A New Three-Dimensional Capacitor Model for Accurate Simulation of Parasitic Capacitances in Nanoscale MOSFETs

Jyh-Chyurn Guo, *Senior Member, IEEE*, and Chih-Ting Yeh

***Abstract*—A new 3-D gate capacitor model is developed to accurately calculate the parasitic capacitances of nanoscale CMOS devices. The dependences on gate length and width, gate electrode and dielectric thicknesses, gate-to-contact spacing, and contact dimension and geometry are fully incorporated in this model. The accuracy is certified by an excellent match with the 3-D interconnection simulation results for three structures with strip, square, and circular contacts. The features of being free from fitting parameters and proven accuracy over various geometries make this model useful for nanoscale MOSFET parasitic capacitance simulation and analysis. Furthermore, the developed capacitor model in the form of multidimensional integral can easily be deployed in general circuit simulators. This model predicts that the parasitic capacitance C_{of} dominates around 25% of the intrinsic gate capacitance (C_{gint}) in 80-nm MOSFETs and that the near nonscalability with gate length brings the weighting factor $C_{\mathrm{of}}/C_{\mathrm{gint}}$ above 30%/40%/60% in 65-/45-/32-nm devices. It actually exceeds the limitation defined by the most updated ITRS and reveals itself as a show-stopper in high-speed and high-frequency circuit design.**

***Index Terms*—MOSFET, nanoscale, parasitic capacitance, 3-D capacitor model.**

I. Introduction

THE advancement of CMOS technology to sub-100-nm scale has driven intrinsic gate delay below 10 ps ($1p = 10^{-12}$) and cutoff frequency f_T well above 100 GHz [1]–[3]. However, the parasitic capacitances, which are generally not scalable, may dominate a significant portion of MOSFET's capacitances and limit further speed improvement in continuous scaling. The impact of parasitic capacitances on circuit performance has been an important subject of concern not only for high-speed-driven logic circuits but also for analog and RF integrated circuits (RF ICs) in which the gain bandwidth product, f_T and $f_{\max}$, noise figures, and linearity are all strongly influenced by the parasitic capacitances [3]–[5]. Over the past two decades, many research works have been conducted on parasitic capacitance model development [6]–[10]. Kamchouchi and Zaky [6] adopted the Schwartz–Christofel transformation method and numerical iterations to derive a semianalytical model. However, the required numerical iterations hinder its application in circuit simulation. Elmasry [7] assumed that the incremental sidewall capacitance could be described as the parallel plane capacitance between the sidewall and the substrates and derived a simple sidewall capacitance model. Unfortunately, the assumption proposed by Elmasry, such as electric field flux in parallel with the sidewall surface, is not justified. In fact, the electric field flux impinges perpendicular to the surface. Shrivastava and Fitzpatrick [8] focused an effort by using a conformal mapping method to fix the problem appearing in Elmasry's approach, and they published an improved sidewall capacitance model that ensures electric field lines perpendicular to the surface. However, the assumption of constant potential near the gate electrode between the silicon substrate and gate electrode's bottom is prone to a significant error. Then, Suzuki [9] presented a modified Shrivastava model with corrected boundary conditions.

Even though many models have been developed, almost all of them lack an extensive verification over real structures in miniaturized MOSFETs. Most importantly, an analytical model in a mathematically closed form without numerical iterations can facilitate circuit simulation and design. As a result, increasing effort has been switched toward this approach. Mohapatra *et al.* [10] followed the conformal mapping method [1], [3], [4] to derive an analytical model in which the presence of source/drain electrodes was considered. However, a number of correction factors were introduced to fit the Monte Carlo simulation results, and the physical meaning was left as an open question. Without using the correction factors, Mohapatra's model suffered significant errors as large as 20%–30% in comparison with the Monte Carlo simulation.

In this paper, a new 3-D capacitor model, accompanied with an extensive verification, will be presented. The derived mathematical expressions with the form of multidimensional integral are relatively simple compared to the conformal mapping method. The favorable features of being free from numerical iterations and elimination of fitting parameters or correction factors enable an easy implementation of this new model in circuit simulators. Furthermore, the consideration of source/drain electrodes, as well as contact plugs, and the added parasitic capacitances arising from their coupling with the gate are among the more important features in our model for assuring the accuracy and prediction capability. The model accuracy

Manuscript received September 18, 2008; revised April 27, 2009. First published June 23, 2009; current version published July 22, 2009. This work was supported in part by the National Science Council under Grant NSC 97-2221-E009-175. The review of this paper was arranged by Editor J. Woo.

J.-C. Guo is with the Department of Electronics Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: jcguo@mail.nctu.edu.tw).

C.-T. Yeh is with the SoC Technology Center, Design Automation Technology Division, Department of Circuit Design, Industrial Technology Research Institute, Hsinchu 31040, Taiwan.

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.2009.2022679

0018-9383/$25.00 © 2009 IEEE

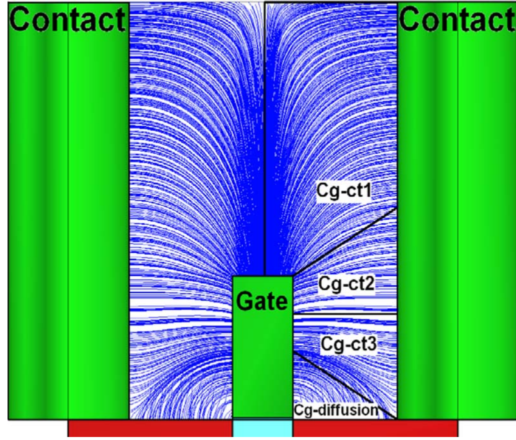


Fig. 1. Device structure with gate, contact plug, and substrate. Field line contours are partitioned into four regions with capacitance components defined as $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$.

and scalability will be extensively verified over 3-D device structures and geometries in a comprehensive spectrum of gate lengths and widths, gate electrode and dielectric thicknesses, gate-to-contact space, and contact geometries and dimensions.

II. Device Structures and 3-D Capacitive Coupling Analysis

A 3-D interconnection simulator, namely, Raphael was employed to investigate the electric field line distribution responsible for the capacitive coupling between any two conducting electrodes. Three structures with different contact geometries, such as strip, square, and circular contacts, were implemented in this paper. Fig. 1 shows the front view of the device structure and the simulated field line contour. Note that the field lines started from or terminated at the conducting plates like the gate, as well as the source/drain diffusion region (Diff), and contact plugs (CT) always keep perpendicular to the surface of the conducting plates. The most important feature demonstrated by the 3-D interconnection simulation is a specific division line revealed between the gate and the contact plug. This division line is exactly parallel to the substrate surface and serves as the boundary for separating two symmetric trapezoid regions ($C_{\mathrm{g-ct2}}$ and $C_{\mathrm{g-ct3}}$), with the field lines originating from the gate sidewall and terminating at the CT, but with opposite curvature directions, i.e., one upward and another one downward. As the trapezoid regions are determined, the area below the horizontal division line is divided into two regions, namely, $C_{\mathrm{g-ct3}}$ and $C_{\mathrm{g-diff}}$, along a known diagonal line, in which the field lines start from the same electrode, i.e., the gate, but terminate at different ones, i.e., CT and Diff. Similarly, the area above the horizontal division line is separated into two regions ($C_{\mathrm{g-ct1}}$ and $C_{\mathrm{g-ct2}}$) along a diagonal line, which is a mirror reflection of the previous one.

