# Two-Pulse $C$–$V$: A New Method for Characterizing Electron Traps in the Bulk of $SiO_2$/high-$\kappa$ Dielectric Stacks

W. D. Zhang, B. Govoreanu, X. F. Zheng, D. Ruiz Aguado, M. Rosmeulen, P. Blomme, J. F. Zhang, and J. Van Houdt

***Abstract*—$SiO_2$/high-$\kappa$ dielectric stack is a candidate for replacing the conventional $SiO_2$-based dielectric stacks for future Flash memory cells. Electron traps in the high-$\kappa$ layer can limit the memory retention via the trap-assisted tunneling, and there is a pressing need for their characterization. A new two-pulse $C$–$V$ measurement technique is developed in this letter, which, for the first time, allows us to probe the discharge of electron traps throughout the $SiO_2$/high-$\kappa$ stack. It complements the charge pumping technique, which can only probe near-interface traps. It is demonstrated that a large number of electron traps, indeed, exist in the bulk of high-$\kappa$ layer. Bulk electron traps also have different discharge characteristics from those near the $SiO_2$/high-$\kappa$ interface.**

***Index Terms*—$Al_2O_3$, electron trap, energy distribution, Flash memory, floating gate, high-$\kappa$ dielectrics, interpoly dielectric (IPD) layer, pulsed $C$–$V$.**

## I. Introduction

THE SCALING of the conventional tunnel and control $SiO_xN_y$ in Flash memory technologies is fast approaching its limits, as the increased leakage current makes the data loss unacceptable [1], [2]. Introduction of $SiO_2$/high-$\kappa$ stacks has been proposed as a potential solution [1]–[3]. Higher dielectric constant will increase the capacitance without reducing the physical thickness, help reducing the leakage, and maintain the coupling ratio to allow the downscaling of the cell size. A large amount of works [2]–[5] have been carried out to investigate the capabilities and limits of using high-$\kappa$ layers in Flash memory cells.

It has been reported that the density of electron traps in high-$\kappa$ layers is orders of magnitude higher than that in conventional $SiO_2$ [6], [7]. Simulation results [8] show that traps in the high-$\kappa$ interpoly dielectric (IPD) stack play a dominating role in the low-field leakage through trap-assisted tunnelling and hence determine the memory retention. Traditionally, discharging from the dielectric layer is measured from the shift of conventional $C$–$V$ [9] on capacitors and transfer characteristics on MOSFETs [10]. These measurements generally take from several seconds to tens of seconds and are too slow for probing traps in thin dielectric layers. To gain a better understanding of these traps, several groups have recently used various charge pumping methods to probe the traps in high-$\kappa$ layers [11]–[15]. However, charge pumping can only reach the traps within 1 nm into the high-$\kappa$ layer when the thickness of interfacial $SiO_2$ layer is around 1 nm [14]. High-$\kappa$ layers used in Flash memory cells are much thicker, usually in the range of 5–15 nm. The interfacial $SiO_2$ layer in Flash memory cells may also be thicker than 1 nm [2], [3], which in turn reduces the depth in the high-$\kappa$ layer that can be probed by charge pumping measurements. There is an urgent need for developing a technique that is capable of probing traps throughout the whole $SiO_2$/high-$\kappa$ stack. In this letter, a new two-pulse $C$–$V$ technique is developed to meet this need.

## II. Devices

$SiO_2/Al_2O_3$ MOS capacitors were fabricated at IMEC using a process flow similar to that used for forming the IPD stacks in floating gate Flash memory devices. A 2-nm-thick high-temperature oxide was deposited as a bottom layer on an n-type Si substrate without p-n junctions. A 6-nm $Al_2O_3$ layer was then deposited by atomic layer chemical vapor deposition, followed by a conventional post deposition anneal. Device fabrication was completed by deposition and etching of a TiN gate. The size of capacitors is $9 \times 10^{-4}$ cm$^2$.

## III. New Technique and Typical Results

### A. Shortcomings of Conventional Techniques

The charge pumping technique relies on the measurement of a recombination current, which is proportional to frequency. The trapping/detrapping times depend exponentially on the tunneling distance [9]. This requires lower frequency in order to be able to probe traps situated farther from the Si substrate. The practical limitation of the frequency is typically around 100 Hz, giving a tunneling time of less than 10 ms. We will show that the traps probed within this time are less than 10% of the total in

Manuscript received March 13, 2008; revised May 30, 2008. This work was supported in part by the EPSRC under Grant EP/C508793/2 and in part by the HEFCE PRF scheme. The review of this letter was arranged by Editor J. Cai.

