T_{63\%}$	Characteristic parameter of a Weibull statistic distribution	s
V'_{FB}	Non-ideal flat band voltage including volume oxide charges	V
V_{dd}	Supply voltage	V
V_{ds}	Voltage between drain and source (with source as reference level)	V
$V_{\mathrm{d,sat}}$	Saturated drain voltage of MOS transistor	V
V_{FB}	Flatband voltage	V
V_{g}	Electrical potential at the gate	V
V_{th}	Threshold voltage of a MOS transistor	V
$V_{\mathrm{t,lin}}$	Threshold voltage of a MOS transistor in the linear regime	V
$V_{\mathrm{t,sat}}$	Threshold voltage of a MOS transistor in the saturated regime	V
V_{poly}	Band bending in the poly-silicon	V
W_{g}	Gate width of a MOS transistor	µm
β	Shape parameter of a Weibull statistic distribution	
ε_0	Vacuum dielectric permittivity, $\varepsilon_0 = 8.854 \cdot 10^{-12}$	F/m
$\varepsilon_{\mathrm{ox}}$	Dioxide dielectric permittivity, $\varepsilon_{\mathrm{ox}} = 3.9$	
$\varepsilon_{\mathrm{si}}$	Silicon dielectric permittivity, $\varepsilon_{\mathrm{si}} = 11.9$	
μ_{eff}	Effective carrier mobility in the semiconductor	$\mathrm{cm^2/Vs}$
μ_{n}	Electron mobility in the semiconductor	$\mathrm{cm^2/Vs}$
μ_{p}	Hole mobility in the semiconductor	$\mathrm{cm^2/Vs}$
Φ_{b}	Barrier height for electron tunneling	eV
Φ_{F}	Substrate volume potential resulting from doping	V
Φ_{ms}	Electron work function between gate and semiconductor substrate	V
Φ_{n}	Potential in the neutral region of an *n*-type semiconductor	V
Φ_{p}	Potential in the neutral region of an *p*-type semiconductor	V
Ψ_{S}	Surface potential	V

Abbreviations

Abbreviations	Description
AHI	Anode Hole Injection
AQL	Acceptable Quality Level
BEOL	Back End Of Line
BTI	Biased Temperature Instability
CDM	Charged Device Model
CMOS	Complementary Metal-Oxide Semiconductor
CSR	Space Charge Region
C-V	Capacitance Voltage characteristic
CVS	Constant Voltage Stress
CP	Charge Pumping
DIBL	Drain Induced Barrier Lowering
DT	Direct Tunneling
DUT	Device Under Test
ECB	Electron Conduction Band
EMMI	EMission MIcroscopy
EOS	Electrical OverStress
EOS/ESD	Electrical OverStress/Electrostatic Discharge
ESD	Electrostatic Discharge
ESDA	Electrostatic Discharge Association (professional voluntary association)
ESD PE	ESD Protection Element
EVB	Electron Valence Band
FET	Field Effect Transistor
FN	Fowler-Nordheim
GIDL	Gate-Induced Drain Leakage
GOX	Gate Oxide
HBD	Hard Breakdown
HBM	Human Body Model
HC	Hot Carrier
HCI	Hot Carrier Injection
I/O	Input/Output
IC	Integrated Circuit
ISSG	In-situ Steam Generation
JEDEC	Joint Electron Device Engineering Council
LDD	Lightly Doped Drain

Continued on next page

Abbreviations	Description
LOCOS	LOCal Oxidation of Silicon
MM	Machine Model
MOS	Metal Oxide Semiconductor
NBTI	Negative Bias Temperature Instability
NFET	*n* channel field-effect transistor
PCM	Process Control Monitoring
PFET	*p* channel field-effect transistor
RF	Radio Frequency
RTN	Random Telegraph Noise
R&D	Research and development
SBD	Soft Breakdown
SCR	Silicon Controlled Rectifier, thyristor
SDM	Socketed Device Model
SEM	Scanning Electron Microscopy
SILC	Stress Induced Leakage Current
SLBK	Silicide Blocked
STI	Shallow Trench Isolation
STM	Scanning Tunneling Microscopy
SiO_2	Silicon dioxide
SiO_xN	Silicon oxynitride
TDDB	Time Dependent Dielectric Breakdown
TLP	Transmission Line Pulsing
ULSI	Ultra Large Scale Integration
VLSI	Very Large Scale Integration
VRS	Voltage Ramp Stress
ggNMOS	Grounded-Gate nMOS
ppm	Parts Per Million

Part I

Dissertation

Introduction

The technologic boom from the XX's century and the impressive advancements in semiconductor technology has radically changed our life style. Electronic devices are omnipresent in our daily life; at home as well as at work. Particularly communication in modern life without achievements such as mobiles or high-speed internet is purely unthinkable.

The development of new electronic systems is basically driven by only two key factors. The major factor up to now is the continuous ongoing miniaturization of the electronics components [1]. The second trend, which is becoming today as important as the first one, is the diversification of the technologies applications and features [2]. The decreased components size reduces the price of the products; the primary reason is that more devices can be processed per silicon area. However, in order to achieve steadily better performances at lower prices, IC design and process innovations are resulting in an complexity increase which consequently increases directly the development and fabrication costs. In the semiconductor industry, the financial loop between the research and development costs (R&D) and the productive phase which ensures benefits is in accordance with Moore's Law [1]. Moore's law was first introduced in 1965 by Gordon Moore and is still verified nowadays by the leading foundries. The well-known statement of that law is that the integration complexity doubles every 2 years. Thus, large benefits are only provided by low-cost production and highly integrated ICs in a huge production volume. To fulfil this requirement, technical innovations are required which permit a continuous integration of more functionalities with better performances on a single chips, which then should require less area. The semiconductor market is deeply dedicated in this intense integration run which forces the constant reduction of the elementary electrical functional cells (transistors) and generates technologies mutations.

The other side of the coin is that these key success factors described above are increasing the sensitivity of ICs to reliability threats and are considerably impacting the complexity of the solution for their protection during life time. Electrostatic Discharges (ESD) are one of the major reliability threats to integrated circuits (ICs). Since more than three decades, ESD is a constant challenge for semiconductor companies; and even today the cost impact of external and internal ESD protection measures is enormous. Without appropriate ESD protection concepts, ICs processed in today's sub-micron CMOS technologies can hardly withstand any electrostatic discharge; ESD would cause an irreversible damage of the IC. Without these external and internal

preventives measures, ESD would be a killing costs for semiconductor industries as the production yield would be dramatically poor and no ESD-safe ICs could be guaranteed in the field - where the end-users insist on reliable products.

The increasing architecture complexity (with mixed signal applications, multiple power domains, over-tolerant pads, RF pads, USB systems, flip chips, ...) is strongly challenging the ESD protection task. The down-scaling of transistors dimensions are going oppositely to the electrical constraints applied to and to the electrical parameters maintain. In advanced CMOS technologies, the strong decrease in the dielectric thicknesses does not go along with the reduction of supplies voltages which results into extremely high field conditions. Consequences are drastically intensification of reliability issues such as hot carriers (HC) degradation, negative bias temperature instability (NBTI), and ESD. With the miniaturization of the ICs and consequently the evolution of the processes, the vulnerability of the ICs to ESD events has dramatically increased. This has two different root causes, basically the continuous increase of the boundary constraints and the process restriction impact on the ESD protection device development.

One pre-dominant boundary limit for today's ESD protection concepts comes from the shrinking thicknesses of gate oxide isolation layers in advanced CMOS processes which is almost approaching its physical limit. The oxide thickness planned in advanced technology nodes is coming down to 10 Angstrom which represents just 4 mono atomic layers. One can easily understand the extreme sensibility of such thin oxides faced to ESD stresses in the range of few to several Amperes causing easily voltage drops of more than 10 V (around 10 times the operational voltage!). The development of robust and safe ESD protections is a real challenge and gate oxide breakdown is becoming one major issue for ESD designers. The characterization of the thin gate oxide under ESD stress condition and the understanding of all impacting parameters on the oxide breakdown and degradation is essential for the protection of integrated circuits. In this work, the problem of the gate oxide reliability which gains significant importance during the last years has been investigated. The main goal is to provide ESD engineers with an useful guidance how to achieve ESD-robust products.

Outline of the thesis

The focus of this thesis is the reliability of thin oxides in advanced sub-microns Complementary Metal Oxide Semiconductor (CMOS) technologies under ESD stress. At the beginning of the work, only very little work was done in this field, although it was considered to be a critical issue particularly for processes with ultra-thin gate oxides. In contrast, gate oxide reliability theory in the long-time range is a well established topic. During the investigations of this thesis, the fundamental dielectric reliability theories have been revised, the post-breakdown modes, the percolation theory, the breakdown mechanisms and the modeling. The special ESD focus on thin oxides has been embedded into this reliability picture evolution to support the robust protection development.

Chapter 1

The first chapter gives an introduction to the dielectric reliability in MOS technologies. The MOS physics, device parameters and leakage conduction modes through oxides will be described to define the basic ideas required in the following of the work. Then, the dielectric reliability under electrical stress is developed considering the oxide degradation statistics and mechanisms. Finally, the dielectric lifetime extrapolation will be presented with the different breakdown theories and acceleration models.

Chapter 2

In Chapter 2, an overview of ESD phenomena and their impacts on the semiconductor industries is discussed, starting from the ESD events as "real world threat" to ICs towards the ESD robustness product qualification by means of electrical models. Next, the "'go/no-go"' standardized qualification methodology and the attributed failure mechanisms are extended with the description of the ESD characterization procedures required for the ICs' protection development against ESD stress. The concept of a physically-based on-chip ESD protection will be exposed by introducing protection strategies and the corresponding elements – but also the manifold limitations that make the ICs protection a hard task are addressed. This topic will be detailed with the discussion of the constraints coming along with the miniaturization and integration as well as the applications diversity; resulting in a drastically increased challenge to achieve ESD safe and robust products. More precisely, the boundary condition resulting from the gate oxide isolation layer integrity in advanced CMOS technologies will be emphasized.

Chapter 3

Chapter 3 is exposing the CMOS technologies investigated in the thesis, as well as the experimental characterization set-up and test structures.

Chapter 4

In Chapter 4, after a brief literature review on the gate oxide modeling towards the nanoseconds, a new characterization approach for the gate oxide testing and breakdown evaluation in the ESD time range is introduced. The verification of the statistical degradation of thin oxides under ESD stress according to the percolation theory is exposed. The impacts of testing procedures and test structures on the time-to-breakdown evaluation are quantified, resulting in the selection of the best accurate oxide breakdown methodology for time-to-dielectric-breakdown (TDDB) investigations. In a second part of that chapter, the thin oxides breakdown from the DC down to the ESD regime is characterized in order to allow an accurate modeling of the oxide breakdown voltage acceleration. As the extrapolation laws from process qualification data in

the direction of ESD robustness estimation are quite important, a full characterization of the parameters impacting the TDDB models are investigated.

Chapter 5

In Chapter 5, impacts of non-destructive ESD stress on gate oxides are investigated for various oxide thicknesses. The oxide to device degradation induced by moderate ESD stress requires a precise knowledge to guarantee reliable ESD safe designs. The impacts of oxide degradation under ESD-like stress on the device functionality as well as on the device lifetime are reported by means of changes in DC characterization and Hot Carrier Injection on nFET (n-channel Field Effect Transistor) devices.

Chapter 6

Chapter 6 concludes the experimental findings in Chapter 4 and 5 and gives precise rules for the definition of safe ESD products designs. A selection of the failure criteria is discussed for the determination of the ESD design window. The GOX breakdown modeling down to the ESD regime is established for the ESD guidelines and methodology flow.

This thesis helps to gain a better understanding of the oxide degradation and reliability concerns of thin oxides exposed to ESD events. The methodology and extrapolation packages characterized in this thesis provide powerful tools for ESD engineers who have to define appropriate guidelines for ESD-robust designs. For the first time, fails and degradations of thin gate oxides under ESD stress is consistently included in the protection concept development. Furthermore, the methodology presented in the thesis can be transferred to future process nodes and, hence, supports the development of future CMOS technologies. This is a big step towards first-time-right and robust product design.

Bibliography

[1] Moore G.E. Progress in digital integrated electronics. In *International Electron Device Meeting Technical Digest*, pages 11–13, 1975.

[2] Ziebart Wolfgang (Infineon Technologies President & CEO). Technical and economical trends in micoelectronics. In *European Solid-State Device Research Conference / European Solid-State Circuits Conference*, pages 1–10, 2007.

Chapter 1

Gate Oxide reliability in CMOS technologies

1.1 MOSFET technologies

State-of-the-art integrated circuits (ICs) often include hundreds of millions of transistors; without transistors the miniaturization and integration of functionality into one single IC is hardly imaginable. The metal-oxide-semiconductor (MOS) structure is the base of the success of the high density integration of transistors at comparably low cost. This thesis focuses on MOS field-effect transistor (MOSFET) technologies which are processed by silicon and the derived silicon oxide. An introduction to the MOSFET transistor element with its basic functionalities will be described in the following section.

1.1.1 MOSFET transistor description

The main function of a transistor is to modulate the carrier concentrations of a semiconductor material (silicon) by a field-effect control and, thereby, to control the electrical current flowing between a source of carriers (source) and a sink of carriers (drain). The source and drain regions are processed by a local high doping concentration of carriers in the semiconductor. These diffusions regions can be either negatively doped (n or n^+ in case of high doping) or positively doped (p and p^+, respectively). The semiconductor region between these two doped diffusions is called channel. There are two types of transistors in complementary MOSFET (CMOS) technologies. First, the n-channel transistor (nFET) with n^+ doped diffusions embedded into a p well and, secondly, the p-channel transistor (pFET) which, in opposite, comprises p^+ doped diffusions in an n well, see Figure 1.1.

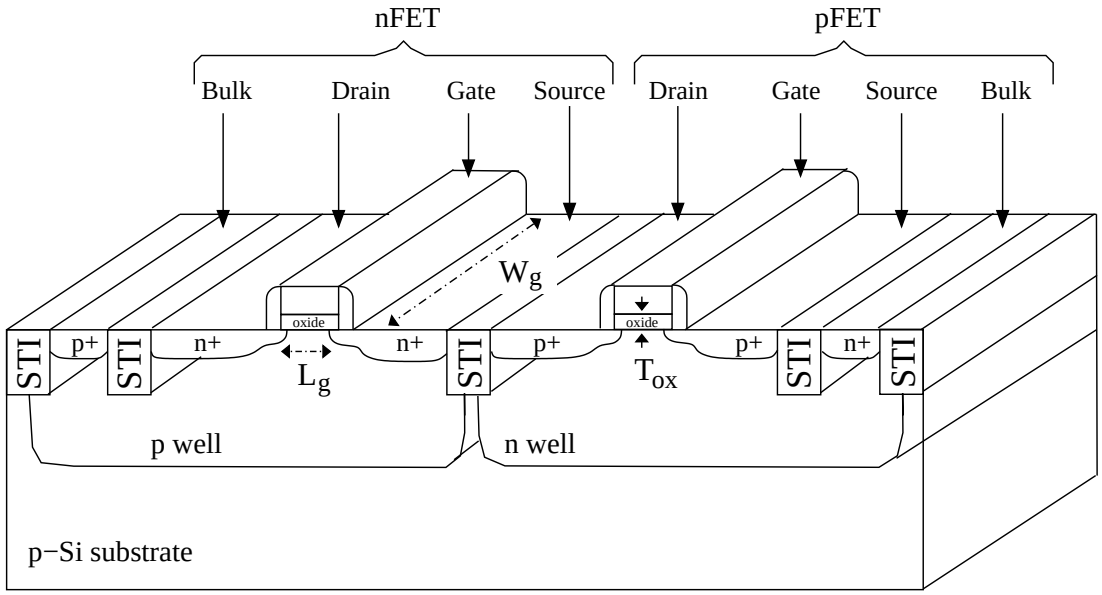

Figure 1.1: Schematic cross section of an advanced CMOS technology with twin-wells and shallow-trench isolation (STI).

The current flow caused by the carrier transport from the source to the drain is controlled by the differential potential between the two diffusions. The intensity of this current flow is modulated by a field effect provided by a third electrode potential, the so-called gate. A thin insulating oxide layer between the channel and the gate avoids carrier flow from the channel into the gate. The gate provides the electrical potential control in the channel by capacitive coupling which permits the regulation of the conduction between source and drain. The role of the gate and the gate oxide is essential for the functionality of the MOSFET transistors.

A MOSFET transistor is in general a four-terminal structure. In addition to source, drain and gate terminals, access to the potential of the doped well (*n* well or *p* well) is enabled by a fourth connection, the so-called "bulk". Lateral isolations are implemented between the transistors to electrically isolate the devices. These isolations could be processed by means of LOCOS (local oxidation of silicon) or shallow trench isolation (STI). The basic design parameters of a transistor which define its electrical characteristics are the gate length (L_g), the gate width (W_g), and the oxide thickness (T_{ox}).

1.1.2 Gate oxide isolation layers

1.1.2.1 Silicon dioxide and oxynitride characteristics

One major factor for the economic success of MOSFET technologies in microelectronics is the easy process implementation of silicon oxides (SiO_2) and oxynitrides dielectrics in the silicon process. This is combined with their extremely good insulating properties (high resistivity in

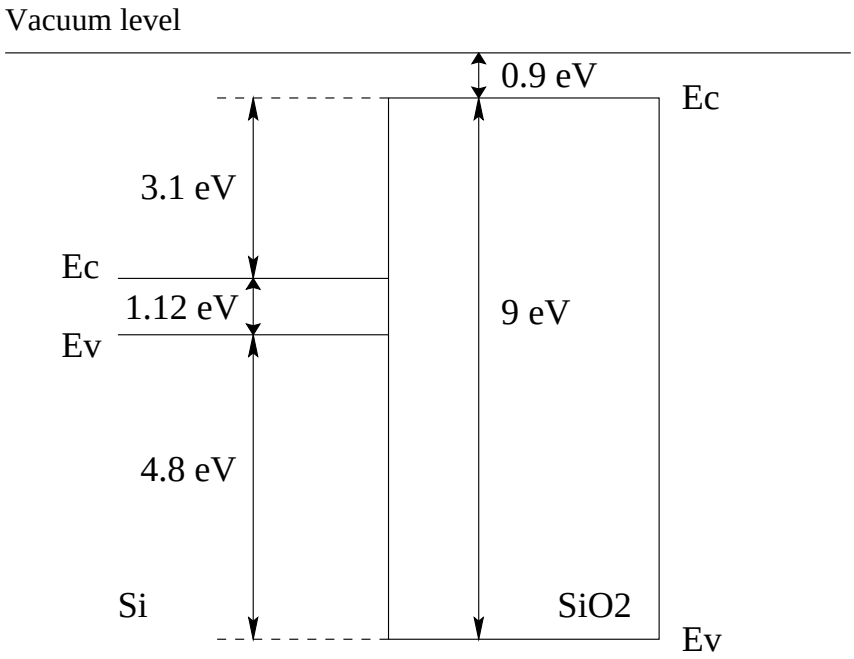

Figure 1.2: Energy diagram of the Si/SiO_2 system without external electrical field. E_c corresponds to the bottom level of the conduction band and E_v to the top level of the valence band.

the range of 10^{15} to 10^{17} Ωcm) and their large forbidden energy band gap of 9 eV [1]. A further benefit of these oxides is their high intrinsic breakdown field strength which is above 15 MV/cm [2].

However, with the ongoing miniaturization of ICs, the transistor operating conditions are becoming extreme. The decrease of the oxide thickness below 3 nm implies already that at maximum operating conditions (maximum ratings) the electric fields across the gate oxide approach 10 MV/cm. This aggressive scaling demands for strongly enhanced oxide reliability considerations. As one counter measure, SiO_2 is strongly nitrided in CMOS processed with thin oxides. The nitridation step provides a further improved isolation property of ultra-thin oxide layers and extenuates some device reliability issues by filling oxide defects.

Without electrical fields, the ideal energetic diagram of the Si/SiO_2 system is sketched in Figure 1.2.

The structure of SiO_2 solids are based on the tetrahedron structure SiO_4. In the case of the SiO_2 amorphous oxide, the angle of the Si–O–Si compounds are in the range of 115° to 180° [3, 4, 5]. The inter-atomic distance from a silicon atom to the central oxygen atom of the SiO_4 structure is 0.162 nm. This means that in ultra-large scale integration (ULSI) technologies, ultra-thin oxides are only composed of few atomics layers; an 1 nm oxynitride oxide just consists of approximately 4 monolayer. A vertical cross section of a 90 nm gate length transistor with a close-up on the silicon/oxide/poly-silicon gate interface structure of an 2.3 nm thin oxide is exposed in Figure 1.3.

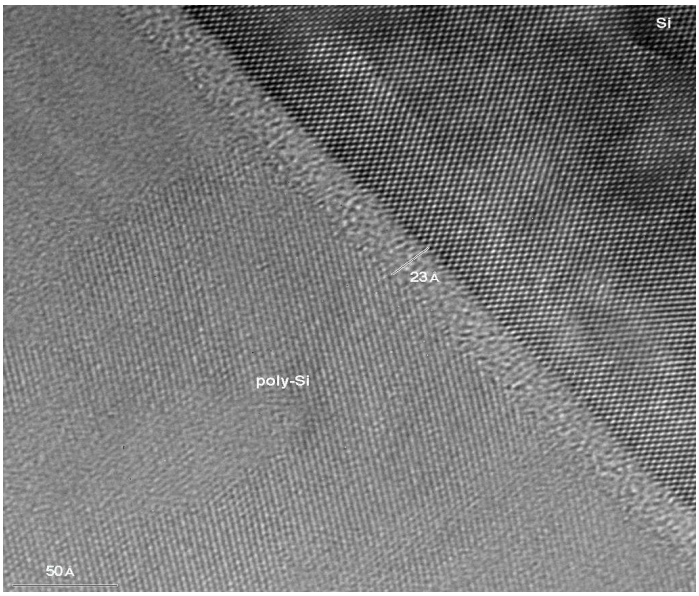

Figure 1.3: Vertical cross section of a 90 nm gate length MOS transistor, a focus on the SiO_2 oxide layer of the MOS device is shown in the picture.

1.1.2.2 Gate oxide quality and defects

The silicon (Si) substrate structure is crystalline and based on a cubic face centered atomic arrangement. At the Si/SiO_2 interface, the abrupt mismatch between the crystallin order and the tetrahedron oxide structure is inducing some distortion of electrical bonds [6]. The structural mismatching at the interface results in broken bonds (Si–O, Si–Si or Si–H). It can results also into oxygen vacancy configurations ($E'\alpha$ centers) which involve Si–Si bonds formation. The possible oxide defects at the interface are compiled in Figure 1.4.

These defects introduce energetic levels in the dielectric electrical scheme which could then be electrically activated and enable trapping or providing of carriers. Their energetic states are located in the forbidden gap of the substrate silicon. There are two types of single traps:

- the acceptor type, which have the capability to trap one electron or to release one hole;
- the donor type, which can provide one electron or trap one hole.

There exists three different charged states of traps, either negatively or positively charged or neutral. A donor type is positively charged when empty and neutral when filled by a hole. An acceptor type is negatively charged when filled and neutral when empty. There exist also amphoteric sites. For example the P_b center resulting from trivalent silicon dangling bond has

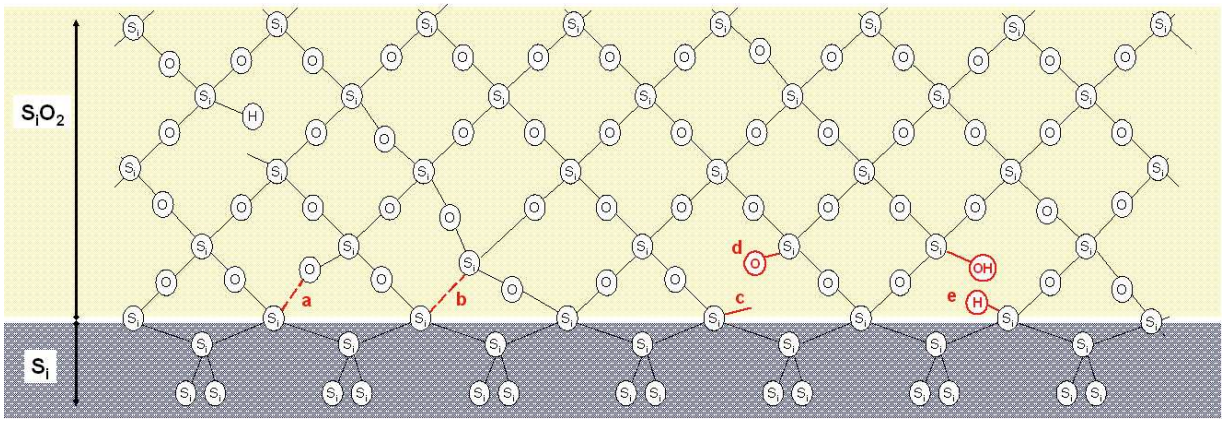

Figure 1.4: Si/SiO_2 interface defects: (a) weak Si–O bond; (b) weak Si–Si bond resulting from an oxygen vacancy ($E'\alpha$ centers); (c) Si dangling bond, called P_b center; (d) non-bridging oxygen, (e) Si dangling bonds, saturated by hydrogen species (done during the passivation step).

the particularity to be either acceptor or donor, meaning positively or negatively charged as well as neutral. In order to fill the dangling bonds which could be electrically active, a thermal passivation step is required after the oxide grown process. This aims to saturate the defects sites with hydrogen species.

The atomic arrangement in these oxides is not dense and the diffusion capability of impurities is then not negligible. The oxide structure itself comprise also defects in its volume:

- Fixed positive charges (Q_f) at the interfaces (Si, N atoms in excess);
- Positive mobile charges (Q_{mob}) due to impurity contamination during the process (Na, K^+, H^+);
- Bulk oxide trapped charges (Q_{ot}) which could be positive or negative (impurities, broken bonds);
- Interface traps charges (Q_{it}) due to the mismatch of the Si/SiO_2 lattice.

The defects in the oxides are characterized by the energetic level in the dielectric, their capture cross section coefficient σ, which is relating their effective speed to interact with carriers and their trap types. A characteristic distinction between the interface traps and the bulk (or border traps) is their speed ability to trap or de-trap charges. The interface states are in general rather “fast” traps whereas the deep traps are described as “slow” traps [7, 8].

1.1.3 Metal-Oxide-Semiconductor structure

1.1.3.1 MOS energy band diagram and MOS capacitor regimes

In this section, the functionality of the Metal-Oxide-Semiconductor structure under gate polarization will be described [9]. In this configuration the structure is acting as a MOS capacitor. In advanced CMOS technologies the gate electrode is not anymore processed in metal but with a highly doped ($> 10^{20}$ cm^{-3}) poly-silicon which provides the same properties as a metal.

The energetic band diagram of a MOS structure is exposed in Figure 1.5. The case of an n-channel device will be developed in this paragraph.

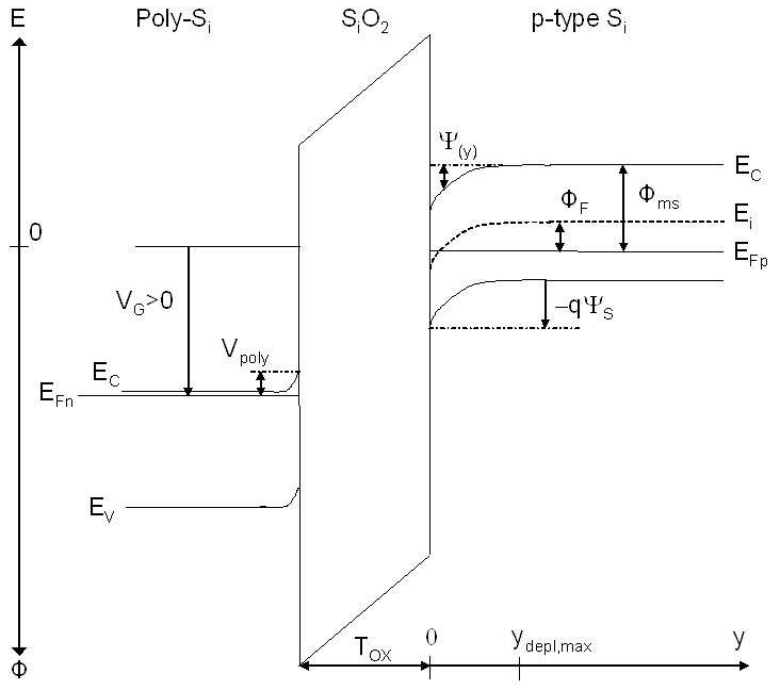

Figure 1.5: Band diagram from an NMOS in strong inversion.

E_c is the bottom level of the conduction band, E_v is the top level of the valence band, E_F is the Fermi level, E_i is the intrinsic level of the silicon (un-doped). V_{poly} is the band bending resulting from the depletion of the poly-silicon gate during inversion stress. $\Psi(y)$ is the band curvature in the silicon and Ψ_S is defined as the surface potential, $\Psi_S = E_i(\infty) - E_i(0)$.

The volume potential in the substrate depends on the doping concentration and is given by:

$$\Phi_F = \frac{E_i - E_{F,p}}{q} = \frac{kT}{q} \ln\left(\frac{N_A}{n_i}\right) \tag{1.1}$$

where n_i is the intrinsic carriers concentration in silicon and N_A is the acceptor doping concentration of the p-type semiconductor substrate. $E_{F,p}$ is the quasi-Fermi level in the p-doped semiconductor.

Under gate polarization (V_g), the MOS capacitor structure can be described by five different states which are function of the surface potential in the channel Ψ_S.

- The accumulation regime (Figure 1.6), $V_g < 0$, $\Psi_S < 0$
  The surface potential is negative and the majority carriers are attracted by this potential. An accumulation of free holes at the SiO_2/Si interface is occurring.

- The flat band condition (Figure 1.6), $\Psi_S = 0$
  The gate bias is compensating the gate-silicon work function (Φ_{ms}), all bands are flat. This gate bias is called flat band voltage V_{FB}. The ideal flat band is $V_{FB} = \Phi_{ms}$, for an n^+ poly-Si gate on a p-doped type Si semiconductor, $\phi_{ms} = -\frac{E_G}{2} - \phi_F$.

- The depletion regime, $V_g > 0$, $0 < \Psi_S < \Phi_F$
  The surface potential is positive and causes the repulsion of the holes at the interface. The minority electron concentration is increasing but is still below the hole concentration. A space charge region is induced at the interface above the "mid gap" condition $\Psi_S = \Phi_F$.

- The inversion regime (Figure 1.7)

  – The weak inversion, $V_g > 0$, $\Phi_F < \Psi_S < 2\Phi_F$
    The weak inversion starts when the electron concentration is becoming higher than the hole concentration in the channel close to the interface. The nature of the semiconductor in the channel is inverted.

  – The strong inversion, $V_g > 0$, $2\Phi_F < \Psi_S$
    Above the condition $\Psi_S = 2\ \Phi_F$, corresponding to the so called threshold voltage ($V_g = V_{th}$), the concentration of electrons is higher than the holes at the SiO_2/Si interface and creates an inversion layer composed of free electrons.

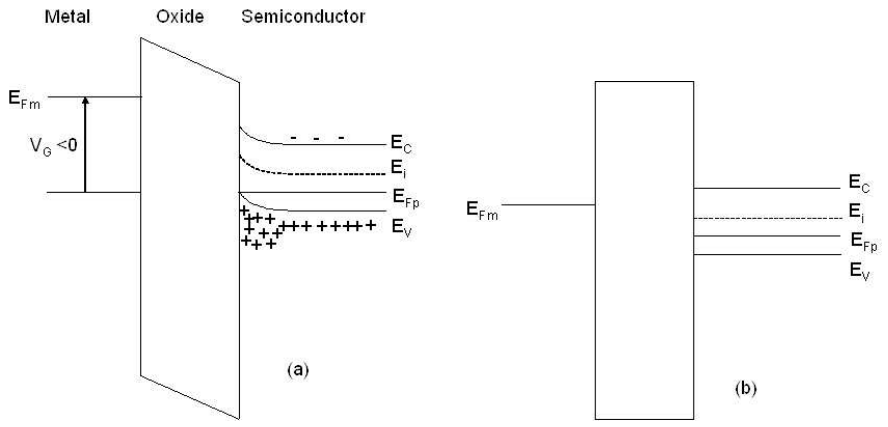

Figure 1.6: NMOS energetic band diagram described for a metal gate: (a) accumulation, (b) flat band.

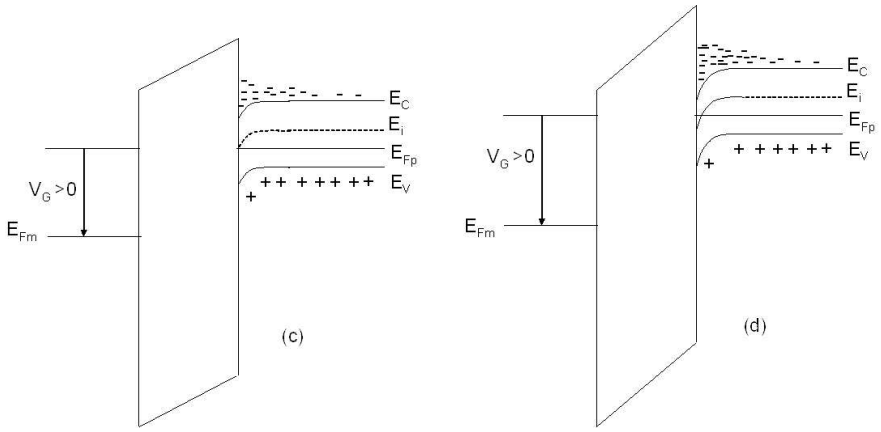

Figure 1.7: NMOS energetic band diagram described for a metal gate: (c) weak inversion, (d) strong inversion.

1.1.3.2 Inversion charge layer and threshold voltage

In this part the inversion charge involved in the MOSFET current conduction as well as the threshold voltage will be developed. Under gate biasing the charge is conserved,

$$Q_{\mathrm{G}} = -(Q_{\mathrm{SC}} + Q_{\mathrm{ox}} + Q_{\mathrm{it}}) \tag{1.2}$$

Q_{sc} is the charge in the semiconductor, Q_{ox} and Q_{it} are representing the charges created by defects which are present in the oxide and at the semiconductor/oxide interface, respectively,

$$(V_{\mathrm{G}} - \Psi_{\mathrm{S}} - \phi_{\mathrm{ms}})C_{\mathrm{ox}} = -Q_{\mathrm{SC}}(\Psi_{\mathrm{S}}) - Q_{\mathrm{ox}} - Q_{\mathrm{it}}(\Psi_{\mathrm{S}}) \tag{1.3}$$

$$\text{where } C_{\mathrm{ox}} = \frac{\varepsilon_0 \varepsilon_{\mathrm{si}}}{T_{\mathrm{ox}}} \text{ and } Q_{\mathrm{SC}} = Q_{\mathrm{inv}} + Q_{\mathrm{depl,max}}$$

$Q_{\mathrm{depl,max}}$ is the maximal depletion charge resulting from the Charged Space Region (CSR) and Q_{inv} is the inversion charge which participates in the current conduction.

In the strong inversion regime, for $\Psi_{\mathrm{S}} = 2\Phi_{\mathrm{F}}$,

$$Q_{\mathrm{inv}} = -C_{\mathrm{ox}}\left(V_{\mathrm{G}} - \left(\phi_{\mathrm{ms}} - \frac{Q_{\mathrm{ox}}}{C_{\mathrm{ox}}}\right) - 2\Phi_{\mathrm{F}}\right) - Q_{\mathrm{depl,max}} - \frac{Q_{\mathrm{it}}}{C_{\mathrm{ox}}} \tag{1.4}$$

The maximal depletion charge is obtained for the maximal depletion depth layer ($y_{\mathrm{depl,max}}$),

$$y_{\mathrm{depl,max}} = \sqrt{\frac{2\varepsilon_0 \varepsilon_{\mathrm{si}}}{qN_{\mathrm{A}}}(2\Phi_{\mathrm{F}} - V_{\mathrm{BS}})} \tag{1.5}$$

$$Q_{\mathrm{depl,max}} = -qN_{\mathrm{A}} y_{\mathrm{depl,max}} = -\gamma_{\mathrm{N}} C_{\mathrm{ox}} \sqrt{(2\Phi_{\mathrm{F}} - V_{\mathrm{BS}})} \tag{1.6}$$

$$\text{where the body factor } \gamma_N \text{ is defined as } \gamma_N = \frac{\sqrt{2\varepsilon_{ox}\varepsilon_{Si}qN_A}}{C_{ox}} \tag{1.7}$$

The threshold voltage at which the inversion layer is formed can be consequently express as,

$$V_{th} = V'_{FB} + 2\Phi_F + \gamma_N\sqrt{2\Phi_F - V_{BS}} - \frac{Q_{it}}{C_{ox}} \tag{1.8}$$

where the non ideal flat band voltage due to the charges in the oxide volume V'_{FB} is defined as

$$V'_{FB} = \phi_{ms} - \frac{Q_{ox}}{C_{ox}} \tag{1.9}$$

1.1.4 MOSFET transistor device parameters

1.1.4.1 MOSFET transistor conduction

The MOSFET device output functionalities are based on the conduction of the minority carriers from the inversion layer in the semiconductor. The source potential is established as the reference potential. The carriers flow from source to drain (I_{ds}) is dependent on the channel state controlled by the gate polarization (V_{gs}) and on the carrier acceleration induced by the lateral Field resulting from the drain potential (V_{ds}). When the inversion layer is formed, the conductive current between source and drain I_{ds} can be split into two regimes, a linear regime and a saturation regime. The drain current can be cut in three different output regimes (I_{ds} - V_{gs}),

1. **The sub-threshold regime**. In this regime no free minority carriers form the inversion layer in the channel. There is no conduction current,

$$I_{DS} = 0 \text{ when } V_{GS} - V_{th} < 0 \text{ (cut-off condition)}$$

However, in the sub-threshold regime, a weak drain current dominated by carriers diffusion exponentially increases until the formation of the inverted layer in the semiconductor.

$$I_{DS} = \phi_t W_g \mu_0 \frac{dQ_i}{dx} \quad \text{where} \quad \phi_t = \frac{kT}{e} \tag{1.10}$$

with μ_0 the carrier mobility in the channel and Q_i the inversion charge in the semiconductor per unit area. The diffusion current can be express as,

$$I_{DS} = \mu_0 C'_{ox}\frac{W_g}{L_g}(n-1)(\phi_t)^2 \exp\left[\frac{(V_{GS}-V_{th})}{n\phi_t}\right]\left[1-\exp\left(-\frac{V_{DS}}{\phi_t}\right)\right] \tag{1.11}$$

$$\text{with } n = 1 + \frac{C_{depl}+C_{it}}{C_{ox}} \text{ and } C_{depl} = \frac{\varepsilon_0\varepsilon_{ox}}{y_d}$$

with C_{depl} the capacitance from the depletion layer and y_{d} the depth of the depletion layer in the semiconductor. C_{it} is the parasitic capacitance resulting from defects at the semiconductor/oxide interface. In a $\log(I_{\text{ds}})$ - V_{gs} plot, this exponential current increase is defined as the sub-threshold slope (see Figure 1.8). This regime is characterized via the sub-threshold swing (S) express in mV/decade and defined as the inverse of the sub-threshold slope:

$$S = \ln(10)\,\frac{\partial V_{\text{GS}}}{\partial \ln I_{\text{DS}}} = \ln(10)\,\phi_{\text{t}}\left[1 + \frac{C_{\text{depl}} + C_{\text{it}}}{C_{\text{ox}}}\right] \tag{1.12}$$

2. **The linear regime** operates under the condition that ($V_{\text{DS}} \ll V_{\text{GS}} - V_{\text{th}}$). The drain current in this regime can be expressed as the flow of the mobile charges Q_{inv}, which are present in the inversion channel layer, within a transit time τ to reach the drain electrode:

$$I_{\text{DS}} = \frac{Q_{\text{inv}}}{\tau} \tag{1.13}$$

When V_{ds} is polarized, the electrical potential is varying along the channel, inducing a variation of threshold voltage by a potential $\mathrm{V_c}(x)$.

$$V_{\text{th}} = V_{\text{t0}} + V_{\text{c}}(x) + \gamma_{\text{N}}\sqrt{2\Phi_{\text{F}} + V_{\text{c}}(x)} \text{ with } V_{\text{t0}} = V'_{\text{FB}} + 2\Phi_{\text{F}} \tag{1.14}$$

$$V_{\text{th}} = V_{\text{t0}} + \alpha_{\text{N}} V_{\text{c}}(x) \text{ with } \alpha_{\text{N}} = 1 + \frac{\gamma_{\text{N}}}{2\sqrt{2\Phi_{\text{F}} - V_{\text{BS}}}} \tag{1.15}$$

Consequently, the mobile charge in the channel depends on the position x in the channel,

$$Q_{\text{inv}} = C'_{\text{ox}} W_{\text{g}} (V_{\text{GS}} - V_{\text{t0}} - \alpha_{\text{N}} V_{\text{c}}(x))\, dl \tag{1.16}$$

The transit time of carrier in the channel is given by,

$$\tau = \frac{\mathrm{d}\,l}{V_{\text{drift}}} = \frac{\mathrm{d}\,l}{\mu_0 F_{\text{lat}}} = \frac{\mathrm{d}\,l}{\mu_0 \frac{\mathrm{d}V}{\mathrm{d}l}} = \frac{\mathrm{d}\,l^2}{\mu_0 \mathrm{d}V} \tag{1.17}$$

where V_{drift} is the drift speed of the charges, μ_0 is the mobility of the charges and F_{lat} is the lateral field. From equation 1.13 and 1.17 the drain current can be written as,

$$I_{\text{DS}} = \frac{1}{\mathrm{d}\,l}\mu_0 C'_{\text{ox}} W_{\text{g}} (V_{\text{GS}} - V_{\text{t0}} - \alpha_{\text{N}} V_{\text{c}}(x))\mathrm{d}V \tag{1.18}$$

Considering $V_{\text{c}}(x)$ being constant for an infinitesimal channel part $\mathrm{d}\,l$,

$$I_{\text{DS}} \int_0^{L_{\text{g}}} \mathrm{d}\,l = \mu_0 C'_{\text{ox}} W_{\text{g}} \int_{V_{\text{c}}=V_{\text{S}}}^{V_{\text{c}}=V_{\text{D}}} [V_{\text{GS}} - V_{\text{t0}} - \alpha_{\text{N}} V_{\text{c}}(x)]\,\mathrm{d}V \tag{1.19}$$

$$I_{\text{DS}} = \mu_0 C'_{\text{ox}} \frac{W_{\text{g}}}{L_{\text{g}}}\left[(V_{\text{GS}} - V_{\text{t0}})V_{\text{DS}} - \alpha_{\text{N}}\frac{V_{\text{DS}}^2}{2}\right] \tag{1.20}$$

The vertical field resulting from the drain biasing V_{ds} causes a reduction of the mobility due to the increase of carriers scattering and in parallel results in a saturation of the carrier drift speed V_{max} in the channel. These effects are taking into account in the effective mobility μ_{eff},

$$\mu_{eff} = \frac{\mu_s}{1 + \frac{\mu_s V_{DS}}{V_{max} L_{eff}}} \text{ where } \mu_s = \frac{\mu_0}{1 + \theta(V_{GS} - V_{t0})} \tag{1.21}$$

where θ is the mobility reduction factor. Finally. the drain current in the linear regime is given by, .

$$I_{DS} = \mu_{eff} C'_{ox} \frac{W_g}{L_g} \left[(V_{GS} - V_{t0}) V_{DS} - \alpha_N \frac{V_{DS}^2}{2} \right] \tag{1.22}$$

In the linear regime, the channel acts as a resistance; the drain current is proportional to the drain voltage and the channel resistance is modulated by the gate voltage.

3. **In the saturated regime** the inversion condition is not valid anymore near the drain side and the current is limited by the pinch-off effect at the drain side of the channel. The current is given by,

$$I_{DS} = \mu_{eff} C'_{ox} \frac{W_g}{L_g} \frac{(V_{GS} - V_{t0})^2}{2\alpha_N} \tag{1.23}$$

with the saturated drain voltage,

$$V_{Dsat} = \frac{V_{GS} - V_{t0}}{\alpha_N} \tag{1.24}$$

The drain equations given in this part are based on the model SPICE level 3 developed at UC Berkeley [10]. Fine models including quantum-mechanic and parasitic effects should be taken into account for a precise modeling of MOS devices. However, the models exposed here are sufficient for the investigation of parameter drifts due to functional stress or induced by an ESD event.

1.1.4.2 MOSFET device characteristics

The main parameters of the electrical transistor characteristic which define the device functionalities and which are monitored to evaluate the degradation impact of electrical stresses are:

- S: the sub-threshold swing parameter;
- I_{off}: the 'off' current defined for $V_{gs} = 0$V and $V_{ds} = V_{dd}$;
- $V_{t,lin}$: the threshold voltage in the linear regime;
- $V_{t,sat}$: the threshold voltage in the saturated regime;

- G_m: the transconductance (gain) of the transistor,

$$G_m = \left(\frac{\partial I_{DS}}{\partial V_{GS}} \right)_{V_{DS},V_{BS}=\text{Const}} \tag{1.25}$$

- I_{ds} - V_{ds}: for different V_{gs} polarity (drain current family).
- I_{on}: the "on"-current of a MOS transistor defined for $V_{gs} = V_{ds} = V_{dd}$.
- $V_{d,sat}$: the saturated drain voltage.

The output drain conduction currents (I_{ds}) as a function of the gate voltage (V_{gs}) characteristics in a 90 nm CMOS technology node for an nMOS transistor are exposed hereafter in the three different modes, first, the sub-threshold (Figure 1.8), secondly, the linear (Figure 1.9) and, thirdly, the saturated regimes (Figure 1.10). The I_{off} parameter as well as the sub-threshold slope ($1/S$) in the linear and saturated modes can be seen in Figure 1.8. The parallel increase in the sub-threshold slope with the drain biasing is typical from short channel transistors. This short channel effect is known as "Drain-Induced Barrier Lowering" (DIBL) [11]. An other increased leakage phenomenon can be observed in the sub-threshold regime close to the transistor "off" state for high V_{ds} levels. This effect called "Gate Induced Drain Leakage" (GIDL) results from a local tunneling current at the drain contact side between the valence and conduction band in the silicon [12].

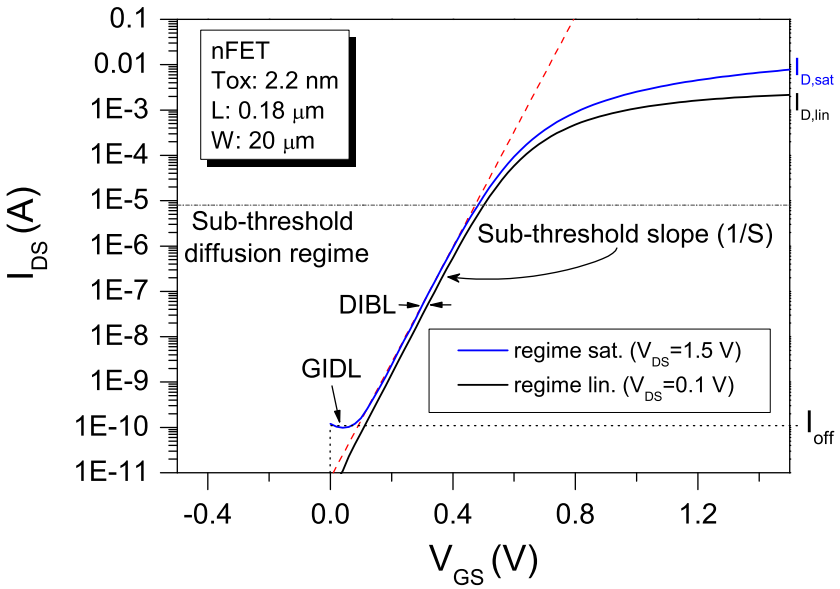

Figure 1.8: Sub-threshold drain current diffusion regime as a function of the gate voltage, plotted as $\log(I_{ds})$ - V_{gs}.

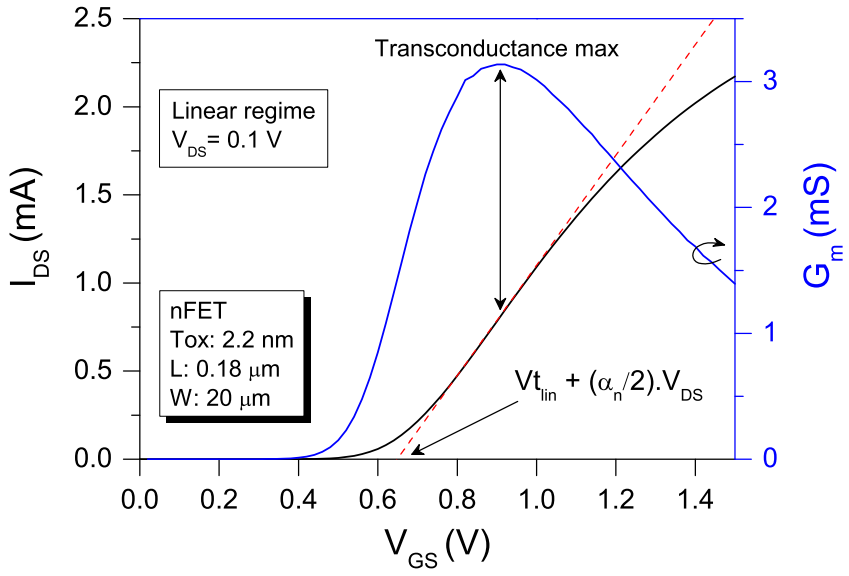

Figure 1.9: Drain current and transconductance versus gate voltage characteristics (I_{ds} - V_{gs} and G_m - V_{gs}) in the linear regime.

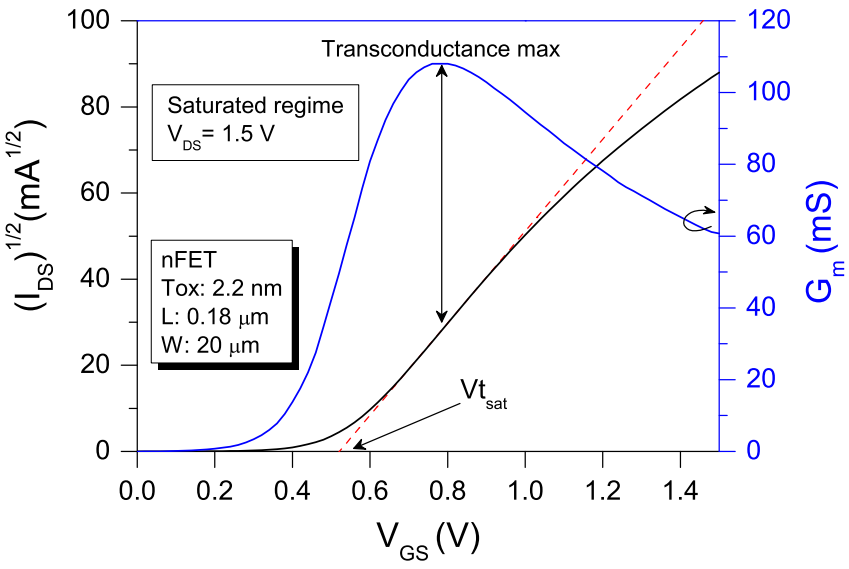

Figure 1.10: Drain current and transconductance versus gate voltage characteristics (I_{ds} - V_{gs} and G_m - V_{gs}) in the saturated regime.

From the $I_{ds} - V_{gs}$ plot, the threshold voltage can be extracted as the following,

- In the linear case, for $I_{ds} = 0$,

$$\text{From equation 1.22, } V_{t0}(I_{DS} = 0) = V_{GS}(I_{DS} = 0) - \alpha_N(V_{DS}/2) \tag{1.26}$$

  The threshold voltage ($V_{t,lin}$) can be extracted from the axis intercept from the tangent of the drain current at the maximum transconductance ($G_{m,max}$).

- In the saturated case, for $\sqrt{I_{DS}} = 0$,

$$\text{From equation 1.23, } V_{t0}(I_{DS} = 0) = V_{GS}(\sqrt{I_{DS} = 0}) \tag{1.27}$$

  The fit of the drain current tangent at the maximum transconductance from the graph $\sqrt{I_{DS}} - V_{GS}$, gives straight forward the $V_{t,sat}$ parameter.

The typical drain current characteristic as a function of the drain voltage at different gate bias voltages is drawn in Figure 1.11. The two distinct linear and saturated regimes are shown in the plot, delimited by the $V_{d,sat}$ values obtained for the different gate voltages.

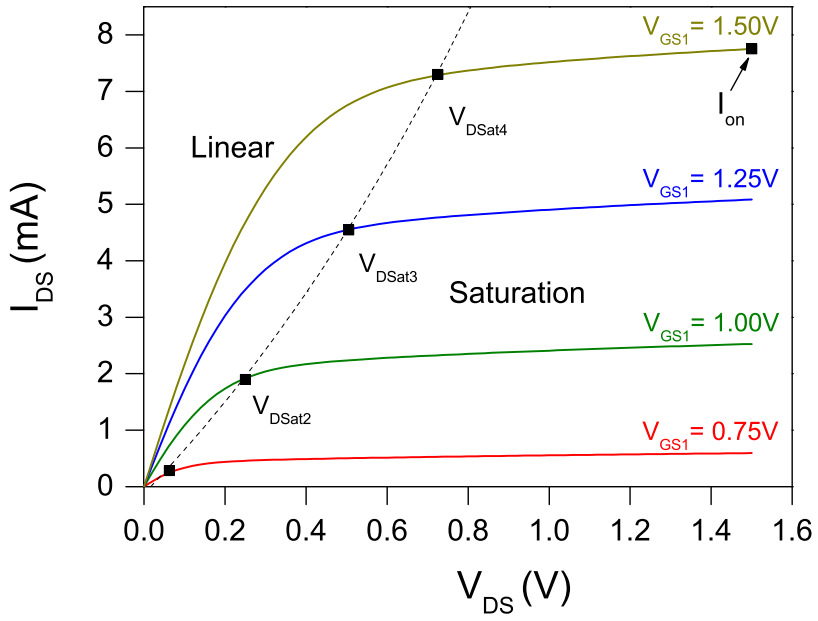

Figure 1.11: The MOS drain current versus the drain voltage is exposed for different gate biasing for an 2.2 nm nFET with L_g = 0.18 µm and W_g = 20µm.

1.2 Dielectric reliability

1.2.1 Transport in dielectrics

Although silicon dioxide dielectric is an insulator, there exists conductive current modes through oxides. Trough the Si/SiO_2 (and Si/SiO_xN) system there are mainly three different current leakage mechanisms (Figure 1.12), either resulting from high thermal energy, or caused by tunneling through the oxide or due to trap-assisted conduction in the oxide or at its interfaces.

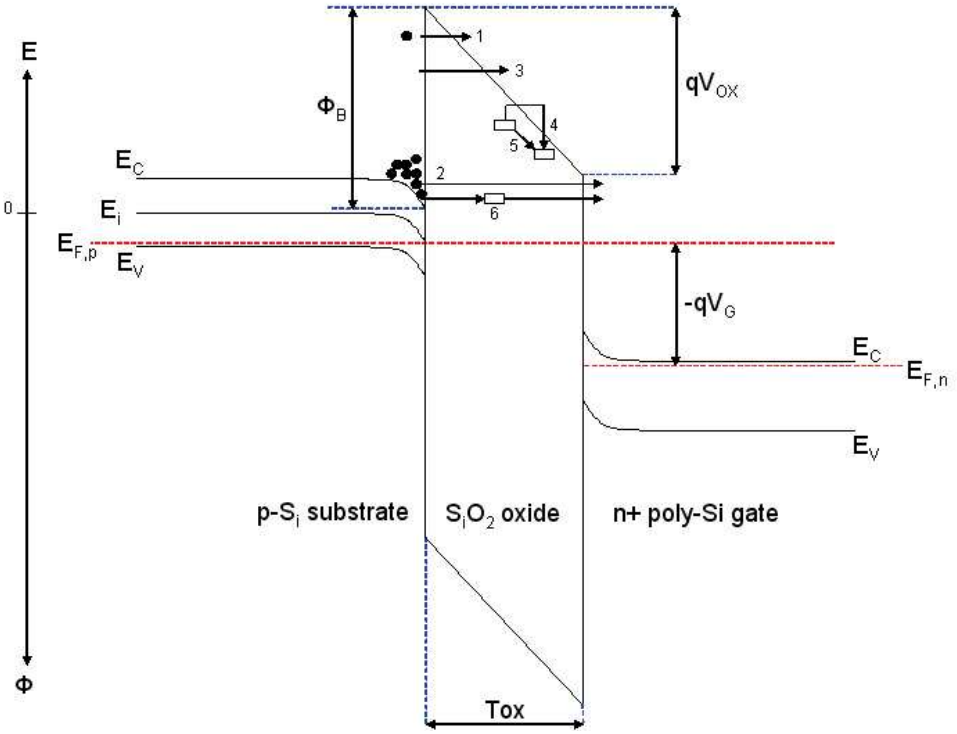

Figure 1.12: Conduction modes through the Si/SiO_2 system. (1) Thermoionic conduction; (2) Direct tunneling (DT); (3) Fowler-Nordheim (FN) tunneling; (4) Local trap to trap thermoionic conduction; (5) Hopping conduction; (6) Direct tunneling assisted by trap called "Stress Induced Leakage Current" (SILC).

A) Thermoionic conduction (1): current resulting from carriers having an higher energy than the potential barrier (ϕ_B) and which are injected in the conduction band. Basically, the energy required for this conduction mode could be provided by thermal energy (elevated temperature), optical excitation with photon radiation or by high electric fields which aims to increase carriers energy. This conduction mode is called Schottky current and is modeled with the Richardson-Schottky equation [13]:

$$J = A\,(k_B T)^2\, exp\left(\frac{B\sqrt{E_{ox}}}{k_B T}\right) exp\left(-\frac{q\phi_B}{k_B T}\right) \tag{1.28}$$

$$\text{with } A = \frac{4\pi q m^*_{ox}}{h^3} \text{ and } B = \sqrt{\frac{q}{4\pi\varepsilon_{ox}}}$$

m^*_{ox} is the effective masse of an electron tunneling in the oxide. k_B is the Boltzmann constant, T the temperature, h is the Planck constant and E_{ox} the field across the oxide.

B) Tunneling currents: they result from the energetic carriers distribution susceptible to pass through the oxide. These currents depend on the Fermi-Dirac function $f(E)$ and on the transparency of the dielectric $T(E)$ which gives the probability for a carrier with the energy E to tunnel trough the oxide potential barrier ϕ_B.

$$J = \frac{4\pi m_{Si}}{h} \int_{E_C}^{\phi_B} T(E) f(E) dE \tag{1.29}$$

with the Fermi-Dirac filling function,

$$f(E) = \frac{1}{1 + exp\left(\frac{E - E_F}{k_B T}\right)} \tag{1.30}$$

The current depends on the shape of the tunneling barrier, if the barrier is triangular then the Fowler-Nordheim [14] tunneling conduction operates; otherwise, in the case of a trapezoidal barrier a direct tunneling process takes place [15, 16]. The transition from the direct tunneling to the Fowler-Nordheim tunneling occurs when the potential energy over the oxide layer qV_{OX} becomes higher than the barrier hight ϕ_B (see Figure 1.12). For oxide thickness below 4 nm the direct electron tunneling is the dominant conductive mechanism. Neglecting confinement effects and using the WBK (Wentzel-Kramer-Brillouin) approximation those simplified tunneling current are expressed hereafter:

- Direct tunneling (DT) current (2).

$$J = \frac{A E_{ox}^2}{\left[1 - \left(\frac{\phi_B - qV_{ox}}{\phi_B}\right)^{1/2}\right]^2} \exp\left(-\frac{B}{E_{ox}} \frac{\phi_B^{3/2} - (\phi_B - qV_{ox})^{3/2}}{\phi_B^{3/2}}\right) \tag{1.31}$$

$$\text{with } A = \frac{q^3}{16\pi^2 \hbar \phi_B} \text{ and } B = \frac{4\sqrt{2m^*_{ox}}}{3q\hbar} \phi_B^{3/2}$$

- Fowler-Nordheim (FNT) tunneling (3).

$$J = A E_{ox}^2 \exp\left(-\frac{B}{E_{ox}}\right) \tag{1.32}$$

$$\text{with } A = \frac{q^3}{8\pi h \phi_B} \text{ and } B = \frac{8\pi\sqrt{2m^*_{ox}}}{3hq} \phi_B^{3/2} \text{ and where, } E_{ox} = \frac{V_g - V_{ox} - \Psi_s}{t_{ox}}$$

Parameters A and B can be extracted from the experimental Follower-Nordheim current density, plotting J/E_{ox}^2 as a function of $1/E_{ox}$ [17], see Figure 1.13. The coefficient B comes from the straight line slope and A is the intercept with the y axis.

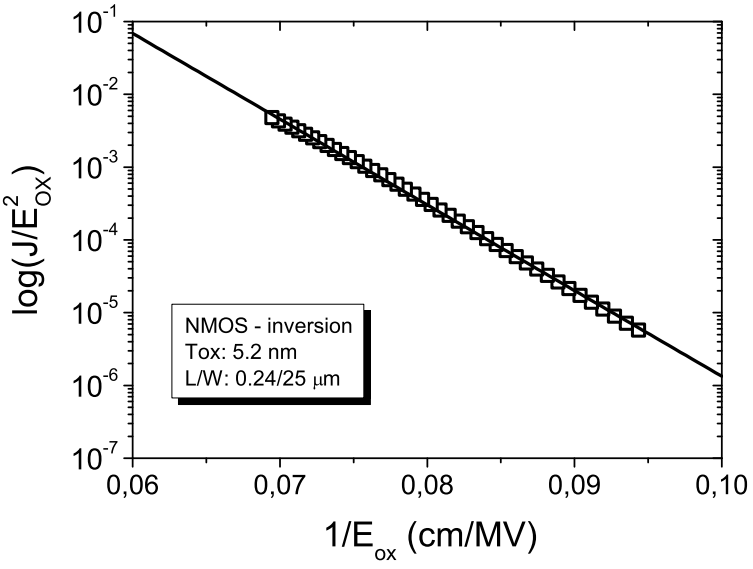

Figure 1.13: Extraction of Fowler-Nordheim conduction parameters from an experimental gate leakage measurement.

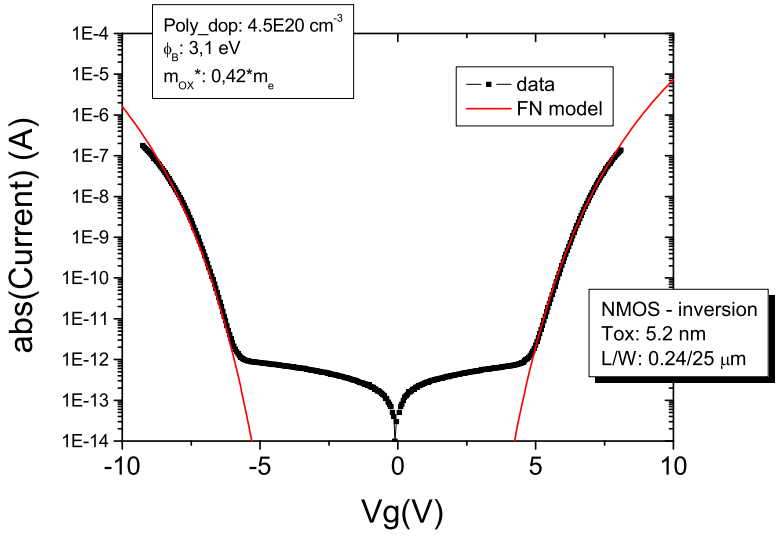

Figure 1.14: Theoretical Fowler-Nordheim leakage conduction fit.

In case of poly-silicon gates from a dual work CMOS process, the poly depletion effect occurring in the inversion stress case should be taken into account in the calculation of V_{ox} [18].

$$|V_{ox}| = |V_g| - |V_{Poly}|, \text{ with } V_{Poly} = \frac{\varepsilon_{ox}^2 E_{ox}^2}{2q\varepsilon_{si}N_{Poly}} \tag{1.33}$$

until pining occurs at 1.12 V due to the saturation of the band bending. N_{Poly} correspond to the poly silicon doping concentration. For accumulation stress, no depletion effect occurs in the gate and $|V_{ox}| = |V_g| - \frac{1}{q}|E_g| = |V_g| - 1.12$ V.

An overview of tunneling currents data obtained for different oxide films thicknesses between 1.5 nm and 6.85 nm is shown in Figure 1.15. For oxides in the range of 3 nm to 6 nm, quantic interaction phenomena also occurs [17, 15, 16] due to the interference of incident and partially reflected electron waves, propagating between the conduction band edge and the anode interface. For detailed modeling these effects should be taken into account. These phenomena could be seen in the gate leakage oscillations of the 3.6 nm oxide in Figure 1.15.

C) Currents assisted by traps:

- Local trap to trap thermoionic conduction (4) express by the Poole-Frenkel tunneling equation [19]:

$$J = qN_c\mu_{ox}(T)E_{ox}\exp\left(-\frac{\phi_B}{k_BT}\right)\exp\left(\frac{\beta\sqrt{E_{ox}}}{k_BT}\right) \tag{1.34}$$

$$\text{with } \beta = \sqrt{\frac{q}{\pi\varepsilon_{ox}}}$$

- Hopping conduction mechanism (5) which is a trap-to-trap tunneling current [13],

$$J = \frac{q^2a^2n^*}{\tau_0 k_B T}E_{ox}\, exp\left(-\frac{4\pi m_{ox}^*}{h}\phi_m a\right) \tag{1.35}$$

  τ_0 is the average hopping frequency, a is the distance between traps, n^* the number of trapping sites and ϕ_m the barrier high between to traps.

- Trap assisted tunneling (TAT) (6), direct tunneling conduction via traps in the oxide. This gate leakage conduction is known as Stress Induce Leakage Current (SILC) [20, 21, 22].

- Interface Trap assisted tunneling (ITAT), direct tunneling conduction via oxide interface traps. It is the gate leakage conduction involved in the Low-Voltage-SILC (LV-SILC) [23, 24, 25]. This mechanism could only occur in the range of ± 1 V around the flat band voltage.

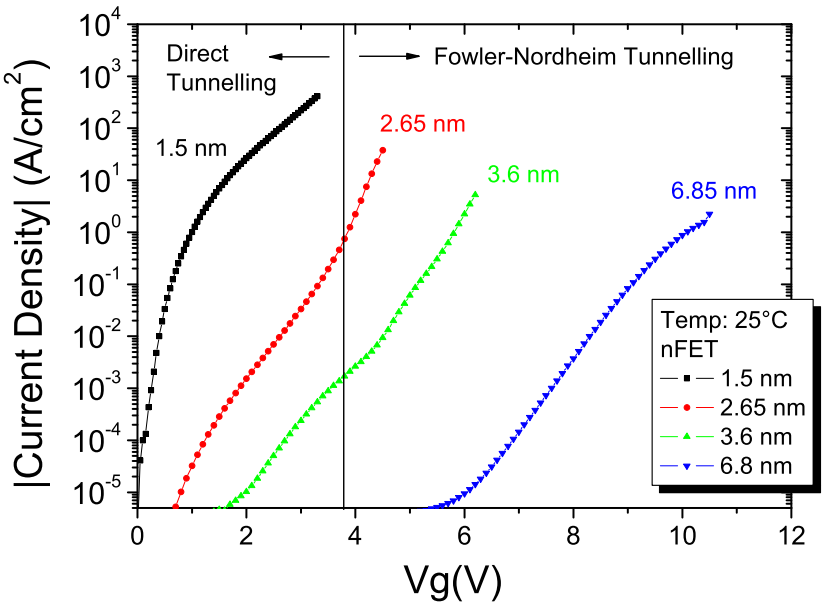

Figure 1.15: Gate leakage current as a function of the gate voltage showing both regimes, the Direct Tunneling current and the Fowler-Norheim Tunneling

1.2.2 Oxides post breakdown modes

Historically, the dielectric breakdown has been defined as a sudden loss of the insulating properties of the dielectric layer, resulting in a large increase of the gate leakage current of several orders of magnitude. This local low ohmic (1 to severals Ω) conductive path was defined as the hard breakdown (HBD). Then, with the introduction of thinner silicon dioxide films, another post breakdown mode has been observed, it is called soft breakdown (SBD). This event causes an increase in the gate leakage current which follows a power law as a function of the gate voltage ($I_g = K \cdot V^{\alpha}$) with an exponent α around 5 [28, 29] (Figure 1.16). Several SBD can occur before the final thermal breakdown (HBD) event [30]. In latest CMOS technologies with ultra-thin dielectric layers, preceding to the two breakdown modes exposed previously a progressive increase of the gate leakage is seen, starting with an instable gate leakage noise [31, 32, 33]. This instable leakage noise is also known as Random Telegraph Noise (RTN) and is exposed in Figure 1.17 in phase II. Then, there is a low increase in current which is called progressive breakdown (phase III in Figure 1.17) [34, 27]. Recent findings seem to explain that the different breakdown modes are only accessible under certain experimental conditions (time resolution, voltage range, compliance circuitry from set-up) [35, 36]. The apparent breakdown modes dependence on the oxide thickness is an artifact and different models have been proposed to explain the soft breakdown wear-out [37, 38, 29]. The sophisticate oxide breakdown wear-out will be

discussed in Chapter 4; this chapter is dedicated to the post breakdown behavior of oxides under ESD stress.

One main characteristic of the oxide breakdown, beyond its irreversibility, is the non-scaling level of the post breakdown leakage with the oxide area contrarily to the fresh leakage conduction mechanisms exposed in the previous section.

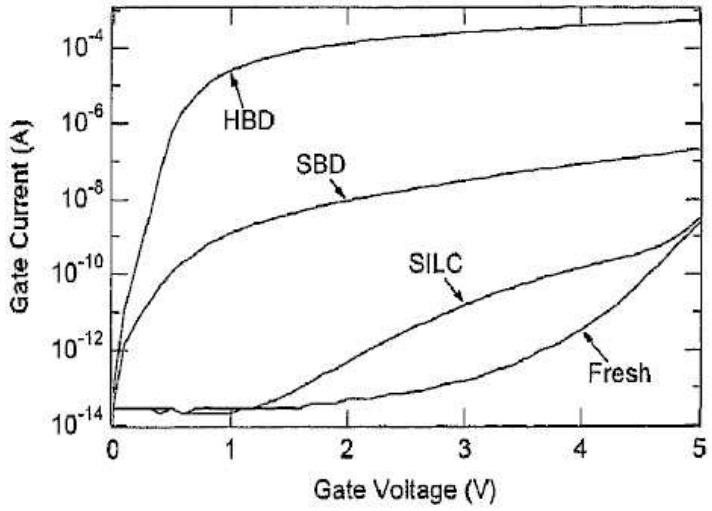

Figure 1.16: Gate leakage current of an 4.5 nm oxide. A comparison between the leakage of the fresh devices with the SILC, soft breakdown (SBD) and hard breakdown (HBD) successive modes after Degraeve *et al.* [26].

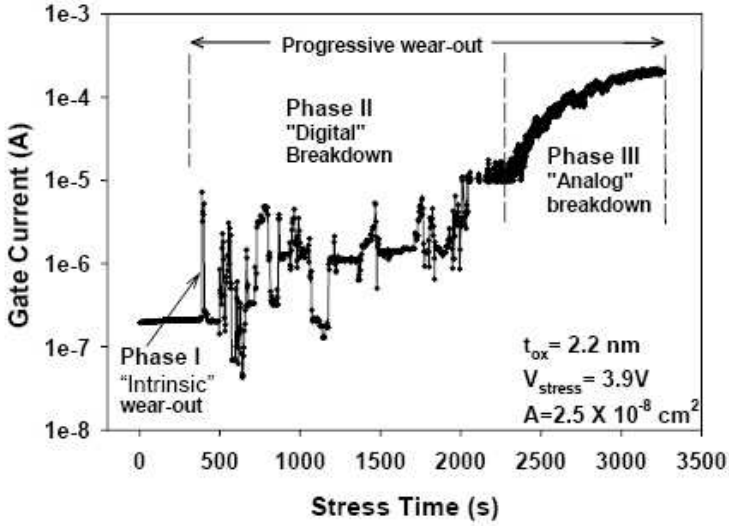

Figure 1.17: Ultra-thin oxide breakdown, with the Random Telegraph Noise (RTN) and progressive breakdown phase after Suehle *et al.* [27].

1.2.3 Dielectric reliability definition and tests methods

The dielectric reliability could be described as the capability of the insulator to keep its properties under operating condition for a defined period of time. The dielectric breakdown used as the failure criterion is defined as a local irreversible gate leakage current path.

To ensure the quality and the functionality of products over their dedicated operational lifetime, accelerated reliability tests have to be done to qualify products for the requested customers' specifications and to permit the release of good quality products. The qualification step should be performed rapidly and obviously much faster than the real product lifetime (typically in the order of years). To meet a competitive time-to-market development, methods have been developed which allow extrapolation of the lifetime and stress limits from accelerated electrical stress tests. Accelerated electrical stress means in general higher stress than standard operating modes. To end up with relevant predictive models, the degradation parameters as the underlying physical mechanisms should be understood and correctly modeled. A special attention is required in testing, as electrical over-stresses could result in different failure mechanisms than the ones operating under operational conditions.

The typical acceleration stresses used for dielectric reliability investigations are [39]:

- Constant Voltage Stress (CVS): during a CVS, the gate voltage stress is kept constant and the gate leakage is measured as a function of time. To monitor the breakdown event, generally sampling gate leakage measurements are also performed at a monitoring condition.

- Constant Current Stress (CCS): the gate current stress is kept constant and the gate voltage is measured as a function of time. This stress methodology was extensively used in the past for time-to-breakdown investigations of thick oxides. For thin oxides, the high leakage levels and the different breakdown behavior make this method less applicable to thin oxides [40].

- Voltage Ramp Stress (VRS): a voltage ramp is applied with a certain ramp rate factor and the gate leakage is measured as a function of the time. A lower level monitoring current leakage is generally sensed between each voltage step.

- Exponential Current Ramp Stress: an exponential current ramp is applied and the voltage is measured as a function of the time. (This procedure is also used as an ESD characterization tool to extract protection element parameters [41]).

The useful parameters obtained out of these electrical stress are the time-to-breakdown T_{BD}, the charge-to-breakdown Q_{BD}, the breakdown voltage V_{bd} and the breakdown field F_{BD}. In the following chapters the reference electrical stress is the CVS.

For CVS, the charge to breakdown is derived from the gate leakage measured during the stress and is given by:

$$Q_{\mathrm{BD}} = \frac{1}{A}\int_{t_0}^{T_{\mathrm{BD}}} I_{\mathrm{g,meas}}(t)\mathrm{d}t \tag{1.36}$$

where A is the gate oxide area (cm^2) and $I_{\mathrm{g,meas}}$ (A) is the gate current measured during the stress. Q_{BD} is expressed in $\mathrm{C/cm}^2$.

1.2.4 Oxide degradation under electrical stress

Under electrical stress conditions, the oxide is gradually getting deteriorated as a consequence of structural damage. The degradation is dependent on the process and on the quality of the oxide. Defects or precursor sites in the oxide exposed in Section 1.1.2.2 take part in the degradation mechanism. During high-voltage stressing, atomic bonds are broken, resulting in traps being generated inside of the oxide and at the interfaces.

1.2.4.1 Interface trap creation

Interface trap creation occurs at the substrate/oxide interface when the oxide is exposed to high field conditions [42, 7, 20]. The interface trap density (D_{it}) or the number of charges per unit area (N_{it}) generated during the operational device modes can be monitored with different characterization technics, capacitance-voltage characterization (C-V) [43], charge pumping (CP) [44, 45], low-voltage SILC [23, 24, 25].

The number of interface traps per area unit can be described by means of the occupation of acceptor and donor traps at the interface as a function of the surface potential Ψ_{S},

$$N_{\mathrm{it}}(\Psi_{\mathrm{S}}) = \int_{gap} e\cdot N_{\mathrm{it,Don}}(E)\cdot(1-f_{\mathrm{s}}(E,\Psi_{\mathrm{S}}))dE + \int_{gap}(-e)\cdot N_{\mathrm{it,Acc}}(E)\cdot f_{\mathrm{s}}(E,\Psi_{\mathrm{S}}) \tag{1.37}$$

with f_{s} the filling function calculated in Equation 1.30, referred to the Fermi level at the silicon/oxide surface. $N_{\mathrm{it,Don}}$ corresponds to the donor interface traps and $N_{\mathrm{it,Acc}}$ to the acceptor trap density.

The generated interface traps impact the threshold voltage in strong inversion for ($\Psi_{\mathrm{S}} = 2\Phi_{\mathrm{F}}$),

$$V_{\mathrm{th}} = V'_{\mathrm{FB}} + 2\Phi_{\mathrm{F}} + \gamma_{\mathrm{N}}\sqrt{2\Phi_{\mathrm{F}} - V_{\mathrm{BS}}} - \frac{q\cdot N_{\mathrm{it}}\Phi_{\mathrm{F}}}{C_{\mathrm{ox}}} \tag{1.38}$$

The interface traps result in transistor characteristic drifts, and, consequently, degrade the reliability of the device. Hot carrier stress interface traps creation and interaction with ESD stress will be discussed in detail in Section 5.1.

1.2.4.2 Oxide charge trapping

Electrical stress induced the trapping of oxide charges in the oxide volume, which could result in negative, positive or neutral states. In case of thick oxides under constant electrical stress, the gate leakage current density has been observed to change during the stress. This change has been attributed to the hole or electron trapping, leading to a modified electrical field in the oxide near the cathode, causing an increase or a decrease in the current gate leakage. This direct observation of oxide trapping can be monitored in the gate leakage current density under a CVS (or in the monitoring voltage level of an CCS). At first, the gate leakage under CVS typically increases as a function of the time due to positive trapping, then, in a second step, the leakage level decreases resulting from electron trapping. Finally, an abrupt increase of the leakage occurs which corresponds to the breakdown of the oxide [42, 46, 47]. The hypothesis that a number of critical defects generated in the oxide was leading to the oxide breakdown was then naturally proposed. However, the nature of the traps causing the breakdown is still a subject of discussion.

A review of the nature of the traps and their reported/proposed generation mechanisms will be outlined in the following section.

1.2.4.3 Oxide trap generation Mechanisms

- **Band gap impact ionization:** Historically, hole trapping have been explained by impact ionization for thick oxides [20]. Direct electron-hole recombination can occur during the steady state Fowler-Nordheim tunneling when the injected carrier energy exceeds the 9 eV SiO_2 band-gap (gate voltage above 12 V).

- **Anode Hole Injection (AHI):** AHI can occur when the voltage across the oxide is above 6 V. An electron entering into the anode could create an electron-hole pair from impact ionization. The hole thereby generated is injected back in the oxide and generates oxide and interface traps [48]. The AHI is known to be the origin of creation of positive and neutral traps in the oxide. The AHI requires an energy threshold of about 6 eV, this corresponds to the energy required to generate an electron-hole pair (E_g) plus the valance band offset between the silicon and oxide (4.8 eV).

  Hole fluence

  The substrate current of thick oxides at high field is attributed to this anode holes injection and can be experimentally observed by means of carrier separation technique [49].

  – Thick oxides

    For thick oxide (> 5 nm) the substrate current could be modeled with a hot hole current, resulting from the impact ionization caused by the high energetically electron at the anode interface [50, 51]. This current is described by the mode (1) in

Figure 1.18. This generated hot hole current was found to be proportional to the gate electron current in the Fowler-Nordheim tunneling regime.

$$J_{\mathrm{h}} = \alpha_{\mathrm{h}} \cdot J_{\mathrm{n}} \tag{1.39}$$

with J_{n} representing the gate electron current density and J_{h} the hole current density. α_{h} is the hole generation efficiency which is field dependent.

$$\alpha_{\mathrm{h}} = \alpha_0 \exp\left(-\frac{H(T_{\mathrm{ox}})}{E_{\mathrm{ox}}}\right) \tag{1.40}$$

where $H(T_{\mathrm{ox}})$ is a constant value depending on the oxide thickness. For oxide thicker than 10 nm this constant has been found to be around 80 MV/cm [52], for thinner oxide this constant is increasing and was estimated in the range of 130 MV/cm for 5 nm oxides [52]. This hot hole current is presented in the Figure 1.19 for an 5.2 nm nFET.