Through the analysis, four regions of field lines are partitioned, as shown in Fig. 1, to account for four components of capacitance, namely, $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$. The physical property of each component is defined as follows. $C_{\mathrm{g-ct1}}$ represents the gate-top-to-CT-coupling capacitance. $C_{\mathrm{g-ct2}}$ and $C_{\mathrm{g-ct3}}$ are of equal amount but separated by the division line, and both are contributed from the gate-sidewall-to-CT coupling. $C_{\mathrm{g-diff}}$ is the gate-sidewall-to-Diff-coupling capacitance. This four-region capacitance partitioning is adequate to the structure with strip contacts having a uniform extension along the gate width. As for more common structures with square or circular contacts in an array, the coupling between every two CTs along the gate width should be taken into account, and two more four-region partitions are required to simulate the capacitances under a more complicated condition. Fig. 2 shows the structure with circular contacts and the corresponding top view of the potential line contour. The potential and field line contours from Raphael simulation reveal the capacitive coupling not only through the space between the gate sidewall and CT sidewall but also through the space between every two adjacent CTs in case of circular or square contacts. The field line analysis provides us a useful guideline to develop a multidimensional capacitor model that is valid for structures with different contact schemes.

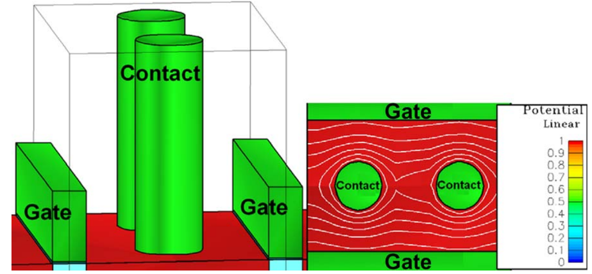


Fig. 2. Three-dimensional device structure with circular contacts and the top view of the potential line contour calculated by 3-D Raphael simulation.

III. 3-D Capacitor Model Development

Fig. 3 shows the full structure with three four-region capacitance partitions corresponding to the 3-D capacitive coupling. Herein, the central domain defined by the contact dimension along the gate width (denoted as 123) is equivalent to the structure with strip contact and represents the gate-to-CT-direct-coupling capacitance. As for the two ends along the gate width (denoted as $1''2''3''$) with a dimension by subtracting the contact size from the total width, a different four-region partition with the division line at a different location was conducted to simulate the gate-to-CT coupling through the space between two adjacent CTs, as shown in Fig. 2. The locations of division lines along the y-axis, i.e., the height of the gate and CT, are defined as h_1 corresponding to the domain (123) in the middle of the gate width and h_2 for the domain ($1''2''3''$) at the two ends, as shown in Fig. 3. Through an analysis on the field line distribution and device structures, h_1 can be calculated as a function of gate thickness (T_g), gate dielectric thickness (T_{ox}), and gate-to-CT space (L_{gct}). As for h_2 corresponding to the domain ($1''2''3''$), a more complicated expression is derived, incorporating two more parameters, namely, contact width (L_{ct}) and gate width (W_g). The mathematical formulas for calculating h_1 and h_2 will be described in the following for different contact schemes, respectively.

According to the four-region capacitance partitioning scheme shown in Fig. 1, four components of parasitic

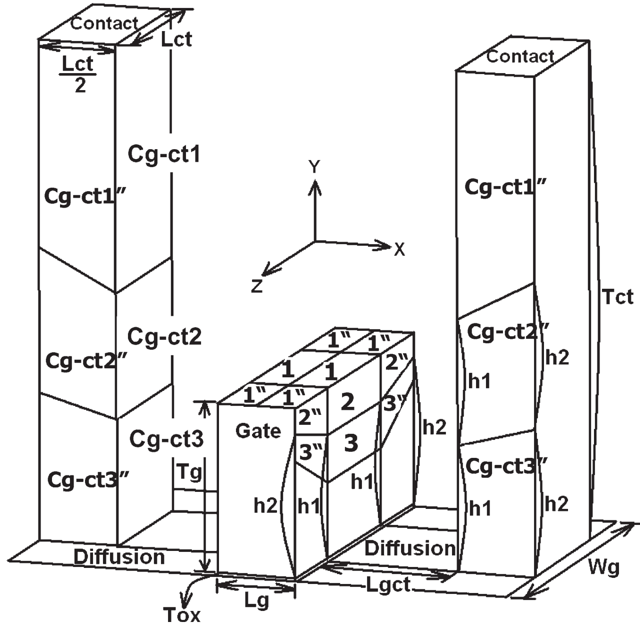


Fig. 3. Three-dimensional structure of a MOSFET and the corresponding four-region capacitance partitioning for 3-D capacitor model development. The domain (123) in the central portion represents $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$ due to capacitive coupling through the space between the gate and CT sidewalls. The domain ($1''2''3''$) at the two ends incorporates $C''_{\mathrm{g-ct1}}$, $C''_{\mathrm{g-ct2}}$, $C''_{\mathrm{g-ct3}}$, and $C''_{\mathrm{g-diff}}$ contributed from the capacitive coupling between the CT and gate through the space between two adjacent CTs.

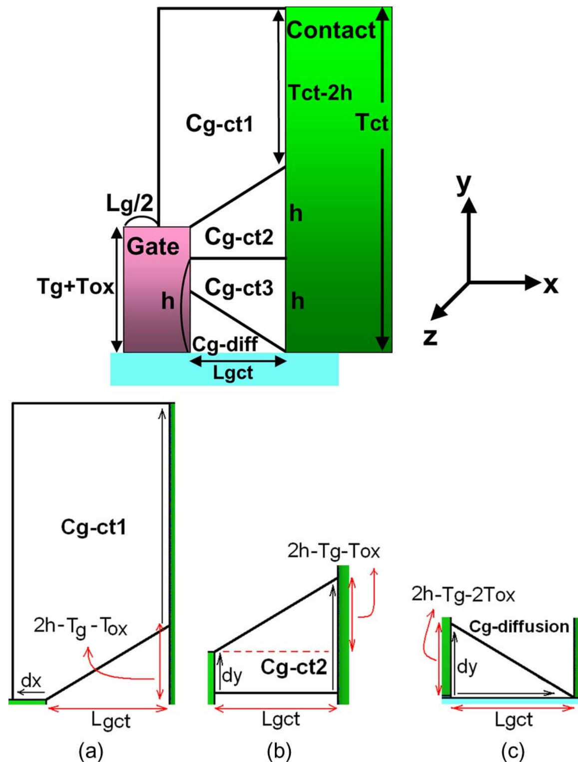


Fig. 4. Four-region partitioning for parasitic capacitances, such as (a) $C_{\mathrm{g-ct1}}$, (b) $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and (c) $C_{\mathrm{g-diff}}$. All the dimensions required for calculating capacitances are specified in the full structure and each capacitance component. L_g: gate length, L_{gct}: gate-to-CT space, T_g: gate electrode thickness, T_{ox}: gate dielectric thickness, T_{ct}: CT plug thickness, and h: position of the division line from the gate's bottom.

capacitance, such as $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$, can be drawn with dimensions specified for capacitance calculation, as shown in Fig. 4. Each component can be treated as a trapezoid-parallel-plate capacitance, in which two conducting plates are not identical to each other. In the following, the 3-D capacitor models for calculating $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}} = C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$ will be derived for strip, square, and circular contacts, respectively.

