PERGAMON

Solid-State Electronics 46 (2002) 1163–1167

SOLID-STATE ELECTRONICS

www.elsevier.com/locate/sse

The extraction of MOSFET gate capacitance from *S*-parameter measurements

Jiong-Guang Su [a], Shyh-Chih Wong [b], Chun-Yen Chang [a,*], Tiao-Yuan Huang [a]

[a] *Institute of Electronics, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan*

[b] *Research & Development, Taiwan Semiconductor Manufacturing Co., 9 Creation Road 1, Science-Based Industrial Park, Hsinchu 300, Taiwan*

Received 21 November 2000; received in revised form 26 July 2001; accepted 18 January 2002

Abstract

A new gate capacitance extraction method from *S*-parameter measurements is proposed. The distributed nature of MOS transistor and the in-series substrate resistance and in-parallel gate conductance are taken into consideration in the gate capacitance extraction by using high-frequency *S*-parameter measurements. The error due to dissipation factor can be more effectively reduced by this method, compared to the conventional *C–V* measurements. Successful extraction of gate capacitance from test transistors with designed test pads has been demonstrated. © 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

Gate capacitance measurement of MOS transistor is an important technique to monitor the quality and also to determine the electrical thickness of gate dielectric. In addition, its role is also crucial in gaining physical insights into the MOSFET operation, and for developing suitable large- and small-signal models. For deep-submicron devices, due to the short- and narrow-channel effects, the electric field and charge distribution can no longer be treated by one-dimensional analysis. Hence, direct gate capacitance measurement of MOS transistors should be performed to reflect the exact physical features of the transistors themselves. However, existing techniques in direct gate capacitance measurements of MOS transistors are not sufficiently adequate in this regard.

Conventional methods for determining the gate capacitance of MOS transistor are to use the quasi-static or high-frequency capacitance–voltage (*C–V*) measurements. However, even with the so-called "high-frequency" *C–V* measurement, a universally accepted frequency of no more than 1 MHz is usually employed [1]. Worse, test structures with large capacitor area are normally required to minimize possible instrument errors due to the large dissipation factor when the gate capacitance is too small, and/or the dielectric conductance is too large [2]. These large test structures, however, cannot be used to reflect the short- and/or narrow-channel effects existing in small-size real MOS transistors. Instead, two-dimensional and/or three-dimensional device simulators are usually adopted for calculating capacitance of small dimensional devices. This approach, however, can only yield qualitative results due to improper device models and difficulties in accurate parameter calibration. To make things worse, the capacitance measured by these conventional techniques may not faithfully predict the behavior of the MOS transistor when the device is under high-frequency operation (e.g., more than 1 GHz).

Recently, several literatures have reported the capacitance extraction of MOS transistor by using *y*-parameter that was obtained from high-frequency *S*-parameter measurements [3–5]. The capacitance was extracted by using the equivalent small-signal model. However, the capacitance was based on the lump elements of small-signal equivalent circuit, and second-order approximation of frequency was used. Besides, the distributed

* Corresponding author. Tel.: +886-3-517-8083; fax: +886-3-572-1500.

E-mail address: cyc@cc.nctu.edu.tw (C.-Y. Chang).

0038-1101/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.

PII: S0038-1101(02)00022-9

nature of the high-frequency behavior was not taken into account in these methods.

In this paper, we proposed a new method to extract the gate capacitance, using a simple test transistor structure, from the *S*-parameter measurements and transmission line theory. This method is more robust to errors caused by dissipation factor, and can be used to more accurately extract the small gate capacitance in small (i.e., short- and narrow-channel) test transistor without resorting to two-dimensional numerical simulation [1,6].

2. Experiment and measurement

MOS transistors used in this study were fabricated on 6-in. wafers with a 0.35 μm logic baseline technology. The gate oxide thickness was 70 Å. The as-grown polysilicon gate thickness was 0.25 μm. A sidewall spacer with a width of 0.21 μm was used in the self-aligned silicidation of source/drain and gate (i.e., salicidation) with TiN material. The final gate sheet resistance was 2–5 Ω/□ after salicidation.

