An Accurate Capacitance–Voltage Measurement Method for Highly Leaky Devices—Part II

Y. Wang, Kin P. Cheung, *Senior Member, IEEE*, R. Choi, and B.-H. Lee

***Abstract*—In Part I, an accurate C–V measurement based on time-domain reflectometry (TDR) for MOS capacitors in the presence of a high level of leakage across the gate dielectric was presented. This new method is expected to have high accuracy even in the presence of a very high level of leakage current. In this paper, the basic TCR-based C–V measurement is extended to handle the parasitic, allowing the overlap capacitance to be extracted simultaneously and accurately without the need for additional measurement. In addition, a detailed error analysis is provided to complete the description of the TDR C–V measurement method.**

***Index Terms*—C–V, leakage, RF capacitor, time domain, time-domain reflectometry (TDR), ultrathin oxide.**

I. Introduction

In PART I, a detailed description of a time-domain-reflectometry (TDR)-based C–V measurement and its theoretical foundation was given. Briefly, a fast voltage step is sent from the TDR scope via a transmission line to the wafer-level device under test (DUT; a MOS capacitor). The impedance mismatch causes the step waveform to be reflected back toward the TDR scope which makes a record of it. The basic setup is shown in Fig. 1. The bias tee inserted between the TDR scope and the transmission line is for the introduction of a dc bias. Fig. 2 shows the captured step waveforms reflected from an open circuit, a MOS capacitor under depletion (with negligible leakage), and the same MOS capacitor under strong accumulation (with high leakage). The inset shows the equivalent circuit of the MOS capacitor.

A rigorous mathematical expression was derived to extract the capacitance from these curves

$$C = \frac{1}{2Z_0 V_{\text{step}}} M \int_0^{\infty} \left[\left(\frac{R_0 - Z_0}{R_0 + Z_0} \right) V_{\text{Open}}(t) - V_{\text{TDR}}(t) \right] dt \tag{1}$$

where $M = (R_0 + Z_0)^2 / R_P^2$.

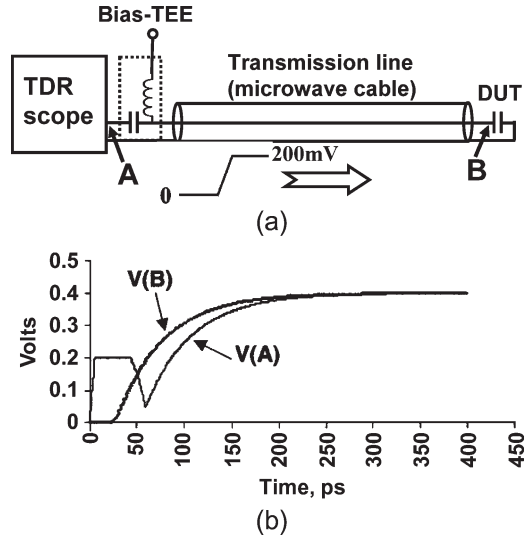

Fig. 1. (a) Basic experimental setup for the TDR measurement. A TDR scope connects to the DUT through a bias tee, a microwave cable, and an RF probe. The TDR scope sends out a fast step voltage and monitors both the outgoing and returning waveforms. (b) The step-voltage (200-mV) waveform at points A and B. The TDR scope monitors the waveform at point A which is the combined result of the outgoing waveform and the reflected waveform.

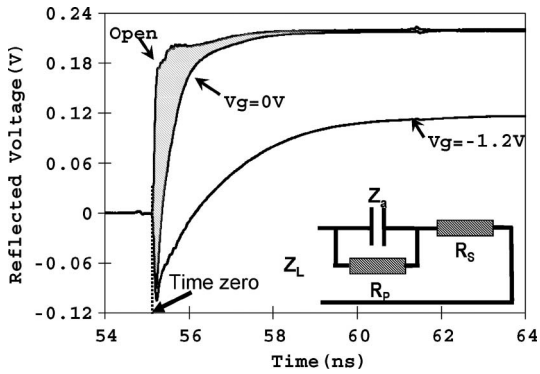

Fig. 2. Reflected waveforms from open circuit (reference) and MOS capacitor (SiO2, 2 nm) at depletion ($V_g = 0$ V) and accumulation ($V_g = -1.2$ V). The depletion capacitance is smaller and therefore has a shorter charging time. In accumulation, high leakage causes the final voltage level to be much lower than the reference. The shaded area represents the total stored charge in the fully charge capacitor for the depletion case. The inset shows the equivalent circuit of the capacitor with thin oxide. Step height = 200 mV.