A. Trapezoid-Parallel-Plate Capacitance Model

Fig. 5(a) shows a 3-D structure of a typical trapezoid-parallel-plate capacitor that is applicable to strip contact, in which the top and bottom plate lengths are specified as L_1 and L_2, the plate width is fixed at W, and the plate-to-plate distance is given by d. Then, the capacitance can be calculated through an integral formula in (1), which is a transformation from a series-capacitance sum rule under infinitively small segmentation along the distance between two plates. As for square or circular contacts with capacitive coupling at the two ends along the width denoted as ($1''2''3''$) in Fig. 3, the simple trapezoid-parallel-plate capacitor limited to a fixed width for two parallel plates, as shown in Fig. 5(a), is no longer suitable. In fact, the trapezoid parallel plate becomes more complicated with different lengths and widths shown in Fig. 5(b), and the capacitance is computed through an integral with the variable along the plate-to-plate distance appearing in both length and width. The formulation and result are expressed in

$$C = \frac{1}{\int \delta\left(\frac{1}{C_i}\right)} = \left\{\int_0^d \left[\frac{1}{k\varepsilon_0 \cdot \left(L_1 + \frac{L_2 - L_1}{d}\cdot x\right)\cdot W}\right] dx\right\}^{-1} = \frac{k\varepsilon_0 W \cdot \frac{L_1 - L_2}{\ln(L_1/L_2)}}{d} \tag{1}$$

$$C = \frac{1}{\int \delta\left(\frac{1}{C_i}\right)} = \left\{\int_0^d \left[\frac{1}{k\varepsilon_0 \cdot \left(L_1 + \frac{L_2 - L_1}{d}\cdot x\right)\cdot \left(W_1 + \frac{W_2 - W_1}{d}\cdot x\right)}\right] dx\right\}^{-1} = \frac{k\varepsilon_0 \cdot \frac{L_2 W_1 - L_1 W_2}{\ln\left(\frac{L_2 W_1}{L_1 W_2}\right)}}{d}. \tag{2}$$

B. Capacitance Model for MOSFET With Strip Contacts

For MOSFETs with strip contacts, the 3-D parasitic capacitances are simply contributed from the domain (123) in the central portion of the 3-D full structure in Fig. 3, which correspond to four components of capacitance defined as $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$. To calculate $C_{\mathrm{g-ct1}}$, $C_{\mathrm{g-ct2}}$, $C_{\mathrm{g-ct3}}$, and $C_{\mathrm{g-diff}}$, the first key parameter to be determined is the location

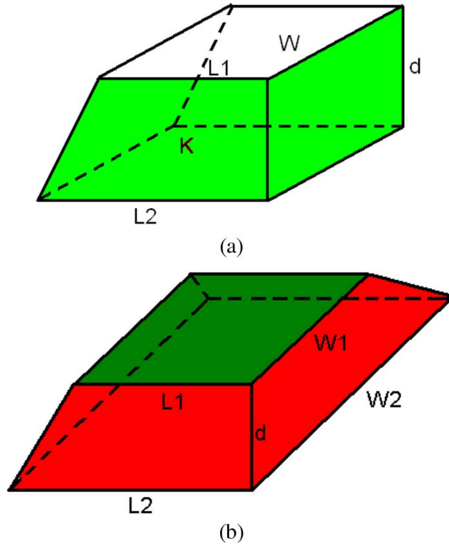


Fig. 5. Three-dimensional structures of trapezoid-parallel-plate capacitors. (a) Simple structure with different lengths L_1 and L_2 but the same width for two conducting plates. (b) More complicated structure with different lengths L_1 and L_2 and widths W_1 and W_2 for two plates in parallel.

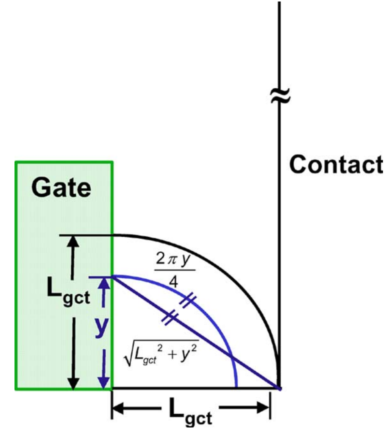


Fig. 6. New method for determining the position of division line h along the gate height. In this improved method, h is defined as the average of L_{gct} and y. The parameter y is determined under the condition that $(2\pi y/4) = \sqrt{L_{\text{gct}}^2 + y^2}$.

of division line along the height of the gate, which is specified as h in Fig. 4. In Mohapatra's work [10], h was set equal to L_{gct} under a simple assumption that all the field lines that started from the gate in the region below h would be terminated at Diff and contributed to $C_{\text{g-diff}}$. However, the assumption cannot be justified through a careful analysis on the field line distribution from Raphael simulation in Fig. 1. The field lines below h are actually divided into two regions, i.e., one terminated at Diff and another one at CT. It means that h and $C_{\text{g-diff}}$ are overestimated in this way. As shown in Fig. 6, a modified method to calculate h is assuming an equal distance in two paths, i.e., one along the quarter circle and another one along the diagonal line connecting the points at h on the gate sidewall and the bottom of the CT sidewall. Then, h can be derived from (3). However, this approach actually underestimates h. Following the argument, a new method is developed, taking the average of the aforementioned two methods for deriving this key parameter h corresponding to typical device dimensions.

To realize a scalable model, h should be implemented as a geometry-dependent function incorporating all relevant geometry parameters, such as T_g, T_{ox}, and L_{gct} for the case of strip contact. Considering two extreme conditions, **h** can be derived as follows. For infinitively large L_{gct}, defined as $L_{\text{gct}}/T_g \gg 1$, the gate-sidewall-to-CT coupling is negligible, and all of the capacitances are contributed from the gate-sidewall-to-Diff region. Thus, the horizontal division line moves to the top of the gate, i.e., $h = T_g$. One the other hand, for infinitely small L_{gct}, denoted as $L_{\text{gct}}/T_g \ll 1$, the gate-sidewall-to-Diff coupling vanishes, and all of the capacitances are contributed from the gate sidewall to CT. Under this condition, h moves to the middle of the gate, i.e., $h = T_g/2$. To observe the aforementioned boundary conditions, h, namely, h_1, for the strip contact is modeled as a function of T_g, T_{ox}, and L_{gct} given by (4), shown at the bottom of the next page. Note that α is a constant extracted from (4) corresponding to the typical geometry parameters (T_g, T_{ox}, and L_{gct}), in which h_1 has been calculated from the average of L_{gct} and y in (3). Then, according to the four-region capacitance notation shown in Fig. 3, the geometrical structure and dimensions for each component in Fig. 4, and the basic formula derived for a trapezoid-parallel-plate capacitance in (1), $C_{\text{g-ct1}}$, $C_{\text{g-ct2}} = C_{\text{g-ct3}}$, and $C_{\text{g-diff}}$ can be calculated by (5)–(7), shown at the bottom of the next page. Herein, the integral is performed in one dimension corresponding to a simple trapezoid box with fixed sizes in the other two dimensions, e.g., $x = L_{\text{gct}}$ and $z = L_{\text{ct}}$ for $C_{\text{g-ct2}} = C_{\text{g-ct3}}$. The integral is performed along the x-axis (i.e., the direction of gate length) for $C_{\text{g-ct1}}$ and along the y-axis (the direction of gate and CT thickness) for the other three components, i.e., $C_{\text{g-ct2}}$, $C_{\text{g-ct3}}$, and $C_{\text{g-diff}}$.

The geometrical parameters in (3)–(7) are defined as follows: L_g—gate length, L_{gct}—gate-to-CT space, T_g—gate electrode thickness, T_{ox}—gate dielectric thickness, T_{ct}—contact plug thickness (height), and h—division line's position from the gate's bottom.

