

Captura y liberación de carga en dispositivos MOS con dieléctricos de alto-K

L. Sambuco Salomone¹, S. H. Carbonetto^{1,3}, M. A. García Inza^{1,3}, J. Lipovetzky^{1,3}, E. G. Redín^{1,3}, F. Campabadal⁴, A. Faigón^{1,2,3}

¹Laboratorio de Física de Dispositivos – Microelectrónica
Universidad de Buenos Aires, Facultad de Ingeniería

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

³Instituto de Ciencias de la Ingeniería (INTECIN)

⁴Instituto de Microelectrónica de Barcelona – Centro Nacional de Microelectrónica – Consejo Superior de Investigaciones Científicas (CSIC)
email: lsambuco@fi.uba.ar

Abstract—Los procesos de captura y liberación de carga en dispositivos MOS con Al_2O_3 crecido por Atomic Layer Deposition (ALD) como aislante de puerta son estudiados y caracterizados eléctricamente en relación a su dependencia con la tensión aplicada y el tiempo, mediante diferentes experimentos. Un modelo físico basado en un frente de túnel es propuesto con el fin de explicar los resultados obtenidos y poder extraer los parámetros físicos asociados, como la densidad y distribución energética de las trampas responsables de dichos fenómenos de captura.

MOS devices; High-K gate dielectrics; Electron traps.

I. INTRODUCCIÓN

El aumento en la complejidad de los circuitos electrónicos ha llevado a un aumento en el nivel de integración de los dispositivos, que se traduce en una reducción de las dimensiones de los mismos. En el caso particular de la estructura MOS (Metal-Oxide-Semiconductor), esta reducción en sus dimensiones ha llevado a los procesos de fabricación a un límite crítico. A medida que el espesor de la capa aislante se reduce por debajo de un valor de aproximadamente 2 nm, los niveles de rendimiento y confiabilidad se ven afectados [1].

Uno de los temas más investigados durante la primera década de este siglo en el área de la fabricación de dispositivos semiconductores, ha sido la posible utilización de nuevos materiales en reemplazo del SiO_2 como capa aislante en las mencionadas estructuras MOS. Dichos materiales, se caracterizan por poseer una constante dieléctrica significativamente mayor a 3.9 del SiO_2 . Este aumento en la constante dieléctrica, permite construir dispositivos MOS que mantengan los niveles de capacidad, y en consecuencia, de acoplamiento puerta-canal, sin la necesidad de reducir el espesor de la capa aislante más allá de lo considerado como límite crítico por la magnitud de las corrientes de túnel.

En particular, la Alúmina (Al_2O_3) se encuentra entre los principales candidatos para formar parte de las estructuras MOS como capa aislante. Entre las aplicaciones tratadas recientemente en la literatura, pueden mencionarse la posibilidad de incorporar Al_2O_3 como interpoly dielectric (IPD) en memorias flash de puerta flotante [2-8], como parte de la

capa aislante en Ion-Sensitive Field Effect Transistors (ISFET) [9], o su utilización en dosímetros MOS [10].

A pesar de los beneficios asociados con el aumento en la constante dieléctrica, numerosos problemas han surgido al tratar de utilizar dichos materiales como capa aislante, principalmente vinculados con la presencia de un gran número de defectos en su estructura, los cuales pueden comportarse como centros de captura de carga, causando algunas dificultades, incluso en condiciones normales de operación [11-12]:

- Inestabilidades en la tensión de umbral, causadas por la dependencia temporal de la carga atrapada.
- Reducción de la movilidad en el canal debido a scattering Coulombiano con la carga atrapada.
- Aumento en la corriente a través del aislante, debido a que los defectos mencionados favorecen la conducción.

Desde el punto de vista de la caracterización eléctrica, la captura de carga en los defectos mencionados genera una histéresis en las típicas características C-V (Capacitance-Voltage) o I-V (Current-Voltage) del dispositivo. Alternativamente, también puede observarse una dependencia temporal en el valor de capacidad medida a una tensión fija [13].