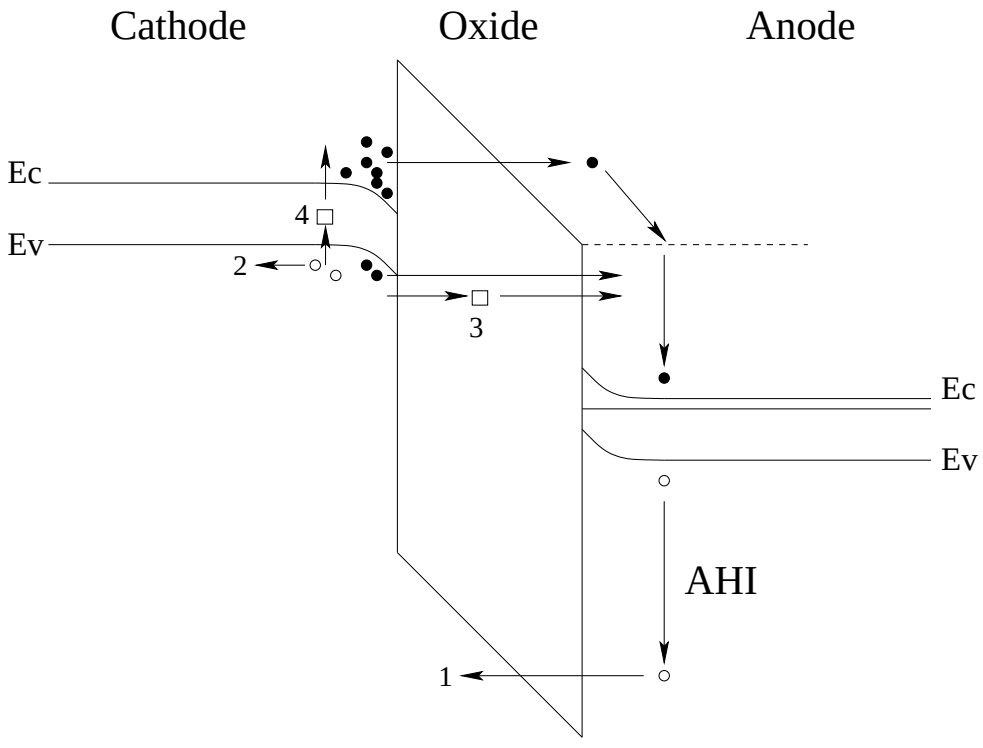

Figure 1.18: Substrate current sensed at the anode of an nFET operating in inversion. This current is generated by the hole fluence resulting from (1) hot hole created by impact ionization (AHI) for thick oxide under high field; (2) resulting from electron tunneling from the valence band of the cathode for thin oxide; (3) trap-assisted electron valence band tunneling; (4) a minority recombination current at the cathode interface.

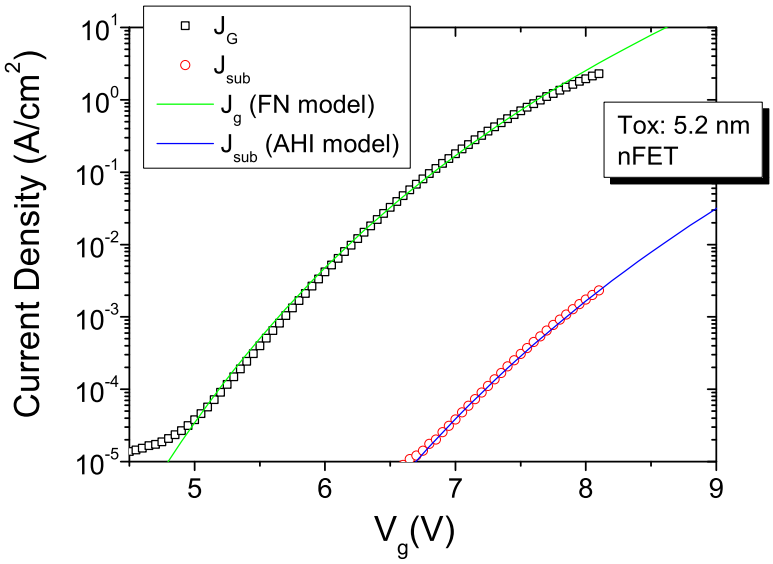

Figure 1.19: Gate and substrate leakage of an 5.2 nm nFET. Gate current is modeled with the Folwer-Nordheim Tunneling equation and the substrate current accordingly to the AHI model.

- Thin oxides

  For medium to thin oxides, the substrate current is a composition of electron valence band tunneling and backward hole injection and minority carriers recombination at the interface [51]. In case of an DT current for thin oxides, the substrate current is attributed mainly to the electron tunneling current from the valence band [49, 12], see Figure 1.18. The substrate current can be modeled considering the electron current flow with the DT tunneling equation, the effective masse of electrons in the oxide coming from the valence band ($m_{ox} \sim 0.35 \cdot m_{e-}$) and the effective barrier high ϕ_B in the order of 3.9 eV [12], see Figure 1.20.

  The majority ionization still occurs for $V_{ox} < 6$ V, but at lower rate. However, the extension of the AHI model has been proposed based on the minority impact ionization [53, 54] in case where holes are present in the anode. This theory predicts the possibility to get AHI until a gate voltage of around 3.5 V.

- **Anode hydrogen Release (AHR):**
  Anode hydrogen desorption occurs when the kinetic energy of the injected carriers exceeds 2 eV (5 V) [51]. This release of hydrogen species into the oxide is generating traps. Both, interface trap generation and hydrogen transport in the oxide has been experimentally revealed to occur above this threshold voltage [55]. The only presence of an hydrogen plasma in the environment of the oxide has been seen to result in an increase gate leakage current similar to the current during voltage stress [51, 56].

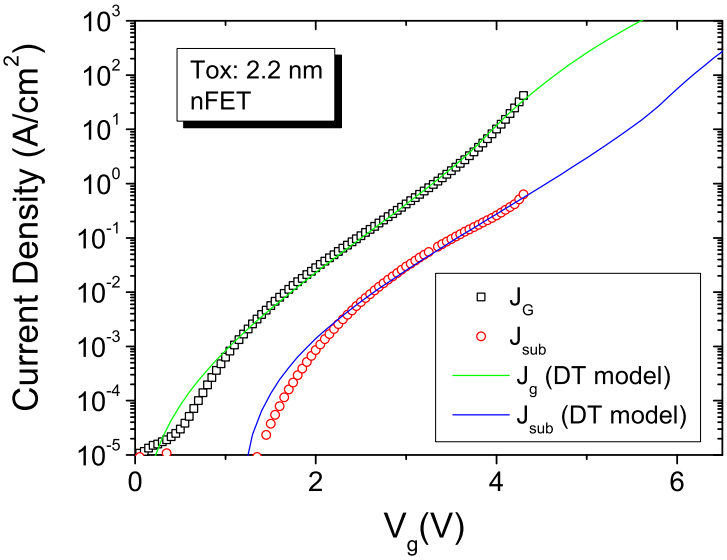

Figure 1.20: Gate and substrate leakage of an 2.2 nm nFET device. Gate and substrate current are modeled with the Direct-Tunneling equation.

- **Multi-Vibrational Hydrogen Release (MVHR):**
  Hydrogen bonds can be also broken via vibrational excitations. It has been observed directly under Scanning Tunneling Microscopy (STM) excitation [57, 58, 59, 60, 61, 62, 63, 64]. There are two silicon hydrogen bond desorption mechanisms, the Electrical Excitation (EE) and the Vibrational Excitation (VE).

  – EE occurs via field emission [59, 65] with a threshold energy of around 5 to 6 eV. This mechanism is weakly dependent on the energy and current.

  – VE consists in the desorption of the hydrogen via excitation phonon modes [57, 58, 59, 61]. The threshold energy activation is about 2.5–3.0 eV and in this case the mechanism is strongly dependent on the voltage [66]. Below 2.5 eV, bond breakage resulting from multi-vibrational excitations of at least two electrons is proposed [67, 65, 68].

1.2.4.4 Stress Induce Leakage Current

For oxides in the range between 3–7 nm, an increased leakage current is observed prior to the soft-breakdown mode. This increase in leakage is called SILC and is based on a conduction mechanism via oxide traps, see Section 1.2.1. For thin oxides below 3 nm, this conductive trap

assisted mechanism is not seen anymore as the possible charges trapped in the oxide are de-trapped as the oxide thickness is in the range of the direct tunneling mechanism. However, some interfacial trap assisted tunneling could still be seen below the flat band voltage (LV-SILC). The SILC generation has been extensively studied [69, 21] as it causes some retention problem in memories (non-volatile memory). The SILC has been reported to be linked with neutral or positive traps [52, 69]. A possible link between a critical amount of SILC generated and the breakdown of dielectric has been proposed in different correlation studies [70, 71, 25]. However, although these correlations are discussed controversially [72], the occurence of SILC indicates an evident proof of the oxide degradation.

1.2.4.5 Oxide trap generation rate

The oxide defect generation rate have been studied by DiMaria and Stathis over a wide range of oxide thicknesses by means of *C-V* and SILC measurements [73, 74]. They have found that the defect generation rate (P_g) is voltage dependent and strongly decreases below a threshold voltage value of around 6 V, (Figure 1.21 below). A similar behavior has been also recently reported from Suñé and Wu considering this defect probability generation (ζ) plotted as a function of the maximal energy released at the anode by the injected electron [75]. This is exposed in the Figure 1.22.

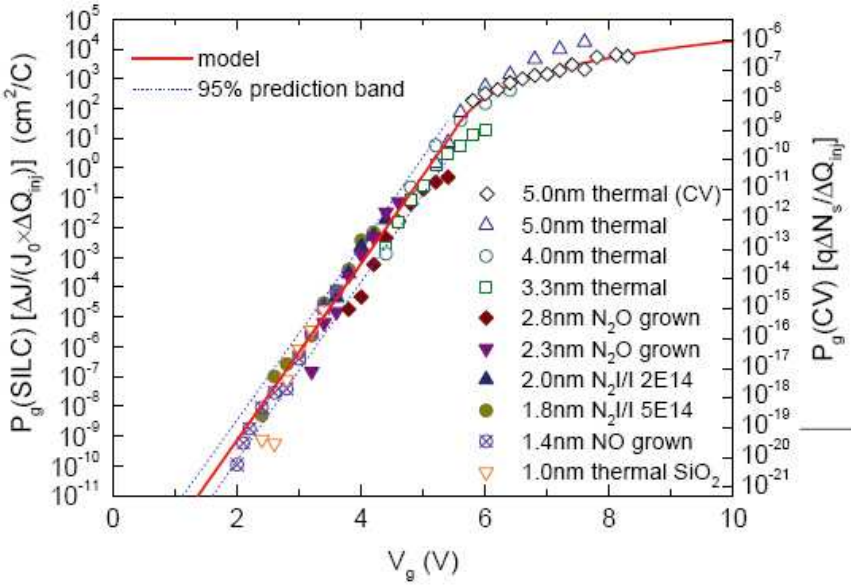

Figure 1.21: Trap and interface-state generation probability as a function of gate voltage under FN stress on nFETs with gate oxides of varying thickness from [73].

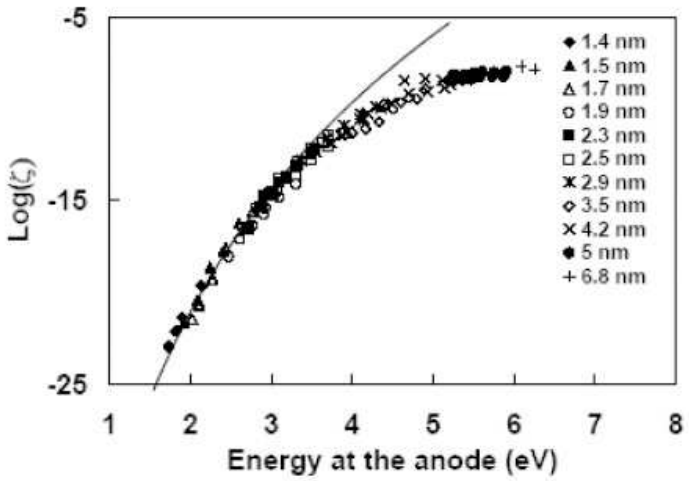

Figure 1.22: Defect generation efficiency ζ as a function of the maximal energy after [75].

1.2.4.6 The traps that cause breakdown

Derived from the different trapping behavior observed, several hypothesis have been formulated [1]. Some searchers have proposed a breakdown driven by the hole fluence with a critical density of holes as leading to the breakdown [50]. This was suggested as it was observed that the total charge of holes at breakdown is independent on the field. Then, by means of Substrate Electron Injection (SHEI) technique [76], the filling of neutral traps by electrons has shown that the density of neutral traps is proportional to the field and that a critical electron trap density is required to trigger the breakdown. Moreover, as the electron trap generation was reported to be constant at all fields, a correlation between the hole fluence and electron trapping has been deduced. This finding has lead to the establishment of the percolation theory idea which is based on the random generation of electron neutral traps up to a critical density, provoking the oxide breakdown [76, 77]. Nowadays the percolation idea is well accepted even if the nature of the traps causing the breakdown is still unclear. For thin oxides the role of hydrogen species is extensively pointed out for oxide degradation. A more appreciable overview on the oxide breakdown mechanisms will be exposed with the dielectric breakdown models.

1.2.5 Dielectric breakdown statistics

1.2.5.1 Statistical reliability basis

The probability that a device will fail in the time interval $[0,t]$ is given by the cumulative failure distribution $F(t) = 1 - R(t)$ where $R(t)$ corresponds to the probability that the device will still be functional under operating condition for a given time period.

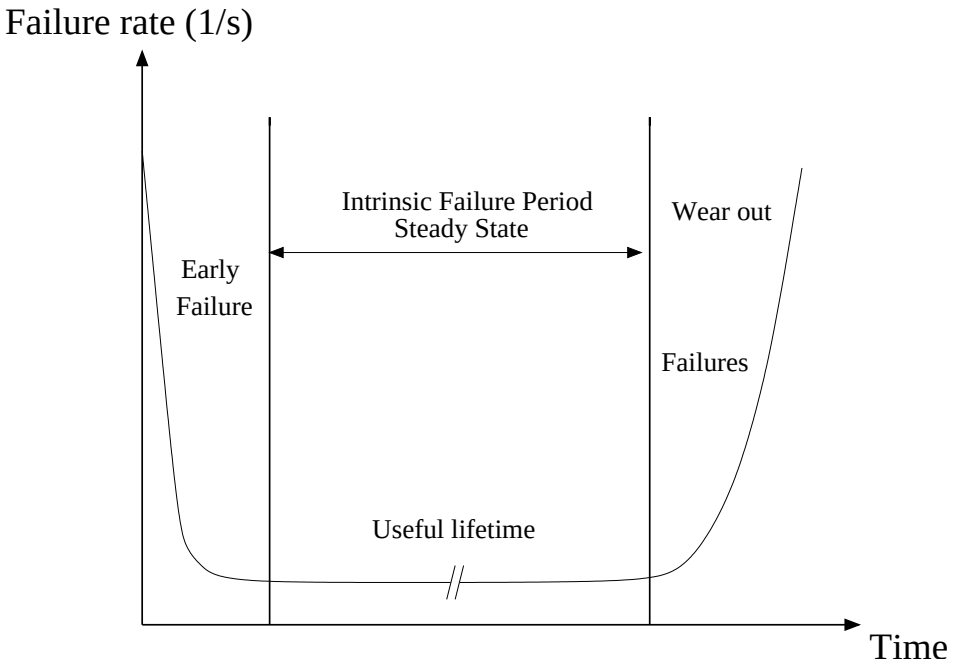

Figure 1.23: Typical failure rate as a function of the time.

The probability density function $f(t)$ of a device to fail in the time interval dt at time t, when the device was operating at $t = 0$ is given by:

$$f(t) = \frac{\mathrm{d}F(t)}{\mathrm{d}t} \tag{1.41}$$

The failure rate $\lambda(t)$ of devices having survived to time t (considering only non-repairable populations) is defined as:

$$\lambda(t) = \frac{f(t)}{(1-F(t))} \tag{1.42}$$

Independently from the nature of the defective entity of interest, across a wide variety of stressing configurations (mechanical, electrical, ...) the empirical failure rate of the population under studies within the times has always shown the same characteristic "bath" curve shape sketched in Figure 1.23. This model is universally valid for animal lifetimes as well as for the dielectric breakdown behavior under electrical stress. The breakdown phenomenon has a statistical nature which should be accounted for.

1.2.5.2 Dielectric breakdown statistics

The dielectric breakdown under electrical stress is driven by the so called percolation theory [78, 76, 77, 79]. The gate oxide breakdown is caused by the generation of defects in the dielectric. As soon as a critical amount of defects is reached, a conductive percolation path between the anode and the cathode via clusters of charges provokes the breakdown of the insulator (Figure 1.24).

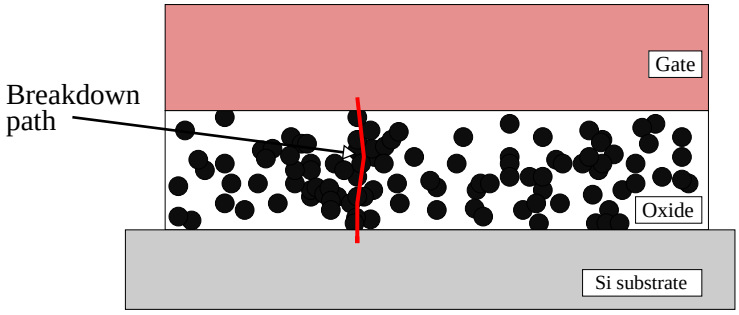

Figure 1.24: Schematic percolation breakdown path between electrode.

Under electrical stress, the time-to-breakdown depends on the stress level. For a fix stress level, a CVS for example, the time-to-breakdown distribution resulting from a population composed of identical devices stressed at the same temperature can be modeled by a statistical law. The statistical nature of time-dependent dielectric breakdown (TDDB) is well described by the Weibull extreme value statistic [78, 80, 81, 12, 82, 83, 84]. The Weibull distribution function is based on the concept that the reliability of a system depends on the lifetime of the weakest link. For a dielectric under electrical stress, this weakest link corresponds to the formation of the percolation path resulting from a random defect generation process in the oxide. The cumulative distribution of time to formation of the first percolation path, which is considered to be the first irreversible and localized leakage current path, is consequently described by the Weibull distribution:

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\tau}\right)^{\beta}\right] \tag{1.43}$$

$$\text{where, } F(t=\tau) = 0.632 = F(T_{63\%})$$

The parameter τ is usually referred as $T_{63\%}$ and corresponds to the time where 63.2 % of the samples have failed. This is the characteristic parameter of the Weibull distribution. The parameter β is the shape parameter of the distribution called Weibull slope. For graphical linearization, Equation 1.43 is written as follows, which results in the so called Weibull plot:

$$\ln(-\ln(1-F)) = \beta\,\ln(t) - \beta\,\ln(\tau) \tag{1.44}$$

The probability density function, $f(t)$ based on the Weibull model is:

$$f(t) = \frac{\beta}{\tau}\left(\frac{t}{\tau}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\tau}\right)\right] \tag{1.45}$$

Graphical representations of the Weibull characteristic probability density and the cumulative failure density are shown in Fig 1.25.

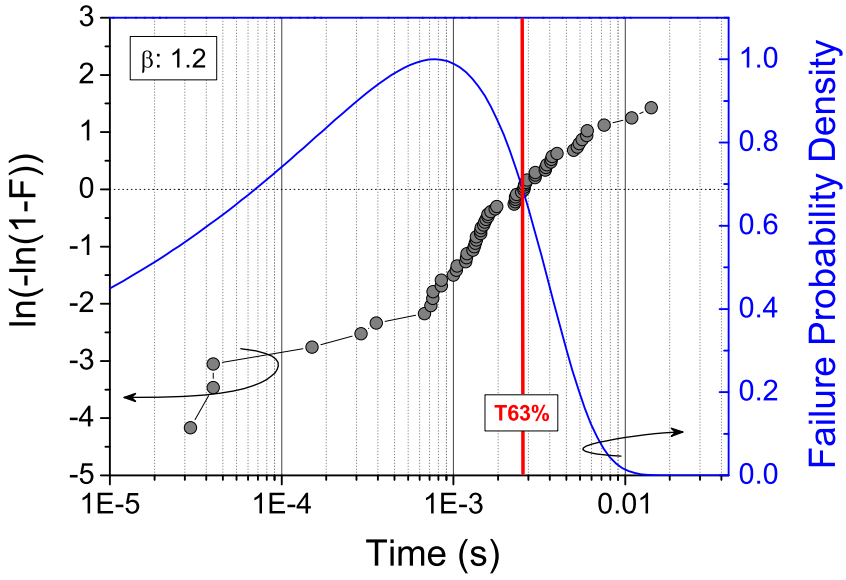

Figure 1.25: A cumulative failure distribution with a slope (β) of 1.2 is plotted in a Weibull plot. The derived probability failure density from this distribution is also depicted in the graph.

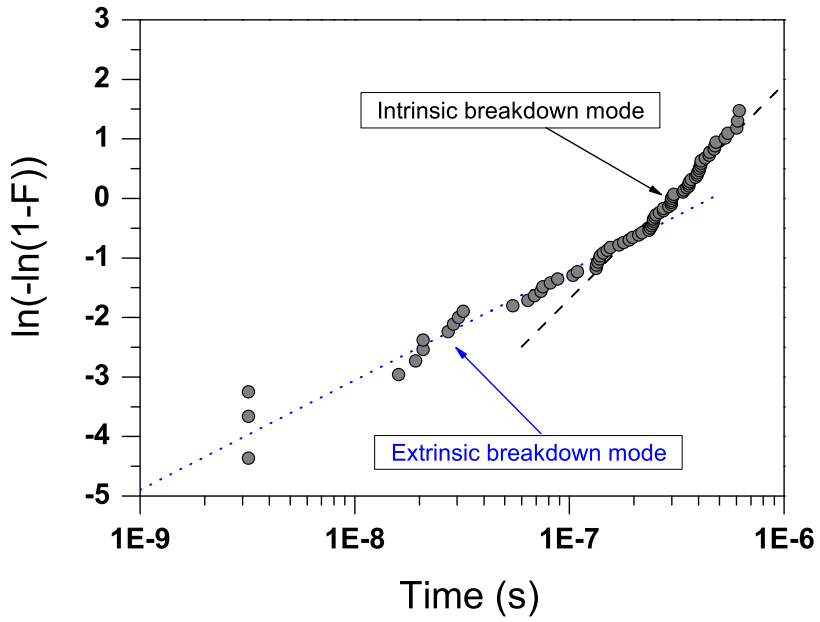

Figure 1.26: Bimodal cumulative failure distribution of an oxide stressed under CVS.

Two distinct cause of breakdown should be distinguished:

- Extrinsic fails, which are caused by defects created by external causes to the process, structural weaknesses, freak devices, impurities or early breakdown during electric stress. These fails compose the early failure mode described in Figure 1.23.
- Intrinsic fails, composed of random failures resulting from the dielectric wear-out under electrical stress and which are inherent to the material and process.

These two distinct modes [85, 26] could be seen in the bimodal cumulative failure distribution shown in the Weibull plot, Figure 1.26. Extrinsic fails or infantile fails have a failure rate which decreases with time and are described by low Weibull slopes, $\beta < 1$. In contrast, the intrinsic Weibull slopes are related to the dielectric wear-out and are characterized by an increasing failure rate, i.e. $\beta \geq 1$.

The intrinsic Weibull slope β is dependent on the oxide thickness and is decreasing for thinner oxides, Figure 1.27 [75]. In this thesis, only intrinsic breakdown modes will be considered, as extrinsic fails are linked to process quality and should not be correlated with the wear-out behavior of the dielectrics material under electrical stress.

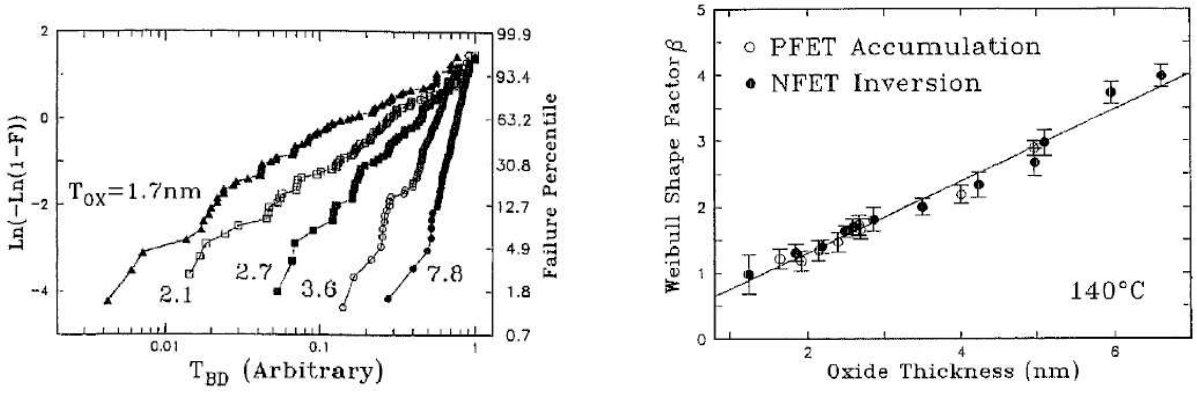

Figure 1.27: Weibull slope dependence with the oxide thickness after Wu *et al.* [86], data at 140 °C.

In the case where no extrinsic defect occurs in the GOX, the Poisson statistic applies to TDDB [78, 80, 40, 87], because the spatial occurrence of the percolation path is a random phenomenon. Thus, the statistical distribution of time to dielectric breakdown depends on the active GOX area of the stressed device. For randomly distributed fails, the probability $P(n)$ of having n defects on the area A with a failure density D is given by the Poisson equation:

$$P(n) = \frac{(AD)^n}{n!}\ \mathrm{e}^{-AD} \tag{1.46}$$

In case of no failure, $\mathrm{e}^{-AD} = R(t) = 1 - F(t)$ and thus $\ln(AD) = \ln(-\ln(1-F)) = W$

For two area A_1 and A_2 with the respective time to fail T_1 and T_2,

$$W_1 - W_2 = \ln\left(\frac{A_1}{A_2}\right) = -\beta \cdot (\ln(T_1) - \ln(T_2)) \tag{1.47}$$

$$\text{which leads then to,} \quad \frac{T_1}{T_2} = \left(\frac{A_1}{A_2}\right)^{-\frac{1}{\beta}} \tag{1.48}$$

Gate oxide time-to-breakdown dependence with active gate oxide area at different stressing voltage level is graphically represented in Figure 1.28 [87]. The time to breakdown of small capacitors is bigger than the time to fail of large capacitor. The time to breakdown of similar structures stressed in the same conditions is dependent on the stress level; higher is the stress and lower is the time to fail, Figure 1.28.

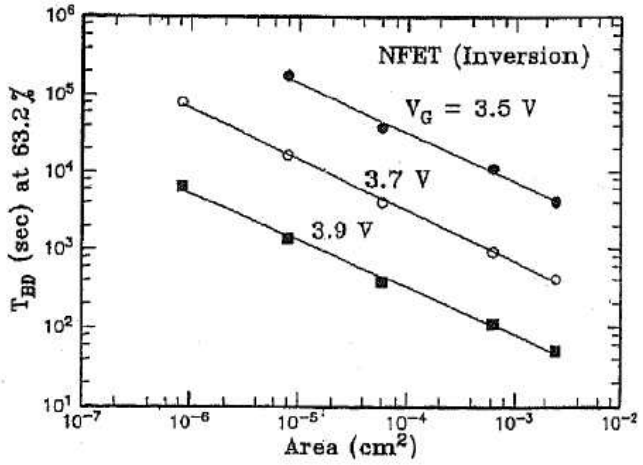

Figure 1.28: Oxide time-to-breakdown area dependence from [87].

The dielectric time to breakdown is temperature dependent. At a fixed electrical stress, an increase of the temperature is found to lead to a reduction in the time to breakdown, Figure 1.29 [86].

The oxide time-to-breakdown parameters required for the understanding of the gate oxide breakdown and therefore for the lifetime prediction from accelerated stress tests are:

- the statistical parameters composed of the shape of the Weibull distribution β and its characteristic parameter $T_{63\%}$;
- the stress condition parameters: voltage, temperature and gate oxide active area;
- the TDDB acceleration models with respect to the stress level, the temperature and gate oxide area.

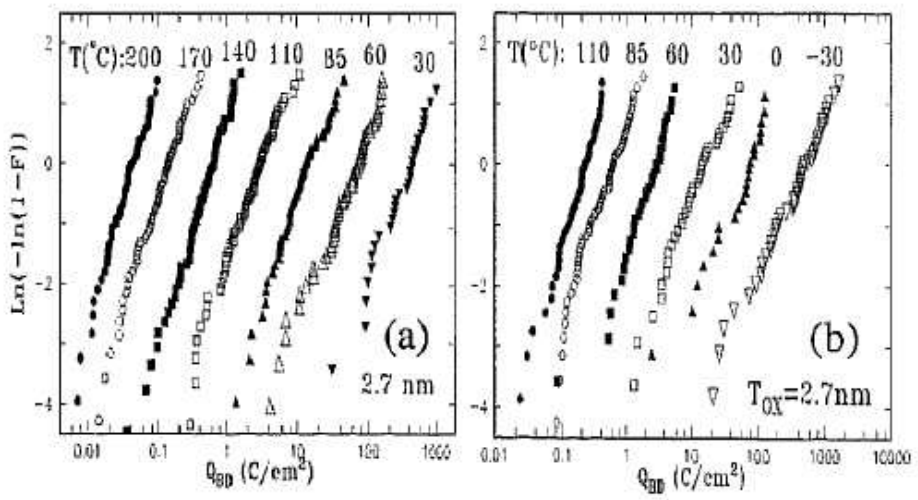

Figure 1.29: Temperature impact on the charge to breakdown Weibull distributions stressed with CVS at 4 V from [86].

1.2.6 Dielectric breakdown modeling to lifetime prediction

1.2.6.1 Percolation statistical breakdown models

In order to describe the experimental observations, models of the percolation breakdown mechanism including the empirical parameters dependence which have an impact in the time to breakdown have been extensively studied. Different approaches have been proposed. Historically, Suñe and co-workers [88] have proposed a two dimensional cell based percolation model in 1990 which was also adopted by Dumin's group [89, 90, 91]. This simple geometrical model was based on 3 items:

- the oxide area A_{ox} was dived in a 2D cells array of area A_0;
- defects points were randomly distributed accordingly to the Poisson's statistics;
- the breakdown was based on the weakest-link phenomenon character and was considered when a critical amount of defects n_{BD} has been generated in one of the 2D cells.

The cumulative breakdown distribution was given as a function of the oxide density defects N_{ot}:

$$\ln(-\ln(1-F)) = \ln\left(\frac{A_{ox}}{A_0}\right) + \ln\left[N_{ot}A_0 t_{ox} - \ln\left[\sum_{k=0}^{n_{BD}-1}\frac{(N_t A_0 t_{ox})^k}{k!}\right]\right] \tag{1.49}$$

This model was able to fit the breakdown considering the defect generation dynamic quite well, but was lacking a predictive modeling concerning the oxide thickness dependency. The geometrical link between the size of defects and the defect number to provoke the breakdown was

introduced by Degraeve *et al.*, with a spherical based defects model [81]. This 3D model reduces the breakdown variable to the unique defect radius r. Using Monte Carlo simulation for the random generation of defects, making the assumption that a conductive path is created when two sphere are in contact and considering the first linked conductive chain joining the two electrode as the breakdown event, this model was successful to verify the Weibull statistical nature of the breakdown distribution. Moreover this model has shown the consistent picture of the Weibull slope dependence with the oxide thickness.

A similar approach with cubes instead of spheres was developed by Stathis [77], which was also successful to simulate the percolation mechanism. However, there was one major point which was different between both models: the size of the defects. In the spherical model, the radius of a defect deduced by Degraeve was 0.45 nm in comparison to the unrealistic cubic sides of 2.7 nm for ultra-thin oxide found by Stathis. Then, an analytic 3D cube based percolation model was proposed by Suñe in 2001 [92]. In this model the oxide is defined by cubic cells with a constant lattice a_0 and the average fraction of defective cells as:

$\lambda = \dfrac{N_s a_0^3}{t_{ox}}$ or $\lambda = N_t a_0^3$, where N_s and N_t are the surface and the volume densities of defects.

The probability for a cell to become a trap is considered to be uniform and writes as:

$$F_{cell}(\lambda) = \lambda \tag{1.50}$$

Assuming independent probability of failure for one column of cells,

$$F_{col}(\lambda) = [F_{cell}(\lambda)]^n = \lambda^n \tag{1.51}$$

where the number of cells in a column is $n = T_{ox}/a_0$. The total number of cells composing the oxide are given by $N = A_{ox}/a_0^2$. The cumulative breakdown function express in an Weibull value form is then:

$$\begin{aligned} W_{BD}(\lambda) &= \ln(-\ln([1-F_{col}(\lambda)]^N)) \\ &= \ln(-\ln([1-\lambda^n]^N)) \\ &= \ln(-N\ln([1-\lambda^n])) \end{aligned} \tag{1.52}$$

Under the assumption that $\lambda \ll 1$ at the very first moment of the breakdown, $\ln(1-\lambda_{BD}) \sim -\lambda_{BD}$, the expression of the Weibull value W_{BD} is:

$$W_{BD}(\lambda) = \ln(N) + n_{BD}\ln(\lambda) \tag{1.53}$$

$$W_{BD}(N) = \ln\left(\frac{A_{ox}}{a_0^2}\right) + \frac{t_{ox}}{a_0}\ln(a_0^3 N) \tag{1.54}$$

From the definition of the Weibull statistics plotted in a Weibull form, the parameter β could be deduced as equivalent to $n_{BD} = T_{ox}/a_0$. The average density of defects N_{BD} at breakdown is obtained from the condition $W_{BD}(N_{BD}) = 0$:

$$N_{BD} = \frac{1}{a_0^3} \exp\left[-\frac{a_0}{t_{ox}} \ln\left(\frac{A_{ox}}{a_0^2}\right)\right] \tag{1.55}$$

Assuming a linear evolution of the defects density with time,

$$N(t) = \frac{\zeta Q}{q t_{ox}} \text{ with } Q = J(t_{ox}) \cdot t \tag{1.56}$$

where ζ is the defect generation efficiency [93, 66], q the electron charge, $J(t_{ox})$ the current density and Q the injected charge per unit of area (cumulative fluency). Consequently, the mean charge to breakdown is given by:

$$Q_{BD} = \frac{q t_{ox}}{a_0^3 \zeta} \exp\left[-\frac{a_0}{t_{ox}} \ln\left(\frac{A_{ox}}{a_0^2}\right)\right] \tag{1.57}$$

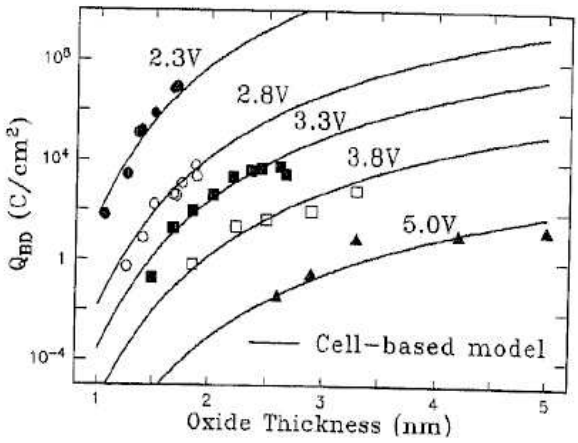

Figure 1.30: Charge to breakdown modeled based on the cell based approach from Wu *et al.* [75]

In this model, the defect trap size is then deduced from the trend of the decreasing Weibull slope versus the oxide thickness. As the Weibull slopes reported for oxide down to 1 nm are equal or above 1, an interfacial layer covering this offset has been proposed in [93, 75],

$$\beta = \frac{m}{a_0} \cdot (T_{ox} + T_{int}) \tag{1.58}$$

assuming m to be 1, a defect size of a_0 = 1.83 nm is deduced and the interfacial layer thickness T_{int} is supposed to be 0.37 nm. This defect size is discussed from a recent theory which suggest that the GOX breakdown requires at least one bulk trap and one interface trap [94, 65]. In this model the coefficient m is derived from trap generation kinetic of hydrogen species. As a

mixture of bulk and interface traps are required for the breakdown, a single defect size cannot be extracted from the Weibull slope β. This model do not consider an interfacial offset but a defect size accordingly to the following equation (1.59), which gives a size of defect of 0.4 nm.

$$a_0 = m \cdot \frac{\beta}{t_{ox}} \text{ with } m = 0.26 \tag{1.59}$$

A more complex analytic extension of the cell-based model has been recently proposed by Krishnana [79] accounting for non-columnar conduction. In this model, the impact of pre-existing defects and oxide lattice misalignment were simulated to evaluate the impact on the Weibull slope.

$$\beta = \frac{\alpha}{a_0} \cdot T_{ox} + G \tag{1.60}$$

Despite of its slight evolution in its mathematical-statistical description, the percolation idea is the base of the breakdown of oxide films. The experimental results reporting oxide degradation and the Weibull slope reduction with decreasing oxide thickness are over all well described by the percolation models. However, remaining questions concerning the mechanisms which drive the degradation and the oxide breakdown are still unclear. This knowledge is, however, required for extrapolation of the accelerated electrical stresses to the reference operation conditions. In the next section the different theories and model will be discussed.

1.2.6.2 Dielectric time-to-breakdown models and mechanisms

Although the root cause of the dielectric breakdown (percolation theory) is nowadays well accepted, the generation and the nature of the traps leading to the breakdown is still a matter of controversial discussions. Different dielectric time-to-breakdown acceleration models have been proposed, coming along with different observations and hypothesis. These models proposed are all consistent with the statistical nature of the dielectric breakdown described by the Weibull statistics, but are based on different explanations and consider different acceleration parameters. A clear cesura splits the breakdown behavior of the thick oxides from the breakdown of ultra-thin ones from advance ULSI CMOS processes. A review of the different models will be chronologically emphasized consequently in two distinct part in the following paragraphs. In the first part, the modeling of the thick oxides breakdown will be presented, then the new breakdown theories for thin oxides will be described.

A - Thick oxides TDDB modeling

In the case of thick oxide ($>$ 7nm), basically two opposite main models have been proposed, either field (E) dependent or inverse of the field ($1/E$) driven.

***a) E*-model**

Historically, the first empirical models were based on a field-dependent acceleration model. An early model was introduced in 1979 by Crook [95, 96] and Anolick [97], describing the TDDB as an exponential function of the electric field:

$$t = t_0 \exp\left(-\alpha \left(E_{\mathrm{ox}} - E_0\right)\right) \tag{1.61}$$

Then McPherson proposed in 1985 the "thermo-chemical model" [98] which gives a physical background for the empirical field exponential dependence. The thermo-chemical model is based on the break of Si–Si bonds due to dipole-field interactions. The breakdown precursor defects are the oxygen vacancies which are leading to high polar O–Si bonds and weak Si–Si bonds. In an electric field, a distortion of the lattice is induced by the polar O–Si bonds which are generating local dipolar fields larger than the applied field across the dielectric. This lowers the enthalpy of the Si–Si bonds up to the breakdown of the bond, leading to local conductive sites which will participate to the percolation path elaboration.

$$\ln(t) \sim \frac{\Delta H}{k_{\mathrm{B}}} - \gamma E_{\mathrm{ox}} \tag{1.62}$$

Kimura *et al* have reported data comfirming the *E* model [99, 100]. The Time Dependent Dielectric Breakdown data from Kimura obtained on 10 nm thick oxide are reported in Figure 1.31. Mizubayashi and co-workers have still reported in 2002 a TDDB *E*-model dependence for thinner oxide than 5 nm [101].

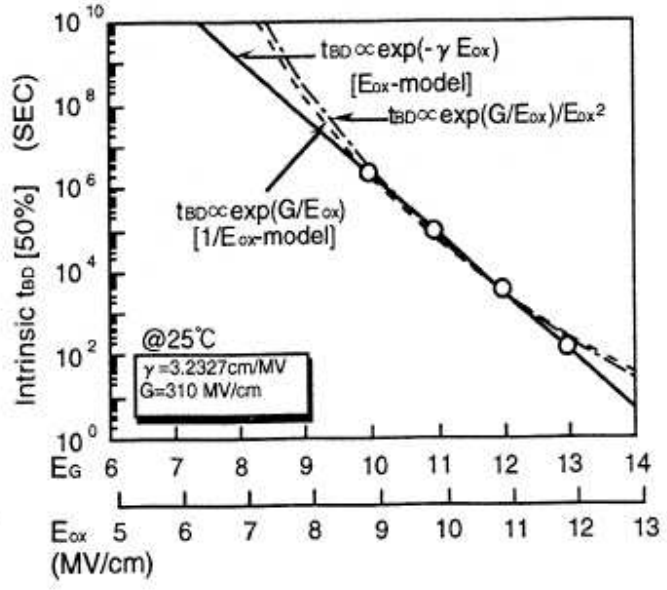

Figure 1.31: TDDB data plotted accordingly to the E-model after Kimura *et al* [100].

Despite the fact that a lot of research work has shown a great correlation between the injected charges or trap generation in the oxide [78, 42, 20, 26, 77] and the breakdown event, this phenomena is neglected in this model. As the physical background of the "thermo-chemical model" is not matching the well accepted percolation theory, McPherson has proposed later a new theory the "molecular model" in 2000, accounting for this phenomenon [102]. In this model, both field and charge transport in the dielectric are considered. For low bond strength (<3 eV), the E-model behavior is valid and for high bond strength, the bond breakage must be catalyzed with hole and then the $1/E$ model dominates for high stress current densities. The injected carriers are accounted for catalyzing the O–Si bond break. More recently, as the role of hydrogen species seems to play an important role in the breakdown of thin oxides, he proposed in 2006 an extension of the molecular model [103] accounting for the critical role of hydrogen release for the breakdown.

b) $1/E$**-model.**

An other empirical model has been introduced in 1985 by the UC Berkley group [104]; it is based on the exponential dependence of the time-to-breakdown regarding the inverse of the field as expressed here:

$$\ln(t) \sim G\frac{1}{E_{\text{ox}}} \tag{1.63}$$

This model has been explained by the release of species at the anode responsible for defects generation in the oxide. Two theories both based on the Fowler-Nordheim tunneling current mode were proposed. The Anode Hole Injection (AHI) model and the Anode Hydrogen Release (AHR) model.

- the Anode Hole Injection Model (AHI) was proposed by the Berkley university in 1988 [105, 106] for oxide thicker than 10 nm. In this model the injected electrons from the cathode into the oxide are supposed to generate electron-hole pairs when reaching the anode by releasing their energy. The hot holes highly energetics generated at the anode side are then injected back into the oxide and are creating defects in there by hole trapping. This hole defect density generated is dependent on the hole fluency and on the hole generation coefficient α_{h} (see 1.39) [50, 26] and can be express as,

$$\begin{aligned} Q_{\text{p}} &= \alpha_{\text{h}}(E_{\text{ox}}) \cdot Q_{\text{n}} \\ &= \alpha_{\text{h}} \cdot J_{\text{FN}} \cdot t \end{aligned} \tag{1.64}$$

  J_{FN} is the Fowler-Nordheim current density given via the formula 1.32 in Section 1.2.1. The breakdown is given when the critical amount of positive charge $Q_{\text{p,crit}}$ is reached in the oxide and the charge to breakdown is given by,

$$Q_{\text{BD}} = \frac{Q_{\text{p,crit}}}{\alpha_0} \cdot \exp\left(\frac{H}{E_{\text{ox}}}\right) \tag{1.65}$$

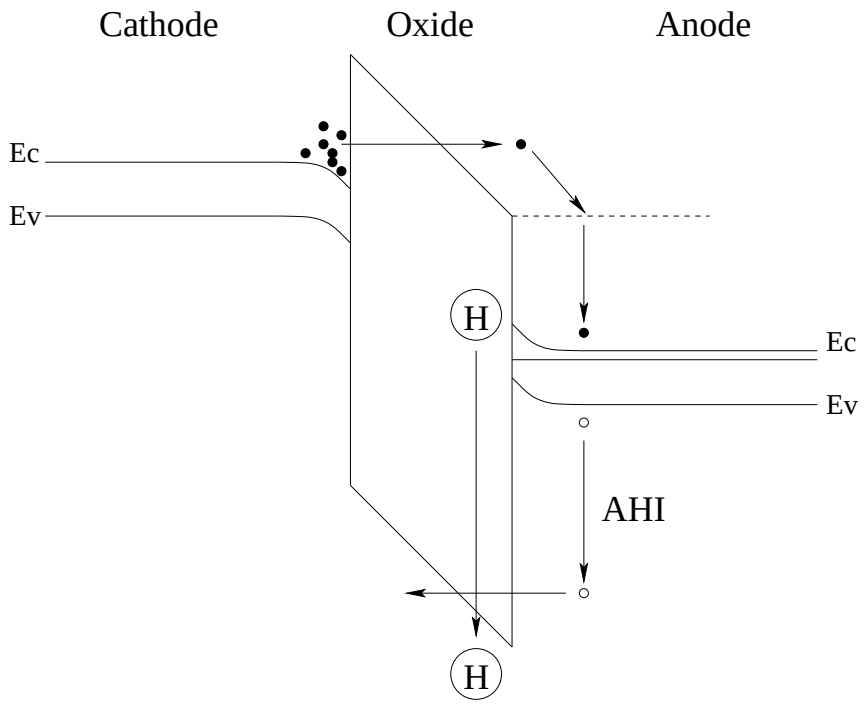

Figure 1.32: Schematic band diagram exposing the two formulated explanation for the $1/E$ trend of TDDB. The Anode Hole Injection (AHI) model and the Anode Hydrogen Release (AHR) model

The critical amount of charge to breakdown was reported to be independent of the oxide thickness and in the order of 10 C / cm^2 [107].

Consequently the time-to-breakdown is

$$t_{\mathrm{BD}} = \frac{Q_{\mathrm{BD}}}{J_{\mathrm{FN}}} = \tau_0 \cdot \exp\left(\frac{G}{E_{\mathrm{ox}}}\right) \text{ with } G = B + H \sim 320\,MV/cm \tag{1.66}$$

TDDB data obtained comforting this theory have been mainly exposed by Schuegraf *et al.* [108, 50, 48], see Figure 1.33. For thinner oxides and at low voltage some question regarding this model have arisen as the impact ionization required at least a voltage across the oxide above 6 V. However to support the $1/E$ model data observed for thin oxides, some searchers have proposed to extend the AHI procedure mechanism to lower conditions, based on the minority carrier impact ionization theory [53]. However, the AHI hole fluence could not be measured directly in the substrate current anymore for thin oxides due to the electron valence band tunneling [50].

- The Anode Hydrogen Release Model (AHR) introduced by IBM Dimaria [20, 51] is based in this case not on the re-injection of holes in the oxide, but is considering the diffusion of hydrogen towards the cathode which causes traps in the oxide and leading to the breakdown path. The hydrogen release triggering is based on the same electron energy loss at the anode side mechanism.

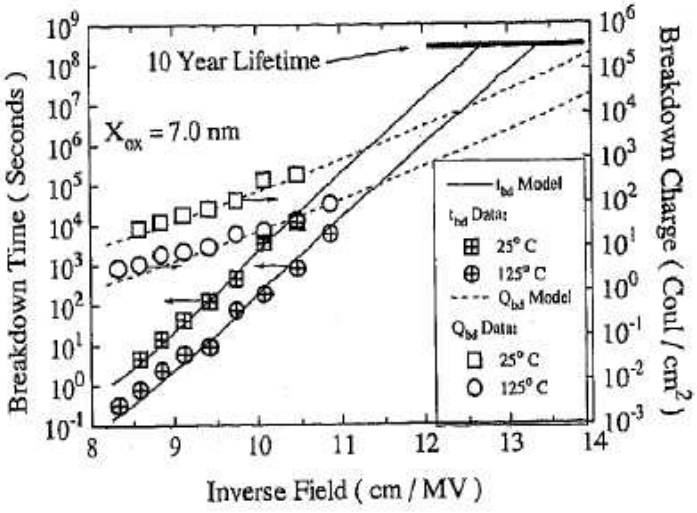

Figure 1.33: TDDB data plotted accordingly to the $1/E$ model after Schuegraf *et al.* [48].

Consensual model, E-model versus $1/E$-model.
With the decreasing oxide thickness, there was a big controversy whether the modeling of the TDDB for lifetime extrapolation should be based on the E-model or the $1/E$-model. As lot of groups have provided consequent results to account for both models [109, 54, 110, 111, 64, 101, 112], consensual models have arisen. A unified model has been proposed by C. Hu [113] to link the E and $1/E$ model based on the assumption that the dielectric breakdown is caused by competitive mechanisms. At high electric fields, the $1/E$ model is dominant while the E-model operates at low fields.

This assumption was also used by McPherson in the thermo-chemical model [102]. However, it seems that this split of the TDDB model dependence on the field strength is too much simplified. Pompl *et al.* have reported a complex TDDB modeling which validates the theory of the competition of different mechanism leading to the breakdown [114]. Using the weakest mechanism, either with the E-model or the $1/E$-model, the data fit the experiment perfectly. It has also been reported a strong impact of the backend of line on the TDDB behavior of the dielectric. The annealing step prior to passivation could change the TDDB behavior from a E-model to a $1/E$-model (Figure 1.34).

Thin oxides (<7 nm)

In case of ultra-thin gate oxides, the TDDB has been found to be independent of the electric field. A clear statement is exposed with the experiment of Nicollian *et al.* performed on 2.6 nm oxides [115]. According to the poly silicon doping, the oxide field could be modulated under a fixed gate voltage stress. In their experiment, the time-to-fail was observed to be constant at fixed voltage stress and totally independent of the poly depletion given by the different poly doping. In contrast, the large variation in the time-to-fail predicted by the E-model was not

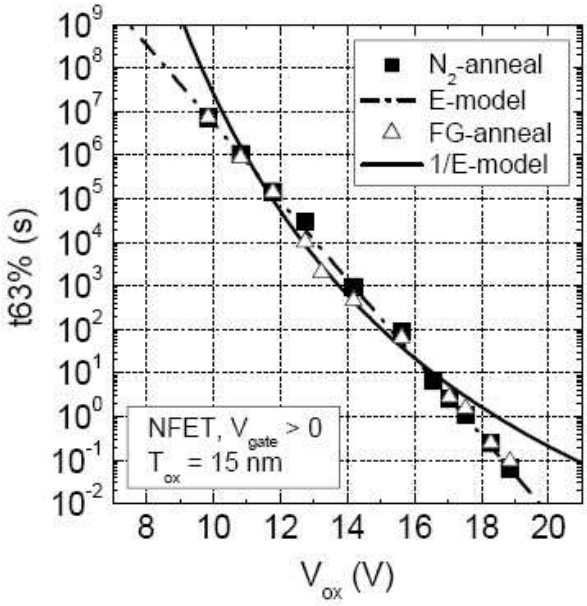

Figure 1.34: A competition of the weakest breakdown mechanism leading to the TDDB E-model or $1/E$-model as a function of the backend of line on process after Pompl *et al.* [114].

seen. It is the evidence that the voltage and not the field is the correct parameter to describe the degradation of ultra-thin oxides. Furthermore this gate oxide time to breakdown dependence on the voltage was exposed by Wu [116] with a very large statistical data base. Through the extensive experimental time to breakdown experiment performed on oxide in the range of 5 nm to 1 nm an empirical power law model was proposed.

$$t = t_0 \left(\frac{V}{V_0} \right)^{-n} \tag{1.67}$$

In the same time, degradation mechanism of ultra-thin gate oxides have been reported to be link with the gate voltage or the electron energy released at the anode [51, 73, 74, 115, 53, 117]. The initial root cause of the voltage driven breakdown was thought to be either the AHI or the AHR. However, coming along with the general acceptance of the power law [118, 119, 120, 75, 121, 122, 123, 121, 123, 35, 124], the hydrogen has been more and more suggested to play an important role in the oxide breakdown and especially towards low voltages [125, 114, 126, 94, 67, 127, 65]. TDDB data described by the power law model are exposed for a 2.2 nm oxide over 8 decades in times after [120] in Figure 1.35.

1.2.6.3 Dielectric lifetime prediction

The TDDB models coupled with the statistical models constitute the base of the dielectric reliability for the lifetime projection. Data collected under conditions exceeding the maximum operation ratings, performed on wafer level (stress range in the ms to hours) and on packaged

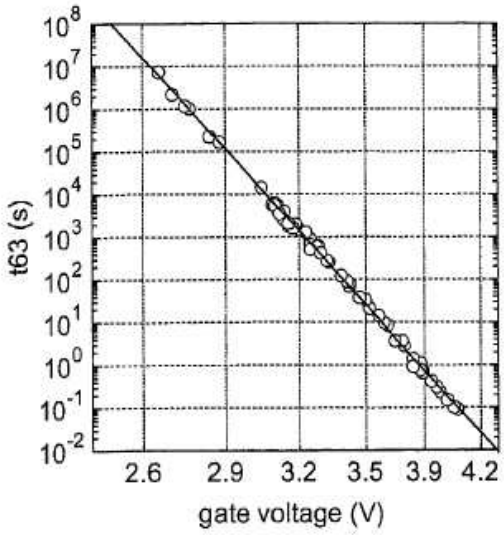

Figure 1.35: TDDB data plotted accordingly to the power law model after Pompl *et al.* [120].

device level for long term stress (from hours to several months) are extrapolated accordingly to the models to the operating conditions (Temperature, voltage, oxide area of the IC). The products could be qualified for a cumulative statistical failure rate criterion defined for the product lifetime.

The reliability focus is the extrapolation of the accelerated stress towards low voltages (supply conditions) and long times (typically for 10 years). The point of interest for ESD developers is of course going into opposite direction, here high voltages and very fast times (nanoseconds range) are the focus. A possible extension of the reliability methodologies towards the ESD time domain would be very beneficial for the ESD engineers. Testing in the nanoseconds range is an extremely tough task which presents a lot of issues and an extrapolation methodology would permit to gain beyond a useful predicting tool for the development, times and money. From these motivations, the goal of this contribution is to understand the oxide degradation mechanisms towards the ESD domain aiming to predict the breakdown behavior via extrapolation models.

1.3 Conclusion

In the first chapter, an overview of the standard dielectric reliability state of the art in CMOS technologies has been presented. However, this reliability topic is also strongly involved in the case where ICs are stressed by electrostatic discharges events. The kind of stresses generated under electrostatic discharges (ESD) differs from the low electrical stresses resulting from the normal operation of ICs. In the next chapter, general overview on ESD phenomena will be discussed to introduce the ESD matters related to thin oxides in the deep sub-µm CMOS technologies.

Bibliography

[1] Green M.L, Gusev E.P., Degraeve R., and Garfunkel E.L. Ultrathin ($<$ 4 nm) Si/SiO2 and Si– O–N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits. *Journal of Applied Physics*, 90(5), 2001.

[2] McPherson J., Kim J.C., Shanware A., and Mogul H. Thermochemical description of dielectric breakdown in high dielectric constant materials. *Applied Physics Letters*, 82(13):2121–2123, 2003.

[3] Mozzi R.L. and Warren B. E. The structure of vitreous silica. *Journal of Applied Crystallography*, 2:164–172, 1969.

[4] Wanabee T. and Ohdomari I. The physics and chemistry of the Si/SiO2 interface. *ECS proceedings*, page 319, 2000.

[5] Helms C. and Pointdexter E.H. The Si/SiO2 system: its microstructure and imperfection. *Rep. Prog. Phys*, 57:791–852, 1994.

[6] Edwards A.H and Fowler W.B. Recent advances in the theory of oxide semiconductor interfaces. *Microelectronics Reliability*, 39:3–14, 1999.

[7] Fischetti M.V., Gastaldi R., Maggioni F., and Modelli A. Slow and fast states induced by hot electrons at Si–Si/SiO2 interface. *Journal of Applied Physics*, 53(4):3136–3144, 1982.

[8] Fleetwood D.M., Winokur P.S., Reber R.A., Meisenheimer T.L., Schwank J.R. Shaneyfelt M.R., and Riewe L.C. Effects of oxide traps, interface traps, and "border traps" on metal-oxide-semiconductor devices. *Journal of Applied Physics*, 73(10):5058–5074, 1993.

[9] Sze S.M. *Physics of Semiconductor Devices*. New York: Wiley, 2nd ed. edition, 1981.

[10] Quarles T.L. *Analysis of Performance and Convergence Issues for Circuit Simulation*. PhD thesis, EECS Department, University of California, Berkeley, 1989.

[11] Hoffman Kurt. *System integration from transistor design to large scale integrated circuits*. John Wiley & Sons Ltd, 2004.

[12] Pompl Thomas. *Gateisolatoren für MOS-Feldeffekttransistoren*. Institut für physik, Universität der Bundeswehr München, 2000.

[13] Hesto P. Nature of electronics conduction. *Instabilities in Silicon Devices*, 1:263, 1986.

[14] Fowler R.H. and Nordheim L. Electron in intense electric field. In *Proc. Soc. London Ser. A*, volume 119, pages 173–181, 1928.

[15] Schuegraf K.F., King C.C., and Hu C. Ultra-thin silicon dioxide leakage current and scaling limit. *VLSI Technology Digest of Technical Papers*, pages 18–19, 1992.

[16] Clerc Raphaël. *Etude des effets quantiques dans les composants CMOS à oxydes de grille ultra minces - Modelisation et caracterisation*. Optique, optoélectronique et microondes, option composants, Institut National Polytechnique de Grenoble, 2001.

[17] Wolters D.R. and Zegers van Duijnhoven T.A. Tunneling in thin Si/SiO2. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 354(1717):2327–2350, 1996.

[18] Schuegraf K.F., King C.C., and Hu C. Impact of polysilicon depletion in thin oxide MOS technology. *International Symposium on VLSI Technology, systems, and applications*, pages 86–88, 1993.

[19] Frenkel J. On pre-breakdown phenomena in insulators and electronic semi-conductors. *Physical Review*, 54(8):647–648, 1938.

[20] DiMaria D.J, Buchanan, Sthatis J.H, and Stahlbush R.E. Interface states induced by the presence of trapped holes near the silicon-silicon-dioxide interface. *Journal of Applied Physics*, 77(5):2032–2040, 1995.

[21] Riess Philip. *Etude de la fiabilité des oxydes minces: analyse des mécanismes de transport et de génération du SILC*. Physique des composants a semiconducteurs, Institut National Polytechnique de Grenoble, 1999.

[22] Goguenheim Didier. *Contribution à l'étude de la fiabilité des oxydes minces dans les structures MOS*. Microélectronique, Université de Provence Aix-Marseille I, 2006.

[23] Nicollian E., Hunter W.R, and Hu J.C. Low voltage stress-induced-leakage-current in ultrathin gate oxides. In *International Reliability Physics Symposium*, volume 37, pages 400–404, 1999.

[24] Petit E., Meinertzhagen A., Zander D., Simonetti O., Fadlallah M., and Maurel T. Low voltage SILC and p- and n-MOSFET gate oxide reliability. In *International Reliability Physics Symposium*, volume 45, pages 479–485, 2005.

[25] Petit Christian. *Contribution à l'étude de la dégradation de films minces et ultra-minces de SiO2 de strucures MOS soumises à des contraintes électriques et à la caractérisation par spectroscopie tunnel inélastique de junction Al-SiO2-Si*. Electronique, Université de Reims Champagne-Ardenne, 2004.

[26] Degraeve R., Kaczer B., and Groeseneken G. Degradation and breakdown in thin oxides layers: mechanisms, models and reliability prediction. *Microelectronics Reliability*, 39:1445–1460, 1999.

[27] Suehle J.S., Zhu B., Chen Y., and Bernstein J.B. Acceleration factors and mechanistic study of progressive breakdown in small area ultra-thin gate oxide. *International Reliability Physics Symposium*, pages 95–101, 2004.

[28] Miranda E., Suñé J., Rodriguez R., Nafria M., and Aymerich X. Detection and fitting of the soft breakdown failure mode in MOS structures. *Solid-State Electronics*, 43:1801–1805, 1999.

[29] Miranda E. and Suñe J. Gate oxide reliability for nano-scale CMOS. *Microelectronics Reliability*, 44(1), 2004.

[30] Pompl T., Wurzer H., Kerber M., and Eisele I. Investigation of ultra-thin gate oxide reliability behavior by separate characterization of soft breakdown and hard breakdown. In *International Reliablility Physics Symposium*, 2000.

[31] Weir B.E., Silverman P.J., Alers G.B., Monroe D., Alam M.A., Sorsch T.W., Green M.L, Timp G.L, Ma Y., Frei M. Bude J.D., and Kirsch K.S. Soft breakdown in ultra-thin oxides. *Materials Research Society*, 567:301–306, 1999.

[32] Cester A., Paccagnella A., Bandiera L., Ghidini G., and Bloom I. Soft breakdown current noise in ultra-thin gate oxides. *Solid-State Electronics*, 46(7):1019–1025, 2002.

[33] Reiner J.C. Pre-breakdown leakage current fluctuations of thin gate oxide. *Microelectronics Reliability*, 43:1507–1512, 2003.

[34] Monsieur F. and. *International Reliability Physics Symposium*, page 45, 2002.

[35] Röhner M., Kerber A., and Kerber M. Voltage acceleration of TBD and its correlation to post breakdown conductivity of n- and p- channel MOSFETs. *International Reliability Physics Symposium*, pages 76–85, 2006.

[36] Stathis J.H. Gate oxide reliability for nano-scale CMOS. *International Conference on Microelectronics*, pages 83–88, 2006.

[37] Suñé J., Miranda E., Nafria M., and Aymerich X. Point contact conduction at the oxide breakdown of MOS devices. *International Electron Device Meeting*, pages 191–194, 1998.

[38] Houssa M., Nigam T., Mertens P.W, and Heyns M. Model for the current-voltage characteristics of ultrathin gate oxides after soft breakdown. *Journal of Applied Physics*, 84(8):4351–4355, 1998.

[39] Martin A., O'Sullivan P., and Mathewson A. Dielectric reliability measurement methods: a review. *Microelectronics Reliability*, 38(1):37–72, 1998.

[40] Nigam T., Degraeve R., Groeseneken G., and Heyns M.M. andd Maes H.E. Constant current charge-to-breakdown: still a valid tool to study the reliability of MOS structures? *International Reliability Physics Symposium*, pages 62–69, 1998.

[41] Esmark Kai. *Device Simulation of ESD Protection Elements*. PhD thesis, 2002.

[42] DiMaria D.J, Cartier E., and Arnold D. Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon. *Journal of Applied Physics*, 73(7):3367–3384, 1993.

[43] Sah C.T., Neugroschel A., and Han K.M. Profiling interface traps in MOS transistors by DC current-voltage method. *Electron Devices Letters*, 17(2):72–74, 1996.

[44] Heremans P., Witters J., Groeseneken G., and Maes H.E. Analysis of charge pumping technique and its application for the evaluation of MOSFET degradation. *Transactions on Electron Devices*, 36(7):1318–1335, 1989.

[45] Groeseneken G. and Maes H.E. Basics and applications of charge pumping in submicron MOSFET's. *International Conference on Microelectronics*, 2:581–589, 1997.

[46] Teh G.L. and Chim W.K. Per-breakdown charge trapping in ESD stressed thin MOS gate oxides. In *International Symposium on Physical and Failure Analysis*, pages 156–161, 1997.

[47] Lombardo S., Stathis J.H., Linder B.L, Pey L., Palumbo F., and Hang Tung C. Dielectric breakdown mechanisms in gate oxides. *Journal of applied physics*, 2005.

[48] Schuegraf K.F. and Hu C. Effects of temperature and defects on breakdown lifetime of thin Si/SiO2 at very low voltages. *Transactions on Electron Devices*, 41(7):1227–1232, 1994.

[49] Weinberg Z.A. and Fischetti M.V. Investigation of the Si/SiO2 - induced substrate silicon field effect transistors. *Journal of Applied Physics*, 57(2):443, 1985.

[50] Schuegraf K.F. and Hu C. Metal-oxide-semiconductor field-effect-transistor substrate current during fowler-nordheim tunneling stress and silicon dioxide reliability. *Journal of applied physics*, 76(6):3695–3700, 1994.

[51] DiMaria D.J. Dependance on gate work function of oxide charging, defect generation, and hole currents in metal-oxide-semiconductor structures. *Journal of Applied Physics*, 81:3320–3326, 1996.

[52] Kimura M. and Ohmi T. Time-dependent dielectric degradation (TDDB) influenced by ultrathin film oxidation process. In *Journal of applied physics*, volume 35, pages 1478–1483, 1996.

[53] Bude J.D., Weir B.E., and Silverman P.J. Explanation of stress induced damage in thin oxides. In *International Electron Device Meeting*, pages 179–182, 1998.

[54] Alam M.A., Bude J., and Ghetti A. Field acceleration for oxide breakdown - can an accurate anode hole injection model resolve the e vs. 1/e controversy? In *International Reliability Physics Symposium*, pages 21–26, 2000.

[55] DiMaria D.J. Trap creation in silicon dioxide produced by hot electrons. *Journal of Applied Physics*, 65:2342, 1985.

[56] DiMaria D.J. Explanation for the polarity dependence of breakdown in ultra-thin silicon dioxide films. *Applied Physics Letters*, 68:3004–3006, 1996.

[57] Walkup R.E., Newns D.M., and Avouris Ph. Vibrational heating and atom transfer with the STM. *Journal of Electron Spectroscopy and Related Phenomena*, 64/65:523–532, 1993.

[58] Walkup R.E., Newns D.M., and Avouris Ph. Role of multiple inelastic transitions in atom transfer with the scanning tunneling microscope. *Physical Review B*, 48:1858–1861, 1993.

[59] Shen T.-C., Wang C., Abeln G.C., Tucker J.R., Lyding J.W., Avouris Ph., and Walkup R.E. Atomic-scale desorption through electronic and vibrational excitation mechanisms. *Science*, 268:1590–1592, 1995.

[60] Avouris Ph., Walkup R.E., Rossi A.R., Akpati H.C., Nordlander P., Shen T.-C., Abeln G.C., and Lyding J.W. Breaking individual chemical bonds via STM-induced excitations. *Surface Science*, 363:368–377, 1996.

[61] Stripe B.C., Rezaei M.A., Ho W., Gao S., Persson M., and Lundqvist B.I. Single-molecule dissociation by tunneling electrons. *Physical Review Letters*, 78(23):4410–4413, 1997.

[62] Stokbro K., Thirstrup C., Sakurai M., Quaade U., Hu B.Y-K., Perez-Murano F., and Grey F. STM-induced hydrogen desorption via hole resonance. *Physical Review Letters*, 80(12):2618–2621, 1998.

[63] Ferry D.K., Goodnick S.M., and Hess K. Energy exchange in single-particle electron-electron scattering. *Physica B*, 272:538–541, 1999.

[64] Chen F., Vollertsen R.-P., Li B., Harmon D., and Lai W.L. A new empirical extrapolation method for time-dependent dielectric breakdown reliability projections of thin Si/SiO2 and nitride-oxide dielectrics. *Microelectronics Reliability*, pages 335–341, 2002.

[65] Nicollian P.E, Krishnan T., Chancellor C.A, Khamankar R.B., Chakravarthi S., Bowen C., and Reddy V. The current understanding of trap generation mechanisms that leads to the power law model for gate dielectric breakdown. In *International Electron Device Meeting*, pages 197–206, 2007.

[66] Suñé J. and Wu E.Y. Hydrogen-release mechanisms in the breakdown of thin Si/SiO2 films. *Thin SiO2 Films*, 92(8):087601–1–4, 2004.

[67] Ribes G., Bruyère S., Denais M., Monsieur F., Huard V., and Roy D. Ghibaudo G. Multi-vibrational hydrogen release: Physical origin of tbd, qbd power-law voltage dependence of oxide breakdown in ultra-thin gate oxides. *Microelectronics Reliability*, 45:1842–1854, 2005.

[68] Suñé J. and Wu E.Y. Mechanisms of hydrogen release in the breakdown of SiO2-based gate oxides. *International Electron Device Meeting*, pages 399–402, 2005.

[69] DiMaria D.J and Cartier E. Mechanism for stress-induced leakage currents in thin silicon dioxide films. *Insulating films on semiconductors*, 78:3883–3894, 1995.

[70] Nigam T., Degraeve R., Groeseneken G., and Heyns M.M. andd Maes H.E. A fast and simple methodology for lifetime prediction of ultra-thin oxides. *International Reliability Physics Symposium*, pages 381–388, 1999.

[71] Ghetti A., Bude J., and Weber G. Tbd prediction from measurements at low field and room temperature using a new estimator. *Symposium on VLSI Technology Digest of Technical Papers*, pages 217–218, 200.

[72] Wu E.Y., Sune J., Nowak E., Lai W., and McKenna J. Weibull slopes, critical defect density, and the validity of stress-induced-leakage current (SILC) measurements. *International Electron Device Meeting*, pages 125–128, 2001.

[73] Stathis J.H. and DiMaria D.J. Reliability projection for ultra-thin oxides at low voltage. *International Electron Device Meeting*, pages 167–170, 1998.

[74] DiMaria D.J. Electron energy dependence of metal-oxide-semiconductor degradation. *Applied Physics Letters*, 75(6):2427–2428, 1999.

[75] Wu Ernest Y. and Suñé Jordi. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45:1809–1834, 2005.

[76] Degraeve R., Groeseneken G., Bellens R., Ogier J.L., Depas M., Roussel P., and Maes H.E. New insights in the relation between electron trap generation and the statistical properties of oxide breakdown. *Transaction on Electron Devices*, 45(4):904–911, 1998.

[77] Stathis J. H. Percolation models for gate oxide breakdown. *Journal of applied physics*, 86(10):5757–5765, 1999.

[78] Wolters D.R. Breakdown and wearout phenomena in Si/SiO2. *Insulating films on semiconductors*, pages 180–194, 1981.

[79] Krishnan A.T. and Nicollian P.E. Analytic extension of the cell-based oxide breakdown model to full percolation and its implications. *Annual International Reliability Physics Symposium*, pages 232–239, 2007.

[80] Wolters D.R. and Verwey J.F. Breakdown and wearout phenomena in Si/SiO2 films. *Instabilities in Silicon Devices*, pages 315–362, 1986.

[81] Degraeve R., Groeseneken G., Bellens R., Depas M., and Maes H.E. A consitent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides. *Annual International Reliability Physics Symposium*, pages 863–866, 1995.

[82] Wu E.Y., Suñé J., and Lai W. on the weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination - part II: Experimental results and the effects of stress conditions. *Transaction of Electron Devices*, 49(12):2141–2150, 2002.

[83] Wu E.Y. and Vollertsen R.P. on the weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination - part I: Theory, methodology, experimental techniques. *Transaction of Electron Devices*, 49(12):2131–2140, 2002.

[84] Weir B.E., Alam M.A., Silverman P.J., Baumann F., Bude J.D., Timp G.L., Hamad A., Ma Y., Brown M.M., Hwang D., Sorsch T.W., gheti A., and Wilk G.D. Ultra-thin gate oxide reliability projections. *Solid-State Electronics*, 46:312–328, 2002.

[85] Sichart K.V. and Vollertsen R.-P. Bimodal lifetime distributions of dielectrics for integrated circuits. *Quality and Reliability Engineering International*, 7:299–305, 1991.

[86] Wu E.Y., Suñé J., Lai W., Nowak E., McKenna J., Vayshenker A., and Harmon S. Interplay of voltage and temperature acceleration of oxide breakdown for ultra-thin gate oxides. *Solid-State Electronics*, 46, 2002.

[87] Wu E.Y., Abadeer W.W, Han L., Lo S., and Hueckel G.R. Challenges for accurate reliability projections in the ultra-thin oxide regime. *International Reliability Physics Symposium*, pages 57–65, 1999.

[88] Suñé J., Placencia I., Barniol N., Farrés E., Martin F., and Aymerich X. On the breakdown statistics of very thin Si/SiO2 films. *Thin SiO2 Films*, 185:347–362, 1990.

[89] Subramoniam R., Scott R.S., and Dumin D.J. High field related thin oxide wearout and breakdown. *International Electron Device Meeting*, pages 135–138, 1992.

[90] Dumin D.J., Mopuri S.K., Vanchinathan S., Scott R.S., Subramoniam R., and Lewis T.G. High field related thin oxide wearout and breakdown. *Transactions on Electron Devices*, 42(4):760–772, 1995.

[91] Dumin D.J. A special issue on oxide reliability: a summary of silicon oxide wearout, breakdown, and reliability. *International Journal of High Speed Electronics and Systems*, 11(3):619–719, 2001.

[92] Suñe J., Jimenez D., and Miranda E. Breakdown modes and breakdown statistics of ultrathin Si/SiO2 gate oxides. *International Journal of High Speed Electronics and Systems*, 11(3):789–849, 2001.

[93] Suñé J. and Wu E.Y. A quantitative two-step hydrogen model of Si/SiO2 gate oxide breakdown. *Solid-State Electronics*, 46:1825–1837, 2002.

[94] Nicollian P.E, Krishnan, Chancellor C.A, and Khamankar R.B. The traps that cause breakdown in deeply scaled SiOn dielectrics. In *International Electron Device Meeting*, 2006.

[95] Crook D.L. Method of determing reliability screens for time dependent dielectric breakdown. In *International Reliability Physics Symposium*, pages 1–7, 1979.

[96] Meyer W.K. and Crook D.L. Model for oxide wearout due to charge trapping. In *International Reliability Physics Symposium*, pages 242–247, 1983.

[97] Anolick E.S. Low filed time dependent dielectric integrity. In *International Reliability Physics Symposium*, page 8, 1979.

[98] McPherson J.W. and Baglee D.A. Acceleration factors for thin gate oxide stressing. *International Journal of High Speed Electronics and Systems*, pages 1–5, 1985.

[99] Kimura M. Oxide breakdown mechanism and quantum physical chemistry for time-dependent dielectric breakdown. In *International Reliability Physics Symposium*, pages 190–200, 1997.

[100] Kimura M. Field and temperature acceleration model for time-dependent dielectric breakdown. *Transaction on Electron Devices*, pages 220–229, 1999.

[101] Mizubayashi W., Yoshida Y., Miyazaki S., and Hirose M. Quantitative analysis of oxide voltage and field dependence of time-dependent dielectric soft breakdown and hard breakdown in ultrathin gate oxides. *Japanese Journal of Applied Physics*, pages 2426–2430, 2002.

[102] McPherson J.W. and Khamankar R.B. Molecular model for intrinsic time-dependent dielectric breakdown in Si/SiO2 dielectrics and the reliability implications for hyper-thin gate oxide. *Semiconductor Science and Technology*, 15:462–470, 2000.

[103] McPherson J.W. Extended mie-grüneisen molecular model for time dependent dielectric breakdown in silica detailing the critical roles of O–Si O3 tetragonal bonding stretched bonds, hole capture, and hydrogen release. *Journal of Applied Physics*, 99:083501–1–13, 2006.

[104] Chen I.C., Holland S., and Hu C. A quantitative physical model for time-dependent breakdown in Si/SiO2. In *International Reliability Physics Symposium*, pages 24–31, 1985.

[105] Lee J., Chen I.H, and Hu C. Statistical modeling of silicon dioxide reliability. In *International Reliability Physics Symposium*, pages 131–138, 1988.

[106] Lee J., Chen I.H, and Hu C. Modeling and characterization of gate oxide reliability. *Transaction on Electron Devices*, pages 2268–2278, 1988.

[107] DiMaria D.J and Stathis J. Anode hole injection, defect generation, and breakdown in ultrathin silicon dioxide films. *Journal of Applied Physics*, 89(9):5015–5024, 2001.

[108] Schuegraf K.F. and Hu C. Hole injection oxide breakdown model for very low voltage lifetime extrapolation. In *International Reliability Physics Symposium*, pages 7–12, 1993.

[109] Teramoto A., Umeda H., Azamawari K., Kobayashi K., Komori J., Ohno Y., and Miyoshi H. Study of oxide breakdown under very low electric field. In *International Reliability Physics Symposium*, pages 66–71, 1999.

[110] McPherson J.W. Physics and chemistry of intrinsic time-dependent dielectric breakdown in Si/SiO2 dielectric. *International Journal of High Speed Electronics and Systems*, 11(3):719–751, 2001.

[111] Takayanagi M., Takagi S.I., and Toyoshima Y. Experimental study of gate voltage scaling for TDDB under direct tunneling regime. In *International Reliability Physics Symposium*, pages 380–385, 2001.

[112] Vollertsen R.-P., Nierle K., Wu E.Y., and Wen S. Product level verification of gate oxide reliability projections using DRAM chips. *International Reliability Physics Symposium*, pages 385–390, 2003.

[113] Hu C. and Lu Q. A unified gate oxide reliability model. In *International Reliability Physics Symposium*, pages 47–51, 1999.

[114] Pompl T., Allers K.-H., Schwab R., Hofmann K., and Röhner M. Change of acceleration behavior of time-dependent dielectric breakdown by the BEOL process: Indications for hydrogen induced transition in dominant degradation mechanism. In *International Reliability Physics Symposium*, pages 388–397, 2005.

[115] Nicollain P.E., Hunter W.R., and Hu J.C. Experimental evidence for voltage driven breakdown models in ultrathin gate oxides. In *International Reliability Physics Symposium*, pages 7–15, 2000.

[116] Wu E.Y., Nowak E., Vayshenker A. Varekamp P., Hueckel G., McKenna J., Harmon D., Han L-K, Montrose C., and Dufresne R. Voltage-dependent voltage-acceleration of oxide breakdown for ultra-thin oxides. In *International Electron Device Meeting*, page 541, 2000.

[117] McKenna J., Wu E., and Lo S.-H. Tunneling current characteristics and oxide breakdown in p-poly gate PFET capacitors. In *International Reliability Physics Symposium*, pages 16–20, 2000.

[118] Duschl R. and Vollertsen R-P. Voltage acceleration of oxide breakdown in the sub-10 nm fowler-nordheim and direct tunneling regime. In *International Integrated Reliability Workshop Final Report*, volume 45, pages 44–48, 2005.

[119] Duschl R. and Vollertsen R-P. Is the power-law model applicable beyond the direct tunneling regime? *Microelectronics Reliability*, 45:1861–1867, 2005.

[120] Pompl T. and Röhner M. Voltage acceleration of time-dependent breakdown of ultra-thin gate dielectrics. *Microelectronics Reliability*, 45:1835–1841, 2005.

[121] Lin C.-L. andKao T., Chen J.-P., Yang J.Y.C, and Su K.C. Ultra-thin gate oxide lifetime projection and degradation mechanism beyond 90 nm CMOS technology. In *International Integrated Reliability Workshop Final Report*, pages 186–189, 2006.

[122] Pic D., Goguenheim D., and Ogier J.-L. Long range statistical lifetime prediction of ultra-thin Si/SiO2 oxides: Influence of accelerated ageing methods and extrapolation models. *International Conference on Microelectronics*, 2006.

[123] Ohgata K., Ogasawa M., Shiga K., Tsujikawa S., Murakami E., Kato H., Umeda H., and Kubota K. Universality of power-law voltage dependence for TDDB lifetime in thin gate oxide PMOSFETs. In *International Reliability Physics Symposium*, pages 372–376, 2005.

[124] Kerber A., Röhner M., Wallace C., O'Riain L., and Kerber M. Wafer-level gate-oxide reliability towards ESD failures in advanced CMOS technologies. *Transaction on Electron Devices*, 53(4):917–920, 2006.

[125] Vollertsen R.-P. and Wu E.Y. Gate oxide reliability parameters in the range 1.6 to 1 nm. *Integrated Reliability Workshop Final Report*, pages 10 – 15, 2003.

[126] Nicollian P.E, Krishnan, Bowen C. T., Chakravarthi S., Chancellor C.A, and Khamankar R.B. The roles of hydrogen and holes in trap generation and breakdown in ultra-thin SiON dielectrics. In *International Electron Device Meeting*, pages 403–406, 2005.

[127] Haggag A., Liu N., Menke D., and Moosa M. Physical model for the power-law voltage and current acceleration of TDDB. *Microelectronics Reliability*, 45:1855–1860, 2005.