The configuration of a conventional *C–V* measurement is shown schematically in Fig. 1(a). The source, drain and bulk of an n-channel MOS transistor are tied together and connected to the Hi-end of an HP4284 LCR meter to avoid substrate noise. The gate terminal is connected to the Lo-end of the LCR meter for sensing the AC signal coming out from the gate [7]. A BSIMPro extractor was used as the controller [7]. The magnitude

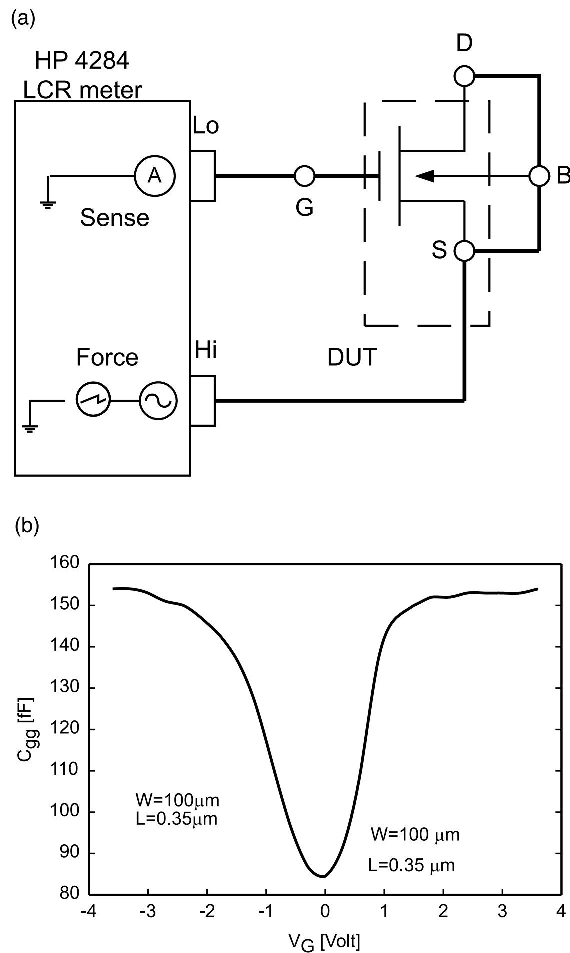

Fig. 1. (a) The setup configuration using a HP4284 LCR meter. (b) Measured gate capacitance from a conventional *C–V* measurement for a device with gate width/length = 100 μm/0.35 μm. The bias range is −3.6 to 3.6 V. Note that a large width is usually required in this method.