Se encuentra relativamente aceptado en la literatura que la captura de carga mencionada se debe a transiciones de túnel entre el semiconductor y las trampas presentes en el dieléctrico [8, 14-15].

En este trabajo se presenta una caracterización sistemática de dichos procesos de captura, desde el punto de vista de su dependencia con la tensión aplicada y su evolución temporal, junto con un modelo físico, propuesto con el fin de explicar los resultados obtenidos y poder caracterizar físicamente los defectos presentes en el material.

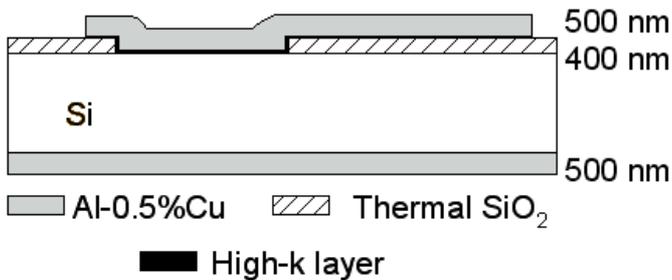


Figura 1. Esquema de las estructuras fabricadas.

II. DETALLES EXPERIMENTALES

A. Descripción de las muestras

Los capacitores MOS utilizados en el presente trabajo fueron fabricados utilizando la técnica de Atomic Layer Deposition (ALD) en el Instituto de Microelectrónica de Barcelona (IMB) – Centro Nacional de Microelectrónica (CNM).

La capa de Al_2O_3 fue depositada sobre un sustrato de silicio tipo n con una resistividad de 1-12 Ωcm , que se corresponde con un nivel de dopaje de 4×10^{14} - 5×10^{15} at/cm^3 .

Inicialmente, se depositó una capa de 400 nm de SiO_2 por oxidación térmica, luego de lo cual, se abrieron por litografía y grabado húmedo las ventanas para la deposición de las capas dieléctricas. Luego, las muestras fueron limpiadas con $\text{H}_2\text{O}_2/\text{H}_2\text{SO}_4$ durante 10 minutos y con HF durante 10 segundos. El proceso de ALD fue llevado a cabo en 95 ciclos a una temperatura constante de 200 $^\circ\text{C}$, resultando en un espesor físico de 11.7 nm. Finalmente, ambas caras fueron metalizadas con 0.5 μm de Al – 0.5% Cu. La figura 1 muestra un esquema de la estructura obtenida.

Las muestras fueron caracterizadas utilizando TEM (Transmission Electron Microscopy). El EOT (Equivalent Oxide Thickness) medido fue de 5.8-6.3 nm, resultando en un valor de constante dieléctrica (K) entre 7.2 y 7.9.

B. Experimentos

A fin de caracterizar las inestabilidades eléctricas, se realizaron mediciones de ciclos de capacidad-tensión (C-V) del dispositivo, utilizando el medidor vectorial de impedancias HP 4277A, comandado por una computadora a través de un software diseñado para tal tarea. La frecuencia de medición fue de 1 MHz.

Cada ciclo de medición consta de dos barridos consecutivos, el primero iniciando en inversión y finalizando en acumulación, y el segundo en el sentido opuesto. La presencia de trampas dentro del material, capaces de capturar carga genera un desplazamiento en tensión entre los dos barridos de un mismo ciclo de medición. La figura 2 muestra el resultado obtenido al medir un ciclo de curvas C-V entre una tensión mínima de -0.5 V y una tensión máxima de 1 V. Como puede apreciarse en el recuadro de la figura, el valor de la histéresis, V_H , alcanza un valor máximo de aproximadamente 150 mV, en el rango 25-55 pF.