Chapter 2

Electrostatic discharge (ESD)

2.1 Electrostatic Discharge phenomena

2.1.1 What is an ESD event?

Electrostatic discharge (ESD) is a common physical phenomenon that anybody could be confronted to within daily life. Everyone knows this nasty, but more or less harmful, electrical shock when touching a metallic door knob after walking on a carpet, or touching the door of his car while stepping out. This unpleasant experience is caused by an electrostatic discharge. It is rather nonhazardous for human beings. But it has to be considered very carefully for the semiconductor industry as electrostatic discharges are drastically critical for integrated circuits (ICs). Since more than three decades it has always been a major reliability topic [1] and its consideration is strongly growing with the continuous down-scaling of technologies.

An ESD is an electric current, resulting from a re-balancing of static charges between two objects. This phenomenon occurs when the dielectric environment, separating the two surface in an unbalanced electrical condition, is breaking down due to an exceeded field strength. There are different ways to generate charged objects, most of them are due to triboelectricity, getting into contact with a charged object or by induction [2].

In general, the discharge event is fast, typically discharge times are in the order of some hundreds of picoseconds to some micro seconds. Although its duration is that fast, its current pulse level, which is in the range of several Amps, is largely high enough to damaged microscopic structures in ICs irreversibly. The power dissipated during the discharge can reach several kilowatts with an energy density of up to 10 J/cm^3 [3]. For an more understandable picture in a macroscopic scale, a discharge leads to an energy density as high as when a lightening strike hits a big tree [4]. For electronic components the heat dissipation could even rise above the meting point of the semiconductor material and, thus, leads to local melting sites and hence irreversible damage of the structure.

2.1.2 ESD and ICs industry

Despite ESD is not a new problem for IC industries and a lot of effort have been done in the last 20 years to reduce its impacts, ESD protection in ULSI technologies is still gaining importance due to increased complexity and sensitivity of the processes. This challenge has to be overcome in order to reduce the costs of ESD on production yield and, even worse, product returns. Losses due to electrical overstress (EOS) and ESD events are evaluated in the range of 20 % to 65 % of all field returns [5, 6, 7, 8, 9, 10, 11]. This number implies an annual cost to semiconductor industries in the range of billions of dollars. Although it is hard to distinguish between ESD and EOS fails, the portion of ESD failures has recently been evaluated to at least around 14 % of returns (Figure 2.1). From this numbers, the increasing demand of industries for ESD solutions is obvious, as the loss per company is estimated to several millions of dollars per year. Beyond this economical reason, a hot reliability topic is concerned for sensitive applications as in the safety, automotive or aerospace domains. In case where an ESD event would occur to an insufficient protected IC, possible mal functions of electronic systems could result in a disaster. Just as an example, one could easily imagine what could happen if an airbag is suddenly blowing up in the car while driving on the highway. As the portion of electronic functionalities in all products in our daily life is increasing steadily, reliability issues are brought to the primary interest of manufacturers and consumers.

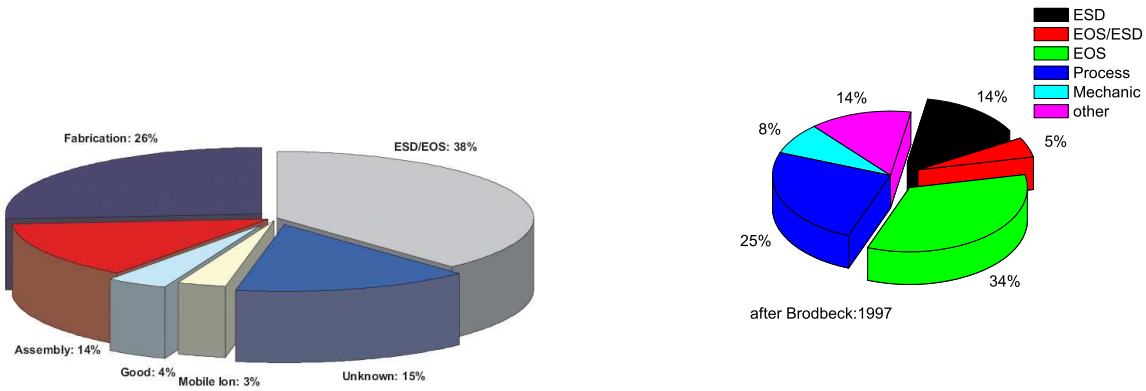

Figure 2.1: Distribution of the causes for field returns after Wagner [9] (left) and Brodbeck [11] (right).

During the life of an IC, the risk to suffer from an ESD event is permanent from its fabrication until its assumed operational end of life. In production lines, even if some specific precautions are taken to avoid electrostatic discharges, the possibility to have an ESD event generated from the processing lines and tools environment is considerably high. As an example, a supplied processing tool which is not well isolated or grounded can lead to an unintentional IC charging during any fabrication step. This would generate a discharge later on, as soon as the IC will come into contact with the ground potential [2]. The ESD occurrence risk is at its maximum when the IC is packaged, before it is assembled in a system [12]. At each step where ICs

are manipulated or even handled manually or in a automated manner, there is a potential ESD risk. From processing dies, packaging, test, assembly on boards to final operation conditions, ICs could be damaged by an ESD event. On the last phase, the ESD generation risk level is obviously dependent on applications and users. An EOS due to a non-proper use of the IC is also increased and the ESD event should be considered here at system level. The IC cannot be considered anymore as stand-alone, as the entire system in which it is integrated has a severe impact on the probability and on the stress levels itself of ESD events.

There are two different ways to keep an IC safe from an ESD event.

- The first one consists in controlling the IC environment and preventing the risk of the occurrence of an ESD [2]. This can be done by introducing protected areas in which all conductors in the environment, including personnel, are grounded. This can be done e.g. with wrist straps connected to ground, conducting ESD shoes, static dissipative work surfaces, ESD safe packaging, adapted floors, furniture and clothes. Air ionizers can be used to compensate the static charges and the relative humidity can be also regulated. This prevention measures requires discipline and has to be established at different process steps as for example fabrication lines, assembly lines, test facilities and system assembly lines, in order to exclude the ESD risk [2].

- The second way consists in integrating ESD protection elements on chip which enables the IC to withstand a certain ESD discharge level [13]. Without this special safety on-chip protection elements, the yield of CMOS would be dramatically low.

In order to guarantee the safeness of ICs, a first ESD robustness standard was developed by the U.S. army in 1980 (MIL standard). Nowadays several organizations and industries have defined some ESD robustness standards (ESDA, JEDEC, IEC, ...). Those different standards tempts to reproduce real ESD events and are based on stress models which aim to cover the major cases of possible ESD discharge on ICs. The robustness level from those tests consists in relative values which prevent from ESD damage in the field. These standard models used by semiconductor industries to qualify ESD robustness of ICs in DATA sheets are presented in the next section.

2.2 ESD robustness qualification models

In the field, there might occur a lot of different discharge wave forms, corresponding to the nature of the event. Parameters as air moisture, contact surface, speed of approach and other specific stress configurations are impacting the current pulse wave form and its level [2]. To define reliable testing conditions of ESD robustness is a complex topic. It is difficult to simulate a well defined and controlled “real ESD event”, moreover in the nanosecond regime and at

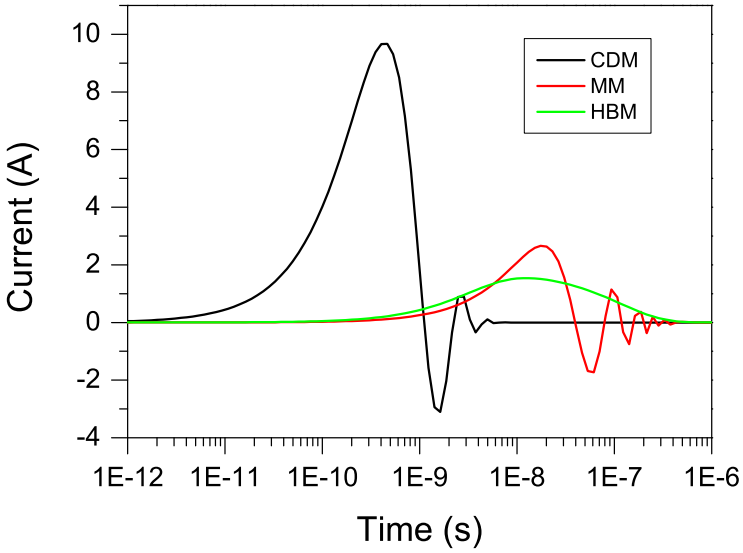

Figure 2.2: Typical ESD discharges that could occur on ICs.

such high current levels. Therefore, standardization of appropriate ESD testing methods is a challenging task which now keeps several international standardization bodies busy for decades. Today, basically three models are established as qualification references at ICs level. These models are thought to reproduce typical "real-life" ESD events which historically are thought to cover the main ESD risks of ICs.

- The Human Body Model (HBM) simulates a charged person stressing one pin of an IC with one or more pins of the IC at ground potential.
- The Machine Model (MM) is based on the same occurrence possibility as the HBM, with the only difference that the stress is generated by a machine (or a metallic tool handled by an operator) getting into contact with one pin.
- The Charge Device Model (CDM) reproduces an ESD event occurring when a charged device comes into (metallic) contact with the ground.

The stress pulse duration of these ESD discharges is shown in Figure 2.2. The ESD events mentioned above are in a typical time range from some hundreds of picoseconds to several hundreds of nanoseconds.

Based on those models, the previous mentioned standards are defining the test procedures and test conditions, including the tester calibration, the number of samples to be tested, pins

combinations to be tested, number of zaps, polarities, etc. ESD stress is performed at increasing stress level until the devices fails. It is important to mention that the pre-stress and post-stress read-out must cover the entire specification of the device under test (DUT). The highest stress level which the devices passes is the so-called ESD withstand voltage. In order to avoid "window fails", the stress step size should be sufficiently small, e.g. for an HBM target robustness of 2.5 kV, stress at 500 V, 1.0 kV, 1.5 kV, 2.0 kV and 2.5 kV should be applied.

2.2.1 Human Body Model

From an electrical point of view, a charged human being which is discharging itself could be described by a charged capacitance storing its accumulated charge, and a resistor representing the resistivity of the body. Theses values are of course dependent on the morphology of the person. For standardization, the typical values of 100 pF and 1.5 kΩ are used to model the electrical properties of an human body.

Historically, these models come from the nineteenth century and has been developed for investigating explosive gas mixtures in mines. During an electrostatic event, the accumulated voltage in the body is going trough this simple "human body model circuit" and gives a high current spike followed by the characteristic exponential discharge (exponential decay 150 ns). The level of the charging voltage is in the range of one to several kilovolts and the maximal current peak value given by $V_{\mathrm{HBM}}/R_{\mathrm{HBM}}$ (equivalent to a current amplitude of 0.67 A/kV). To explain this charging voltage levels, some typical static charged voltage values resulting from different activities [14] are compiled in Table 2.1. The sensitivity threshold for a person to feel an electrostatic discharge is evaluated in the range of 3 kV [12]. However, being charged-up to voltages far below this pain threshold value, ICs can get damaged in the case of an ESD. Internal protection and grounded operators are inevitable to avoid any mysterious fails.

Relative Humidity	10 %	40 %	55 %
Walking across a carpet	35 kV	15 kV	7.5 kV
Walking across a vinyl floor	12 kV	5 kV	3 kV
Motions of bench employee	6 kV	800 V	400 V
Removing dual in-line packages (DIPs) from plastic tubes	2 kV	700 V	400 V
Removing (DIPs) from polystyrene foam	14.5 kV	5 kV	3.5 kV

Table 2.1: Electrostatic voltage generated from human activities [14]

Testers are not ideal and some parasitic elements have to be taken into account to explain the real stress generated by the tester. The effective parasitic inductance from the tester determines the rise time of the waveform, this value should be in the range from 2 ns to 10 ns (evaluated time from 10 % to 90 % of I_{MAX}). The stress rise time has an impact on the ESD robustness of devices [15, 16] and can result in different HBM thresholds for different testers.

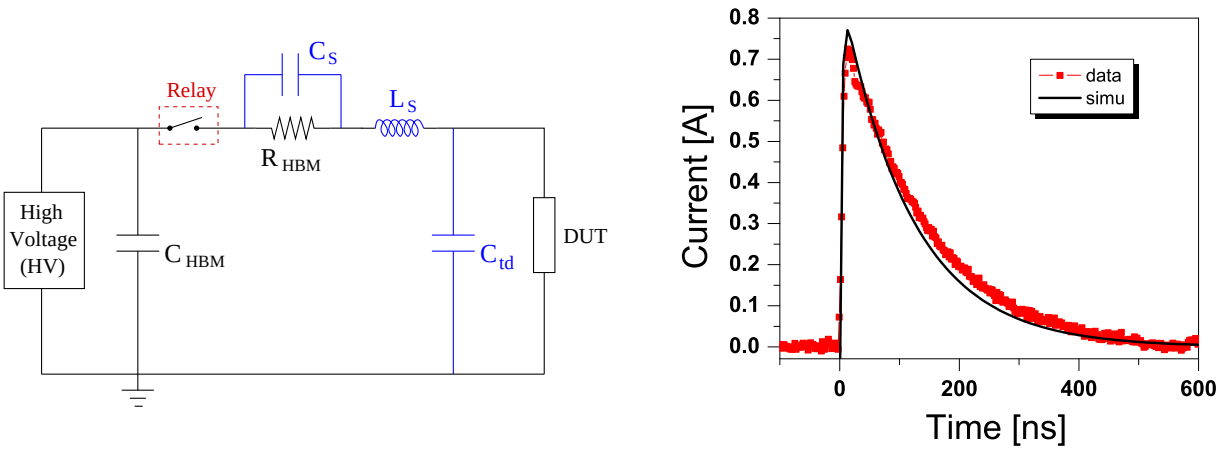

Figure 2.3: HBM equivalent electrical circuit and current waveform for a 1 kVdischarge.

Two other major parasitics elements are the board capacitance and a serial capacitance resulting from the HBM tester [17, 18]. The equivalent HBM testers circuits including the parasitic elements is shown in Figure 2.3. The typical parasitic inductance (L_s) is about 5.000 nH, the serial capacitance (C_s) is in the range of 5 pF and the board capacitance (C_{tb}) is typically lower than 30 pF. The real discharge wave form is described by an fourth order analytical solution. For the knowledge of the exact stress applied, a calibration procedure is required for the extraction of the HBM parasitic testers. A typical discharge waveform of an 1 kV HBM stress generated from the KeyTek ZapMaster tester is shown in Figure 2.3. Neglecting the parasitic capacitances of the tester, the HBM current stress results in a second order *RLC* circuit which is also well describing the event. In this case, the simplified differential equation is:

$$L_s \cdot \frac{d^2 i}{dt} + (R_{HBM} + R_{DUT}) \cdot \frac{di}{dt} + \frac{1}{C_{HBM}} \cdot i = 0 \tag{2.1}$$

and consequently the HBM stress current can be described by

$$I_{HBM}(t) = V_{HBM} C_{HBM} \frac{\omega_0^2}{\sqrt{\alpha^2 - \omega_0^2}} \cdot e^{-\alpha t} \sinh(\sqrt{\alpha^2 - \omega_0^2}\, t) \tag{2.2}$$

$$\text{where } \alpha = R_{HBM}/2L_{HBM} \text{ and } \omega_0 = 1/\sqrt{L_{HBM} C_{HBM}}$$

Even if the probability of the occurence of an ESD from a human being is strongly decreased with the increasing degree of automation of production as well as testing, HBM is the most commonly used model in the microelectronic industry. It is for almost all products mandatory to go through a HBM qualification. It is *the* model worldwide used as comparative reference. The reason is that currently it is the best defined and most reproducible ESD model as a lot of

standardization works have been done for it. The different international standards, i.e., MIL–STD–883x method 3015.7 from 1989 [19], JEDEC JESD22–A114–B [20], ESDA STM5.1–2001 [21], AEC Q100–02 REV D [22], IEC 61340–3–1 [23], EIAJ ED–4701/304 [24] enable a quite repeatable and reproducible evaluation of one ESD robustness for the semiconductor industry.

However, all of these standards include some differences and are not really fully comparable, the values from one to another could differ slightly. The most popular standards which are close to each other, JEDEC (Joint Electron Device Engineering Council) and ESD Association are very similar, at least with respect to wave form definition. The only main difference between the JEDEC and ESDA standard concerns the pin combination testing. In ESDA STM5.1-2001, supply pins should be grounded together if the resistance between them is less than a – more or less arbitrary chosen – value of 2 Ω. In contrast, no value is given in JEDEC JESD22–A114–B. Recently, activities have been initiated by the both standardization bodies ESDA and JEDEC to harmonize the HBM standards and define a "common" HBM standard.

Infineon Technologies started to use the ESDA standard for qualification, however, now since 6 years the JEDEC standard is used in order to harmonize the IC qualification standards in the company. The typical target HBM stress voltage level for products is 2 kV ; this value can be significantly higher for specific applications, e.g. automotive or chip card applications. It means that devices have to resist to an 2 kV stress (accompanied by a peak current of I_{HBM} of around 1,3 A) which is becoming more and more problematic with the down-scaling of technologies. Recently, some discussions between the mains semiconductor industries since the EOS/ESD symposium 2006 are ongoing in order to lower this target to a value of 1 kV for the future CMOS technology nodes (65 nm, 45 nm, ...) which is still acceptable for handling as it is far above all possible HBM risks in EPAs [25]. In fact, for these advanced technologies the protection development is becoming much more critical, complex and expensive and, thus, it is an absolute must to avoid over-design.

HBM test procedure

Prior to qualification runs, calibration of testers should be done. Waveforms have to be recorded for discharges into a short and into a 500 Ω resistance for tester verification. The HBM stress is a two point stress configuration. The testing procedure was recently revised. One zap per stress level and polarity should be applied for the pin stress configurations mention below:

- all pins once at a time versus each power pin group
- all pins once at a time versus non-power pins ganged together

For a fast evaluation whether the device passed or failed the stress, *I-V* characterization can be performed. Before and after each stress level, DC curves or DC spots are recorded to monitor any increase in leakage or power-down current. For single devices leakage current is often the

only failure criterion and, therefore, a full functional read-out can be omitted. The stress level at which the DC characteristic of the DUT violates the specification gives the HBM robustness threshold level for the pin configuration under stress. The weakest pins of the IC is defining the ESD robustness of the chip. Typically, three samples are required for the ESD robustness evaluation to increase the statistic.

Tester artifacts

One may also notice that HBM testers are quite complex machines which can also generate some unintentional artefact stresses. Artefact effects, like HBM stress rise time modification caused by the capacitance of relays especially for high pin counts [26, 27] or the non connecting pins fails due to the board capacitance are complicating the evaluation of the ESD robustness and can lead to irregular values. Trailing pulses reported in [28, 29] for different testers may also cause some suspicious ESD fails. The long duration of the trailing pulse in the range of several hundreds of µs is generating a severe unintentional EOS stress, particularly critical for gate oxides of advanced sub-micron processes if the current of the trailing pulse is too low to re-trigger the protection elements. Basically, this trailing pulse is induced by the leakage current of the plasma relay used in testers. Trailing pulses from a Zap Master HBM tester for different clamping element are shown in Figure 2.4. This stress depends strongly on the HBM tester system architecture and on the relay configuration involved in the stress. The HBM tester systems built before this effect was considered to be critical may suffer from this artefact. Pulse waveforms from different HBM testers available at Infineon (IMCS tester, device tester Verifier, Zap Master and Zap Master II) have been observed. All of them are generating a second EOS stress which differ from each other and is specific to the tester. The trailing pulse depends on

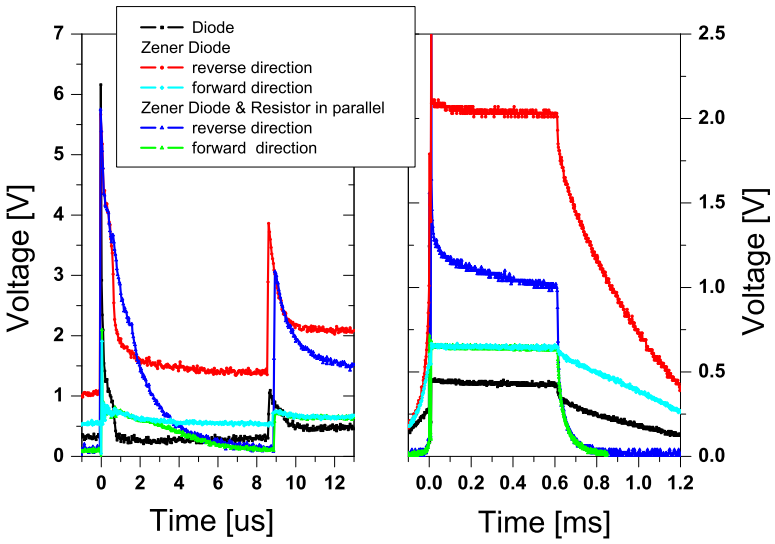

Figure 2.4: Trailing pulses from a Zap Master HBM tester.

the combination of relays used in the stress configuration and on the DUT.

It is important to mentioned that the results from oxide reliability studies done in the past using standard HBM testers may not be correct due to this non negligible artifact stress.

2.2.2 Machine Model

The Machine Model is mainly used for automotive qualification or sometimes for RF products particularly in the Japanese and US market [1, 30], but is rarely asked and used for other types of applications. This model is similar to HBM, only capacitance and resistor values are changed to about 200 pF and 0 Ω. Therefore, it is sometimes called as "worst-case HBM". The very low resistor value is a sever technical problem for the measurement because the results are mainly determined by the test set up and by the device parasitics. For practical testing, the main standards ESDA STM5.2–1999 [31], JEDEC JESD22–A115–A [32], AEC Q100–00 REV D [33], IEC 61340–3–2 [34] define the parasitic resistance to 10 Ω and the lumped inductance from the tester to 750 nH. The total effective parasitic resistance from the tester as well as the DUT resistance should be considered in the damping coefficient α [17]. MM discharge calibration for a 500 Ω is discussed in the standard and is needed to deduce the parasitic impedance of testers.

MM is described by an oscillating discharge waveform at the frequency ω_0 and with a damping coefficient α. A typical MM waveform for a pre-charged level of 200 V is shown in Figure 2.5 and simulated with the following equation:

$$I_{\mathrm{MM(t)}} = V_{\mathrm{MM}} C_{\mathrm{MM}} \frac{\omega_0^2}{\sqrt{\omega_0^2 - \alpha^2}} \cdot \mathrm{e}^{-\alpha \mathrm{t}} \sin((\sqrt{\omega_0^2 - \alpha^2})\mathrm{t}) \tag{2.3}$$

$$\text{where } \alpha = (\mathrm{R_{MM}} + \mathrm{R_{DUT}})/2\mathrm{L_{MM}} \text{ and } \omega_0 = 1/\sqrt{\mathrm{L_{MM}C_{MM}}}$$

The typical target value for MM robustness is 200 V, but the strong dependence of the threshold robustness voltage achieved from testers to testers could not permit good ESD robustness comparisons. However, the current peak given by $\mathrm{I_{MM}} = \mathrm{V_{MM}}\sqrt{\mathrm{C_{MM}/L_{MM}}}$ which results from the MM stress enables better results comparability. As MM addresses the same failure mode as HBM (mainly thermal fails), the peak current is the decisive parameter.

Depending on testers and devices the ESD threshold level is five to twenty times lower than HBM [35, 3]. However, the same failure mechanism as in MM is in general addressed by HBM [1]. Cost attributed for qualification and the lower confidence of the correlation in the MM threshold level due to the important impact of tester and DUT on the discharge waveform lead to a disregard of MM in favor to the Human Body Model. As the HBM standards, discussion on the MM target values are ongoing to reduce the typical 200 V to 30 V [25].

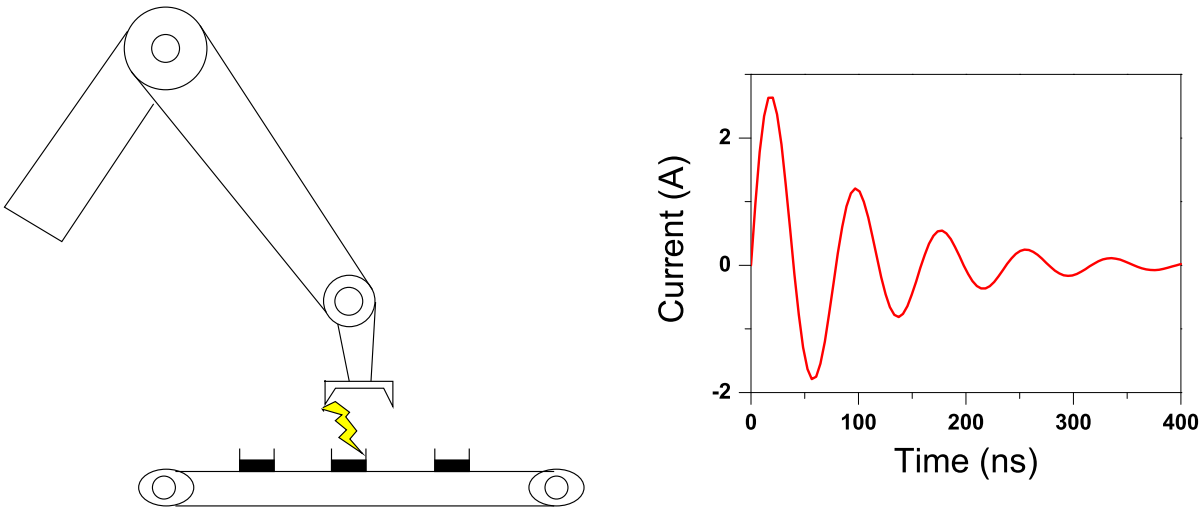

Figure 2.5: MM discharge scheme and waveform for an 200 V charge level.

2.2.3 Charged Device Model and Socket Discharge Model

The CDM is a good simulation of real events occurring during IC fabrication and in assembly and testing. However, this model poses a real challenge with regard to measurement and characterization due to the associated time scales and current levels. The CDM is characterized by a very rapid current pulse of about 1 ns with a very fast rise time from 100 ps to 300 ps, reaching high current levels up to several tenth of Amps ($\sim$ 10 A/kV).

The current discharge has been modeled also with a second order *RLC* circuit (see Chapter 1 from Gieser in [1] and [36]), considering the low ohmic series resistance from the discharge path in the range of 10 Ω, the inductance from the tester and the chip (several pH) and the capacitance of the device (few pF).

An example of an CDM waveform is drawn in Figure 2.6. The CDM event is resulting either from a direct charging or from a field-induced charging. Historically, the first CDM testing was using direct charging and discharging via a relay systems [37]. Nowadays, an alternative testing method based on the field-induced CDM [38] is mainly applied, commercial testers built for high pin-count testing always use this solution. The products at Infineon Technologies are qualified with a FCDM tester according to the JEDEC standard JESD22-C101C [39]. The charging of the chip is done via a charged field plate and a discharge pin (called pogo pin) is used to connect the pin of the IC to discharge the DUT to ground (Figure 2.6).

2.2.3.1 CDM testing issues and SDM as alternative method

CDM discharges depend mainly on parasitic parameters of the DUT and are thus strongly depending on the size and the type of component package. Moreover, a main issue concerning

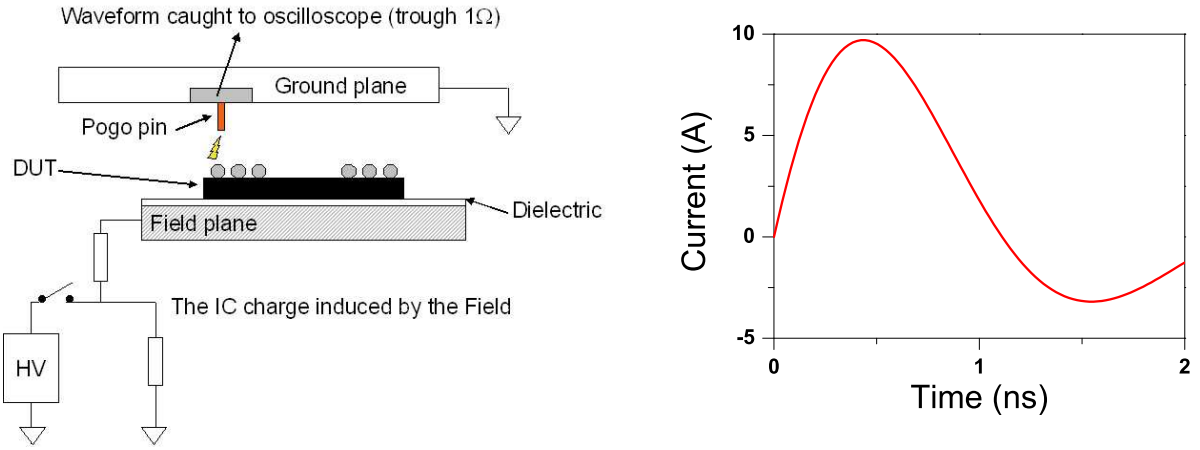

Figure 2.6: DUT under Field CDM test and CDM discharge waveform for an 500 V charge level

the CDM robustness evaluation is seen due to the different testing modes and definition given in CDM standards ESDA STM5.3.1–199 [40], JESD22-C101C [39], AEC Q100-011 [41], EIAJ ED-4701/305 [42]. The standards define the CDM robustness based on the charging voltage level, but the failures obtained from a CDM event type are caused by the peak current [43], not by the charge stored on the device after charging. Furthermore, in standards the recording of the current waveforms is not required, even if this is the main parameter to monitor.

The standards also are not homogeneous and very different robustness values are achieved from one standard to an other. This can be explained also by the distinct peak current per kilovolt pre-charge voltage defined in the standard [44]. As the discharge current is dependent on the package and on the test procedure itself, this can lead to a large variation of the CDM results. Actually, the discharge itself depends on the air moisture, on the shape of the pogo pin, on the discharge pin geometry and on the test procedure. For high charging voltage (>1.5 kV), the peak current of the stress is reduced due to the spark occurring during the fast discharge [44].

CDM testing is suffering from a lack of reproducibility and the complexity in attempting to reproduce a “real world event” has lead to the development of alternative methods. As a work-around to the challenging CDM, the so-called Socketed Device Model (Figure 2.7) is sometimes used instead of CDM.

SDM is a socketed tester which permits a good reproducibility and offers standard test equipments, nevertheless the drawback of this procedure is that it is not as close to the real world event as CDM event and that its discharge characteristic is to some extent dependent on the test system [45]. However, this fast controlled discharge between one hot terminal versus a group of pins is mostly addressing the same failure mechanism than CDM events [46]. SDM is usually more sensitive than CDM, and, therefore, it is used as a screening to identify CDM weak products. The typical value that devices have to sustain under SDM is 500 V, the same value as for CDM.

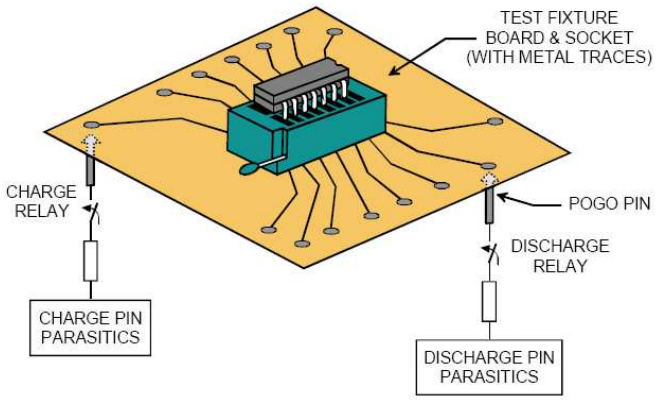

Figure 2.7: SDM test procedure after [45]

In advanced CMOS technologies, the CDM event can be observed in automated lines. It is becoming the major ESD issue and intense work on this topic have been recently done on simulation [36] and on alternative characterization methods as e.g. ultra fast-TLP [47, 48].

It is important to mention that the failure caused by CDM stresses is a typical voltage overstress. Even on small parasitic resistances like metal wiring, the extremely high peak currents can cause significant voltage drops. One failure mode obviously is thus gate oxide fails, and a better understanding of this breakdown mode is needed to cope the strong oxide breakdown voltage decrease with the high density integration and the large stress waveforms variations caused by the lack of reprehendsibility of the CDM testing.

2.2.4 System Level Models

2.2.4.1 System Level HBM

Even for complete systems like mobiles, automotive control circuits, chip cards, etc. ESD is a continuous severe threat. A common test methodology to check the ESD robustness of *systems* is the IEC-61000-4-2 [23]. In the European Community, this test is part of the CE EMC compliance test and, thus, obligatory for all electronic products.

This standard is based on the HBM discharge network but with a capacitor value of 1.5 pF and an 330 Ω resistor. In the standard itself only 3 points are defined to characterize the stressing waveform (Figure 2.8) which could of course result in different ESD withstand voltages for DUTs for different test systems, even if the test systems are compliant with the current specification. Moreover several companies (Thomson, VW, BMW, GM, Mercedes, ...) are using their own standards, varying the capacitor and resistance values and the discharge mode, either by air-discharge or contact [49]. A system level ESD discharge is composed of two different stress region, at first a high CDM-like current peak (1 ns) is applied, followed by an HBM-type

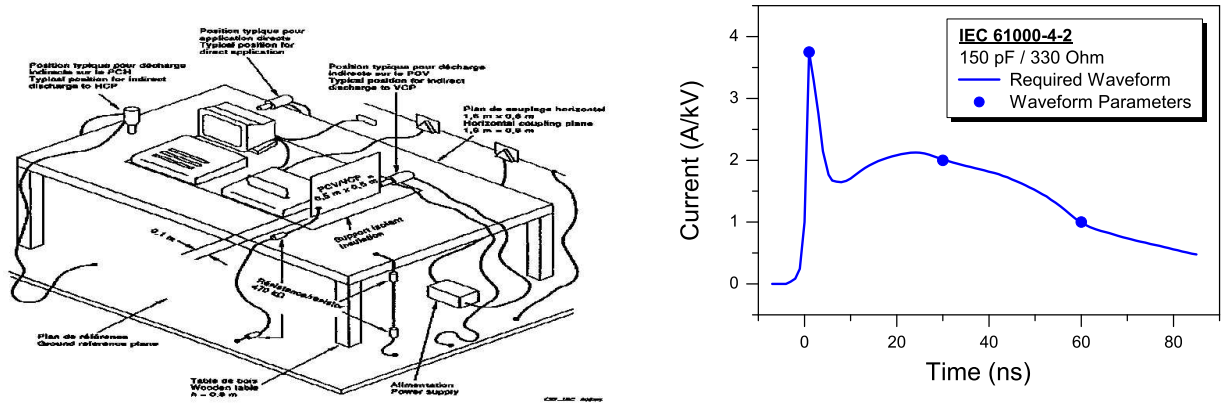

Figure 2.8: IEC 61000-4-2 testing procedure and stress waveform of the gun discharge

stress around 100 ns. After these two stresses, a trailing EOS is also occurring in the range of 10 µs. The stress is applied to the system via an ESD gun and accordingly to the environment description given in the picture 2.8.

2.2.4.2 Recently discussed ESD models

With the ongoing miniaturization and integration, there is an increasing probability that ICs are stressed in a system more or less directly, without "guarding" influence of the board or the ESD protection measures on board. Thus, it is not astonishing that customers are requesting ESD models which reproduce system level stress applied to products. As the ESD system level stress was originally only defined for *systems* and particularly *not* for devices, new standards have to be defined in order to serve industry's demands.

Three models have been recently focused on, namely the Cable Discharge Event (CDE), the Human Metal Model (HMM), and very recently new ideas concerning the Charged Board Model (CBM) have been discussed.

The Cable Discharge Event reproduces the real world event in that an end-user plugs a charged cable into a connector [50]. Currently, the standardization work is focusing on test of systems, but it can be foreseen that similar stress will be applied in future to devices or components. The CDE gives a typical square pulse. The characteristic parameters of the pulse can be calculated by means of theoretical considerations [51].

A long-lasting problem for ESD system-level protection designers is the missing link between ESD robustness on device level (characterized by the conventional device level models HBM, MM, CDM) and the ESD system level robustness (e.g. IEC). Therefore, there is a strong demand from system designers to the IC vendors to characterize the ICs itself with system level stress. The basic idea behind this characterization is that if all devices withstand a certain stress

level, the entire system composed from those devices will withstand at least the same level. Although this idea is not necessarily true and there are several examples which show the contrary, ESDA started to define a system level-type ESD stress methodology to devices. This new model is called Human Metal Model as it reproduce the real world event that a charged person touches an IC with a metallic tool. Besides defining the pulse wave form of such an event, the main purpose of the standardization work is to define a reproducible test environment (i.e. test board) which is as close as possible to the real application. An approach could be the "Zwickau" boards which are used to characterize LIN and CAN bus devices in automotive applications.

In analogy to HMM, the Charged Board Model reproduces a CDM event on system level [52]. The charge is now no longer stored on the IC, but on the board capacitance/system capacitance. Such discharges can occur in automated final assembly lines as reported by Nokia [53]. It is clear that the stress severity could be orders of magnitude higher compared to traditional CDM as the board capacitance is significantly larger. However, the same problems apply as for HMM, and standardization work has yet not started.

2.3 ESD characterization

In the previous section, the different ESD qualification models and procedures have been discussed. However, these verification testing methods give just a "pass" or "fail" information and are therefore not sufficient for the development of ESD protection elements. For the development of ESD protection concepts, a detailed characterization of the DUT in the high-current regime is inevitable. Parameters like turn-on voltage, snap-back voltage and current, differential resistance in the on-state and failure current/failure voltage are important for the definition of an optimized ESD protection concept. Beside these static parameters, dynamic information as e.g. the turn-on time and the voltage evolution during the pulse are required.

In order to avoid a damage of the DUT during the high-current stress, a short time pulse approach is used. Typically, the pulse duration is in the same range as the corresponding ESD qualification methodology. Contrarily to the qualification tests, a pulsed high-current characterization tool allows detailed insight into the physics involved during an ESD stress by monitoring the impedance behavior of the device under short pulses during the stress.

2.3.1 High square pulse testing

A ramp of short current square pulse stress in the time domain of the ESD event is applied to the DUT. For each single pulse, the DUT's transient voltage and current response is recorded, see Figure 2.9. For each voltage pulse applied to the DUT, the average pulse value integrated from the plateau (i.e. the steady state) of the pulse is taken as the DUT's response voltage value. Typical time windows for the voltage/current averaging are between 60 % to 90 % of the pulse

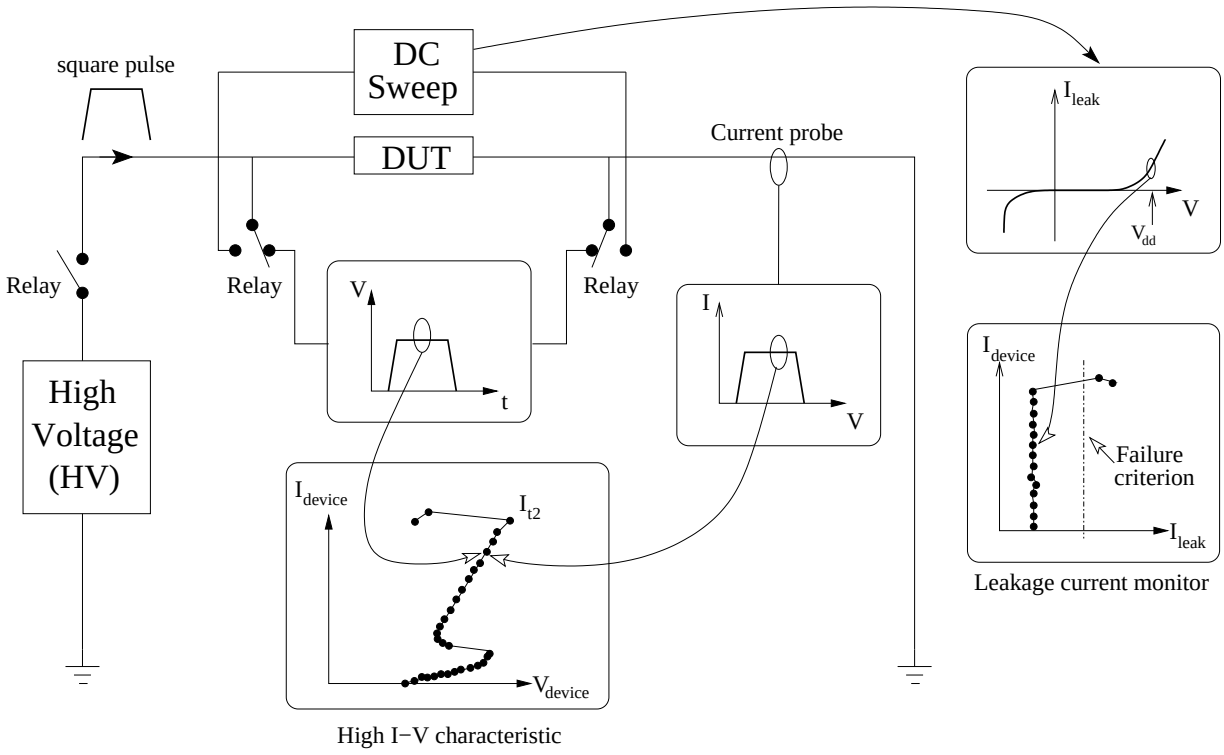

Figure 2.9: Schematic of the square pulsing characterization used for ESD characterization

length. From these two values one point of the (high-current) *I-V* characteristic is derived. The ramping is performed up to the failure of the protection device. A final high current *I-V* trace is compiled from up to 500 pulses, depending on the failure threshold.

The breakdown detection of the device is typically performed by means of leakage measurements (DC sweep or spot criterion). From a DC sweep performed on the fresh device, leakage failure criteria are defined. After each pulse stress the DC criteria are monitored and evaluated. If the DUT exceeds the specification, the device is rated as fail and the high-current characterization stops. On the schematic drawn in Figure 2.9, the usual leakage failure criterion used is defined for an increase in the current at V_{dd}. The failing current of the structure is called I_{t2}.

In the set-ups used for this thesis, the DC measurements are performed either with Keithley 2400 SMUs or with a Keithley 4200 SCS, both integrated into the set-up by relays. When the pulse is applied to the DUT, the source meter is disconnected, the sense channels are in the measurement state and the current probe determines the current forced into the device. After the pulse, the relays switch and the source meter unit is connected to the device for the DC sweep.

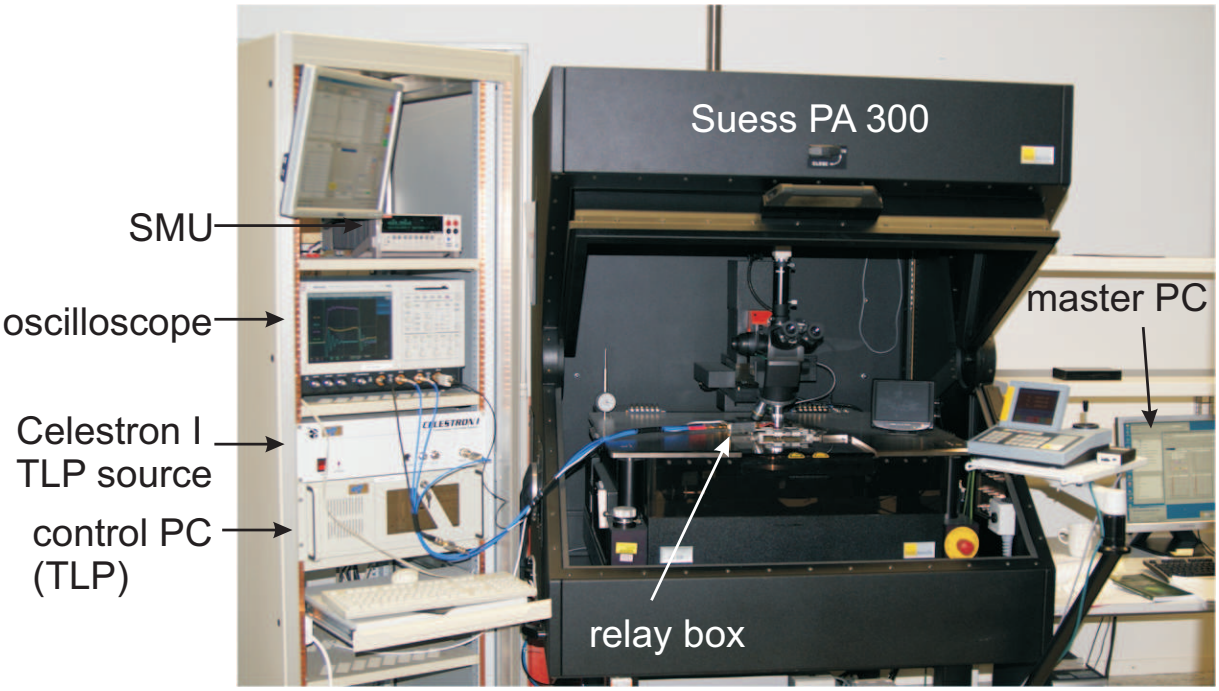

Figure 2.10: Transmission Line Pulse (TLP) used for the high-current *I-V* characterization of ESD protection devices.

2.3.2 Pulse generation in the HBM regime, TLP/Pulser set-up

To perform this *I-V* characterization by pulsing, a special equipment, called TLP (Transmission Line Pulse) has been introduced for ESD time domain by Maloney [54]. TLP systems are today state-of-the-art and this technique is one of the standard tools for ESD engineers. The TLP setup consist of a pre-charged transmission line which is rapidly discharged trough a 1 kΩ resistor to force a well defined current pulse through the device under test. The value of the current pulse is controlled by the pre-charge voltage and the pulse's width by the length of the transmission line. The pulse width for a characterization in the HBM time domain is typically defined to 100 ns, with pulse rise times from 0.5 ns to 10 ns. The TLP equipment available in the laboratory (see Figure 2.10) allows stressing of the devices up to 10 A which corresponds to an HBM stress level of around 15 kV.

Alternatives used for the generation of the square pulses are solid-state pulse generators. The maximal stress level obtained is not as high as the levels reached by TLP systems, but nevertheless it is also a good set-up to establish quasi-static *I-V* characteristics. The solid-state pulse generators have the same functionality as the TLP, except the fact that the pulse applied is generated by a semiconductor circuit targeted for a 50 Ω load. The Agilent pulser (8114A) as a typical example for a solid-state pulse generator is capable of generating a ramp of short rectangular voltage pulses from 1 V to 50 V with variable voltage level steps. For standard device characterization, the pulse width is chosen to be about 100 ns, while the rise time (10 %

to 90 %) is 7 ns which advantageously is comparable to HBM pulse rise times.

There exist many possible configurations for square pulse testing, the main set-ups are TDT (Time Domain Transmission), TDR (Time Domain Reflection), TDRT (Time Domain Reflection and Transmission), and direct current [55]. At Infineon Technologies, the set-ups are built in a current source configuration and permit a large configuration flexibility. Particularly the flexibility in the system impedance is useful to obtain a good accuracy of the DUT response in the triggering phase, as the load lines of the TLP system depend on the set-up impedance which directly impacts the *I*-*V* characteristic.

However, one further important configuration is used to increase the performance of the TLP set-ups, it is the independent force and sense technique with four terminal connections. To establish the real behavior of the device, it is necessary to avoid any parasitic resistance coming from the measurement set-up. The use of two needles, one for the pulse and one for the ground is not sufficient to achieve an accurate voltage resolution at test current levels of about 1 A. In this case, the current flowing through the needles and the metal connections cause a voltage drop and, hence, increase the measured voltage level. The use of four-terminal measurement set-up is an important prerequisite for real sensitive tests. To avoid the parasitic resistance, two needles are used for the pulse stress and two additional needles for the almost current-free voltage sensing. The sense needles are independent from the current flow and so their self resistance do not have any impact on the voltage measurement. This measurement technic is commonly known as Kelvin probe sensing measurement. The added resistance value of the needle probes is around 1 Ω which is quite a lot for the *I*-*V* characterization of the device. Actually, in the high current regime the differential resistance of ESD devices is very low, typical values are in the order of a few ohms. Adding 1 Ω parasitic resistance would unacceptably falsify the measurement results.

Despite an unclear correlation with the HBM qualification tests [15], a lot of interest is seen in these test methodologies as similar failure mode are obtained from square pulsing characterization [16]. Furthermore, it is the only technique which allows the electrical characterization and, thus, gives access to the important device parameters under ESD stress.

In order to pass qualification tests first-time right, the characterization of the protection element is essential and the quasi-static testing provides a very efficient tool for this. The TLP measurement is not yet standardized as a standard test method as the device models HBM, MM, and CDM, however, there is already work on a Standard Practice document at the ESDA [56].

2.3.3 Emerging device transient characterization tools towards CDM

TLP testing for device characterization is nowadays efficient and commonly used in development of ESD protection concepts of ICs. However, this tool operates in the HBM time domain regime and there is currently significantly increased interest to do similar device characterization procedure for fast types of event like CDM. Some set-ups based on the square pulse testing

method have been proposed with the very-fast TLP (approx. 5 ns pulse width), uf-TLP (approx. 1 ns pulse duration) and capacitively-coupled TLP [47, 48]. These set-ups, although working in two-terminal configuration, aims to reproduce CDM and SDM fails [57, 48]. In the nanosecond regime with transient fronts in the range of some hundreds of picoseconds, reliable and accurate test is difficult, but heavily required moreover since CDM fails is the real ESD threat in advanced production lines. However the vf/uf-TLP cannot characterize products CDM robustness but the behavior of ESD devices under short transients. The characterization of the ESD protection devices regarding their turn-on speed is becoming a hot topic for the risk of thin oxides. Lots of interests have been focused on this topics during the last years. An alternative method using repetitive pulses and fourier transformation technique based on a sampling oscilloscope has been proposed to access the transient behavior of the device triggering [47].

2.3.4 Imaging analysis tools

Finally, imaging tools like EMMI (EMission MIcroscopy), back side laser interferometry (BLI) [58] and scanning electron microscopy (SEM) are also regularly used for ESD investigations or reverse engineering studies (for a comprehensive summary of the tools, see e.g. [3, 30]).

EMMI is basically used for failure localizations, detection of non-homogenous triggering or mal-functions of protection concepts. Back side spectrometry is mainly use to study filamentation evolution and temperature distribution in a device during ESD stress. SEM is revealing physical pictures, useful for EOS/ESD failure modes correlation.

2.4 ESD failure modes

From products failures and investigation on ESD fails [59, 60, 61, 1, 3, 30, 62], the typical failures generated by ESD stresses in CMOS technologies observed are:

- Melted junction due to a local temperature increase above the silicon melting temperature. During the ESD stress, the temperature increase is generated by the high power dissipated locally in silicon during the pulse duration.
- Filamentation failure which results in a local meting path corresponding to the current filament discharge course.
- Junction spiking caused by an aluminium filament which intersects the junction. The aluminium filament comes from melt contact metal caused by the current path which solidifies afterwards.
- Thin metallization burn-out which generally results in an open circuit. The overload current during the pulse causes resistive heating which melts the metal thin lines.

- Gate oxide breakdown caused by an excess voltage drop across the oxide layer. It leads to the loss of the oxide insulator property.
- Gate oxide degradations that could occur during an ESD stress due to the injection of charges in the oxide. This results into lower device reliability performance.

Down to the 100 nm technology node, it was clearly possible to separate ESD events onto two main categories which were causing different damages:

- The very fast ESD discharges with an high intensity spike and fast rise time as well as generated by CDM or SDM.
- The fast ESD discharges, with a typical pulse duration of 100 ns and a rise time about 10 ns which can be caused by the HBM or the MM.

The extremely fast discharges of CDM and SDM have a very high current peak which in reverse do not have a lot of energy because of the very short duration (1 ns). However, the associated high voltage drop due to the resistive elements on the discharge path can cause gate oxide breakdown (Figure 2.11)[59, 43]. In general, these gate oxide failures are often located at the periphery or at the interfaces between power networks [1, 63]. Some failures have also recently been reported even in the circuit core [64]. The gate oxide failures are due to over-voltage stress.

The second group of stress which includes HBM and MM is well known to cause some physical damages caused by thermal over-stress due to the comparably large energy (W_{el}) [49].

$$W_{\mathrm{el}} = \int_0^{t_{\mathrm{end}}} V_{\mathrm{dev}}(t) I_{\mathrm{dev}}(t) dt \tag{2.4}$$

The devices heat in temperature which leads to thermal damages of e.g. *p-n* junction or layers, see Figure 2.11 below.

A distinction between the medium or fast ESD events (HBM, MM) and EOS stress is distinguished from the post failure modes observed. The EOS (DC stress) are causing massive melting of large parts of the IC, even phase change and epitaxy [62, 65]. The damage resulting from an EOS stress is correlated with the thermal power (P_{el}) involved during the stress.

$$P_{\mathrm{el}} = V_{\mathrm{dev}} I_{\mathrm{dev}} \tag{2.5}$$

In recent technologies, the ESD events HBM and MM could also lead to gate oxide fails [66]. As the gate oxide breakdown voltage level is dependent on the stress duration and it is lower for longer stress; the voltage drop from the protective resistive path could be sufficient to provoke gate oxide damage, too; even if the peak of current could be five times lower compared to a CDM discharge. In the case of ggNMOS protection element, the well known filamentation failure between source and drain is evolving for thin oxide devices to GOX fails [67, 68, 69]. The gate oxide protection against ESD events is not an exclusive problem of very fast discharges, but a more general issue in advance technologies where thin gate oxides sensitivity should be regarded with attention.

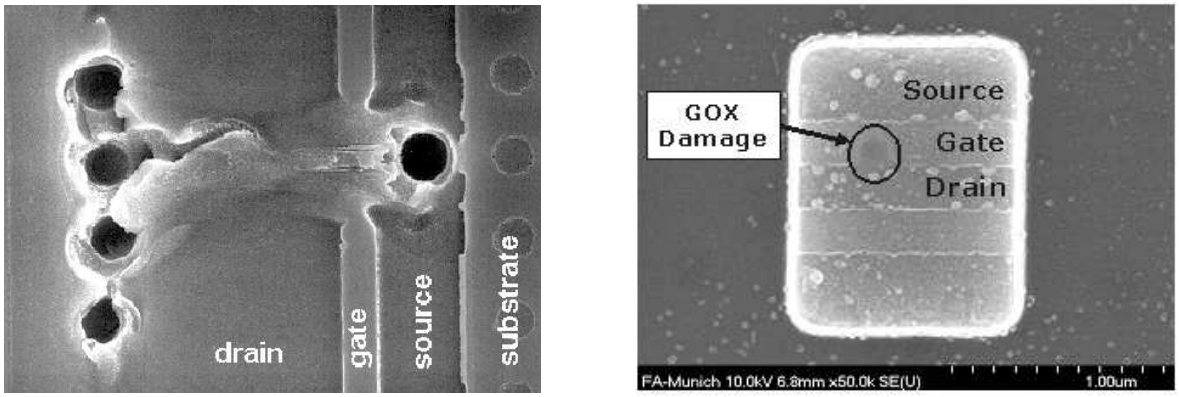

Figure 2.11: ESD failure, left: typical HBM failure mode [3]; right: gate oxide damaged.

2.5 ESD on chip protection

2.5.1 Basic strategy

As an IC is potentially jeopardized by ESD stress, the definition of circuits dedicated to ESD protection on silicon is mandatory. The basic strategy is to shunt the electrostatic discharge current in special ESD paths and, thereby, to guarantee the safety of I/O and core circuits.

All pins which are exposed to the outside have to be protected. A discharge could occur between any pin combination, and in both polarities (as positive discharge or negative discharge). It is the case for supplies pins, ground pins and I/O interface pins. Between all different supplies and ground lines dedicated ESD structures must be included (e.g. ESD power clamps and anti-parallel diodes) in order to enable the discharge current to flow to the grounded pin without forcing its way through the active circuit. Two examples current discharges via low ohmic ESD shunting paths are depicted in Figure 2.12 for an HBM type of stress applied on a chip.

A generic constellation of the ESD protection elements (ESD PE) constituting the low ohmic discharge nets required to prevent from circuit damage from any pin combination stress is described in Figure 2.13. Figure 2.14 describes the two main protection concept strategy for IO pads consisting in

- a local I/O cell (pad) protection approach (using snapback devices see Section 2.5.4).
- or in a rail-based concept which consists in diodes as local pad elements shunting the current flow through a power clamp element. The low ohmic current discharge “ESD path 1” indicated in the Figure 2.12 considers this second shunting concept.

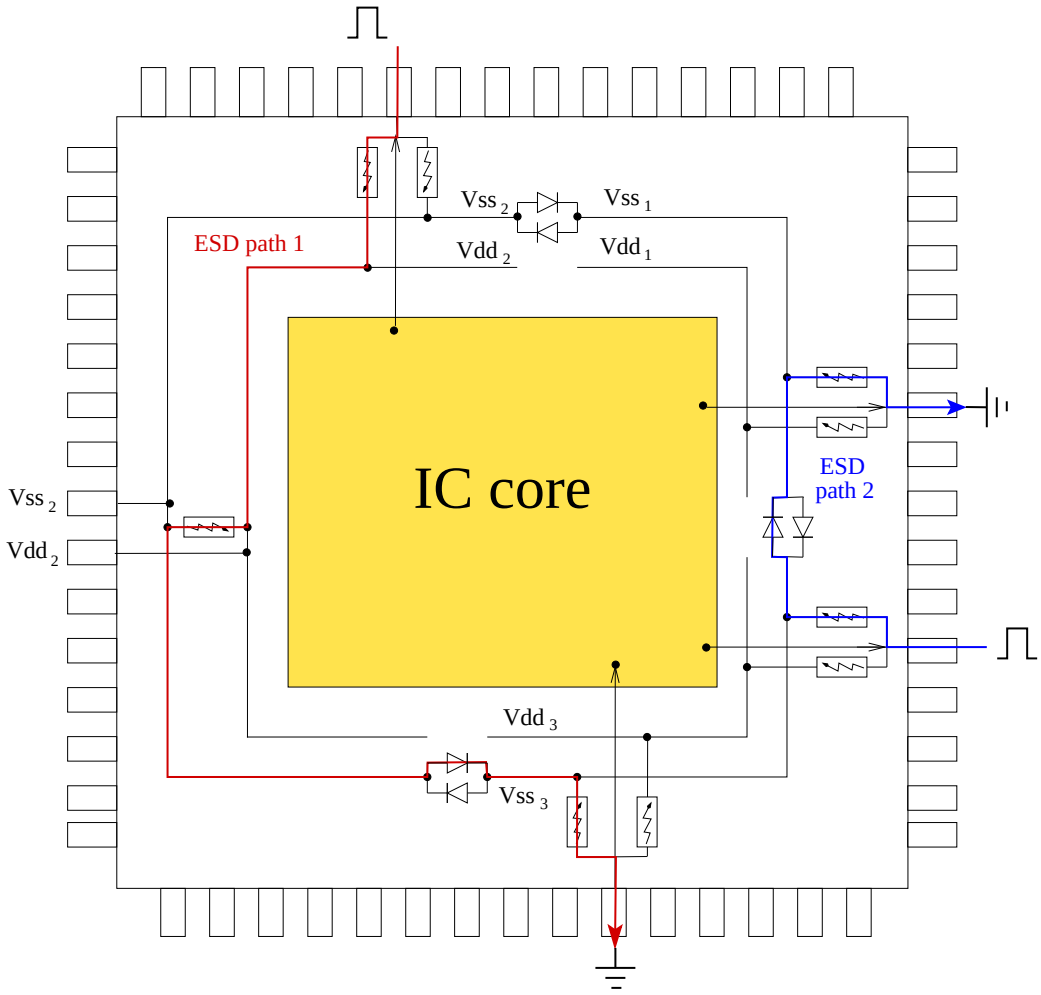

Figure 2.12: Illustration of different low-ohmic ESD current paths between two arbitrary pads. These paths are provided by coupling of signal pads to the supplies via ESD protection elements and by a connection between supplies through anti-parallel diodes.

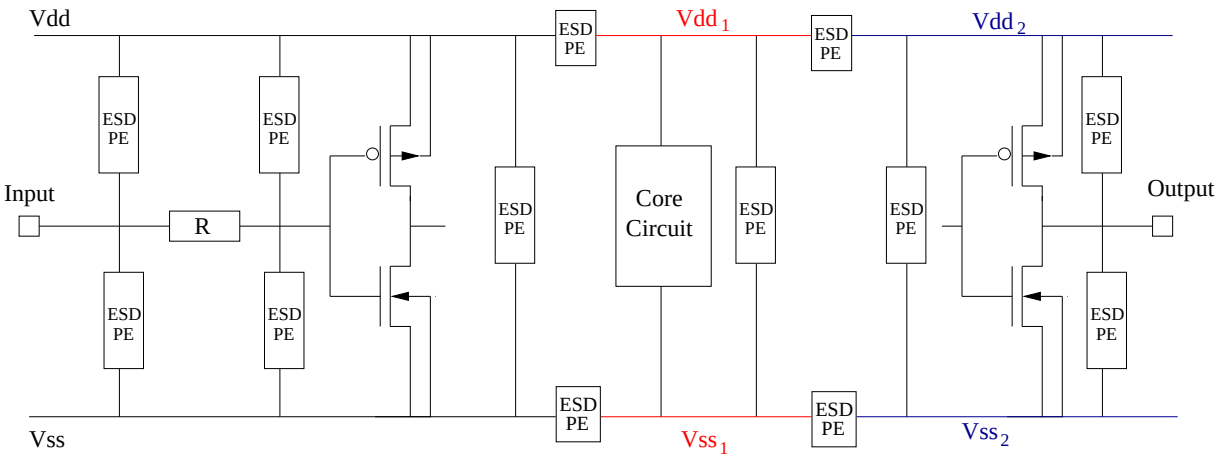

Figure 2.13: ESD protection principe with a typical constellation of ESD protection elements (ESD PE).

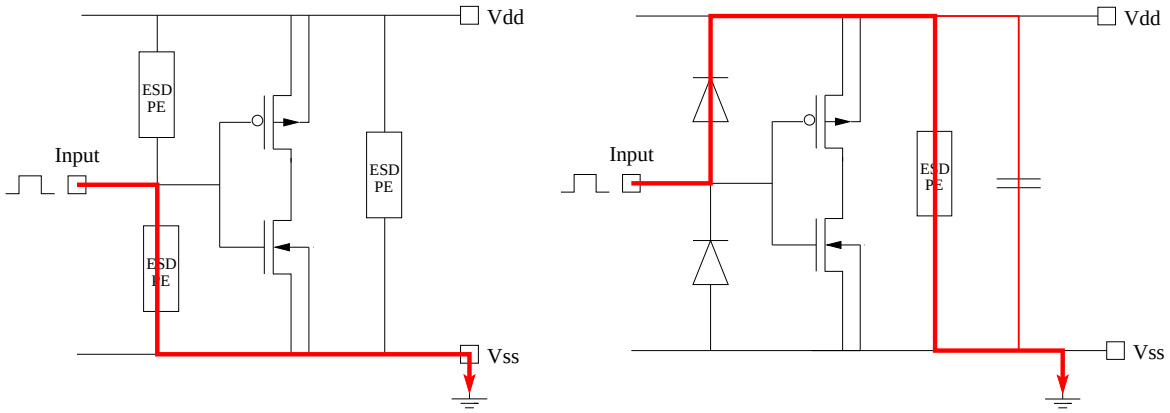

Figure 2.14: ESD protection concepts shown for an input, with the local clamping method (left) and the rail-based approach (right).

2.5.2 Property of an ESD protection element

ESD protection structures must not turn-on during normal IC operation, otherwise device functionality would be heavily distorted. In their non-activated state, ESD protection should be "invisible" with regard to I/O or core circuit functionality. However, during electrical overstress as in the case of an ESD event, they have to trigger and shunt the current to uncritically current paths. In the operational regime it is of a high priority that ESD structures do not cause leakages, which would mean a constant discharge of the batteries of a mobile application or cause a not acceptable high power consumption. In fact, leakages I/O pin protections may pose a severe problem because of the high number of I/O pins that required one ESD structure for each. Additionally, ESD PE must not degrade the performance of the circuitry. As an example, the capacitive load added to the circuit by an ESD PE must not deteriorate the RF performance. Thus, a thorough balance between ESD protection capabilities and parasitic load is required.

2.5.3 ESD design window

In order not to interfere with application functionalities, triggering voltage of protection elements must not be in the range from 0 V to the supply voltage V_{dd}. In addition, usually a safety margin of about 10 % is added to protect the system from unintentional triggering of parasitics structures by noise voltage overshoots in bulk technology (latch-up effect [70, 71]). This results in a forbidden zone in the *IV* characteristic of the protection element, area I in Figure 2.15.

During an ESD event, these structures should resemble short circuits, i.e. provide a very low impedance and non-destructive paths for the discharge in order to limit the stress voltage level at I/O and core circuits. Therefore, the high-current *IV* characteristic from ESD protection should stay considerably below the breakdown of parasitics as gate oxide or junctions, which leads to a

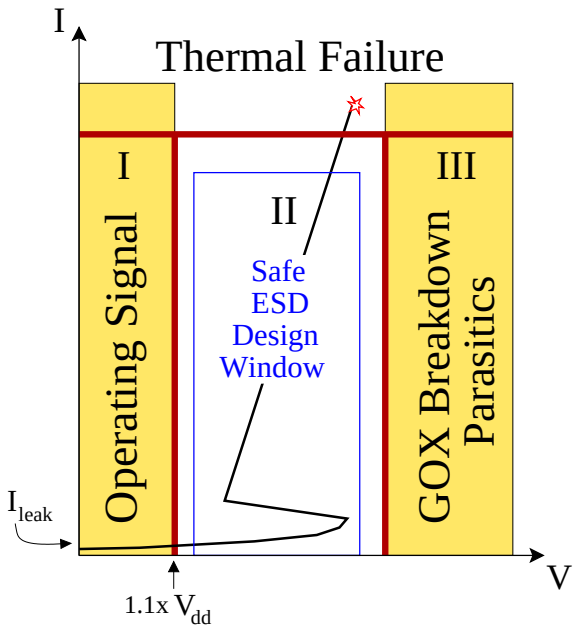

Figure 2.15: Restrictive ESD design window for ESD protection elements quasi-static *IV* characteristic permitted domain.

second forbidden zone, area III in Figure 2.15. It is important to mention that this upper limit is not a fix value and depends on the stress configuration, on the type of the circuit to protect and on other parameters as the stress duration and rise time for example. A detailed review for the correct determination of this limit will be exposed in the Section 6.3.

An ESD protection concept characteristic should be strictly confined in the constrained *IV* characteristic voltage range (area II), also often called "ESD design window". The impact of resistive metal lines or any other voltage drop added to the protection path should be considered and controlled. These overvoltage stresses are dangerous regarding gate oxides damage. ESD protection elements must have their *IV* characteristics supporting very high current values while keeping the voltage drop low.

2.5.4 ESD basics protection elements

Based on the devices available in the technology process or from their parasitics elements, the choice of the best protection concept which fits to the ESD design window and fulfils the circuit requirements has to be carefully determined. The design of protection elements should be optimized to the application. Requirements are various and restrictive, some of these are capability to sustain a high current flow, fast triggering in time, small size, low capacitance value, low leakage, small clamping voltage, wide temperature requirements, ... A brief description of the most prominent basic protection elements which are available in CMOS technologies will be described in the following section.

2.5.4.1 Diodes

Diodes are basic elements and quite frequently used in protection elements for I/O pins in all process technologies. Quite often, "parasitic" diodes are formed in circuit elements, e.g. a n^+p diode is formed between the n^+ diffusion of a driver and the p well. In forward polarity, diodes are able to shunt large discharge currents with low dissipation of power and with a high area efficiency. In that mode, diodes are very robust against ESD discharge, in fact they constitute the most efficient protection in a very simple and clever way. The draw-back of those devices is that they just constitute unidirectional protection and are not sufficient as stand-alone elements. It is necessary to add another structure for the other reverse ESD path or to use a rail-based concept approach.

An easy diode protection example between pad and V_{ss} is shown in the Figure 2.16. For negative discharges at the pad or positive discharges applied at V_{ss} versus pad, the diode will properly shunt the current and does not disturb the operating signal range. The same applies for diodes connected in the similar configuration between pad and V_{dd} so as for V_{dd} and V_{ss}. Nevertheless diodes from pad to V_{dd} cannot be used for over-tolerant pads.

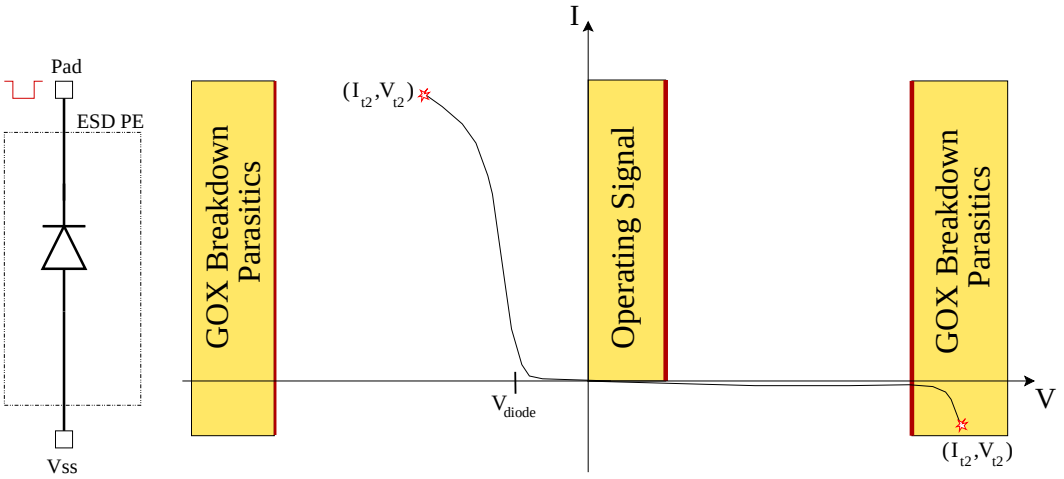

Figure 2.16: Left: example of a diode used as a protection element between pad and V_{ss}. Right: diode high IV characteristic.

2.5.4.2 Grounded gate NMOS (ggNMOS)

From their construction, all CMOS transistors are accompanied with parasitic bipolar transistors, i.e. lateral *npn* or *pnp* transistors built from their active diffusions and wells. For an nFET transistor the parasitic *npn* (Figure 2.17), is constituted by the n^+ diffusion of the drain which acts as the collector, the p well for the base and the source diffusion n^+ which plays the role of the emitter. This bipolar transistor can be used as a protection structure when the drain is

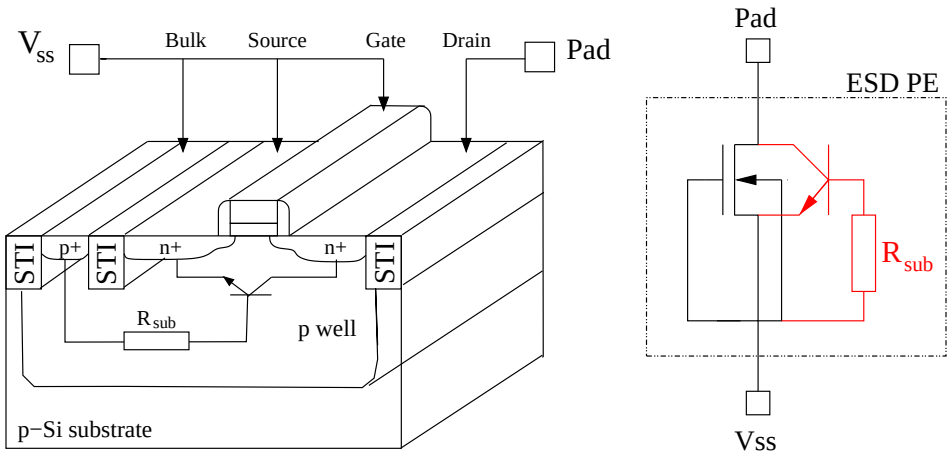

Figure 2.17: Parasitic bipolar of an nFET transistor.

connected to the I/O (or supply) pad to protect and the gate, the collector and the bulk grounded. In this mode the FET is always off and the bipolar transistor is only activated by an ESD event.

ESD operation of the ggNMOS

When an ESD event occurs, the drain level increases and polarizes in reverse mode the collector-base junction (drain-bulk). If the discharge is high enough then it can provoke the avalanche breakdown of the junction. This phenomena occurs at a threshold voltage called V_{bd} (breakdown). A hole current flows towards the ground through the substrate and the voltage drop across the substrate resulting from its resistivity can forward-bias the base-emitter junction (source-bulk). The bipolar effect is triggered and generates a huge electron current through the base collector junction. This current is able to maintain easily the bipolar conduction due to the corresponding voltage drop at the base-emitter, so at a certain current and voltage level (I_{t1}, V_{t1}) the voltage of the bipolar rapidly collapses, leading to the so-called “snap-back” phenomenon. The snap-back leads to a negative differential resistance part, ending in the sustaining point (V_h, I_h). From the sustaining point on to even higher breakdown currents, the bipolar characteristic shows a low ohmic behavior up to the point where the device is destroyed thermally (V_{t2}, I_{t2}). In order to avoid unintentional triggering of the protection element, the minimum value of the sustaining point (V_h) must be higher than the signal voltage.

Snap-back characteristic devices are commonly used as ESD protection elements, as the improvement of the clamping behavior and the reduced power dissipation during an ESD stress allow to achieve robust protection elements.

In contrast to a simple diode, the ggNMOS provides protection in both pulse polarities, as it has a built-in n^+/p well junction diode from drain to bulk. This diode comes into play at negative ESD stress at the I/O pad versus ground. A typical ggNMOS high current-voltage characteristic is exposed in Figure 2.18.

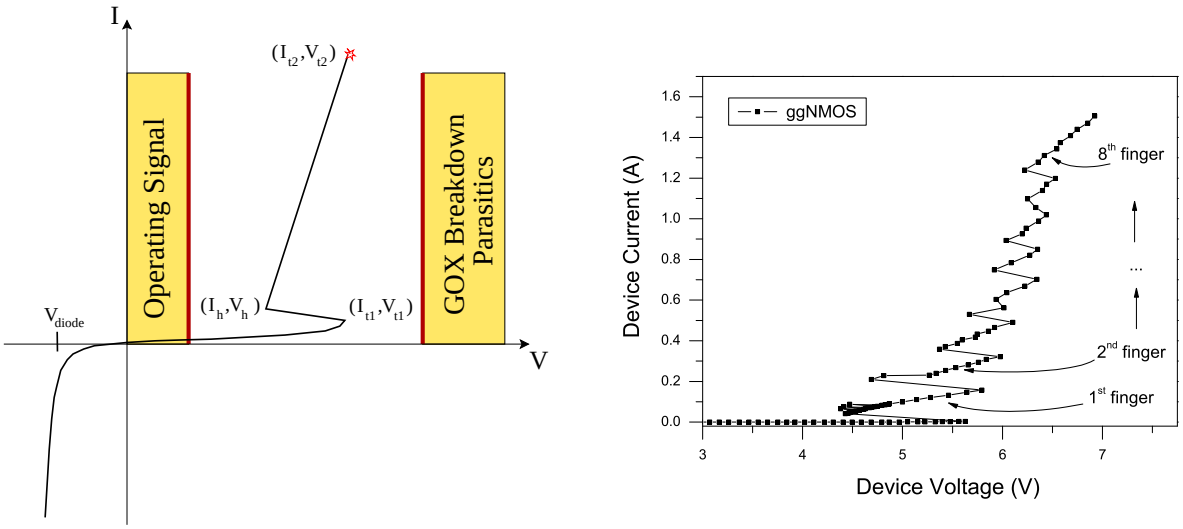

Figure 2.18: Left picture, a ggNMOS high current-voltage characteristic. Right picture, an eight finger Output driver with butted substrate layout high current-voltage characteristic obtained by TLP testing.

It is possible to control the *IV* behavior of a ggNMOS device with layout and process parameters [30, 3]. A rapid list of the main parameters with their impacts is given here after:

- Substrate resistor: determines the value of the triggering current (I_{t1}, V_{t1}). Can be tuned with the distance of the bulk contact.
- Channel length: modifies the point of the sustaining voltage (V_h).
- n^+ diffusion of source/drain sheet resistance: determines the slope (R_{diff}) in the resistive differential regime of the bipolar.

The I/O output drivers which are requiring a large area to provide a high output signal are usually using their own snap-back modes as ESD protection solution. In that case, due to area restrictions and to the sheet poly resistance, the layout of big ggNMOS is decomposed in several small gate widths in parallel (typically in the range of 25 µm). Each gate width is composing one finger of the entire structure. During an ESD stress, if just one finger is driving the discharge and is getting damaged before the shunting contribution of the rest of the device, the protection element will be weak and unusable. This kind of failure was typically seen with the introduction of the silicidation step in advance processes. The root cause of the damage lies in a filamentation of the ESD current and the occurrence of critical hot spots along the ggNMOS device width. To avoid this kind of failure, ballasting resistance is necessary and was reintroduced for a modern CMOS technology by blocking the silicidation process in diffusion areas [72].

Ballasting resistors enables the triggering control of the successive fingers (see Figure 2.18), leading to an uniform distribution of the current flow in the full width of the protection device [73]. Hot spots and filamentation phenomena due to an excessive local current density are then not occurring anymore. This effect is transcribed in the design rules by the fixation of the drain contact holes to gate distance and diffusions to substrate distance. When the source diffusion and the substrate are connected without space in between, the layout style is called butted substrate and the *IV* curve shows pronounced successive finger triggering effect as depicted in the Figure 2.18. These distances combined with the silicide blocking step generate the needed ballasting resistor (via the sheet resistance of the n+ doped diffusion). Nevertheless this design rule is area consuming as the drain to gate distance required for ESD robust design is larger than the minimum spacing defined for the process.

Down to the sub-100 nm node, the ggNMOS is used as a basic brick in the ESD protection constellation. This is due to the good characteristic device control capability under ESD stresses based on the good physic understanding [1, 30, 74] which is enhanced by beneficial TCAD simulations [3, 75]. However, beyond the 90 nm node alternative to ggNMOS are preferred due to the too large area required (dependent on the ballasting resistors) to withstand high current levels.

2.5.4.3 NMOS current driver (BIGFET)

As a clamping element, a sufficient large NFET (big FET) can be used to drive the huge current of the ESD discharge. During the stress, the transistor operates in the saturation regime. This transistor is generally triggered by a *RC* time constant below the µs to shunt over-voltage discharge, see Figure 2.19. However this concept could be used only for a limited signal frequency regime and the element could turn on due to high frequency noise. In case of an basic protection concept the bigFET is consuming a lot of area. Several derived concepts [8, 76, 77, 78] using big PFET and feedback circuits have been proposed too minimize the size, reduce the *RC* constant [79] and avoid the inopportune FET breakdown. Some variation with reduced big NFET size (RC triggered nFET) operating first in the MOS current driver and then entering in the bipolar mode smoothly without snapback as shown in the Figure 2.19 gives also a good alternative for the area consumption versus ESD fast triggering and robustness.

2.5.4.4 Silicon Controlled Rectifiers (SCR)

Silicon Controlled Rectifiers (SCRs) also known as thyristors [1, 30, 80, 81, 82], are devices which are used extensively in power device applications because of their capability to switch from a very high impedance state to a very low one. That is also the reason why this special lateral *npnp* layer element is used as ESD protection element.