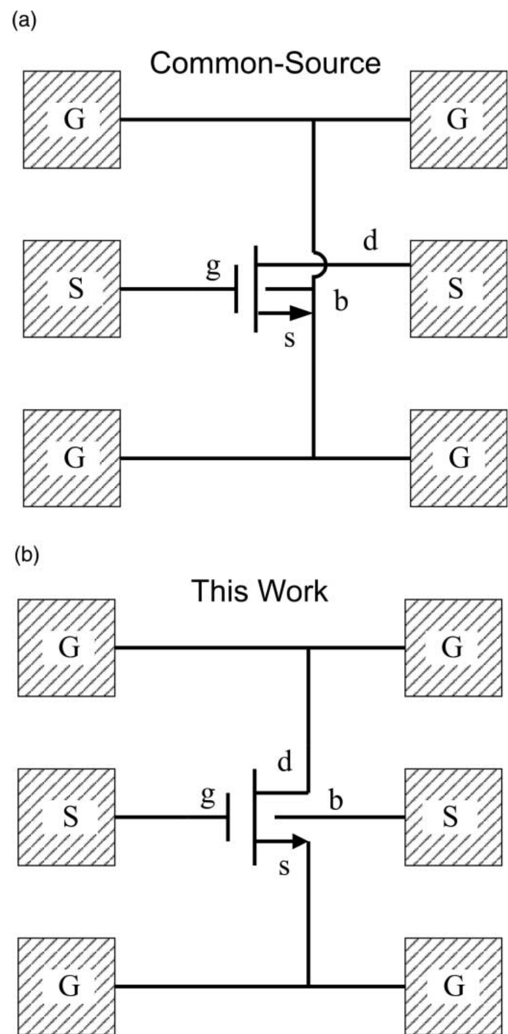

Fig. 2. Test pads configuration for (a) conventional common-source configuration for RF testing, and (b) special G–S–G pad configuration for the proposed method. Note that the bulk is connected to port-2 of the network analyzer in the new method.

of the input AC signal was 50 mV with 100 KHz frequency in the LCR meter setup. Fig.1(b) shows an example of the measured C_{gg} versus the gate bias for an n-channel transistor. The gate bias is swept from −3.6 to +3.6 V. The masked gate width and length of the transistor are 100 and 0.35 μm, respectively. It is worthy to note here that a large-width transistor is normally necessary in this approach to ensure a sufficient C_{gg} value.

The proposed method of extracting C_{gg} from *S*-parameter measurements, on the other hand, allows the use of test transistor with smaller dimensions. In addition, higher frequency (i.e., GHz) is used, which can more faithfully represents the capacitance behavior under high-frequency operation. Fig. 2(a) shows the commonly used common-source configuration of RF testing pattern, while Fig. 2(b) shows the test pads design for the new method. The bulk of the test transistor in the proposed method is connected to port-2, while the source and drain are connected to the ground pads. In addition, the pads are configured in ground–signal–ground of each port for *S*-parameter measurements. The *S*-parameter measurements are executed using an HP8510C network analyzer, and the sweep frequency range is 1.25–10.05 GHz. The controller is a HP ICCAP. The network analyzer performed short-open-load-through (SOLT) calibration in ISS (i.e., impedance-standard-substrate) before *S*-parameter measurements [8]. Fig. 3 shows an extracted gate capacitance by using the proposed method. The gate bias range here is −2.5 to 2.5 V with the *bulk bias equal to zero volts* (i.e., bias of port-2 is 0 V). The masked gate width and length of the test transistor are 20 and 0.35 μm, respectively. All the extracted values by using *S*-parameter measurements were de-embedded by using the "two-step de-embedding" procedure before extraction [9,10]. The proposed gate-capacitance extraction method will be discussed in more detail in the next section.

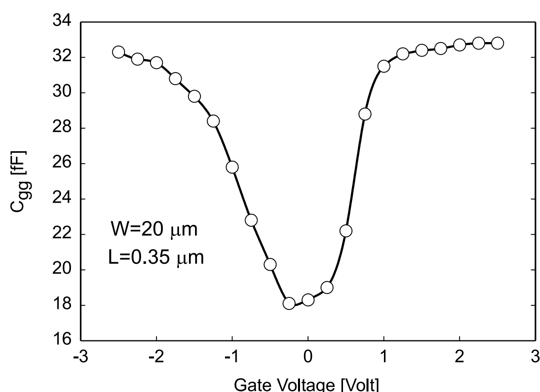

Fig. 3. Gate capacitance versus gate voltage curve extracted by using *S*-parameter measurements. The gate width and length are 20 and 0.35 μm, respectively.

3. Extraction method

Since port-2 of nMOS transistor used in *S*-parameter measurements is the bulk terminal, the RF power flows mainly along the gate width direction and toward the substrate. As shown schematically in Fig. 4, the distributed elements along the gate width direction as seen from the gate contact consist, in per unit length, of in-series gate resistance (R_g), in-series gate inductance (L_s), in-parallel gate capacitance (C_p), in-parallel gate conductance (G_p) and in-series substrate resistance (R_b). Hence, the gate stack can be modeled using a set of distributed elements as shown in Fig. 4. Here we assume that the in-parallel capacitance C_p contains all pertinent capacitance such as C_{ox}, C_{semi}, C_{poly} [2] and *the outer-fringing and overlap capacitance*.