In the expression, $V_{\text{Open}}(t)$ is the reflected step waveform from an open-circuit reference, $V_{\text{TDR}}(t)$ is the reflected step waveform from the DUT (the capacitor), V_{step} is the height of the step, Z_0 is the impedance of the transmission line (typically 50 Ω), and R_0 is the sum of series resistance (R_S) and shunt

Manuscript received May 27, 2008. The review of this paper was arranged by Editor C. McAndrew.

Y. Wang is with the Department of Electrical and Computer Engineering, Rutgers University, Camden, NJ 08102-1519 USA.

K. P. Cheung is with the Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899 USA (e-mail: kin.cheung@nist.gov).

R. Choi and B.-H. Lee are with SEMATECH, Austin, TX 78741 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2008.927659

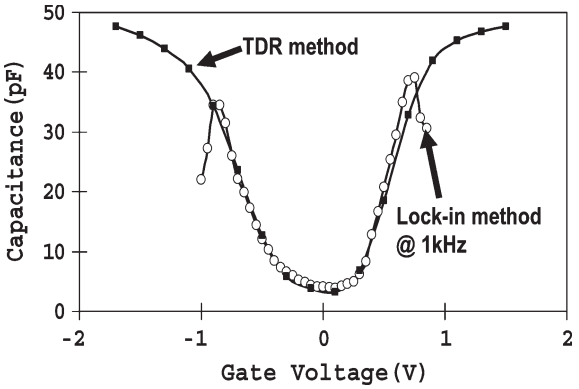

Fig. 3. Measured C–V curves of a MOS capacitor with TiN gate and high-κ gate dielectrics (3-nm HfO_2 and 1-nm SiO_2, EOT = 1.2 nm) from the TDR method and the conventional lock-in method. In the depletion region where the conventional method is accurate, good agreement is evidenced. In the region where leakage is severe, only the TDR method produces the proper C–V curve. Step height = 200 mV.

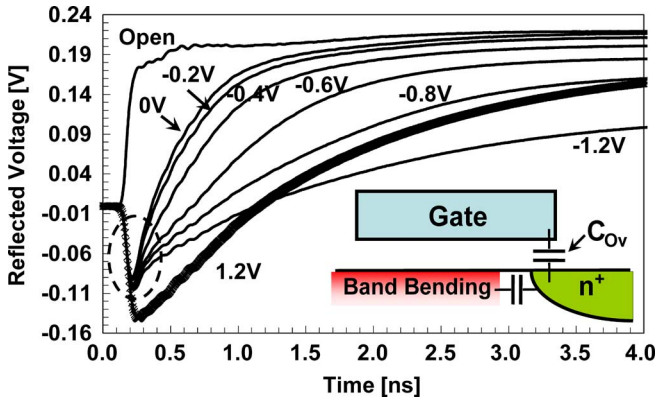

Fig. 4. Reflected voltage curves from the SiO_2 capacitor. Only the negative-bias curves are included to highlight the nonsingle time constant charging behavior. The curves from open circuit and from strong inversion are also shown for reference. (Inset) Illustration of two additional capacitors existing due to the presence of source and drain. One is the overlap capacitor $C_{\rm ov}$, and the other is band-bending capacitor $C_{\rm Bb}$.