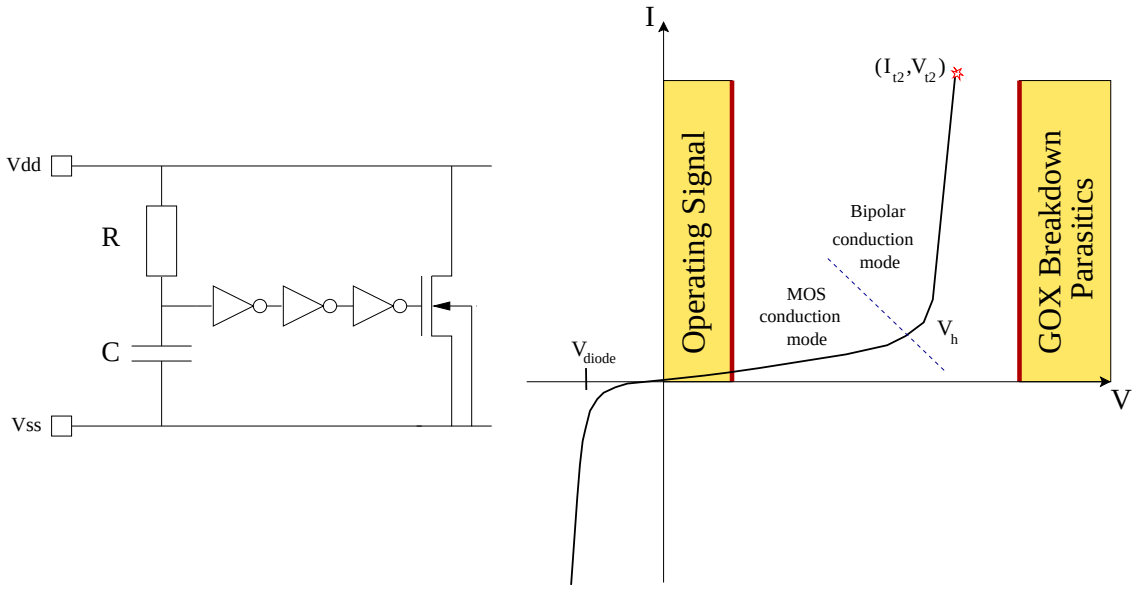

Figure 2.19: Left: protection concept based on a bigFET driver. Right: *RC* triggered nFET *IV* characteristic showing first, the MOS current driving mode and then, the bipolar current regime after (V_{t1}-I_{t1}).

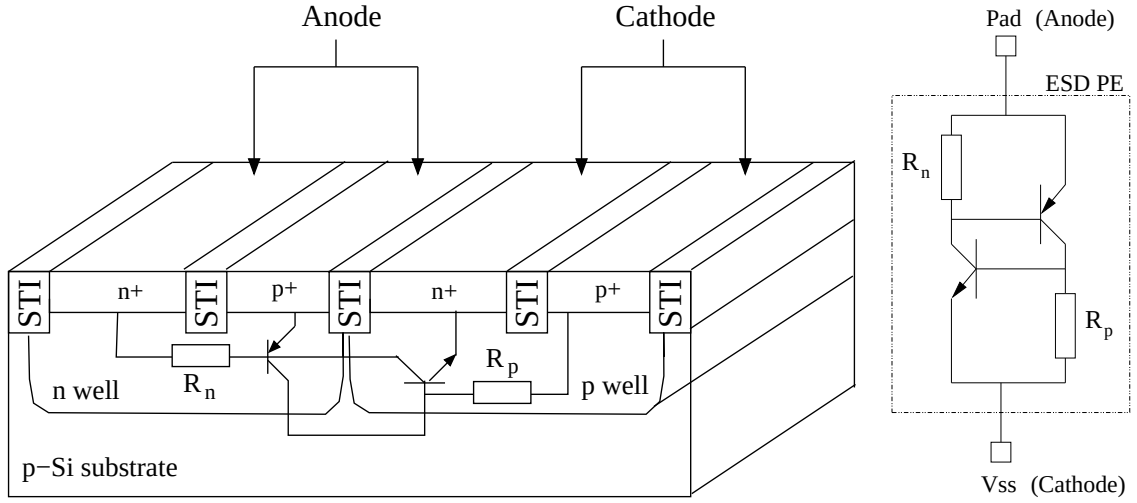

Figure 2.20: SCR cross section and equivalent electrical circuit model

The p^+ diffusion in the n well forms the anode of the thyristor where holes are injected in the n well. The n^+ diffusion in the p well constitute the cathode of the SCR from where electrons are injected into the p well. The contact for the n well is realized by the n^+ diffusion which is linked to the SCR core through the n well resistance R_n. The p well contact as well is established through a p^+ diffusion and linked by an effective resistance R_p. The SCR is usually considered as two bipolar transistors with two access resistors R_n and R_p one lateral *npn* and one both vertical and lateral *pnp* as shown in Figure 2.20.

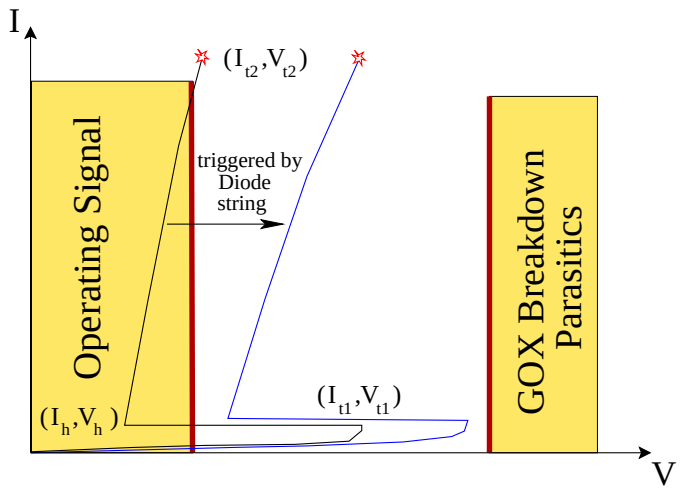

Figure 2.21: SCR high *IV* characteristic.

There are several functional modes for an SCR depending on the connection of the anode, cathode, *n* well and *p* well. Used as an basic ESD protection element, the anode is connected to the node to protect and the cathode is grounded. In normal IC operation the SCR protection element is to be turned off, i.e. base-emitter potential of the SCRs NPN and PNP device has to be zero, which means that the *n* well contact should be tied to the anode and the *p* well contact to the cathode. The SCR used as a basic protection element is described in the Figure 2.20.

As for the ggNMOS the SCR high current *IV* trace shows a snap-back behavior and a differential resistance high current branch. However, the physics involved is a little different and some more parameters can be use in the control of the SCR behavior. To some extent the SCR characteristic is easier to control due to the possibility to adjust critical parameters in process or layout. SCR are used in all process technologies and with various triggering modes. Nevertheless, SCRs show a very low holding voltage which might cause latch-up problem [70, 71] and support just an unidirectional polarity protection. In that case, an other element (usually a diode) should be placed in parallel to shunt the non covered discharge.

The use of SCRs in deep sub-nanometer technologies has strongly increased due to their area efficiency and their capability to clamp the voltage to very low values. SCRs concept are very flexible and a lot of optimization work have been done to provide aggressive protection concepts, especially for power clamp elements. Nevertheless a draw back is seen in their triggering speed which is critical for very fast discharges. This lack of reactivity provokes over-shoots and lead to a critical issue for the protection of the thin oxides from fast discharges [83]. This is furthermore emphasis as the SCR are commonly triggered via diodes string to avoid latch-up by shifting the SCR characteristic in the ESD design window (see Figure 2.21). The snapback part of the curve during the triggering of the device is a major concern in advanced CMOS technologies in respect to CDM type of events.

2.6 ESD protection in deep sub-micron MOS technologies

2.6.1 ESD challenges

2.6.1.1 Process challenges

With the continuous shrinking of CMOS processes and the target to reach better performances, ESD protection development task is getting more complicated [84, 75, 13, 85, 86]. Technologies evolutions like introduction of LDD implants, salicided diffusion and gates, STI, SOI (Silicon On Insulator), higher doping concentrations, are changing the behavior of the protection devices under ESD stress. This impacts strongly the ESD robustness, see Figure 2.22. Protection elements and concepts are not easily transferable from one technology to the next one [87], a continual development and improvement is needed to face the challenging ESD robustness achievement. The process complexity induces more sensitivity to parasitics effects [88] and the technologies evolution is bringing stronger boundary limits (section 2.6.2) with the down scaling of the oxide thickness and the increase substrate doping concentration (GOX breakdown and latch-up).

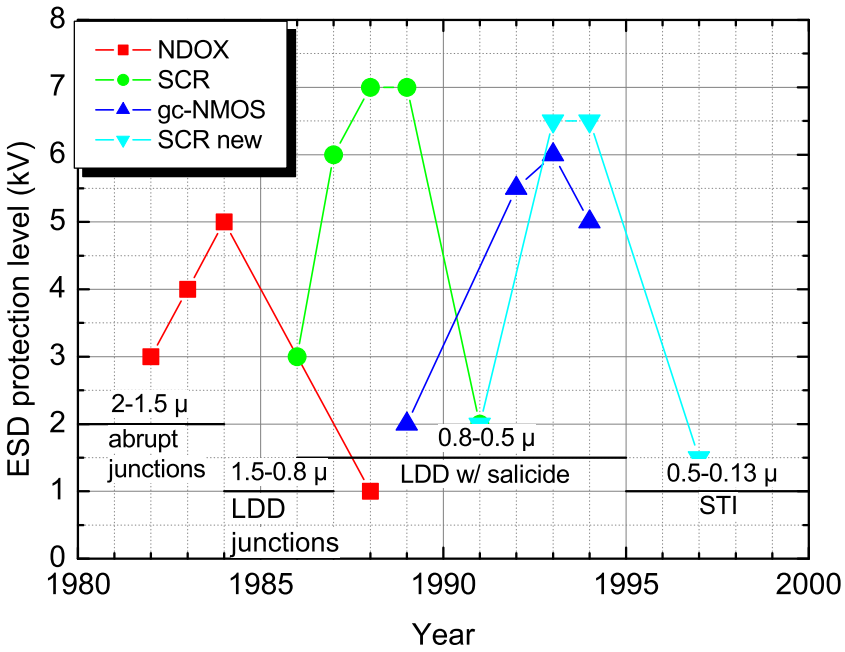

Figure 2.22: History of the ESD Robustness trend as a function of the CMOS process evolution and protection devices after [1].

2.6.1.2 Application requirements

The processes changes coming along with the new technologies is not the only constraining factor, the application requirements are also participating a lot in the limitations. With the diversification of the application, fast interfaces and hot plug interface like HS-USB, FireWire, ethernet, the challenge is also strongly reenforced. For high frequency applications, as the ESD elements are inducing some parasitics effects, the complexity degree of the protection concepts is very high as ESD co-design counter measures should be done to preserve the high performances and protect thin oxides devices needed for RF performances and furthermore connected directly at the I/O pad [89, 90]. For USB connection, it is even worth as it is not an interface but an entire system which can be powered during an ESD event [86, 91].

Another aspect to take into account is the product's high pin count. Advance CMOS ICs have from several hundreds to few thousands pins. This implies a complex ESD discharge networks counting around 1000 protection devices. ESD Analysis methodology at chip level (ESD full-chip simulation) become important for complex products verification [92]. The package solution for modern ICs poses also a challenge regarding the CDM robustness performance as the package size trends to increase with flip-chip solution for example. The package size is linked to the capacitance value and, hence, to the current peak during the CDM discharge [43]: in general the bigger the size of the package is, the higher is the discharge current level.

In advanced technologies, a large variety of constraints is valid. In this context it is essential to acquire a precise knowledge on the operating margin left for the conception of robust designs as the use of critical circuits architecture will be more and more needed.

2.6.1.3 Protection strategy, robustness, area and complexity trade-off

As a matter of cost saving, the size dedicated to ESD protection should be minimum. The optimization task between performance and area efficiency leads to different strategy for the protection. In advanced CMOS technologies, the choice between a local pad protection and rail-based protection concept is the first typical question which arises. With rail-based concepts, it is possible to save area on high pin products uses diodes instead of bigger structures as local I/O protection. However, this strategy relies on a good power clamp element with a fast triggering and requires low ohmic connections in order to avoid additional voltage drops. Moreover, the main critical protection element for sub 100 nm technologies is certainly the core supply power clamp element. In case of an ESD stress directly between the supply and the ground, the critical large oxide from buffer capacitors are endangered and can not tolerate high voltage overshoots. This is also especially a challenge in case of an rail-based concept because input gates voltage limitations could become problematic. Let's consider an example focused on a SCR power clamp element designed for the protection of an 2.2 nm oxide thickness. For an ESD stress occurring on an input stage versus the ground line the voltage drop seen at the nFET gate oxide can then be expressed as,

- $V_{\text{Input}} = V_{\text{diode}} + I_{\text{ESD}} \cdot R_{\text{Diode}} + I \cdot R_{\text{VDDbus}} + V_{\text{PowerClamp}} + I_{\text{ESD}} \cdot R_{\text{VSSbus}}$, in case of a railed-base approach.
- $V_{\text{Input}} = V_{\text{LocalClamp}} + I_{\text{ESD}} \cdot R_{\text{VSSbus}}$, with a local protection concept.

Regarding the safeness of the thin oxides, the choice of a railed-based concept can be critical in comparison to a local protection concept due to the added voltage drops. This is exposed in the Figure 2.23 for an 2 kV HBM stress. For very fast discharges as CDM events, these voltage drops are even more critical and pronounced as the current peak is in the range of 10 A. This is also purely cumulative with over-shoots due to slow triggering of the protection element that can be induced under fast transient stresses. This general voltage drop (*IR* drop) issue is more severe in low complexity technologies or low cost variations for specific products where minimal metal stack layers are used leading to high *IR* drops from the bus systems.

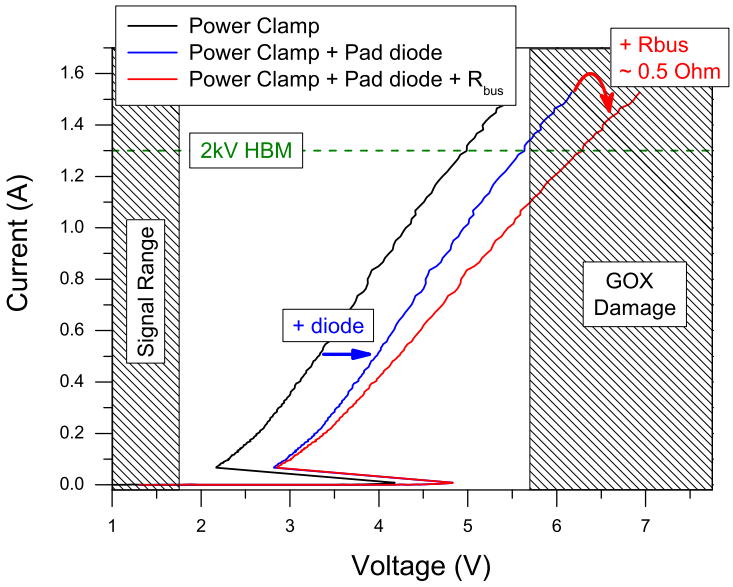

Figure 2.23: TLP curve of an SCR protection element embedded into a railed based protection concept. The impact of the diode and buss resistance are described.

Concerning the protection element portfolio, the four basic cells previously discussed (diodes, ggNMOSs, triggered big FET and SCRs) are basically used. However, protection concepts are not necessarily using these cells exclusively as single elements. In order to meet better ESD performance and area efficiency, combinations and cascaded protection stages are often employed. A two stage protection is commonly used for input protection for example to ensure the safety of input gate oxides for example. Parasitic or lateral bipolar transistors are also often implemented as ESD protection elements as well. To achieve the imposed constraining performances, protection concepts are then increasing in sophistication along the down scaling of technologies. To

over come the requirements with the minimal area cost and maximal performances, declinations of the basic elements are preferred; this has lead to some advanced concepts listed hereafter:

- diode concepts: diode strings [88], ...
- nFET concepts: zener diode triggering nMOS, capacitor triggering nMOS (gcNMOS), RC triggering circuitry, distributed power clamps [93], ...
- SCR concepts: Diode Triggering SCR (DT-SCR), Low-Voltage Triggering SCR (LVT-SCR), TT–SCR [89], gg-SCR [94], NPN triggering SCR, ...
- bipolar concepts: lateral bipolar [95], LVT-PNP, ...

Plenty of small variations and optimizations of protection concepts could be found in the numerous publications from the EOS/ESD symposiums over a wide technology range and application specific concerns. For the specific topic of ultra-thin oxides protection, more and more sophisticated solutions are proposed to face the issue (designs, systems) [94, 96, 97, 98].

2.6.2 ESD design window trend

The determination of the correct ESD design window trend is a crucial point as it constitutes the road-map for the ESD protection development. With the aggressive down-scaling of the oxide thickness, the breakdown of these thin oxides is becoming a major issue for the ESD design [85]. The technologies evolution lead to the approaching diode junction breakdown close to the supply voltage due to the increasing doping concentration. The triggering voltage of bipolar parasitic V_{t1} is in the same range as the GOX breakdown below the 90 nm CMOS technology node. Furthermore the desire of supplying the advanced technology at over operating condition contributes also to the drastically reduction of the ESD design window. The basic ggNMOS protection element used up to 0.1 µm gate length technologies node cannot be used anymore as a stand-alone element in sub-µm CMOS technologies. Up to now, the characteristic shifts of the *IV* protection concept due to *IR* drops were not a big issue as the ESD margin regarding the parasitics breakdown was sufficiently large and the *IV* variations could still be maintained in the ESD design window. In advanced CMOS technologies, only a very narrow safety margin is left for ESD protection development, see Figure 2.24. This increases the probability to generate overshoots on protected GOX areas during an ESD stress. The question of non-destructive stresses on thin oxides requires an important consideration as the clamping voltage are really very close to oxide breakdown values. On the circuit side, the precise knowledge of the GOX breakdown voltage is needed also for deciding of products architectures and circuits released. This meets particulary the interfaces which mix device voltage class as level shifters, voltage regulators, current mirrors where the thin oxide devices are in an hostile environment.

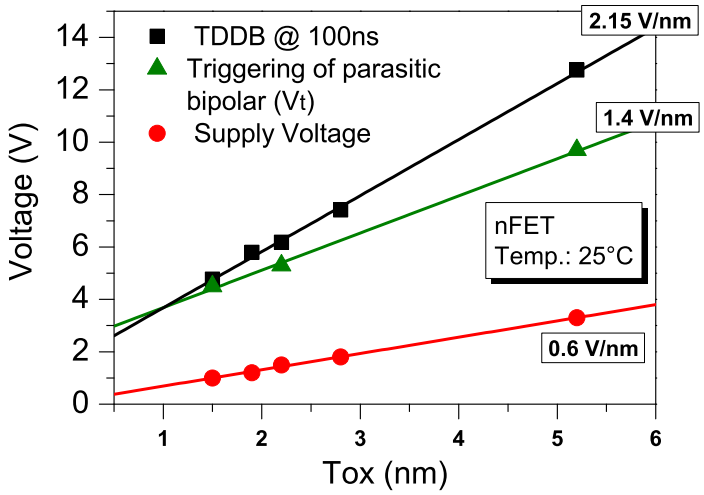

Figure 2.24: ESD design window trend in the HBM time domain as a function of the oxide thickness. The oxide breakdown as well as the supply voltage and the turn on of the parasitic bipolar described the ESD design window and the safety margin available.

In advanced CMOS technologies, the trend of CDM oxide failures is increasing [99] and underlines the necessity to take care also of over-shoots and of the transient triggering speed of the protection elements under CDM events. The gate oxide reliability is becoming a much more serious concern as it has been in the past. Gate oxide has always been a serious reliability topic for device and dielectric lifetime. Lots of investigations have been done in advanced CMOS technologies in that field (Hot Carrier, NBTI, TDDB, Plasma Induced Damage (PID) ...). However, thin gate oxides degradation behavior under short pulse stress is still unclear and needs to be fully investigated. Especially the lack of knowledge about parameters impacting the gate oxide reliability in the ESD regime has to be fulfilled. The risks to meet products fails with thin gate oxides are becoming too big to keep approximative approach and consider just 100 ns TLP measurements on few GOX structures. A clean and controlled safe design should be guaranteed.

2.7 Conclusion

This second chapter has given beyond the basic introduction to the ESD issues in CMOS semiconductors, the clear challenge of ESD protection counter-measures to achieve reliable products in advanced technologies with a high manufacturing yield. One pre-dominant limitation to the ESD robustness of thin oxide processes is the failure of this thin insulator films. This major issue will be investigated in the following chapters. First, the experimental details and the technologies investigated will be developed in Chapter 3.

Bibliography

[1] Amerasekera Ajith and Duvvury Charvaka. *ESD in Silicon Integrated Circuits*. John Wiley & Sons, LTD, Chichester, England, second edition, 2002.

[2] Baumgärtner H. and Gärtner R. *ESD, Elektrostatische Entladungen (in German)*. Oldenbourg, Munich, Germany, 1997.

[3] Esmark Kai, Gossner Harald, and Stadler Wolfgang. *Advance Simulation Methods for ESD Protection Development*. ELSEVIER Science Ltd, Oxford, England, 2003.

[4] Walther A. Internal ESD siemens report. Technical report, 1986.

[5] Green T.J. and Denson W.K. A review of EOS/ESD field failures in military equipment. In *EOS/ESD Syposium Proceedings*, pages 7–14, 1988.

[6] McAteer O. ESD: A decade of progress. In *EOS/ESD Symposium*, pages 1–6, 1988.

[7] Euzent B.L., Maloney T.J., and Donner J.C. Reducing failure rate with improved EOS/ESD design. In *EOS/ESD Symposium*, pages 59–64, 1991.

[8] Merill R. and Issaq E. ESD design methodology. In *EOS/ESD Symposium*, pages 233–237, 1993.

[9] Wagner R.G., Soden J.M., and Hawkins C.F. Extent and cost of EOS/ESD damage in an IC manufacturing process. In *EOS/ESD Symposium*, pages 49–55, 1993.

[10] Shumway S. Keynote address. In *EOS/ESD Symposium*, 1995.

[11] Brodbeck T. Correlation between ESD hardness of semiconductor devices and field failures. In *ESD Forum*, 1997.

[12] Brodbeck T. ESD tutorial. Internal Workshop, 2005.

[13] Gossner H., Domanski K., Drüen S., Esmark K., Pessl P., Russ C., Stadler W., and Zängl F. SoC - a real challenge for ESD protection? In *EOS/ESD Symposium*, 2005.

[14] Dangelmayer T. ESD program management; a realistic approach to continuous measurable improvement in static control. New York. Van Nostrand Reinhold.

[15] Musshoff C., Wolf H., Gieser H., Egger P., and Guggenmos X. Risetime effects of HBM and square pulses on the failure thresholds of ggNMOS transistors. In *Microelectronics Reliability*, volume 36, pages 1743–1746, 1996.

[16] Stadler W., Guggenmos X., Egger P., Gieser H., and Musshoff C. Does the ESD - failure current obtained by transmission–line pulsing always correlate to human body model tests. In *Microelectronics Reliability*, pages 1773–1780, 1998.

[17] Roozendaal L.J., Amerasekera A., Bos P., Baelde W., Bontekoe F., Kersten P., Korma E., Rommers P., Krys P., Weber U., and Ashby P. Standard ESD testing. In *EOS/ESD Symposium*, 1990.

[18] Veraege K., Roussel P.J., Groeseneken G., Maes H.E., Gieser H., Russ C., Egger P. Guggenmos X., and Kuper F.G. Analysis of HBM ESD testers and specifications using 4th order lumped element model. In *EOS/ESD Symposium*, pages 233–237, 1993.

[19] Military Standard for Test Methods and Procedures for Microelectronics. ESD sensitivity classification. MIL STD 883.C/3015.7 notice 8, 1989.

[20] JEDEC Solid State Technology Association. Electrostatic discharge (ESD) sensitivity testing human body model (HBM), 2006.

[21] Working Group WG 5.1 ESD Association. ESD standard test method for electrostatic discharge sensitivity testing - human body model (HBM) component level, 2001.

[22] Automotive Electronics Council (AEC). Human Body Model (HBM) Electrostatic Discharge Test. AEC - Q100-002 - Rev-D.

[23] International Electrotechnical Commission (IEC). Electrostatic Discharge Immunity Test, 1996.

[24] Electronic Industries Association of Japan (EIAJ). Environmental and endurance test methods for semiconductor devices, (HBM).

[25] Industry Council on ESD Target Levels. White Paper 1: A Case for Lowering Component Level HBM/MM ESD Specifications and Requirements, 2007.

[26] Brodbeck T. and Gaertner R. Experience in HBM ESD testing of high pin count devices. In *EOS/ESD Symposium*, pages 1–6, 2005.

[27] Stadler Wolfgang. State-of-the-art in ESD standards. In *International Electrostatic Discharge Workshop*, pages 127–147, 2007.

[28] Meuse T., Ting L., Schichl J., Barrett R., Bennett D., Cline R., Duvvury C., Hopkins M., Kunz H., Leiserson J., and Steinhoff R. Formation and suppression of a newly discovered secondary EOS event in HBM test systems. In *EOS/ESD Symposium*, 2004.

[29] Duvvury C., Steinhoff R., Boselli G., Reddy V., Kunz H., Marum S., and Cline R. Gate oxide failures due to anomalous stress from HBM ESD testers. In *EOS/ESD Symposium*, 2004.

[30] Russ Christian. *ESD Protection Devices for CMOS Technologies: Processing Impact, Modeling, and Testing Issues*. Shaker Verlag, Aachen, 1999.

[31] Working Group WG 5.2 ESD Association. ESD standard test method for electrostatic discharge sensitivity testing - machine model (MM) component level, 1998.

[32] JEDEC Solid State Technology Association. Electrostatic Discharge (ESD) Sensitivity Testing Machine Model (MM).

[33] Automotive Electronics Council (AEC). Machine Model (MM) Electrostatic Discharge Test. AEC - Q100-002 - Rev-E.

[34] International Electrotechnical Commission (IEC). Electrostatics part 3-2: Methods for simulation of electrostatic effects – Machine model (MM) electrostatic discharge test waveforms).

[35] Veraege K. Component level esd-testing. In *ESREF smposium*, 1996.

[36] Etherton Melanie. *Charged Device Model (CDM) ESD in ICs : physics, modeling, and circuit simulation*. Technische wissenschaften, Eidgenössische Technische Hochschule ETH Zürich, 2006.

[37] Bossard P.R., Chemelli R.G, and Unger B.A. ESD damage from triboelectrically charged IC pins. In *EOS/ESD Symposium*, pages 17–22, 1980.

[38] Renninger R.G., Jon M.-C., Lin D.L., Diep T., and Welsher T.L. A field-induced charged-device model simulator. In *EOS/ESD Symposium*, pages 59–71, 1989.

[39] JEDEC Solid State Technology Association. Field-induced charged device model test method for electrostatic discharge withstand thresholds for microelectronic components, 2000.

[40] Working Group WG 5.3.1 ESD Association. ESD standard test method for electrostatic discharge sensitivity testing - charged device model (CDM) component level, 1999.

[41] Automotive Electronics Council (AEC). Charged Device Model (CDM) Electrostatic Discharge Test. AEC - Q100-011.

[42] Electronic Industries Association of Japan (EIAJ). Environmental and endurance test methods for semiconductor devices, (CDM).

[43] Brodbeck T. and Kagerer A. Influence of the device package on the results of CDM tests - consequences for tester characterization and test procedure. In *EOS/ESD Symposium*, pages 320–327, 1998.

[44] Brodbeck T., Esmark K., and Stadler W. Reproducibility of field failures by ESD models - comparison of HBM, socketed CDM and non-socketed CDM. In *EOS/ESD Symposium*, pages 1–8, 2007.

[45] Chaine M., Verhaege K., Avery L., Kelly M., Gieser H., Bock K., Henry L.G., Meuse T., Brodbeck T., and Barth J. Investigation into socketed CDM (SDM) tester parasitics. In *EOS/ESD Symposium*, 1998.

[46] Gossner H. and Brodbeck T. Dangerous parasitics of socketed CDM ESD testers. In *EOS/ESD Symposium*, 1997.

[47] Wolf H., Gieser H., Stadler W., and Wilkening W. Capacitively coupled transmission line pulsing cc-TLP - a traceable and reproducible stress method in the CDM-domain. In *Microelectronics Reliability*, volume 45, pages 275–285, 2005.

[48] Wolf H., Gieser H., and Walter D. Investigating the CDM susceptibility of IC's at package and wafer level by capacitive coupled TLP. In *EOS/ESD Symposium*, pages 297–303, 2007.

[49] Stadler W. and Brodbeck T. Basics of ESD on systemlevel. Internal IFX (COM BTS LIB) PhD seminar, 2005.

[50] USB – ESD – Problem bei manchen Pentium – 4 – Mainboards [in german, USB ESD problem in some pentium 4 main boards]. *C"t, Magazin für Computer Technik*, December 2005.

[51] Stadler W., Brodbeck T., Gärtner R., and Gossner H. Cable discharges into communication interfaces. In *EOS/ESD Symposium*, pages 144–151, 2006.

[52] Olney A., Gifford B., Guravage J., and Righter A. Real-world charged board model (CBM) failures. In *EOS/ESD Symposium*, 2003.

[53] Tamminen P. and Viheriäkoski T. Characterization of esd risks in an assembly process by using component-level cdm withstand voltage. In *EOS/ESD Symposium Proceedings*, pages 202–211, 2007.

[54] Maloney T.J. and Khurana N. Transmission line pulsing techniques for circuits modeling of ESD phenomena. In *EOS/ESD Symposium*, pages 49–54, 1985.

[55] Grund Evan. Tutorial on TLP theory and advanced TLP subjects. In *International Electrostatic Discharge Workshop*, pages 63–104, 2007.

[56] Voldman S.H., Ashton R., Barth J., Bennett D., Bernier J., Chaine M., Daughton J., Grund E., Farris M., Gieser H., Henry L.G., Hopkins M., Hyatt H., Natarajan M.I., Juliano P., Maloney T.J., McCaffrey B., Ting L., and Worley E. Standardization of the transmission

line pulse (TLP) methodology for electrostatic discharge (ESD). In *EOS/ESD Symposium*, 2003.

[57] Stadler W., Esmark K., Reynders K., Zubeidat M., Graf M., Wilkening W., Willemen J., Qu N., Mettler S., Etherton M., Nuernbergk D., Wolf H., Gieser H., Soppa W., De Heyn V., Natarajan M., Groeseneken G., Morena E., Stella R., Andreini A., Litzenberg M., Pogany D., Gornik E., Foss C., Konrad A., and Frank M. Test circuits for fast and reliable assessment of CDM robustness of I/O stages. In *Microelectronics Reliability*, volume 45, pages 269–277, 2005.

[58] Pogany D., Byschikhin S., Fürböck C., Litzenberger M., Gornik E., Gross G., Esmark K., and Stecher M. Quantitative internal thermal energy mapping of semiconductor devices under short current stress using backside laser interferometry. 49:2070, 2002.

[59] Kitamura Y. and Kitamura H. Breakdown of thin gate-oxide by application of nanosecond pulse as ESD test. In *International Symposium on Testing and Failure Analysis*, pages 193–199, 1989.

[60] Vinson J.E. and Liou J.J. Electrostatic discharge in semiconductor devices: An overview. In *proceedings of the IEEE*, volume 86, pages 399–417, 1998.

[61] Kelly M. A comparison of electrostatic discharge models and failure signatures for CMOS integrated circuit devices. In *EOS/ESD Symposium*, pages 95–175, 1995.

[62] Tung C.H., Cheng C.K., Radhakrishnan M.K., and Natarajan M.I. Physical failure analysis to distinguish EOS and ESD failures. In *International Symposium on the Physical and Failure Analysis of Integrated Circuits*, pages 65–69, 2002.

[63] Brodbeck T., Bauch H., Gugenmos X., and Wagner R. CDM tests on interface test chips for the verification of ESD protection concepts. In *Microelectronics Reliability*, volume 36, pages 1719–1722, 1996.

[64] Ito C. and Loh W. A new mechanism for core device failure during CDM ESD events. In *EOS/ESD Symposium*, 2006.

[65] Tang L.J, Pey K.L, Tung C.H, Rahakrishnan M.K, and Lin W.H. Gate dielectric-breakdown-induced microstructural damage in NMOSFETs. In *Transactions on Device and materials Reliability*, pages 38–45, 2004.

[66] Chaine M. Unique ESD failure mechanisms during negative to Vcc HBM tests. In *EOS/ESD Symposium*, pages 346–355, 1997.

[67] Amerasekera A., Gupta V., Vasanth K., and Ramaswamy S. Analysis of snapback behavior on the esd capability of sub 0.2 um NMOS. In *International Reliability Physics Symposium*, pages 159–166, 1999.

[68] Salman A., Gauthier R., Stadler W., Esmark K, Muhammad M., Putnam C., and Ioannou D. Characterization and investigation of the interaction between hot electron and electrostatic discharge stress using NMOS devices in 0.13 ţm CMOS technology. In *International Reliability Physics Symposium*, pages 219–225, 2001.

[69] Salman A., Gauthier R., Wu E., Riess P., Putnam C., Muhammad M., Woo M., and Ioannou D. Electrostatic discharge induced oxide breakdown characterization in a 0.1ţm CMOS technology. In *International Reliability Physics Symposium*, pages 170–174, 2002.

[70] Troutman R.R. *Latch-up in CMOS Technology*. Kluwer Academic,New York, 1986.

[71] Domanski K., Heer M., Esmark K., Pogany D., Stadler W., and Gornik E. External (transient) latchup phenomenon investigated by optical mapping (TIM) technique. In *EOS/ESD Syposium Proceedings*, pages 347–356, 2007.

[72] Polgreen T. and Chaterjee A. Improving the ESD failure threshold of silicided n- MOS output transistors by ensuring uniform current flow. *Transactions on Electron Device*, 39(2), 1992.

[73] Voldman Steven H. *ESD Circuits and devices*. John Wiley & Sons, LTD, Chichester, England, 2006.

[74] Beebe Stephen G. *Characterization, modeling, and design of ESD protection circuits*. PhD thesis, Stanford University, 1998.

[75] Gossner H. ESD protection for the deep sub micron regime - a challenge for design methodology. In *Conference on VLSI Design*, pages 809–818, 2004.

[76] Torres C., Miller J., Stockinger M., Akers M., Khazhinksy M., and Weldon J. Modular, portable and easily simulated ESD protection networks for advanced CMOS technologies. In *EOS/ESD Symposium*, pages 82–95, 2001.

[77] Li J., Gauthier R., and Rosenbaum E. A compact, timed-shutoff, MOSFET based power clamp for on-chip ESD protection. In *EOS/ESD Symposium*, pages 273–279, 2004.

[78] Maloney T., Poon S., and Clark L. Methods for designing low leakage power supply clamps. In *EOS/ESD Symposium*, pages 27–33, 2003.

[79] Smith J.and Boselli G. A MOSFET power supply clamp with feedback enhanced triggering for ESD protection in advanced CMOS processes. In *EOS/ESD Symposium*, pages 8–16, 2003.

[80] Caillard Benjamin. *Le Thyristor Parasite en technologie CMOS : Application à la Protection contre les Décharges Electrostatiques*. PhD thesis, Université Montpellier II, 2003.

[81] Caillard B., Azais F., Nouet P., Dournelle S., and Salomé P. STMSCR: A new multi-finger SCR-based protection structure against ESD. In *EOS/ESD Symposium*, pages 223–241, 2003.

[82] Azais F., Caillard B., Dournelle S., and Salomé P.and Nouet P. A new Multi-Finger SCR-based structure for Efficient On-Chip ESD Protection. *Microelectronics Reliability*, 45:233–243, 2005.

[83] Smedes T. and Guitard N. Harmful voltage overshoots due to turn-on behaviour of ESD protections during fast transients. In *EOS/ESD Symposium*, pages 366–375, 2007.

[84] Amerasekera A. and Duvvury C. The impact of technology scaling on ESD robustness and protection circuit design. *Transactions on Components, Packaging, and Manufacturing Technology*, 18(2), 1995.

[85] Boselli G., Rodriguez J., Duvvury C, and Smith J. Analysis of ESD protection components in 65nm CMOS technology: Scaling perspective and impact on ESD design window. In *EOS/ESD Symposium*, 2005.

[86] Duvvury Charvaka. Challenges of soc ESD design. In *International Electrostatic Discharge Workshop*, pages 1–29, 2007.

[87] Zängl F, Gossner H., Esmark K., Owen R., and Zimmermann G. Case study of a technology transfer causing ESD problems. In *Microelectronics Reliability*, volume 42, pages 1275–1280, 2002.

[88] Glaser Ulrich. *Complex ESD Protection Elements an Issues in Decananometre CMOS Technologies*. PhD thesis, 2007.

[89] Soldner W., Streibl M., Hodel U., Tiebout M., Gossner G., Schmitt-Landsiedel D., Chun J.H., Ito C., and Dutton R.W. RF ESD protection strategies: Co-design vs. low-c protection. In *EOS/ESD Symposium*, pages 8–16, 2005.

[90] Langguth G., Gossmann T., Rauch S., Kreppold B., and Wendel M. A self protecting RF Output with 2 kV HBM hardness. In *EOS/ESD Symposium*, pages 9–18, 2007.

[91] Kim M.-J., Langguth G., Esmark K., Gossner H., and Lee T. High-voltage tolerant fail-safe ESD protection for USB 2.0 compliant I/O circuits. In *International Electrostatic Discharge Workshop*, pages 289–299, 2007.

[92] Drüen Stephan. *Virtual ESD Test – An ESD Analysis Methodology at Chip Level*. PhD thesis, 2007.

[93] Stockinger M. and Miller J. Advance ESD Rail Clamp Network Design for High Voltage CMOS Applications. In *EOS/ESD Symposium*, page xxx, 2004.

[94] Russ C., Mergens M., Verhaege K., Armer J., Jozwiak P., Kolluri G., and Avery L. GGSCRs: GGNMOS triggered silicon controlled rectifiers for ESD protection in deep sub-micron CMOS process. In *EOS/ESD Symposium*, 2001.

[95] Streibl M., Esmark K., Sieck A., Stadler W., Wendel M., Szatkowski J., and Gossner H. Harnessing the base-pushout effect for ESD protection in bipolar and BiCMOS technologies. In *EOS/ESD Symposium*, 2002.

[96] Mergens M., Armer J., Jozwiak P., Keppens B., De Ranter F., Verhaege K., and Kumar R. Active-source-pump (ASP) technique for ESD design window expansion and ultra-thin gate oxide protection in sub-90nm technologies. In *Custom Integrated Circuits Conference*, pages 251 – 254, 2004.

[97] Okushima Mototsugu. ESD protection for mixed-power domains in 90 nm CMOS with new efficient power clamp and GND current trigger (GCT) technique. In *EOS/ESD Symposium*, pages 205–213, 2006.

[98] Linten D., Thijs S., Scholz M., Sawada M. Trémouilles D., Nakaei T., Hasebe T., and Groeseneken G. Characterization and modeling of diodes in sub-45 nm CMOS technologies under HBM stress conditions. In *EOS/ESD Symposium*, pages 158–164, 2007.

[99] Industry Council on ESD Target Levels. White Paper 2, 2008, released planed for April.

Chapter 3

Experimental set-up and test structures

3.1 CMOS technologies investigated

The advanced CMOS technologies investigated during this PhD were mainly from the 0.13 µm, 90 nm, 65 nm and 45 nm gate length nodes. Nevertheless, in order to access a wider range of oxides thicknesses, longer gate length CMOS technologies (1.3 µm, 0.5 µm, 0.35 µm, 0.25 µm, 0.18 µm) have been also characterized. The gate oxide dielectric used in all of these CMOS processes is silicon dioxide (SiO_2). Concerning ultra-thin oxides thinner than 3 nm, gate dielectrics are processed with oxynitrides (nitrided silicon dioxide) to improve the leakage behavior and device reliability. The principal technologies characteristics relevant for ESD-oxide concerns are exposed in Table 3.1. The strong decrease in the oxide thickness coming along with the multiplication of the power supply domains and the uncorrelated decrease of the supply voltage with the decrease of the thickness as discussed in the challenges that ESD designers should face (Section 2.6.1) are clearly observable.

Technology node (L_g)	Oxide thickness (T_{ox})	Supply (V_{dd})	Isolation type	Salicide
1.3 µm	27.5 nm	5.5 V, 2.7 V	LOCOS	no
0.5 µm	16 nm	5.5 V	LOCOS	no
0.35 µm	8.7 nm	5.5 V, 3.3 V	LOCOS	yes, SLBK* enable
0.25 µm	7.2 nm 4.7 nm	5 V 3.3 V	STI	yes SLBK* enable
0.18 µm	6.8 nm 3.6 nm	3.3 V, 2.5 V 1.8 V	STI	yes SLBK* enable
0.13 µm	5.2 nm 2.2 nm	3.3 V, 2.5 V 1.5 V, 1.2 V	STI	yes SLBK* enable
90 nm	5.2 nm 2.2 nm 1.5 nm	2.5 V 1.5 V 1.2 V	STI	yes SLBK* enable
65 nm	5.2 nm 2.8 nm, 2.2 nm 1.8 nm, 1.25 nm	2.5 V 1.8 V, 1.5 V 1.2 V, 1 V	STI	yes SLBK* enable
45 nm	2.8 nm, 2.6 nm 1.8 nm, 1.15 nm	1.8 V, 1.5 V 1.2 V, 1 V	STI	yes SLBK* enable

Table 3.1: Technologies characteristics investigated in the thesis. SLBK*: gate to diffusion salicide blocked during the process to enable ballasting (see Section 2.5.4.2).

3.2 Gate oxide characterization down to the nanosecond regime

3.2.1 Stress configuration

As a general stress configuration to investigate gate oxide dielectric time dependent breakdown or degradation induced by ESD-like pulses, the devices were connected as capacitors; the drain, source and bulk terminals of the transistors are commonly grounded and the gate connection is forced with a stress bias. This stress configuration ensures an uniform stress across the whole gate area and simulates the sensitive ESD stress case occurring on the gate of an input stage or on a capacitor connected between the supply and the ground lines. The stress is either qualified as inversion or accumulation accordingly to the state of the device under stress, see Table 3.2.

Device type	Gate bias	
	$V_g > 0$	$V_g < 0$
nFET	inversion	accumulation
pFET	accumulation	inversion
nCAP *	accumulation	accumulation

Table 3.2: Stress mode as a function of the device type and gate bias.

*: a nCAP device is constituted of a nFET device in a *n* well, therefore there is no possibility to generate an inversion layer channel from the diffusions.

3.2.2 Test structures

3.2.2.1 TDDB test structures

A dedicated layout for ultra-thin oxides TDDB investigation is required. The test structures used were FET-arrays of small active GOX area. The small oxide area is an important parameter to observe intrinsic breakdowns and to prevent from high gate leakage current during the constant voltage stress (CVS) due to direct tunneling currents. The test structures were optimized in terms of low resistance path from the pads to the gate, drain, source, and well terminals of the FET-array. Each gate was contacted to a pad, all source and bulk contacts were tied together to one common terminal and one pad was dedicated to all drain contacts. Tie-down protection diodes were connected from gate to bulk in order to avoid plasma charging and oxide pre-damage. In case of accumulation stresses, tie-down diodes were removed or the use of single devices was enforced.

Two-terminals (gate - source, drain, bulk) and 4-terminals FETs have also been used under the condition that an appropriate layout without too much parasitics effects has been designed.

3.2.2.2 Oxide degradation test structures

4-terminals structures with variable geometries with separate drain, source, gate and bulk contacts were used to investigate the impact of degradation and breakdown states. During the stress, drain, source and bulk were connected together via an external hardware allowing to switch from a two-pole stress configuration into a 4-terminals accessible system for DC measurements.

3.2.3 Set-up

The set-up used for gate oxide characterization is based on one Agilent (8114A) pulse source unit which can generate pulse stresses from 20 ns to 1 s. The voltage waveforms are captured by a 1 GHz Tektronix oscilloscope TDS 7104, using high ohmic voltage probes HP 6139A. The current during the pulse is extracted from the voltage drop across a 100 Ω resistor as described in Figure 3.1. The voltage is sensed at the gate using an independent force and sense system (Kelvin technique) in order to avoid parasitic effects from the set-up and to increase the voltage sensitivity. Alternating to the stress, gate oxide leakage measurements of the DUT (in the capacitor configuration) were performed with a source meter unit Keithley 2400 connected to a relay matrix.

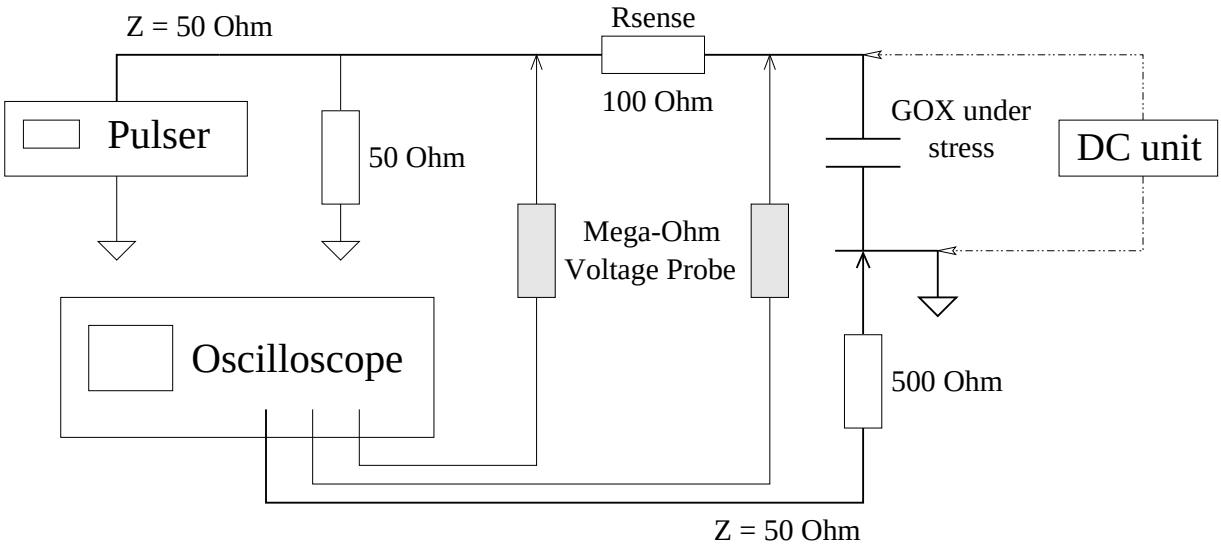

Figure 3.1: Set-up configuration used for the gate oxide time dependent dielectric breakdown experiments

The temperature of the wafer is settled via a thermo-chuck and the stress and leakage measurements sequence as well as the 200 mm wafer probe station Süss MicroTec are controlled via software. Pictures of the set-up are shown in Figures 3.2 and 3.3. Added to the basic GOX set-up discussed above, an other DC instrument is used to perform device characterization after GOX stresses phase. For this a Keithley 4200 system is used. As the capacitor stress configuration defined requires to shorten the drain, source and bulk during the pulse and as the DC

Figure 3.2: Set-up configuration used for the gate oxide time dependent dielectric breakdown experiments and DC characterization.

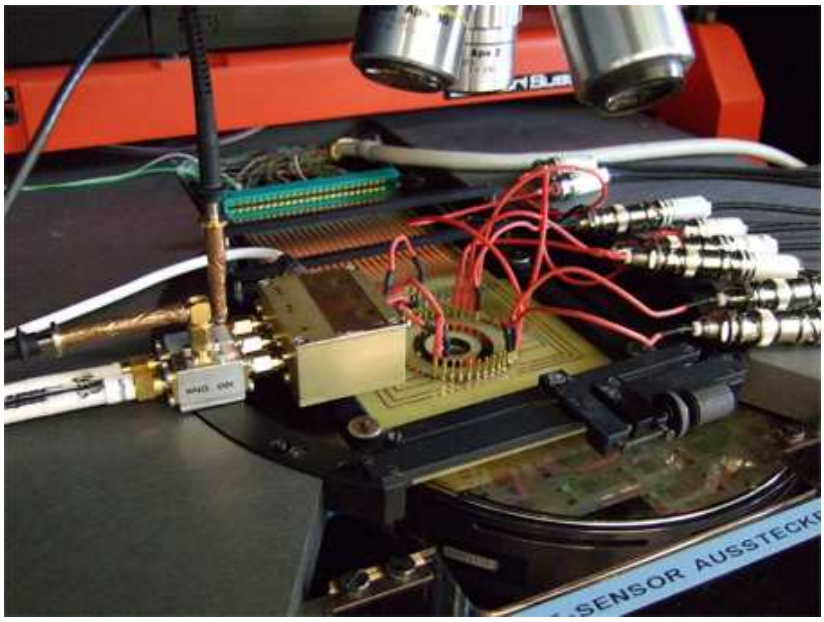

Figure 3.3: GOX standard pulser (box on the left) and external DC system (SMU on the right) connected to a 28 needles card for oxide stress / device characterization measurements.

characterization (I_{ds} - V_{ds}, I_{ds} - V_{gs}, G_m, ...) requires an independent pad access, a switching hardware composed of 50 Ω high-frequency relays is used to alternate from the DC to the stress configuration. DC measurements and ESD-like stress are both performed in a sense and force configuration. Eight relays were split in two switching boards for the permutation of each needle to the allocated SMU or stress configuration. However, the use of this switching hardware could

induce some ringing in the pulse waveform due to the added inductances from the cables and hardware system. Another method using one high-count probecard could be used to decouple both stress and DC connection by an automatical position stepping from one contact to the other (Figure 3.3).

In Figure 3.2, the standard GOX characterization set-up can be observed in the rack located on the left hand of the probe station. A close-up of the set-up connection is shown in Figure 3.3, the box on the left incorporates the hardware for the stress and DC monitoring. On the right hand side of the pictures, the external DC characterization system Keithley 4200 with the eight force and sense SMUs can be seen.

3.2.4 Stress methodologies

Basically, three different stress methodologies can be applied with this set-up, either long CVS (Constant Voltage Stress), or repetitive CVS, or a Voltage Ramp Stress (VRS) with a fixed pulse width and increasing voltage level, see Figure 3.4. This last method corresponds to the "classical" square pulsing procedure commonly used for ESD characterization. An *IV* sweep is applied on the fresh devices and two voltage values are defined for monitoring the gate leakage current. After each pulse the spot values are checked in order to detect the breakdown event. The breakdown event criterion is defined as an increase in the monitoring leakage current sensed in the direct tunneling region at $V_g = V_{dd}$. Actually, detection of the breakdown for thin oxides is relatively clear in the high voltage range. The energy dissipation during the breakdown is sufficient to melt the silicon through the gate percolation path which leads to a low resistive state and results in a drastic increase in the leakage current by several orders of magnitude. This breakdown event could be easily seen in the sensed voltage when the voltage drops drastically due to the deterioration of the electrical insulator properties.

For each stress methods the gate oxide time to failure determination was done according to the following definition.

- VRS: the threshold voltage is taken from the voltage average of the highest voltage pulse which does not cause any damage. The time to breakdown is taken as the whole effective pulse width.
- Long CVS: the methodology is done in a way that the pulse applied to the DUT is longer than the time that the gate oxide can withstands at the stressing voltage level. Consequently the hard breakdown event could be easily seen in the sensed voltage when the voltage drops drastically due to the deterioration of the electrical insulator properties. The effective time-to-fail was extracted from the pulse as described in Figure 3.5.
- Repetitive CVS: the time-to-failure is the sum of the total stress time of the pulses before detection of breakdown (number of pulses $\times$ pulse width) plus the part of the damaging pulse before breakdown.

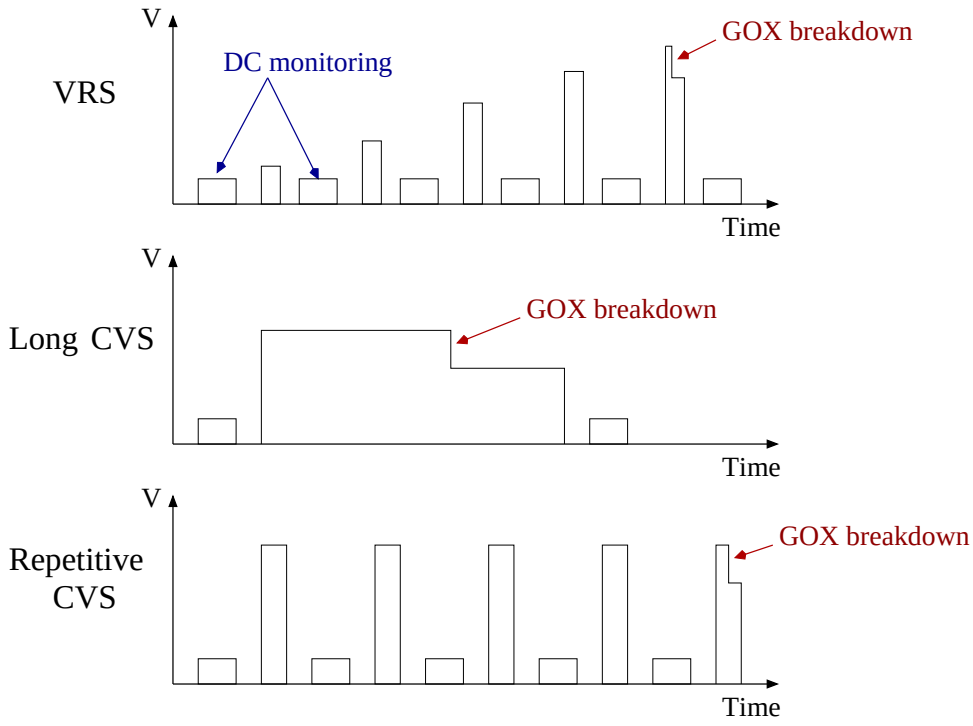

Figure 3.4: Gate oxide stress methodologies.

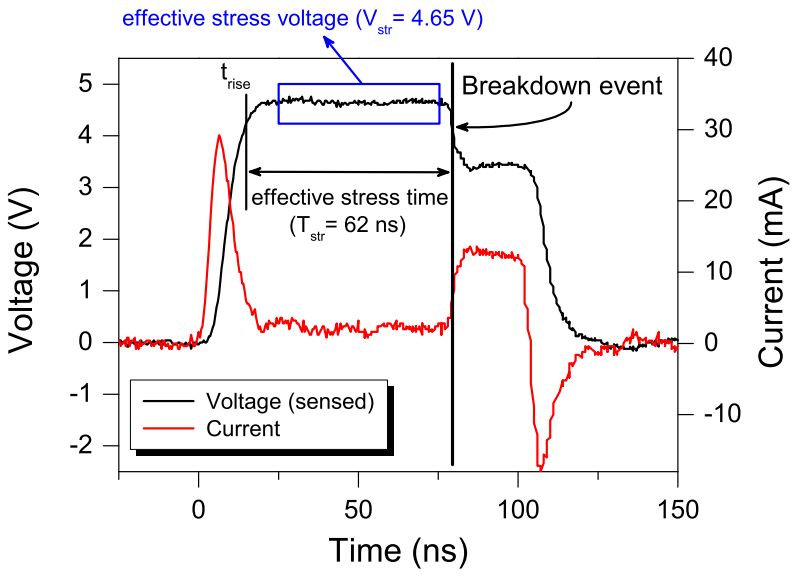

Figure 3.5: Pulse waveform for a 4.65 V stress, sensed at the gate. The detection of the breakdown event during the pulse is exposed for the determination of the effective time-to-breakdown ($T_{\mathrm{BD,eff}} = 62$ ns). The current spikes at the beginning and at the end of the pulse are artifacts due to reflections.

The rise time in the short pulse regime (< 1 µs) is not negligible and the real duration of the stress is less than the full pulse width. This effect is taken into account for the determination of the TDDB for all methods. The efficiency start of the stress was settle to the typical end of the rise time criterion, i.e. 90 % of the pulse plateau. The derived voltage value corresponds to the voltage sensed at the gate on which an averaging voltage level from the plateau was done on each sample stressed. The effective voltage stress level reference is derived from the median value of all single averaging pulse.

3.2.4.1 Automated stress and extraction

In order to investigate the gate oxide breakdown nature with large statistics, automatization of the stress measurement and consecutive extraction was mandatory. The measurements were performed using a Süss probe station with an automated stepping programm. The development of an extraction data programm was done in Perl to analyze the results. During the data extraction an averaging of the stressing voltage sensed across the DUT and the calculation of the effective time to breakdown accordingly to the previous definition was performed.

3.2.5 Gate leakage measurement

3.2.5.1 Current set up accuracy

One difficulty for GOX characterization under high frequency testing is to achieve good current accuracy sensing during the stress. Indeed, the TDDB structures used for monitoring oxide intrinsic fails are especially designed with small layout parameters to prevent from high tunneling current which would lead also to problematic detection of the SBD in the DC range and to parasitic voltage drop during high voltage stress. In standard TLP/pulser set-up used for ESD characterization, an entire 50 Ω system using a current probe CT1 is used. This probe is limited in frequency to the MHz range which is not flexible enough for long pulse stress up to the second time regime. The accuracy of the CT1 integrated in the set-up (also limited by the oscilloscope sensibility) is in the range of hundreds of micro amperes.

To enable studies on the pulse width dependence and to avoid data mismatch from set-up artifacts, measurement of the leakage via a resistor was chosen. With the GOX characterization set-up defined in Section 3.2.3, the gate leakage current accuracy sensed during the stress is basically limited by the sensing resistor and by the oscilloscope. Moreover, the value of this resistor is critical for the stress itself as a high value needed for a good current accuracy would totally degrade the square pulse waveform due to the *RC* circuitry provided by the MΩ voltage probes and the GOX capacitance. In the standard oxide characterization set up a value of 100 Ω which enable good short pulse quality done to the 20 ns regime and up to high voltages has been

fixed. However, the resolution from the set up does not permit to reach better current values than some hundreds of micro amperes.

3.2.5.2 High voltage gate leakage measurements

The determination of the gate leakage for high voltage is needed for several topics:

- the estimation of possible self-heating during an ESD stress.
- the evaluation of the charge-to-breakdown.
- the extraction of the internal parasitic resistance of the DUT which cause appreciable voltage drops in the extreme high tunneling current provided in the ESD voltage range.

This high voltage-current curve can be sensed using the GOX characterization set-up. Using a voltage ramp stress with a pulse width in the 50 to 200 ns range allows to reach far higher voltage than from standard DC measurements where the GOX breakdown is occurring rapidly. For high gate leakage characterization measurements, the value of the 100 Ω resistor used for gate oxide dielectric breakdown investigation (Figure 3.1) could be increase up to 10 kΩ to exclude scope averaging artifact and get a smooth curve for the matching with DC cures. An example of the parasitic resistance extraction using gate leakage tunneling equations and high gate leakage measurement is exposed in the Figure 3.6 below.

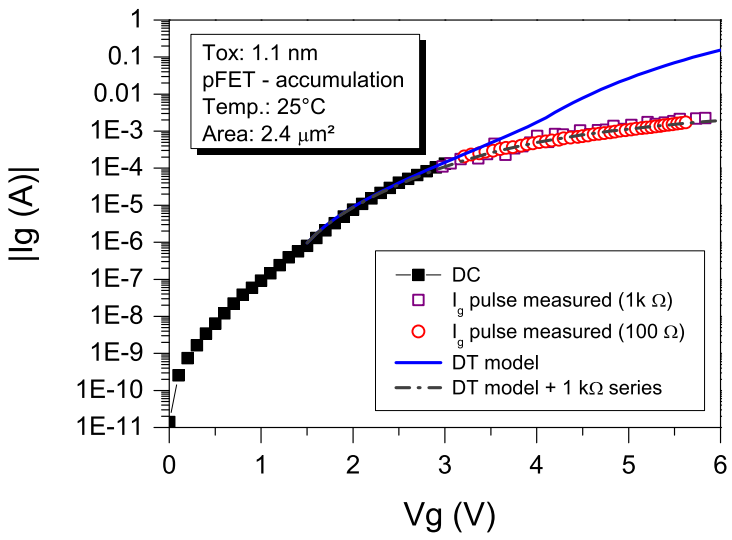

Figure 3.6: 1.1 nm pFET I_g - V_{gs} characteristic obtained from DC and short pulse measurement. The fitting of the leakage current reveals internal device voltage drops.

3.3 Conclusion

In this chapter, the experimental basis of the thesis for the characterization of the thin oxides degradation under ESD stress were detailed. The mastering and control of the set-up is mandatory to be able to distinguish accurately the main degradation from experimental artifacts. This is a hard task due to the high voltages applied down to 20 ns with fast rise times. We have developed an accurate Pulser set-up to adjust the precision of the degradation detection. This step is a mandatory for a correct extraction and modeling. Beyond the primordial importance of a good set-up, automatization efforts have been performed in order to obtain accurate statistical results. In the next chapter this last point will be underlined during the presentation of the experimental results.

Chapter 4

Thin oxides breakdown characterization and modeling down to the ESD regime

4.1 Gate Oxide Time-to-Breakdown modeling in the ESD regime: state of the art

In this section an historical review of the time-to-breakdown modeling of thin gate oxides in the direction towards the nanosecond regime will be given.

At the EOS/ESD symposium 1999, Kin P. Cheung has introduced the high importance of thin gate oxide reliability to the ESD community by means of his studies done on plasma charging investigation on 2.3 nm oxide devices [1]. Based on the Weibull statistic and on the percolation model, his message was clear: as the density of defect to trigger the breakdown is drastically reduced for thinner oxide and as the Weibull slopes are getting shallower, even very fast high level stresses (as plasma charging or ESD) could lead to oxide degradation. Based on the strong logarithmic trap generation rate dependence on the voltage reported by Stathis and DiMaria [2], a clear warning on the oxide trap generation during the transient triggering of the ESD protection elements (snap-back) was established [3, 4].

4.1.1 $1/E$ model TDDB behavior towards ESD time-scale

With the introduction of thinner oxides than 6 nm, preoccupation about the gate oxide breakdown modeling under ESD stress has grown. Before 2000, the model commonly used in the high-field regime was the $1/E$ model [5, 6]. This model was proposed for CDM extrapolation by Beebe in 1998 [7] after investigations done on a 0.25 µm SRAM process. In 2000, charge to breakdown from pulse-stress in the range of 200 ns to 1 ms and CDM results mismatch with

the $1/E$ model obtained by Leroux and co-workers on 7 nm oxide has been explained by current crowding and self-heating effects generated during the ESD event [8]. They suggested to include these two limiting phenomena in the analytical models. Contrarily to their findings, Jie Wu *et al.* [9] made the assumption that the power density in the oxide under ESD stress is too low to generate consequent self-heating. Moreover, for the fist time results of stress experiments over the wide range of 1 ns up to 10^4 s were performed on 2.2 nm, 3.7 nm and 4.7 nm pMOS capacitors using vf-TLP , TLP and solid-state pulse generator set-ups. Confirmed by further investigation in 2004 done on 3.5 nm pMOS with an optimized layout [10], it was stated that:

- the gate oxide time to breakdown is driven by the $1/E$ model (Figure 4.1).
- no prediction of gate oxide breakdown in the ESD regime as suggested by Cheung [4] could be done from DC measurements as the oxide trap generation was found to be dependent on the pulse-width stress.

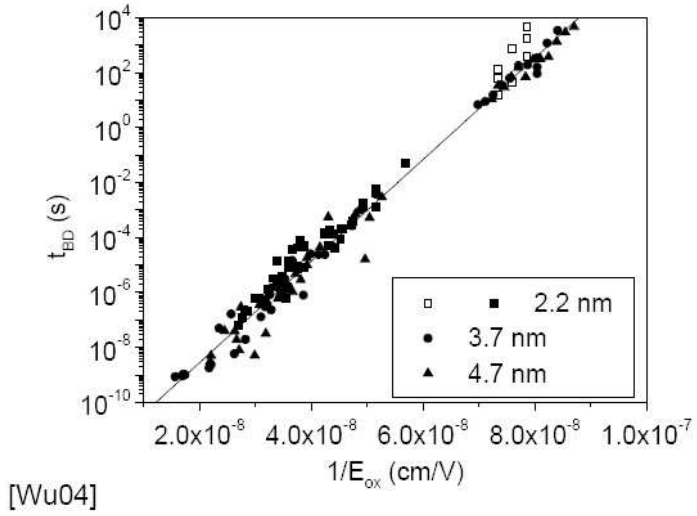

Figure 4.1: Gate oxide time to breakdown versus $1/E$ from Wu *et al.* [9, 10].

4.1.2 Oxide breakdown towards ESD regime accordingly to the AHI model

In 2003, Matsuzawa *et al.* [13] have linked the number of TLP pulses (100 ns) to breakdown with Anode Hole Injection simulation. Despite the good agreement found between the simulation and their 3 nm oxide data collected from 100 ns to 10 µs, the predicted model was not conclusive for their 6 nm data. They have claimed that breakdown dependence with the polarity of the gate bias could be explained by gate electrode depletion phenomena and that non uniformities of gate oxide breakdown probability should be considered due to transistor edge effects. However, a major point of gate oxide reliability has not been considered so far, the real importance of

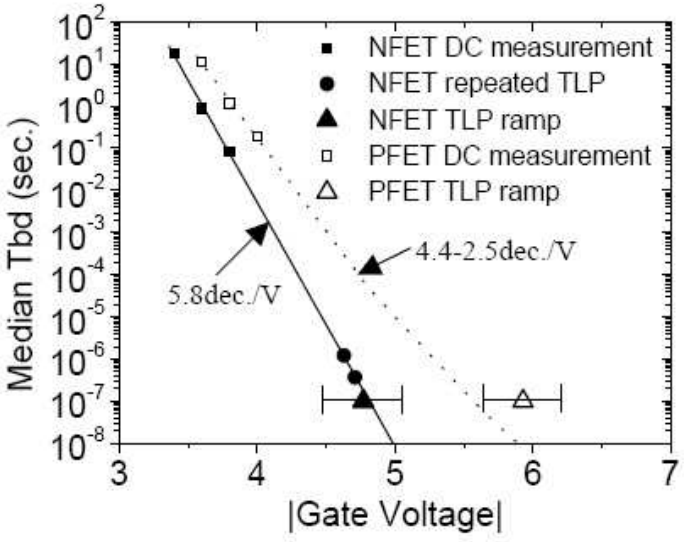

Figure 4.2: Gate oxide time to breakdown versus the gate voltage. The Anode Hole Injection model accordingly to the E-model derived is plotted from DC towards short pulse data from Weir [11, 12].

statistics to generate reliable data for thin oxides, in general few samples were used to cover the short time scale with the TLP methodology. Then in 2004, Bonnie Weir *et al.* [11, 12] have presented medium time reliability and short pulse data from 1.5 nm thin oxide using a better statistical methodology than all previous studies. They made the assumption of the AHI model exposed hereafter:

$$T_{\mathrm{bd}} = \frac{1}{J_{\mathrm{h}}} \left(\frac{N_{\mathrm{bd}}}{k} \right)^{1/m} \text{, where } J_{\mathrm{h}} = \alpha \, T_{\mathrm{h}} \, J_{\mathrm{e}} \tag{4.1}$$

k corresponds to the trap generation efficiency or the probability that a hole get trapped in the dielectric. m account for the dependence of the trap generation rate on the stress voltage. α is the probability of an electron-hole pair being generated and T_{h} is the transmission probability of holes.

- for $V_{\mathrm{g}} > 6V$, $\log(T_{\mathrm{bd}}) \sim 1/E$, because impact ionization is plainly occurring and α and T_{h} are equal to 1.
- for $4V < V_{\mathrm{g}} < 6V$, $\log(T_{\mathrm{bd}}) \sim V$, because in this regime the impact ionization is not dominant and both factors α and T_{h} depend on the field.
- for $V_{\mathrm{g}} < 4V$, the assumption that $T_{\mathrm{bd}} \sim V^{-n}$ is done based on the hypothesis that α becomes strongly dependent on the voltage.

Accordingly to this AHI model (E-model driven in the range of interest), they have also found that gate oxide breakdown could be extrapolated from the DC domain towards the ESD

time-scale, see Figure 4.2. The power law relation of the time to breakdown with the voltage was stated to be important in the direction to low voltages for lifetime extrapolation but was dismissed in the direction to nanosecond stresses. Last but not least they have claimed the hypothesis that the statistical reliability laws should still be valid in the ESD regime.

4.1.3 ESD TDDB modeling, state of the art summary

In the reported gate oxide reliability investigation under short pulses, one failure mode was observed: the Hard Breakdown. Two models have been proposed for the Time to Hard Breakdown of thin oxide (< 6 nm) in the ESD time-scale range:

- the $1/E$ model.
- the E -model.

4.1.4 Scope

The motivations for TDDB studies on the entire voltage range covering the operational domain, over-stress to ESD are evident with the aggressive reduction of the ESD design window for thin oxides. Investigations on TDDB modeling towards the ESD regime should be done to understand whether any extrapolation of oxide breakdown under ESD stress could be provided from the low voltage reliability model projections. The essential inherent question is, are the breakdown mechanisms occurring at low voltages still the ones which lead to the breakdown in the high Field domain?

In the previous works aiming to model the TDDB towards the ESD regime, no continuous set of data were provided in the wide voltage range staring from the long term reliability to the high voltage domains. The mix of experimental measurements and simulation to cover the unknown gap located between this two far domains is relatively hazardous. For an accurate TDDB modeling, assumption and manipulation of data to fit one model should be avoid as much as possible.

Contrarily to other works, a pure experimental approach towards the TDDB modeling is chosen in this thesis. A first part of the thesis is dedicated to an exhaustive investigation of thin oxides breakdown in the ESD regime in order to provide an accurate breakdown model based on highly reliable data. The choice of the set up and of the methodology to collect data in the ESD regime as well as the identification of the relevant parameters which have an impact on the oxide breakdown in the high Field domain are crucial and will be developed in the next Section 4.2. It is only on this solid base that a reliable modeling of the TDDB could be established from the empirical picture thereby obtained.

4.2 Statistical GOX breakdown modeling under CVS stress

Gate oxide time-to-breakdown of nFET under Constant Voltage Stress (CVS) have been performed with high statistics (45 samples at least) in the time range of 10 ms down to several ns. The source, bulk and drain of the structures under stress were connected together to the ground potential and a positive voltage pulse-stress was applied to the gate terminal. In this configuration, the state of the nFET under stress is the strong inversion ($V_g \gg V_{th}$).

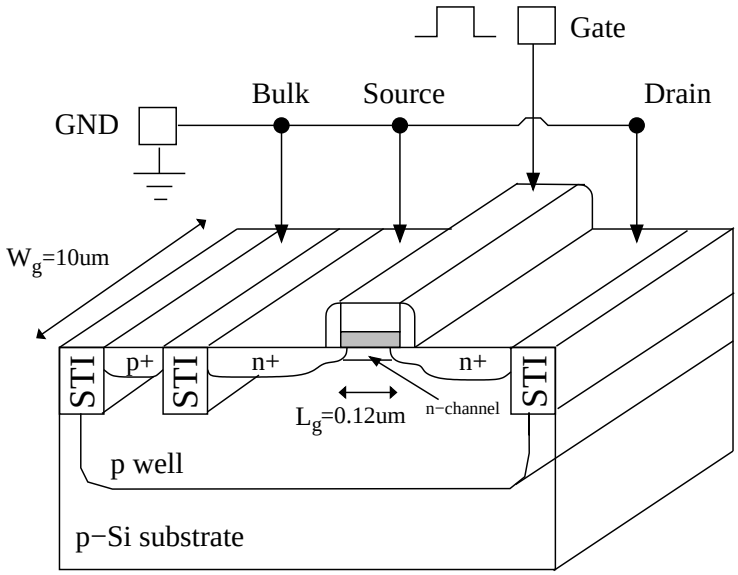

Figure 4.3: Schematic of the 1.5 nm nFET structures investigated under inversion CVS stress.

The structures investigated were *n* channel MOSFETs processed in a 90 nm gate length technology from the state-of-the-art CMOS process. The physical oxide thickness of the nFET determined from accumulation leakage measurement was 1.5 nm and the electrical gate oxide thickness in inversion regime is 1.6 nm. The test structures were designed with a 0.12 µm channel length and 10 µm width, resulting in an effective gate oxide area of 1.2 μm^2.

The stresses were generated by the set-up described in Section 3.2.3 and applied accordingly to the long Constant Voltage Stress methodology. Prior to the CVS, a calibration step is required to evaluate the adequate stress voltage levels and to ensure the condition that the pulse width chosen for the voltage level is higher than the time to fail of the oxide. This calibration step could be done rapidly with Voltage Ramp Stress for example. The typical breakdown observed in the long Constant Voltage Stress of 4 different voltage levels (4 V, 4.35 V, 4.65 V and 4.8 V) are exposed in the figures hereafter.

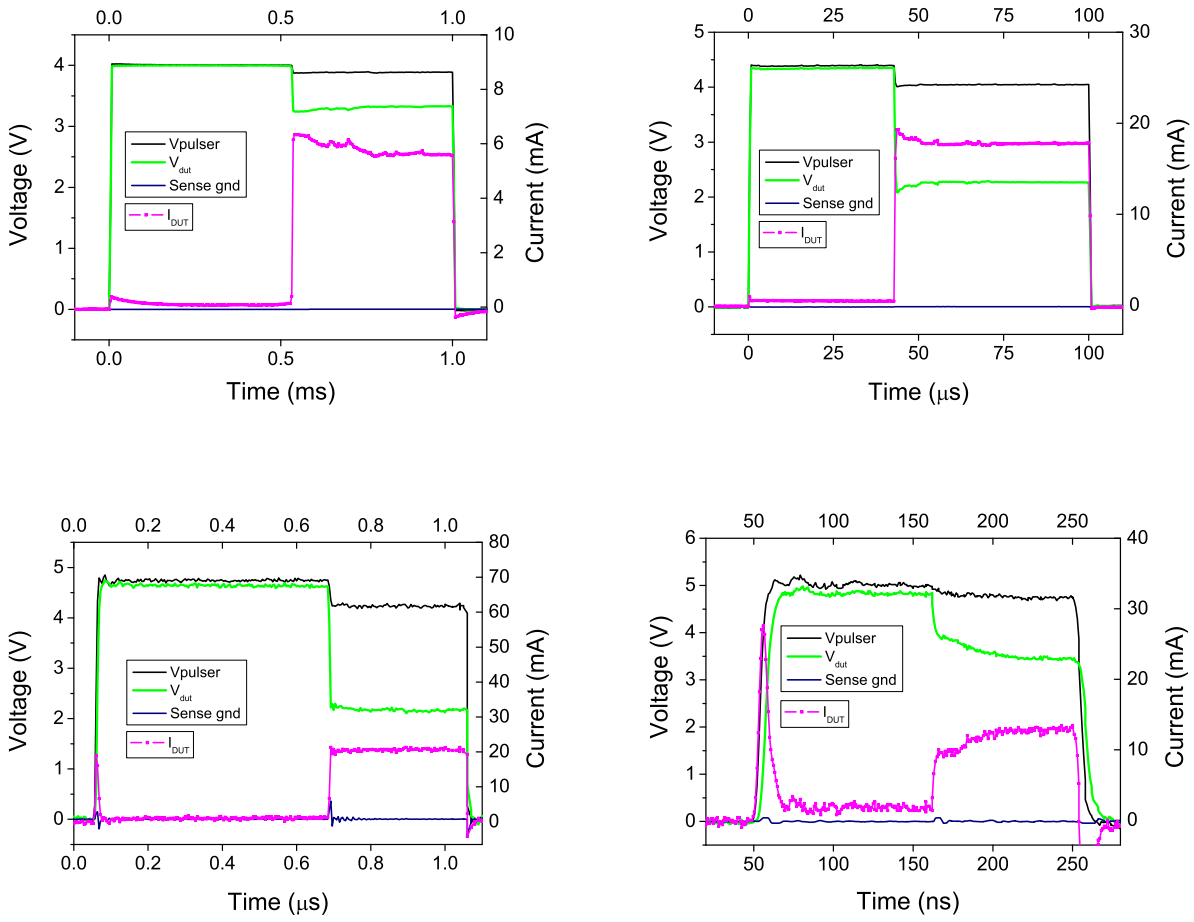

Figure 4.4: Pulses recorded from different long CVS levels. The voltage across the oxide corresponds to the voltage sensed at the gate terminal minus the sense ground signal and the current is given through a 100 Ω resistor placed after the pulse generated by the pulse generator (see 3.1).

The recorded pulses show a breakdown event, noticeable with a sharp decrease in the sensing voltage across the DUT. The breakdown event was also confirmed after each stress by a leakage failure criterion. An *IV* sweep on the fresh structure before the stress was applied and two voltage values were defined for monitoring the leakage current. In all cases, an important increase of several orders of magnitudes in the gate leakage current sensed in the direct tunneling regime at $V_g = V_{dd}$ was observed. As shown from the bottom right hand graph in Figure 4.4, under high level and short pulses, the effective stress applied should be properly evaluated. In this pulse time regime the set up capacitive parasitics, the oxide capacitance and the high tunneling current increase at high voltage stress, leading to a lower rise time caused by the set-up *RC* circuitry.

The distributions of the cumulative oxide failure level as a function of the time to breakdown resulting from the long CVS are plotted in a Weibull plot in Figure 4.5. The distributions show

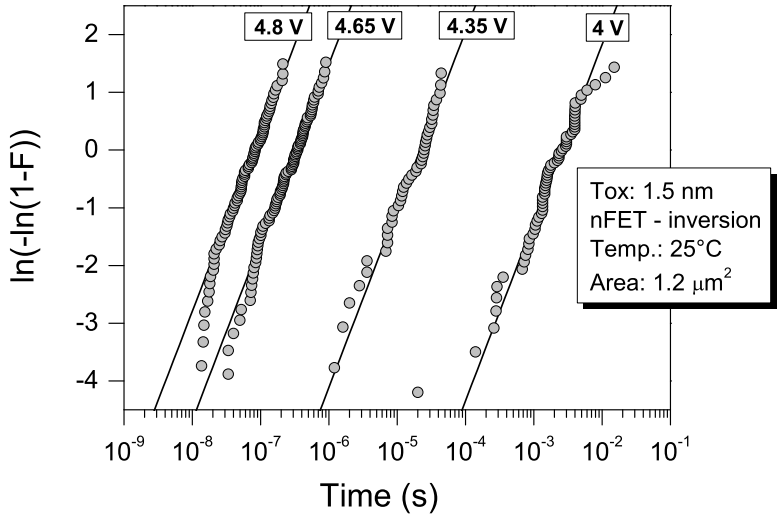

Figure 4.5: Cumulative oxide failure distribution obtained for 4 different long CVS level performed on 1.5 nm nFET devices stressed in inversion at room temperature

the same behavior down to the nanosecond regime and are well described by the Weibull statistics. The global breakdown picture resulting from a percolation path based on the weakest link event is valid until the ESD time regime. The shape of the Weibull distributions is found to be independent of the voltage level and of the stress time range, even in the short-time investigated. A common Weibull slope (β) value of 1.3 is very well fitting the distributions; this observation is especially evident from the very large statistics used for the 4.8 V and 4.65 V CVS, where more than 85 samples have been used. This value is in accordance with the percolation theory and is in really good agreement with published values [14, 15, 16]. The cumulative failure distributions obtained do not show any extrinsic fails which are characterized with a front tail having a slope below 1 (see section 1.2.5.2). The use of proper small test vehicles [17, 18, 19] has been an successful approach to avoid the monitoring of process-induced defects.

Independently from the device type, oxide thickness and polarity of the stress, the cumulative failure distributions of the devices under stress have been observed to follow the Weibull statistic in the ESD time domain over the entire technologies investigated in the thesis (table 3.1). An other example of cumulative failure distributions obtained on 2.2 nm pFET devices under accumulation ($V_g > 0$) is depicted in Figure 4.6.

Accordingly to the percolation picture, the reduction of the Weibull slope with the decreasing oxide thickness has been observed also in the nanosecond regime for oxide thickness in the range of 6.8 nm down to 1.1 nm. The average values of the Weibull slopes as a function of the oxide thickness are summarized in Figure 4.7 and compared to values extracted from the literature.