Based on the distributed structure of Fig. 4, the input impedance, Z_{in}, as seen from the gate contact, can be derived by using the transmission line theory [11], and we obtain:

$$Z_{in} = \frac{\rho_{sheet} W}{3L} + \frac{G_p + R_b G_p^2 + \omega^2 R_b C_p^2}{W(G_p^2 + \omega^2 C_p^2)} + j\left[\frac{\omega W L_s}{3} - \frac{\omega C_p}{W(G_p^2 + \omega^2 C_p^2)}\right], \quad (1)$$

where ρ_{sheet} is the sheet resistance of the resistive gate, W and L are the gate width and length, respectively. ω is the operation radius frequency, and R_b, L_s, G_p, C_p are the substrate resistance, series inductance, parallel conductance and capacitance per unit length along the gate width direction, respectively. It is interesting to note that

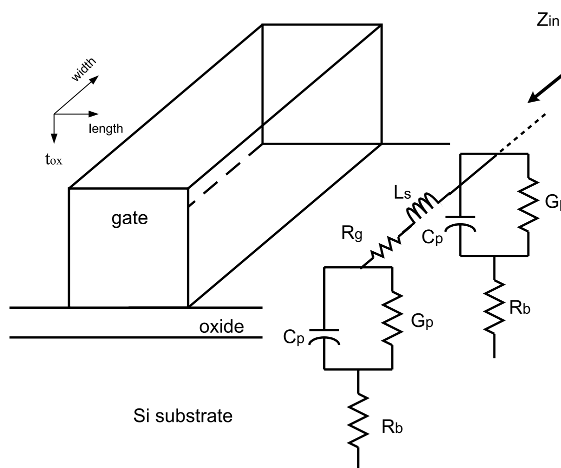

Fig. 4. Schematic showing the distributed nature of MOS gate stack. R_g, R_b, L_s, G_p, C_p are the series gate resistance, substrate resistance, series inductance, parallel conductance and capacitance per unit length along the gate width direction, respectively.

the contribution of the substrate resistance is incorporated into the real part of the input impedance. The term we are most interested is the imaginary part of Eq. (1). By using only the imaginary term, the substrate resistance that causes the measured capacitance degradation in [2] can be excluded in our capacitance extraction. The first term in the imaginary part of Eq. (1) can be ignored in a few GHz frequency range, because of the small value of L_s (in the order of pico-H/μm range) [5]. Then we can express the imaginary part of input impedance in a simplified form:

$$\mathrm{Im}(Z_{\mathrm{in}}) \cong -\frac{\omega C_{\mathrm{p}}}{W(G_{\mathrm{p}}^2 + \omega^2 C_{\mathrm{p}}^2)} = -\frac{1}{W(\omega C_{\mathrm{p}})}\frac{1}{1 + G_{\mathrm{p}}^2/\omega^2 C_{\mathrm{p}}^2} \approx -\frac{1}{W(\omega C_{\mathrm{p}})}. \quad (2)$$

Eq. (2) neglects the term $(G_{\mathrm{p}}/\omega C_{\mathrm{p}})^2$, which will be discussed in the next section. By taking the logarithm on both sides of Eq. (2), we obtain the following equation:

$$\log[-\mathrm{Im}(Z_{\mathrm{in}})] = -\log(\omega W C_{\mathrm{p}}) = -\log f - \log(2\pi C_{\mathrm{gg}}). \quad (3)$$

Here ω is equal to $2\pi f$ and WC_{p} is equal to C_{gg}, which represents the capacitance seen from the gate contact. Eq. (3) is a linear function with the slope equal to -1, and its interception equals $\log(2\pi C_{\mathrm{gg}})$, if the variables are $Y = \log[-\mathrm{Im}(Z_{\mathrm{in}})]$ and $X = \log f$, respectively. By using the relationship shown in Eq. (3), the effective capacitance at the gate contact can be determined from the interception.