Para obtener una caracterización más completa de los procesos de captura y liberación de carga en el aislante, se diseñó un experimento donde se registran evoluciones temporales. El experimento consiste en tres etapas, en la primera, se lleva a cabo la medición de una curva C-V comenzando en algún valor de tensión determinado y finalizando cuando la capacidad alcanza un valor dado, que denominaremos V_C , en la segunda etapa, que comienza inmediatamente después de la primera, se sigue la evolución de la tensión V_G que debe aplicarse sobre el dispositivo para mantener constante el valor de capacidad alcanzado al final de la primera etapa de la medición; finalmente, luego de un tiempo de seguimiento dado, se continúa el barrido iniciado en la primer etapa hasta alcanzar un valor de tensión final.

La figura 3 muestra los resultados del registro temporal de ΔV_C en 4 tipos de condiciones diferentes, cada una de las cuales está caracterizada por su tensión inicial del barrido, -0.5 V (Figura 3a) y 1 V (Figura 3b), y por su historia previa, diferenciando el caso en que el dispositivo estuvo algunas horas en reposo, de aquel en el cual el experimento es llevado a cabo inmediatamente después de terminar otro similar. Respecto a la dependencia de la dinámica con la tensión inicial del barrido, puede observarse que, para el caso en el que la tensión inicial de la medición es -0.5 V (inversión), la dinámica inicial genera un desplazamiento de V_C positivo, mientras que en el caso de una tensión inicial de 1 V (acumulación), dicho desplazamiento de V_C es negativo. Este tipo de comportamiento es consistente con la captura y liberación de carga en el aislante durante la medición. Respecto a la dependencia con la historia previa del dispositivo, puede notarse para ambas tensiones iniciales, que la dinámica varía para tiempos mayores a aproximadamente 500 segundos. En ambos casos, la realización del experimento con el dispositivo en reposo agrega un desplazamiento positivo de V_C para dicha región temporal, respecto del caso en que el experimento se realice apenas terminado otro similar.

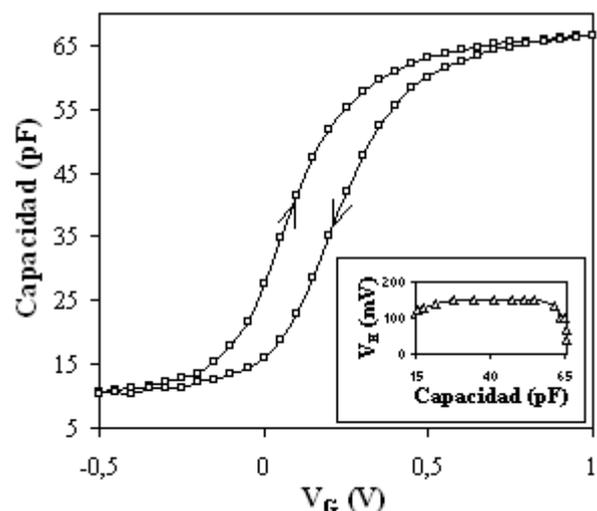


Figura 2. Curvas C-V experimentales sobre capacitores MOS con Al_2O_3 como aislante de puerta. El recuadro de la figura muestra el valor de la histéresis en función de la capacidad.

Por todo lo antes dicho, puede concluirse que el comportamiento de ΔV_C con el tiempo está formado por la combinación de dos dinámicas independientes: Una dinámica rápida, que depende exclusivamente de la tensión inicial del barrido, domina el comportamiento general para tiempos cortos ($t < 500$ s); otra dinámica más lenta que la anterior, depende exclusivamente de la historia previa del dispositivo y domina el comportamiento general para tiempos largos ($t > 500$ s).