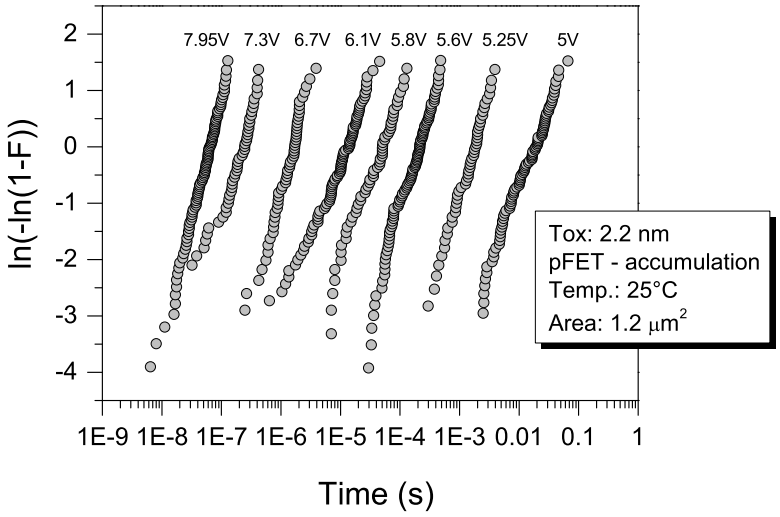

Figure 4.6: Cumulative oxide failure distribution obtained for 8 different long CVS level performed on 2.2 nm pFET devices stressed in accumulation at room temperature

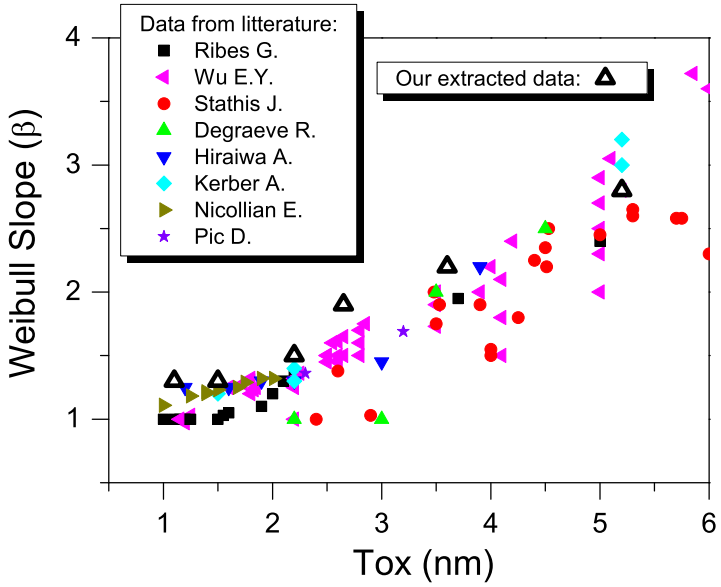

Figure 4.7: Measured and literature values of the Weibull slope versus the oxide thickness.

This trend rise a major topic for the thin GOX breakdown estimation, the use of large statistics is imperative. As depicted in Figure 4.7, the shape of the cumulative failure distribution (β) for thin oxides is getting shallower and failing samples spread widely over time decades. The

extraction of the time-to-fail from few samples can lead severe errors.

4.2.1 Area dependence

As the statistical nature of the gate oxide breakdown under electrical stress is seen to be similar in the short time range and in the DC or long reliability term, the question of the TDDB area dependence arises. This dependence is known to be a random statistical process occurring on the whole gate oxide area at low voltage and for long time stress [20, 21, 22] (Section 1.2.5.2). However, no time-to-fail dependence with the oxide area has been reported yet in the ESD time frame and it is not clear whether the wear-out of the oxide can still be described during an ESD stress by a random area occurrence.

To investigate the influence of the gate oxide area on the time-to-breakdown in the ESD range of time, long CVS tests were performed on small 1.5 nm nFET structures with different active area (0.12 μm^2, 0.36 μm^2 and 1.2 μm^2). The voltage stress level was set to 4.7 V in order to access a time-to-breakdown window in the hundreds of nanoseconds. The time-to-breakdown cumulative failure distributions are plotted in the Figure 4.8. Experimentally, the verification of the Poisson statistics does not involve any knowledge about the Weibull slope; the distributions from the 0.12 μm^2 and 0.36 μm^2 area, have been vertically shifted with equation 1.48. The good agreement obtained from the extrapolated $T_{63\%}$ from the different devices area with the $T_{63\%}$ value obtained from the distribution of reference, confirms the validity of Poisson statistic down to the nanoseconds and thus its application for the ESD GOX breakdown estimation.

However, the Weibull slope deduced from the measurements exposed in Figure 4.8 is around 1.7 which is a higher value as expected [24, 18]. The Weibull slope has been nevertheless confirmed by a CVS stress at 4.5 V performed in the microseconds regime as exposed in the Figure 4.9.

Basically, a deviation from Poisson scaling could be induced in presence of thickness or stress voltage level variation and could be possibly explained by insufficient statistics [22, 25]. This last reason has a clear impact, as the Weibull slope extracted from high statistical data in Figure 4.5 was significantly lower (β = 1.3). Bad design of test vehicle could also cause internal voltage drops [26]. Keeping the same voltage stress level exactly at one level becomes problematic in the direction to shorter time scales. The error bar relative to the stressing voltage applied on the different devices area is dependent on the accuracy of the set up at different stressing time window. A variation in the stress voltage is consequently leading to incorrect evaluation of the Weibull slope from the scaled area. Over estimation of the voltage stress in the high frequency domain due to artifact overshoots resulting from impedance mismatches [27] and internal voltage drops in large structures could result in a higher Weibull slope extraction from the area scaled devices under the same theoretical voltage level.

The experiment was reproduced in a 45 nm CMOS technology on ultrathin 1.1 nm oxide

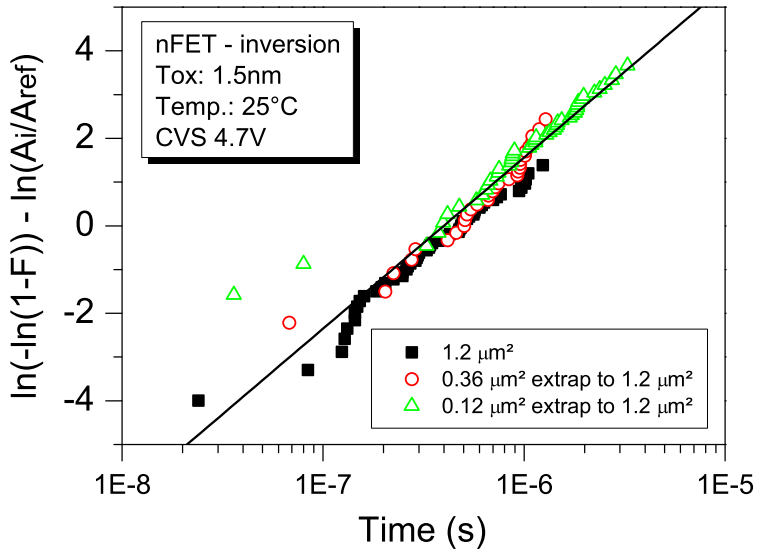

Figure 4.8: The time-to-breakdown distribution of 1.5 nm nFET devices at 25 °C from different active gate oxide area is plotted for CVS at 4.7 V. The normalized size is 1.2 µm^2 (black squares), the 0.12 µm^2 and 0.36 µm^2 TDB distributions are vertically shifted (open symbols) [23].

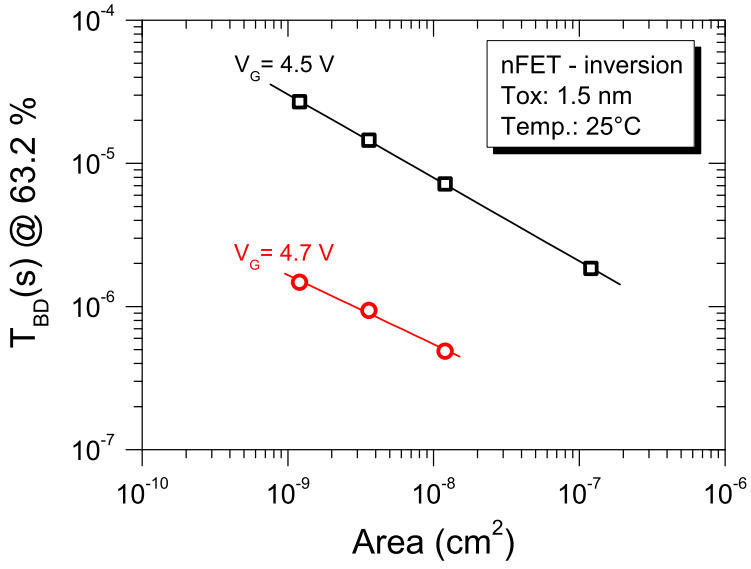

Figure 4.9: The extracted $T_{63\%}$ observed from different area at one CVS level are plotted as a function of the GOX area for 1.5 nm nFET devices stressed in inversion.

nFET devices in the same stress configuration. Once again a very good dependence of the time-to-breakdown with the area accordingly to the Poisson statistics is observed (Figures 4.10 and 4.11). Nevertheless, a slightly higher but fairly acceptable value of the Weibull slope (β = 1.35) is found in comparison to reported values in literatures [24, 15, 28]. Experimentally, the Weibull slopes reduction observed for oxide thinner than 2 nm is saturating around 1.1 to 1.3 as depicted in the Figure 4.7.

4.2.2 Methodologies confrontation

In the previous TDDB GOX studies under short-time pulses (< 1 µs) published in the literature, different stressing methodologies have been used:

- CVS was performed for the evaluation of the number of critical defects generated in the oxide under an ESD stress [4].
- Repetitive CVS has been used to quantify the oxide endurance robustness under repetitive number of stress [29], to investigate the number of pulse to breakdown (TDDB) [13, 30] or to evaluate the charge to breakdown [8].
- VRS has been the main used testing procedure used for oxide breakdown evaluation as a function of the pulse width (TDDB) [7, 11, 12, 9, 10, 31]. This methodology is generally used in ESD laboratory to determine the gate oxide robustness in the HBM (or CDM) regime with a TLP system (or vf-TLP) with a fixed pulse-width of 100 ns (or 5 to 2.5 ns).

Expect for few studies [4, 12], the evaluation of the gate oxide breakdown in the nanoseconds domain for one oxide thickness in one technology was extracted from few samples and often regardless of the active oxide area tested. As exposed in the two previous sections, the lack of statistics is critical for meaningful evaluations and the oxide area should be extrapolated in consequence. In the past, the GOX stressing methodology itself was totally disregarded and the question of the comparability of the results should be asked especially for thin oxides. However, satisfying the two previous criteria cited above, the GOX stressing procedure will be studied in this part.

A comparative investigation between the three methodologies presented in Section 3.2.4 has been carried out on nFET devices from an 1.5 nm oxide process. A positive stress was applied on the gate versus the other terminals grounded. To avoid any hardware variation, the comparative measurements have been done on the same wafer with high statistics. In this experiment the hard breakdown event criterion was defined as a 15 % increase in the monitoring leakage current sensed in the direct tunneling region at $V_g = V_{dd}$.

These different methodologies results in time-to-breakdown values in the same range at first order but differ at second order. For the establishment of the correct gate oxide time to fail, no

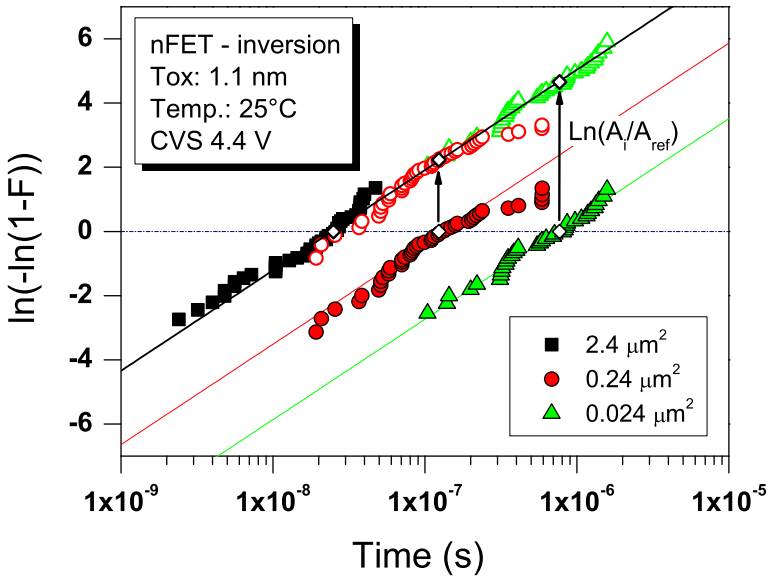

Figure 4.10: The time-to-breakdown distribution of 1.1 nm nFET devices at 25 °C from different active gate oxide area is plotted for CVS at 4.4 V.

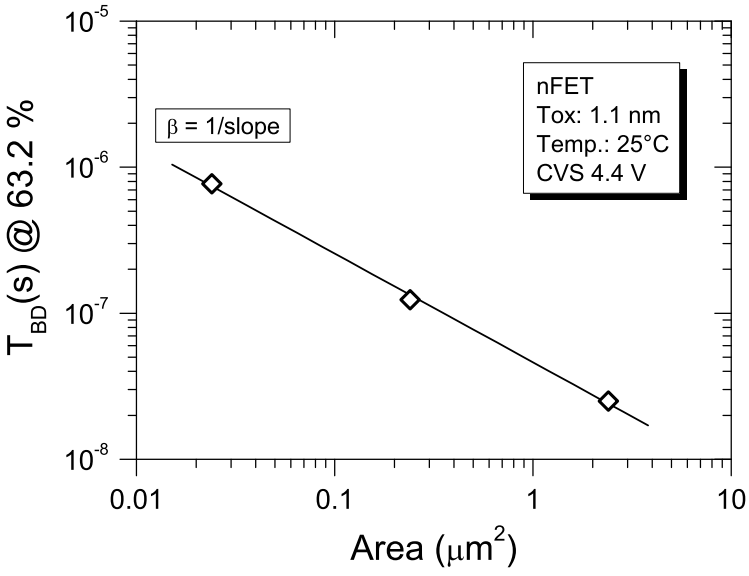

Figure 4.11: The extracted $T_{63\%}$ obtained from the 4.4 V CVS (Figure 4.10) are plotted as a function of the GOX area.

artifact from the methodology should be incorporated. Due to the cumulative nature of dielectric breakdown event (imposed by the percolation theory) the test methodologies need to be carefully investigated, as it will be discussed in the following.

1. **Repetitive CVS** This methodology implying the repetition of fixed pulse-width stress in a large number leads to a longer time-to-breakdown. A comparison between long CVS and repetitive CVS is shown in Figure 4.12.

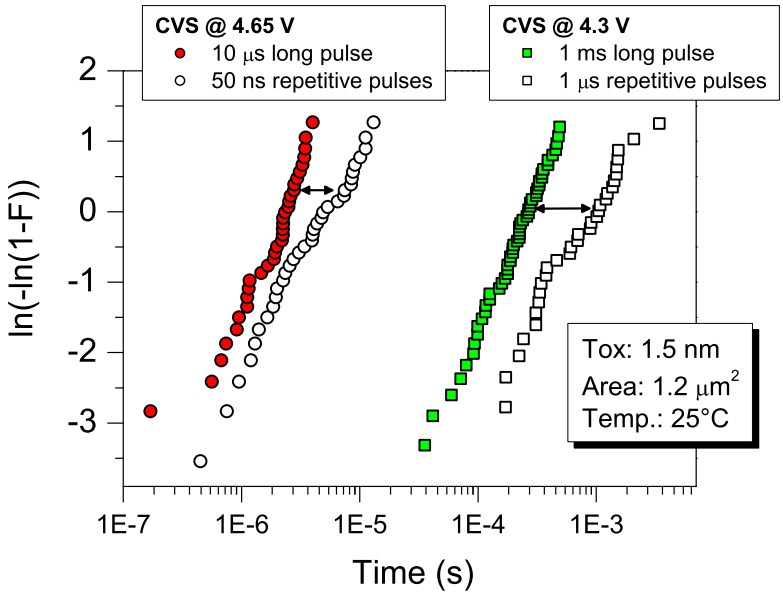

Figure 4.12: TBD distributions from long CVS and repetitive CVS (4.3 V & 4.64 V). A factor 2 in the time-to-fail is caused by the CVS stress methodology [23].

The degradation of the oxide from repetitive stresses is dependent on the defined pulse width [9], on the stressing frequency [32, 33] and on the stressing voltage. The evaluation of the GOX breakdown therefore does not lead to equivalent results and, thus, this methodology should be avoided. This methodology is directly revealing the cumulative effect of the GOX degradation mechanism trough the efficiency of short repetitive stresses to provoke GOX breakdown in a consistent statistical picture.

2. **VRS methodology** This procedure results in lower time to breakdown levels. The cumulative effect of this model causes an error in the evaluation of the gate oxide breakdown with the TLP-like increase voltage level stress methodology. The last previous pulses before the breakdown event are effectively counting in the total stress applied to the devices and thus the "TLP methodology" (VRS) is too critical for GOX breakdown evaluation.

   The error from this method can be calculated and breakdown voltage distributions can be converted into effective time-to-breakdown distributions. Considering the power-law

TDDB model, this transformation could be done with the following formula [34]:

$$t_{\text{eff}} = \sum_{i=1}^{N_{\text{str}}} \frac{t_{\text{mess}}}{N_{\text{str}}} \left(\frac{V(i)}{V_{\text{eff}}} \right)^n \tag{4.2}$$

$$\begin{aligned} \lim_{N_{\text{str}} \to \infty} t_{\text{eff}} &= \int_0^{t_{\text{mess}}} \mathrm{d}t \left(\frac{V(i)}{V_{\text{eff}}} \right)^n \text{ with, } \mathrm{d}t = \frac{t_{\text{mess}}}{N_{\text{str}}} \\ &= \frac{V_{\text{eff}}}{RR \cdot (n+1)} \left(\frac{V}{V_{\text{eff}}} \right)^{n+1} \text{ with, } V = V_0 + RR \cdot t_{\text{mess}} \end{aligned} \tag{4.3}$$

RR is the Ramp Rate corresponding to the voltage step chosen for the VRS measurement. V_{eff} is the effective voltage of the CVS and t_{mess} corresponds to the total stressing time, N_{str} is the number of stress applied and V_0 is the first voltage floor.

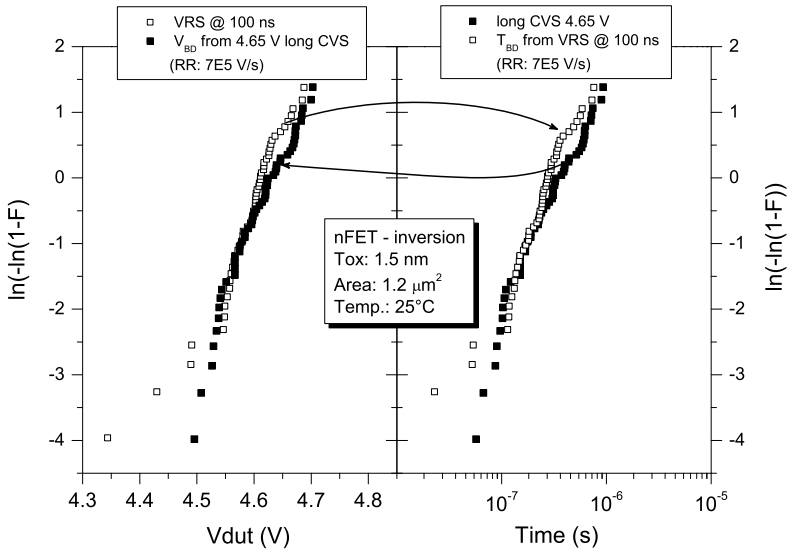

Figure 4.13: Transformation of a breakdown voltage distribution into a time-to-breakdown one and its reciprocal accordingly to the equation 4.3, exposed for long CVS at 4.65 V and VRS at 100 ns on 1.5 nm nFETs.

The validity of this formula in the ESD domain is shown in Figure 4.13 above with the good data correlation obtained from VRS and long CVS applied on 1.5 nm nFET devices. The conversion of an VRS into time-to-breakdown distributions depends on the voltage Ramp Rate (RR) which is hardly held constant on the DUT at such high voltages in the Fowler-Nordheim regime due to the logarithmic increase of the gate leakage current. In case of high DT or in the FN current regime, the voltage steps are decreasing due to the

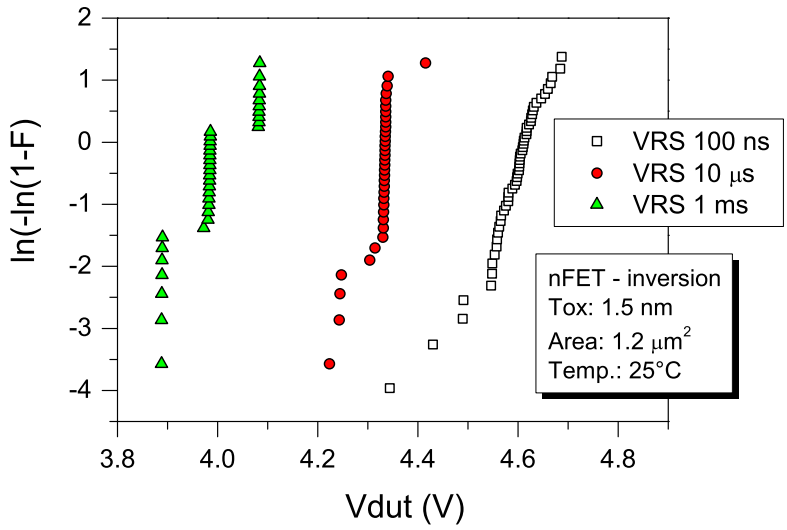

Figure 4.14: Breakdown voltage distributions of 1.5 nm nFET obtained from VRS for three different pulse-width, 100 ns, 10 µs and 1 ms

increase voltage drops resulting from the set-up and device parasitics resistances. The cumulativeness effect of the last stresses is enforced also by this effective modulation of the ramp rate. In Figure 4.13, to soften this effect, the effective ramp rate sensed on the DUT around the breakdown value (7.10^5 s/V) was used.

To investigate more precisely the difference in the time-to-breakdown obtained from the two different methodologies, VRS measurements at three different pulse widths (Figure 4.14) were performed. For their transformation into time-to-breakdown distributions, exposed in the Figure 4.15, the total stress time was considered as

$$t_{\text{mess}} = t_{\text{width}} \cdot (nb_{\text{str}} - 1) + t_{\text{eff}}(last)$$

where nb_{str} is the number of stress applied before the failure detection in the leakage current, t_{width} is the effective pulse width of the stresses and $t_{\text{eff}}(last)$ is the effective stress time in the last pulse.

The net difference between the long CVS method and VRS is shown in Figure 4.16. An error in the time-to-fail about a factor 5 is initially found. The transformation from the VRS distributions to TBD distributions allow an acceptable correlation but the voltage ramp rate introduces an other uncertainly level in comparison to direct TDDB estimation from CVS.

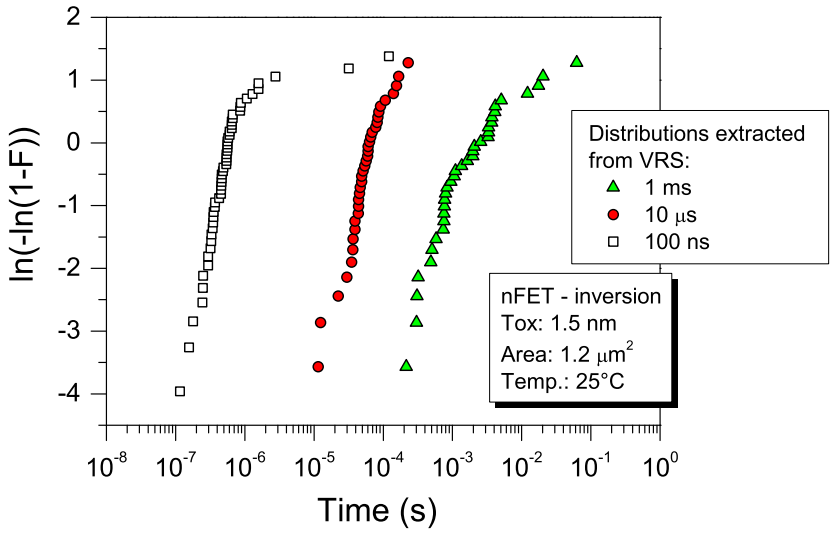

Figure 4.15: Time-to-breakdown distributions obtained from the transformation of the breakdown voltage distributions exposed in Figure 4.14 using the equation 4.3.

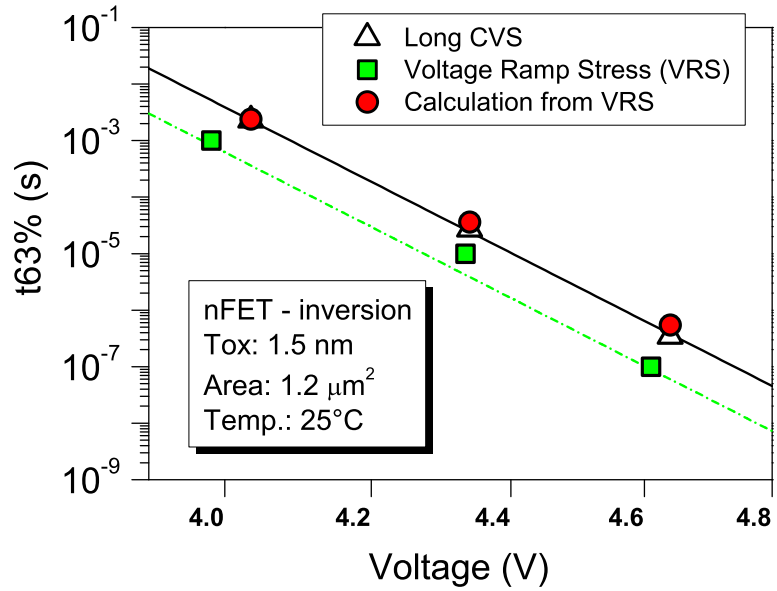

Figure 4.16: TDDB data at 25 °C obtained by long CVS and VRS. The $T_{63\%}$ from the transformed VRS distributions are also plotted (circles) [23].

3. **Long CVS** This method is closer to the stress methodology used for standard oxide reliability qualification which is using a logarithmic increase of the pulse stress width. It is also the closest stressing procedure with a real ESD event.

   This testing procedure present the interest to be only dependent on the stressing voltage and set-up sensitivity and to be de-correlate to other parameters at it is the case with the VRS or repetitive CVS (ramp rate, frequency, relaxation effects, ...).

Due to the non-homogeneity of the TDDB results obtained by different stressing methodology, the methodology used to characterize the GOX breakdown has to be adequately chosen and clearly specified. For dielectric hard breakdown monitoring, the long CVS procedure appears to be the proper methodology for the TDDB evaluation. This latter will be considered as the reference stressing methodology in the following of the thesis.

4.2.3 Test structures artefacts

A proper design and the use of small active GOX area are needed to observe intrinsic breakdowns [18]. In addition, long channel devices or poorly designed structures can give wrong TDDB results due to current crowding effects [35, 36, 26]. Artifacts and derived effects have to be known and integrated in the evaluation.

Big structures related reliable TDDB testing issues:

- lead to the monitoring of extrinsic defects.
- lead to problematic evaluation of the breakdown detection. The first breakdown event (progressive digital part or soft breakdown) could be missed due to the too low leakage level involved in this breakdown phase in comparison to the high tunneling current generated by such big structures [19].
- induce a strong loss of voltage level control across the oxide under stress. The voltage across the oxide is reduced due to internal voltage drop in the large structures. Under high field stress (ESD regime), the parasitic voltage drops are really significant due to the extreme high tunneling current.
- an other effect coming along with the internal parasitic voltage drop occurs, the non homogenous stressing on the whole gate oxide area. Current crowding operates and only some portions of the gate endured an effective high stress whereas other portions are stressed to lower levels [26, 8]. Design with large gate oxide area, long poly-Si lines, small and far bulk contacts, insufficient metal lines are heavily critical for the correct determination of the oxide breakdown and this furthermore under ESD stress conditions.

- cause also a large deviation of the stress pulse wave form and voltage level control from the applied stress. The high capacitance of large gox area plus the use of high Ohmic voltage probes provoke a considerable RC damping on the generated pulse. Under short and high stress levels the logarithmic domination of the gate leakage provokes a saturation in the maximal stressing voltage that the set up can provide at the DUT. Moreover the RC damping coefficient becomes to strong and deteriorates too much the pulse waveform.

A difference in the cumulative time-to-fail can be observed resulting from geometry and layout variation. TDDB structures need to be very small and optimized in terms of parasitic resistances, an area split in several cells permit also a better coupling. Long channels, big structures, far bulk contacts and narrow metal lines effects results into increased time-to-breakdowns and do not obey to the Poisson statistics anymore. In Figure 4.17, this impact is revealed for different W_g/L_g ratio and layout style. An over-estimated GOX time-to-fail by factor 5 can be induced from the test structures.

The effect of bad structures designs and stress methodology combined together result in a large error which is increasing towards the high current-voltage regime and short time domain.

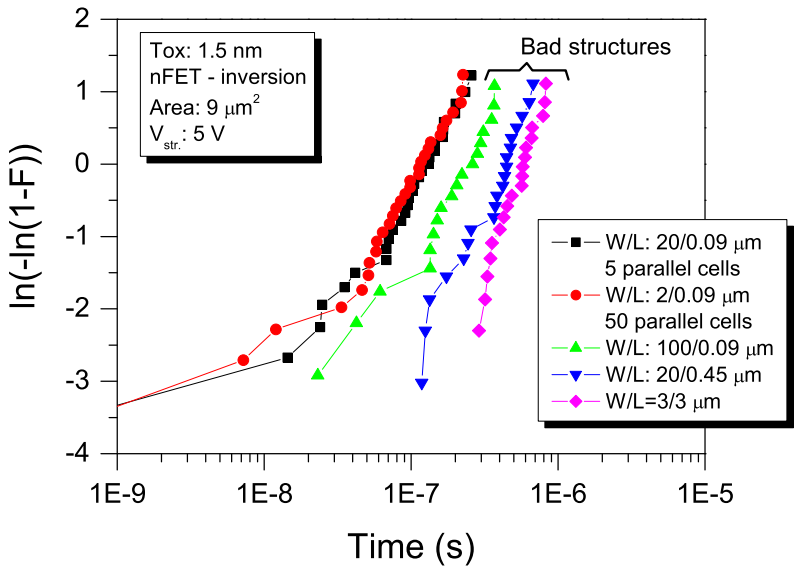

Figure 4.17: TDDB of 1.5 nm nFET devices stressed at 5 V. The impact of bad layout and geometry for the TDDB evaluation is exposed via poor test structures leading to over-estimated time to fail.

4.3 TDDB acceleration modeling towards ESD

The GOX breakdown voltage dependence plotted as a function of the time-to-breakdown is a crucial parameter for the ESD developers. From this model, GOX breakdown voltage in the HBM and CDM time domain and impacts of pre-stress and/or multi-stress can be derived. An extensive characterization work will be developed in this part aiming to provide a useful tool for ESD developers.

4.3.1 An universal TDDB law

4.3.1.1 Voltage acceleration at 140 °C.

The determination of the correct model from experimental data requires access of a wide range of voltages to avoid the use of a local fitting model which would fail to describe the real behavior of the oxide breakdown over the entire time domain. In order to accurately extract the GOX breakdown model in the ESD range, one set of experiment have been performed in the 1 s - 20 ns time range in the same condition than for the long reliability qualification tests procedure in order to observe the breakdown behavior over the whole range of time, covering the operational condition, DC, EOS and ESD. For lifetime projection, all qualifications set are done at high temperature (typically 125 °C to 140 °C) to integrate the real temperature of the chip in the operational state. The investigation done in the short-pulse time regime have been carried out at 140 °C on the same structures and from the same hardware as the experiment performed by Kerber *et al.* in the DC domain and towards the long reliability term [37].

The experiment has been performed on a 1.5 nm and 2.2 nm oxides CMOS process with the set up described in Chapter 3. Long CVS were applied to devices accordingly to the procedure described in the previous part and with the use of high statistics; in most cases between 45 and 90 samples were used per stress level. The devices were stressed as capacitors in inversion. The test structures were optimized in terms of low resistance path from the pads to the gate, drain, source, and well terminals of the FET-array. An example of the cumulative GOX failure distributions obtained in the range of seconds to nanoseconds at 140 °C on the 1.5 nm oxide are exposed in Figure 4.18 for nFET and Figure 4.19 for pFET [23].

The Time Dependent Dielectric Breakdown is extracted from the $T_{63\%}$ values out of the GOX time to fail distributions. For the Weibull characteristic parameter $T_{63\%}$ the data confidence bounds level is maximal [25] and the use of high statistics allow to reach reliable values. The time to breakdown extracted from Figure 4.18 and Figure 4.19 are plotted as a function of the stress voltage in Figures 4.20 and 4.21, respectively. The obtained values for the 2.2 nm oxide as well as the lower CVS stress performed in the reliability range [37] are also plotted in the same graphs. The TDDB from the active gate oxide area of the 2.2 nm oxide test structures were

normalized from their originate value (102 μm^2) to 1.2 μm^2 using the Poisson statistics (defined in Section 1.2.5.2).

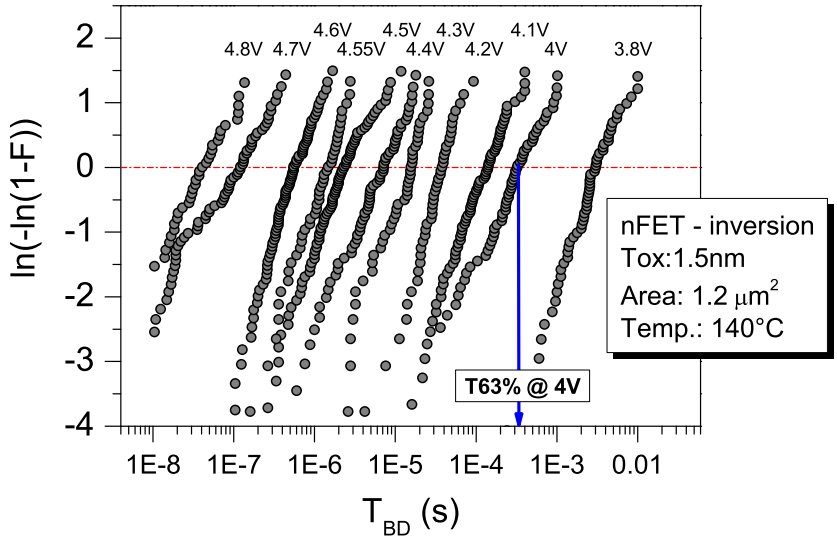

Figure 4.18: Gate oxide time to breakdown distribution of 1.5 nm nFET devices at 140 °C for CVS measurements from 3.8 V to 4.8 V using the set-up exposed in Section 3.2.3 [23].

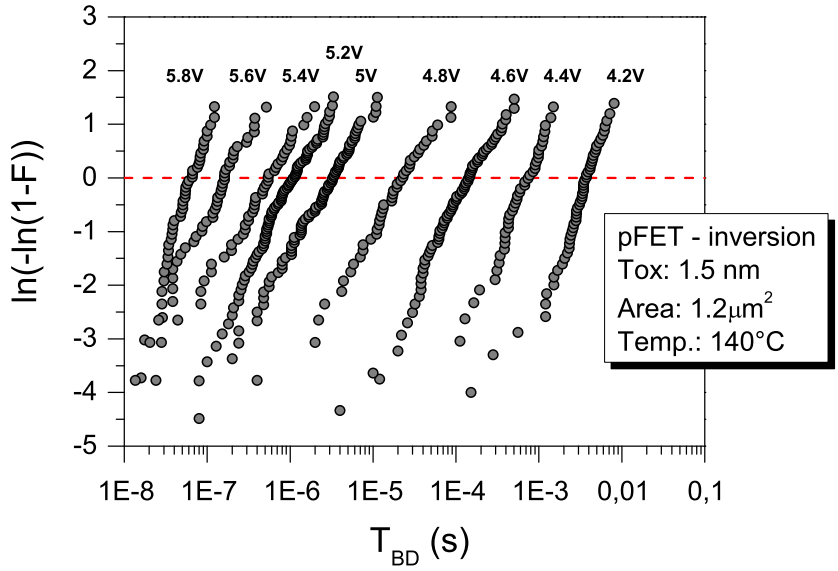

Figure 4.19: Gate oxide time to breakdown distribution of 1.5 nm pFET devices at 140 °C for CVS measurements from 4.2 V to 5.8 V using the set-up exposed in Section 3.2.3.

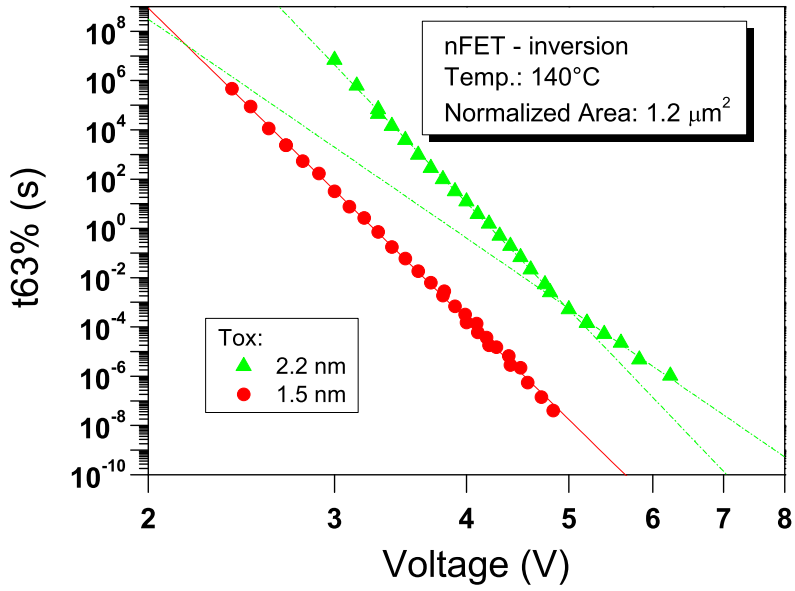

Figure 4.20: Voltage acceleration for 1.5 nm and 2.2 nm nFET stressed in inversion at 140 °C [37, 23].

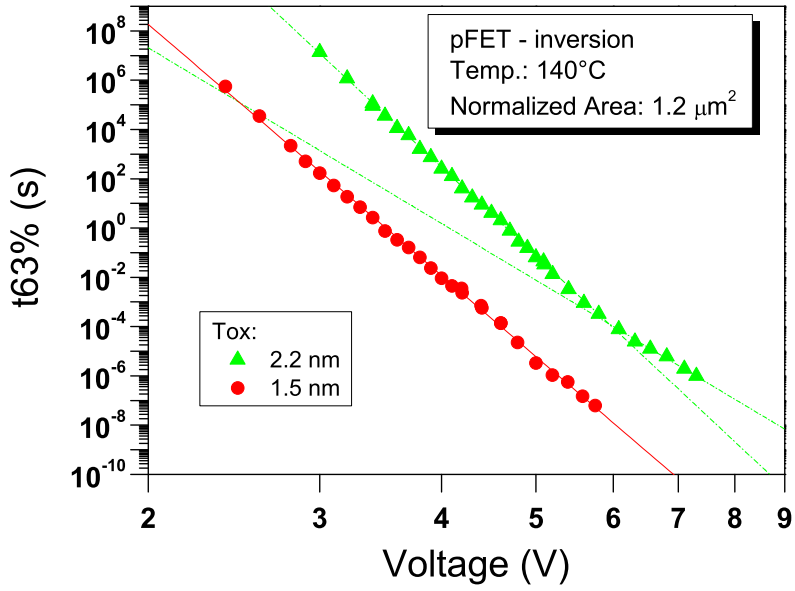

Figure 4.21: Voltage acceleration for 1.5 nm and 2.2 nm pFET stressed in inversion at 140 °C [37, 23].

The TDDB data obtained (Figure 4.20 and 4.21) by different measurements set-up are covering, with real good agreement, 14 decades in time. From the long-term and medium-term reliability [37] down to the ESD regime, the data are following a power-law [38] behavior described by the equation 1.67 below:

$$t = t_0 \left(\frac{V}{V_0} \right)^{-n}$$

This confirms that the thin GOX breakdown is a voltage driven phenomenon [39, 40, 41, 38] and not a thickness or field dependent effect as it was thought in the past [42, 43, 44, 45, 46, 10, 12]. Anyway, for the 2.2 nm oxide thickness, there is a clear change in the voltage acceleration factor around 5 V for nFETs and around 6 V for pFETs. Recent studies [24, 47, 19] have reported this power-law exponent transition for thicker oxides. It is the first time that this changing phenomenon in the breakdown mechanism is so clearly exposed for thinner oxides and through the same continuous TDDB experiments. This transition region is believed to be related to the changing behavior in the gate leakage current from direct tunneling regime to the Fowler Nordheim regime. This is in agreement with the threshold in the defect generation rate reported in [2, 48] and to recent studies reporting a change in the hydrogen excitation modes [49].

An universal TDDB modeling for thin oxides including this kink effect was proposed by Duschl and Vollertsen [47]. The same degradation mechanism holds on the whole voltage range but a voltage acceleration change around 5.3 V should be to accounted for. The breakdown can be described by

$$\left(\frac{t}{t_0} \right) = \left(\frac{V}{V_t} \right)^{-n} + \left(\frac{V}{V_t} \right)^{-m} \tag{4.4}$$

with the typical values at 140 °C, $n = 44$ and $m = 30$. The voltage acceleration factor γ is expressed as,

$$\gamma = -d\ln(t)/dV = \frac{n \cdot (V_t/V)^{(n-m)} + m}{V \cdot (1 + (V_t/V)^{(n-m)})} \tag{4.5}$$

The GOX TDDB voltage acceleration follows the same power law from long-term/low-voltage regime up to the high-level/short-time stress domain. This continuity of the TDDB model and of the reliability theories in the ESD domain allows the extrapolation of the oxide breakdown values in the HBM and CDM regimes from the data generated at low voltages done in all technologies for the process qualification (dielectric lifetime evaluation).

The continuity of the theories and TDDB behavior from the low voltages to the ESD domain gives a strong motivation for an extensive oxide breakdown characterization in the short-time domain. The knowledge of the GOX breakdown dependence with temperature, stress polarity and device type is essential to complete the global oxide breakdown picture model in the ESD regime. In the goal of providing a powerful extrapolation tool for gate oxide robustness facing fast over-shoots, all the parameters cited above will be characterized in the next following sections.

4.3.1.2 Voltage acceleration at 25 °C.

ESD "handling" event occurs in general at room temperature, therefore further TDDB measurements have been performed at 25 °C. The TDDB voltage acceleration of 1.5 nm pFET in inversion obtained at 25 °C is shown in Figure 4.22. The $1/E$ model and the E-model expressed in their voltage acceleration forms, $t \cdot \exp(G/V)$ and $t \cdot \exp(-\alpha \cdot V)$, respectively, are also plotted in Figure 4.22.

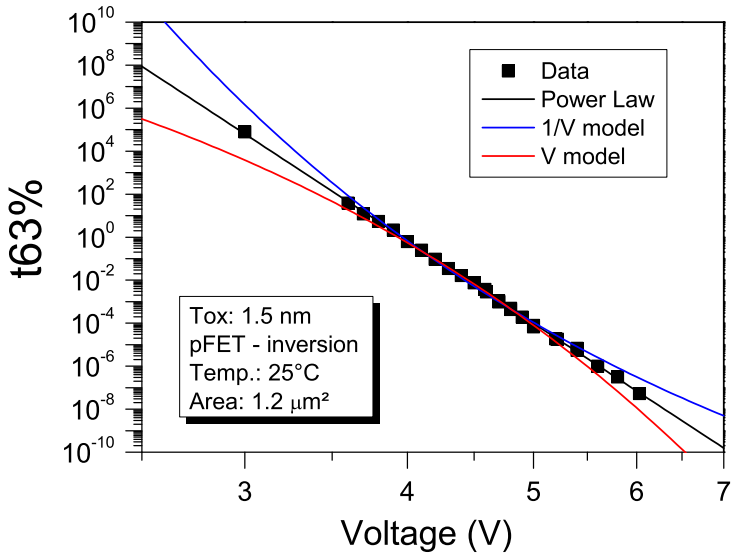

Figure 4.22: TDDB of 1.5 nm pFET stressed in inversion at 25 °C [50].

It is obviously exposed that the E-field or the inverse of the E-field are no more valid parameters to describe the TDDB acceleration for thin GOX, neither towards the lifetime direction [19] nor towards the ESD regime [50].

Extended TDDB experiments have been performed on state-of-the-art logic sub-micron CMOS processes. Eight oxide thicknesses were tested: 1.1 nm, 1.5 nm, 1.97 nm, 2.2 nm, 2.6 nm, 3.6 nm, 5.2 nm and 6.8 nm. The devices under test were nFETs and pFETs operating in inversion condition ($V_g > 0$ for nFET, and $V_g < 0$ for pFET, respectively). The TDDB are plotted in Figure 4.23 and Figure 4.24. Once again, the results are perfectly described by a power-law model and the same transition regimes are seen. In the case of thick oxides as 3.6 nm, 5.2 nm or 6.8 nm, this transition regime shifts towards very long times. In the other way round, for thin gate oxides in the range of 1.1 nm or 1.5 nm, this transition can experimentally not be observed because it would imply an evaluation of the TDDB below the picoseconds regime. For the ESD community, the strong dependence of the gate oxide breakdown with the voltage is a

constringent model which is somehow relived by the reduction of the voltage acceleration after the transition regime.

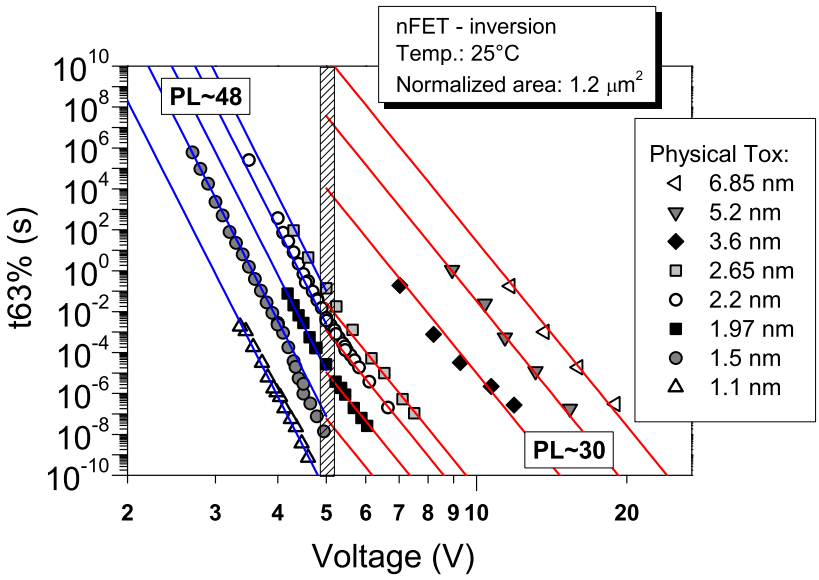

Figure 4.23: Voltage acceleration for nFET stressed in inversion regime at 25 °C normalized to the area of 1.2 μm^2 [51].

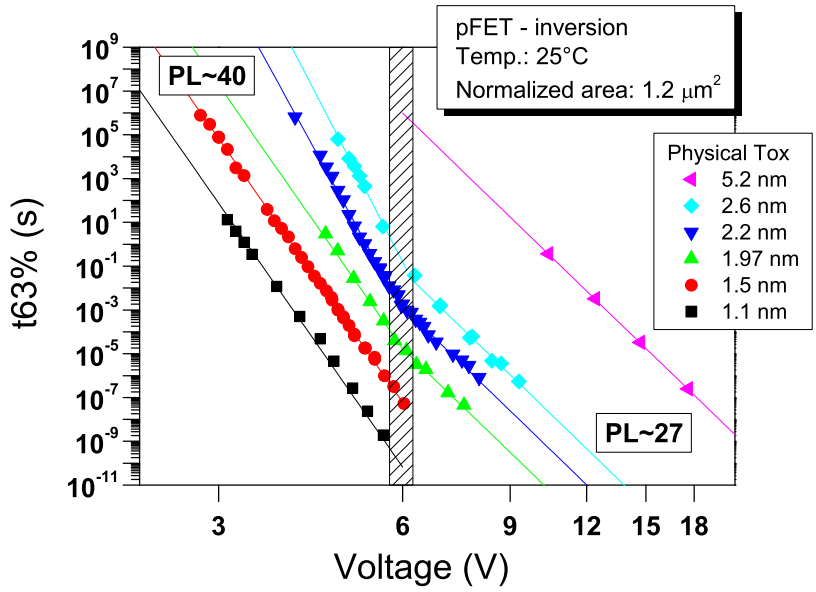

Figure 4.24: Voltage acceleration for pFET stressed in inversion regime at 25 °C normalized to the area of 1.2 μm^2.

The GOX TDDB voltage acceleration of oxide thinner than 7 nm stressed in inversion is described by a power law. Nevertheless, the voltage accelerations of the TDDB at room temperature are higher than at 140 °C. Regardless of the technology and thickness, the same acceleration factor can be used. In the case of nFET stressed in inversion regime, a voltage acceleration of 48 provides a good fit for the nFET data below the kink and a lower acceleration of 30 is found above it. For the pFET devices stressed in inversion lower voltage acceleration values are observed. An power law exponent of 40 is well describing the data below the kink whereas an exponent of 27 is modeling the data above the kink. This transition in the voltage acceleration is occurring around 5 V for nFET stressed in inversion. For pFET devices the kink is seen 1 V later around 6 V.

4.3.2 Post breakdown modes

As previously shown, the time-to-breakdown is following the same acceleration law as a function of the stressing voltage between the long and medium term reliability and the ESD time range. However, a difference can be seen in the gate oxide's post breakdown modes, which can be soft or progressively increasing in conductivity until being hard. The time resolution of the measurement set-up as well as device type, stress polarity, and active GOX area play a role in experimentally evaluating the different modes [52]. Contrarily to long and medium DC time range (EOS) in case of 100 ns pulses, not the first soft breakdown is detected, but oxide is stressed into a harder breakdown mode. TDDB tests performed on nFETs for oxide thickness down to 1.1 nm allow a good data matching between the ESD and the long/medium time domains, because there is no significant difference in time between soft and hard breakdown observed in these experiments. For the 1.1 nm pFET stressed in the inversion biasing mode, a difference is measured between the long/medium time and the ESD range (Figure 4.25) [51], because the time difference between soft and hard breakdown becomes remarkable long. The use of the first breakdown event as a failure criterion in the DC range leads to a severe mismatch with the data obtained in the nanoseconds regime whereas the consistent use of the hard breakdown criterion in the DC range results in a perfect data correlation. This is an experimental artifact, which needs to be considered to correctly interpret the data and to avoid wrong conclusions for the ESD design window.

For pFET in accumulation biasing mode. The progressive increase in the gate leakage is shown in Figure 4.26 and Figure 4.27; it demonstrates how clearly soft breakdown in the progressive phase and a hard breakdown can be experimentally distinguished for pFET. The power law characteristic of post breakdown gate leakage [53] can be used for identifying the post breakdown mode as:

$$I_{\mathrm{G}} = K \cdot V^{\alpha} \tag{4.6}$$

Hard breakdown results in an ohmic characteristic, which means $\alpha = 1$. Values larger than 1 indicate that the post breakdown mode is in the progressive phase. The larger the value of α,

the closer is the status of the breakdown spot to the initial percolation path which forms the first SBD [54].

In the case of pFET triggering on soft breakdown becomes a delicate task for TDDB tests in the ESD range, because during the applied voltage pulse, a low voltage drop on linear scale is caused by a small current increase on logarithmic scale, which itself already represents the progressive phase. In addition to this experimental problem in accumulation biasing mode, the well resistance may cause a voltage drop during high voltage breakdown tests in the ESD time range. In Figure 4.28 the I_g - V_{gs} characteristic of gate leakage reveals a resistance of 1 kΩ. This allows correcting measured time to breakdown (Figure 4.29) and plotting the real difference in time between medium and ESD test range as discussed above.

The soft breakdown of thin oxides can cause GIDL in narrow transistors [55] and characteristic shifts but may not have damaged the IC functionality in some logic domains [56, 57, 58]. A net increase in the leakage gate current is seen and this event will be succeeded later on by a HBD after the residual time, which is independent of the pre-stress [52, 34]. However, the thin oxides used in the logic parts can take benefit of this soft breakdown event as the residual time considerably extends the lifetime of the thin oxides under normal operation conditions [16].

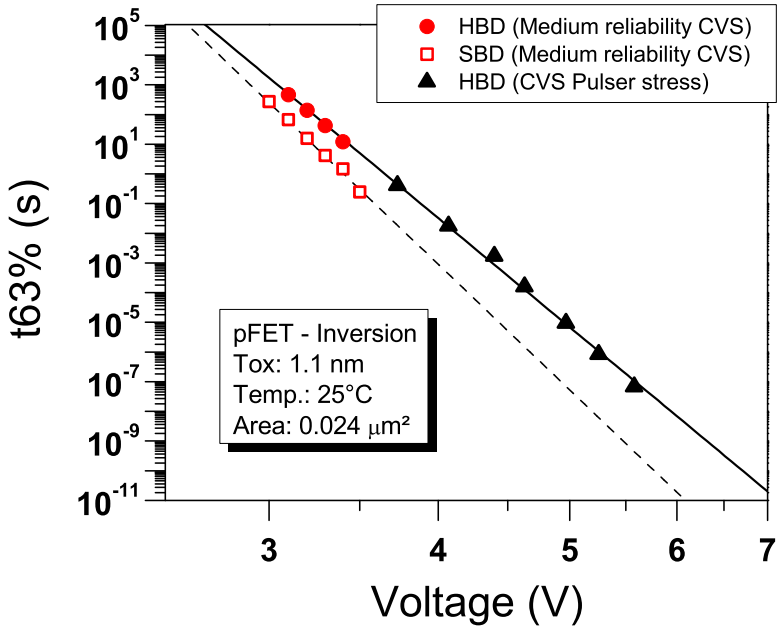

Figure 4.25: Voltage acceleration of 1.1 nm pFET in inversion obtained from conventional TDDB and short-time measurements set-ups at room temperature [51].

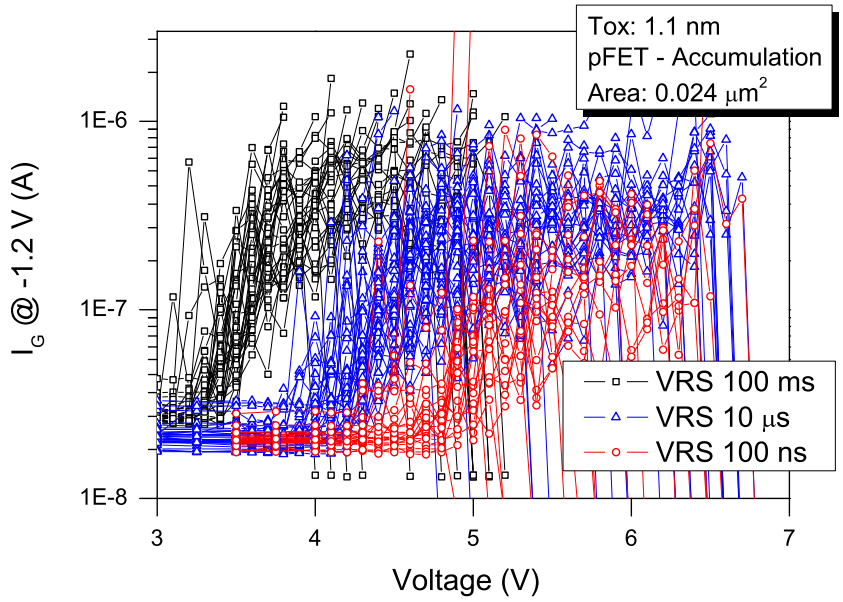

Figure 4.26: Progressive increase in the monitored gate leakage current of 1.1 nm pFET stressed in accumulation by Voltage Ramp Stress (VRS) [51].

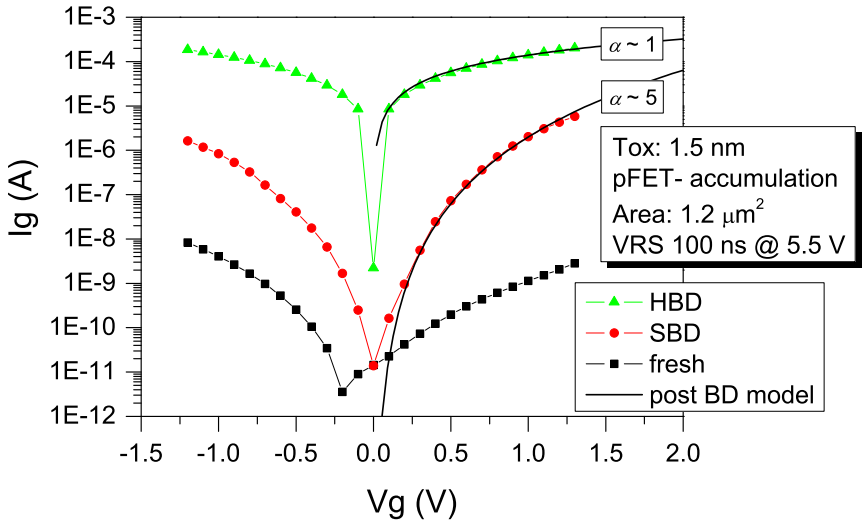

Figure 4.27: Post breakdown gate leakage of a 1.5 nm pFET device stressed in accumulation with CVS [51].

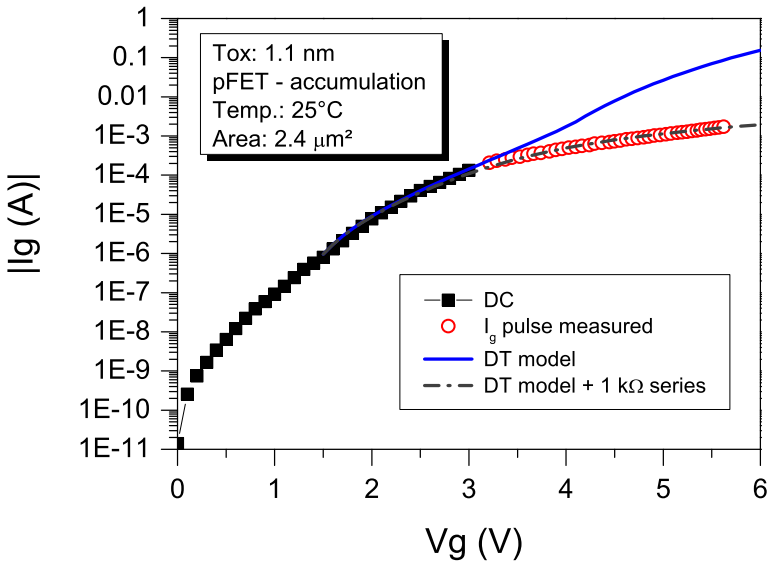

Figure 4.28: 1.1 nm pFET $I_g - V_g$ characteristic obtained from DC and short pulse measurement. The fitting of the tunneling current reveals internal devices voltage drops [51].

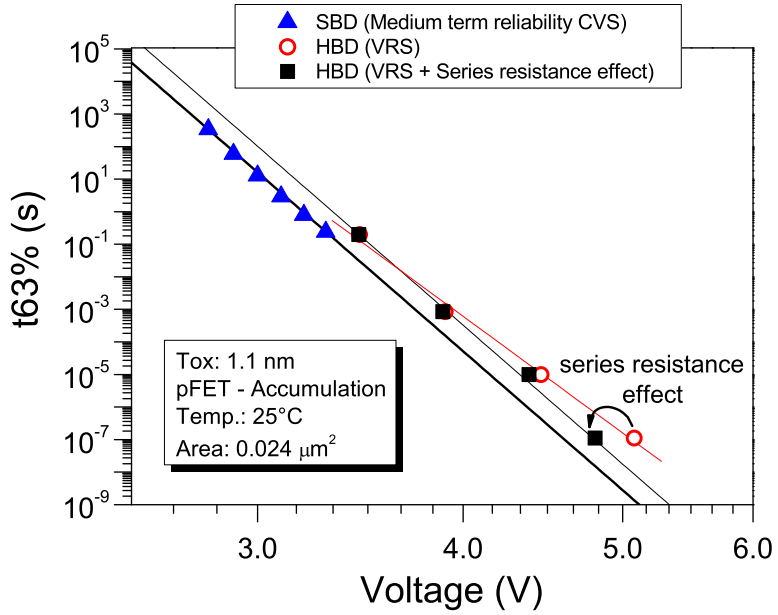

Figure 4.29: Voltage acceleration for 1.1 nm pFET stressed in accumulation obtained from conventional TDDB and VRS short-time measurements [51].

4.3.3 Polarity dependence

ESD events can lead to a stress in accumulation, for example in buffer capacitors. Thus, the same characterization work was done in accumulation biasing mode and has revealed different voltage acceleration power law compared to the inversion biasing mode. Comparative examples are given for the 1.1 nm nFET in Figure 4.30 and for the 2.2 nm in Figure 4.31.

An overview of the TDDB results obtained for nFET and pFET as a function of the polarity is summarized in Figures 4.32 and 4.34.

The oxide time-to-fail of devices stressed in accumulation is also described by a power law but which differs from the inversion stress one. The oxide robustness hierarchy a function of the stress polarity and device type can be quantitatively evaluated in the ESD-stress regime. An universal change in the acceleration voltage behavior occurs around 5-6 V. It is observed that this transition in the acceleration voltage factor is indeed occurring around 6 V at the exception case of the nFET devices stressed in inversion for which the kink occurs in the range of 5 V. This transition region is exposed for nFET and pFET devices stressed in accumulation in the Figure 4.34. Thin oxide below 2 nm are not affected anymore of the kink in the nanoseconds time range and the difficulty of the thin oxide protection is increased due to the strong voltage power law dependence.

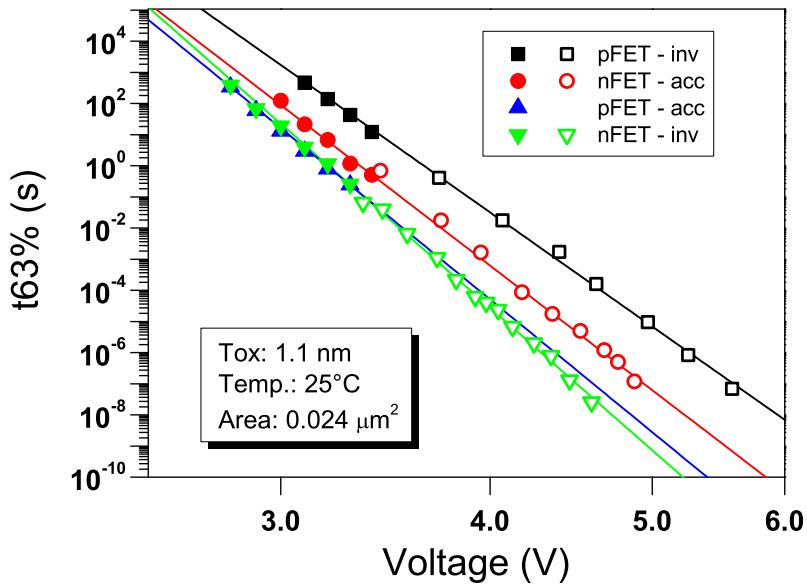

Figure 4.30: Voltage acceleration for 1.1 nm nFET and pFET stressed at room temperature obtained from conventional TDDB and short-time measurements set-ups.

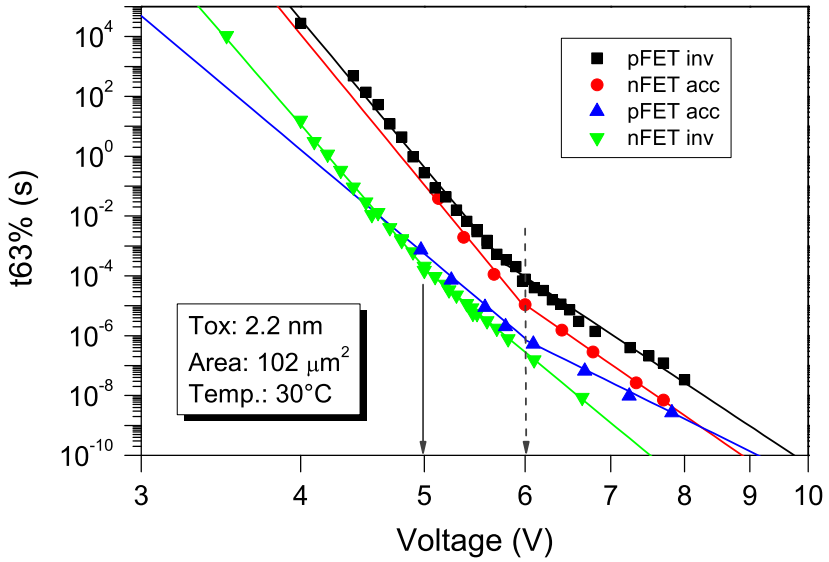

Figure 4.31: Voltage acceleration for 2.2 nm nFET and pFET stressed at room temperature obtained from conventional TDDB and short-time measurements set-ups.

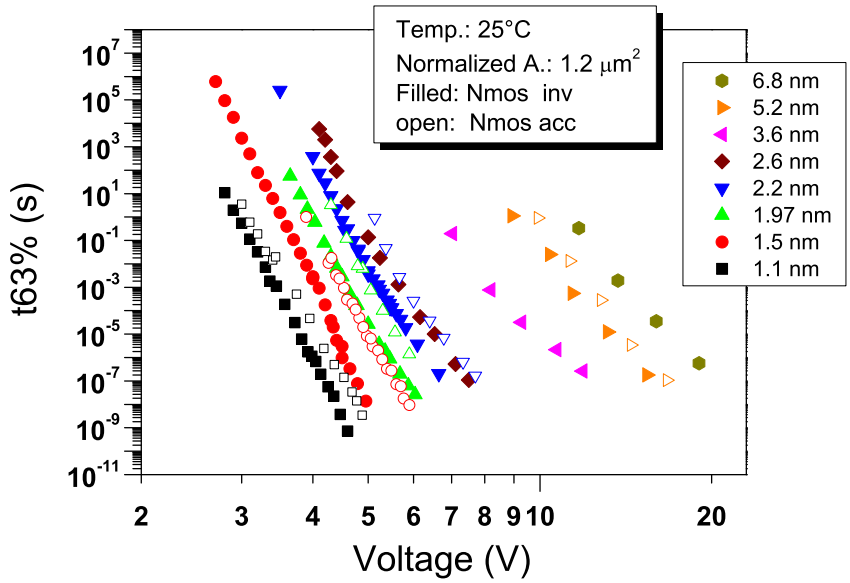

Figure 4.32: Voltage acceleration of nFET devices at 25 °C under inversion stress (filled symbols) and accumulation stress (open symbols).

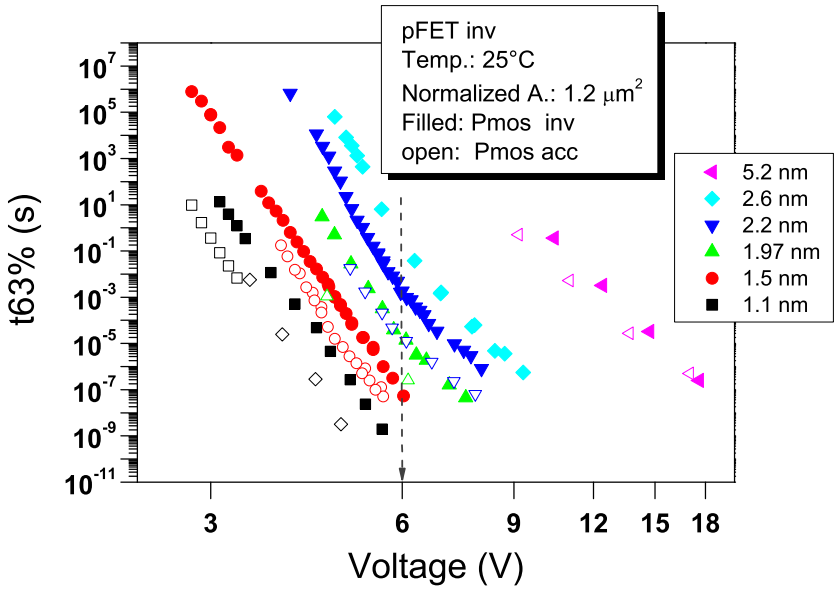

Figure 4.33: Voltage acceleration of pFET devices at 25 °C under inversion stress (filled symbols) and accumulation stress (open symbols).

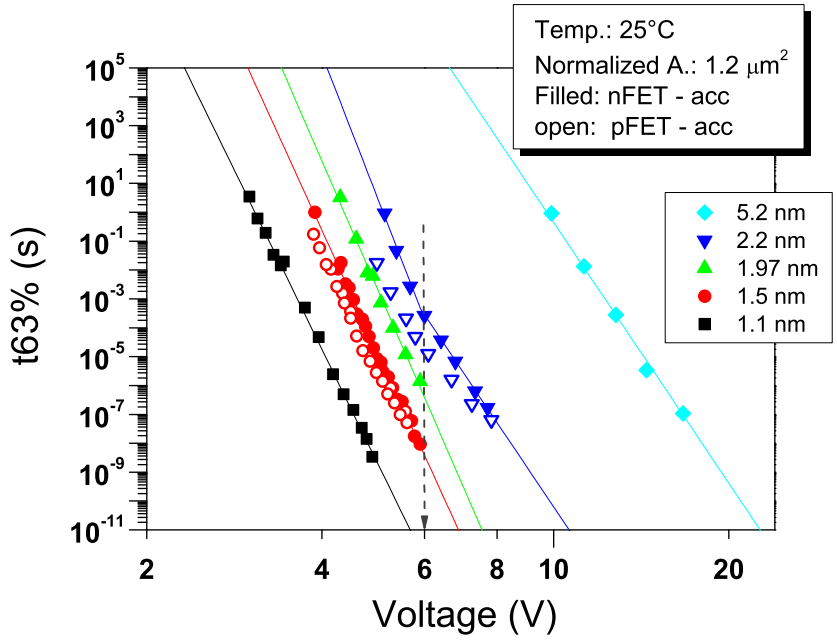

Figure 4.34: Voltage acceleration of nFETs (filled symbols) and pFETs (open symbols) devices at 25 °C under accumulation stress.

4.3.4 Physical origin of the transition regime observed in the TDDB voltage acceleration

The occurrence of the reported kink in the TDDB voltage acceleration leading to two distinct domains which are both characterized via a power law equation is up to know not clear. The first fact observed is the independence of the kink value with the temperature, see Section 4.3.1.1 and later Section 4.3.6. This transition regime was claimed to be linked with the change of the gate leakage conduction mechanism from the direct tunneling to the Fowler-Nordheim tunneling [47]. However, concerning the case of the pFET stressed in inversion regime, the kink is showing off around 6 V which already corresponds to the steady state of the Fowler-Nordheim tunneling current. Moreover this transition regime is seen for the same voltage value independently of the oxide thickness. It seems that this phenomenon is not related with the field across the oxide (E_{ox}) but rather to the voltage across the oxide (V_{ox}).

Due to the strong correlation of the breakdown with the voltage [41, 24] and the clear impact of hydrogen species involved in the thin oxides breakdown [39, 49, 59], the AHR hypothesis based on the hydrogen bond breaking at the anode due to the release of the maximal energy of the electron injected from the cathode seems to be a fair theory. The direct hydrogen bond breaking threshold by electrical excitation has been reported to be 5 eV [60]. However below this energy indirect breakdown mechanisms based on vibrational mechanisms have been reported to still create hydrogen defects. Between 2.5 eV and 5 eV an indirect excitation of the hydrogen bonds by single injected electron gives a good explanation for the strong voltage dependence of the time to breakdown [61] and for the defect probability generation voltage dependence as well [62, 48]. Below 2.5 eV, bond breaking via multi-vibrational effect [63] occurs. This process requires at least two electrons and is dependent of the fluency.

In the kink voltage domain, already two phenomena are taking place, the change in the oxide conduction leakage mechanism and in the excitation mode of the hydrogen bonds. Nevertheless, we will focus on the value of the oxide barrier height (V_{ox}) at which the kink is seen.

- In the case of devices stressed in accumulation, V_{ox} can be deduced from V_g from the following equation [46, 64, 65],

$$\begin{aligned} q|V_g| &= q|V_{ox}| + |E_g| \\ &= q|V_{ox}| + 1.12 \end{aligned} \tag{4.7}$$

  The kink is seen to occur at V_g = 6 V for devices stressed in accumulation. That is to say for an voltage across the oxide of $V_{ox} \sim 5$ V.

- For devices stressed in inversion, $qV_{ox} = qV_g - qV_{poly} \sim qV_g$. If we consider the nFET stressed in inversion the same V_{ox} value of 5 V is found as the kink occurs for V_g = 5 V.

- However, the pFET stressed in inversion presents a more complex case due to the p^+ doped poly-silicon. Under inversion, the electrons are injected from this poly-Si gate. In this particular case, the electron are indeed not injected from the valence band but from a level closer to the conduction band [64, 65]. The barrier height Φ (eV) for injected electrons is voltage dependent accordingly to:

$$\Phi(V) = \Phi_{\mathrm{b}} + E_{\mathrm{g}} - qV_{\mathrm{poly}} \tag{4.8}$$

  The effective oxide barrier height seen from the injected electron is therefore $\mathrm{V_{ox,eff}} = V_{\mathrm{ox}}$ - E_{g}/q. The similar effective oxide barrier height of 5 V also appears for the pFET stressed in inversion as $qV_{\mathrm{ox}} = qV_{\mathrm{g}} - qV_{\mathrm{poly}} \sim qV_{\mathrm{g}}$ in inversion.

It seems that the dissipation energy mechanism changes when the injected electrons are facing an oxide barrier bigger than 5 V. The transition from the elastic electrons tunneling through the oxide giving at the anode side their maximal energies to the inelastic release of their energies when passing trough a trapezoidal oxide barrier, triggered for V_{ox} = 5 V, gives an good explanation for the lower voltage acceleration factor seen after the kink and then for the physical origin of the kink itself.

4.3.5 Thickness dependence

The oxide voltage acceleration behavior of nFETs stressed in inversion at room temperature for eight distinct thicknesses was summarized for the same normalized size in Figure 4.35. From this general behavior, the effective time to breakdown scaling factor as a function of the physical oxide thickness can be extracted. The TDDB dependence of the oxide thickness is consistent with the percolation cell model picture [24]. A TDDB oxide thickness scaling factor around 6 decades in time per nanometer is found for oxides in the thickness range of 1.1 nm to 2.5 nm (Figure 4.35). Based on this characterization data, GOX breakdown voltage values in the ESD and EOS time domains for oxide thicknesses in the range of 6.85 nm to 1.1 nm stressed in inversion at room temperature can be extrapolated.

The GOX breakdown in the HBM regime as a function of oxide thickness is plotted in Figure 4.36 for NMOS and PMOS devices in accumulation and inversion stress. It should be mentioned that this trend is specific to the area and failure criteria chosen, as the TDDB voltage acceleration is dependent on the stress polarity and device type. The GOX breakdown as a function of oxide thickness is different for other areas and is also dependent of the choice of a meaningful statistical failure level (see Chapter 6). We are close to the physical limit of SiO2 (or SiON) dielectric material with an important reduction in the gate oxide breakdown voltage. Identification of critical ESD cases is becoming crucial for the design of ESD robust ICs.

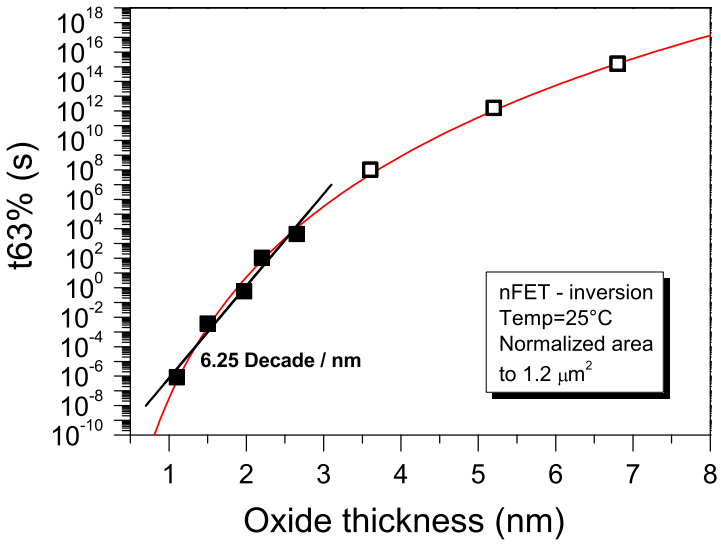

Figure 4.35: TDDB cross section from Figure 4.23 at 4 V showing the acceleration factor from the GOX time to fail as a function of the oxide thickness [51].

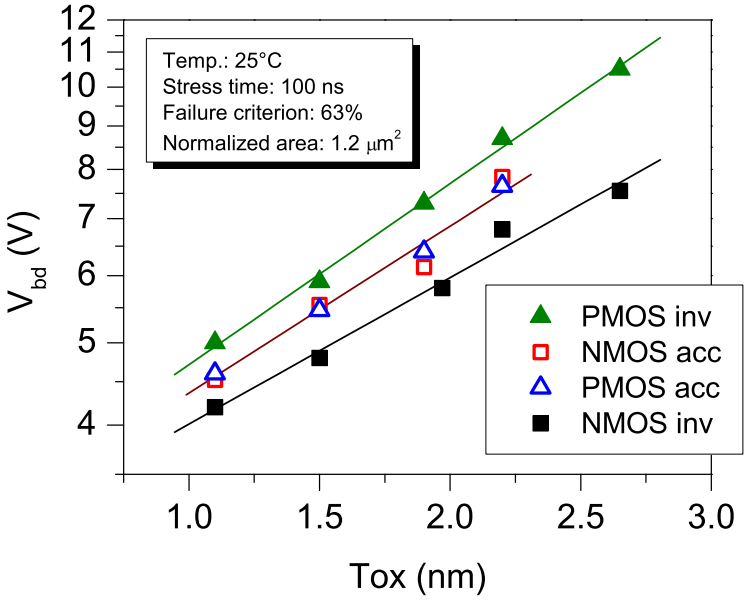

Figure 4.36: Breakdown voltage acceleration as a function of the physical oxide thickness for nFET and pFET stressed at 25 °C in the HBM range of time for a normalized size of 1.2 μm^2 [51].

4.3.6 Temperature dependence

The knowledge of the GOX breakdown values extrapolated from the reliability qualification data is an essential topic for ESD groups. As the universal power-law model established to qualify the gate oxide lifetime can be used in the ESD range, this results in a good alternative method to determine the GOX breakdown which does not need so much time, cost and effort. The qualification measurements for lifetime estimations are done at elevated temperature. To fulfill this interest, a focus on the temperature influence on the TDDB is developed here.

The TDDB measurements of 1.5 nm and 2.2 nm thin oxides at 25 °C and 140 °C are plotted together in Figure 4.37 for nFET devices and in Figure 4.38 for pFET devices. At high voltages the TDDB model from the gate oxide stressed at 25 °C is converging towards the elevated temperature power-law; we note that both data are in the same range below 10 µs. The TDDB at high voltages for the nFET seems to be weakly temperature dependent. For a better understanding of the temperature influence, the ratio between the $T_{63\%}$ value at 25 °C and 140 °C versus the voltage is shown in Figure 4.39 for pFET devices. This plot reveals the strong decreasing behavior of the thermal activation energy ratio function with increasing stress voltage. The gate voltage dependence for the thermal acceleration parameter was already discussed in the past [66, 67, 68]. It was stated that this phenomenon was not a thickness effect but rather resulting from a voltage-dependent voltage acceleration and a temperature-independent voltage acceleration within a fixed time-to-breakdown (TBD) window [67]. From the larger picture established in this work, it is evident that this weakly temperature dependent acceleration behavior occurs at low voltage and in a narrow voltage window to first order which is not correctly extended to the full range especially for high voltages. To enable extrapolations from two different domain regimes, that is to say between the elevated temperature (125 °C - 140 °C) in the low voltage regime from the reliability department and the ESD world which is generally occurring at room temperature with high voltage drops, the accuracy of the extrapolation model has to be verified. The voltage dependence of the thermal activation energy has a serious impact on the temperature extrapolation model towards the ESD regime. This parameter needs to be incorporated in the thermal model, otherwise the derived data obtained from the lifetime estimation towards ESD regime would be too optimistic for the ultra-thin gate oxide.

Further investigation on the TDDB dependence with the temperature of thin oxide has been performed on a 1.1 nm oxide from a 45 nm CMOS process. Constant voltage stress have been performed at different temperature. No dependence of the Weibull slope as a function of the temperature has been noticed. The time to breakdown distribution of the 1.1 nm nFET stressed in inversion at high temperature is exposed in Figure 4.40. A Weibull slope around 1.2 is found, which is in good agreement with the Weibull slopes obtained at room temperature. The deviation observed for the distribution obtained at 50 °C is resulting from the set-up extraction accuracy.