(leakage) resistance (R_P). When $R_P \gg R_S$, and $R_P \gg 50\,\Omega$, (1) reduces to

$$C = \frac{1}{2Z_0 V_{\rm Step}} \int_0^\infty [V_{\rm Open}(t) - V_{\rm TDR}(t)]\, dt. \qquad (2)$$

The integral in (2) is the shade area in Fig. 2 between the reflected waveforms from the open-circuit reference and from the MOS capacitor under depletion, which can be found with high accuracy using numerical integration. The integrand in (1) multiplies the open-circuit waveform by a fraction to handle the general case where leakage can be high. With this multiplication, the magnitude of the reflected waveform from the open-circuit reference is normalized to the magnitude of the reflected waveform from the DUT to create an enclosed area for integration. Intuitively, it is clear that the resulting integral is not the total charge flowed into the capacitor. However, the correction factor M in front of the integral restores the capacitance to the value in the absence of leakage. Fig. 3 shows the extracted C–V curve using the TDR method. Also shown is the C–V curve, as measured by a lock-in amplifier at the frequency of 1 kHz. The agreement is very good in the depletion region where the lock-in method is known to be accurate. The agreement in strong accumulation and strong inversion regions are poor. In these regions, only the TDR method produced the expected C–V shape.

In this paper, the basic TDR method is extended to handle a more complex but realistic situation of transistorlike capacitors. The accuracy of the TDR method is discussed in detail and is experimentally verified.

II. Extending the Method—Handling the Overlap Capacitance

Inversion capacitance is part of the C–V curve that is highly important for CMOS technology. When leakage is high, the traditional low-frequency method for producing the inversion portion of the C–V curve does not work because the little inversion charges that can be generated are leaked across the dielectric to the gate. For highly leaky dielectrics, transistorlike capacitors with source and drain tied to the substrate are needed to support the inversion charge [1]–[7]. The capacitor that produces the C–V curve, as shown in Fig. 3, is transistorlike, and the measured C–V curve includes the overlap capacitance. In most C–V measurements, we need to separate the channel capacitance from the overlap capacitance. Measuring the overlap capacitance using a conventional LCR meter suffers from the same high level of leakage problem. Thus, it is important to be able to extract the overlap capacitance using the TDR method and preferably not requiring a separate test structure (the RF design for such a structure will be extremely difficult).

Fortunately, the TDR method can extract both the channel capacitance and the overlap capacitance simultaneously and accurately from the same reflected voltage waveform. Thus, it accomplishes the task without requiring a separate measurement.

Fig. 4 shows the reflected voltage waveforms for the SiO_2 capacitor under negative biases ranging from depletion to accumulation. It is clear that, under negative bias, there is a faster rising section at the beginning of the capacitor charging curves. These waveforms behave like two capacitors being charged in parallel, with the smaller one charging up much faster than the larger one. What is this additional capacitor?

Parasitic capacitances due to probe pads and cables are not possible candidates because we would see them in the reference waveform as well. An immediate suspect for the small capacitor is the overlap capacitor $C_{\rm ov}$. However, an additional capacitor due to surface band bending also exists, as shown in the inset of Fig. 4. On the other hand, this capacitor is much smaller. Thus, the source of the additional small capacitor is the overlap capacitor.

The equivalent circuit of the oxide capacitor $(C_{\rm gc})$ plus the overlap capacitor $(C_{\rm ov})$ is shown in the inset of Fig. 5. The overlap capacitor and the oxide capacitor will both be charged by the step function. In the case of accumulation, at the beginning of charging, because the series resistance of the substrate contact $(R_{s,\rm acc})$ is larger than the series resistance of the source/drain contact $(R_{s,\rm inv})$, the overlap capacitor path has lower impedance and therefore dominates the overall impedance and the reflectivity. With larger current and smaller

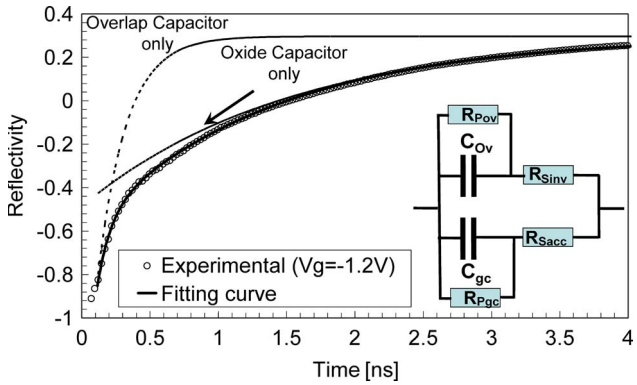

Fig. 5. Comparison of the fitting curve with the experimental reflectivity curve of the capacitor at $V_g = -1.2$ V. A simulation with a single capacitor charging is also included to show that it provides a poor fit. (Inset) The equivalent circuit of the oxide capacitor with the overlap capacitor.