4. Results and discussions

Fig. 5 shows the measured $\log[-\mathrm{Im}(Z_{\mathrm{in}})]$ versus $\log f$ with different gate biases of -1, 0 and 2.5 V. The device under test (DUT) is an n-type MOS transistor with the gate width and length equal to 20 and 0.35 μm, respectively. The input impedance can be obtained from the measured S-parameters of DUT. It shows that the slope approaches -1, and all the curves are in parallel. This result is consistent with Eq. (3). Hence, the interception can be determined by extrapolation of the lines and the capacitance C_{gg} can be extracted. The extracted C_{gg} by using our method is shown previously in Fig. 3. The range of gate bias is from -2.5 to 2.5 V, with 0.25 V increments. Despite the high-frequency operation, the n^+ source/drain serves to supply abundant channel charge that follows the applied gate signal, as the source/drain of the DUT is connected to the AC ground of the network analyzer [12]. Hence the capacitance increases with increasing gate bias in strong inversion region. This is regardless of the fact that the thermal recombination–generation rate of carriers in the bulk depletion region cannot follow the fast excitation of the input signal, and thus does not contribute to the inversion layer charges.

As mentioned previously, the dissipation factor causes instrument error in LCR meter. The error increases when the gate oxide conductance is larger or the capacitance is smaller [1,2]. Eqs. (4) and (5) represent the dissipation factor and the corresponding instrument error, respectively.

$$D = \frac{G}{\omega C}, \quad (4)$$

$$\%\mathrm{error} = 0.1\sqrt{1 + D^2}, \quad (5)$$

where G and C are the effective in-parallel conductance and capacitance of the measured capacitor, respectively.

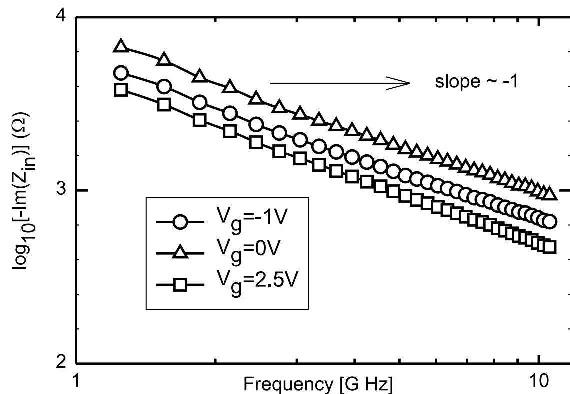

Fig. 5. Imaginary part of input impedance (Z_{in}) versus frequency with different gate biases of -1, 0 and 2.5 V.

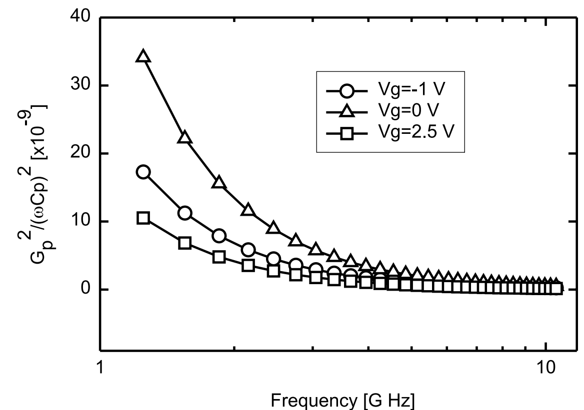

Fig. 6. The error term as a function of operation frequency with different gate biases of -1, 0 and 2.5 V.