C. Capacitance Model for MOSFET With Square and Circular Contacts

As for MOSFET with square or circular contacts, the 3-D parasitic capacitances are contributed from both domain (123) in the central portion and domain ($1''2''3''$) at the two ends, referring to Fig. 3. Corresponding to the domain ($1''2''3''$) appearing at the two ends, four components of capacitances, namely, $C''_{\text{g-ct1}}$, $C''_{\text{g-ct2}}$, $C''_{\text{g-ct3}}$, and $C''_{\text{g-diff}}$, are formulated in (8)–(10), shown at the bottom of the next page.

Note that the model equations are derived following the basic formula in (2), an integral in two dimensions corresponding to a more complicated trapezoid box, as shown in Fig. 5(b) with a fixed size in only one dimension. The division line location h is actually a variable of z, which is given by

$$h(z) = h_1 + \frac{h_2 - h_1}{(W_g - L_{ct})/2} \cdot z$$

$$h(z = 0) = h_1$$

$$h\left(z=\frac{W_g-L_{ct}}{2}\right)=h_2. \tag{11}$$

Herein, $h(z)$ is formulated as a linear composition of h_1 in (12) and of h_2 in (13), which represent $h(z)$ at two boundary conditions of $z=0$ and $z=(W_g-L_{\mathrm{ct}})/2$, respectively. Note that the model equations for a structure adopting circular contacts are derived with formulas similar to those having square contacts but with a major change on h_1 and h_2 by introducing a geometry-dependent parameter β in (12) and (13). Herein, β is calculated by (15) as a function of the gate-to-CT space L_{gct} and CT width L_{ct}

$$h_1=\frac{T_g}{2}\left[1+\exp\left(-\alpha\cdot\beta\frac{T_g}{L_{\mathrm{gct}}}\right)\right]+T_{\mathrm{ox}} \tag{12}$$

$$h_2=\frac{T_g}{2}\left[1+\exp\left(-\alpha\cdot\beta\frac{T_g}{L_{\mathrm{gct}}}\frac{L_{\mathrm{ct}}}{W_g}\right)\right]+T_{\mathrm{ox}} \tag{13}$$

$$\beta=1,\quad \text{for square contacts} \tag{14}$$

$$\beta=\sqrt{1-\left(\frac{L_{\mathrm{ct}}/2}{L_{\mathrm{gct}}+L_{\mathrm{ct}}/2}\right)^2},\quad \text{for circular contacts} \tag{15}$$

$$\frac{2\pi y}{4}=\sqrt{L_{\mathrm{gct}}^2+y^2}|_{y=h} \tag{3}$$

$$h_1=\frac{T_g}{2}\left[1+\exp\left(-\alpha\frac{T_g}{L_{\mathrm{gct}}}\right)\right]+T_{\mathrm{ox}},\quad \alpha=0.686188 \tag{4}$$

$$C_{\mathrm{g-ct1}}=2\int_0^{\frac{L_g}{2}}\left\{\frac{k\varepsilon_0\frac{1-\frac{T_{\mathrm{ct}}-2h_1}{L_g/2}}{\ell n\left(\frac{L_g/2}{T_{\mathrm{ct}}-2h}\right)}}{\sqrt{(L_{\mathrm{gct}}+h_1)^2+\left[(2h_1-T_g-T_{\mathrm{ox}})+\frac{x\cdot(T_{\mathrm{ct}}-2h_1)}{L_g/2}\right]^2}}\right\}dx \tag{5}$$

$$C_{\mathrm{g-ct2,3}}=2\int_0^{T_g+T_{\mathrm{ox}}-h_1}\left\{\frac{k\varepsilon_0\frac{1-\frac{h_1}{T_g+T_{\mathrm{ox}}-h_1}}{\ell n\left(\frac{T_g+T_{\mathrm{ox}}-h_1}{h_1}\right)}}{\sqrt{(L_{\mathrm{gct}})^2+\left[y\cdot\left(\frac{h_1}{T_g+T_{\mathrm{ox}}-h_1}-1\right)\right]^2}}\right\}dy \tag{6}$$

$$C_{\mathrm{g-diff}}=2\int_0^{2h_1-T_g-2T_{\mathrm{ox}}}\left\{\frac{k\varepsilon_0\frac{1-\frac{L_{\mathrm{gct}}}{2h_1-T_g-2T_{\mathrm{ox}}}}{\ell n\left(\frac{2h_1-T_g-2T_{\mathrm{ox}}}{L_{\mathrm{gct}}}\right)}}{\sqrt{\left[y\cdot\left(\frac{L_{\mathrm{gct}}}{2h_1-T_g-2T_{\mathrm{ox}}}\right)\right]^2+(T_{\mathrm{ox}}+y)^2}}\right\}dy \tag{7}$$

$$C''_{\mathrm{g-ct1}}=4\int_0^{\frac{W_g-L_{\mathrm{ct}}}{2}}\int_0^{\frac{L_g}{2}}\left\{\frac{k\varepsilon_0\left[\frac{\frac{T_{\mathrm{ct}}-2h(z)}{L_g/2}-\frac{L_{\mathrm{ct}}/2}{(W_g-L_{\mathrm{ct}})/2}}{\ell n\left(\frac{(T_{\mathrm{ct}}-2h(z))\cdot(W_g-L_{\mathrm{ct}})/2}{L_g L_{\mathrm{ct}}/2}\right)}\right]}{\sqrt{\left(L_{\mathrm{gct}}+\frac{L_{\mathrm{ct}}/2}{(W_g-L_{\mathrm{ct}})/2}\cdot z+x\right)^2+\left[(2h(z)-T_g-T_{\mathrm{ox}})+\frac{T_{\mathrm{ct}}-2h(z)}{L_g/2}\cdot x\right]^2+z^2}}\right\}dxdz \tag{8}$$

$$C''_{\mathrm{g-ct2,3}}=4\int_0^{\frac{W_g-L_{\mathrm{ct}}}{2}}\int_0^{T_g+T_{\mathrm{ox}}-h(z)}\left\{\frac{k\varepsilon_0\left[\frac{\frac{h(z)}{T_g+T_{\mathrm{ox}}-h(z)}-\frac{L_{\mathrm{ct}}/2}{(W_g-L_{\mathrm{ct}})/2}}{\ell n\left(\frac{h(z)\cdot(W_g-L_{\mathrm{ct}})/2}{(T_g+T_{\mathrm{ox}}-h(z))L_{\mathrm{ct}}/2}\right)}\right]}{\sqrt{\left(L_{\mathrm{gct}}+\frac{L_{\mathrm{ct}}/2}{(W_g-L_{\mathrm{ct}})/2}\cdot z\right)^2+\left[\left(\frac{h(z)}{T_g+T_{\mathrm{ox}}-h(z)}-1\right)\cdot y\right]^2+z^2}}\right\}dydz \tag{9}$$

$$C''_{\mathrm{g-diff}}=4\int_0^{\frac{W_g-L_{\mathrm{ct}}}{2}}\int_0^{2h(z)-T_g-2T_{\mathrm{ox}}}\left\{\frac{k\varepsilon_0\left[\frac{1-\frac{L_{\mathrm{gct}}+L_{\mathrm{ct}}/2}{2h(z)-T_g-2T_{\mathrm{ox}}}}{\ell n\left(\frac{2h(z)-T_g-2T_{\mathrm{ox}}}{L_{\mathrm{gct}}+L_{\mathrm{ct}}/2}\right)}\right]}{\sqrt{\left(\frac{L_{\mathrm{gct}}+L_{\mathrm{ct}}/2}{2h(z)-T_g-2T_{\mathrm{ox}}}\cdot y\right)^2+(T_{\mathrm{ox}}+y)^2}}\right\}dydz \tag{10}$$

where

- W_g gate width;
- L_g gate length;
- $L_{\rm gct}$ gate-to-CT space;
- $L_{\rm ct}$ CT width;
- T_g gate electrode thickness;
- $T_{\rm ox}$ gate dielectric thickness;
- $T_{\rm ct}$ contact plug thickness (height);
- h_1 division line's position from the gate's bottom for domain (1, 2, 3);
- h_2 division line's position from the gate's bottom for domain $(1''2''3'')$.