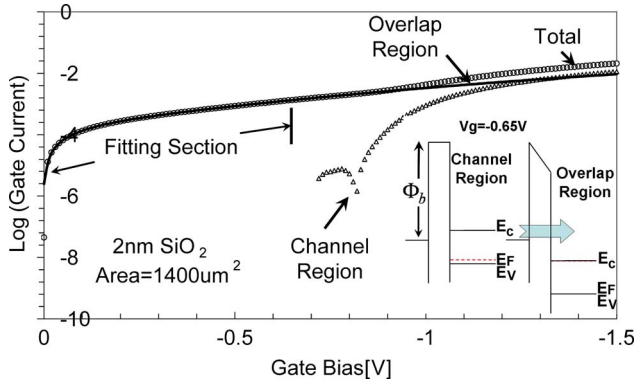

Fig. 6. I–V curve of a 2-nm SiO_2 MOS capacitor. Leakage current from the zero to the flatband (-0.65-V) gate bias range was fitted with a known tunneling function and then extrapolated to -1.2 V to get the leakage component from the overlap region. The leakage current from the channel region is obtained by subtracting the leakage current of the overlap region from the measured total leakage current. (Inset) Band diagram of the overlap and gate-to-channel regions, showing that the overlap region completely dominated the leakage at low gate bias.

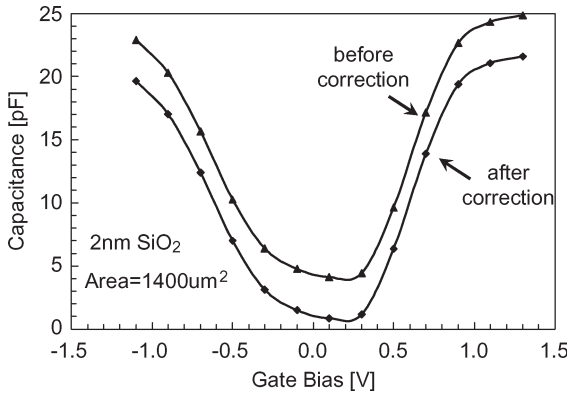

Fig. 7. (Upper curve) Measured C–V curve for the 2-nm SiO_2 capacitor with 1400-μm^2 area and (lower curve) the corrected C–V curve after the removal of the overlap capacitance.

capacitance, the overlap capacitor charges up quickly. The rapid charging of the overlap capacitor continues until the current ratio of the two paths reflects the capacitance of the two capacitors. At this point, the impedance of the oxide capacitor path is much lower and dominates the overall impedance.

The aforementioned parallel charging process can be simulated [8] if all the circuit components are known. Conversely, we should be able to extract the overlap capacitance from the reflectivity curve extracted from the TDR measurement using the equivalent circuit, as shown in the inset of Fig. 5.

The equivalent circuit, as shown, has six elements. If we let them all be fitting parameters, the outcome will not be very meaningful. Therefore, we must pin down the value of as many elements as possible independently. Using the methodologies discussed in Part I, we can extract the total capacitance ($C_T = C_{\rm gc} + C_{\rm ov}$) and series resistances $R_{s,\rm acc}$ and $R_{s,\rm inv}$. The ratio of $C_{\rm ov}$ and C_T is simply the ratio of the area of the overlap region over the total capacitor area f. This area ratio can, in theory, be known from the mask layout. However, in practice, the ratio in an as-fabricated device is quite different from the designed ratio. We are now down to three parameters.