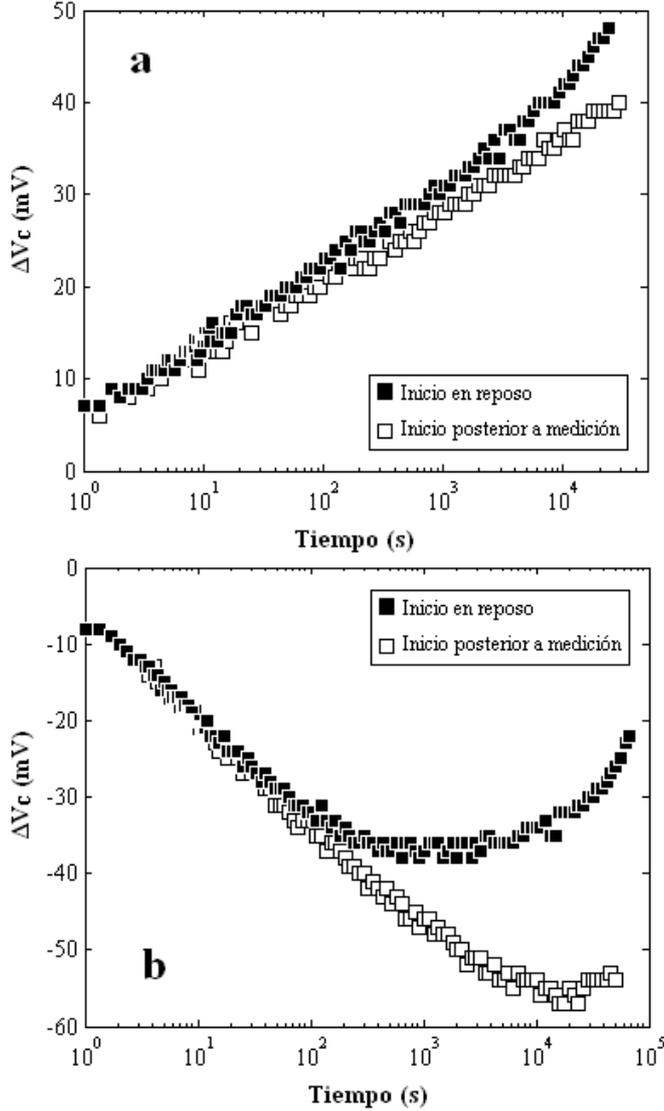


Figura 3. Seguimientos de la tensión para una capacidad constante, V_C , a partir de barridos iniciando en -0.5 V (a), o 1 V (b).

III. MODELO FÍSICO PARA LA CAPTURA DE CARGA

A fin de explicar cuantitativamente los resultados obtenidos, se propone un modelo físico conocido como frente de túnel [16-18], basado en la hipótesis de que la carga y descarga de trampas de electrones se deben a transiciones de túnel desde y hacia el sustrato.

El desplazamiento en V_C medido, ΔV_C , está relacionado con la densidad de carga capturada a través de la siguiente ecuación.

$$\Delta V_C(t) = \frac{q}{C_{OX}} \int_0^{t_{OX}} n_t(x,t) \left(1 - \frac{x}{t_{OX}}\right) dx \quad (1)$$

donde n_t es la densidad de electrones capturados, C_{OX} es la capacidad por unidad de área y t_{OX} es el espesor de la capa dieléctrica.

De acuerdo al modelo del frente de túnel, n_t tiene un valor constante, N_t , desde la interfaz hasta una distancia $x_m(t)$, y cero para valores mayores de x . La dependencia de $x_m(t)$ es dada por

$$x_m(t) = \frac{1}{2K} \ln\left(\frac{t}{\tau_0}\right) \quad (2)$$

donde τ_0 y K son constantes asociadas al proceso de túnel. Reemplazando, de acuerdo a lo descripto, n_t en (1), obtenemos

$$\Delta V_C(t) = \frac{qN_t}{C_{OX}} \left[\frac{1}{2K} \ln\left(\frac{t}{\tau_0}\right) - \frac{1}{8K^2 t_{OX}} \ln^2\left(\frac{t}{\tau_0}\right) \right] \quad (3)$$

lo cual es válido desde $t=\tau_0$ hasta $t=t_f$, el tiempo que tarda el frente de túnel en alcanzar una distancia x_f , correspondiente con las trampas más lejanas.