The mathematical formulas for deriving β has been performed based on a geometrical analysis on circular contact and described in the Appendix.

IV. Results and Discussion

The developed 3-D capacitor models were applied to calculate the parasitic capacitances for MOSFETs with different contact schemes and various geometrical parameters. The typical one follows the 130-nm high-speed CMOS design rule and device/process parameters, such as $L_g = 80$ nm, $T_g = 170$ nm, $T_{\rm ox} = 2.4$ nm, $L_{\rm gct} = 135$ nm, and $L_{\rm ct} = 160$ nm. A wide range of variations around the typical rule has been conducted to investigate the scaling effect. The accuracy of the 3-D capacitor models was verified through an extensive comparison with 3-D interconnection simulation results carried out by Raphael.

The calculated capacitances are presented as the gate-to-CT capacitance $C_{\rm G-CT}(= C_{\rm g-ct1} + C_{\rm g-ct2} + C_{\rm g-ct3})$, gate-to-source/drain-diffusion capacitance, $C_{\rm G-Diff} = C_{\rm g-diff}$, and the total parasitic capacitance that is equal to the sum of these two terms, which is denoted as $C_{\rm of} = C_{\rm G-CT} + C_{\rm G-Diff}$. Fig. 7(a)–(d) shows $C_{\rm G-CT}$, $C_{\rm G-Diff}$, and $C_{\rm of}$ calculated for MOSFETs with strip contact corresponding to various geometry parameters, such as L_g, $L_{\rm gct}$, T_g, and $T_{\rm ox}$ over a wide range of variations. A good match is achieved between the 3-D capacitor model and Raphael simulation. All the parasitic capacitances indicate a weak dependence on L_g over a wide range of variations from 65 to 250 nm, as shown in Fig. 7(a). The results suggest that the parasitic capacitances $C_{\rm of} = C_{\rm G-CT} + C_{\rm G-Diff}$ are not scalable with L_g and that the impact on high-frequency performance will go up with L_g scaling. For $L_{\rm gct}$ fixed at 135 nm, according to a 130-nm CMOS typical rule, $C_{\rm G-Diff}$ dominates around two-thirds of the total parasitic capacitance $C_{\rm of}$, and $C_{\rm G-CT}$ contributes the remaining portion of around one-third. Fig. 7(b) shows the parasitic capacitances subject to varying $L_{\rm gct}$ and a good agreement between the calculation by the 3-D capacitor models and Raphael simulation. Interesting results are demonstrated with a significant dependence on $L_{\rm gct}$ that is in contrast with the weak dependence on L_g. The reduction of $L_{\rm gct}$ by 100 nm leads to an increase of $C_{\rm G-CT}$ by around 0.06 fF/μm, while a decrease of $C_{\rm G-Diff}$ leads to an increase by around 0.026 fF/μm. The resulting $C_{\rm of}$ is dominated by the change of $C_{\rm G-CT}$ and increased by 0.034 fF/μm. Note that the increase of parasitic capacitances with $L_{\rm gct}$ scaling highlights the impact from the interelectrode space shrinkage in miniaturized devices on high-frequency or high-speed circuit performance. Fig. 7(c) shows the gate electrode thickness (T_g) scaling effect on parasitic capacitances. Apparently, a thinner T_g helps reduce $C_{\rm G-CT}$ and $C_{\rm of}$ but has a nearly negligible effect on $C_{\rm G-Diff}$. The reduction of $C_{\rm G-CT}$ from thinner T_g contributes to smaller $C_{\rm of}$ and can be beneficial to high-frequency or high-speed circuits. Interestingly, gate oxide thickness $(T_{\rm ox})$ scaling reveals an opposite trend against that of T_g scaling. As shown in Fig. 7(d), the thinner $T_{\rm ox}$ leads to the penalty of significantly larger $C_{\rm of}$, due to a dramatic increase of $C_{\rm G-Diff}$, even a minor effect on $C_{\rm G-CT}$. Note that $T_{\rm ox}$ scaling from 50 to 1 nm results in around four times larger $C_{\rm G-Diff}$ and 83% increase in $C_{\rm of}$. Again, the tremendous increase of $C_{\rm of}$ from $C_{\rm G-Diff}$ will degrade CMOS circuit performance under high-frequency operations. A good match with Raphael simulation in terms of sensitivity and absolute values (mismatch $< 4\%$) can justify the accuracy of the developed 3-D capacitance model.

In the following, a comprehensive certification has been extended to the more complicated structures with square and circular contacts. Both demonstrate similar trends in the dependence on lateral dimensions defined by $(L_g, L_{\rm gct})$ and vertical dimensions specified by $(T_g, T_{\rm ox})$, and the circular contacts shown in Fig. 8 are selected to represent the results. The scaling of $L_{\rm gct}$ and $T_{\rm ox}$ results in larger $C_{\rm of}$ due to an increase of $C_{\rm G-CT}$ and $C_{\rm G-Diff}$ versus shorter $L_{\rm gct}$ and thinner $T_{\rm ox}$, respectively. Note that $T_{\rm ox}$ scaling from 50 to 1 nm increases $C_{\rm of}$ by nearly 100% higher, owing to increasing weighting factor played by $C_{\rm G-Diff}$ in square and circular contact schemes. On the other hand, the thinner T_g helps reduce $C_{\rm G-CT}$ and then $C_{\rm of}$. Again, the accuracy of the 3-D capacitor model is justified by a close match with Raphael simulation, which is even better than those presented for strip contact. Fig. 9 makes a comparison of parasitic capacitances like $C_{\rm G-CT}$, $C_{\rm G-Diff}$, and $C_{\rm of}$ in 80-nm MOSFET with different contact schemes, i.e., strip, square, and circular contacts under varying $L_{\rm gct}$. This comparison reveals the largest $C_{\rm G-CT}$ versus the smallest $C_{\rm G-Diff}$ for strip contact. One more important observation is that $C_{\rm G-CT}$ and $C_{\rm G-Diff}$ are going in opposite directions under varying $L_{\rm gct}$. The shrinkage of $L_{\rm gct}$ leads to larger $C_{\rm G-CT}$ but smaller $C_{\rm G-Diff}$. The resulting $C_{\rm of}$ is dominated by $C_{\rm G-CT}$ in terms of scaling trend versus $L_{\rm gct}$. It is interesting to notice that the circular contact reveals around 13%–15% smaller $C_{\rm G-CT}$ but 3%–4% larger $C_{\rm G-Diff}$ than the square contact. The resulting difference in total parasitic capacitance $C_{\rm of}$ is below 3%. The developed 3-D capacitor model can accurately predict the interesting results and help understand the mechanism. The simultaneous occurrence of $C_{\rm G-CT}$ suppression and $C_{\rm G-Diff}$ growth in the circular contact is governed by the division line locations h_1 and h_2 and the resulting redistribution of field lines from gate to CT and Diff. According to our model for h_1, h_2, and β in (12)–(15), the smaller β for the circular contact compared to the square contact leads to the larger h_1 and h_2 and then decreasing/increasing portion of field lines from the gate-to-CT/Diff region. In this respect, the accuracy of the 3-D capacitor model is proven by a good agreement with Raphael simulation, with a small deviation maintained within 4%.