For the shunt resistances $R_{p,\rm gc}$ and $R_{p,\rm ov}$, their sum at accumulation (-1.2 V) can be obtained from the current–voltage (I–V) measurement of the capacitor. According to the band diagram (inset of Fig. 6), for the gate bias from 0 V to the flatband voltage (-0.65 V), the channel region is at depletion, whereas the full gate voltage drops across the gate dielectric in the highly doped overlap region. The leakage current is therefore completely dominated by the overlap region. We can thus fit confidently the leakage current in this range with the known tunneling function [9] and then extrapolate the leakage current from the overlap region up to -1.2 V. With the overlap region's leakage current for the entire voltage range being known, it can be subtracted from the measured leakage current to obtain the leakage current from the channel region, as shown in Fig. 6. With that, we are down to one fitting parameter which is the area ratio f.

The area ratio f and, therefore, the overlap capacitance can be found from the best fit of the equivalent circuit to the extracted reflectivity curve (Fig. 5). The reflectivity curve at strong accumulation ($V_g = -1.2$ V) was selected for the fitting procedure because it produces the largest difference in capacitance between the channel and the overlap region. The best fit was found for $f = 0.142$, and the fit is excellent. Also shown in Fig. 5 is the result for fitting the reflectivity curve to a single capacitor. Clearly, the two-capacitor charging process is necessary to explain the experimental data.

With 23 pF as the total capacitance under this condition (2-nm SiO_2 at $V_g = -1.2$ V; see Part I), the overlap capacitance works out to be 3.27 pF. We can now remove the overlap capacitance from the measured C–V curve of the capacitor to obtain the channel capacitances. The result is shown in Fig. 7.

As discussed in the section of series resistance extraction (Part I), the reflectivity curve is particularly prone to noise near time zero. We can (and did) avoid using the reflectivity data at very early time in our extraction of series resistance. For overlap capacitance extraction, because the faster charging event happens at early time, we must use as much of the early time data as possible. Thus, before we leave this section, we need to discuss an important experimental detail.

As in most RF measurements, the reference device must be designed carefully. In our case, we have an open-circuit

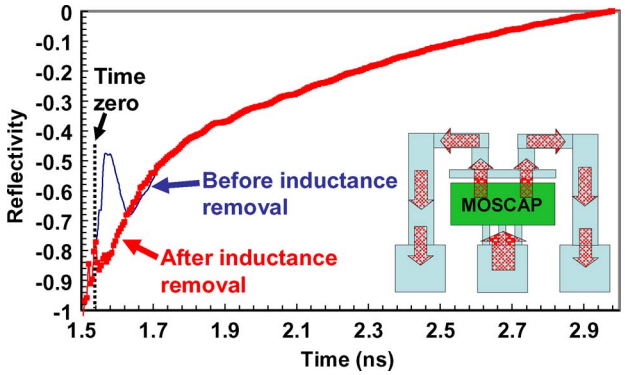

Fig. 8. Reflectivity of DUT before and after correction with inductance is shown. The bump in the reflectivity is due to the inductance, and its disappearance after correction indicates the success of the correction procedure. (Inset) Inductance arises from the current loop, as indicated by the arrows.

reference that contains merely the ground–signal–ground probe pads. When the step function is reflected, the return current flows directly across the probe pads. For the actual capacitor, the return current travels an additional loop, as shown in the inset of Fig. 8. This introduces an additional small inductance that affects the reflected waveform at very early time. This additional feature in the reflectivity curve will impact the extraction of overlap capacitance and can lead to serious error.

To accurately determine this additional inductance and to remove it, we introduced a hard breakdown in a capacitor (on a neighboring die) and used it as a short-circuit reference. Between this short-circuit reference and the open-circuit reference, the inductance effect on the reflectivity can be quantitatively measured. Once the inductance effect is quantified, it can be removed from the reflectivity curves of our experiments, as shown in Fig. 8. Of course, if we have a properly designed open-circuit reference, this would not have been necessary.

III. Error Analysis

Accurate C–V characteristics are very important for device parameter extraction, such as EOT, substrate doping, interface state density, and so on. Therefore, it is of great importance to address the accuracy of this new measurement method, particularly in the presence of a very high leakage current.