En el presente caso, para poder dar cuenta de las dos dinámicas observadas, hará falta considerar dos tipos de trampas diferentes, unas responsables del comportamiento inicial de $\Delta V_C(t)$, que denominaremos “rápidas”, y otras responsables del comportamiento para tiempos largos ($t > 500$ s), que denominaremos “lentas”. La dinámica de carga y descarga en ambos tipos de trampas responderá a un modelo como el presentado, resultando la dinámica general, la combinación de ambas dinámicas independientes. En relación con lo anterior, la diferencia entre las dinámicas que sólo difieren en la historia previa del dispositivo puede ser explicada, asumiendo que en aquellas mediciones llevadas a cabo inmediatamente luego de otra medición similar, aquellas trampas con una dinámica lenta aún permanecen cargadas, por lo cual no contribuirán al comportamiento de $\Delta V_C(t)$, siendo en tales casos posible explicar el comportamiento general considerando sólo la dinámica de las trampas rápidas.

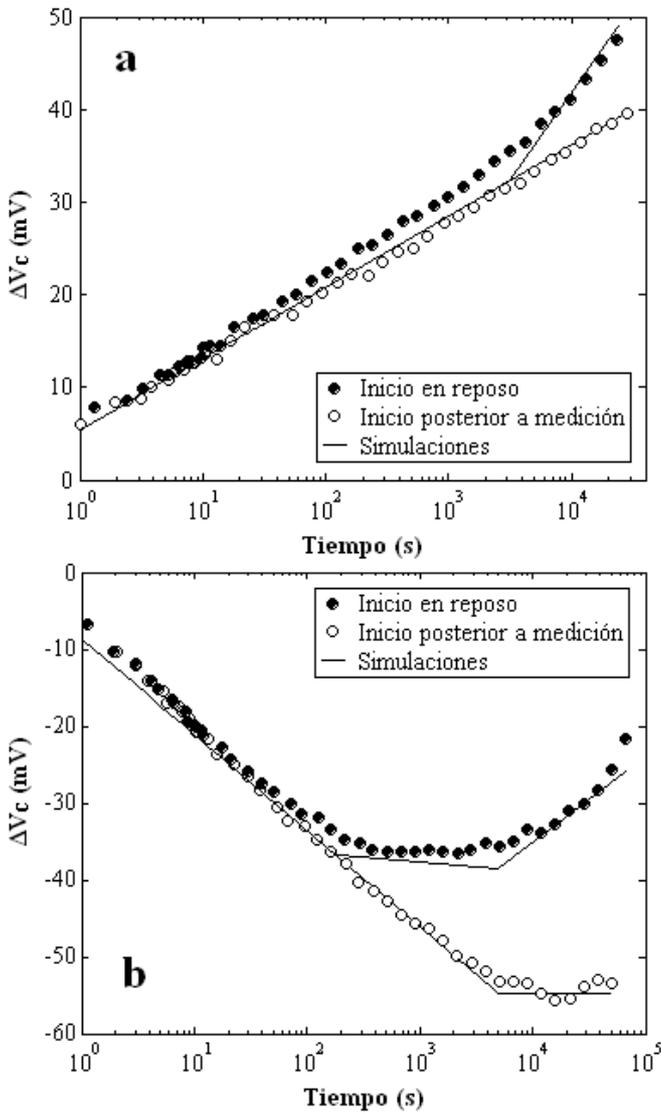


Figura 4. Comparación entre los experimentos y las simulaciones realizadas de acuerdo al modelo propuesto para el caso de barridos iniciando en -0.5 V (a), o 1 V (b).

Las curvas resultantes del modelo propuesto para el comportamiento general de $\Delta V_c(t)$ son mostradas en la figura 4, junto con las curvas experimentales. Asimismo, en la figura 5 se muestran, para ambas tensiones iniciales de barrido (-0.5 V y 1 V), la diferencia entre la curva que se inicia en reposo y aquella que se inicia luego de realizada una medición similar. Teniendo en cuenta la consideración antes mencionada de que dicha diferencia sería el resultado de la contribución de trampas con una dinámica lenta, en la misma figura 5, se muestra el resultado del modelo cuando sólo se considera la contribución de dichas trampas. Como puede observarse en ambas figuras, el modelo propuesto reproduce satisfactoriamente las mediciones realizadas.