W. D. Zhang, X. F. Zheng, and J. F. Zhang are with the School of Engineering, Liverpool John Moores University, L3 3AF Liverpool, U.K. (e-mail: W.Zhang@ljmu.ac.uk).

B. Govoreanu, M. Rosmeulen, P. Blomme, and J. Van Houdt are with the RDO/PT Division, Interuniversity Micro-Electronics Center, 3001 Leuven, Belgium.

D. Ruiz Aguado is with the RDO/PT Division, Interuniversity Micro-Electronics Center, 3001 Leuven, Belgium, and also with the Catholic University of Leuven, 3001 Leuven, Belgium.

Digital Object Identifier 10.1109/LED.2008.2001234

0741-3106/$25.00 © 2008 IEEE

Authorized licensed use limited to: IEEE Xplore. Downloaded on April 6, 2009 at 07:11 from IEEE Xplore. Restrictions apply.

a typical sample for Flash memory application. Consequently, charge pumping is not suitable for probing the bulk traps in thicker high-$\kappa$ layers.

The traditional measurements using quasi-dc $C$–$V$ [9] or transfer characteristics [10] take at least seconds and are not suitable for characterizing the stack because of the fast detrapping near the interface [6], [7]. It will be shown in Section III-C that even with a 2-nm interfacial layer used in this letter, detrapping can reach about 30% of the total within 1 s. Quasi-dc techniques are too slow for characterizing thin stacks.

To overcome the inherent shortcoming of conventional techniques, we will develop a technique based on the fast pulsed measurement technique that is not relying on the pulse frequency. This allows the detrapping time varying from values short enough for probing defects close to the interfacial layer to those long enough for probing traps up to the gate.

### B. Two-Pulse $C$–$V$ Measurements

The pulsed $C$–$V$ technique [16]–[18] is used to measure the flatband voltage. Similar to the conventional $C$–$V$ measurements, the rising and falling slopes of a pulse signal applied to the gate will give rise to a displacement current proportional to the capacitance and the pulse ramp rate. The ramp rate is set at a high level of 10 kV/s to minimize the trapping/detrapping during the ramp. As a calibration of the pulsed $C$–$V$, it was compared with the conventional $C$–$V$, and a good agreement was obtained on a reference $SiO_2$ capacitors [18]. Further experimental details can be found from earlier works [17], [18].

Since there is a minimum 4-s delay between each measurement due to the equipment initialization, single-pulse $C$–$V$ cannot be used to accurately measure the discharging-induced $V_{\mathrm{FB}}$ shift. A two-pulse $C$–$V$ technique was therefore developed to measure the flatband voltage shift caused by charging and discharging the traps in dielectric. A two-pulse signal is generated by the pulse generator so that the delay between the two pulses can be accurately controlled to the level of milliseconds. As shown in the inset of Fig. 1, the first pulse is used to charge the capacitor, and the flatband voltage after electron trapping can be determined from its ramp-down trace. The capacitor is then discharged for a period of time $T_{\mathrm{discharge}}$ under a gate bias of $V_{\mathrm{base}}$ until the beginning of the second pulse. The amount of discharge can be determined by the flatband voltage shift between the ramp-up trace of the second pulse and the ramp-down trace of the first pulse. In this letter, $T_{\mathrm{discharge}}$ and $V_{\mathrm{base}}$ are varied to achieve different levels of discharge. $V_{\mathrm{top}}$ and pulsewidth are kept constant to ensure the same charging level for each pulse measurement.

Fig. 1 shows a typical result of several two-pulse $C$–$V$ measurements with variable $T_{\mathrm{discharge}}$ at $V_{\mathrm{base}} = -1$ V. The first up-trace is from the first pulse on a fresh device without trapping. The difference between this up-trace and the down-traces gives the total amount of trapping. The rest of the up-traces are from the subsequent second pulses. They move toward the left when $T_{\mathrm{discharge}}$ increases, indicating an increase of discharge. $V_{\mathrm{FB}}$ does not return to its fresh value even after 100 s of discharge, indicating that some electrons are still trapped in the high-$\kappa$ layer.