A. Accuracy of Capacitance Extraction

For the accuracy of capacitance measurement, we can measure known capacitors (control experiment). A 221-pF ceramic capacitor (1% tolerance) is used in this procedure. The tests are done by adding series and shunt resistances to this capacitor to see if the TDR method can accurately determine the capacitance of the capacitor. The capacitor alone was first measured by the lock-in method and was found to be 221.1 pF. Because it agrees with the factory specification, we consider this to be the true value of the capacitor. We note that the adding of series and parallel resistors to the capacitor inevitably adds an inductive loop. However, this inductive loop should not affect the capacitance extraction in the TDR method.

Table I shows results extracted by TDR in the presence of various shunt/series resistances. Results measured by the lock-in method under the same condition are also listed for comparison. Error is defined by comparing the TDR result to 221.1 pF. As expected, under low-shunt-resistance (high-leakage-current) conditions, the conventional C–V method becomes problematic, whereas the TDR method can still measure the capacitance to within 1% error. Only when the shunt resistance drops to below 20 Ω did the error increase rapidly. A 20-Ω shunt resistance is equivalent to a 4000-A/cm^2 leakage current density on our 1400-μm^2 MOS capacitor with 1 V across the oxide. Because we can reliably measure capacitance down to a few picofarads (small area), the upper limit of leakage current density for the TDR method can reach 100 000 A/cm^2 if the required accuracy is maintained at 1%.

B. Accuracy of Series Resistance Extraction

Series resistance extraction was discussed in Part I. It is based on the fact that the capacitors behave like a short circuit at time zero. Thus, to find the series resistance is to find the reflectivity at time zero. For nonideal step waveform used in actual TDR measurement, time-zero reflectivity cannot be measured directly and must be extracted. We will now analyze the accuracy of the series resistance extraction method. For series resistance extraction, we do not have the means to perform a controlled experiment and must therefore rely on analysis. Because we use a known function to fit a good range of relatively low-noise reflectivity data, the basic method of finding the reflectivity at time zero is highly dependable and extremely accurate. The main source of error is from the determination of time zero. As shown in Fig. 9 (same as the inset in Fig. 9 of Part I), which has an expanded time scale, the shaded area in the figure is the introduced error. It represents the total charge already flowed into the capacitor at time zero, as defined by our method. This is an error because the capacitor acts like a short circuit only when it has not been charged to any degree.

To see how much error this introduces to the determination of time zero, we need to keep in mind that we are seeking the time zero of an ideal step function. For the ideal step function, the reflected waveform at time zero has the same magnitude as the step function. The charge flowed into the capacitor per unit time at time zero is therefore proportional to twice the step function's magnitude. To find the time-zero error, all we need to do is to find the area (magnitude $\times$ time) of the shaded region and divide it by twice the step function's magnitude. Thus, the small amount of charge in the shaded area translates into an extremely small time-zero error that is smaller than the timing jitter associated with the TDR instrument, which is about 10 ps. From the slope of the reflectivity curve near time zero, we can estimate that the 10-ps jitter will contribute to 1% error in the extracted reflectivity. The error in the extracted series resistance is therefore roughly 1%.

It is instructive to see how much a 1% error in series resistance extraction affects the measurement of the capacitance using the TDR method. From Fig. 6 of Part I, the error in capacitance extraction is about twice the $\Delta R_S/(R_S + R_P)$. Because R_P is larger than R_S in almost all cases, the effect on capacitance extraction accuracy is less than 1%.

TABLE I
RESULTS OF THE CONTROL EXPERIMENT BY TDR AND LCR MEASUREMENTS

220pF Capacitor with different R_P and R_S	Obtained by Lock in Amp C_{Lock}	Extracted in TDR result C_{TDR}	Percentage of Error (221.1 pF-C_{TDR})/221.1 pF
R_S=43 Ω , R_P=1K Ω	221.1 pF	221.3 pF	0.09 %
R_S=43 Ω , R_P=500 Ω	212 pF	221.5 pF	0.18 %
R_S=43 Ω R_P=220 Ω	185 pF	221.0 pF	0.04 %
R_S=43 Ω R_P=100 Ω	135 pF	221.3 pF	0.09 %
R_S=10 Ω , R_P=47 Ω	N/A	221.5 pF	0.18 %
R_S=35 Ω , R_P=47 Ω	N/A	221.7 pF	0.27 %
R_S=5 Ω , R_P=20 Ω	N/A	219.1 pF	0.90 %
R_S=20 Ω , R_P=20 Ω	N/A	218.9 pF	0.99 %
R_S=10 Ω , R_P=10 Ω	N/A	201.1 pF	9.04 %