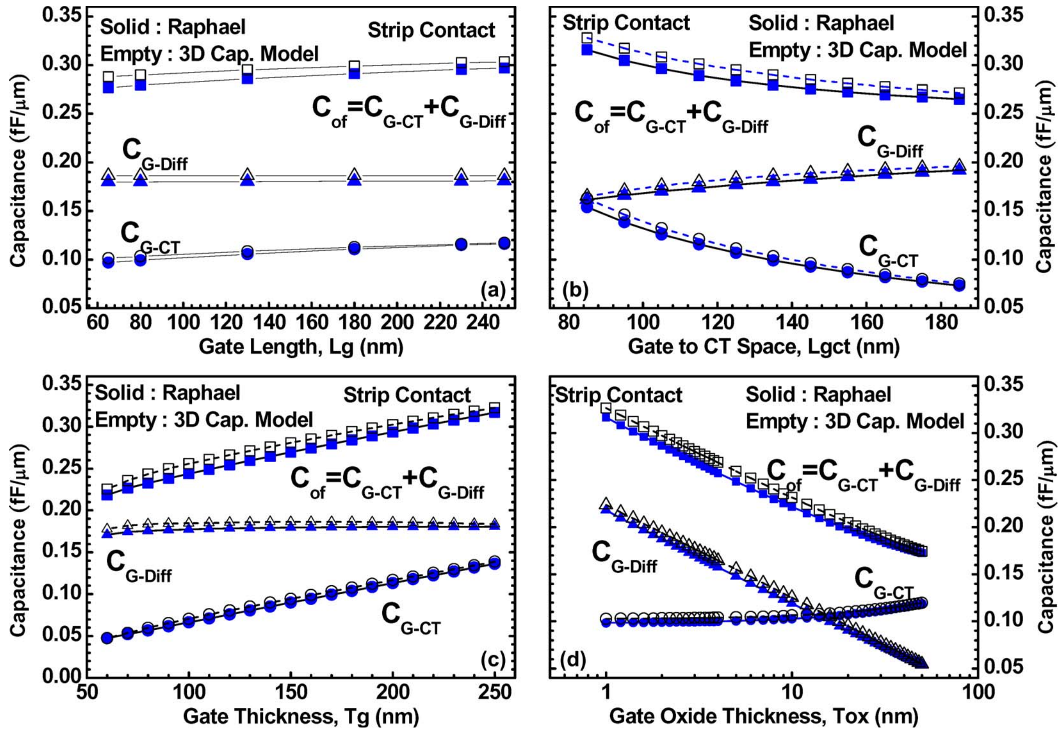


Fig. 7. Parasitic capacitances $C_{\mathrm{G-CT}}$, $C_{\mathrm{G-Diff}}$, and C_{of} calculated by the 3-D capacitor model for MOSFETs with strip contact versus device geometry parameters (a) gate length L_g (b) gate-to-CT space L_{gct}, (c) gate electrode thickness T_g, and (d) gate dielectric thickness T_{ox}. A comparison with Raphael simulation results demonstrates a good agreement. (Solid symbols) Raphael simulation. (Empty symbols) Three-dimensional capacitor model.

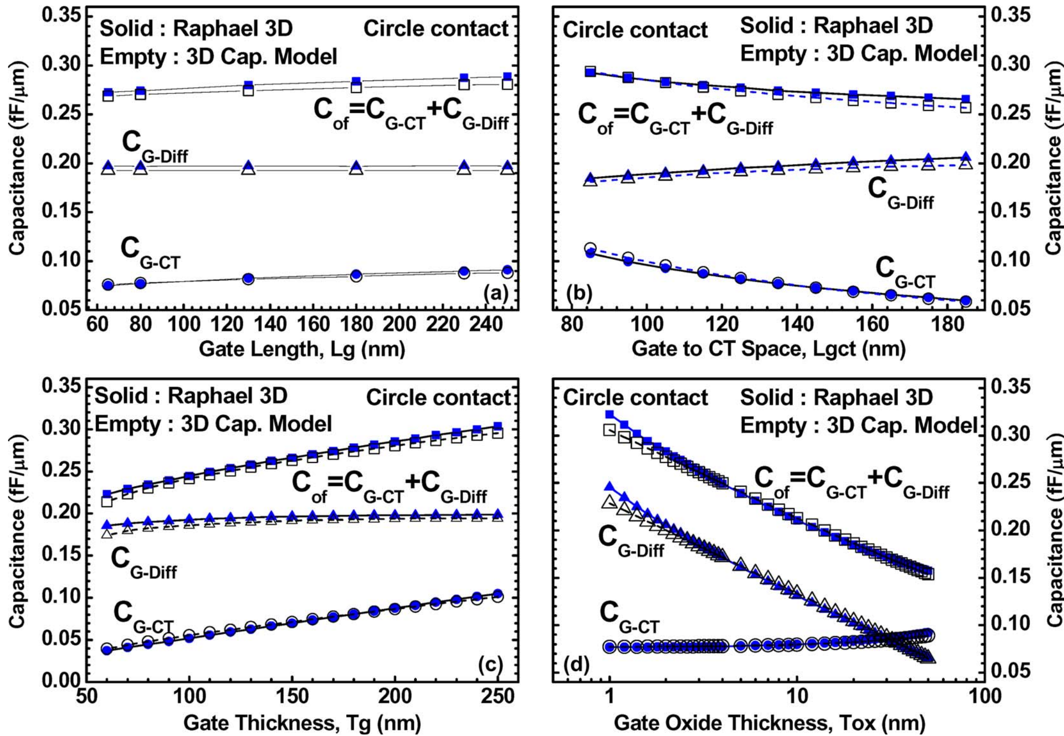


Fig. 8. Parasitic capacitances $C_{\mathrm{G-CT}}$, $C_{\mathrm{G-Diff}}$, and C_{of} calculated by the 3-D capacitor model for MOSFETs with circular contacts versus device geometry parameters (a) gate length L_g, (b) gate-to-CT space L_{gct}, (c) gate electrode thickness T_g, and (d) gate dielectric thickness T_{ox}. A comparison with Raphael simulation results indicates an excellent agreement and justifies the accuracy. (Solid symbols) Raphael simulation. (Empty symbols) Three-dimensional capacitor model.

The dependences on contact size L_{ct} and gate width W_g are two more interesting factors arising from structures with square and circular contacts to be verified. As shown in Fig. 10 for circular contacts, the downscaling of L_{ct} can help reduce $C_{\mathrm{G-CT}}$ by 0.032 fF/μm per 100 nm L_{ct} shrinkage but increases $C_{\mathrm{G-Diff}}$ in a rate of around 0.015 fF/μm. Accordingly, C_{of}