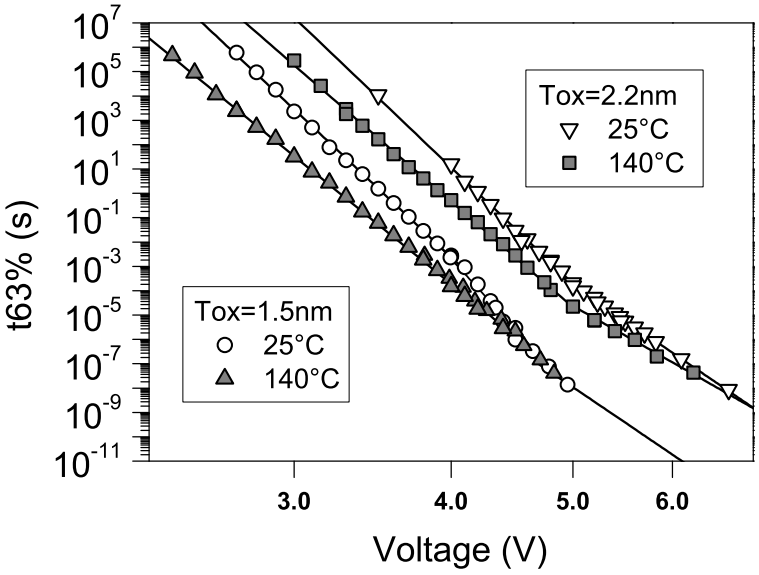

Figure 4.37: Voltage acceleration for the 1.5 nm, 2.2 nm nFET at 25 °C and 140 °C [23].

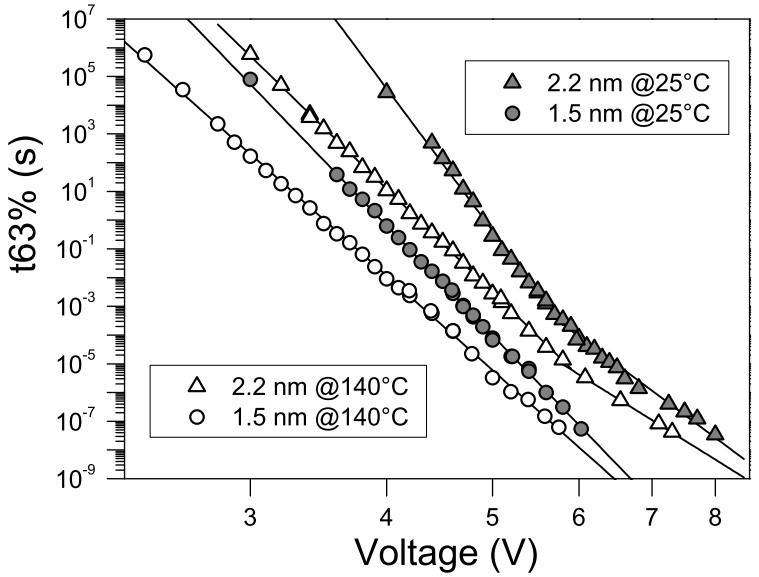

Figure 4.38: Voltage acceleration for the 1.5 nm, 2.2 nm pFET at 25 °C and 140 °C.

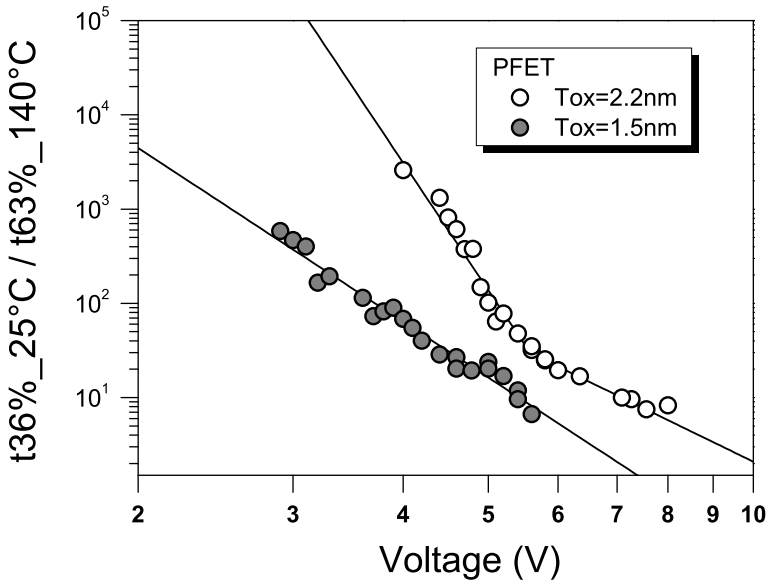

Figure 4.39: Thermal activation energy ratio of the TBD function of the stress voltage for 1.5 nm and 2.2 nm pFET [51].

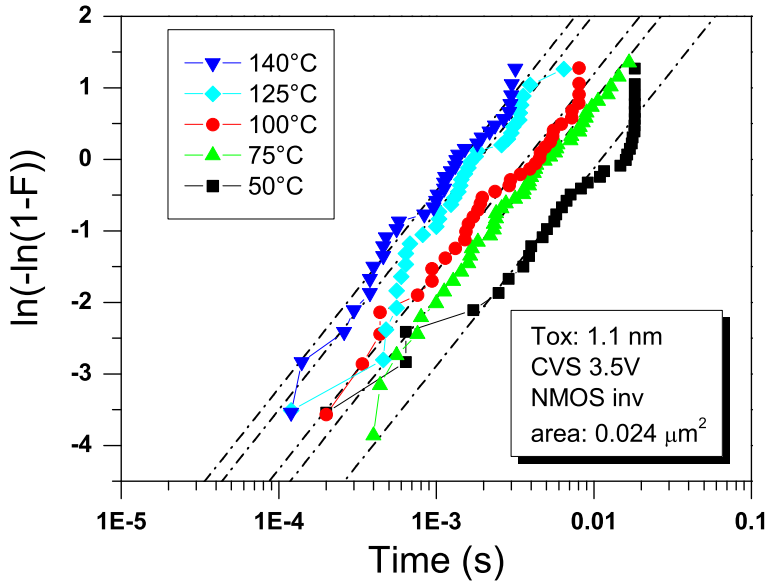

Figure 4.40: Gate oxide time to breakdown distribution as a function of the temperature for 1.1 nm nFET under 3.5 V constant voltage stress.

The TDDB obtained for the different temperature at 4 CVS levels are exposed in Figure 4.41 for nFETs and in Figure 4.42 for pFETs. A remarkable reduction of the voltage acceleration with the temperature increase is observed, this phenomenon has been also reported at low voltages [67]. This effect is plotted for different nFET oxide thickness stressed in inversion in the Figure 4.43. A linear reduction with a slope of $4.7 \times 10^{-2}\ ^{\circ}\mathrm{C}^{-1}$ is well describing of the power law exponent decrease as a function of the temperature (as described in the Figure 4.43).

The time to breakdown of nFET and pFET can be indifferently plotted in linear plot versus the temperature or in an Arrhenius plot $(1/T)$. In the literature, it was recently reported that a non-Arrhenius tendency can be observed due to the dependence of the activation energy with the voltage [69, 66]. However the TDDB results are plotted with correct agreements in an Arrhenius form in the Figures 4.44 and 4.45. This permits to extract the thermal activation energy via the Arrhenius equation 4.9.

$$t_{\mathrm{BD}} = t_0 \cdot \exp\left[\frac{E_{\mathrm{a}}(V)}{k_{\mathrm{B}}} \cdot \left(\frac{1}{T} - \frac{1}{T_0}\right)\right] \tag{4.9}$$

This activation energy can be plotted as a function of the stress voltage. Very low values are reported at high voltages in Figure 4.46. For the nFET devices, a very strong reduction of the thermal activation leads to a quasi negligible effect with a coefficient around 0.1 eV at 4.5 V. Similar low values have been reported by Kerber *et al.* [16] in the range of 3 V to 4 V for a 1.5 nm oxide.

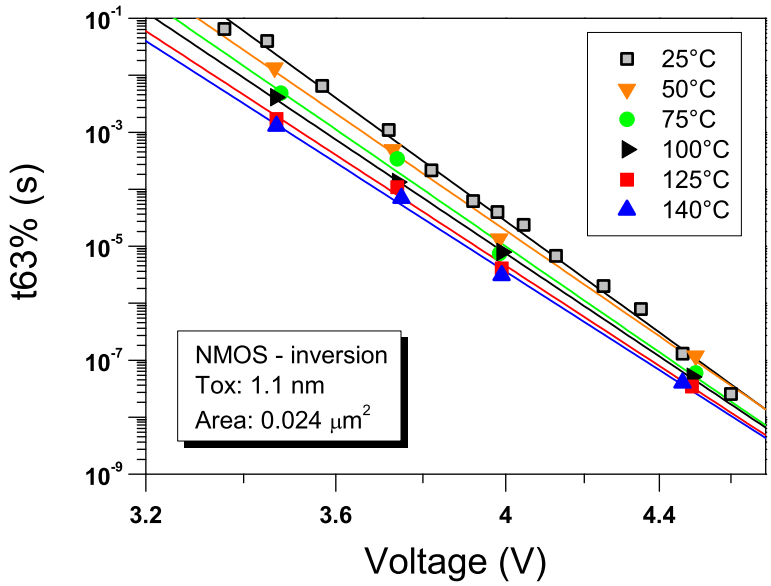

Figure 4.41: Voltage acceleration for the 1.1 nm nFET stressed in inversion as a function of the temperature. CVS at 3.5 V, 3.75 V, 4 V and 4.5 V.

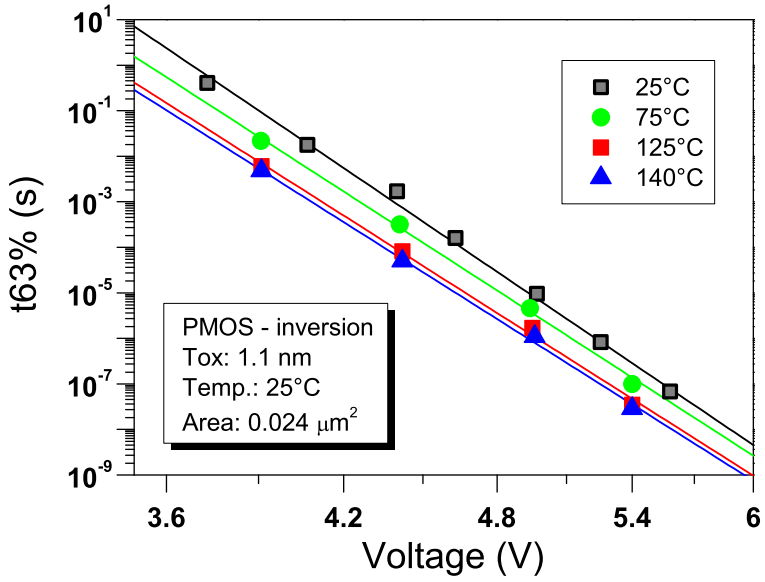

Figure 4.42: Voltage acceleration for the 1.1 nm pFET stressed in inversion as a function of the temperature.

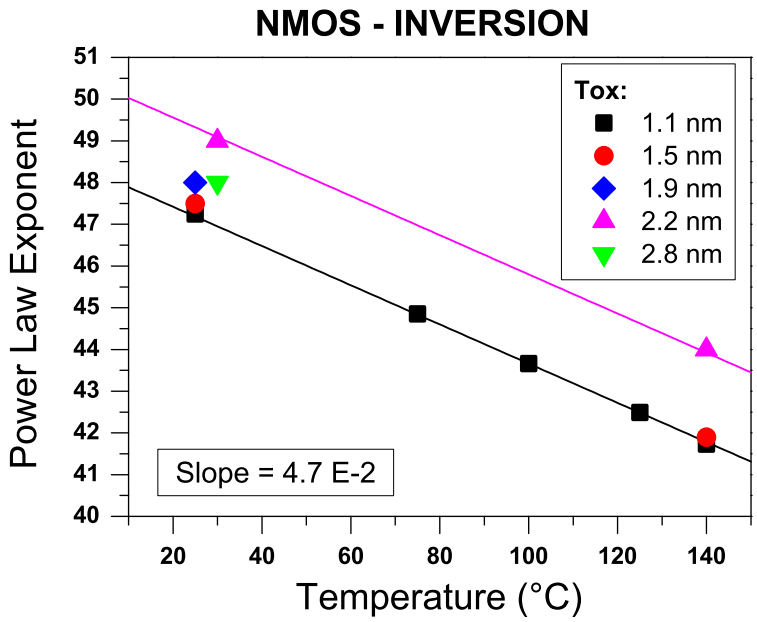

Figure 4.43: TDDB voltage acceleration factor dependence with the temperature of nFET stressed in inversion. CVS at 3.9 V, 4.4 V, 4.9 V and 5.4 V.

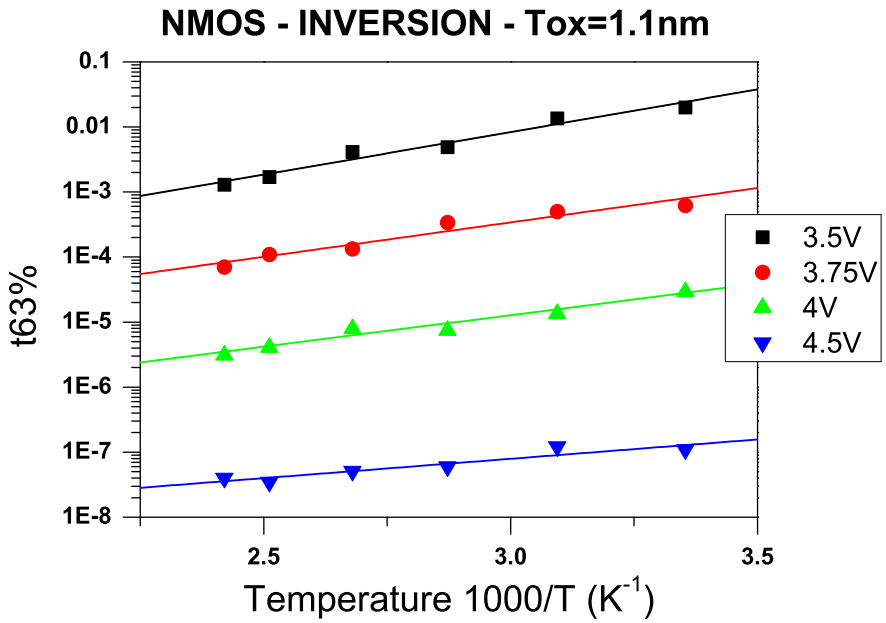

Figure 4.44: Time-to-breakdown of 1.1 nm nFET stressed in inversion as a function of the temperature for different CVS levels

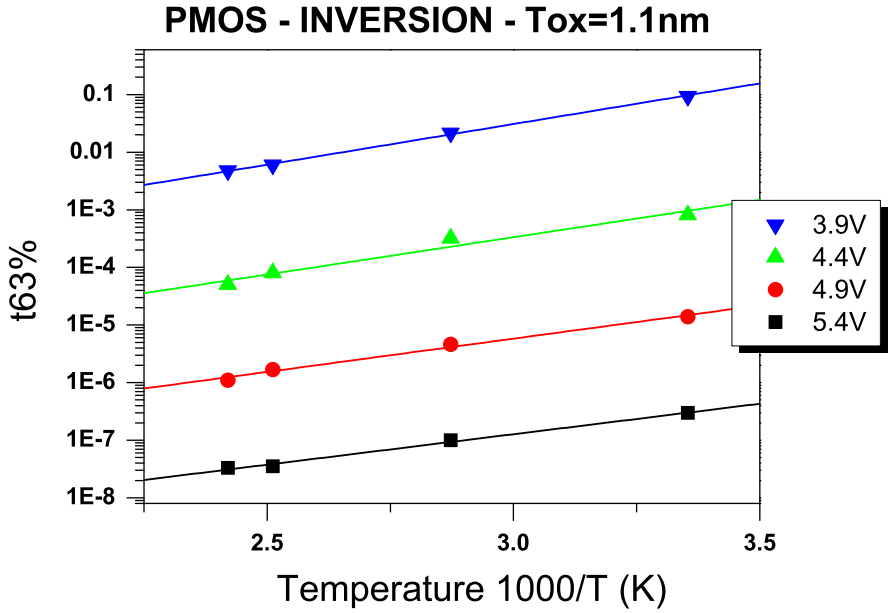

Figure 4.45: Time-to-breakdown of 1.1 nm pFET stressed in inversion as a function of the temperature for different CVS levels

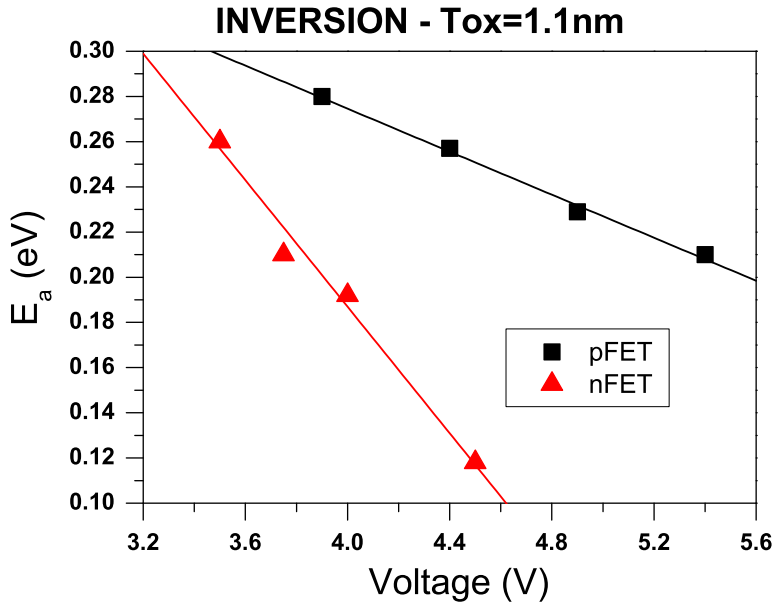

Figure 4.46: Thermal activation energy of 1.1 nm FETs as a function of the stressing voltage.

A strong dependence of the thin oxide TDDB dependence with the temperature has been reported from the reliability community at low voltages [24, 16]. However, a very strong dependence of the thermal activation energy with the voltage has been exposed in a wide range of voltage. In the ESD regime the impact of the temperature on the TDDB is becoming negligible for thin oxide. The voltage acceleration of the time-to-fail is decreasing with the temperature elevation. This very low activation energy in the high voltage range and the higher power law exponent at room temperature than at 125 °C should be taken into account for the correct extrapolation of the oxide breakdown from the long-term reliability data.

4.3.7 Charge to breakdown

The time-to-breakdown is the useful parameter for dielectric lifetime failure criterion and for ESD limits establishment. However, the physical breakdown picture considering the weakest link breakdown mechanism is better exposed with the charge-to-breakdown. The modeling of the GOX breakdown in respect with Q_{BD} will be expressed in this section.

The gate leakage currents density obtained from DC measurements for the different oxide thickness stressed are exposed in the Figure 4.47. From the DC measurements, high gate leakage data obtained from short pulse testing and via the DT and FNT fits, the deduce charge-to-breakdown calculated are summarized in the Figure 4.48 for nFET devices and in Figure 4.49

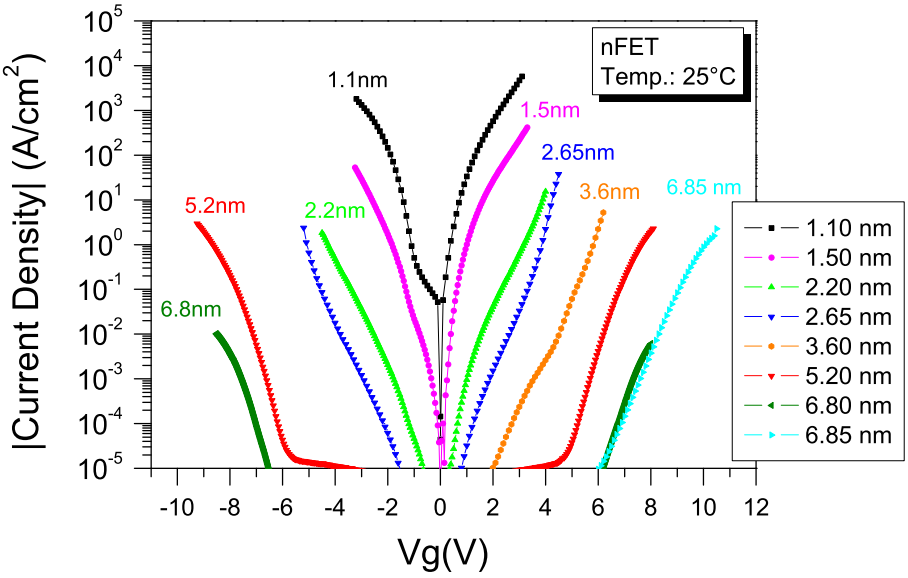

Figure 4.47: Gate leakage current density curves measured from different oxide thicknesses (5.2nm to 1.1nm).

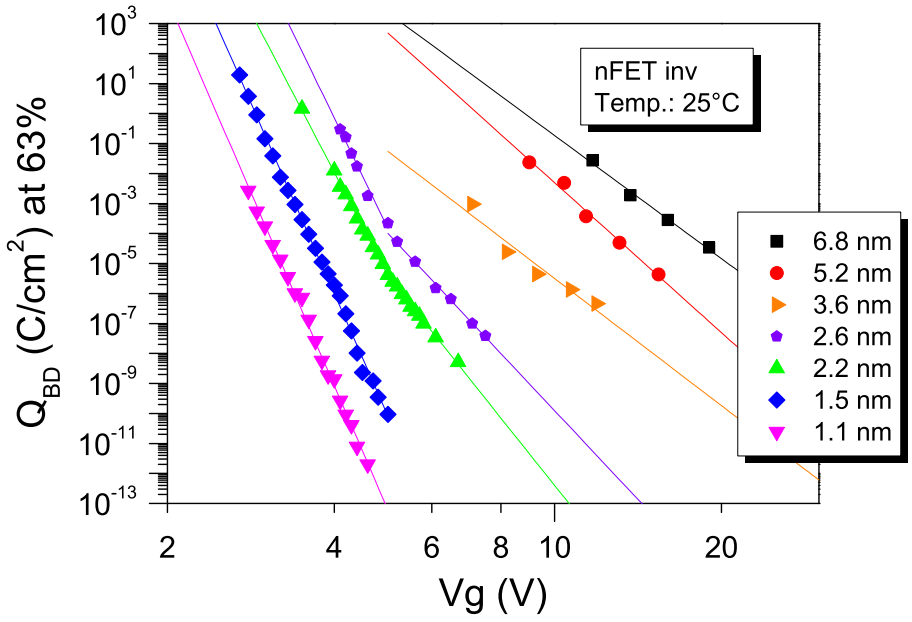

Figure 4.48: Charge to breakdown of nFET devices stressed in inversion at room temperature.

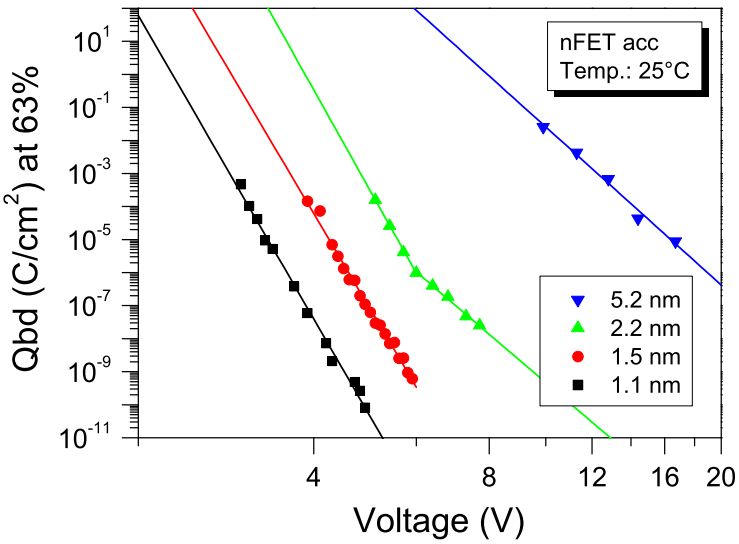

Figure 4.49: Charge to breakdown of nFET devices stressed in accumulation at room temperature.

for pFETS. The charge to breakdown is calculated from the electron fluence. The charge to breakdown Q_{BD} was obtained from the multiplication of the normalized TDDB data to the area of 1 cm^2 with the current density (J_g). Charge-to-breakdown (Q_{BD}) power laws are observed as a function of the voltage. As for the TDDB, the two different voltage acceleration are observable. The lower exponent value above the kink is around 15 and a value in the range of 30 to 36 are reported below the kink.

In the initial percolation theory [70, 71], the saturation of n_{BD} as a function of T_{ox} is predicted due to the lower oxide thickness values than the defect size (a_0). No saturation effect for oxides thinner than 2 - 3 nm is reported here in the charge to breakdown. For thin oxides, a more precise model relating the saturation of the Weibull slope should be accounted for. The critical defect density model at the breakdown from Suñé [62] was demonstrated to be valid experimentally also down to ultra-thin oxides [24] according to the equation (4.10),

$$N_{BD} = \frac{t_{ox}}{a_0^3} \exp\left[-\frac{1}{\beta(t_{ox})} \ln\left(\frac{A_{ox}}{a_0^2}\right)\right] \tag{4.10}$$

$$\text{where, } \beta(t_{ox}) = \frac{T_{ox} + T_{int}}{a_0} \tag{4.11}$$

The Weibull slope dependence with the oxide thickness found in this work is plotted in Figure 4.50, where the theoretical models from Suñé and Wu are plotted as well for comparison. The critical defect size a_0 and the interfacial defect size T_{int} observed from the characterization

are 2.17 nm and 1.1 nm respectively. Based on these parameters the critical defect density have be simulated in the Figure 4.51.

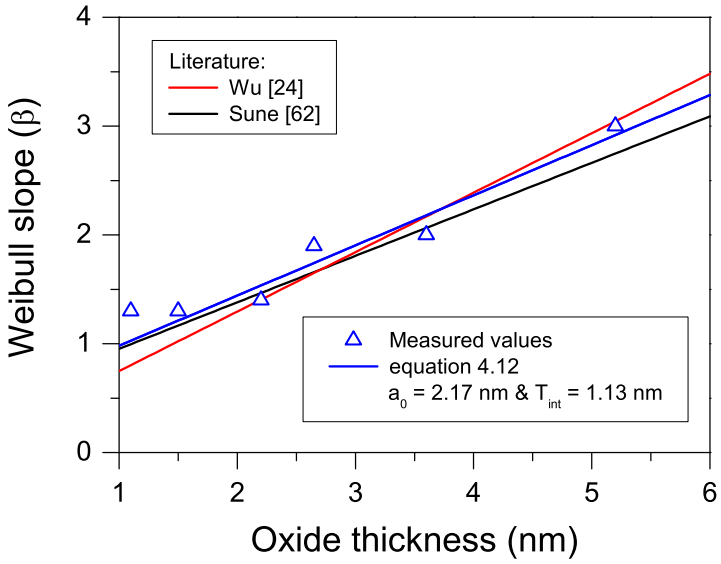

Figure 4.50: Experimental Weibull slopes β plotted as a function of the physical oxide thickness.

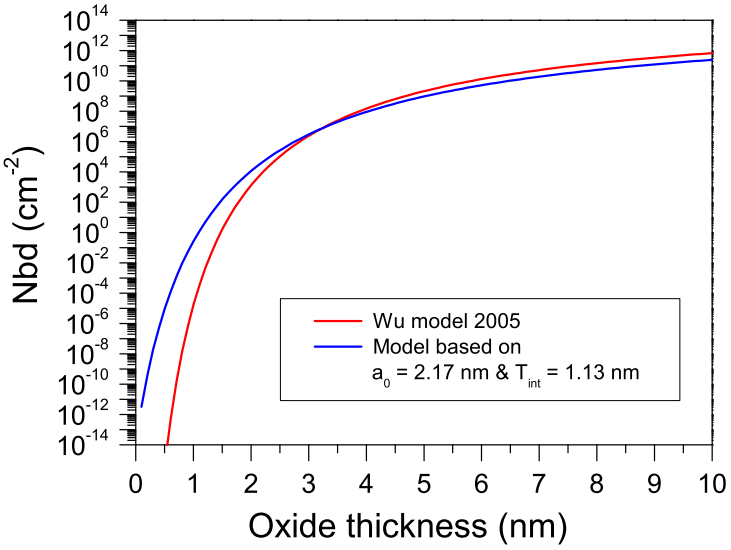

Figure 4.51: Critical defect density (n_{BD}) simulated as a function of the physical oxide thickness (T_{ox}).

From the charge to breakdown and the critical defect density, it is possible to express the defect generation probability (P_g) according to the following relation [72],

$$Q_{BD} = \frac{q \cdot N_{BD}}{P_g} \tag{4.12}$$

The defect generation probability (P_g) dependence with the voltage based on the experimental a_0 and T_{int} is plotted in Figure 4.52. This trend is depicted for nFET devices stressed in inversion and accumulation. Two different regimes in the defect generation are clearly exposed, above about 6 V, a saturation effect is seen. A similar behavior was reported based on SILC studies by Stathis and DiMaria [2]. This defect generation trend was also plotted as a function of the maximale energy released at the anode by Suñé in [62], where this universal saturation effect was reported around 5 to 6 V. Below the saturation regime P_g is described via a power law as,

$$P_g = E_{max}^{\alpha} \tag{4.13}$$

In the case of nFET devices, the maximum electron energy released at the anode is given by the gate voltage. A lower P_g power law acceleration as a function of the maximale energy released is found; an exponent α = 30 is well describing the data in comparison of the exponent 38 reported in [72]. Figure 4.52 shows some dispersion in the data at high voltages. At first glance this was thought to be an experimental issue due to the set-up sensitivity in solving the stress current and also in the use of current modeling based on FN equation which do not fully integrate the internal parasitics resistances incoming from the devices. However, the same data are plotted once again in the Figure 4.53 and Figure 4.54 using the $\beta(T_{ox})$ dependence from Wu [24]. The a_0 parameter is set to 1.83 nm in this model and the defect interface is 0.37 nm. The defect density at breakdown is also reported for these values in the Figure 4.51.

The defect probability generation reported in both Figure 4.53 and 4.54 shows a net alignment of the data. The clear trend obtained permits also to derived the power law exponent of P_g as a function of the voltage below the saturation regime. For nFET stressed in inversion, the exponent alpha is totally matching the reported value from Wu [24]. Nevertheless a distinction can be seen here between the nFET stressed under inversion regime and under accumulation. First of all, a lower defect probability generation rate is found for the accumulation polarity with an exponent α around 32. Then exactly as for TDDB or Q_{BD} data, the saturation regime concerning nFET devices seems to occur more around 5 V for the inversion polarity in comparison of accumulation stresses where a value about 6 V can be observed.

The charge-to-breakdown of nFET devices at 4 V and 5 V are reported as a function of the oxide thickness in both stress polarities in Figure 4.55. A really good matching of the cell-based model [24] is observed with the data presented here. Experimentally, the breakdown data obtained at 4 - 5 V for the thin oxides presented are corresponding to the range of the ESD characterization. This underlines the similitude of the physical oxide degradation mechanisms between the classical reliability and the ESD types of events.

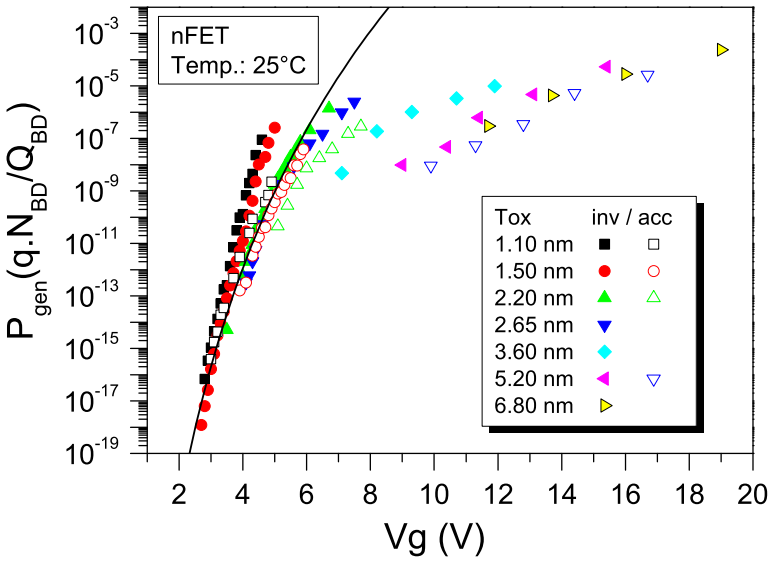

Figure 4.52: Gate voltage dependence of the defect generation probability (P_g) for nFET stressed at room temperature in inversion and in accumulation regime based on the experimental values of a_0 = 2.17 nm and T_{int} = 1.1 nm.

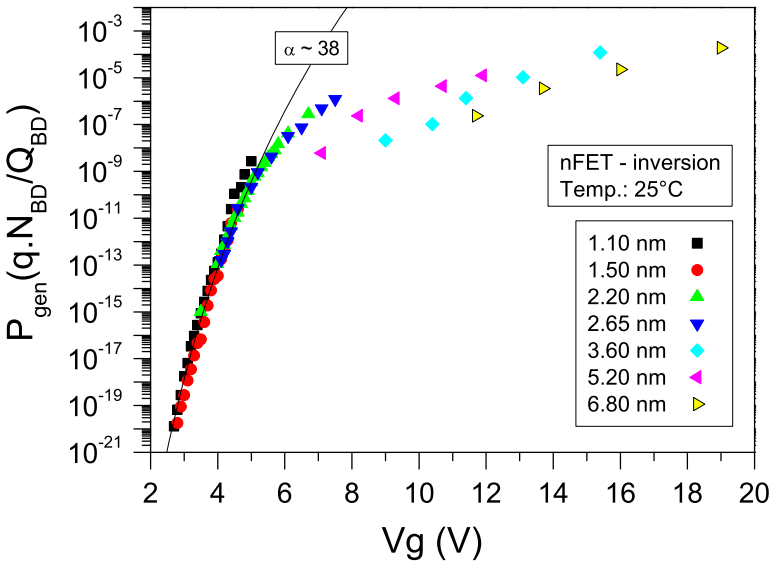

Figure 4.53: Gate voltage dependence of the defect generation probability (P_g) for nFET stressed in inversion at room temperature and based on a_0 = 1.83 nm and T_{int} = 0.37 nm.

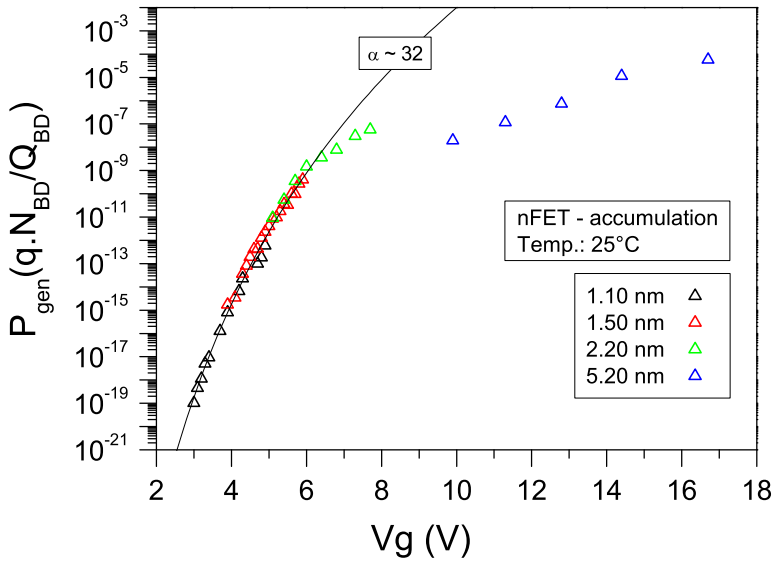

Figure 4.54: Gate voltage dependence of the defect generation probability (P_g) for nFET stressed in accumulation at room temperature and based on a_0 = 1.83 nm and T_{int} = 0.37 nm.

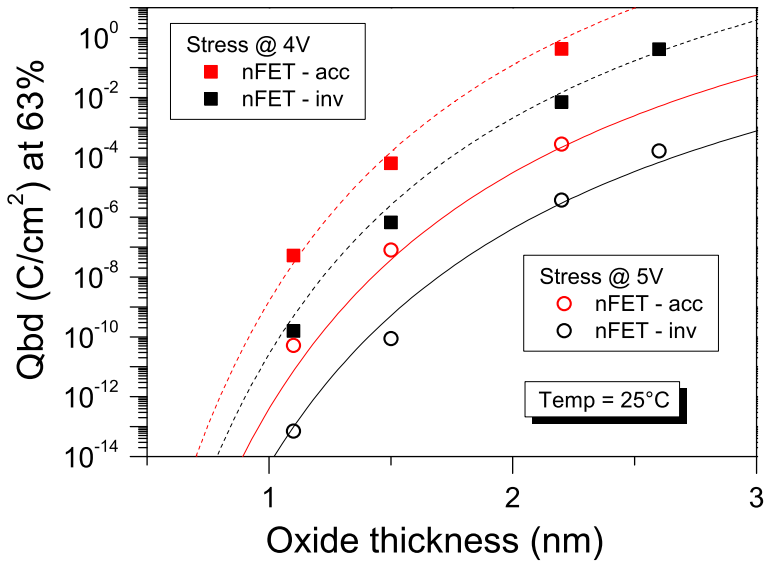

Figure 4.55: Charge to breakdown (Q_{BD}) of nFET devices as a function of the oxide thickness for two selected voltages under the TDDB acceleration voltage kink.

This really good agreements to the cell-based model implies that the experimental $\beta(T_{ox})$ trend extracted suffers from uncertainties in the estimation of the β values. In the experiments driven for the TDDB voltage acceleration law, the focus was on the extraction of the time-to-fail at the $T_{63\%}$ criterion were the experimental extraction of the time to fail has the best confidence level. The studies done have covered the verification of the Poisson statistics but the intensive characterization work required for the precise determination of beta has not been performed for all oxide thicknesses. The estimation of this parameter needs really a very large samples size for each thickness [25, 24]. The absolute needs of high statistical data is once again pointed out and emphasis the uncertainty of the methodology used up to now for the evaluation of the gate oxide breakdown in the ESD time domain.

4.3.8 Process impact on TDDB and singularity of the 1.5 nm nFET

A particular behavior has been observed for the 1.5 nm nFET TDDB at 25 °C [23], namely a high voltage acceleration (around 64) is reported above 4 V. This deviation from the straight power-law is unexpected and critical for the ESD development as HBM and CDM breakdown voltages expected for thin oxides as 1.5 nm or below are exactly in the region between 4 V and 5 V where this GOX breakdown acceleration voltage is maximal. This kink has been observed with different set-ups, testing methodologies, different test structures (geometry and also nCAP), different wafers and hardware processed from different fabrication locations. There is no doubt that a physical effect is taking place triggered at 4 V.

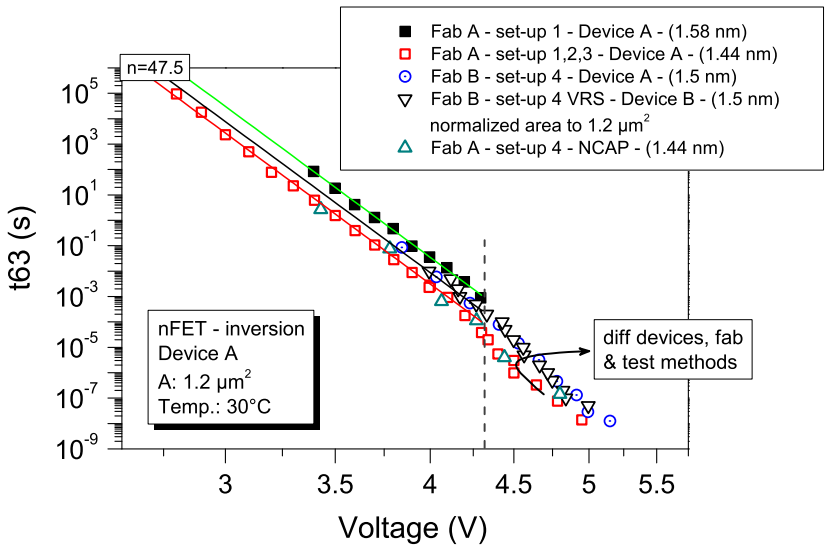

Figure 4.56: Voltage acceleration of 1.5 nm oxide nFET with its particular kink occurring around 4 V.

The origin of the kink was thought at first glance to come from self-heating generated dur-

ing the ESD stress. As the room temperature TDDB seems to converge towards the elevated temperature power-law (Figure 4.37), it was a quite obvious hypothesis. The power density obtained for the 4 V to 5 V stress domain is in the range of few mW/μm^2. In the case where the temperature dissipation is homogeneously distributed along the whole gate oxide area, calculations performed estimate the self-heating to around 10 K for these conditions. Moreover, this result confirms a previous study [9] where it was also demonstrated that self-heating do not have to be integrated in the TDDB acceleration. No satisfying explanation has been yet found for this critical kink phenomenon and this is still an open question up to now which is beyond the scope of this study. The power-law is an empirical model and the physical explanation of the exponent coefficient responsible for the voltage breakdown driven acceleration is not yet clearly understood. It is possibly resulting from different breakdown mechanisms which are voltage activated. For nFET devices, one transition regime has been identified around 5 V, but it is not to exclude that other dominant mechanisms in ultra-thin GOX could be triggered earlier around 4 V (multi-vibrational bonds modes, ...). Due to the set-up sensitivity no data could be obtained above 5 V. As expected from the general behavior and starting from this limit, a power-law with a reduced acceleration factor around 30 is expected.

Unexpected GOX TDDB behaviors could also nevertheless be observed in some special cases related to process steps and hydrogen containments. One interesting effect of the BEOL impact on the NBTI degradation and TDDB operating mechanisms has been reported by Pompl *et al.* in [73] for thick oxides. The gate oxide reliability seems to be driven by the weakest mechanism from many degradation mechanisms operating in parallel. For this particular case of the 1.5 nm nFET a possible process step could activate a more sensitive mechanism that the one which is normally involve in the breakdown. From the entire oxides tested in this work, this 1.5 nm oxide is the only one processed In-situ Steam Generation (ISSG) in a wet atmosphere and possibly the impact of the hydrogen concentration involved in this methodology could be responsible for this weakest behavior.

4.4 Conclusion

In this chapter we have presented an intensive and extensive thin oxide hard breakdown characterization towards the ESD time domain for the extraction of physically based TDDB models. This is the first time that a pure continuous experimental study is used for this topic through a large thickness range. The methodology presented has allowed to demonstrate the following new results:

- The Gate oxide TDDB in the ESD regime follows the Weibull statistics and matches the percolation theory scheme.
- Intrinsic gate oxide TDDB under uniformed ESD like stress is a random process occurring on the whole gate oxide area and which can be described by the Poisson statistics.

- Methodology to accurately determine the TDBB in the ESD regime lies on statistics and methodology choice.
- Proper gate oxide test structures are inevitable for accurate measurements. Intrinsic breakdown failures requires small gate oxide area and an optimized layout to avoid any parasitics voltage drops in the device.
- The reported ESD post breakdown mode for small GOX active areas is always hard, except for thin pFET stressed in accumulation where soft breakdowns could occur.
- Thin GOX TDDB voltage acceleration is following a power law model from the long time stress to the nanosecond regime. A perfect matching in the TDDB law and model with reliability proofs the continuity of the same operating mechanisms. It was shown that the breakdown of thin GOX in the ESD range can be accurately described by power law behaviors even down to 1.1 nm oxides. GOX breakdown projection in the ESD regime towards the limit of SiO_2 (or SiO_2N) dielectric thickness is reported without any deviation.
- The TDDB depends not only on the area but also on the device type, stress polarity and temperature. Identification of ESD worst case (nFET in inversion) and precise devices ESD robustness hierarchy with respect to the stress polarity is provided quantitatively.
- In the ESD regime, the impact of the temperature on the TDDB is becoming negligible for thin oxides. The voltage acceleration of the time-to-fail is decreasing with the temperature elevation with a very low activation energy in the high voltage range.
- Accordingly to the parameters characterized, the GOX breakdown can be predicted in the ESD regime via the TDDB power law extrapolation of the process qualification data.

The impact of this new findings on the ESD protection development and design will be discussed in detailed in the last chapter. Nevertheless, the thin oxide hard breakdown which was the focus up to now, is not the only failure modes and the oxide degradation under non destructive ESD stress is also a major issue to studied. This will be the focus of the next chapter.

Bibliography

[1] Cheung K.P. Plasma-charging damage and ESD, help each other? In *EOS/ESD Symposium*, pages 38–42, 1999.

[2] Stathis J.H. and DiMaria D.J. Reliability projection for ultra-thin oxides at low voltage. *International Electron Device Meeting*, pages 167–170, 1998.

[3] Ministry K.R., Krakauer D.B., and Doyle B.S. Impact of snapback-induced hole injection on gate oxide reliability of N-MOSFET's. *Electron Device Letters*, 11(10), 1990.

[4] Cheung K.P. Impact of ESD protection device trigger transient on the reliability of ultra-thin gate oxide. *Microelectronics Reliability*, 41:745–749, 2001.

[5] Tunnicliffe M.J., Dwyer V.M., and Campbell D.S. Experimental and theoretical studies of ESD/ESD oxide breakdown in unprotected MOS structures. *EOS/ESD Symposium*, pages 162–168, 1990.

[6] Tunnicliffe M.J., Dwyer V.M., and Campbell D.S. The integrity of gate oxide related to latent failures under EOS/ESD conditions. *Journal of Electrostatics*, 31:91–110, 1993.

[7] Beebe Stephen G. *Characterization, modeling, and design of ESD protection circuits*. PhD thesis, Stanford University, 1998.

[8] Leroux C., Andreucci P., and Reimbold G. Analysis of oxide breakdown mechanism occurring during ESD pulses. In *International Reliability Physics Symposium*, pages 276–282, 2000.

[9] Wu J., Juliano P., and Rausenbaum E. Breakdown and latent damage of ultra-thin gate oxides under ESD stress conditions. In *EOS/ESD Symposium*, 2000.

[10] Wu J. and Rausenbaum E. Gate oxide reliability under ESD-like pulse stress. *IEEE Transactions on Electron Devices*, 51(7), 2004.

[11] Weir B.E., Leung C.C, Silverman P.J., and Alam M.A. Gate dielectric breakdown: a focus on ESD protection. In *International Reliability Physics Symposium*, 2004.

[12] Weir B.E., Leung C.C, Silverman P.J., and Alam M.A. Gate dielectric breakdown in the time-scale of ESD events. *Microelectronics Reliability*, 45:427–436, 2005.

[13] Matsuzawa K., Satake H., C. Sutou, and Kawashima H. Gate oxide reliability under ESD-like pulse stress. In *Simulation of Semiconductor Processes and Devices*, pages 129 – 132, 2003.

[14] Röhner M., Kerber A., and Kerber M. Voltage acceleration of TBD and its correlation to post breakdown conductivity of n- and p- channel MOSFETs. *International Reliability Physics Symposium*, pages 76–85, 2006.

[15] Nicollian P.E., Krishnan T., Chancellor C.A., and Khamankar R.B. The traps that cause breakdown in deeply scaled SiON dielectrics. In *International Electron Device Meeting*, pages 1–4, 2006.

[16] Kerber A., Rompl T., Duschl R., and Kerber M. Lifetime prediction for CMOS devices with ultra thin gate oxides based on progressive breakdown. In *International Reliability Physics Symposium*, pages 217–220, 2007.

[17] Martin A., O'Sullivan P., and Mathewson A. Dielectric reliability measurement methods: a review. *Microelectronics Reliability*, 38(1):37–72, 1998.

[18] Degraeve R., Kaczer B., and Groeseneken G. Degradation and breakdown in thin oxides layers: mechanisms, models and reliability prediction. *Microelectronics Reliability*, 39:1445–1460, 1999.

[19] Pompl T. and Röhner M. Voltage acceleration of time-dependent breakdown of ultra-thin gate dielectrics. *Microelectronics Reliability*, 45:1835–1841, 2005.

[20] Wolters D.R. and Verwey J.F. Breakdown and wearout phenomena in Si/SiO2 films. *Instabilities in Silicon Devices*, pages 315–362, 1986.

[21] Nigam T., Degraeve R., Groeseneken G., and Heyns M.M. andd Maes H.E. Constant current charge-to-breakdown: still a valid tool to study the reliability of MOS structures? *International Reliability Physics Symposium*, pages 62–69, 1998.

[22] Wu E.Y., Abadeer W.W, Han L., Lo S., and Hueckel G.R. Challenges for accurate reliability projections in the ultra-thin oxide regime. *International Reliability Physics Symposium*, pages 57–65, 1999.

[23] Ille A., Stadler W., Kerber A., Pompl T., Brodbeck T., Esmark K., and Bravaix A. Ultra-thin gate oxide reliability in the ESD time domain. In *EOS/ESD Symposium Proceedings*, pages 285–294, 2006.

[24] Wu Ernest Y. and Suñé Jordi. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45:1809–1834, 2005.

[25] Wu E.Y. and Vollertsen R.P. on the weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination - part I: Theory, methodology, experimental techniques. *Transaction of Electron Devices*, 49(12):2131–2140, 2002.

[26] Pio F. Sheet resistance and layout effects in accelerated tests for dielectric reliability evaluation. In *Microelectronics Journal*, volume 27, pages 675–685, 1996.

[27] Trémouilles D., Thijs S., Roussel Ph., Natarajan M.I., Vassilev V., and Groeseneken G. Transient voltage overshoot in TLP testing - real or artifact? In *EOS/ESD Symposium*, 2005.

[28] Ribes G., Rafik M., Barge D., Kalpat K., Denais M., Huard V., and Roy D. New insights on percolation theory and the origin of oxide breakdown thickness and process deposition dependence. In *International Reliability Physics Symposium*, pages 578–579, 2007.

[29] Russ C., Mergens M., Verhaege K., Armer J., Jozwiak P., Kolluri G., and Avery L. GGSCRs: GGNMOS triggered silicon controlled rectifiers for ESD protection in deep sub-micron CMOS process. In *EOS/ESD Symposium*, 2001.

[30] Kinoshita A., Mitani Y., Matsuzawa K., Kawashima H., Sutoh C., Kurihara J., Hiraoka T., Hirano I., Muta M., Takayanagi M., and Shigyo N. In *Breakdown Voltage Prediction of Ultra-Thin Gate Insulator in Electrostatic Discharge (ESD) Based on Anode Hole Injection Model*, pages 623–624, 2006.

[31] Tseng J.C. and Hwu J.-G. Effects of electrostatic discharge high-field current impulse on oxide breakdown. *Journal of Applied Physics*, 101:014103–1–6, 2007.

[32] Riess Philip. *Etude de la fiabilité des oxydes minces: analyse des mécanismes de transport et de génération du SILC*. Physique des composants a semiconducteurs, Institut National Polytechnique de Grenoble, 1999.

[33] Tan Y.N., Chim W.K., and Lim P.S. Stress-induced leakage current in thin oxides under high-field impulse stressing. In *International Symposium on the Physical & Failure Analysis of Integrated Circuits*, pages 228–233, 2001.

[34] Kerber A., Pompl T., Röhner M., Mosig K., and Kerber M. Impact of failure criteria on the reliability prediction of CMOS devices with ultrathin gate oxides based on voltage ramp stress. *Electron Device Letters*, 27(7):609–611, July 2006.

[35] Pio F., Ravazii L., and Riva C. Series resistance effects in thin oxide capacitor evaluation. In *Electron Device Letters*, volume 13, pages 544–546, 1992.

[36] Pio F., Ravazii L., and Riva C. Influence of series resistance in oxide parameter extraction from accelerated tests data. In *Microelectronics Journal*, volume 24, pages 445–451, 1993.

[37] Kerber A., Röhner M., Wallace C., O'Riain L., and Kerber M. Wafer-level gate-oxide reliability towards ESD failures in advanced CMOS technologies. *Transaction on Electron Devices*, 53(4):917–920, 2006.

[38] Wu E.Y., Vayshenker A., Nowak E., Suñé J., Vollertsen R-P., Lai W., and Harmon S. Experimental evidence of Tbd power-law for voltage dependence of oxide breakdown in ultrathin gate oxides. *Transaction on Electron Devices*, 49(12):2244–2253, 2002.

[39] DiMaria D.J, Buchanan, Sthatis J.H, and Stahlbush R.E. Interface states induced by the presence of trapped holes near the silicon-silicon-dioxide interface. *Journal of Applied Physics*, 77(5):2032–2040, 1995.

[40] Bude J.D., Weir B.E., and Silverman P.J. Explanation of stress induced damage in thin oxides. In *International Electron Device Meeting*, pages 179–182, 1998.

[41] Nicollain P.E., Hunter W.R., and Hu J.C. Experimental evidence for voltage driven breakdown models in ultrathin gate oxides. In *International Reliability Physics Symposium*, pages 7–15, 2000.

[42] McPherson J.W. and Baglee D.A. Acceleration factors for thin gate oxide stressing. *International Journal of High Speed Electronics and Systems*, pages 1–5, 1985.

[43] Chen I.C., Holland S., and Hu C. A quantitative physical model for time-dependent breakdown in Si/SiO2. In *International Reliability Physics Symposium*, pages 24–31, 1985.

[44] Kimura M. Field and temperature acceleration model for time-dependent dielectric breakdown. *Transaction on Electron Devices*, pages 220–229, 1999.

[45] Mizubayashi W., Yoshida Y., Miyazaki S., and Hirose M. Quantitative analysis of oxide voltage and field dependence of time-dependent dielectric soft breakdown and hard breakdown in ultrathin gate oxides. *Japanese Journal of Applied Physics*, pages 2426–2430, 2002.

[46] Schuegraf K.F. and Hu C. Hole injection oxide breakdown model for very low voltage lifetime extrapolation. In *International Reliability Physics Symposium*, pages 7–12, 1993.

[47] Duschl R. and Vollertsen R-P. Is the power-law model applicable beyond the direct tunneling regime? *Microelectronics Reliability*, 45:1861–1867, 2005.

[48] DiMaria D.J. Electron energy dependence of metal-oxide-semiconductor degradation. *Applied Physics Letters*, 75(6):2427–2428, 1999.

[49] Shen T.-C., Wang C., Abeln G.C., Tucker J.R., Lyding J.W., Avouris Ph., and Walkup R.E. Atomic-scale desorption through electronic and vibrational excitation mechanisms. *Science*, 268:1590–1592, 1995.

[50] Ille A., Stadler W., Gossner H., Brodbeck T., Pompl T., and Bravaix A. Thin gate oxides time-to-breakdown in the ESD time domain and consequences for applications. In *International Electrostatic Discharge Workshop*, pages 176–186, 2007.

[51] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. In *EOS/ESD Symposium Proceedings*, pages 328–337, 2007.

[52] Pompl T., Kerber A., Röhner M., and Kerber M. Gate voltage and oxide thickness dependence of progressive wear-out of ultra-thin gate oxides. *Microelectronics Reliability*, 46:18603–1607, 2006.

[53] Miranda E., Suñé J., Rodriguez R., Nafria M., and Aymerich X. Detection and fitting of the soft breakdown failure mode in MOS structures. *Solid-State Electronics*, 43:1801–1805, 1999.

[54] Miranda E. and Suñe J. Gate oxide reliability for nano-scale CMOS. *Microelectronics Reliability*, 44(1), 2004.

[55] Pompl T., Wurzer H., Kerber M., Wilkins R.C., and Eisele I. Influence of soft breakdown on NMOSFET device characteristics. In *International Reliability Physics Symposium*, 1999.

[56] Guitard N., Trémouilles D., Alves S., Bafleur M., Beaudoin F., Perdu P., and Wislez A. ESD induced latent defects in CMOS ICs and reliability impact. In *EOS/ESD Symposium*, pages 174–181, 2004.

[57] Mason P.W., La Duca A.J., Holder C.H., Alam M.A., and Hwang D.K. A methodology for accurate assessment of soft-broken gate oxide leakage and the reliability of VLSI circuits. In *International Reliability Physics Symposium*, pages 430–434, 2004.

[58] Stathis J. Impact of ultra thin oxide breakdown on circuits. In *International Conference on Integrated Circuit and Technology*, pages 123–127, 2005.

[59] Avouris Ph., Walkup R.E., Rossi A.R., Akpati H.C., Nordlander P., Shen T.-C., Abeln G.C., and Lyding J.W. Breaking individual chemical bonds via STM-induced excitations. *Surface Science*, 363:368–377, 1996.

[60] DiMaria D.J, Cartier E., and Arnold D. Impact ionization, trap creation, degradation, and breakdown in silicon dioxide films on silicon. *Journal of Applied Physics*, 73(7):3367–3384, 1993.

[61] Nicollian P.E, Krishnan T., Chancellor C.A, Khamankar R.B., Chakravarthi S., Bowen C., and Reddy V. The current understanding of trap generation mechanisms that leads to the power law model for gate dielectric breakdown. In *International Electron Device Meeting*, pages 197–206, 2007.

[62] Suñé J. and Wu E.Y. A quantitative two-step hydrogen model of Si/SiO2 gate oxide breakdown. *Solid-State Electronics*, 46:1825–1837, 2002.

[63] Ribes G., Bruyère S., Denais M., Monsieur F., Huard V., and Roy D. Ghibaudo G. Multi-vibrational hydrogen release: Physical origin of tbd, qbd power-law voltage dependence of oxide breakdown in ultra-thin gate oxides. *Microelectronics Reliability*, 45:1842–1854, 2005.

[64] Pompl T., Kerber M., Wurzer H., and Eisele I. Contribution of interface traps to valence band electron tunneling in pmos devices. In *ESSDERC*, 2000.

[65] Pompl Thomas. *Gateisolatoren für MOS-Feldeffekttransistoren*. Institut für physik, Universität der Bundeswehr München, 2000.

[66] Suehle J.S., Vogel E.M., Wang B., and Bernstein J.B. Temperature dependence of soft breakdown and wear-out in sub-3 nm SiO2 films. In *International Reliability Physics Symposium*, pages 33–39, 2000.

[67] Wu E.Y., McKenna M., Lai W., Nowak E., and Vayshenker A. The effect of change of voltage acceleration on temperature activation of oxide breakdown for ultrathin oxides. *Electron Device Letters*, 22(12):603–605, 2001.

[68] Wu E.Y., Suñé J., Lai W., Nowak E., McKenna J., Vayshenker A., and Harmon S. Interplay of voltage and temperature acceleration of oxide breakdown for ultra-thin gate oxides. *Solid-State Electronics*, 46, 2002.

[69] DiMaria D.J. and J.H. Stathis. Non-arrhenius temperature dependence of reliability in ultrathin silicon dioxide films. *Applied Physics Letters*, 74(12):1752–1754, 1999.

[70] Degraeve R., Groeseneken G., Bellens R., Depas M., and Maes H.E. A consitent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides. *Annual International Reliability Physics Symposium*, pages 863–866, 1995.

[71] Stathis J. H. Percolation models for gate oxide breakdown. *Journal of applied physics*, 86(10):5757–5765, 1999.

[72] Suñé J. and Wu E.Y. Mechanisms of hydrogen release in the breakdown of SiO2-based gate oxides. *International Electron Device Meeting*, pages 399–402, 2005.

[73] Pompl T., Allers K.-H., Schwab R., Hofmann K., and Röhner M. Change of acceleration behavior of time-dependent dielectric breakdown by the BEOL process: Indications for hydrogen induced transition in dominant degradation mechanism. In *International Reliability Physics Symposium*, pages 388–397, 2005.

Chapter 5

Gate oxide to device degradation under ESD

5.1 Device reliability

5.1.1 Device reliability and ESD crossed interactions

In the operating conditions of ICs, the electrical conditions applied burden the oxide to different phases of stresses, non conducting stress, gate biased stress (BTI), hot carrier stress(HC). On the long term, the impact of these altering stresses are altering the oxide. Interface states and volume traps are generated during the stresses and impact directly the transistor parameters and device functionality. The guarantee of IC's functional integrity for a targeted lifetime require to learn about the degradation mechanisms and their impacts on the device parameters (I_{off}, V_{th}, G_{m}, I_{on}, ...). With the down-scaling of technologies, without the corresponding reduction in the supplies voltage, the electrical fields in the devices have strongly increase (Figure 5.1). The field across the oxide is reaching the 10 MV/cm for the ultra thin 1.1 nm nitrided oxides with a lateral field in the range of 250 kV/cm. The high field in the oxide lead to dielectric reliability problems and induced also an extreme high leakage level, see Figure 5.1. The exponential increase of the gate leakage as a function of the oxide thickness can lead to the end of the MOS integration based on silicon dioxide dielectrics towards 1 nm. The vertical field in the silicon gave rise to hot carriers degradation issues.

Although the aggressive down-scaling of the technologies leads to more and more challenging reliability issues, the device lifetime specification must be kept. To counter act a too fast degradation of the device performances within the time, technological solutions as implants or new process steps are developed. An efficient manner to reduce the amount of physical damage generated during HC stress is to improve the oxide quality for example by the nitridation of thin oxides or the incorporation of fluorine in the oxide or replacing the hydrogen with deuterium.

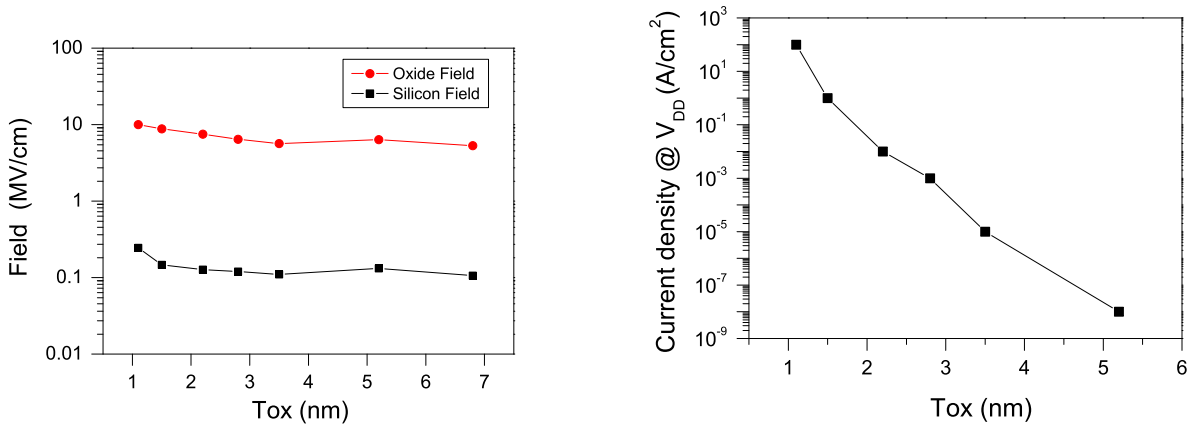

Figure 5.1: Impact of the non-proportional scaling of the supply voltage with the oxide thickness reduction pushed by the ICs integration run: (left) evolution of the field in the oxide and in the silicon as a function of the oxide thickness, (right) increase in the gate leakage current with the oxide thickness.

The second solution to improve the HC robustness is to proceed to a drain engineering. However, the ESD robustness is strongly process sensitive but is considered only to a lower level in the technologies development. Process modifications can be highly critical for the ESD protection requirements, it has been seen in the past with the introduction of LDD junction and salicide or with products technology transfers [1, 2, 3, 4]. The appearance of mixed voltage applications, which challenge the ESD protection complexity with new problems, is also one consequence of hot carrier issues combined with the on-chip power dissipation. The hot carrier issues are globally responsible of a restriction of the ESD robustness capability and required also an extra consideration for the ESD protection design. Beyond the protection of ICs to hard failures, the ESD protection design should also ensure the non-degradation of the circuits' reliability. In this challenging context, the impact of non-destructive ESD stress on the protected circuits needed to be investigated to settle a clean ESD design and avoid the possibility of problematic latent damage occurrence.

5.1.2 ESD and reliability focus

The issue of reliability crossed with ESD stresses has always been an important point of focus and has been extensively studied in the past for thick oxides. Strong motivations were based on the observation of some latency effects induced by ESD stress on output drivers.

5.1.2.1 Outputs

The big pull-down nFETs at the output drivers designed to provide a powerful signal can also be used as ESD self-protection elements in the ggNMOS configuration. In this case it is essential that the electrical as well as the reliability properties of the drivers do not change.

The output drivers degradation studies began with the observation of ggNMOS under TLP stress where the snap-back regime has been reported to induce some oxide defects. Some injection and trapping of electron causing an I_{off} increase has been first reported in [5]. Then some hole charges trapping and interface states creation were also reported [6, 2, 7]. This oxide degradation observation under the snap-back condition lead to further studies to evaluate the impacts of these defects on the reliability. Most of the investigations were performed by injection of current pulses at the drain to force the snap-back and observe its effects on the oxide and device parameters [8, 6, 9, 2, 10, 11, 12]. An increase in the off drain current was reported without any device functional fails [5, 2, 10, 11, 12]. This effect was called "soft fail" in comparison to hard fails causing the irreversible total loss of device control. The nature of this leakage increase has been attributed to GIDL effect caused by localized electron trapping [11], fragmentation between source drain [2, 9, 8] or molten silicon under the drain-to-gate overlap [13, 14, 12]. The latency of this soft fails have been discussed with different lifetime tests. The consequent impact of the ESD stress on HC lifetime degradation has been reported for thick oxides (25 nm and 8 nm) at device level [8] and on 64 Mb DRAM product level [7]. TDDB test have also shown a reduction of the lifetime for a 19 nm thick oxide [10] so as High Temperature Operating Life (HTOL) in a 0.35 µm technology node [12].

In the case of thin oxide ggNMOS, the failing mechanism occurring is not the drain-to-source fragmentation, but the oxide breakdown at the drain overlap side [15]. In this case the oxide degradation of the protection device matters. The ageing of the structures by hot electron stress have been reported to impact the robustness of the thin oxides protection elements (1.5 nm and 2.2 nm) I_{t2} [16, 17]. Contrarily, it is not the case when the failure mode is driven by filamentation [8, 16]. An acceleration in the transconductance degradation under post CVS stress of about a factor of 2 as been reported for a 2.5 nm nFET devices pre-ESD stressed in the ggNMOS configuration by Cester *et al*. [18]. In case of thin pFET drivers it has been reported by Mishra *et al.* [19] that short channel ggPMOS stressed to 50% of the I_{t2} value suffer more under NBTI than fresh devices.

5.1.2.2 Inputs and decoupling capacitors

The critical configurations for oxide devices during an ESD stress are the input stages, where thin oxides can be connected, and decoupling capacitors which are directly exposed to the stress between supply and ground lines. During an ESD event, the failing parts are not exclusively restricted to protection elements, the core composed of thin oxides can be also hit. In advanced

CMOS technologies below the 100 nm nodes, ggNMOS are not really popular anymore and some protection devices without oxides likes SCRs, bipolars or diodes are mainly used. The degradation of protection elements itself is not an issue, but the turn-on behavior and the clamping voltage of the protection concept impacting the circuitry behind the protection circuit is now coming the main reliability issue.

5.2 Oxide and device degradation under ESD

In this section, the degradation of the dielectric and its impact on the device integrity will be discussed. It basically depends strongly on the oxide thickness. The reliability of nFET devices under inversion ESD stress will be the focus. Actually, this case was exposed in the previous chapter to be the critical one in term of robustness. During an ESD event, the probability to have this concrete case is also very high, the direct GOX input buffers and capacitors between the supplies could be hit by such stresses. Furthermore, devices stressed as capacitor bring also a "knowledge" about output degradation as the homogeneous oxide stress can be seen as the worst case configuration.

Oxides under electrical stress degrade as a results of oxide traps or/and interface generation. These defects degenerate the device parameters and can lead to circuit malfunction if the drift in the parameter is too large. In a first step, the effect of the trapped charges and interface traps on the device characteristic will be recall.

The threshold voltage is one of the most sensitive parameters influenced by the oxide defects, the charges trapped in the oxide volume directly impact the threshold voltage of a MOS transistor by modifying the flat band voltage (see Section 1.1.3.2),

$$\text{From equation 1.9, } \Delta V_{\text{th,Not}} = -\frac{Q_{\text{ox}}}{C_{\text{ox}}} \tag{5.1}$$

Positive charges in the oxide will consequently lower the threshold voltage, whereas negative charges will increase it. The charge in the volume of the oxide can be described by a density of charges located at a distance z of the oxide / silicon interface,

$$\Delta V_{\text{th,Not}} = -\frac{\int_0^{T_{\text{ox}}} Q_{\text{ox}}(z)\left(1-\frac{z}{T_{\text{ox}}}\right)dz}{C_{\text{ox}}} \tag{5.2}$$

For thick oxides, the charges could be located deep into the oxide contrarily as for thin oxides were the location of the traps is lower to the direct tunneling distance ($z < 3$ nm) implying the de-trapping of the carriers by DT.

The interface traps impacts also the threshold voltage depending on their energy,

$$\Delta V_{\text{th,Nit}} = -\frac{q \cdot N_{\text{it}}(\Psi_{\text{S}} - \Phi_{\text{F}})}{C_{\text{ox}}} \tag{5.3}$$

The interface traps are characterized via their traps density,

$$D_{\text{it}} = \frac{1}{q}\frac{dQ_{\text{it}}}{dE} \tag{5.4}$$

The interface degradation can be monitored also though the swing deviation ΔS expressed as a function of the interface trap density (D_{it}),

$$\Delta S = \frac{kT}{q} ln(10)\,\frac{qD_{\text{it}}}{C_{\text{ox}}} \tag{5.5}$$

However, in case of non-uniformed interface degradation, this method is not reliable as one localized charge has a similar impact as an uniformed interfacial charge widely spread along the channel [20]. Interface traps are commonly characterized via the charge pumping technique [21, 22]. For the mid-gap condition ($\Psi_{\text{S}} = \Phi_{\text{F}}$), the donor states are filled and acceptor states are empty and consequently no interface charges are observable for this condition $\Delta V_{\text{th,Nit}}(\Phi_{\text{F}}) = 0$.

The degradation nature of the gate oxide is cumulative (Chapter 4) and the question arises, what would be the impact of non-destructive ESD stress on a device? A lot of work has already been done in that field for thick oxides (mainly on outputs) [23, 8, 24, 6, 25, 26, 27, 28, 29, 9, 2, 30, 10, 13, 7, 31, 11, 32, 14, 31, 33, 12, 34, 17, 35], but there is very little work on thin oxides below 3 nm [36, 16, 17, 37, 38, 19]. To answer this question, positive rectangular pulse stresses (100 ns) were applied to nFET gate transistors with drain, source and bulk grounded (capacitors in inversion).

One has to distinguish between thin oxides ($T_{\text{ox}} < 2.5$ nm), medium oxides (2.5 nm to 7 nm) and thick oxides ($T_{\text{ox}} > 7$ nm). For thick and medium oxides, it is possible to monitor a straight and clear influence of the ESD stress on device characteristics due to a strong trapping, while it is not the case anymore in the thin oxides where the direct tunneling of the oxide traps occurs. The conventional monitoring of the oxide and device parameters degradation induced by an ESD event will be presented for a large range of oxide thickness to distinguish the particularity behavior of each thickness domain.

5.2.1 Thick Oxides ($T_{\text{ox}} > 7$ nm)

As a starting point, the impact of non-destructive ESD stress on thick oxide devices will be presented in order to better understand the difference in the degradation behavior between thick and thin oxides. Three different oxide thicknesses and technologies have been investigated (27.5 nm, 16 nm, 8.7 nm). For all of them, prior to moderate ESD-like stress, an evaluation of the breakdown voltage was performed by means of VRS. For that thick oxides, the spreading of the distribution is very narrow due to large Weibull slope values. The breakdown value corresponds to the breakdown voltage obtained for 63% of the samples.

5.2.1.1 Thick oxide, $T_{ox} = 27.5$ nm

The breakdown voltage (V_{bd})of the investigated structure (area of 60 μm^2) has revealed a clear hard breakdown value around 36 V. A stress close to the breakdown voltage leads to a very strong oxide trapping, easily observable in the device characteristic. An example of the device parameters degradation for a 100 ns pulse at 85 % of V_{bd} is exposed in Figure 5.2 and 5.3. A parallel and uniformed shift in the sub-threshold slope and in the saturated drain current was observed towards negative values. This is a result of a flat band voltage shift induced by positive charges trapped into the oxides. Figure 5.3 shows the extraction of the saturated threshold voltage, a shift than more than 1 V can be easily seen. The transistor is in a punch through state and consequently always on, leading to a strong leakage in the I_{off}.

5.2.1.2 Thick oxide, $T_{ox} = 16$ nm

For the 10 μm^2 nFET stressed in inversion, the breakdown value obtained by means of VRS is 25.5 V. A moderate ESD stress to 85% of V_{bd} leads to the same trend as for the 27.5 nm oxide. A strong positive trapping into the oxide is observed uniformly along the channel. The impact of this traps on the I_{ds} -V_{ds} curve for different V_{gs} bias is plotted in Figure 5.4. A drastic increase in the I_{on} current can be noticed. The clear impact on the drain saturated current and transconductance is also shown in Figure 5.5.

As depicted in this part, the impact of short pulses applied on thick gate oxides have a consequent impact on the transistor characteristics. These large effects have been reported for thicker oxides than 7 nm with *IV* characterization, and mainly for capacitor devices via *C-V* in several publications [23, 6, 27, 28, 2, 10, 32, 31, 33]. Concerning the case of thick nFET devices stressed in inversion, the data exposed show in all cases a net positive oxide trapping resulting from the pre-ESD stress. It cannot be excluded that a part of neutral traps were also generated from the stress. This kind of traps have been evidenced with a post CVS step which allows to refill the neutral traps with electrons [32, 31]. During the ESD stress not only oxide traps are created, border traps [35] and interface traps [2, 34] are also generated.

5.2.1.3 Relaxation effect and parameter recovery

The drifts in parameters induced by the ESD charge trapping is however partially compensated by a relaxation effect in time. Carriers trapped in the bulk oxide or near the interface responsible of device characteristic drifts could be de-trapped with time or annealed at high temperature [10, 32, 33, 30]. In the relaxation process observed at room temperature, two different time constants are revealed (Figure 5.6). Initially, the superficial and border traps are de-trapped and then the deeper level traps are relaxing much more slowly. The de-trapping of the interface and border traps are observable in Figure 5.6 in the recovery effect of the I_{off} parameter plotted a linear scale

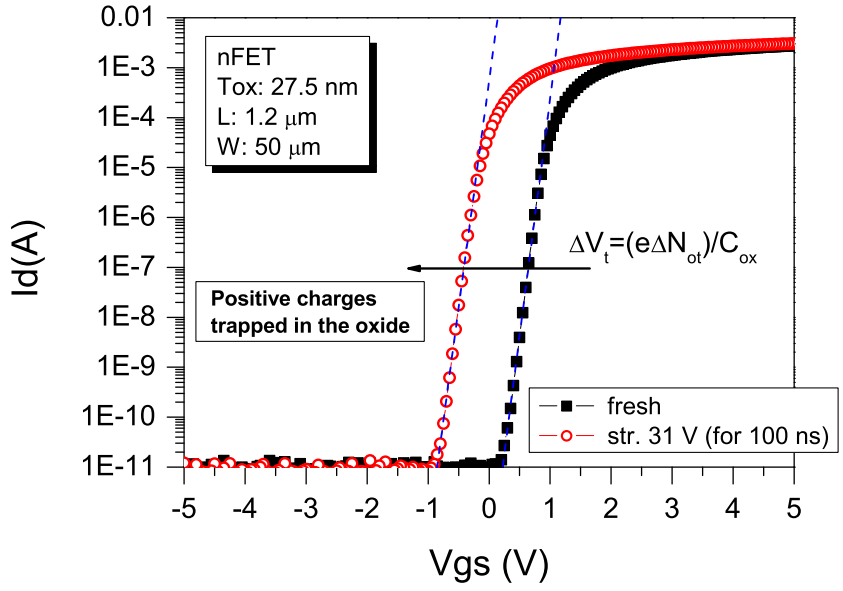

Figure 5.2: Subthreshold degradation of a 27.5 nm thick oxide nFET induced by a 100 ns stress at 85 % of V_{bd}.

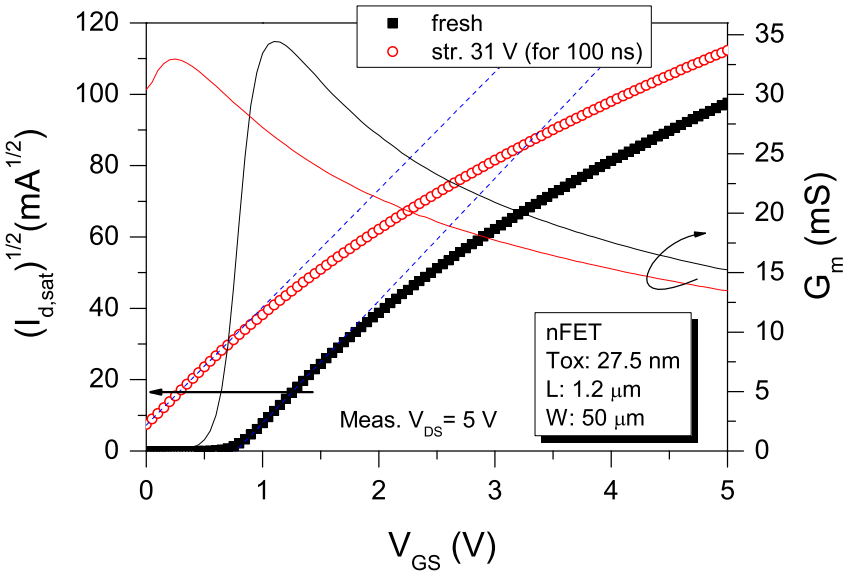

Figure 5.3: Saturated threshold voltage and transconductance of a 27.5 nm thick oxide nFET induced by a 100 ns stress at 85 % of V_{bd}.

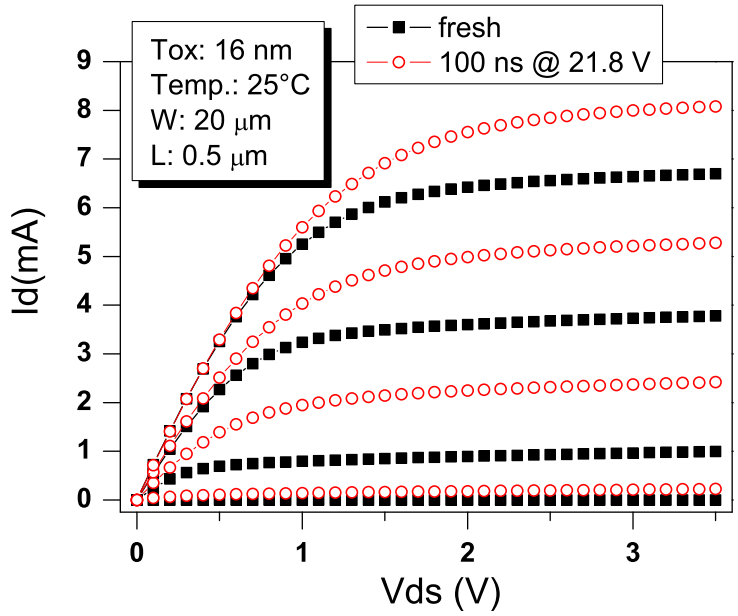

Figure 5.4: Drain current as a function of the drain voltage for different gate biasing of a 16 nm thick oxide nFET. A fresh characteristic as well as a pre-stressed with a 100 ns stress at 85 % of V_{bd}.