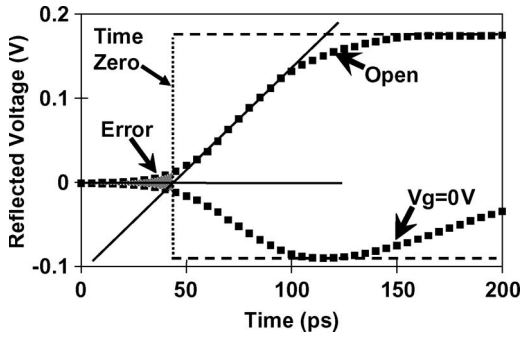

Fig. 9. Reflected waveforms, expanded in time scale, showing time zero are determined by extrapolating the steep rising edge of the open-circuit reflection to 0 V. The shaded area represents charges already flowed into the capacitor at time zero and, therefore, a small error.

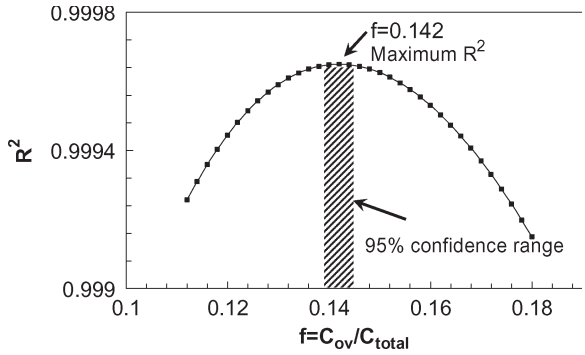

Fig. 10. R^2 value of the fitting with different ratios of overlap capacitance to total. The best fit is at the maximum R^2 condition with $f = 0.142$. The shaded range is the 95% confidence limit.

C. Accuracy of Overlap Capacitance Extraction

In the overlap capacitance extraction procedure, we perform a fitting of the equivalent circuit to the measured reflectivity data. Because we independently determine five out of the six parameters of the equivalent circuit and the accuracy of their determination has already been discussed earlier, we assume them to be accurate. In other words, in this analysis, we ignore error propagation and focus on assessing the accuracy of the area ratio extraction that has a direct bearing on the accuracy of the overlap capacitance extraction. The main task is therefore to determine how well the fitting process can pin down the value of the area ratio f. The goodness of fit (R^2 value) is a good tool for this purpose.

The R^2 as a function of the f value used for fitting is shown in Fig. 10. The best fit (maximum R^2) is $f = 0.142$. The 95% confidence limits are $f = 0.142 \pm 0.003$ or 2%. The uncertainty (95% confidence) will, of course, depend on each capacitor design. We expect that the uncertainty will decrease with the larger value of f. While it is more desirable to design transistors with lower f, the trend in practical reality is the opposite. Thus, we expect the overlap capacitance extraction to remain highly accurate as CMOS technology continues to advance. It is instructive to note that the 2% error is referring to the area ratio, not the actual overlap capacitance. The total capacitance can be determined to better than 1%. The small 2% error in separating the channel capacitance and overlap capacitance essentially leaves the determination of the overlap capacitance with the same (percentage) accuracy as the total capacitance.

IV. CONCLUSION

In summary, the new TDR-based method to measure C–V can be extended to handle the inevitable parasitic such as overlap capacitance. Thus, the TDR method can extract capacitance, series resistance, and overlap capacitance simultaneously from the same measurement. The accuracy of the extracted capacitance, series resistance, and overlap capacitance are very high even when the capacitor is extremely leaky. The ability to accurately measure the series resistance at inversion raises the possibility of measuring the series resistance of a transistor—a long standing challenge. If this could be done, the transistor effective channel length measurement will be greatly simplified. The ability to measure both the channel capacitance and overlap capacitance allows the split C–V-based measurement of mobility to be done quickly and reliably.