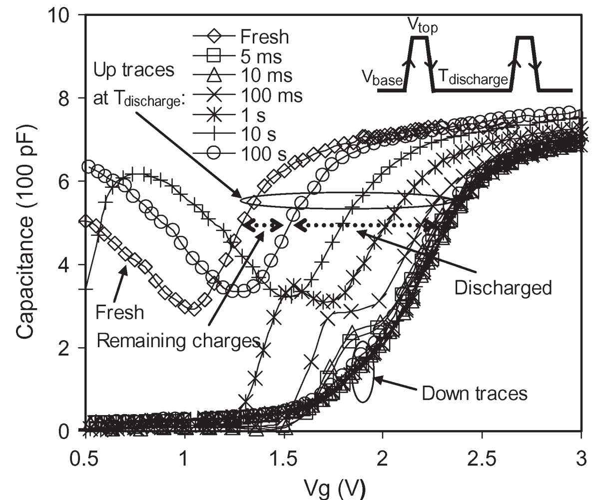


Fig. 1. Inset shows the waveform of the two gate pulses in a typical two-pulse $C$–$V$ test. The first pulse is used to charge the capacitor. The flatband voltage determined from its ramp-up trace indicates the fresh $V_{\mathrm{FB}}$ value without charge trapping, and the flatband voltage after electron trapping can be determined from the ramp-down trace. The capacitor is then discharged for a period of time, $T_{\mathrm{discharge}}$, under a gate bias of $V_{\mathrm{base}}$ until the beginning of the second pulse. The amount of discharge can be determined by the flatband voltage shift between the ramp-up trace of the second pulse and the first ramp-down trace. In this case, the top voltage level of the pulses is 5 V, $V_{\mathrm{base}} = -1$ V, and the pulsewidth is 1 ms. The ramp-up/ramp-down rate on the rising/falling edges is 10 kV/s.

Note that there is a peak or kink on the up-traces at $V_g$ smaller than $V_{\mathrm{FB}}$. This is caused by the minority carriers (holes) generated in the substrate during $T_{\mathrm{discharge}}$ at $V_{\mathrm{base}} = -1$ V. For $T_{\mathrm{discharge}} = 100$ s, the device is biased at $V_{\mathrm{base}} = -1$ V long enough so that equilibrium and inversion are approached before the ramp-up, and the up-trace behaves like the quasi-static $C$–$V$. This is because the recombination of holes with electrons is a rapid process that can follow the change in $V_g$ and contributes to the displacement current, unlike the slow generation of minority carriers (holes). The shorter the $T_{\mathrm{discharge}}$, the less holes are available and the weaker the inversion becomes, and the smaller the kink. For the down-trace, the fast ramp rate does not allow the minority carrier generation during the ramp-down, and the down-trace starts from accumulation at $V_{\mathrm{top}} = 5$ V and ends in deep depletion at $V_{\mathrm{base}} = -1$ V. Therefore, inversion is not achieved for the down-trace, regardless of the $T_{\mathrm{discharge}}$. As $T_{\mathrm{discharge}}$ reduces, the up-trace approaches the down-trace. The complication introduced by the thermal nonequilibrium at short time makes the inversion region unsuitable for measuring discharge. In this letter, we always measure the discharge from the flatband voltage shift.

It is also noted that all the ramp-down traces in Fig. 1 are overlapped, suggesting that the total trapping level changes little during each measurement. Similar results have been observed and confirmed by using conventional HF $C$–$V$ measurements. This can also be clearly seen in Fig. 2, where a number of two-pulse measurements were carried out with $V_{\mathrm{base}}$ sweeping from $-3$ to 1.7 V at a step of 0.1 V and $T_{\mathrm{discharge}} = 1000$ s. While the flatband voltage on the up-traces of the second pulse decreases when $V_{\mathrm{base}}$ gets more negative due to the increased discharge, the flatband voltage on the down-traces and, therefore, the total trapping level are not affected by the

Authorized licensed use limited to: IEEE Xplore. Downloaded on April 6, 2009 at 07:11 from IEEE Xplore. Restrictions apply.