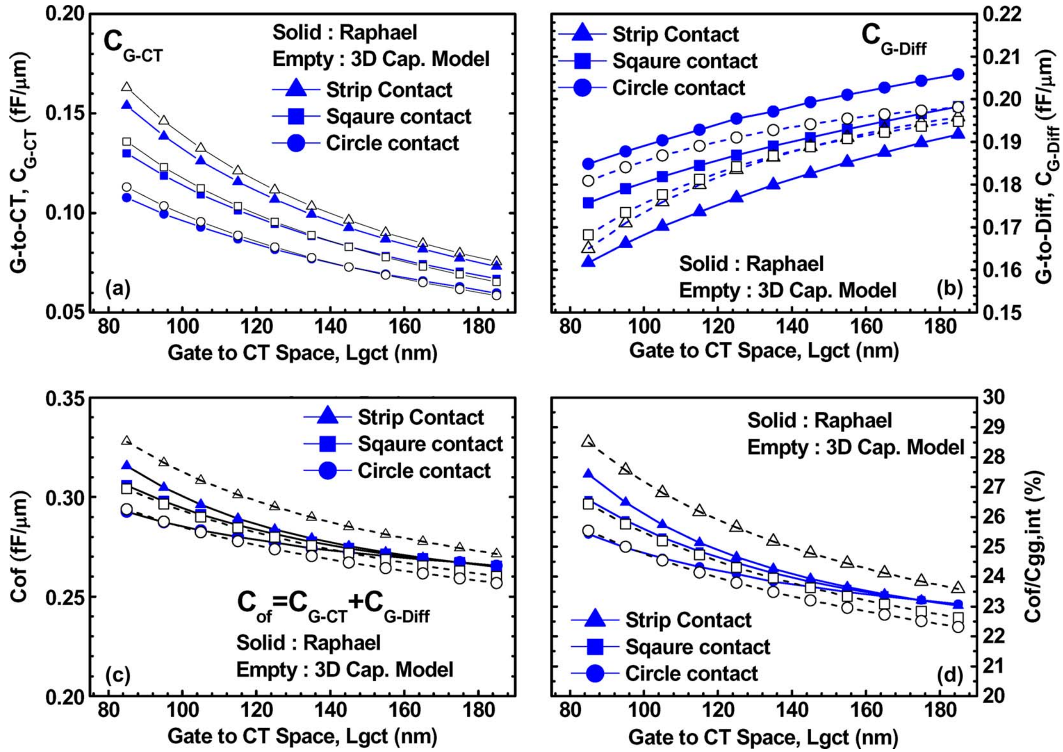


Fig. 9. Comparison of parasitic capacitances in 80-nm MOSFET with different contact schemes, such as strip, square, and circular contacts. (a) $C_{\mathrm{G-CT}}$, (b) $C_{\mathrm{G-Diff}}$, (c) C_{of}, and (d) $C_{\mathrm{of}}/C_{\mathrm{gint}}$ under varying L_{gct}. ($C_{\mathrm{of}} = C_{\mathrm{G-CT}} + C_{\mathrm{G-Diff}}$, $C_{\mathrm{gint}} = C_{\mathrm{ox(inv)}}L_g$, and $L_g = 80$ nm). (Solid symbols) Raphael simulation. (Empty symbols) Three-dimensional capacitor model.

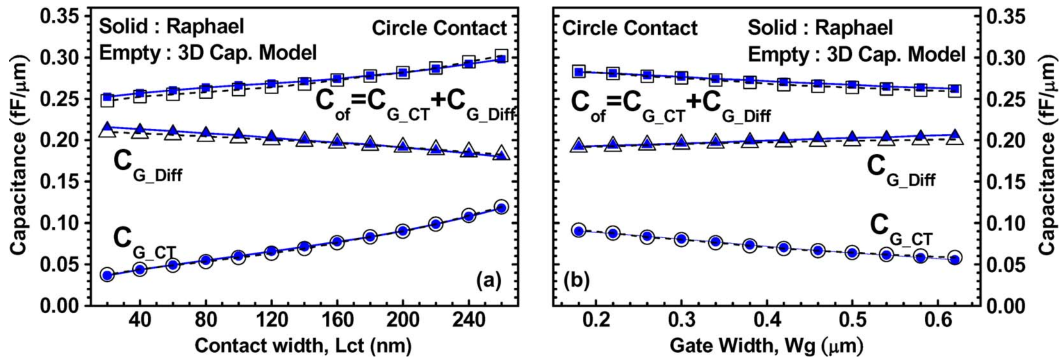


Fig. 10. Parasitic capacitances $C_{\mathrm{G-CT}}$, $C_{\mathrm{G-Diff}}$, and C_{of} calculated by the 3-D capacitor model for MOSFETs with circular contacts versus device geometry parameters (a) contact width L_{ct} and (b) gate width W_g. A comparison with Raphael simulation results indicates an excellent agreement and justifies the accuracy. (Solid symbols) Raphael simulation. (Empty symbols) Three-dimensional capacitor model.

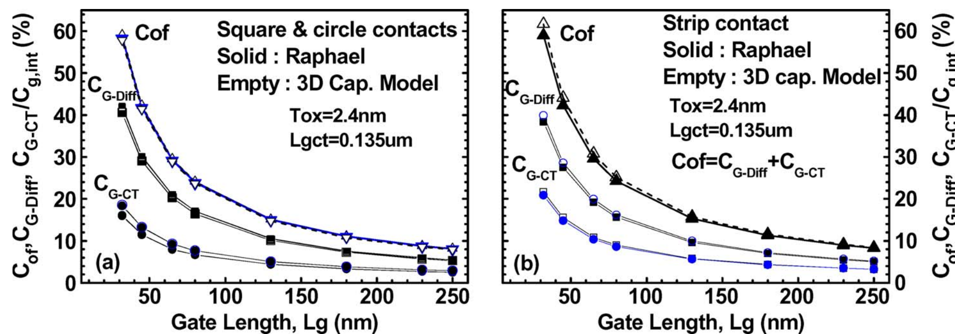


Fig. 11. Weighting factor of the total parasitic capacitances C_{of}, $C_{\mathrm{G-CT}}$, and $C_{\mathrm{G-Diff}}$ over the intrinsic gate capacitances $C_{\mathrm{of}}/C_{\mathrm{gint}}$, $C_{\mathrm{G-CT}}/C_{\mathrm{gint}}$, and $C_{\mathrm{G-Diff}}/C_{\mathrm{gint}}$ ($C_{\mathrm{of}} = C_{\mathrm{G-CT}} + C_{\mathrm{G-Diff}}$ and $C_{\mathrm{gint}} = C_{\mathrm{ox(inv)}}L_g$) corresponding to a wide range of gate lengths, i.e., $L_g = 32-250$ nm, under fixed L_{gct} (135 nm) and T_{ox} (2.4 nm). (a) Square and circular contacts. (b) Strip contact. (Solid symbols) Raphael simulation. (Empty symbols) Three-dimensional capacitor model.

TABLE I

SUMMARY OF MOSFET GEOMETRICAL PARAMETERS L_g, $L_{\rm gct}$, T_g, $T_{\rm ox}$, AND $L_{\rm ct}$ AND THEIR SCALING EFFECT ON PARASITIC CAPACITANCES $C_{\rm G-CT}$, $C_{\rm G-Diff}$, AND $C_{\rm of} = C_{\rm G-CT} + C_{\rm G-Diff}$. L_g: GATE LENGTH, $L_{\rm gct}$: GATE-TO-CT SPACE, T_g: GATE ELECTRODE THICKNESS, $T_{\rm ox}$: GATE DIELECTRIC THICKNESS, AND $L_{\rm ct}$: CONTACT DIMENSION

Geometrical parameters		$C_{G=CT}$	C_{G-Diff}	$C_{of}=C_{G=CT}+C_{G=Diff}$	$\Delta C_{of}/C_{of}$	Material features	Remark
Lg	↓	↓	----	↓	-3 ~ -4%/0.1um	Conductor	😐
Lgct	↓	↑ ↑ ↑	↓	↑ ↑	18.15% /0.1um	Inter-conductor space	☹
Tg	↓	↓ ↓	----	↓ ↓	-18.6%/0.1um	Conductor	☺
Tox	↓	↓	↑ ↑ ↑	↑ ↑	27~31%/dec.	insulator	☹
Lct	↓	↓ ↓	↑	↓	-7 ~ -9%/0.1um	Conductor	☺

that is equal to the sum of $C_{\rm G-CT}$ and $C_{\rm G-Diff}$ is suppressed by around 0.0174 fF/μm. Similar results are realized for the square contact scheme. Fig. 11 shows that the weighting factor played by the parasitic capacitance in terms of $C_{\rm of}/C_{\rm g,int}$ ($C_{\rm g,int}$: the intrinsic gate capacitance) is dramatically increased with L_g scaling for strip, square, and circular contacts. The results manifest itself the nonscalability of parasitic capacitances and the increasing impact on continuously scaled MOSFETs. For 80-nm devices, the parasitic capacitance contributes around 25% of the total gate capacitances. As for a more aggressive scaling to 65, 45, and 32 nm, the weighting factor goes up toward 30%, 40%, and 60%, respectively. This dramatic increase of $C_{\rm of}/C_{\rm g,int}$ actually goes beyond the limitation defined by the latest ITRS for CMOS scaling pushed below 45 nm. As shown in Fig. 11 for $L_{\rm gct}$ fixed at 135 nm, i.e., the typical rule, $C_{\rm G-Diff}$ and $C_{\rm G-CT}$ contribute around two-thirds and one-third of the total parasitic capacitance $C_{\rm of}$, respectively.