REFERENCES

[1] K. Ahmed, E. Ibok, G. C.-F. Yeap, Q. Xiang, B. Ogle, J. J. Wortman, and J. R. Hauser, "Impact of tunnel currents and channel resistance on the characterization of channel inversion layer charge and polysilicon-gate depletion of sub-20-angstrom gate oxide MOSFET's," *IEEE Trans. Electron Devices*, vol. 46, no. 8, pp. 1650–1655, Aug. 1999.

[2] D. W. Barlage, J. T. O'Keeffe, J. T. Kavalieros, M. M. Nguyen, and R. S. Chau, "Inversion MOS capacitance extraction for high-leakage dielectrics using a transmission line equivalent circuit," *IEEE Electron Device Lett.*, vol. 21, no. 9, pp. 454–456, Sep. 2000.

[3] H. Aoki and M. Shimasue, "Leaky large area gate capacitance extraction for nanometer CMOS technology used for RF applications," in *Proc. Int. Meeting Future Electron Devices Tech. Dig.*, 2004, pp. 109–110.

[4] G. A. Brown, "Capacitance characterization in integrated circuit development: The intimate relationship of test structure design, equivalent circuit and measurement methodology," in *Proc. ICMTS*, 2005, pp. 213–217.
[5] D. Rideau, P. Scheer, D. Roy, G. Gouget, M. Minondo, and A. Juge, "Series resistance estimation and $C(V)$ measurements on ultra thin oxide MOS capacitors," in *Proc. ICMTS*, 2003, pp. 191–196.
[6] J. Schmitz, F. N. Cubaynes, R. J. Havens, R. De Kort, A. J. Scholten, J. Tiemeijer, and F. Luuk, "Test structure design considerations for RF-CV measurements on leaky dielectrics," *IEEE Trans. Semicond. Manuf.*, vol. 17, no. 2, pp. 150–154, May 2004.
[7] J. Schmitz, F. N. Cubaynes, R. J. Havens, R. de Kort, A. J. Scholten, and L. F. Tiemeijer, "RF capacitance–voltage characterization of MOSFETs with high leakage dielectrics," *IEEE Electron Device Lett.*, vol. 24, no. 1, pp. 37–39, Jan. 2003.
[8] W. C. Lee and C. Hu, "Modeling CMOS tunneling currents through ultrathin gate oxide due to conduction- and valence-band electron and hole tunneling," *IEEE Trans. Electron Devices*, vol. 48, no. 7, pp. 1366–1373, Jul. 2001.
[9] Y. Wang, K. P. Cheung, R. Choi, G. A. Brown, and B. H. Lee, "Accurate series resistance extraction from capacitor using time domain reflectometry," *IEEE Electron Device Lett.*, vol. 28, no. 4, pp. 279–281, Apr. 2007.

Y. Wang, photograph and biography not available at the time of publication.

Kin P. (Charles) Cheung (M'95–SM'02) received the Ph.D. degree in physical chemistry from the New York University, New York, NY, in 1983.

From 1983 to 1985, he was a Postdoc with Bell Laboratories, during which he pioneered terahertz spectroscopy. From 1985 to 2001, he was a Member of Technical Staff with Bell Laboratories, Murray Hill. From 2001 to 2006, he was an Associate Professor with Rutgers University, Camden, NJ. He is currently a Project Leader with the Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD. He is the author of more than 120 refereed journal and conference proceeding papers, with more than 1200 citations; a book on plasma charging damage; and a book chapter. He is the Editor of three conference proceeding papers. He served on the committee of a number of international conferences and has given tutorials in ten international conferences. Recently, he has been working on technology integration, gate oxide reliability, and device reliability. His research interests include VLSI technology, including processing technologies such as metallization and etching. His current interests include CMOS characterization and reliability, MEMS wafer-level packaging, microwave spectroscopy of organic monolayers, and microplasma arrays.

R. Choi, photograph and biography not available at the time of publication.

B.-H. Lee, photograph and biography not available at the time of publication.