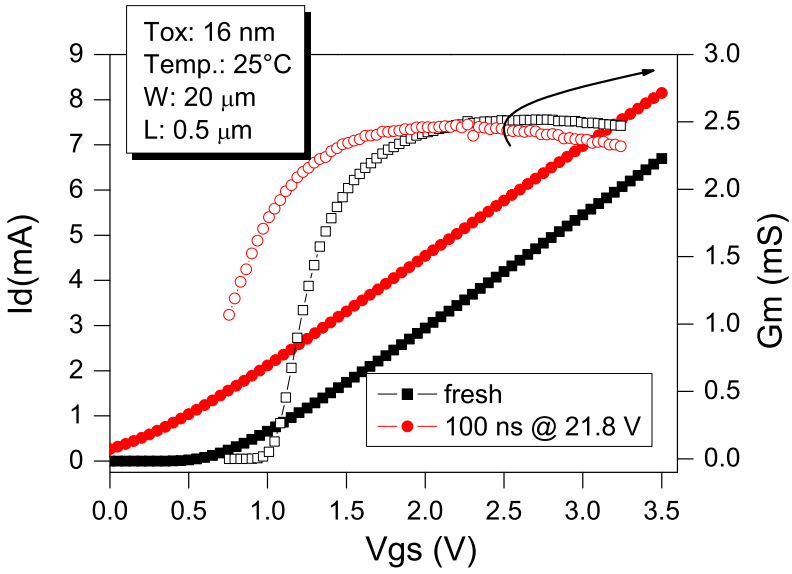

Figure 5.5: Saturated threshold voltage and transconductance of a 16 nm thick oxide nFET induced by a 100 ns stress at 85 % of V_{bd}.

for an oxide T_{ox} = 16 nm. For the long term a lower relaxation effect is initiated with a power law function Figure 5.6. Trapping and relaxation capability are linked to the oxide quality and the processing steps (as the nitridation for example); thus recovery effects are technology and/or device dependent. However even if a relaxation process occurs, the kinetic of this procedure at room temperature is low and will consequently impact the power consumption in case where the circuit is left undamaged. For the 16 nm oxide a low ESD stress to 85% of the breakdown voltage already creates large parameter drifts and high standby leakage as refereed in Figure 5.6.

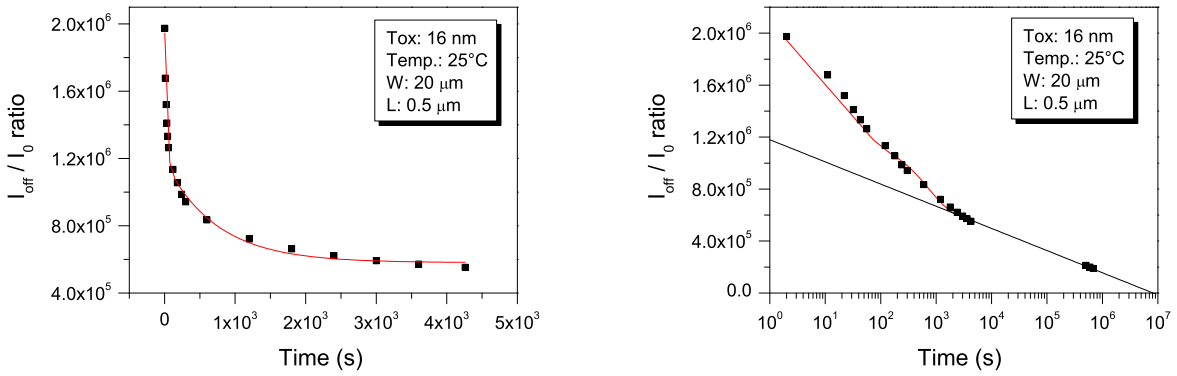

Figure 5.6: Off state drain current recovery effect from a 16 nm thick oxide nFET after an initial 100 ns stress at 85 % of V_{bd}.

5.2.2 Medium oxides (2.5 nm to 7 nm)

Two medium oxides have been tested, one of 6.8 nm and the other of 5.2nm. A similar clear positive trapping as for thicker oxides was noted for the 6.8 nm but a more complex behavior occurs for 5.2 nm devices.

The degradation of 5.2 nm nFETs just after a 100 ns inversion mode stress at a moderated ESD stress level (90 % of V_{bd}), first results in a strong positive trapping seen in the threshold voltage shift towards the negative values (Figure 5.7, state 1). The reduction in the $G_{m,max}$ peak observed is significative of the mobility reduction induced by interface states generation. After some minutes, an opposite effect is revealed, a net effective negative trapping is becoming predominant, see Figure 5.8. This can possibly be explained by the diffusion of positive hydrogen species into the oxide and by their recombination with precursor sites. The negative trapping obtained after some minutes is staying in a stable state (Figure 5.7, state 2). The diffusion of the positive species and their neutralization is occurring within a very short time (10 s), see Figure 5.8.

A charge pumping measurement was performed on the device one day later and has revealed an uniform interface trapping along the channel from the homogenous increase in the I_{cp} peak

(Figure 5.9). No oxide charges are observed in the charge pumping curve. The stable $V_{t,sat}$ shift to the higher voltage in the second phase is resulting from the acceptor interface states. The hypothesis formulated seems to be coherent as the effective negative trapping observed in the $V_{t,sat}$ is engendered by the interface states. A similar first unstable positive trapping behavior has been observed for a 3.2 nm nFET stressed in inversion under TLP [38] but it was suggested that these positive charges were trapped into the LDD spacers and neutralized via electron injection applied from a post CVS stress. The neutralization of positive charges generated from TLP testing via electron injection was also reported on thicker oxides (7 to 8 nm) [32, 31]. In this study, the positive charges generated during the ESD stress are believed to get neutralized directly from their diffusion into the oxide. Concerning the off state leakage, after the neutralization step, the I_{off} parameter is however not fully relaxed due to the interface state which are degrading the sub-threshold slope (Figure 5.10).

Anyway, even if the oxide charges get neutralized, the interface states remains present and cause some parameter degradation such as a reduction of the I_{on} and an increase in the I_{off} (Figure5.7).

In the case of medium oxides, a further degradation parameter can be monitored in the gate leakage via the SILC phenomenon. This leakage current linked to the oxide traps will be developed in the next section with a comparison in the defect generation between FN and ESD stress.

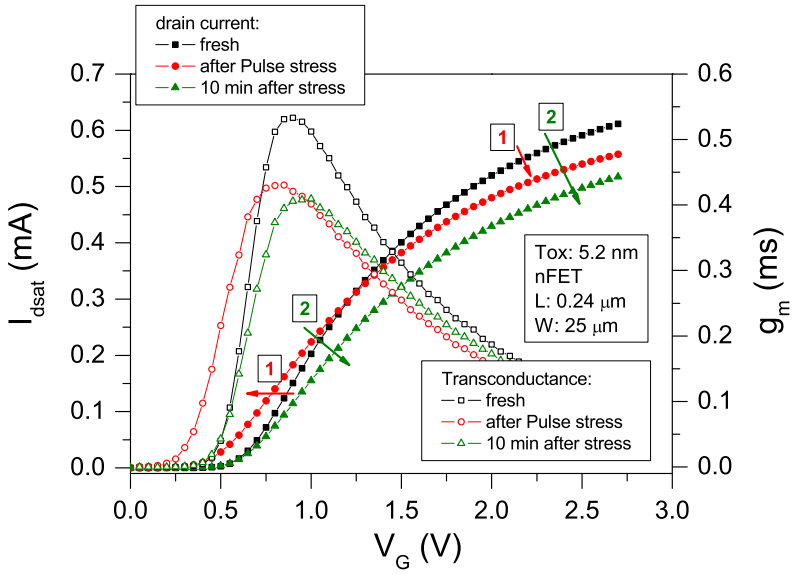

Figure 5.7: Saturated drain current from an 5.2 nm nFET after an 100 ns pulse stress at 90 % of the breakdown voltage. In a first step (1), a threshold voltage (V_{th}) reduction due to unstable positive charges occurs, then in a second phase (2) the degraded characteristic is dominated by a net negative trapping (increased V_{th} and I_{on} reduction) [39].

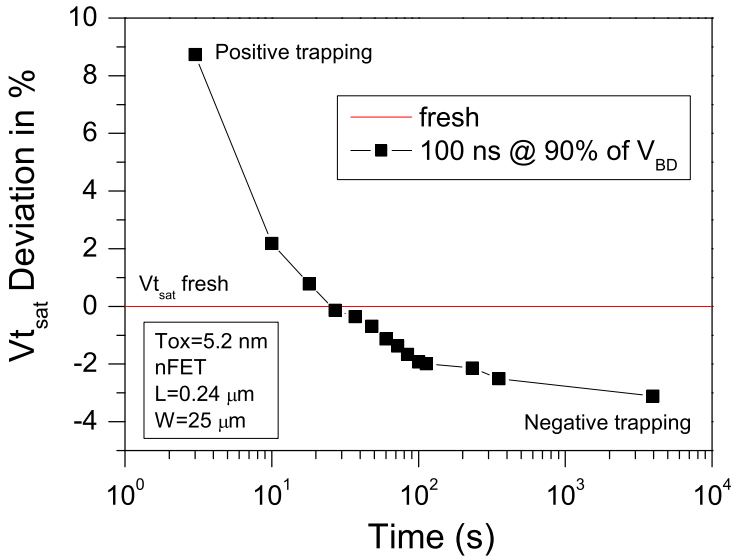

Figure 5.8: Saturated threshold voltage monitored within the time after one 100 ns pulse stress applied at 90 % of the breakdown voltage.

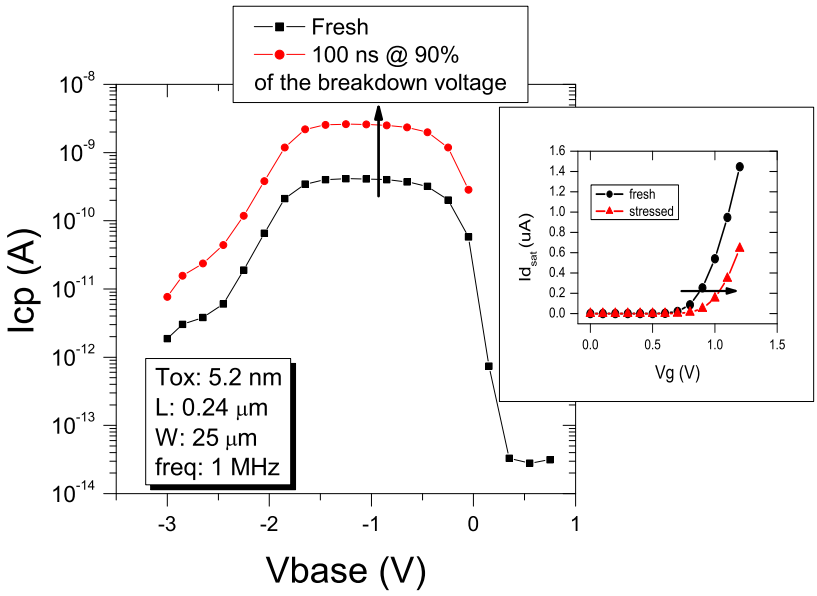

Figure 5.9: Homogenous increase in the charge pumping current on a 5.2 nm oxides stressed at 90 % of its breakdown value. The measure was performed one day after the stress [39].

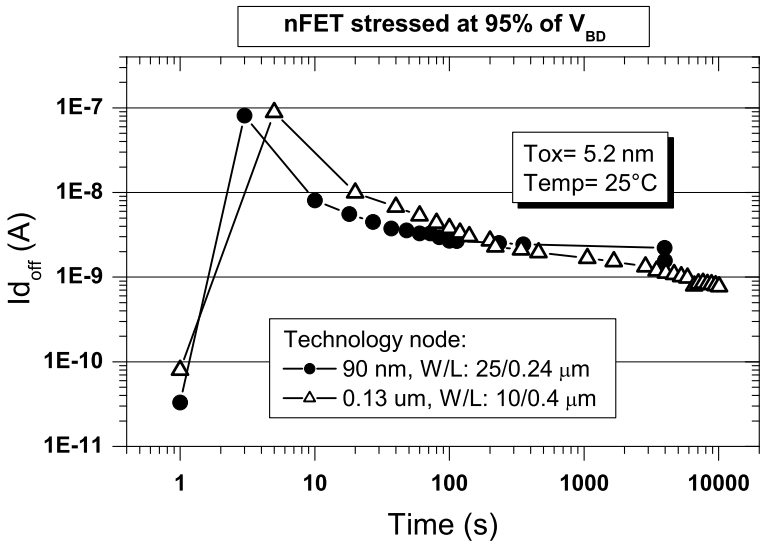

Figure 5.10: Characteristic recovery effect in time at room temperature from stressed devices by a 100 ns pulse stress [39].

5.2.2.1 Trap generation monitoring via SILC

For medium oxide thicknesses, the contribution of the bulk traps to a conductive current based on the tunneling electron via these traps can be monitored. This leakage current introduced in Chapter 1 is the SILC. The investigation on the SILC generation can contribute to the understanding of the oxide degradation [40, 41]. The GOX deterioration under short 100 ns pulses shows the increase in the gate leakage current (SILC) followed by an abrupt hard post breakdown state. The modeling of the post breakdown leakage curves in Figure 5.11 uses Equation 4.6.

A special focus has been put on the SILC generation in 2.65 nm nFET through CVS stress in the "DC" (10 ms) and in the ESD time domain (100 ns and 20 ns). Both stresses are performed in the Fowler-Nordheim tunneling regime for the voltage levels applied. This can be observed in the Figure 5.12 where the FN tunneling model has been fitted to the current density measurements. The trap generation rate under DC or short pulses described the same kinetic behavior (Figure 5.13). The common power law relationship of the SILC generation as a function of the injected charge [41, 42, 43] is also valid in the ESD time range. This confirms the continuity of degradation mechanisms from DC until the 100 ns time range which has already been observed in TDDB measurements. During an ESD stress, there is less injected charges than during a DC stress ($Q_{\mathrm{inj}} = J_{\mathrm{g}} \cdot t_{\mathrm{stress}}$). The current density during an ESD stress (100 ns) is about 2 to 3 orders in magnitude higher than for a DC stress (10 ms), but the stress time is 5 orders in magnitude lower. This emphasizes that at high field, the SILC generation is not fluency-driven but rather voltage-dependent. The efficiency of the defects generated by the injected charge is

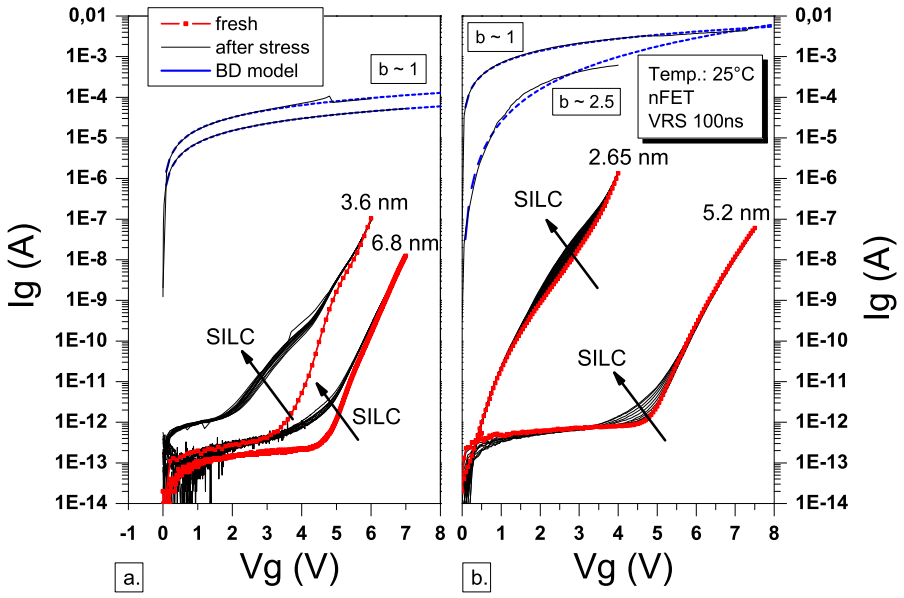

Figure 5.11: nFET gate leakage degradation under 100 ns rectangular voltage ramp stress pulses. This related SILC phenomena induced by short pulses can be also seen down to 2.65 nm oxide thickness [39].

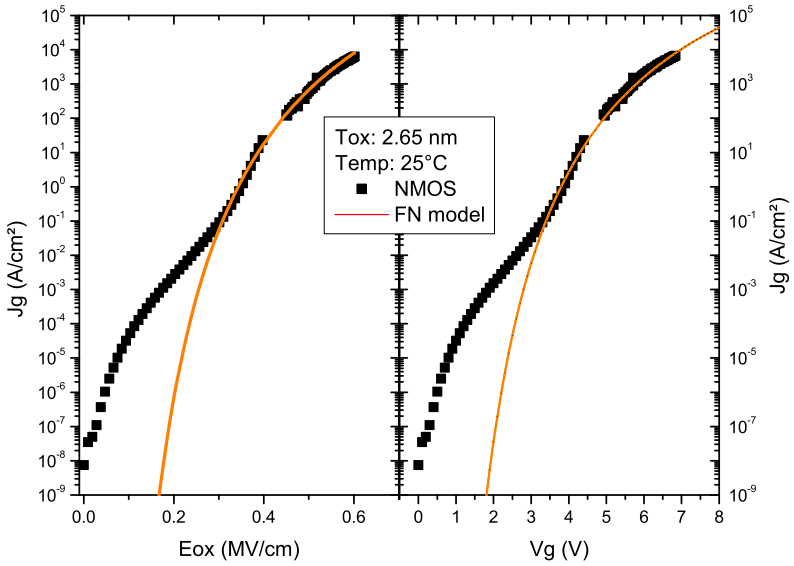

Figure 5.12: Current density of an 2.65 nm nFET at room temperature measured via DC and via high voltage sensing technic (see section 3.2.5.2).

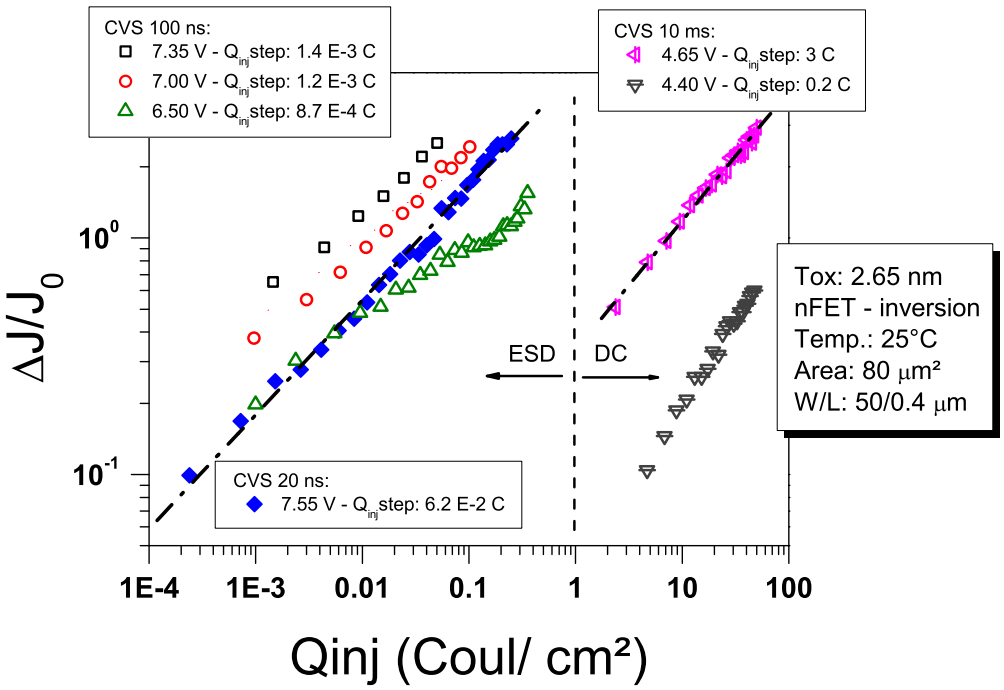

Figure 5.13: SILC generation under CVS injection with various pulse width stresses. From DC to high and short pulse stresses the kinetic of the SILC hold the same power law behavior [39].

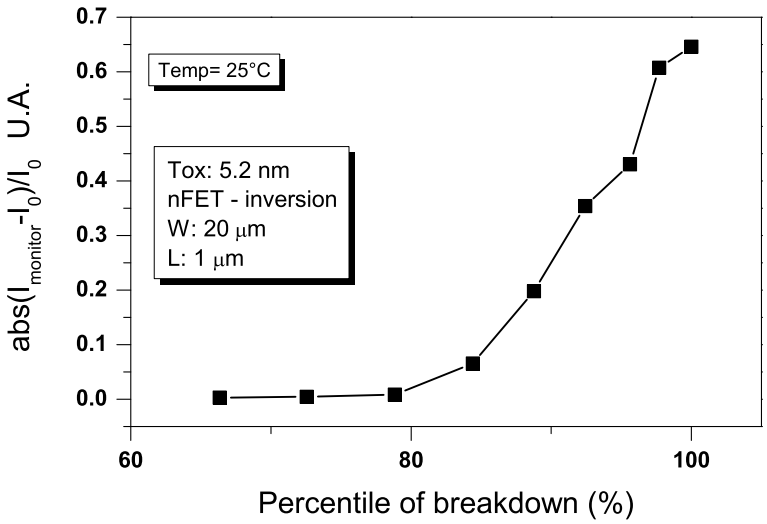

Figure 5.14: SILC generation under CVS injection with various pulse width stresses. From DC to high and short pulse stresses the kinetic of the SILC hold the same power law behavior.

consequently increasing with voltage. This matches the gate-oxide degradation picture based on the release of the maximum energy from injected carriers at the anode side [44, 45, 46]. The SILC generation is dependent on the injection procedure [47, 43, 48, 49, 50, 51, 37, 52], the pulse stress width and the temperature dependence on the SILC generation rate have been reported in the ESD time regime in [36, 53]. Above all, it was concluded with high statistical certainty by Wu *et al.* in [42] that no reliable correlation exists between the critical amount of SILC at breakdown and the critical defect density at breakdown. However, the SILC is still a good indicator of the oxide deterioration induced by ESD stresses as its significant occurrence is coming together with DC characteristic drifts around 80 % of V_{bd} (see Figure 5.14).

5.2.3 Thin Oxides (< 2.5 nm)

In the case of thin gate oxides, it is more complex to understand the degradation because of carrier de-trapping by direct tunneling occur. Therefore, degradation in the bulk of oxide is not possible. Thin gate oxides are at least composed by two interface layers which can be degraded by the non-destructive ESD stress. In DC parameters the effect of the moderate stress is not very pronounced for the 2.2 nm and 1.5 nm nFET oxide tested. The stress can lead to some minor changes in parameters, smaller than 2 % and very close to V_{bd}. For 4-terminals DC measurements a minimal time delay in the range of several seconds is ineluctable and can be too long to catch the degradation on thin oxides due to the very fast de-trapping by direct tunneling. No increase in I_{off} is noticed, the main affected parameters is the on linear drain current which suffers from a weak degradation. The eventual interface traps generated under ESD pulse stress hasn't been observed in the gate current leakage via Low-Voltage SILC [54, 55, 56]. However, charge pumping measurement performed on a 2.2 nm oxide one day after a stress to 95% of V_{bd} has revealed the presence of homogenous interface traps along the channel as well as the neutralization of an positive oxide charge (Figure 5.15). This can be seen from the flat band voltage shift towards the positive values.

5.2.4 Device degradation threshold to ESD

The oxide to device degradation exposed in this part begins around a stress level of 70 %-80 % of the characteristic breakdown voltage (V_{bd}) of the VRS Weibull distribution. The typical parameter drifts $V_{t,sat}$ and I_{off} as a function of an 100 ns VRS monitored immediately after the stress (within 1 s) are summarized in Figures 5.16 and 5.17 for the different oxide thicknesses investigated. From thick to medium oxides, the main parameter undergoing a large deviation is the off state leakage current of the transistor (I_{off}). The large increase observed of several order in magnitude could consequently damage the power consumption for low leakage application and the recovery of this parameter within the time is also rather low. The shifts right after the stress can be sufficient to drift transistor characteristics out of specifications and possibly lead to

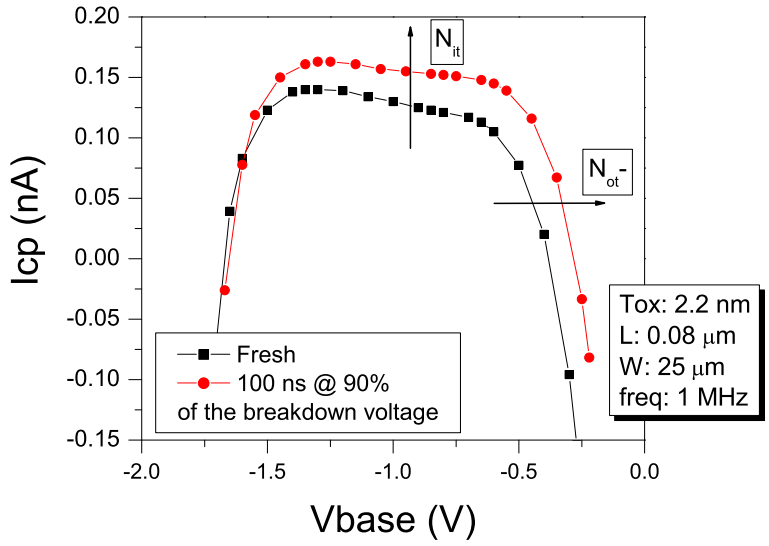

Figure 5.15: Charge pumping current from on a 2.2 nm oxide showing a comparison of a fresh device and a stressed one to 95 % of its breakdown value. The measure was performed one day after the stress.

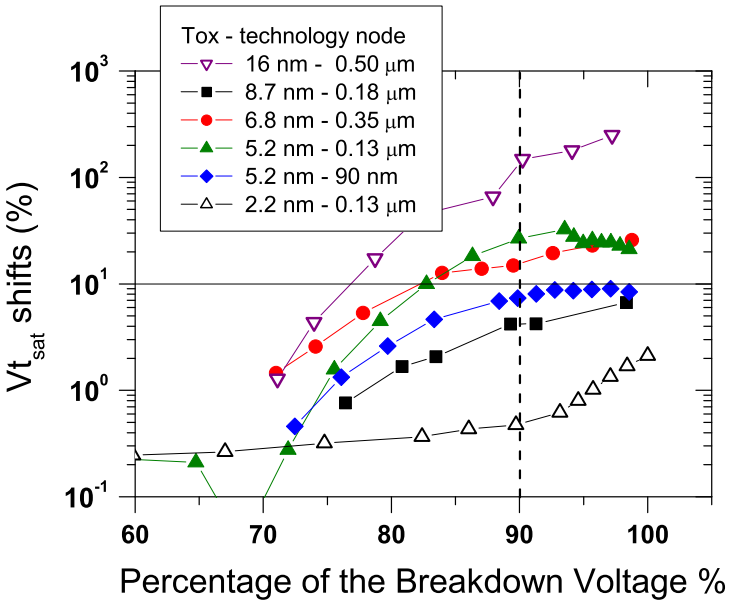

Figure 5.16: Threshold voltage shift induced by ESD-VRS as a function of the breakdown voltage percentage value. (b) Homogenous increase in the charge pumping current on a 5.2 nm oxides stressed at 90 % of its breakdown value [39].

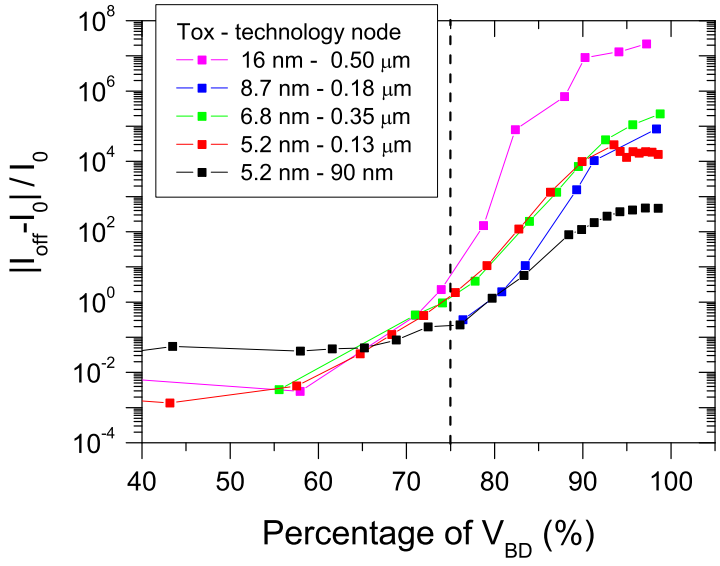

Figure 5.17: Threshold voltage shift induced by ESD-VRS as a function of the breakdown voltage percentage value. (b) Homogenous increase in the charge pumping current on a 5.2 nm oxides stressed at 90 % of its breakdown value.

a circuit failure [57]. If a variation in the saturated threshold voltage of 10 % is defined as a fail, the ESD design must guarantee that no moderate HBM stress above 75 % of V_{bd} should occur on thick to medium GOX. The impact of degraded device parameters on circuits is questionable for digital parts but not for thick analogue I/O transistors for which precisely the ESD charge injected in the oxide shows a pronounced influence. Although a relaxation process occur and some annealing steps could catalyze the recovery, a conservative approach is advised for these thick to medium oxides where the ESD design window is still large enough to cover without too much difficulties this extra margin.This voltage limit at which the ESD charge injection becomes too large should also permit to verify the reliability of the output nFET drivers used as snapback protection element. The gate to drain voltage should not cross this limit at which potentially the injected charges during the snapback condition [6] will affect the reliability of the device as reported in [8, 7].

Concerning the thin oxides, no special issue directly caused by some characteristic shifts is seen as the device degradation occurs very late and close to the breakdown value, see Figure 5.16. The reason for this is the different trapping behavior between thick and thin oxides. For thin oxides in which no deep oxide charges can be fixed, it seems that the interface states generated during the stress are still to low to cause significant device characteristics. However, as for the thick oxides where a clear oxide degradation is seen as for thin oxides where no degradation has been reported, the observation of the device degradation alone is not sufficient. The

impact of the oxide defects generated has to be investigated in regard to long term reliability to ensure no latency effects. This study is perhaps even more important for thin oxides than for thick ones as in the case where a moderate ESD stress would have a consequent impact on the long reliability term it is not possible to notice it from a simple *IV* characterization.

Low ESD stress amplitude lead certainly to an oxide degradation via oxide traps and interfaces states generation on thick and medium oxides. It has been extensively reported [27, 30, 47, 34, 16, 37, 35]. Few long term reliability studies have been performed [8, 7, 34, 10, 12, 16, 17] but none of them have really quantified the device degradation and correlated it with the TDDB or with an ESD design failure criterion. The goal of this oxide to device degradation chapter is to contribute to the elaboration of gradable failure criteria applicable for ESD safe design. The next section will then focus on the interaction between moderate ESD stresses and an accelerated aging tests emulating a long term device degradation procedure, being hot carrier injection.

5.3 ESD interaction with Hot Carrier Injection

Interface traps creation during operating conditions is one of the major reliability topic which is driving the CMOS technology integration capability. Special process steps have been modified with the down scaling of technologies to overcome this problem. Nevertheless, it is has been seen in the past that introduction of special implants in the channel close to the diffusion (LDD, MDD) have changed the physics involved in the ESD protection element and has lead to a drastically reduction in the ESD robustness. As both, HC and ESD are coming crucial for highly integrated technologies, interaction of both effect will be presented here. HCI has a direct impact on the device functionality and is a good test methodology to estimate the device lifetime. Moderate ESD stress have been observed to cause oxide to device degradation and the interaction from a pre-ESD stress device with the HCI occurring under normal operations need to be investigated to guarantee the reaching of the device lifetime criterion targeted.

5.3.1 Medium oxides, $Tox = 5.2$ nm

5.3.1.1 Reference HC stress

For the HCI, the focus was on 5.2 nm nFETs. A typical I/O geometry was chosen with a width of 25 µm and a channel length of 0.24 µm. For this short channel MOSFETs under high field condition, the carriers in the channel and in the pinch-off region of the transistor can gain sufficient energy to surmount via thermo-ionic conduction (see 1.2.1) or tunnel through the energy barrier between the silicon and oxide (Φ_b). This carriers injection generates some traps at the interface and into the oxide, leading to degraded device parameters. For this type of device and oxide range, the degradation generated from these hot carriers are maximum for

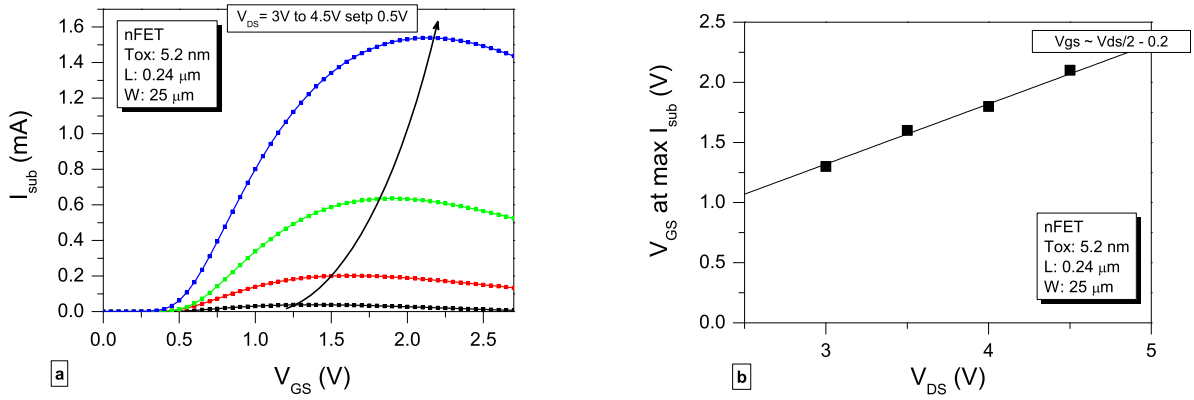

Figure 5.18: (a) Substrate current as a function of the gate voltage for different drain biasing condition. (b) From the extraction of the maximum substrate current peak from Figure (a), the corresponding gate voltages are plotted as a function of V_{ds}.

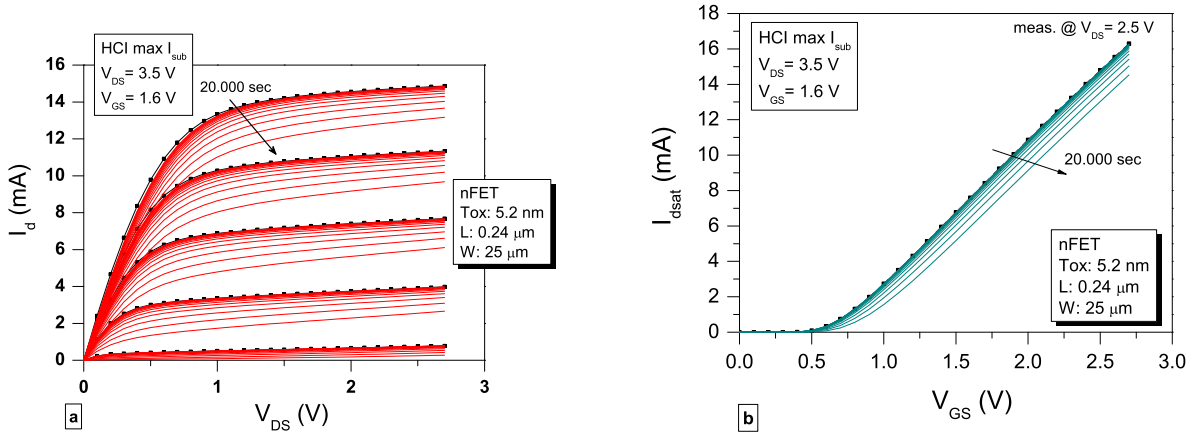

Figure 5.19: NFET device characteristic drifts under an hot carrier injection of 20.000 s at the maximum substrate condition for $V_{ds} = 3.5$ V. (a) I_{ds} current reduction as a function of the drain voltage for different gate bias. (b) Saturated drain current degradation with a lateral $V_{t,sat}$ shift.

the maximal substrate current. The substrate peak of current is dependent on the drain and gate potential and can be extracted from the I_{sub} vs. V_{gs} curve plotted for different V_{ds} bias as exposed in Figure 5.18. At this condition the field and impact ionization phenomena induced the possibility for electrons and holes to be injected into the oxide [58, 59, 3]. For nFET devices, the HC degradation lead to the increase of the threshold voltage V_{th} and to the diminution of the transconductance and MOSFET drive current. This parameters degradation can be observed for an hot carrier stressed at V_{ds} = 3.5 V (for the max I_{sub} current at V_{gs} = 1.6 V) in the Figures 5.19 and 5.20. The degradation are mainly due to negative trapping and acceptor interface generation.

The parameters drifts under HCI are following a power law as a function of the time. The reference parameter in this investigation is the drain current in the saturated regime for $V_{gs} = V_{dd}$ (I_{on}). The degradation of this parameter can be described by:

$$\frac{\Delta I_{on}}{I_0} = A \cdot t^n \tag{5.6}$$

where t is the stressing time and A is a technological dependent parameter depending on V_{ds}. The exponent n for stresses at the maximum of I_{sub} is in the range of 0.5 to 0.7. The I_{on} degradation in the linear and saturated regimes as a function of the stressing time is exposed for a device stressed at V_{ds} = 3.5 V in Figure 5.21. The kinetic of the degradation is found with a slope equal to 0.5. This power law degradation as a function of the time is correlated with the interface state generation, via the "Lucky Electron model" from Hu *et al.* [60],

$$\Delta N_{it}(t) = C_1 \left[\frac{I_{DS}}{W_G} \cdot t \cdot \exp\left(-\frac{\Phi_{it,e}}{q\lambda_e E_m} \right) \right]^n \tag{5.7}$$

where $\Phi_{it,e}$ is the energy that an electron must possess in order to create an interface trap, λ_e is the hot electron mean free path. A measure for the electrical field E_m is the multiplication factor M = I_{sub} / I_{ds},

$$M = \frac{I_{sub}}{I_{DS}} = C_2 \exp\left(-\frac{\Phi_i}{q\lambda_e E_m} \right) \tag{5.8}$$

where Φ_i is the impact ionization energy. The lifetime defined as the time to reach a fixed amount of damage, is consequently given by,

$$\frac{t I_{DS}}{W_G} = C_3 \left(\frac{I_{sub}}{I_{DS}} \right)^{-\frac{\Phi_{it,e}}{\Phi_i}} = C_3 M^{-m} \tag{5.9}$$

A simplified form from Equation 5.7 is the lifetime related to the drain voltage as proposed by Takeda *et al.* [61],

$$\tau_{DC} = C \exp \frac{B}{V_{DS}} \tag{5.10}$$

At a fixed failure criterion, here defined as 10 % increase in the I_{on} parameter, the stressing time elapsed until the reaching of this limit defines the period in which the conformance is still acceptable. This time is defined as the device lifetime with respect to DC Hot Carrier degradation. Higher the drain potential is and shorter the device lifetime is. This effect is shown in Figure 5.22 for four V_{ds} bias (3 V, 3.5 V, 4 V, 4.5 V). From different high drain bias, it is possible to extrapolate, accordingly to the model 5.10, the device lifetime towards the supply biasing condition. The stress considered here is pure DC static and doesn't correspond to the real stress that a device endured in a product. Nevertheless, this test is a good indicator for the study of device degradation in the long term and permit to reveal possible latency effect from pre-electrical stress.

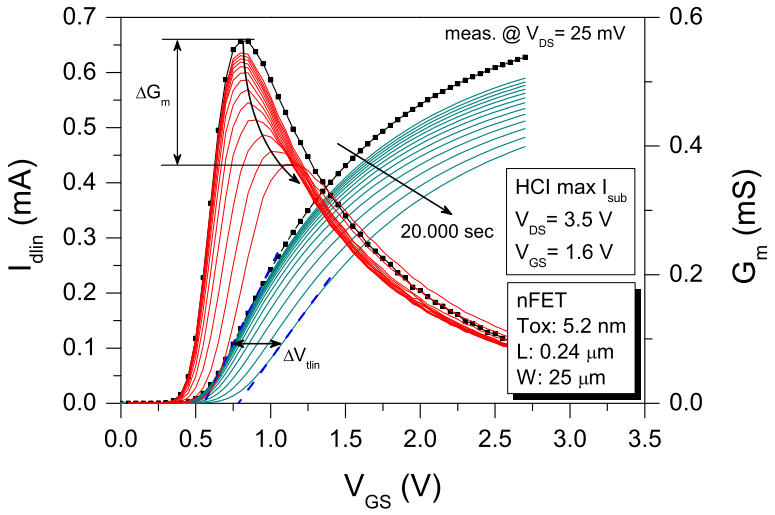

Figure 5.20: $I_{\mathrm{d,lin}}$ and transconductance degradation under an HCI of 20.000 s at the maximum substrate condition for $V_{\mathrm{ds}} = 3.5$ V. A reduction in the transconductance peak ($G_{\mathrm{m,max}}$) and in the on current (I_{on}) as well as a lateral shift of $V_{\mathrm{t,lin}}$ to positive values can be observed.

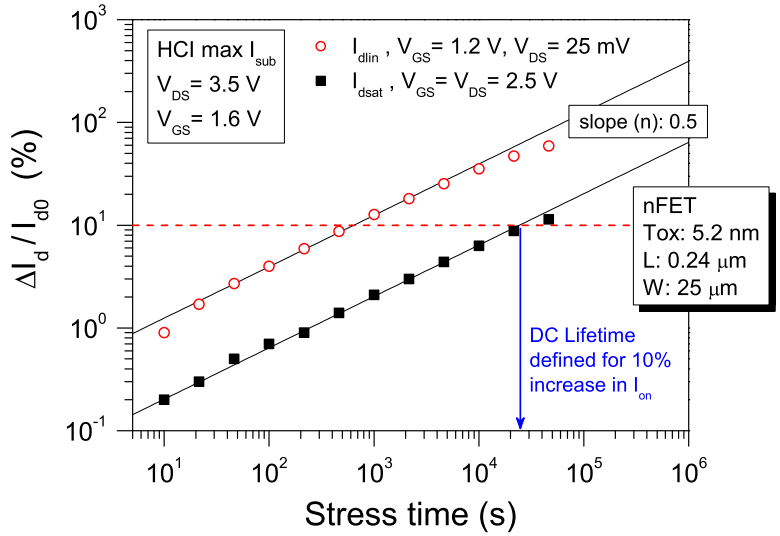

Figure 5.21: On drain current degradation (I_{on}) drifts monitored in the linear and saturated regimes, under an HCI at the maximum substrate condition for $V_{\mathrm{ds}} = 3.5$ V. Both degradation are following a power law as a function of the stressing time with an exponent value of 0.5. The device lifetime extraction corresponding to 10% increase in the saturated I_{on} is indicated.

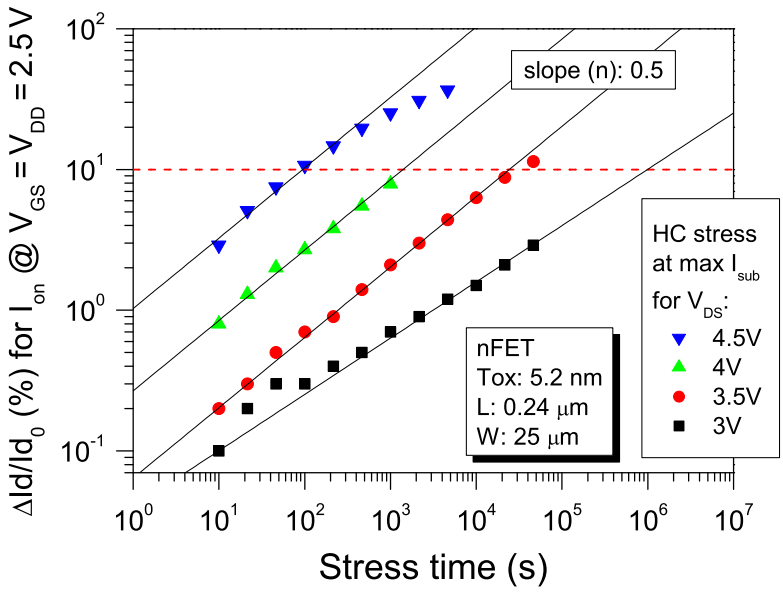

Figure 5.22: Degradation of the I_{on} current in the saturated regime under HCI performed for the maximal substrate condition at different V_{ds}.

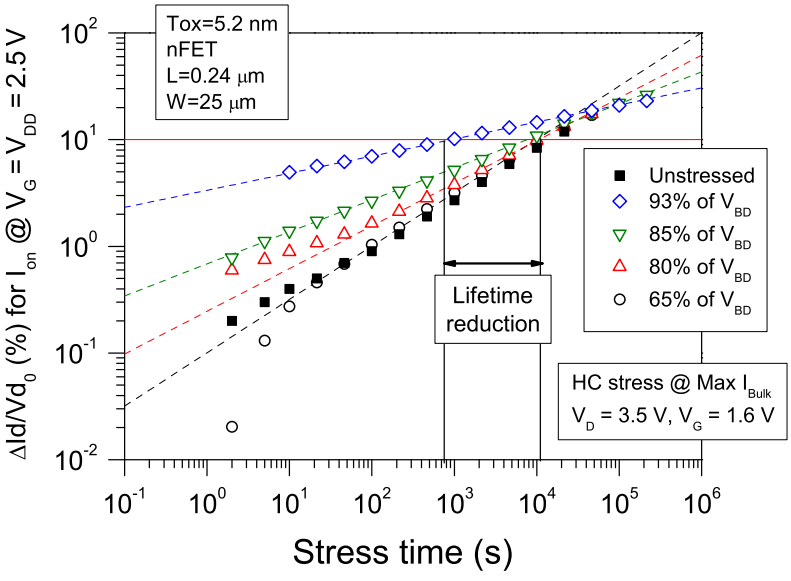

Figure 5.23: Impact of the ESD pre-stress level on the hot carrier degradation kinetic. A pure cumulative effect of the charged induced from the ESD stress is revealed. No consequent long term lifetime impact is noticed for stress levels below 85 % of the breakdown voltage [39].

5.3.1.2 ESD stress prior to HC

Pre-stressed structures with an 100 ns moderate pulse were performed before HC injection at the maximal I_{sub} condition. The devices tested are the same as used for device degradation studies in 5.2.2. The post HC stresses were performed about 10 min after the stress only at the point where device characteristics have reached the second stable phase (Figure 5.7, state 2). The ESD-like stress and the HC stress seem to be purely cumulative; the HC kinetic degradation of stressed devices is slightly lower as fresh ones but shifted by a strong initial trapping which is dependent on the pre-stress level (Figure 5.23). At long term, the charges generated by the ESD stress are merged in the high quantity of carriers generated by the HC stress. No reduction in the device lifetime is noted for ESD-like pre-stress up to 85 % of V_{bd}. On the other hand, the strong trapping induced by a stress within 90 % of V_{bd} leads to a clear effect on the device lifetime. The power law degradation exponent is decreasing as a function of the pre-ESD stress, for a stress to 90% of V_{bd} the slope is reduced to 0.15.

To fully evaluate the impact on the device lifetime, HC stresses were performed on per-stressed devices at the same stressing drain voltage than fresh devices and for a fixed stress level of 90 % of V_{bd} (Figure 5.24). The ESD stressed devices have the lifetime reduced by a factor 15 (Figure 5.25). As device degradation due to a moderate uniform ESD stress applied to the gate is purely dependent on the ESD stress level (in relation to the amount of trapped charges, see Figure 5.23), the question is if these stressed devices can be annealed by charges de-trapping, leading to the partial recovery of the device reliability criterion. To investigate whether an annealed device could recover its initial fresh device lifetime, a stressed device at 90 % of the breakdown voltage value was stored at room temperature for one week and then put at 200 °C for 10 h. An HC stress was performed on this device afterwards. However, only a slow relaxation process was observed and the amount of traps was slightly smaller but still very high. This confirms that bulk oxide traps dominate on the hot carrier degradation mechanisms which is observed by a typical kinetic with a small time exponent reducing the device lifetime. The assumption that an annealed oxide results in a consequent recovery of the device lifetime has not been seen, however recovery effects are technology dependent.

Moderate ESD stresses can result in immediate failure due to characteristic drifts out of specification (for the failing criterion of 10 % increases in the saturated drain current for example). As ESD stresses and HC degradation mechanisms seem to be purely cumulative, no latent damage is seen here. Anyway, the strong trapping effect itself induced by the ESD stress could significantly reduce the device lifetime. A safe ESD design margin covering the oxide trapping degradation range (characteristic drifts, SILC, HC lifetime reduction) should be settled for sensitive circuits in order to avoid any malfunction issues.

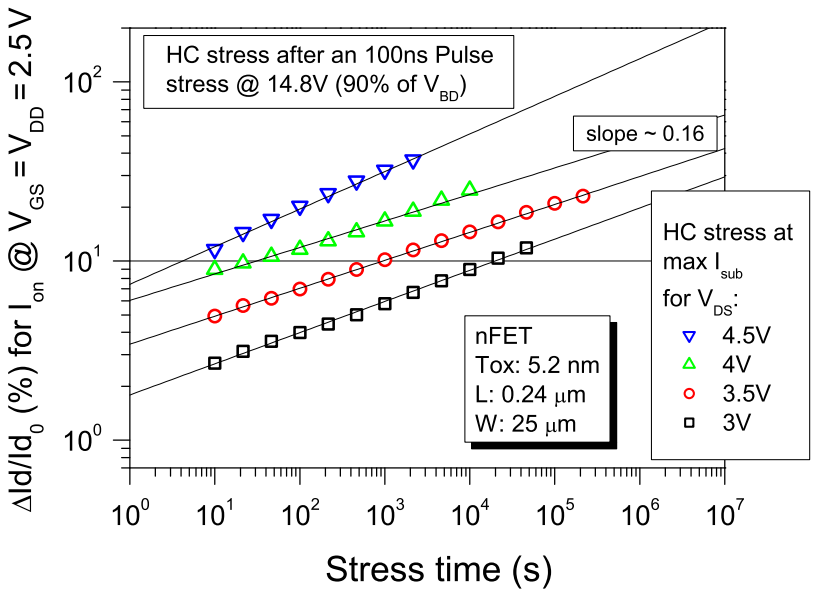

Figure 5.24: HC stresses at the maximal substrate current condition are applied on fresh and pre-stress devices at 90 % of V_{bd}. A clear impact on the hot carrier degradation kinetic can be observed.

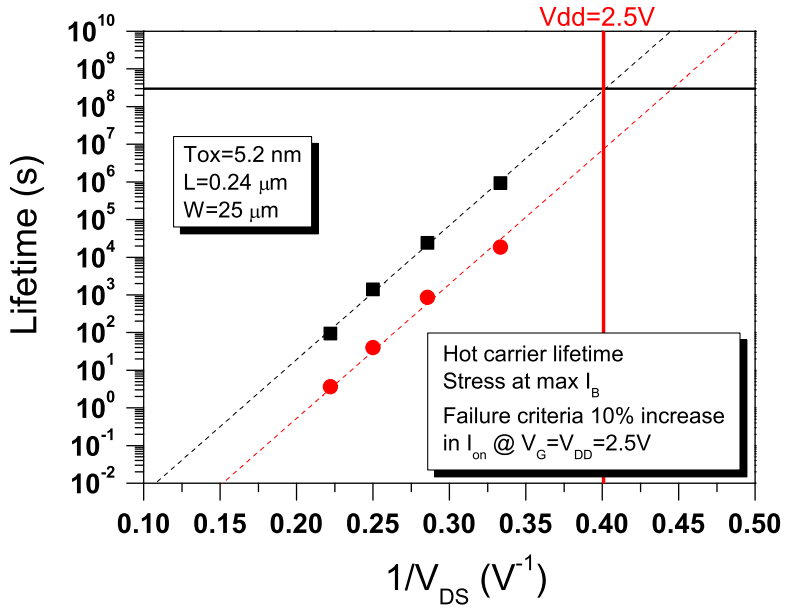

Figure 5.25: Hot carrier lifetime deduced from the Figures 5.22 and5.24. The pre-ESD stressed devices at 90 % of V_{bd} have a reduced device lifetime of 1.5 decade in time [39].

5.3.2 Thin Oxides ($T_{ox} < 2.5$ nm)

In order to investigate also device lifetime degradation caused by moderate ESD stress on thin oxide devices, HC stresses have been performed after a moderate ESD-like pulse stress.

Two different HC stress conditions have been performed, the worst degradation case for thin oxide which is occurring for the maximal field condition at $V_g = V_d$ and the maximum substrate current bias condition. Hot carrier stresses after a 100 ns stress at 90 % of the GOX breakdown value has minor impact on the HC degradation kinetic for the tested 2.2 nm oxide (Figure 5.26). The weak trapping induced by the ESD stress is rapidly dominated by the HC degradation mechanisms and in the long term no HC lifetime criterion reduction has been noticed. Similar behavior was also observed for drain ESD stress configuration [16]. The ESD stresses and HC are purely cumulative, fewer amounts of defects generated or of charges trapped in thin oxide (no occupied volume traps due to the direct tunneling distance) does not lead to any specific device reliability issues. For an ESD stress in the range of 90 % of V_{bd}, no degradation could be monitored but the dielectric lifetime will be consequently affected 4.2.2 due to the cumulative nature of the dielectric. A possible latent damage is seen here if the lifetime margin criteria fixed basically by the power law voltage acceleration is not taken into account.

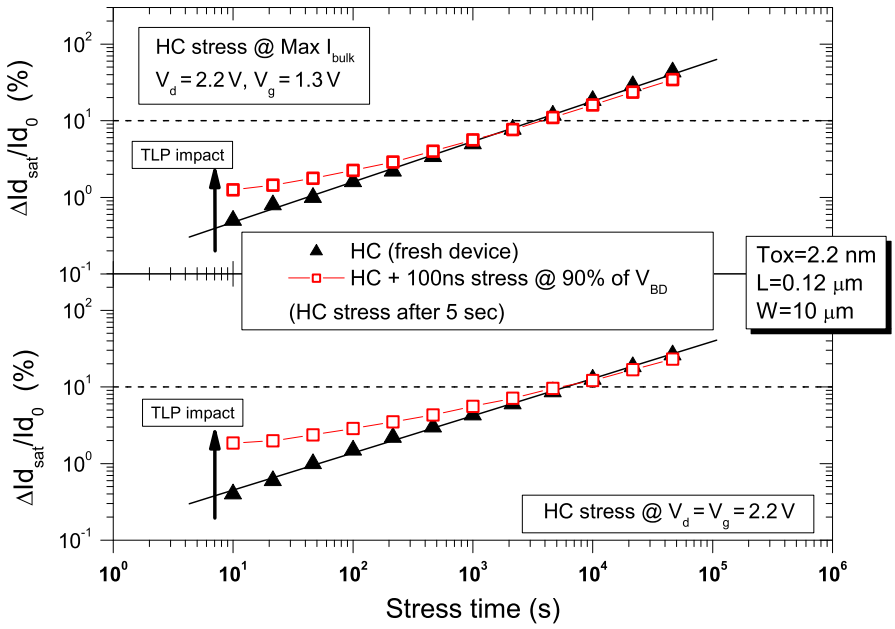

Figure 5.26: Impact of an 100 ns ESD-like pulse on the device hot carrier stress degradation for a 2.2 nm nFET oxide [39].

5.4 Conclusion

In this chapter, the effects induced by moderate ESD stresses on the device integrity of nFETs have been developed. It has been observed that the device degradation is strongly thickness dependent.

- In thick and medium oxides, oxide bulk traps, border traps and interface states are generated during the ESD stress. The effect of the deep oxide and border traps induced some consequent drifts in device parameters and can degrade the device lifetime in case of strong oxide charge trapping. A safety margin should be accounted for in the ESD design of sensitive analogue I/O pins.
- In case of thin oxides, the impossibility to trap charges in the oxide due to direct tunneling avoid significant drifts in the device parameters. No latency damage has been reported concerning thin oxides.

The consequences of the oxide degradation under ESD stresses on the device lifetime and integrity as well as the statistical oxide breakdown nature presented in both Chapters 4 and 5 will be developed and discussed for a proper determination of a safe ESD flow for advance CMOS technologies in the last chapter 6.

Bibliography

[1] Amerasekera A. and Duvvury C. The impact of technology scaling on ESD robustness and protection circuit design. *Transactions on Components, Packaging, and Manufacturing Technology*, 18(2), 1995.

[2] Kurachi I., Fukuda Y., Miura N., and Ichikawa F. Analysis of soft breakdown failure with ESD on output buffer nMOSFETs and its improvement. *Transaction on Industry Applications*, 30(2):358–364, 1994.

[3] Groeseneken Guido. Hot carrier degradation and ESD in submicrometer CMOS technologies: How do they interact? In *Transaction on Device and Materials Reliability*, volume 1, pages 23–32, 2001.

[4] Zängl F, Gossner H., Esmark K., Owen R., and Zimmermann G. Case study of a technology transfer causing ESD problems. In *Microelectronics Reliability*, volume 42, pages 1275–1280, 2002.

[5] Khurana N., Maloney T., and Weh W. ESD on CHMOS devices-equivalent circuits, physical models and failure mechanisms. In *International Reliability Physics Symposium*, pages 212–223, 1985.

[6] Ministry K.R., Krakauer D.B., and Doyle B.S. Impact of snapback-induced hole injection on gate oxide reliability of N-MOSFET's. *Electron Device Letters*, 11(10), 1990.

[7] Huh Y., Lee M.G., Jung H.C., Li T., Song D.H., Lee Y.J., Hwang J.M., Sung Y.K., and Kang S.M. A study of ESD-induced latent damage in CMOS integrated circuits. In *International Reliability Physics Symposium*, 1998.

[8] Aur S., Chattejee A., and Polgreen T. Hot-electron reliability and ESD latent damage. *Transaction on Devices*, 35(12):2189–2193, 1988.

[9] Manku T. The identification and analysis of latent ESD damage on CMOS input gates. *Electronics Letters*, 30(24):2074–276, 1994.

[10] Song M., Eng D.C, and MacWilliams K.P. Quantifying ESD/ESD latent damage and integrated cirecuit leakage currents. In *EOS/ESD Symposium*, pages 304–310, 1995.

[11] Wada Tetsuaki. Characterization of charge accumulation and detrapping processes related to latent failure in CMOS integrated circuits. *Microelectronics Reliability*, 36:1707–1710, 1996.

[12] Reiner J.C., Keller T., Jäggi H., and Mira S. Impact of ESD-induced soft drain junction damage on CMOS product lifetime. *Microelectronics Reliability*, 40:1619–1628, 2000.

[13] Reiner J.C. Latent gate oxide defects caused by CDM-ESD. In *EOS/ESD Symposium*, pages 311–321, 1995.

[14] Salome P., Leroux C., Mariolle D., Lafond D., Chante J.P., Crevel P., and Reimbold G. An attempt to explain thermally induced soft failures during low ESD stresses: study of the differences between soft and hard NMOS failures. *Microelectronics Reliability*, 38:1763–1772, 1998.

[15] Amerasekera A., Gupta V., Vasanth K., and Ramaswamy S. Analysis of snapback behavior on the esd capability of sub 0.2 um NMOS. In *International Reliability Physics Symposium*, pages 159–166, 1999.

[16] Salman A., Gauthier R., Stadler W., Esmark K, Muhammad M., Putnam C., and Ioannou D. Characterization and investigation of the interaction between hot electron and electrostatic discharge stress using NMOS devices in 0.13 ţm CMOS technology. In *International Reliability Physics Symposium*, pages 219–225, 2001.

[17] Salman A., Gauthier R., Wu E., Riess P., Putnam C., Muhammad M., Woo M., and Ioannou D. Electrostatic discharge induced oxide breakdown characterization in a 0.1ţm CMOS technology. In *International Reliability Physics Symposium*, pages 170–174, 2002.

[18] Cester A., Gerardin S., Tazzoli A., Paccagnella A., Zanoni E., Ghidini G., and Meneghesso G. ESD induced damage on ultra-thin gate oxide MOSFETs and its impact on device reliability. In *International Reliability Physics Symposium*, pages 84–90, 2005.

[19] Mishra R., Mitra S., Gauthier S., Ioannou D., Kontos D., Chatty K., Seguin C., and Halbach R. On the interaction of ESD, NBTI and HCI in 65nm technology. In *International Reliability Physics Symposium*, pages 17–22, 2007.

[20] Vuillaume D., Marchetaux J., Lippens P.-E., Bravaix A., and Boudou A. A coupled study by floating-gate and charge-pumping techniques on hot-carrier-induced defects in sub-micrometer LDD n-MOSFET's. *Transaction on Electron Devices*, 40(4):773, 1993.

[21] Heremans P., Witters J., Groeseneken G., and Maes H.E. Analysis of charge pumping technique and its application for the evaluation of MOSFET degradation. *Transactions on Electron Devices*, 36(7):1318–1335, 1989.

[22] Groeseneken G. and Maes H.E. Basics and applications of charge pumping in submicron MOSFET's. *International Conference on Microelectronics*, 2:581–589, 1997.

[23] Vigrass W.J. and D.J. Dumin. Comparison of oxide wearout resulting from DC and pulsed charge injection. In *Southeastcon apos*, pages 225–227, 1988.

[24] Kitamura Y. and Kitamura H. Breakdown of thin gate-oxide by application of nanosecond pulse as ESD test. In *International Symposium on Testing and Failure Analysis*, pages 193–199, 1989.

[25] Tunnicliffe M.J., Dwyer V.M., and Campbell D.S. Experimental and theoretical studies of ESD/ESD oxide breakdown in unprotected MOS structures. *EOS/ESD Symposium*, pages 162–168, 1990.

[26] Greason W.D. and Chum K. Characterization of charge accumulation and detrapping processes related to latent failure in CMOS integrated circuits. *Transactions on Industry Application*, 1:586–593, 1991.

[27] Greason W.D. and Chum K. The integrity of gate oxide related to latent failures under EOS/ESD conditions. In *EOS/ESD Symposium*, pages 106–111, 1992.

[28] Tunnicliffe M.J., Dwyer V.M., and Campbell D.S. The integrity of gate oxide related to latent failures under EOS/ESD conditions. *Journal of Electrostatics*, 31:91–110, 1993.

[29] Colvin Jim. The identification and analysis of latent ESD damage on CMOS input gates. *Journal of Electrostatics*, 33:291–311, 1994.

[30] Greason W.D., Kenneth W., and Chum K. Characterization of charge accumulation and detrapping processes related to latent failure in CMOS integrated circuits. *Transaction on Industry Applications*, 30(2):350–357, 1994.

[31] Chim W.K. and Teh G.L. Neutral electron trap generation and hole trapping in thin oxides under electrostatic discharge stress. *Journal of Applied Physics*, 37:1671–1673, 1998.

[32] Teh G.L. and Chim W.K. Per-breakdown charge trapping in ESD stressed thin MOS gate oxides. In *International Symposium on Physical and Failure Analysis*, pages 156–161, 1997.

[33] Lim P.S. and Chim W.K. Latent damage investigation on lateral non-uniform charge generation and stress-induced leakage current in silicon dioxides subjected to low-level electrostatic discharge impulse stressing. In *International Symposium on Physical and Failure Analysis*, pages 156–161, 1999.

[34] Chim W.K. and Lim P.S. Latent damage generation in thin oxides of metal-oxide-semiconductor devices under high-field impulse stress and damage characterization using low-frequency noise measurements. *Japanese Journal of Applied Physics*, 40:6770–6777, 2001.

[35] Tseng J.C. and Hwu J.-G. Effects of electrostatic discharge high-field current impulse on oxide breakdown. *Journal of Applied Physics*, 101:014103–1–6, 2007.

[36] Wu J., Juliano P., and Rausenbaum E. Breakdown and latent damage of ultra-thin gate oxides under ESD stress conditions. In *EOS/ESD Symposium*, 2000.

[37] Chim W.K. and Lim P.S. Hole injection with limited charge relaxation, lateral nonuniform hole trapping, and transient stress-induced leakage current in impulse-stressed thin (<5 nm) nitrided oxides. *Journal of Applied Physics*, 91(3):6770–6777, 2002.

[38] Cester A., Gerardin S., Tazzoli A., and Meneghesso G. Electrostatic discharge effects in ultrathin gate oxide MOSFETs. *Transaction on Device and Materials Reliability*, 6(1):87–94, 2006.

[39] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. In *EOS/ESD Symposium Proceedings*, pages 328–337, 2007.

[40] Nicollain P.E., Hunter W.R., and Hu J.C. Experimental evidence for voltage driven breakdown models in ultrathin gate oxides. In *International Reliability Physics Symposium*, pages 7–15, 2000.

[41] DiMaria D.J, Buchanan, Sthatis J.H, and Stahlbush R.E. Interface states induced by the presence of trapped holes near the silicon-silicon-dioxide interface. *Journal of Applied Physics*, 77(5):2032–2040, 1995.

[42] Wu E.Y., Sune J., Nowak E., Lai W., and McKenna J. Weibull slopes, critical defect density, and the validity of stress-induced-leakage current (SILC) measurements. *International Electron Device Meeting*, pages 125–128, 2001.

[43] Esseni D. and Bude J.D. On interface and oxide degradation in VLSI MOSFETs - part II: Folwer-nordheim stress regime. *Transaction on Electron Devices*, 49:254–263, 2002.

[44] Wu Ernest Y. and Suñé Jordi. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45:1809–1834, 2005.

[45] Schuegraf K.F. and Hu C. Hole injection oxide breakdown model for very low voltage lifetime extrapolation. In *International Reliability Physics Symposium*, pages 7–12, 1993.

[46] Suñé J. and Wu E.Y. A quantitative two-step hydrogen model of Si/SiO2 gate oxide breakdown. *Solid-State Electronics*, 46:1825–1837, 2002.

[47] Lim P.S. and Chim W.K. Anomalous positive charge trapping in thin nitrided oxides under high-field impulse stressing. *International Symposium on Physical and Failure Analysis*, 77(17):2719–2721, 2000.

[48] Tan Y.N., Chim W.K., and Lim P.S. Stress-induced leakage current in thin oxides under high-field impulse stressing. In *International Symposium on the Physical & Failure Analysis of Integrated Circuits*, pages 228–233, 2001.

[49] Riess P., Ghibaudo G., Pananakakis G., Brini J., and Ghidini G. Impact of constant current stressing procedure on stress induced leakage current generation in thin oxides. In *European Solid-State Device Research Conference*, pages 544–547, 2007.

[50] Riess P., Ghibaudo G., Pananakakis G., and Brini J. Reversibility of charge trapping and SILC creation in thin oxides after stress/anneal cycling. *Microelectronics Reliability*, 38:1057–1061, 1998.

[51] Riess Philip. *Etude de la fiabilité des oxydes minces: analyse des mécanismes de transport et de génération du SILC*. Physique des composants a semiconducteurs, Institut National Polytechnique de Grenoble, 1999.

[52] Ghidini G., Capolupo C., Giusto G., Sebastiani A., Stragliati B., and Vitali M. Tunnel oxide degradation under pulsed stress. *Microelectronics Reliability*, 45:1337–1342, 2005.

[53] Rosenbaum E. and Wu J. Trap generation and breakdown processes in very thin gate oxides. *Microelectronics Reliability*, 41:625–632, 2001.

[54] Nicollian E., Hunter W.R, and Hu J.C. Low voltage stress-induced-leakage-current in ultrathin gate oxides. In *International Reliability Physics Symposium*, volume 37, pages 400–404, 1999.

[55] Petit Christian. *Contribution à l'étude de la dégradation de films minces et ultra-minces de SiO2 de strucures MOS soumises à des contraintes électriques et à la caractérisation par spectroscopie tunnel inélastique de junction Al-SiO2-Si*. Electronique, Université de Reims Champagne-Ardenne, 2004.

[56] Petit E., Meinertzhagen A., Zander D., Simonetti O., Fadlallah M., and Maurel T. Low voltage SILC and p- and n-MOSFET gate oxide reliability. In *International Reliability Physics Symposium*, volume 45, pages 479–485, 2005.

[57] Mohamedi S.Z., Chan V.-H., Park J.-T., Nouri F., Scharf W., and Chung J.E. Hot-electron-induced inupt offset voltage degradation in CMOS differential amplifiers. In *International Symposium on Testing and Failure Analysis*, pages 76–80, 1992.

[58] Goguenheim Didier. *Contribution à l'étude de la fiabilité des oxydes minces dans les structures MOS*. Microélectronique, Université de Provence Aix-Marseille I, 2006.

[59] Groeseneken G., Bellens R., Van den bosch G., and Maes H. Hot carrier degradation submicrometer MOSFETs: from uniform injection towards the real operating conditions. In *Semiconductor Science Technologies*, volume 10, pages 1208–1220, 1995.

[60] Hu C., Tam S.C., Hsu F.C., Ko P.K., and Terrill K.W. Hot electron model induced MOSFET degradation - model, monitor and improvement. *Transaction on Electron Devices*, 32:375, 1985.

[61] Takeda E. and Suzuki N. An empirical model for device degradation due to hot carrier injection. *Electron Device Letters*, 4:111, 1983.

Chapter 6

Thin oxides dielectric & device reliability impacts on ESD designs

6.1 Introduction

In this chapter, the extensive characterization work done in the two previous chapters will be applied to the ESD protection concept development. The strong motivation to prevent thin GOX degradation during voltage overshoots caused by ESD events is the main focus of this thesis.

There is one major critical cause responsible for GOX failure which is the high voltage drops induced by very fast and high current peaks discharges such as CDM or ESD system level during the first peak of the stress. For this kind of events, the clamping voltage and overshoots resulting from the protection concepts constitute a major concern. In advanced sub-micron CMOS technologies, this problematic is driven by the need to protect sensitive thinner oxides and is also enforced by the ever-changing nature of the protection elements which are basically not easily re-useable from one technology node to the next scaled ones. The reasons are various, operation at lower voltages, leakage consumption, increase of application features and complexity .... In CMOS technologies below the 100 µm node, low voltage concepts are mainly based on thyristors derivative concepts and diodes. The ability of these protection elements to turn-on fast enough to avoid rapid high overshoots is currently one hot topic [1, 2] in the ESD community. The SCR elements are known to be relatively slow reactive structures to rapid discharges, consequently the expected steady-state clamping voltage may not be reached during an ESD event in the first part of the pulse. It has also been recently reported that in sub-micron technologies even for basic elements as gated-diodes, overshoots can be observed for fast rise times [3, 4]. Their long reaction in time, which is due to the forward recovery effect, matters.

The need of a careful ESD characterization and relevant analysis of the protection elements with respect to over-shoots and reactivity is now strongly required [5]. To assure the efficiency of the ESD protection until the robustness target level, ESD developers need to incorporate the ex-

act stressing conditions to provide adequate protection concepts and evaluate potential weakness risks. As the GOX breakdown constitutes one major boundary condition for the ESD design, the understanding of the GOX breakdown driving parameters and the precise estimation of the stress impacts regarding their amplitudes and pulse widths are required. For this, the TDDB characterization of GOX in the ESD time domain constitutes the elementary basic investigation. The discovery of an common thin oxide breakdown voltage acceleration behavior from the long DC stresses down to short ESD pulses results in a powerful ESD design tool based on the universal properties of the thin GOX degradation. For thinner oxides as 7 nm, this ESD design tool can be described by the power law TDDB voltage acceleration and its set of inflating parameters:

- the oxide thickness
- the active GOX area
- the statistical nature of the oxide breakdown
- the stress polarity
- the device type
- the temperature

For the development of the ESD protection, the precise oxide breakdown knowledge as a function of the time stress is explicitly mandatory for the accurate choice of the protection strategy and for the selection of the protection concept. This is for example the case for fast discharges where the failing oxide monitors provoked by over-stress voltage from protection elements under vf-TLP can be correlated with the TDDB laws found down to the nanoseconds. The oxide breakdown values in the range of few to several picoseconds from the literature [2] are in very good agreement with the TDDB laws found in this study.

The impact of the fast overshoots induced by slow protection structures on thin gate oxides can be also quantified by means of the power law described previously. The ESD design is nowadays focusing on the rising importance of very fast transient phenomena and the precise information of the oxide breakdown as a function of the ultra-short stresses is becoming inevitable in the selection of the adapted counter measures in case of slow elements (e.g. local second stages to protect thin oxides, additional fast power clamps in parallel of the major one, *RC* elements, ...). However, beyond the obvious reasons of the ESD protection choice, the restrictive ESD spacing rules as a consequence of hazardous voltage drops from the wiring and busses impacts strongly the whole IC layout. For I/O cell library and ESD concept optimization, the precise evaluation of this boundary condition through the TDDB laws permits the definition of aggressive small concepts which have a considerable impact on the I/O cell library area as well as on the IC layout, and, thus, finally on the costs. Doubtless, optimized ESD protection concepts allow small chip areas which gives a clear competitive edge.

For the implementation of ESD concepts, the first step that ESD designers have to care of is the estimation of the oxide time-to-fail as a function of the voltage in the various short times ranges which could occur during the discharges. Evaluating and obeying the gate oxide TDDB laws is the starting point of the ESD development flow leading to a safe ESD design window. In that target, there are different ways to derive the required TDDB voltage acceleration laws,

- The first and most accurate one, consists in the full characterization of the thin GOX breakdown behavior obtained by the state of the art reliability testing methodology over a very wide range of time/voltage and consequent statistics.
- Another smart way is the direct extrapolation of the TDDB voltage acceleration from the long term DC qualification data characterized by reliability departments. The advantage of this methodology is that no additional testing efforts and costs are needed and moreover reliable data should be available for all technologies and oxides thicknesses.
- The third approach is based on extrapolations from published data scaled accordingly to the adequate parameters (cited previously) to the need of the technology of interest.

In a first part, the determination of the oxide time-to-fail voltage acceleration laws in the ESD regime will be described. In a second part, the degradation nature of the oxides with their failing criteria modes will be considered and confronted in order to give an ESD flow guidance for the achievement of a safe, reliable and robust ESD design window. In a third part a focus on the application of this ESD GOX characterization package and ESD design flow will be exposed.

6.2 Gate oxide TDDB laws determination in the ESD time domain

6.2.1 Gate oxide TDDB laws from extrapolation methodologies

For semiconductor companies that do not can spend time and money (silicon testing, resources and set up) for an accurate oxide characterization in the ESD domain, an alternative to GOX testing could be the extrapolations methodologies.

6.2.1.1 TDDB extrapolation based on long-term reliability data

In the lucky case where an internal reliability department can provide long-term GOX breakdown data, the oxide breakdown values in the ESD regime could be deduced through the extrapolation of the TDDB empirical models available in the low voltage range.

For the extrapolation procedure based on the long term reliability data, the main concerns come from the difference in the temperature settled for the testing. The elevated temperature used during the oxide lifetime qualification reduces significantly the TDDB values in the DC domain at low voltages. However, the effect of the temperature on the time-to-breakdown of thin oxides at high voltages (in the short-time stress range) is found to be considerably less than for low voltages. If the temperature extrapolation is done in the DC regime according to the DC reliability models, the TDDB will strongly increase towards the room temperature and then, if the power law acceleration deduced from the data at 140 °C is used, the TDDB extrapolated values in the ESD regime will be too high in comparison to the real values. The temperature effect has to be handled properly towards the ESD domain to correct this too optimistic estimation.

An exact determination of the power law in the ESD regime for 25 °C extrapolated from the reliability data is not straight forward. For a fixed voltage level, the temperature acceleration of the time to breakdown can be modeled via an Arrhenius law with an fixed activation energy. Over wide voltage ranges, this thermal activation energy is a function of the voltage and thus the temperature impacts the voltage acceleration parameter n (power law exponent, see Equation 6.3). The power law acceleration factor has a net impact on the long-range extrapolation and the error in this case is very large so as the risk to generate oxide breakdowns. The activation energy decreases towards larger voltages, implying an higher voltage acceleration for lower temperatures. Two approach can be used:

1. either extrapolating first the 140 °C data at low voltages to the room temperature values and then using an adapted lower voltage acceleration exponent to adequately models the TDDB at room temperature in order to extend the extrapolation in the short times range.

2. or extrapolating first the 140 °C data to higher voltages, and then using the appropriate temperature extrapolation valid in the ESD regime. This implies the exact knowledge of the activation energy dependence with the voltage in the full extrapolation range.

Unfortunately, for an exact voltage acceleration modeling at room temperature, data over a wide voltage domain covering the DC domain are required. However, these data are generally not available as they are out of the interest range for the lifetime qualification. Generally the TDDB scaling law with temperatures used for the reliability estimations are based on interpolation and not on extrapolation. The TDDB dependence law as a function of the temperature is estimated in a temperature window located around the product heating value under operation. The extreme extrapolation towards room temperature using the models give correct values for a small voltage range. For the wide voltages/times extrapolation required, either an exact thermal time-to-breakdown model coupled with the thermal voltage dependence or a conservative way should be chosen to cover the lack of accuracy generated from the conventional extrapolation models which increase a lot the error bar at the end of the evaluation.

The second point to consider for the extrapolation is the universal kink seen in the TDDB voltage acceleration. The extrapolation from the reliability data are valid until this kink only.

At that point, a lower voltage acceleration should be used. In practise the methodology (1) is better suitable and do not need complex temperature modeling. Once the first part of the TDDB power law is established at room temperature, a power law with the lower exponent seen at 25 °C can be directly used in the continuity. The second voltage acceleration domain can be also disregarded as first conservative approach, but this can lead to drastic over-design especially for medium oxides and furthermore regarding very short pulses (CDM range).

As a general procedure to achieve a safe ESD design, a good oxide breakdown extrapolation procedure considering worst case conditions for the modeling of the breakdown can be based on the following elements,

(a) the voltage acceleration below the kink can be fixed to the worst-case power law exponent of the nFET stressed in inversion, $n \sim 50$.

(b) the voltage at which the voltage acceleration decreases comes around 6 V. Only in the case of the nFET stressed in inversion, the occurrence of the kink can be fixed to 5 V.

(c) above the kink the TDDB voltage acceleration at room temperature can be estimated with the worst-case value of 30.

The application of the method (1) based on the worst case condition is demonstrated in Figure 6.1.

6.2.1.2 TDDB extrapolation from theory and empirical models

For small companies that do not have access to any reliability data, a simple alternative approximative methodology would be to make use of the characterization data published in the literature and to use the powerful reliability theories. Indeed, there is already a huge empirical data base available in the literature on which the designers can rely on [6, 7, 8]. The main task is to adapt the oxide thickness to the correct value, using the TDDB thickness acceleration reported. However, with an higher uncertainty risk, it is suggested to use as reference the worst-case conditions and to extract the data exposed for nFET devices stressed in inversion. The wide range of data characterized in this work, see Chapter 4. [9] could provide a good base as well as the numerous reliability publications which can be used accordingly to the previous extrapolation procedure.