Usando los datos $C_{OX} = 5.71 \times 10^{-7}$ F/cm² y $t_{OX} = 11.7$ nm, pueden obtenerse los valores de los parámetros τ_0 y N_t/K asociados al modelo, los cuales se resumen en la Tabla I para cada caso analizado.

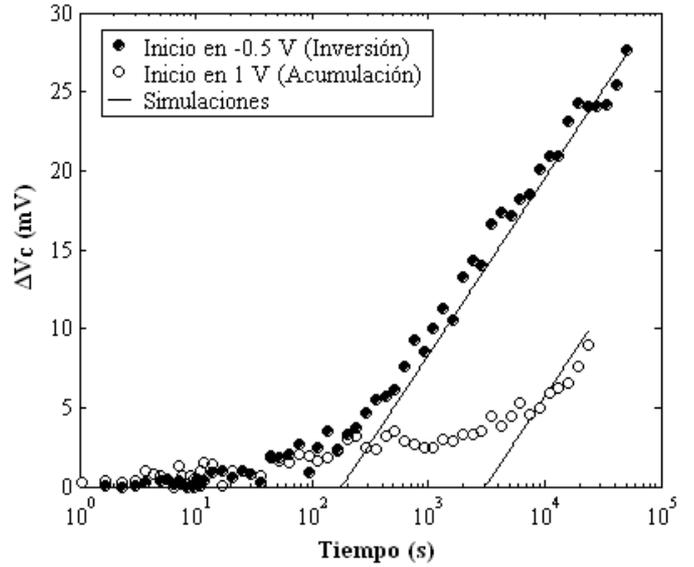


Figura 5. Comparación entre los experimentos y las simulaciones realizadas de acuerdo al modelo propuesto para el caso de las curvas que resultan de restar los comportamientos de $\Delta V_c(t)$ durante mediciones iniciadas con el dispositivo en reposo y luego de finalizada otra similar, para barridos iniciados en -0.5 V y 1 V, respectivamente.

TABLA I. VALORES OBTENIDOS PARA LOS PARÁMETROS DE AJUSTE.

Dinámica	Inicio	N_t/K (cm ⁻²)	τ_0 (s)	τ_f (s)
Rápida	-0.5 V	2.38×10^8	2×10^{-1}	-
	1 V	3.86×10^8		5×10^3
Lenta	-0.5 V	3.46×10^8	3.1×10^3	-
	1 V		1.8×10^2	-

IV. DISCUSIÓN

Un importante punto de discusión acerca del modelo propuesto es la validez de los parámetros físicos obtenidos. A partir de (3), puede observarse que el comportamiento de la dinámica asociada a cada tipo de trampas, tendrá una dependencia lineal con $\log(t)$ siempre que el segundo término en dicha ecuación sea notablemente menor que el primero, lo cual es una suposición razonable en la mayoría de los casos. Lo anterior explica satisfactoriamente la dependencia lineal con $\log(t)$ de las curvas experimentales $\Delta V_c(t)$, para tiempos muy cortos y muy largos, en los cuales, cada una de las dinámicas domina el comportamiento general de la curva.

Teniendo en cuenta lo anterior, tendremos para cada dinámica, 4 parámetros físicos, τ_0 , τ_f , N_t y K , cuyos valores resultan del proceso de ajuste entre las curvas simuladas y el comportamiento general de $\Delta V_c(t)$ para ambas regiones de tiempo. Sin embargo, a partir de las curvas experimentales, sólo 3 parámetros pueden ser extraídos para cada región temporal, el tiempo inicial, el tiempo final y la pendiente de la curva. Los tiempos inicial y final están representados por los

parámetros físicos τ_0 y τ_f , mientras que la pendiente está representada por la relación N_t/K , la cual es proporcional a la relación $N_t/\sqrt{E_t}$, debido a la siguiente expresión para K.

$$K = \left(\frac{2m_{eff}E_t}{\hbar^2} \right)^{1/2} \quad (4)$$

donde E_t es el nivel energético de la trampa medido respecto de la banda de conducción del dieléctrico, m_{eff} es la masa efectiva del electrón en Al_2O_3 , la cual toma un valor de 0.35 veces la masa del electrón libre [7], y \hbar es la constante reducida de Planck.