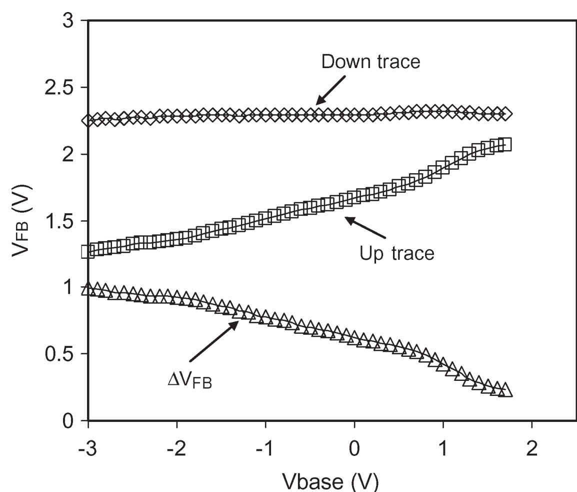


Fig. 2. Trapping levels during a number of charging and discharging cycles with $V_{\rm base}$ sweeping from $-3$ to 1.7 V at a step of 0.1 V and $T_{\rm discharge} =$ 1000 s. Symbol "◇": The flatband voltages measured on the down-traces represent the total trapping level after the charging. Symbol "□": The flatband voltages measured on the up-traces represent the trapping level after the discharging. Symbol "Δ": The difference between the corresponding flatband voltages on the up- and down-traces represents the amount of discharge.

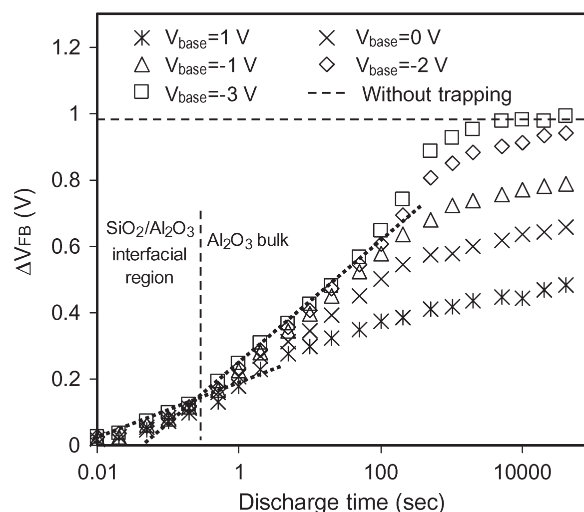


Fig. 3. Discharge-induced $\Delta V_{\rm FB}$ against discharge time at various $V_{\rm base}$. The test procedure is the same as that in Fig. 2. The dashed horizontal line is the total trapping level. The two dotted lines are linear fittings to the data at $V_{\rm base} = -3$ V before and after the slope changes. Their intersection at 200 ms indicates the transition time from $SiO_2$ to $Al_2O_3$.

number of pulses, indicating that the traps are as-grown, rather than generated by the stress. This allows us to simplify the test procedure by using the $V_{\rm FB}$ taken from the down-trace of the second pulse instead of that from the first pulse as the charged value. The difference between the up- and down-traces, $\Delta V_{\rm FB}$, will be used as a measure of the discharged trap density in this letter.

Typical two-pulse $C$–$V$ test conditions used in the letter are as follows. The top voltage level of the pulse is 5 V and the pulsewidth is 1 ms, to emulate the voltage drop across the IPD layer in floating gate memory cells during a typical programming session. The base voltage level $V_{\rm base}$ is varied between $-3$ V and the flatband voltage. The minimum $T_{\rm discharge}$ is limited to milliseconds, but there is no limit on the maximum $T_{\rm discharge}$. This provides an effective way to probe defects across the whole high-$\kappa$ stack.

### C. Electron Traps Probed by the Two-Pulse $C$–$V$ Technique

In the following tests, it will be demonstrated that there are, indeed, a large amount of electron traps located deep in the high-$\kappa$ bulk that can be probed by the two-pulse $C$–$V$. A number of two-pulse $C$–$V$ with various $V_{\rm base}$ and $T_{\rm discharge}$ were measured, and the $\Delta V_{\rm FB}$ induced by discharge is given in Fig. 3. Several features can be observed, as detailed next.