6.2.2 TDDB laws determination based on the GOX testing methodology

As proven in Chapter 4, the statistical nature of gate oxide breakdown and voltage acceleration dependence in the ESD regime are similar for DC and long-term stresses. This fact implies to transfer (or adapt) the extensive oxide degradation know-how developed by the reliability community into the ESD world. The state-of-the-art testing flow and methodology used for

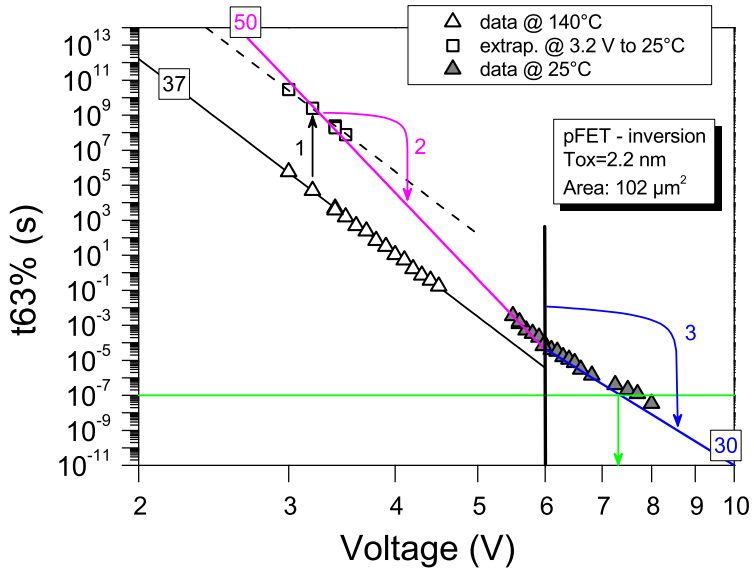

Figure 6.1: Oxide breakdown voltage extraction from long-term reliability data extrapolation. 1) extrapolation of the 140 °C process qualification data to 25 °C at one low voltage using the reliability DC temperature extrapolation law. 2) modification of the voltage acceleration from 140 °C process to 25 °C. 3) lower voltage acceleration occurring at 6 V for pFET stressed in inversion. Finally, the oxide breakdown voltage can be extracted.

the reliability qualification of thin oxides process should especially be considered in the GOX characterization procedure under ESD stress. The correct stressing procedure is the first brick to apply for the accurate determination of the oxide failures. That is why a lot of effort is spent in reliability departments for the understanding of the testing environment impacts on the results. From the investigation done in the nanoseconds regime on comparing the different stressing methodologies, the preferred stressing procedure is the constant voltage stress. This one has less parameters uncertainty than the other stressing procedures and enables the direct use of the powerful statistical package well characterized by the reliability community, as defined by

1. the use of the Weibull statistics to describe the cumulative failure distributions as a function of the time-to-breakdown

2. the Poisson statistics law for the time-to-fail dependence with the gate oxide area

Nevertheless, the use of the VRS methodology to determine the TDDB in the ESD regime leads also to the correct breakdown voltage acceleration, but is shifted to lower time values due to the cumulation of the increasing voltage steps during the run, see Figures 6.3 and 6.2. The impact of this effect depends on the power law voltage acceleration and is thereby more

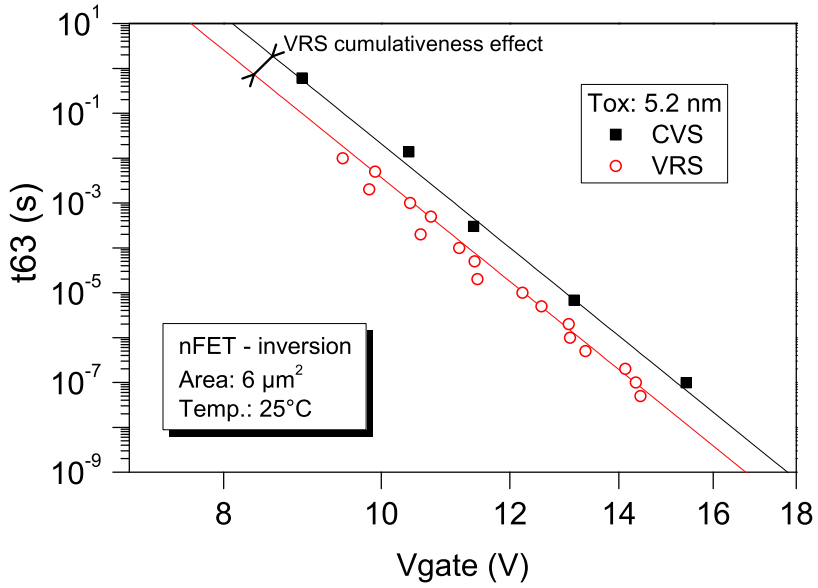

Figure 6.2: Time Dependent Dielectric Breakdown of 5.2 nm nFET devices stressed in inversion by means of VRS and long CVS.

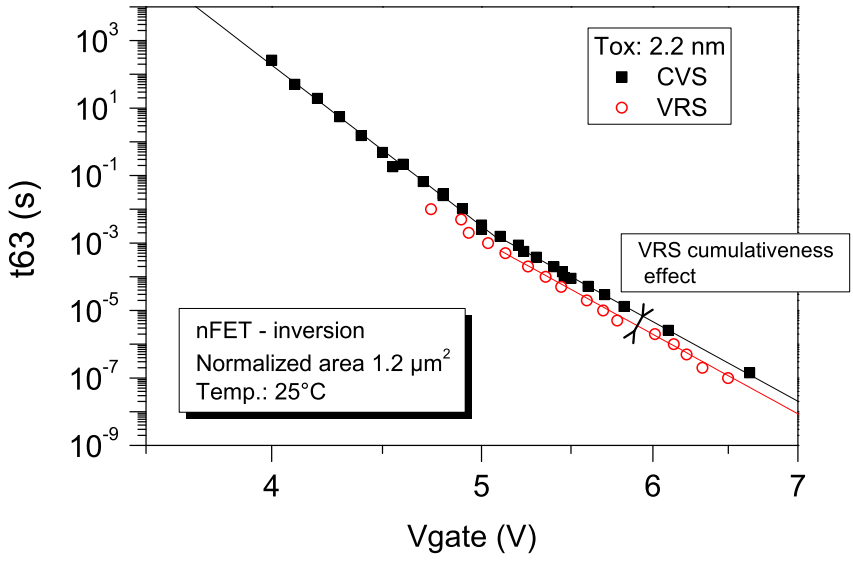

Figure 6.3: Time Dependent Dielectric Breakdown of a 2.2 nm nFET stressed in inversion with VRS and long CVS. The time-to-breakdown from the CVS are normalized to the area of the GOX devices tested by VRS, being 1.2 µm^2.

important in the lowered acceleration exponent voltage domain (above 5-6 V). Although, the transformation from breakdown voltage distributions to time-to-breakdown ones is possible and gives a correct correction for the cumulativeness of the ramping methodology, the uncertainly level introduced by the ramp rate and possible relaxation effects can be avoid by the direct and simplest long CVS method.

The correct oxide breakdown evaluation in the ESD regime is deduced from the extraction of the $T_{63\%}$ from large cumulative failure distributions obtained via long CVS. Nevertheless, short repetitive CVS with sensitive current accuracy must be employed for ultra-thin oxides where progressive breakdowns occur. This case has been reported in 4.3.2 for the pFET stressed in accumulation for example. The use of large statistical data is a mandatory pre-requisite for the GOX breakdown estimation and the sample size should be fixed in proportion to the time-to-breakdown spreading of the cumulative failure distributions imposed by the oxide thickness. A large samples size is especially required for ultra-thin oxides due to the shallow Weibull slopes modeling the distributions. The accurate extraction of the Weibull slope, which is primordial to avoid large error in the extrapolations, should be determined from the harsh evaluation of time-to-breakdown resulting from different area scaled structures layouted in conformity with the reliability testing requirements.

6.3 ESD development flow to the safe and robust ESD design window

The gross gate oxide time-to-fail voltage acceleration is not sufficient for fixing a safe ESD design window, although it constitutes the first basis required. The gate oxide degradation nature needs to be considered in details for the definition of failure criteria.

6.3.1 GOX failure criteria under ESD stress

6.3.1.1 Statistical failure criteria

From the correct testing methodology previously exposed, the GOX hard failure trend as a function of the physical oxide thickness can be sketched from the characterization data exposed in the chapter 4. This trend is exposed in the CDM regime for nFET and pFET devices and for accumulation and inversion stress in Figure 6.4. The similar figure has been depicted in the HBM regime in a semi-log plot 4.36, see Figure 4.3.5. The ESD robustness hierarchy of nFET and pFET with respect to the stress polarity can be deduced and allows the identification of an ESD weakness, for the choice of appropriate and optimized ESD protection concepts. It is shown that nFET in inversion is the most critical case and should be carefully considered. The breakdown voltage trend of thin oxides accordingly to the oxide thickness is very challenging

with the exponential decrease of the breakdown voltage coming along with the diminution of the oxide thickness. However, it should be mentioned that this trend is specific to the area and failure criteria chosen, as the TDDB voltage acceleration is dependent on the stress polarity and on the device type. Consequently the voltage boundary trend depends on the area and on the choice of a meaningful statistical cumulative failure level. Indeed a net distinction should be done regarding the $T_{63\%}$ parameter and the oxide breakdown limit for the ESD design window. From a testing point of view the GOX breakdown definition from the extracted $T_{63\%}$ value is justified for two reasons:

1. the $T_{63\%}$ characterizes the Weibull cumulative failure distribution
2. for the $T_{63\%}$ value, the confidence bound level of the experimental data is maximal [10]

However use of the $T_{63\%}$ criterion is not correct for the ESD design as it would implies that 63% of the ESD stressed devices will fail. For a robust ESD design, an other statistical failure criterion must be fixed.

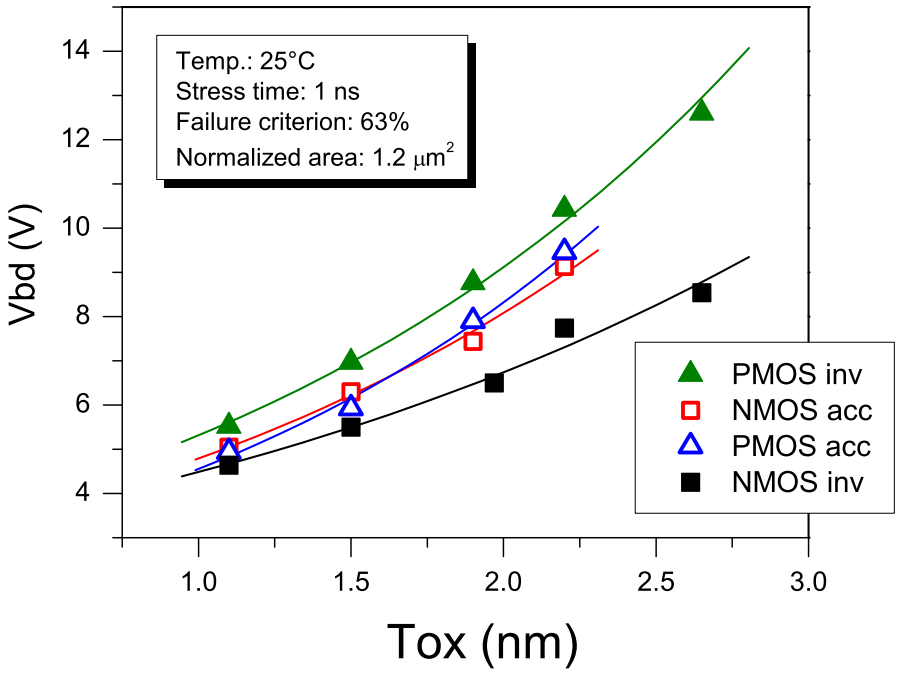

Figure 6.4: Breakdown voltage acceleration as a function of the physical oxide thickness for nFET and pFET stressed at 25 °C in the CDM range of time for a normalized size of 1.2 μm^2.

6.3.1.2 The cumulative failure level

Derived from the Weibull statistics equation 1.44, the time to breakdown at the failure criterion F can be extrapolated from the $T_{63\%}$ value according to the following equation:

$$\log(t_F) = \log(t_{63\%}) + \delta_{(\beta,F)} \quad \text{with} \quad \delta_{(\beta,F)} = \frac{\ln(-\ln(1-F))}{\beta \cdot \ln(10)} \tag{6.1}$$

The impact on the TDDB of a lower statistical failure criterion is exposed in the Figure 6.5. The TDDB power law of 2.2 nm nFET stressed in accumulation is scaled to a cumulative failure criterion of 1% which causes a lower shift of the TDDB power law of 0.9 V. The failed samples resulting from the different CVS are plotted with red bars. The spreading of the about 100 samples are kept above the 1% failure criterion power law.

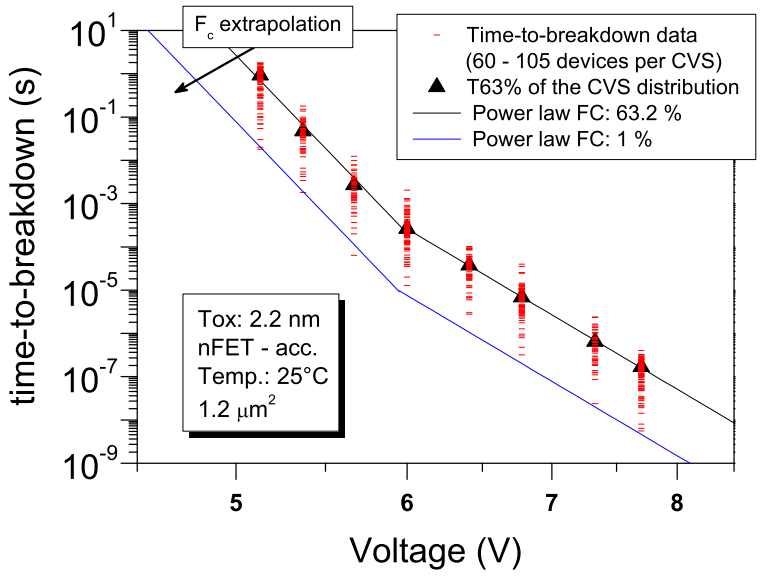

Figure 6.5: TDDB voltage acceleration of 2.2 nm nFET devices stressed in accumulation. The failed samples per CVS level as well as the $T_{63\%}$ criterion extraction are detailed with scattered data. The $T_{63\%}$ power law as well as its extrapolation to 1 % cumulative failure are drawn.

The linear shape of the distributions plotted in a Weibull plot is expressed with the parameter β (Weibull slope). These Weibull slopes are getting shallower with thinner oxide, equivalent in a spreading of the cumulative failure distributions through the time. The experimentally extracted Weibull slope as a function of the physical oxide thickness is depicted in the Figure 6.6. These slopes are observed to decrease with the oxide thickness accordingly to the following equation 6.2. This trend is in good agrement with the reported data and recent work done on the

modeling of the percolation theory [6, 11, 12]. However, an even more precise dependence of β with the oxide thickness can be used as detailed in the section 4.3.7.

$$\beta \simeq 0.46 * T_{\text{ox}} + 0.52 \tag{6.2}$$

At one stress voltage level and for a fixed cumulative failure criterion, the reduction of the time-to-fail imposed by the shape of the distribution increases for thinner oxides (Figure 6.7). A reduction in the time to fail also means a reduction in the gate oxide voltage limit (Figure 6.5) and thus should be accounted for the ESD design. Theoretically the reduced time-to-fail for thinner oxides, results in a faster shrinking of the ESD design window for thinner oxides if the same acceptable statistical cumulative failure criterion is kept.

If we consider the TDDB dependence of the cumulative failure level given by the equation 6.1 and the time-to-breakdown voltage acceleration described with the following power law expression,

$$T_{\text{BD}} = K_0 V^{-n} \tag{6.3}$$

the reduction in the voltage ($\Delta V_{(\beta,\text{F,n})}$) required for the ESD safe design consequently imposed by the choice of an acceptable cumulative failure level can be expressed by:

$$\Delta V_{(\beta,\text{F,n})} = 10^{\left[\frac{-\log(t)-K_0}{n}\right]} \left[1 - 10^{\frac{\delta_{(\beta,\text{F})}}{n}}\right] \tag{6.4}$$

where n is the power law exponent from equation 6.3, β is the Weibull slope and F, the cumulative failure level.

The figure 6.8, illustrate the impact of the cumulative failure level and oxide thickness parameters on the voltage reduction ($\Delta V_{(\beta,\text{F,n})}$) that should be considered for an reliable ESD design with respect to hard HBM fails of 1.2 μm^2 gate inputs. The shrinking of the ESD window for thinner oxides is somehow relief by the higher power law exponent of the TDDB voltage acceleration in the lower voltage domain (below 5-6V). However, this voltage reduction should be considered accordingly to the ESD design window available in order not to under-evaluate the reduction impact for the design. The boundary voltage ($V_{\text{ESD,limit}}$) applicable for the design regarding this hard breakdown criterion is given by

$$V_{\text{ESD,limit}} = V(T_{63\%}) - \Delta V_{(\beta,\text{F,n})} \tag{6.5}$$

The choice of the statistical failure criterion is not straight forward, The criterion should be carefully chosen to ensure a safe ESD design but should not be too restrictive to avoid over-design costs. The cumulative failure limit obtained from the testing is given for an entire population of ESD stressed devices and is not corresponding to the entire product population as it should be. Obviously, not all products or pins will be stressed by an ESD event and we need to include the probability of an ESD occurrence. This independent criterion is a considerable

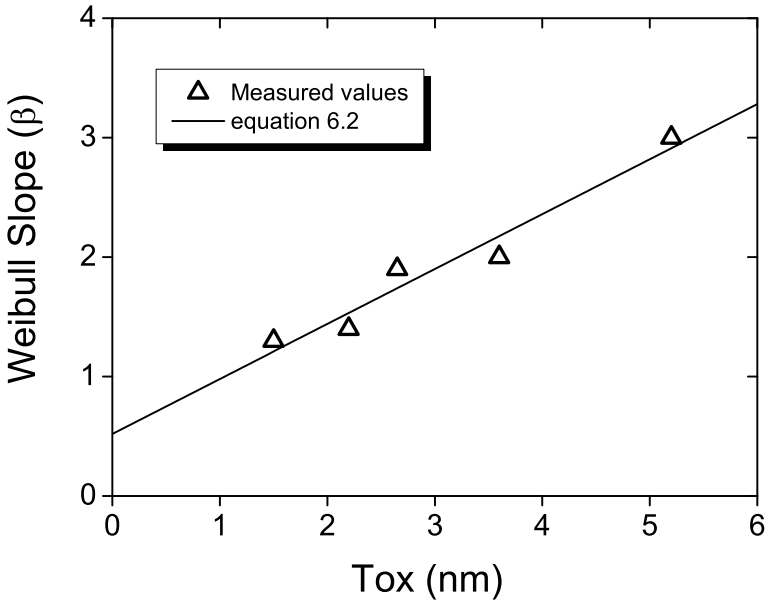

Figure 6.6: Experimental Weibull slopes β plotted as a function of the physical oxide thickness.

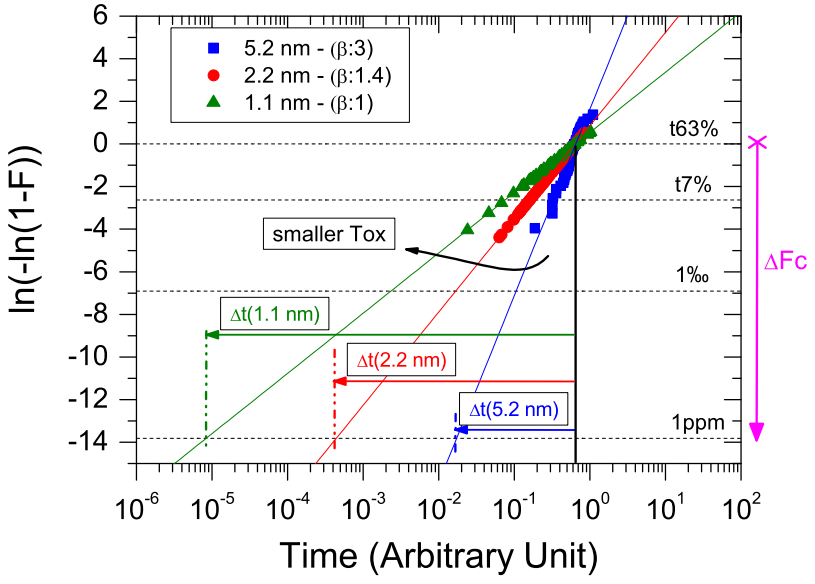

Figure 6.7: Impact of the Weibull slopes β (oxide thickness related) on the time to fail as a function of the statistical Failure Criteria (FC), which is the cumulative failure level [9].

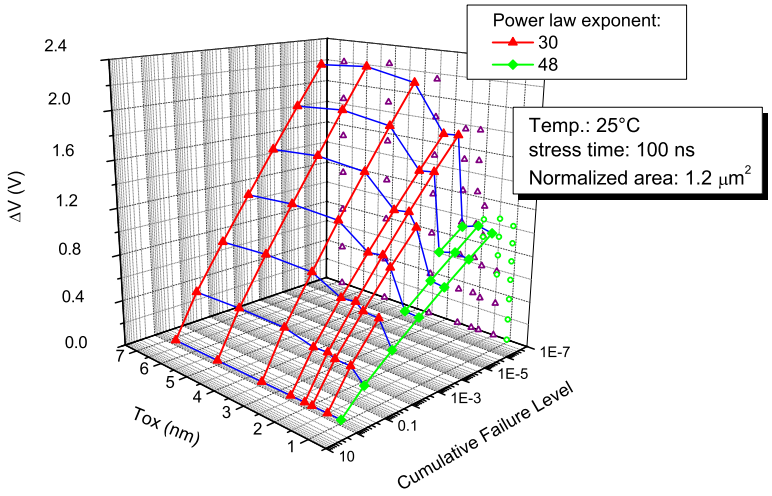

Figure 6.8: Voltage margin from the hard breakdown imposed for the ESD design window in the HBM time domain for an oxide are of 1.2 µm^2 as a function of the oxide thickness and cumulative failure level. This voltage margin integrates also the variation of the TDDB acceleration factor adequately to the voltage domain involved.

benefit for the ESD gate oxide robustness because the probability of an ESD occurrence in the field is low. Considering not just the stressed devices but the ESD stressed elements out of all products sold; it will reduce the total effective GOX breakdown failure rate for the definition of the upper limit of the ESD design window. Anyway a value matching the political quality of the company is required for the design and a discussion on its choice will be covered later after the presentation of all impacting variables.

6.3.1.3 The statistical influence of the Area on the breakdown

As proven in the chapter 3, the gate oxide breakdown can be described as a random process under ESD stresses. This statistical occurrence accordingly to the Poisson statistics imposes the incorporation of the area parameter in the GOX breakdown evaluation. Under testing condition, the GOX structures used are very small to avoid parasitics effects and to monitor intrinsic breakdowns. An extrapolation towards bigger areas, meaning earlier time to fail and reduced ESD design margin, is then required for the protection of the typical oxides areas present in input and output stages, so as capacitors. Concretely, for a stress at one pin, all the sensitive gate oxide area present in the discharge pathes should be summed-up before applying an acceptable cumulative statistical failure criterion.

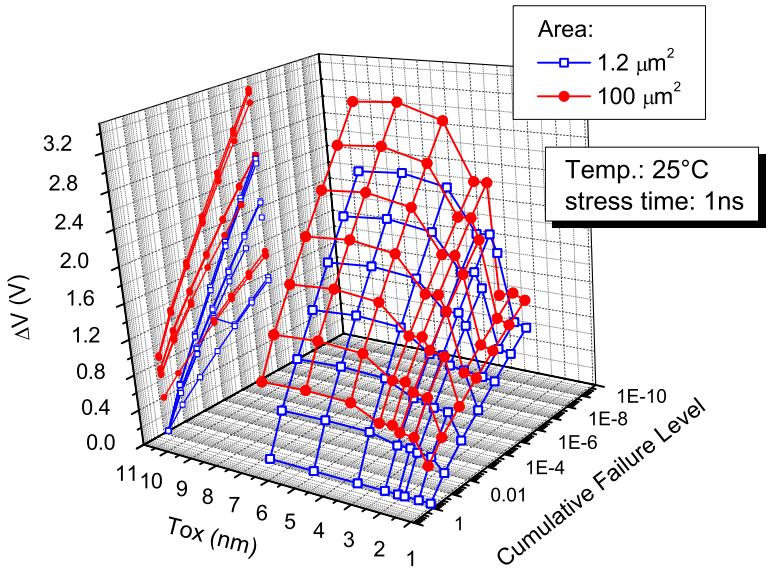

Figure 6.9: Voltage margin from the hard breakdown imposed for the ESD design window in the CDM time domain as a function of the oxide thickness and cumulative failure level. The impact of the gate oxide area is exposed for for 1.2 µm^2 and 100 µm^2.

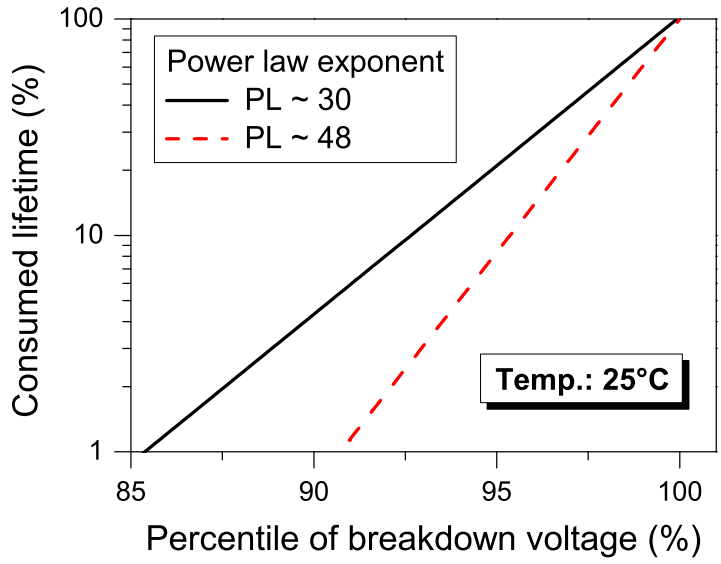

Figure 6.10: GOX lifetime consumption as a function of the stress level for the two typical voltage acceleration factors of nFET stressed in inversion [9].

The gate oxide area impact on the margin to consider for the CDM safe design is exposed in the plot 6.9 where an increased area about a factor 100 is calculated.

6.3.1.4 Cumulativeness dielectric degradation criterion

Let's define the GOX breakdown voltage (V_{bd}) as the value obtained for the $T_{63\%}$ failure criterion and assume that over-voltage stresses degrade the GOX dielectric lifetime with a pure cumulative process (at least to a first order). This hypothesis seems to be fair considering the VRS correction methodology exposed in 4.2.2. Based on the TDDB power voltage acceleration law, the calculated dielectric reduced lifetime as a function of the stress level is shown in Figure 6.10 for the two typical acceleration values of nFET stressed in inversion, being 48 and 30.

- A deduced safe margin of about 10 % to 15 % from V_{bd} should be considered to ensure no significant dielectric lifetime consumption. This margin is only attributed to the TDDB power law voltage acceleration.

6.3.1.5 Device degradation criteria

The observation of the oxide to device degradation should contribute to the establishment of a failure criterion taking into account the limit allowed for the device parameters drifts. This criterion can be settled as a percentile of the breakdown voltage (establishment of a margin) where the device degradation is considered too high for a specific analogue application. An exact limit is hard to define because of the relaxation effects, specificities of the application and also because of the fear to generate possible latent damages. Long-term reliability impacts should be integrated in this device degradation criterion. The impact of moderate ESD stresses is dependent on the oxide thickness and on the stress configuration. For pull-up and pull-down as well as for drivers an intensive amount of work have been investigated on degradation related to this particular "gg-MOS" stress configuration (see section 5.1.2.1). Concerning the issue leading to soft fails related with drain to source filamentation, investigations on that topic constitute one essential part of the ESD protection development engineering. These elements need special and careful ESD layout, such as blocking salicide or non butted substrate layout to enable a better ballasting of the current and therefore prevent filamentation effects for example. Countermeasures solutions exists and are directly related to the ESD concept development. In case where the leakage degradation are emanating from the protection elements or from the drivers (self-protecting or not), it means that the ESD protection concept is simply not adapted and need to be modified. Concerning the injection and trapping of carriers during the snap-back and the effect induced on the devices, the topic is known but is more delicate to deal with. No quantitative rules are generally settled to avoid oxides degradation in that way. This oxide-trapping topic can be included with the inputs and buffer capacitors type of stress, where a related voltage level can be linked with defined failing criteria as AC/DC functional parameters or devices reliability

targets. For a clean ESD design, no effect of the ESD events should be theoretically noticed. A tradeoff between uncritical drifts and the strictly conservative approach is needed to guarantee reliable designs without inconsiderate limitations. The failing criteria should be clarify basically for the sensitive analogue interfaces with the designers.

6.3.2 Failure criteria selection for the ESD design window

6.3.2.1 Failure criteria and ESD design

The failure criteria cited above lead to a multiple restrictions regarding the oxide integrity for the upper-limit of the ESD design window (Figure 6.11). The boundary condition given by these oxide limitations are not fixed and are strongly thickness dependent.

(i) The oxide breakdown depends on the oxide thickness, cumulative failure criterion, area, device type, stress polarity and temperature.

(ii) Concerning the dielectric and device degradation,

1- the cumulativeness impact of the dielectric degradation depends on the TDDB voltage acceleration which is a function of the device type and voltage domain. An indirect thickness dependence can be rated as the thin oxides breakdown in the ESD regime are in an other breakdown acceleration voltage domain than the thick oxides.

2- the oxide to device degradation depends also mainly on the oxide thickness,then on the device type, circuits and application.

As the principal parameter that impact the upper-limit voltage of the ESD design window is the oxide thickness, the confrontation of the different failure criterion will be discussed in the following section as a function of the oxide thickness. For the selection of the limiting boundary parameter, this voltage limit (or the margin voltage from the breakdown voltage) will be expressed in percentage to enable the comparison of the different criteria.

6.3.2.2 Thick oxides

In order to identify the limiting criterion for thick oxides, all constraints defined in the previous part will be reviewed with a focus on a 5.2 nm oxide. For this category of thick oxides, the voltage acceleration at room temperature is in the range of 20 to 30 depending on the device type and stress polarity. In the following, the worst acceleration of 30 is considered.

- First, the voltage reduction settled by this TDDB voltage acceleration regarding the cumulativeness of the dielectric degradation must be accounted for. From the Figure 6.10, a fixed margin of 10 % to 15 % from V_{bd} is mandatory to exclude any reduction of the lifetime potential from the dielectric during an hazardous ESD event.

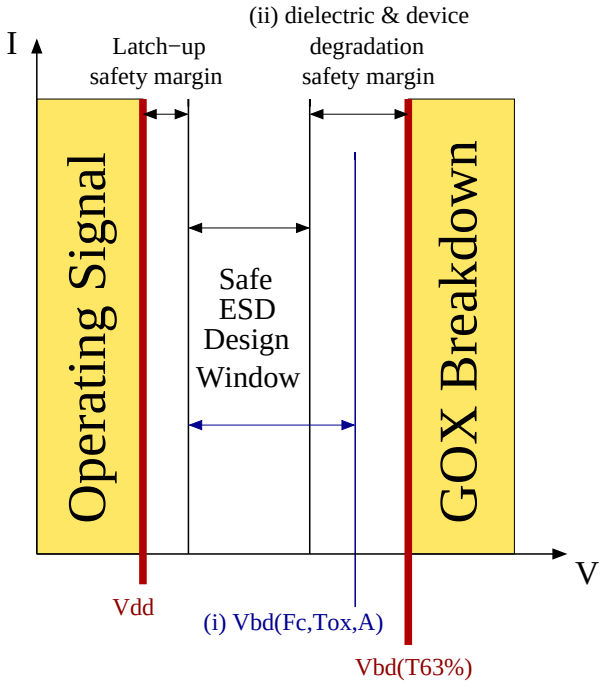

Figure 6.11: ESD design window considering two different limiting criteria regarding GOX degradation: (i) the dielectric breakdown as a function of an acceptable statistical F_c, (ii) the dielectric lifetime consumption and device degradation limiting criteria (given as a percentile of V_{bd}). The order of the two mentioned criteria is arbitrary.

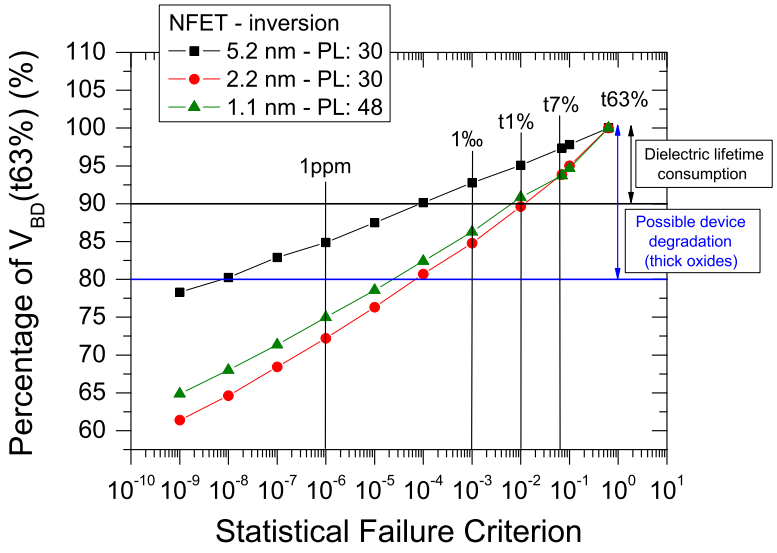

Figure 6.12: Correlation between the statistical failure criteria and the percentage of the breakdown voltage (V_{bd}) for nFET stressed in inversion. Three oxide thicknesses are considered (1.1 nm, 2.2 nm, 5.2 nm) [9].

- For a fair comparison of the statistical criterion with oxide degradation levels, the percentage of the breakdown voltage is plotted versus the cumulative failure level for a 5.2 nm oxide in the Figure 6.12. In the correlation established between the failure criterion resulting from the statistical nature of the GOX breakdown and the percentage of the GOX breakdown value (Figure 6.12), arbitrary area and temperature are taken into account. The appropriate conditions should be considered for the evaluation of the GOX breakdown voltage reference ($T_{63\%}$).

  For the 85 % to 90 % of V_{bd} a low cumulative failure level of 1 failure per thousand is reached. This effect is due to the steep Weibull slopes of thick oxides, which enable to reach low cumulative failure level without a large reduction of the ESD design window. Without including the probability of an ESD occurrence in the field, this statistical failure level is already quite acceptable. For thick oxides, a safe voltage margin seems to be appropriate to ensure a reliable ESD protection.

- In a second step, an extended safety margin from V_{bd} can be fixed to assure the integrity of the device functionality. In case of thick oxides, the oxide to device degradation (V_{t1}, I_{on}, I_{off}, ...) have been reported in the chapter 5 to appear around 80 % V_{bd}, see Figure 5.16. For sensitive analogue transistors, specified drift limits can be used. In a total conservative approach, this limit can be established at the early starting point of parameter drifts in the range of 75 % to 80 % of V_{bd}. For these values a very low cumulative failure level of 100 ppm is achieved. For this criterion, no long-term device reliability are expected as the impact of the DC hot carrier injections performed on nFET devices pre-ESD stressed was not significant.

For the few number of analogue pins per ICs in the communication business, the focus on the starting point of degradation of these thick oxides devices could be used as the failure parameter for the robust ESD design. The target is also accessible, for HBM type of stresses a clamping voltage around 10 V is necessary for reasonable sized oxide area. In case of fast discharges however a more detailed analysis of the triggering behavior from the protection concept is required to guarantee that the clamping voltage or the transient over-shoot is low enough.

6.3.2.3 Thin oxides

As for thick oxides, the selection of the critical boundary limit will be discussed based on the impact of each criterion.

- Let's consider also at first, the voltage reduction settled by the TDDB voltage acceleration regarding the cumulativeness effect of the dielectric degradation. For thin oxides depending on the thickness and on the area, the TDDB voltage acceleration could vary from a low coefficient around 30 to a higher one in the range of 45 to 50. From the Figure 6.10,

a fixed margin of 10 % to 15 % from V_{bd} is mandatory to avoid any dielectric lifetime consumption. For thin to ultra-thin oxides a margin in the range of 10 % is more relevant to consider for two reasons. The first one because the breakdown voltage are more located below the voltage acceleration transition (5-6 V) and secondly as for the thin oxides which are used in logic blocks, a relaxed long-term reliability behavior is reported due to softer breakdowns [13, 14, 15, 16]. Moreover the impact on the dielectric lifetime is anyway not too significant at 90% of V_{bd}.

- Secondly let focus on the oxide to device degradation. For thin oxides, thanks to the de-trapping via direct tunneling, almost no shifts in parameters are seen below 90% of V_{bd}. No latent device degradation in the long reliability term have been observed. This 10% margin from V_{bd} is sufficient.

- In the case of thin oxides the cumulative failure criterion can be the limiting parameter if judged too high. In the Figure 6.12, the cumulative failure level of 1.1 nm and 2.2 nm nFETs stressed in inversion are plotted as a function of their breakdown voltages. For both of them, an upper voltage limit fixed for 90 % of V_{bd} corresponds to a statistical failure level around 1 %. This higher value in comparison to thick oxides comes from the very shallow Weibull slopes.

For the definition of the upper limit of the ESD design window, it appears that a percentage of V_{bd} is a sufficient failure criterion for thick oxide whereas for thin oxide the statistical failure criterion should be carefully considered.

Discussion on the choice of the cumulative failure level for thin oxides

The influence of the cumulative failure level on the voltage margin from V_{bd} is exposed in the Figure 6.13. The severity of this criterion on the required margin is not a straight function of the thickness as possibly deduced in a first thought. Due to the different voltage acceleration above and below the universal kink, a discontinuity appear in the voltage margin increase with the oxide thickness. A clear peak can be observed, resulting from the thin oxides breakdown laws which are switching from a high voltage acceleration to a lower one above 5 to 6 V.

This is not an easy task, neither to evaluate the probability of an ESD occurrence in the field nor to specify an allowed failure limit due to ESD events. An other boundary is given by the width of ESD design window and the feasibility of a protection concept at low costs.

Anyway an other view on this criterion is given from the qualification aspect. The ESD qualification procedure certified the products to resit to high arbitrary ESD levels that have been observed to generate no ESD fails in the field. In the ESD qualification Standards (MIL,JEDEC, ESDA,...) no failure out of 3 devices are required. If we consider the Acceptable Quality Level (AQL) tables which fix an equivalent failure criterion as a function of the sample size and if we use the AQL values exposed in the MIL-STD-105D. It is specified that 0 fails on 3 corresponds

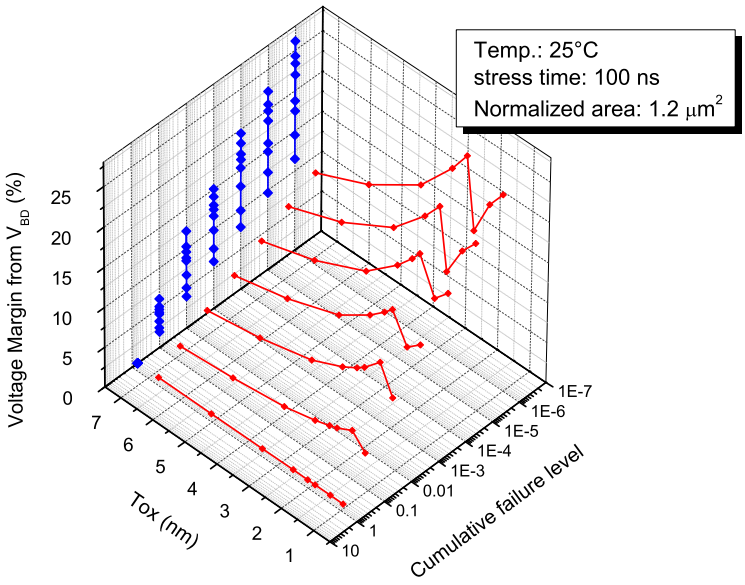

Figure 6.13: Design margin from the hard breakdown expressed in percentage in the HBM time domain as a function of the oxide thickness and cumulative failure level.

to 1 fail over 13, 2/20, 3/32, ... and corresponds to an AQL 4. This Failure criterion level is approximately 7%. This failure rate permitted is considering all pins and multiple stressed pins from the 3 ICs. In the field it is really unlikely that more than one pin would be involved in an ESD discharge event. In the case of qualification, all sensitive GOX present in all the pins will be stressed and this for different polarity and possibly with multiple stresses [17, 18].

The choice of the failure criterion requires at least to pass the qualification which depends on the number of pins of the IC. The number of pins (and crossed pins) stressed is acting as a sample size of gate oxide test structures. The size of the GOX test structures fixes a cumulative failure level. That means that the statistical failure criterion should be at least strictly lower than level fixed by the IC number of pins and by the virtual sample stressed size.

The procedure to follow to pass the qualification is:

1. regroup the similar pins as a function of their sensitive gate oxide area exposed during an ESD discharge. For each group of pins, an scaling of the referenced time-to-breakdown based on the $T_{63\%}$ to the correct area should be done

2. The number of pins in the groups can be interpreted as the number of similar stressed samples, the 3 ICs should be accounted for the total samples evaluation. For each groups a cumulative failure criterion can be extracted fixed by the sample size. This value gives the strict maximal level that can be theoretically fixed for the statistical failure criterion in order not to monitor one failure.

3. the number of stresses per I/O (polarity, power domains) must be accounted for the effective stress time endured by same GOX group domains during qualification

The voltage extracted from the scaled voltage acceleration from the effective stressing time accordingly to this procedure gives the minimum voltage limit to settled in the ESD design window for the group of pins considered. The question of the adequation of the deduced cumulative failure level with quality and reliability is still open.

Let's consider an example with 20 similar pins with 20 μm^2 oxide area. First the $T_{63\%}$ from the worst TDDB law of the nFET stressed in inversion should be selected and then scaled to total effective oxide area of 1200 μm^2 (20 μm^2 $\times 20 \times 3$) which includes the 3 ICs. The $T_{63\%}$ should be scaled to a failure level in the range of 1% (90 pins tested). For the effective stressing time, the 3 positives and negatives pulses applied during the test procedure should be accounted for, meaning an effective stress in the range of 600 ns (6×100 *ns*, a longer stressing value can be chosen also for the HBM effective voltage stress, *i.e.* 200 ns for example). The derived voltage value extracted at 600 ns from the TDDB power law scaled from this procedure will give the safe ESD window limit targeted for the HBM qualification of this group of pins.

The unknown parameter of the probability and reproducibility of ESD stress in the field was sidestep by using high current values. As the physics of the protection elements is quite stable and do not implies strong statistical variation and as the oxides where thick in the past this demarche was fine to guarantee a strong ESD robustness level that does not need regard for the statistical ESD occurrence.

The high pins count ICs and the multiplication of the supplies domains have a direct strong impact on the cumulative failure level and then on the probability to observe one GOX fails during qualification. The correlation of the field return crossed with the current levels fixed in the standards is questionable. During HBM testing a lot of artefact are generated during the test which are not relevant for the field. It is by the way shown in the white paper from the industry council [19]. In an other way the CDM is becoming the major kind of stress leading to failure in advance CMOS technologies.

6.3.2.4 Conclusion for the ESD design window

The ESD design window significant trend for gate inputs is summarized in Figure 6.14 regarding the worst limiting failure criterion that should be considered, 25 % of V_{bd} for thick oxides and for example 1 ppm for the thin oxide cumulative failure level. If we compare the proportion of the area dedicated for the ESD margin with respect to the ESD design window potentially available; this ratio is quasi independent of the oxide thickness. The ESD safety area required from the total ESD design window is about 30 % to 40 %. Of course the precision required in the small voltage range in case of thick oxides make the task really harder and complicated, indeed each small voltage drops matter.

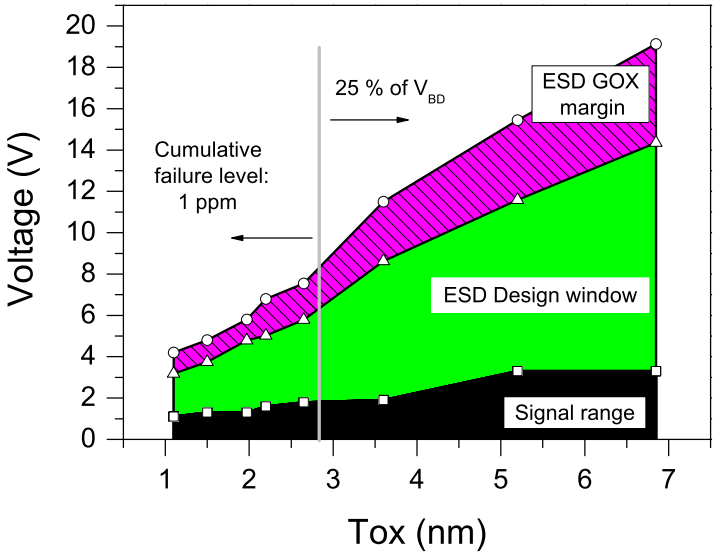

Figure 6.14: ESD design window trend considering the protection of gate oxide inputs.

6.4 GOX characterization package application for ESD safe design

6.4.1 Thin oxide buffer capacitors and ESD safe technology development

In this part the established oxide concerns which has be discussed in the previous sections will be applied on the ESD design needs of thin buffer capacitances.

In advanced CMOS technologies, the trend to decrease the supply voltage and to increase the clock frequency leads to an issue regarding the power supply quality [20]. The role of the decoupling capacitors are to preserve the stability of the power supply in an acceptable margin (around 10 %). For this application the capacitance value should be as high as possible. Consequently, it is less and less attractive to use thick oxide capacitances as the capacitance value as a function of the oxide thickness T_{ox} is given by $C = (\varepsilon_0 \cdot \varepsilon_{ox})/T_{ox}$. In order to achieve an optimum performance-size ratio in future technologies, the use of thin dielectric capacitors is inevitable. Let consider the ESD protection case of buffer capacitances (see Figure 6.15). Typically a diode and a power clamp are used as protection concept. In case of a negative ESD stress on V_{dd} versus V_{ss} (or from a positive stress on V_{ss} versus V_{dd}), the diode will trigger and the oxide will be well protected due to the good diode clamping voltage. But in the case of positive ESD stress on V_{dd} versus V_{ss} (or negative stress on V_{ss} versus V_{dd}), an adapted clamping protection element should

take over the ESD stress. This case is not easy as the protection element requires a fast triggering with a low clamping voltage. This is becoming crucial in advanced CMOS technologies because of the very low GOX breakdown voltage of thin oxides. These elements are clearly ESD sensitive. The ESD risk might be even higher than at input oxides due to the large oxide areas required in the applications as power supply coupling capacitors. Moreover, the direct connection to external pins (V_{dd}, V_{ss}) is obviously critical for ESD as the gate oxides endure plainly the stress. The improved performance evolution of the technologies is possibly limited by the ESD robustness parameter in this case. In the future, a trade-off between technology performance and ESD risk will have to be met concerning the thin oxide buffer capacitors.

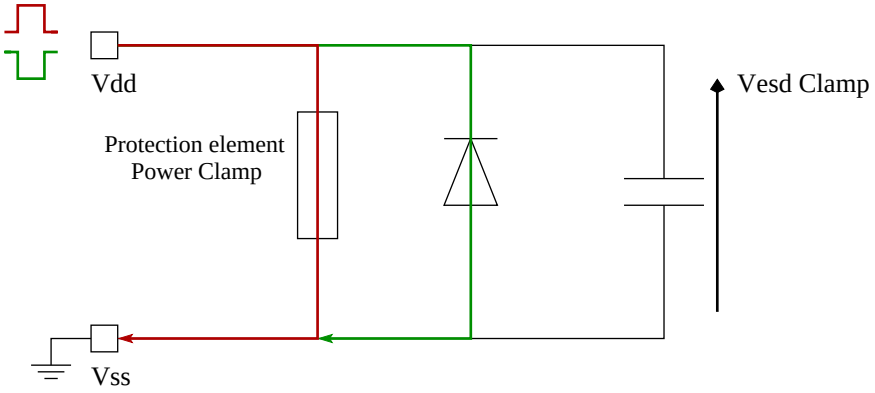

Figure 6.15: Schematic of the ESD protection elements involved in stresses between supplies.

As a case study, the design of buffer capacitors using nMOS, pMOS and nCAP in an 1.5 nm gate oxide CMOS process is discussed. Typical buffer capacitance values are in the range from several fF up to few pF, depending on applications and area availability. Common target values are in the the pF range, which requires an area of thin gate oxide in the order of 100 μm^2. In order to compare the ESD robustness of these different active oxide capacitors, the oxide breakdown normalized to the same useful condition is required. TDDB measurements of the different 1.5 nm elements have been performed in both polarities on small oxide area. The results have been discussed in Chapter 4.

First, the TDDB voltage acceleration laws of each devices for each polarity have been scaled from the 1.2 μm^2 structures used for TDDB characterization to the bigger area required for the 1 pF value. This extrapolation step is done using Poisson's Equation 1.48. An example is depicted for the 1.5 nm nFET stressed in inversion in Figure 6.16. In the same Figure, the similarity of the breakdown behavior of the nCAP devices is also demonstrated. The nCAP devices (nFET in an *n* well) are used in one polarity, the gate tied to V_{dd} vs. the well to ground. In this configuration when a positive stress hit the supply versus the ground, the nCAP driven accumulation mode is comparable to the nFET under inversion stress (electron injection from the substrate).

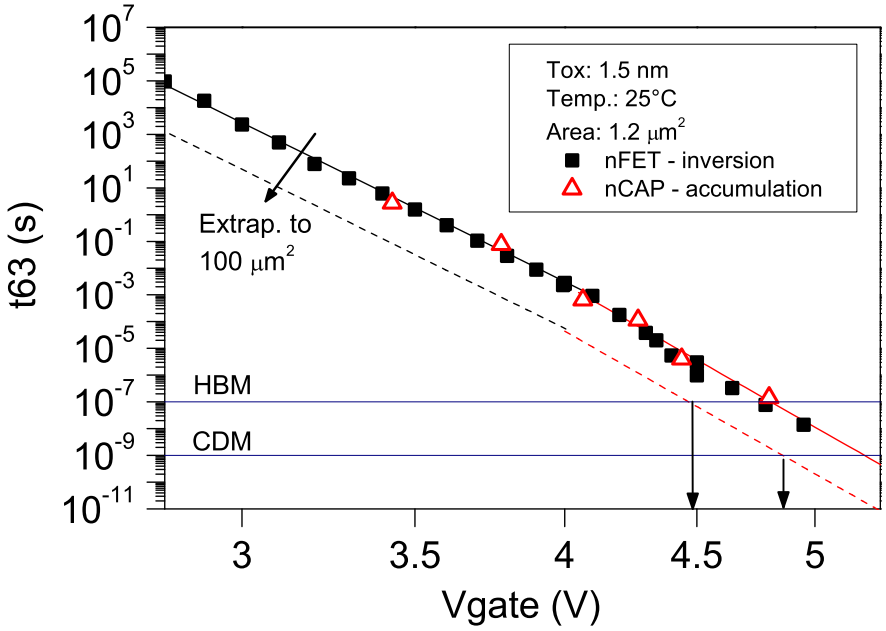

Figure 6.16: Time-to-breakdown of 1.5 nm nFET devices stressed in inversion and 1.5 nm nCAP in accumulation. The plot shows the extrapolation of the voltage acceleration from the 1.2 μm^2 tested structures to 100 μm^2 oxide area.

For large capacitors, however a beneficial effect is obtained from the high capacitance which dampens the applied stress level and leads to a reduced effective voltage across the oxide. The voltage values discussed in the following consist in the real voltage drop across the oxide layer. This damping phenomenon depends on the technology and on the circuitry, consequently, the effective stressing voltage values resulting from an ESD event across the thin oxide layers should be estimated to correlate the ESD risk with a discharge current level. Although the capacitance value increase with the area, their large oxide area is one critical parameter which counters against the ESD robustness.

Secondly, the ESD design requires a safety margin to prevent from the statistical breakdown nature of the oxide. An extrapolation from the $T_{63\%}$ value to an acceptable cumulative failure level is important for a complete comparison of the breakdown voltages in the HBM and CDM domains. As a pragmatic approach a cumulative failure criterion of 1 % will be used for the example. This criterion relates the minimum mandatory safety margin to apply. The successive breakdown voltage acceleration extrapolations to meet the area and statistical failure requirements is shown in Figure 6.17 for pFET devices stressed in inversion.

The choice of a buffer capacitor is governed by several parameters: the capacitance value, area, leakage, lifetime criteria and breakdown under ESD. These criteria are compiled for dif-

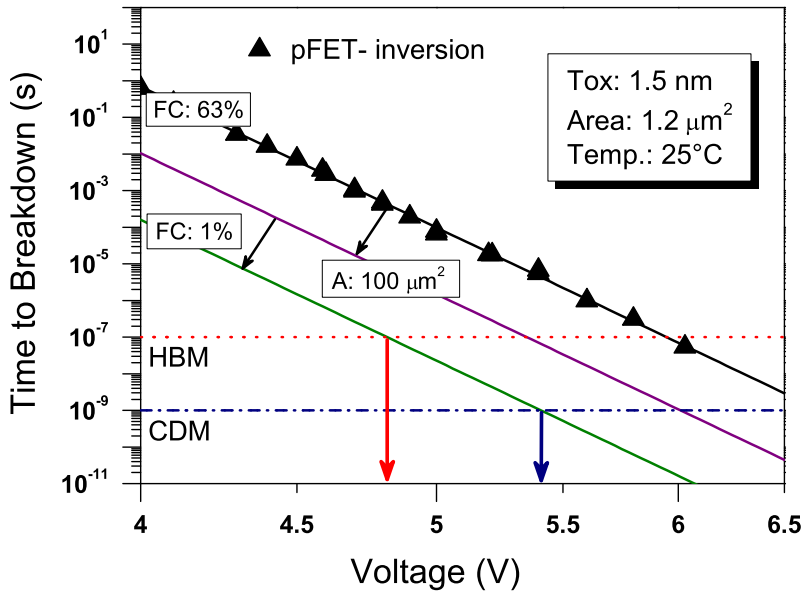

Figure 6.17: Breakdown voltage acceleration of 1.5 nm pFET devices stressed in inversion on 1.2µm^2 test structures extracted for a cumulative failure level of 63.2 %. The TDDB is scaled accordingly (1) to the area of 100 µm^2 and (2) to the cumulative failure criterion of 1 % [21].

ferent device types in a 1.5 nm oxide CMOS process in Table 6.18. Obviously, the choice of a thin buffer capacitor is not straight forward and a compromise is required among the parameters. The nMOS in inversion and nCAP in accumulation lead to best capacitance values but also to low breakdown voltages which limits considerably the ESD design window. The pMOS in inversion is the best candidate with a high capacitance value, highest breakdown voltage under ESD and a relatively low leakage. For low leakage application pMOS in accumulation should be preferred with its extremely low leakage at V_{dd} and its reasonable breakdown value under ESD. The trade-off is that its smaller specific capacitance value implies a larger area consumption. In this configuration one should be aware of the potential latch up risk since the *n* well is connected to V_{ss} instead of V_{dd}.

Based on the GOX reliability approach extended to the ESD time domain and considering the important functional parameters, the optimum choice for active oxide buffer capacitors is the pMOS devices used in inversion. This oxide capacitor is a good compromise between ESD safe operation and buffer capacitance performance requirements.

Stress configuration	Device type	Capacitance ratio	Gate leakage ratio @ V_{dd}	GOX area ratio	ESD robustness @ 25 °C F_c: 1% for 1 pF Cap.	
					HBM	CDM
inversion	nFET	1	1	1	4.2	4.7
	pFET	0.9	0.35	1.1	4.9	5.5
Accumulation	nFET	0.8	0.04	1.25	4.7	5.2
	pFET	0.8	0.03	1.2	4.4	4.9
	nCAP	1	1	1	4.2	4.7

Figure 6.18: Thin buffer capacitance characteristic parameters (1.5 nm gate oxide thickness). The capacitance values are extracted from TITAN simulation based on 10.000 µm^{2}square capacitance.

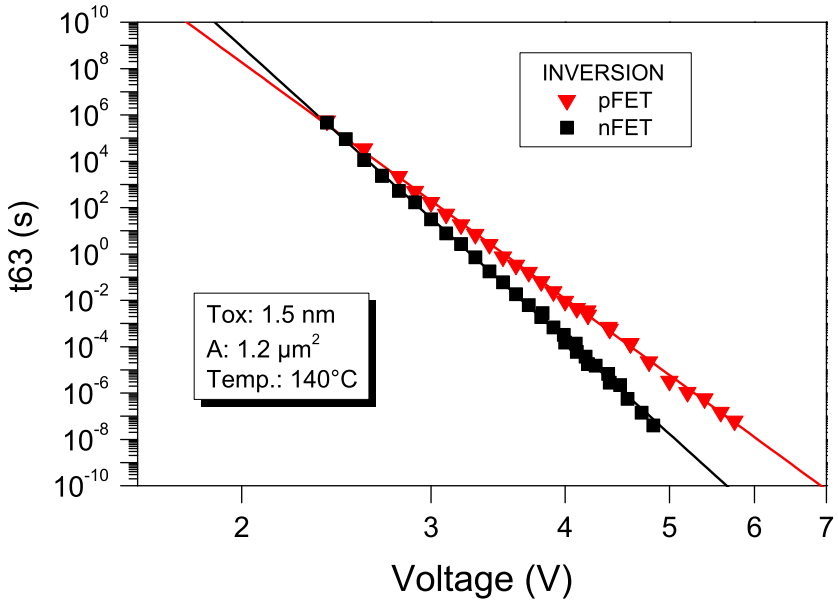

Figure 6.19: Voltage acceleration of 1.5 nm nFET and pFET stressed in inversion at 140 °C.

6.4.1.0.1 Remark: voltage acceleration cross-over Due to the different voltage acceleration between nFET and pFET, the exponent of the power law for stress in inversion is for nFETs larger than for pFETs. For ultra-thin gate oxides (<2 nm), this difference in the acceleration factor between nFET and pFET leads to a cross-over in the TDDB as a function of the stressing voltage between the two device types (Figure 6.19). Depending on the lifetime targeted and the area of the GOX area, it can imply a conflict between the long term reliability requirements and the ESD robustness. The nFET devices stressed in inversion have a better long-term lifetime

than the pFET stressed in inversion. However, for ultra-thin oxides if the residual time of the soft breakdown to generate a hard fail is used, the reliability margin can wins two orders in magnitude for pFET devices [22, 23, 16]. The use of thin nFETs devices connected in inversion directly between the supplies must be really avoid due to the too elevated ESD risk. The gap in performance between the two devices is really considerable in the ESD time domain.

6.4.2 Voltage over-shoot consideration for ESD concept engineering

With the introduction of thin oxides in the deep sub-µm CMOS technologies and the predominance of the oxide failures, fast discharges/rise time are critical threats for thin oxides and requires the use of new development tools as very-fast TLP or systems that enable the observation of the ultra-fast voltage/current transients with an very high time resolution. This is needed to investigate in detail the turn-on voltage of the ESD protection devices directly within the pulse wave form. However, the experimental measurement of very fast over-shoots is a delicate task which is strongly limited by the probes and the oscilloscope. The first over-voltage peak is dependent on the oscilloscope resolution and could significantly vary from one system to the next one. The extraction of the exact value from this high first over-shoot has up to know a quite low confidence level and still constitute one of the challenge to solve.

Basically for thin oxides processes, the common focus done on the energy dissipation within the ESD protection device extracted from the wave forms should now integrates also the evaluation of the entire clamping voltage during the ESD stress. The different voltage phases resulting from the protection device have to be taken into account to guarantee the safety of the thin oxides [5]. The effective stress endured by the oxide can be deduced by the use of the power law in summing all the different type of stresses occurring during the event (fast over-voltage shoot, unstable transient clamping, clamping steady state). Possibly in worst case, the over-shoot resulting from the slow triggering physics involved in the protection device could be sufficient alone to lead to oxide fails [2]. For core class concept development in advance CMOS technologies, this step should be integrated in the standard ESD flow. Development strategies and special stages dedicated for very fast discharges or first transient part is becoming a real hot topic.

6.4.3 ESD Process Control Monitoring (PCM)

In analogy to “standard” Process Control Monitoring, ESD PCM measurements are performed to indicate any possible process variation which can become critical for ESD. Whenever a process step or a processing tool is changed, the consequent effects could be seen in the main ESD parameters of the protection devices as well as on the parasitics structures which are useful for the ESD design window. For example, gate oxide breakdown modules are placed in the ESD PCM macros to monitor any variation that can be harmful for the products. The use of standard GOX PCM modules for ESD purpose can be discussed.

Due to evident measurement time costs, the PCM evaluation should be done really quickly in-line, making use of high sampling impossible. One first methodology, to save some times is the use of oxide thickness measurements which are already done in-line. The physical thickness variation of the lots can be correlated to the ESD robustness sensitivity via the empirical TDDB thickness dependence scaling law (see Chapter 4, Section 4.3.5). The oxide breakdown variation resulting from this thickness effect which is one of the major parameter impacting the TDDB can be directly evaluated by means of the power laws obtained during the process qualification step. However, this approach is not sufficient alone. Indeed a process step change can have a significant impact on the pre-dominant oxide degradation mechanisms which may change the TDDB predictive model [24].

The accuracy of the sensed GOX PCM values is really too low for thin oxides due to the spreading of their time-to-fail distributions over many magnitudes orders. For an efficient GOX PCM methodology, the measurements would requires, first, the full reference characterization available in the ESD regime and, secondly, high statistics to compare the $T_{63\%}$ values in the ESD time domain. This in-line testing procedure is not possible due to the highly specialized set-up for measurements in the nanosecond time domain. Moreover, the testing time in-line is really expensive as it has a direct impact on the production. As a draw back, it is then suggested to use the characterization done by the reliability engineers to monitor the oxide quality and lifetime. Fast Wafer Level Reliability (FWLR) measurements are already performed with DC voltage ramp for the reliability monitoring of extrinsic oxide issues.

Concerning the intrinsic breakdowns, the use of high statistical samples is mandatory and once by quarter the foundries are collecting some wafers to performed standard TDDB lifetime test. Despite these measurements are done in the DC regime, no better useful in-line methodologies can be achieved via GOX ESD PCM. Anyway, this methodology do not replace the intense characterization require in the ESD time domain as some high voltage specific effects could possibly not been observed. This could have been the case for example for the specific 1.5 nm oxide breakdown behavior which shows an unexpected kink emerging from the nFET devices stressed in inversion above a threshold of 4 V.

6.5 Conclusion

In this last chapter, the intensive oxide/device reliability characterization was analyzed for the definition of ESD development flow to come up to a quantitative safe ESD design window. Indeed the statistical aspect of the oxide degradation which was totally disregarded up to now has lead to review the ESD development methodologies approaches which have been used up to now. As discussed in this chapter, some of the common ESD procedures are indeed not anymore valid in advanced deep sub-µm CMOS technologies. The new ESD development kit presented through out this chapter regarding the gate oxide reliability has been practically applied to some examples of the ESD design as the thin oxides buffer capacitance.

Bibliography

[1] Wu J., Juliano P., and Rausenbaum E. Breakdown and latent damage of ultra-thin gate oxides under ESD stress conditions. In *EOS/ESD Symposium*, 2000.

[2] Smedes T. and Guitard N. Harmful voltage overshoots due to turn-on behaviour of ESD protections during fast transients. In *EOS/ESD Symposium*, pages 366–375, 2007.

[3] Linten D., Thijs S., Scholz M., Sawada M. Trémouilles D., Nakaei T., Hasebe T., and Groeseneken G. Characterization and modeling of diodes in sub-45 nm CMOS technologies under HBM stress conditions. In *EOS/ESD Symposium*, pages 158–164, 2007.

[4] Manouvrier J.P., Fonteneau P., Legrand C-A., Nouet P., and Azaïs F. Characterization of the transient behavior of gated/STI diodes and their associated BJT in the CDM time domain. In *EOS/ESD Symposium*, pages 165–174, 2007.

[5] Wybo G., Verleye S., Van Camp B., and Marichal O. Characterizing the transient device behavior of SCRs by means of VFTLP waveform analysis. In *EOS/ESD Symposium*, pages 357–365, 2007.

[6] Wu Ernest Y. and Suñé Jordi. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, 45:1809–1834, 2005.

[7] Kerber A., Röhner M., Wallace C., O'Riain L., and Kerber M. Wafer-level gate-oxide reliability towards ESD failures in advanced CMOS technologies. *Transaction on Electron Devices*, 53(4):917–920, 2006.

[8] Ille A., Stadler W., Kerber A., Pompl T., Brodbeck T., Esmark K., and Bravaix A. Ultra-thin gate oxide reliability in the ESD time domain. In *EOS/ESD Symposium Proceedings*, pages 285–294, 2006.

[9] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. In *EOS/ESD Symposium Proceedings*, pages 328–337, 2007.

[10] Wu E.Y. and Vollertsen R.P. on the weibull shape factor of intrinsic breakdown of dielectric films and its accurate experimental determination - part I: Theory, methodology, experimental techniques. *Transaction of Electron Devices*, 49(12):2131–2140, 2002.

[11] Nicollian P.E, Krishnan T., Chancellor C.A, Khamankar R.B., Chakravarthi S., Bowen C., and Reddy V. The current understanding of trap generation mechanisms that leads to the power law model for gate dielectric breakdown. In *International Electron Device Meeting*, pages 197–206, 2007.

[12] Krishnan A.T. and Nicollian P.E. Analytic extension of the cell-based oxide breakdown model to full percolation and its implications. *Annual International Reliability Physics Symposium*, pages 232–239, 2007.

[13] Guitard N., Trémouilles D., Alves S., Bafleur M., Beaudoin F., Perdu P., and Wislez A. ESD induced latent defects in CMOS ICs and reliability impact. In *EOS/ESD Symposium*, pages 174–181, 2004.

[14] Mason P.W., La Duca A.J., Holder C.H., Alam M.A., and Hwang D.K. A methodology for accurate assessment of soft-broken gate oxide leakage and the reliability of VLSI circuits. In *International Reliability Physics Symposium*, pages 430–434, 2004.

[15] Stathis J. Impact of ultra thin oxide breakdown on circuits. In *International Conference on Integrated Circuit and Technology*, pages 123–127, 2005.

[16] Kerber A., Rompl T., Duschl R., and Kerber M. Lifetime prediction for CMOS devices with ultra thin gate oxides based on progressive breakdown. In *International Reliability Physics Symposium*, pages 217–220, 2007.

[17] Brodbeck T. and Gaertner R. Experience in HBM ESD testing of high pin count devices. In *EOS/ESD Symposium*, pages 1–6, 2005.

[18] Gaertner R., Aburano R., Brodbeck T., Gossner H., Schaafhausen J., Stadler W., and Zaengl F. Partitioned HBM test - a new method to perform HBM tests on complex devices. In *EOS/ESD Symposium*, pages 1–6, 2005.

[19] Industry Council on ESD Target Levels. White Paper 1: A Case for Lowering Component Level HBM/MM ESD Specifications and Requirements, 2007.

[20] Chen H.H. and Schuster S.E. On-chip decoupling capacitor optimization for high-performance VLSI design. In *International Symposium on VLSI Technology, Systems and Applications*, pages 99–103, 1995.

[21] Ille A., Stadler W., Gossner H., Brodbeck T., Pompl T., and Bravaix A. Thin gate oxides time-to-breakdown in the ESD time domain and consequences for applications. In *International Electrostatic Discharge Workshop*, pages 176–186, 2007.

[22] Kerber A., Pompl T., Röhner M., Mosig K., and Kerber M. Impact of failure criteria on the reliability prediction of CMOS devices with ultrathin gate oxides based on voltage ramp stress. *Electron Device Letters*, 27(7):609–611, July 2006.

[23] Pompl T., Kerber A., Röhner M., and Kerber M. Gate voltage and oxide thickness dependence of progressive wear-out of ultra-thin gate oxides. *Microelectronics Reliability*, 46:18603–1607, 2006.

[24] Pompl T., Allers K.-H., Schwab R., Hofmann K., and Röhner M. Change of acceleration behavior of time-dependent dielectric breakdown by the BEOL process: Indications for hydrogen induced transition in dominant degradation mechanism. In *International Reliability Physics Symposium*, pages 388–397, 2005.

Thesis summary

In advanced sub-micron CMOS technologies, the pre-dominant failure type occurring under ESD stressed is currently changing from thermal fails in the junction to gate oxide failures. The latter emerging issue was the focus of this thesis.

The gate oxide reliability under ESD stress has been investigated with the goals to increase the knowledge about oxide degradation mechanisms and to provide a useful guidance for ESD engineers in achieving ESD safe products. This work helps to improve the ESD concept development procedures and to optimize the designs regarding oxide degradations.

The primary target has been the study of hard GOX fails which are known to define the clear boundary condition for ESD designs in thin oxide processes. First of all, the focus of the study has been on the GOX breakdown mechanism in the nanoseconds time regime. Only few works have addressed this issue in the ESD time domain and a new approach has been used in this thesis to approach this topic. With the experimental set-up described in Chapter 3, a new methodology for the gate oxide testing in the ESD time range has been developed in Chapter 4.

Beyond a reliable testing procedure, the main benefit of this methodology is the capability of modeling the GOX breakdown via statistical laws. The statistical nature of the GOX breakdown and its area dependence proven in Chapter 3 were totally disregarded in the ESD regime, although this aspect is the basic element to be considered for the ESD protection. This statistical aspect should bring a re-thinking of the development methodologies approaches which have been used up to now. This applies obviously for the GOX testing procedure as directly experimentally exposed via a large statistical data base. The experimental conditions have a clear impact on the oxide breakdown results and the use of improper methodologies and test structures can lead to wrong results and thus to incoherent and weak ESD designs. Beyond the testing procedure, the ESD development phase including for example the definition of the ESD design window, the choice of the concepts, the silicon verification procedure, and the qualitative derived safety margin must now evolute to integrate this important statistical nature. Indeed, the definition of a new parameter, the statistical oxide failure level, is required in the ESD development procedures and should account for an indisputable reliability criterion for the products.

From the appropriate characterization method (from the test structures, testing procedure to the data extraction), the thin oxides time-to-breakdown (TDDB) has been characterized in Chapter 4 from the DC time regime down to the ESD time regime. A common power law

TDDB voltage acceleration from the long reliability term down to the ESD regime is reported for thinner oxides than 7 nm. This TDDB voltage dependence is reported in the ESD regime for the first time. Up to now, the $1/E$ predominant model was used for ESD GOX breakdown evaluation at high fields. This last model is too optimistic and extrapolation of thin oxides breakdown based on this $1/E$ law leads to dangerous unsafe ESD protections.

The strong power law voltage dependence of the thin oxide breakdown increases the challenge to the ESD community. The extension of this power law model in the nanosecond time range proofs also the continuity of the breakdown mechanisms which occur under DC conditions. This new fact enables then the direct extrapolation of the process qualification data to characterize the oxide breakdown under "short" EOS and ESD events (HBM, CDM). The experimental characterization was extended to quantify the influence of the parameters impacting the TDDB, i.e. the oxide thickness, the stress polarity, the device type, the temperature and the area. The consideration of these physically based models in the TDDB extrapolation provides a powerful tool for the ESD engineers that can precisely determine the strict GOX breakdown limit for their specific designs, regarding the different ESD type of stress.

From the DC extrapolation validity and by means of empirical models and statistical oxide degradation theories, the oxide breakdown value can be derived in the ESD time domain from the long term reliability qualification data without additional efforts and costs required due to silicon area, set-up and testing effort. Beyond these advantages, the increased oxide breakdown estimation accuracy permits to achieve reliable products within a controlled quality. This new practical methodology developed is now adopted as standard within Infineon Technologies. Moreover, this development package increases the confidence level of pre-silicon ESD concept and guidelines for the protection of sensitive thin oxides as well as for the definition of forbidden critical circuits' architectures in the early steps of new technologies developments.

However, the hard breakdown is not the only oxide failure mode. The study of oxide degradation of nFET devices under ESD stress was extended in Chapter 5 to evaluate the possibility of failures or latent damages induced by moderate ESD stress. Due to the decreasing ESD design window for thinner oxides, the gate oxides are more and more susceptible to over-shoots during ESD surges. This type of non destructive stress has thus also been investigated in details. The oxide degradation is reported to be thickness dependent. For thinner oxides than 3 nm, no significant effect has been noticed. For medium to thick oxides, degradation linked with traps in the bulk of oxides are responsible for parameters drift and can lead to device lifetime reduction if a too strong amount of traps was generated by the stress. To avoid any reliability problem in the field, a quantitative margin has to be included in the ESD design window for analogue (I/O) pads. This problem is not really recent, but no practical margin or failure criterion was proposed as a function of the oxide degradation for thick oxides to avoid this reliability issue.

Finally, combining oxide dielectric and device degradation, the characterization and modeling learning from the previous chapters was analyzed in Chapter 6 to ESD purposes. Discussions on important topics as the definition of the ESD design window, ESD testing procedures, ESD

guidelines and GOX breakdown extrapolation alternative were summed-up. This experimental study is contributing to the improved reliability and robustness of ICs regarding ESD stress and permits also the optimization of protection concepts as well as the preparation of the ESD development for further CMOS technologies. The ESD development flow improved by the new package introduced in this thesis enable to save some development time, effort and costs and to achieve ESD robust products in a quality approach. Moreover, area optimized I/O cells resulting from more aggressive ESD protections enables by a better control of the design limitations permit to be competitive in the semiconductor markets.

Outlook

The results presented in this work were somewhat limited by the experimental set-up down to around 20 ns pulse stress. The issue covering the very fast ESD stresses in the nanosecond regime (CDM or system level type of stresses) as well as the ultra-fast ESD transient overshoots should still be investigated to confirm the validity of the universal breakdown behavior of thin oxides reported in this work. We assume the extension of the breakdown laws to be valid for ultra-fast discharges. Despite the validity of the power laws within 14 orders of magnitude in times and over a wide voltage range accessed with the 7 different oxide thicknesses which suggest no further universal deviation at high field, further effects specifically related to thin oxides cannot be excluded *a priori*. Wide band oscilloscopes available nowadays (20 GHz) make the investigation of the first peak down to the several picoseconds also possible. Beyond the characterization of the gate oxide breakdown under high/fast stresses, a concurrent characterization and development of the dedicated ESD protection concepts regarding their speed and capability to clamp the transient voltage over-shoots is highly mandatory for sub-µm CMOS technologies where thin oxide fails under ESD stresses becomes the pre-dominant failure type.

The work presented in this thesis concerns exclusively SiO_2 and SiO_xN oxides layers. However, for the future of the CMOS technologies, there is a strong demand to introduce new dielectric insulator materials. The strong decrease of the oxide thickness is nearly reaching the physical limit of SiO_2 layers with a drastic gate leakage increase. To counter this, alternative solution of high permittivity constant dielectrics (high-k [1]) which are providing an equivalent Electrical Oxide Thickness (EOT) with a reduced leakage have been planned beyond the 45 nm technologies nodes for 2008/2009. More likely, the implemented solutions will be based on stacked dielectric materials. From the recent works done on this topic [1], the favorite new gate insulator candidate is certainly the hafnium oxinitride (HfSiON). Other emergent technologies evolutions as SOI (Silicon on Insulator [2]) or MugFET (Multi-gate FET [3, 4]) processes are coming along with these new dielectric layers.

In order to ensure a correct visibility on the ESD sensitivity and trend in the future, the dielectric reliability of these alternative layers as well as new ESD concepts have to be investigated

in the different promising technologies flavors. New ESD concepts in these radically changed processed are very likely, and a dielectric reliability package as developed in this thesis would be more than an helpful support in this task.

Bibliography

[1] Houssa M., Pantisano L., Ragnarsson L.A., Degreave R., Schram T., Pourtois G., De Gendt S., Groeseneken G., and heyns M.M. Electrical properties of high-k gate dielectrics: Challenges, current issues, and possible solutions. *Materials Science and Engineering*, 51:37–85, 2006.

[2] Gossner H. ESD protection for the deep sub micron regime - a challenge for design methodology. In *Conference on VLSI Design*, pages 809–818, 2004.

[3] Russ C, Gossner H. Schulz T., Chaudhary N., Xiong W., Marshall A., Duvvury C., Schruefer C., and Cleavelin C.R. ESD Evaluation of the Emerging MuGFET Technology. In *EOS/ESD Symposium Proceedings*, 2005.

[4] Gossner H., Russ C., Siegelin F., Schneider J., Schruefer K., Schulz T., Duvvury C., Cleavelin C.R., and Xiong W. Unique ESD Failure Mechanism in a MuGFET Technology. In *International Electron Device Meeting*, 2006.

Part II

Appendix

Personal bibliography

- International EOS/ESD symposium [1, 2],
  ([1] Best paper in 2006, [2] Best Student paper in 2007)
- Reliability Center for Electronic Components of Japan Symposium [3], *(Invited paper)*
- International ESD Workshop [4]
- Infineon Technologies conference - poster [5]
- IEEE Microelectronics Reliability Journal [6], *(ESD special issue)*

Bibliography

[1] Ille A., Stadler W., Kerber A., Pompl T., Brodbeck T., Esmark K., and Bravaix A. Ultra-thin gate oxide reliability in the ESD time domain. In *EOS/ESD Symposium Proceedings*, pages 285–294, 2006.

[2] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. In *EOS/ESD Symposium Proceedings*, pages 328–337, 2007.

[3] Ille A., Stadler W., Kerber A., Pompl T., Brodbeck T., Esmark K., and Bravaix A. Ultra-thin gate oxide reliability in the ESD time domain. In *17th annual Reliability Center for Electronic Components of Japan Symposium*, number ISSN 1349-6379, pages 75–84, 2007.

[4] Ille A., Stadler W., Gossner H., Brodbeck T., Pompl T., and Bravaix A. Thin gate oxides time-to-breakdown in the ESD time domain and consequences for applications. In *International Electrostatic Discharge Workshop*, pages 176–186, 2007.

[5] Pompl T., Ille A., Wendel M., Gossner H., Stadler W., Brodbeck T., Esmark K., and Alvarez D. Gate oxide breakdown and ESD - interdisciplinary know-how exchange towards reliable and competitive products, 2007.

[6] Ille A., Stadler W., Pompl T., Gossner H., Brodbeck T., Esmark K., Riess P., Alvarez D., Chatty K., Gauthier R., and Bravaix A. Reliability aspects of gate oxide under ESD pulse stress. *Microelectronics Reliability*, 2008, In press.

FIABILITE DES OXYDES DE GRILLE ULTRA-MINCES SOUS DECHARGES ELECTROSTATIQUES (ESD) DANS LES TECHNOLOGIES CMOS FORTEMENT SUB-MICRONIQUES

Résumé:

Les décharges électrostatiques (ESD) constituent un problème majeur de fiabilité pour les entreprises de semi-conducteurs. Pour enrayer les défauts générés par les ESD sur les circuits intégrés (ICs), des éléments de protection sont implantés directement dans les puces. La constante poussée de l'intégration des circuits a pour conséquence la réduction des dimensions des cellules technologiques élémentaires ainsi que l'accroissement du nombre d'applications supportées par les ICs. Les conditions restrictives imposées par les procédés technologiques et par la complexité croissante des systèmes entraînent un défi considérablement accru pour le développement de produits robustes aux ESD. Dans ce travail de recherche, le problème émergeant des défaillances des couches d'oxydes minces d'épaisseur Tox = 8 à 1.1nm sous contraintes ESD est adressé dans les technologies CMOS les plus avancées, par une contribution à la compréhension des mécanismes de dégradation de la fiabilité du diélectrique et des dispositifs sous contraintes ESD. Une nouvelle approche de caractérisation des oxydes minces sous des stress à pulses ultra-courts (20 ns) est décrite jusqu'à la modélisation complète de la dépendance temporelle du claquage du diélectrique. Basé sur un ensemble cohérent de modélisations, une nouvelle méthodologie est proposée pour ajuster la détermination de la fenêtre ESD de façon mieux adaptée aux intervalles de tension et d'épaisseur d'oxyde de grille pour l'ingénierie des concepts de protection. Ceci a permis d'améliorer la prise en compte des problèmes ESD pour une meilleure fiabilité et robustesse des produits conçus en technologies CMOS fortement sub-microniques vis-à-vis des décharges électrostatiques.

Mots Clés:

Décharges électrostatiques (ESD), fenêtre de design ESD, CMOS, caractérisation, fiabilité des oxydes de grille minces, modélisation de la dépendance du claquage des diélectriques en fonction du temps (TDDB), injection de porteurs chauds (HCI).

Summary:

Electrostatic Discharges (ESD) is a major reliability concern for semiconductor companies. To prevent the ICs from failures caused by ESD events, on-chip ESD protections concepts are implemented. With the down-scaling of the CMOS technologies, the boundary conditions defined by application and process is getting extremely challenging for the conception of robust protection elements. In this work the emerging issues of the thin oxide failures due to ESD is addressed. The work contributes also to the understanding of thin oxide dielectrics and device reliability degradation mechanisms under ESD events. A new characterization approach for thin gate oxides under short pulse stresses (down to 20 ns) is introduced; it allows complete modeling of the time-to-breakdown. An universal time-to-fail voltage power law acceleration is reported. This is an extremely important result for the ESD designs with regard to all kinds of ESD stress events. From the modeling package established in this work, a novel ESD development kit is described, aiming to improve the ESD robustness and reliability of products based on advanced sub-micron CMOS technologies.

Key words :

Electrostatic Discharges (ESD), ESD design window, Complementary Metal-Oxide Semiconductor (CMOS), characterization, thin gate oxides reliability, Time Dependent Dielectric Breakdown (TDDB) modeling, Hot Carrier Injection (HCI).

Im2np Institut Matériaux Microélectronique Nanosciences de Provence
UMR 6242 CNRS, Universités Paul Cézanne, Provence et Sud Toulon-Var