Como criterio, se decidió fijar el valor correspondiente al nivel energético E_t de ambos tipos de trampas (rápidas y lentas) en 2.6 eV y 3.4 eV, respectivamente, de acuerdo a lo reportado en la literatura [6], obteniendo así $N_t = 1.94 \times 10^{18} \text{ cm}^{-3}$ para el caso de las trampas lentas.

La constante N_t representa la densidad de trampas involucradas en el proceso de carga o de descarga. Si bien depende del posicionamiento energético de las trampas respecto del nivel de Fermi; los valores de N_t para trampas rápidas encontrados para cada proceso resultan proporcionales a los corrimientos de V_C durante los mismos. Esto permite estimar su densidad total en los procesos de carga y descarga que dan lugar a la histéresis. A partir del valor de la histéresis, V_H , en un ciclo C-V entre una tensión mínima de -0.5 V y una tensión máxima de 1 V, se obtiene $N_t \approx 5 \times 10^{19} \text{ cm}^{-3}$ para las trampas rápidas.

Con respecto al tiempo final, a partir de la figura 4b, puede obtenerse un valor $\tau_f = 5000 \text{ s}$ para las trampas rápidas, en el caso de comportamientos $\Delta V_C(t)$ precedidos por barridos iniciados en una tensión de 1 V. Para el rango de tiempos considerados en los experimentos, no resulta posible obtener el valor del parámetro t_f para el resto de los procesos.

En lo que se refiere a la validez de la hipótesis de que las dinámicas observadas se deben a la presencia de trampas con diferentes niveles energéticos, es importante destacar que la presencia de más de un tipo de trampas en Al_2O_3 , diferenciadas por su nivel energético, ya ha sido observada previamente por otros grupos mediante diferentes métodos experimentales [5-8, 19]. En este sentido, la técnica experimental utilizada en el presente trabajo, basada en el seguimiento de la tensión a una capacidad constante, resulta complementaria de las otras técnicas utilizadas, y aporta a la discriminación de diferentes tipos de defectos dentro del material.

V. CONCLUSIONES

Se presentó una caracterización sistemática del fenómeno de captura y liberación de carga en estructuras MOS con Al_2O_3 como aislante de puerta, a partir del seguimiento de la tensión a una capacidad constante, V_C , en función del tiempo.

Se presentó un modelo físico, basado en la teoría del frente de túnel que explica satisfactoriamente los resultados experimentales. De la aplicación del modelo resulta que dos tipos de trampas son responsables de la captura de carga en el dieléctrico: una con una energía de 2.6 eV por debajo de la banda de conducción del dieléctrico, que domina el

comportamiento de V_C para tiempos cortos ($t < 500 \text{ s}$), y otra más profunda ($E_t = 3.4 \text{ eV}$) que domina el comportamiento de V_C para tiempos largos ($t > 500 \text{ s}$).