1) It has been reported [14] that the detrapping time for traps situated at 2 nm from the $Si/SiO_2$ interface is on the order of 100 ms. Fig. 3 shows that the discharging within this time is less than 20% of the total trapping. Trapping is dominated by the high-$\kappa$ layer, which cannot be probed by charge pumping in our case. The discharging within 1 s is about 30% of the total, which will be missed if the slow quasi-dc techniques were used.
2) When the discharging front reaches the $SiO_2/Al_2O_3$ interface, a change in the slope of $\Delta V_{\rm FB}$ versus $\log(T_{\rm discharge})$ can be clearly observed, as illustrated by the two dotted lines, which are linear fittings to the data at $V_{\rm base} = -3$ V for relative short ($\leq$ 200 ms) and long ($\geq$ 500 ms) time. At a constant discharge bias, this slope is proportional to the local trapping density [9], [11]. An increase of the slope when moving into the high-$\kappa$ layer confirms that trapping density in $Al_2O_3$ is higher than that in $SiO_2$. The two lines intersect at about 200 ms, which corresponds to the transition from $SiO_2$ to $Al_2O_3$ and agrees well with the calculation in [14]. The change of the slope is less prominent at a lower $V_{\rm base}$, such as 1 or 0 V where only partial discharge in the high-$\kappa$ layer is achieved.
3) For the highest bias $|V_{\rm base} = -3\ {\rm V}|$, $\Delta V_{\rm FB}$ saturates after 1000 s approximately. To find if this saturation originates from a complete discharge of trapped electrons, the $\Delta V_{\rm FB}$ corresponding to the total amount of trapped charge is plotted as the dashed horizontal line in Fig. 3. The excellent agreement between the discharge saturation and total trapping level confirms that 100% discharge was achieved. This demonstrates the capability of the two-pulse $C$–$V$ technique for probing traps throughout the whole stack. The high workfunction of the TiN gate used here prevents charging from the gate during the discharge at negative $V_{\rm base}$.
4) It is also observed that, at long discharge times, the amount of discharge increases when $V_{\rm base}$ becomes more negative. This indicates that electron traps with deeper energy levels may exist in the bulk of $Al_2O_3$ layer, which can only be discharged at larger gate bias [12]. This awaits further investigation.

In summary, for the first time, a new two-pulse $C$–$V$ measurement technique is developed in this letter. When compared with the charge pumping, the main strength of this technique is that there is no limitation on the maximum discharging time, so that the new technique can probe traps throughout

Authorized licensed use limited to: IEEE Xplore. Downloaded on April 6, 2009 at 07:11 from IEEE Xplore. Restrictions apply.

the whole stack. It is found that electron trapping in the 2-nm $SiO_2/6$-nm $Al_2O_3$ stack is dominated by $Al_2O_3$ layer, which cannot be reached by charge pumping. This makes the two-pulse $C$–$V$ a powerful tool for characterizing traps in high-$\kappa$ stacks with a thickness suitable for future Flash memory applications.

## ACKNOWLEDGMENT

The authors would like to thank their colleagues for the kind support. The authors would also like to thank Dr. R. Degraeve at IMEC, Dr. D. Brunco (Intel assignee at IMEC), Dr. S. P. Sim (Samsung assignee at IMEC), Dr. K. van der Zanden (Infineon assignee at IMEC), and Dr. V. Srividya (Micron assignee at IMEC) for the fruitful discussions. The test sample used in this work was provided by IMEC.

## REFERENCES

[1] *International Technology Roadmap for Semiconductors*, 2005, San Jose, CA. [Online]. Available: http://public.itrs.net

[2] J. Van Houdt, "High-$\kappa$ materials for non-volatile memory applications," in *Proc. IRPS*, 2005, pp. 234–239.

[3] B. Govoreanu, D. P. Brunco, and J. Van Houdt, "Scaling down the interpoly dielectric for next generation flash memory: Challenges and opportunities," *Solid State Electron.*, vol. 49, no. 11, pp. 1841–1848, Nov. 2005.

[4] Y. Y. Chen, C. H. Chien, and J. C. Lou, "High quality $Al_2O_3$ IPD with $NH_3$ surface nitridation," *IEEE Electron Device Lett.*, vol. 24, no. 8, pp. 503–505, Aug. 2003.