Table 1
Five-point on-wafer oxide thickness by using large-area capacitance measurements (first row), their equivalent small-geometry oxide capacitance (second row) and oxide capacitance extracted from *S*-parameter measurements (third row)

Position	Center	Up	Down	Left	Right
t_{ox} (nm)	7.3353	7.352	7.4075	7.3983	7.393
$C_{ox} = (\varepsilon_{ox}/t_{ox})WL$ (fF)	32.937	32.862	32.616	32.657	32.680
C_{gg} in accumulation region (fF)	32.30				

A useful way to reduce the dissipation factor is to increase the operation frequency of AC signal, and hence, reduce the dissipation factor. However, this is difficult to achieve in an LCR meter, and the distribution nature should be taken into account. One can compare the neglected term $(G_p/\omega C_p)^2$ in Eq. (2) with the dissipation factor in Eq. (4). The neglected term in the new method is proportional to D^2, meaning that our method is more effective in reducing the possible errors as long as $\omega C_p > G_p$. This is especially true in GHz operation frequency. Fig. 6 shows the measured neglected term versus frequency. It can be seen that this term indeed approaches zero with increased operation frequency. The maximum value occurs when the applied gate bias equals 0 V. Besides, all the observed values in this figure are less than 10^{-7}, indicating that the effect due to this term can indeed be neglected when we calculate the capacitance seen from the gate contact in our proposed method.

Table 1 shows the five-point oxide thickness on the same wafer by using the large-area capacitance measurement. The oxide thickness shown in Table 1 was determined by using the conventional *C–V* measurement. The calculated oxide capacitance for a device with gate width/length equal to 20 μm/0.35 μm is shown in the second row of the table. The third row is the gate capacitance obtained from our new method. It can be seen clearly that the extracted results from our method are satisfactorily accurate.

5. Conclusion

A method to extract gate capacitance of MOSFET by using *S*-parameter is proposed. This method considers the effects of the in-parallel substrate resistance and gate conductance, and also considers the distributed nature of the MOS transistor. In addition, the proposed method is inherently more robust to errors caused by the dissipation factor, and therefore it is more suitable for measuring small-geometry capacitance than conventional *C–V* measurements.

Acknowledgement

This work was supported in part by National Science Council, Taiwan, ROC, under contract NSC90-2215-B009-072.

References

[1] Arora N. MOSFET models for VLSI circuit simulation. Wien (NY): Springer; 1993.
[2] Henson WK, Ahmend KZ, Vogel EM, Hauser JR, Wortman JJ, Venables RD, et al. Estimating oxide thickness of tunneling oxides down to 1.4 nm using conventional capacitance–voltage measurements on MOS capacitors. IEEE Electron Dev Lett 1999;20:179.
[3] Dambrine G, Cappy A, Heliodore F, Playez E. A new method for determining the FET small-signal equivalent circuit. IEEE Trans Microwave Theory Tech 1988;36: 1151.
[4] Lovelace D, Costa J, Camilleri N. Extracting small-signal model parameters of silicon MOSFET transistors. IEEE Dig MTT-S 1994:865.
[5] Lee S, Yu HK, Kim CS, Ko JG, Nam KS. A novel approach to extracting small-signal model parameters of silicon MOSFET's. IEEE Microwave Guided Wave Lett 1997;7(3):75;
Stathis JH, DiMaria DJ. In: Reliability projection for ultra-thin oxides at low voltage. IEDM Tech; 1998. p. 167.
[6] BSIM3v3 Manual, 1995. p. 4–5.
[7] BSIMPro Manual. 1997.
[8] Kolding TE. On-wafer calibration techniques for gigahertz CMOS measurements. In: International Conference on MTS, 1999 March, vol. 12.
[9] Koolen MCAM, Geelen JAM, Versleijen MPJG. An improved de-embedding technique for on-wafer high-frequency characterization. In: IEEE Proceedings of Bipolar Circuit and Technology Meeting, 1991. p. 188.
[10] Jen SHM, Enz CC, Pehlke DR, Schröter M, Sheu BJ. Accurate modeling and parameter extraction for MOS transistors valid up to 10 GHz. IEEE Trans Electron Dev 1999;46(11):2217.
[11] Kolding TE. Calculation of MOSFET input gate impedance. Technology Report of RISC, 1998.
[12] Chan M, Hui KY, Hu C, Ko PK. A robust and physical BSIM3 non-quasi-static transient and AC small-signal model for circuit simulation. IEEE Trans Electron Dev 1998;45(4):834.