REFERENCIAS

- [1] Gusev et al. "Ultrathin high-k metal oxides on silicon: processing, characterization and integration issues", *Microelectronic Engineering* vol. 51 pp 341-349, 2001.
- [2] B. Govoreanu, R. Degraeve, J. Van Houdt, and M. Jurczak, "Statistical Investigation of the Floating Gate Memory Cell Leakage through High-k Interpolo Dielectrics and Its Impact on Scalability and Reliability," *IEEE International Electron Devices Meeting*, 2008.
- [3] X.F. Zheng, W.D. Zhang, B. Govoreanu, J.F. Zhang, and J. Van Houdt, "A discharge-based multi-pulse technique (DMP) for probing electron trap energy distribution in high-k materials for Flash memory application," *IEEE International Electron Devices Meeting*, 2009.
- [4] X.F. Zheng, W.D. Zhang, B. Govoreanu, J.F. Zhang, and J. Van Houdt, "Impact of PDA temperature on electron trap energy and spatial distributions in SiO_2/Al_2O_3 stack as the IPD in Flash memory cells," *Microelectronic Engineering*, 2009.
- [5] X.F. Zheng, W.D. Zhang, B. Govoreanu, D. Ruiz Aguado, J.F. Zhang, and J. Van Houdt, "Energy and Spatial Distributions of Electron Traps Throughout SiO_2/Al_2O_3 Stacks as the IPD in Flash Memory Application," *IEEE Transactions on Electron Devices*, vol. 57, no. 1, 2010.
- [6] X.F. Zheng, W.D. Zhang, B. Govoreanu, J.F. Zhang, and J. Van Houdt, "A New Multipulse Technique for Probing Electron Trap Energy Distribution in High-K Materials for Flash memory Application," *IEEE Transactions on Electron Devices*, vol. 57, no. 10, 2010.
- [7] D. Ruiz Aguado, B. Govoreanu, W.D. Zhang, M. Jurczak, K. De Meyer, and J. Van Houdt, "A Novel Trapping/De-trapping Model for Defect Profiling in High-k Materials Using the Two-Pulse Capacitance-Voltage Technique," *IEEE Transactions on Electron Devices*, vol. 57, no. 10, 2010.
- [8] M.B. Zahid, D. Ruiz Aguado, R. Degraeve, W.C. Wang, B. Govoreanu, M. Toledano-Luque, V.V. Afanas'ev, and J. Van Houdt, "Applying Complementary Trap Characterization Technique to Crystalline γ -Phase- Al_2O_3 for Improved Understanding of Nonvolatile Memory Operation and Reliability," *IEEE Transactions on Electron Devices*, vol. 57, no. 11, 2010.
- [9] H.-J. Jang, and W.-J. Cho, "High performance silicon-on-insulator based ion-sensitive field-effect transistor using high-k stacked oxide sensing membrane," *Applied Physics Letters*, vol. 99, 2011.
- [10] E. Yilmaz, I. Dogan, and R. Turan, "Use of Al_2O_3 layer as dielectric in MOS based radiation sensors fabricated on a Si substrate," *Nuclear Instruments and Methods in Physics Research B*, vol. 266, 2008.
- [11] J. Robertson, "High dielectric constant oxides," *Eur. Phys. J. Appl. Phys.*, vol. 28, pp. 265-291, 2004.
- [12] N. A. Chowdhury and D. Misra, "Charge Trapping at Deep States in Hf-Silicate Based High-K Gate Dielectrics," *Journal of The Electrochemical Society*, 154(2)G, pp. 30-37, 2007.
- [13] A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H. E. Maes, and U. Schwalke, "Origin of the Threshold Voltage Instability in SiO_2/HfO_2 Dual Layer Gate Dielectrics," *IEEE Electron Device Letters*, vol. 24, no. 2, 2003.
- [14] G. Puzzilli, B. Govoreanu, F. Irrera, M. Rosmeulen, and J. Van Houdt, "Characterization of charge trapping in SiO_2/Al_2O_3 dielectric stacks by pulsed C-V technique," *Microelectronic Reliability*, vol. 47, 2007.
- [15] M. Cho, R. Degraeve, P. Roussel, B. Govoreanu, B. Kaczer, M. B. Zahid, E. Simoen, A. Arreghini, M. Jurczak, J. Van Houdt, and G. Groeseneken, "A consistent model for oxide trap profiling with the Trap Spectroscopy by Charge Injection and Sensing (TSCIS) technique," *Solid-State Electronics*, vol. 54, 2010.
- [16] P. J. McWhorter, S. L. Miller, and W. M. Miller, "Modeling the anneal of radiation-induced trapped holes in a varying thermal environment," *IEEE Transactions on Nuclear Science*, vol. 37, no. 6, 1990.