[5] A. H. Miranda, R. van Schaijk, M. van Duuren, N. Akil, and D. S. Golubović, "Reliability comparison of $Al_2O_3$ and HfSiON for use as interpoly dielectric in flash arrays," in *Proc. ESSDERC*, Montreux, Switzerland, 2006, pp. 234–237.

[6] G. Groeseneken, L. Pantisano, L.-Å. Ragnarsson, R. Degraeve, M. Houssa, T. Kauerauf, P. Roussel, S. De Gendt, and M. Heyns, "Achievements and challenges for the electrical performance of MOSFETs with high-$\kappa$ gate dielectrics," in *Proc. 11th IPFA*, Taiwan, R.O.C, Jul. 5–8, 2004, pp. 147–155.

[7] C. Z. Zhao, J. F. Zhang, M. B. Zahid, B. Govoreanu, G. Groeseneken, and S. De Gendt, "Determination of capture cross sections for as-grown electron traps in $HfO_2$/HfSiO stacks," *J. Appl. Phys.*, vol. 100, no. 9, p. 093 716, Nov. 2006.

[8] B. Govoreanu, D. Wellekens, L. Haspeslagh, J. De Vos, and J. Van Houdt, "Investigation of the low-field leakage through high-$\kappa$ interpoly dielectric stacks and its impact on nonvolatile memory data retention," in *IEDM Tech. Dig.*, 2006, pp. 206–209.

[9] J. F. Zhang, S. Taylor, and W. Eccleston, "A comparative study of the electron trapping and thermal detrapping in $SiO_2$ prepared by plasma and thermal oxidation," *J. Appl. Phys.*, vol. 72, no. 4, pp. 1429–1439, Aug. 1992.

[10] W. D. Zhang, J. F. Zhang, M. Lalor, D. Burton, G. Groeseneken, and R. Degraeve, "Two types of neutral electron traps generated in the gate silicon dioxide," *IEEE Trans. Electron Devices*, vol. 49, no. 11, pp. 1868–1875, Nov. 2002.

[11] M. B. Zahid, R. Degraeve, L. Pantisano, J. F. Zhang, and G. Groeseneken, "Defects generation in $SiO_2/HfO_2$ studied with variable T-charge-T-discharge charge pumping $VT^2CP$," in *Proc. IEEE Int. Rel. Phys. Symp.*, 2007, pp. 55–60.

[12] E. Cartier, B. P. Linder, V. Narayanan, and V. K. Paruchuri, "Fundamental understanding and optimization of PBT1 in nFETs with $SiO_2/HfO_2$ gate stack," in *IEDM Tech. Dig.*, 2006, pp. 57–60.

[13] D. Heh, C. D. Young, G. A. Brown, P. Y. Hung, A. Diebold, E. M. Vogel, B. Joseph, and G. Bersuker, "Spatial distributions of trapping centers in $HfO_2/SiO_2$ gate stack," *IEEE Trans. Electron Devices*, vol. 54, no. 6, pp. 1338–1345, Jun. 2007.

[14] Y. Wang, V. Lee, and K. P. Cheung, "Frequency dependent charge-pumping, how deep does it probe?" in *IEDM Tech. Dig.,* 2006, pp. 491–494.

[15] I. Crupi, R. Degraeve, B. Govoreanu, D. P. Brunco, P. J. Roussel, and J. Van Houdt, "Energy and spatial distribution of traps in $SiO_2/Al_2O_3$ nMOSFETs," *IEEE Trans. Device Mater. Rel.*, vol. 6, no. 4, pp. 509–516, Dec. 2006.

[16] Z. A. Weinberg and M. V. Fischetti, "$SiO_2$-induced substrate current and its relation to positive charge in field-effect transistors," *J. Appl. Phys.*, vol. 59, no. 3, pp. 824–832, Feb. 1986.

[17] M. Rosmeulen, E. Sleeckx, and K. De Meyer, "Electrical characterization of silicon-rich-oxide-based memory cells using pulsed current-voltage techniques," in *Proc. Eur. Solid-State Device Res. Conf.*, 2002, pp. 471–474.

[18] 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," *Microelectron. Reliab.*, vol. 47, no. 4/5, pp. 508–512, 2007.

Authorized licensed use limited to: IEEE Xplore. Downloaded on April 6, 2009 at 07:11 from IEEE Xplore. Restrictions apply.