Table I summarizes the device geometrical parameters in a MOSFET and their scaling effect on the parasitic capacitances such as $C_{\rm G-CT}$, $C_{\rm G-Diff}$, and $C_{\rm of}$, respectively. The results achieved from the developed 3-D capacitance model provide an important guideline that the scaling of conductors like gate electrode thickness T_g and contact width $L_{\rm ct}$ can help reduce the parasitic capacitances, whereas the shrinkage of interconductor space or insulator thickness like $L_{\rm gct}$ and $T_{\rm ox}$ will lead to the penalty of larger parasitic capacitances owing to the increase of interconductor coupling.

V. CONCLUSION

A new 3-D capacitor model has been developed with the advantages of free from numerical iterations, as well as fitting parameters, and an easy deployment in standard circuit simulators. The accuracy has been proven by an excellent match with Raphael simulation results for devices with various geometry parameters and contact schemes, such as strip, square, and circular contacts. This model predicted that the parasitic capacitance $C_{\rm of}$ may dominate more than 30%/40%/60% of the intrinsic gate capacitance ($C_{\rm gint}$) in 65-/45-/32-nm MOSFETs. It suggests that the 3-D parasitic capacitances exceed the limitation defined by the most updated ITRS and reveal themselves as a show-stopper in high-speed and high-frequency circuit design.

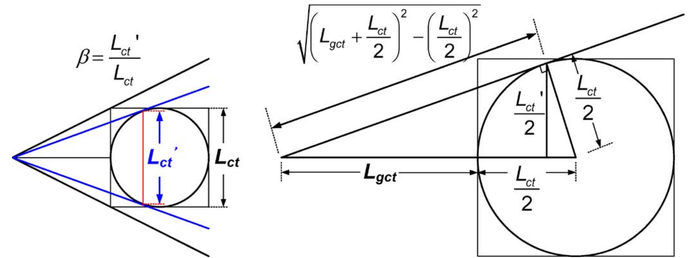


Fig. 12. Circular contact layout analysis using geometrical mathematics for calculating the constant β, which represents the effective dimension of the circular contact compared to the square contact.

APPENDIX

The mathematical formulas for determining the constant β in the key parameters h_1 and h_2 for circular contacts are derived based on the geometrical mathematics and contact layout dimensions such as $L_{\rm ct}$ and $L_{\rm gct}$ shown in Fig. 12.

According to the theory of triangular geometry

$$\frac{L_{\rm ct}}{2}\sqrt{\left(L_{\rm gct}+\frac{L_{\rm ct}}{2}\right)^2-\left(\frac{L_{\rm ct}}{2}\right)^2} = \left(L_{\rm gct}+\frac{L_{\rm ct}}{2}\right)\cdot\frac{L'_{\rm ct}}{2}. \tag{A.1}$$

Then, β, which is defined as $L'_{\rm ct}/L_{\rm ct}$, can be derived from

$$\beta = \frac{L'_{\rm ct}}{L_{\rm ct}} = \frac{\sqrt{\left(L_{\rm gct}+\frac{L_{\rm ct}}{2}\right)^2-\left(\frac{L_{\rm ct}}{2}\right)^2}}{\left(L_{\rm gct}+\frac{L_{\rm ct}}{2}\right)} = \sqrt{1-\left(\frac{\frac{L_{\rm ct}}{2}}{L_{\rm gct}+\frac{L_{\rm ct}}{2}}\right)^2}. \tag{A.2}$$

ACKNOWLEDGMENT

The authors would like to thank Chip implementation Center for providing the simulation environment and the NDL RF Laboratory for the device measurement.

REFERENCES

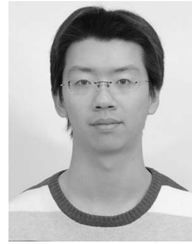
[1] J. C. Guo, W. Y. Lien, M. C. Hung, C. C. Liu, C. W. Chen, C. M. Wu, Y. C. Sun, and P. Yang, "Low-K/Cu CMOS logic based SoC technology for 10 Gb transceiver with 115 GHz f_T, 80 GHz $f_{\max}$ RF CMOS, high-Q MiM capacitor, and spiral Cu inductor," in *VLSI Symp. Tech. Dig.*, Kyoto, Japan, Jun. 2003, pp. 39–40.

[2] C. H. Chen, C. S. Chang, C. P. Chao, J. F. Kuan, C. L. Chang, S. H. Wang, H. M. Hsu, W. Y. Lien, Y. C. Tsai, H. C. Lin, C. C. Wu, C. F. Huang,

S. M. Chen, P. M. Tseng, Y. T. Lin, C. F. Chang, H. J. Lin, J. C. Guo, G. J. Chern, D. Tang, and J. Y. C. Sun, “A 90 nm CMOS MS/RF based foundry SoC technology comprising superb 185 GHz f_T RFMOS and versatile, high-Q passive components for cost/performance optimization,” in *IEDM Tech. Dig.*, 2003, pp. 39–42.
[3] J.-C. Guo, “Low-K/Cu CMOS based SoC technology with 115 GHz f_T, 100 GHz $f_{\max}$, low noise 80 nm RF CMOS, high-Q MiM capacitor and spiral Cu inductor,” *IEEE Trans. Semicond. Manuf.*, vol. 19, no. 3, pp. 331–338, Aug. 2006.
[4] J.-C. Guo and Y.-M. Lin, “A lossy substrate de-embedding method for sub-100 nm RF CMOS noise extraction and modeling,” *IEEE Trans. Electron Devices*, vol. 53, no. 2, pp. 339–347, Feb. 2006.
[5] J.-C. Guo and Y.-M. Lin, “A compact RF CMOS modeling for accurate high frequency noise simulation in sub-100-nm MOSFETs,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 9, pp. 1684–1688, Sep. 2008.
[6] H. Kamchouchi and A. A. Zaky, “A direct method for the calculation of the edge capacitance of thick electrodes,” *J. Phys. D, Appl. Phys.*, vol. 8, no. 12, pp. 1365–1371, Aug. 1975.
[7] M. I. Elmasry, “Capacitance calculations in MOSFET VLSI,” *IEEE Electron Device Lett.*, vol. EDL-3, no. 1, pp. 6–7, Jan. 1982.
[8] R. Shrivastava and K. Fitzpatrick, “A simple model for the overlap capacitance of a VLSI MOS device,” *IEEE Trans. Electron Devices*, vol. ED-29, no. 12, pp. 1870–1875, Dec. 1982.
[9] K. Suzuki, “Parasitic capacitance of submicrometer MOSFETs,” *IEEE Trans. Electron Devices*, vol. 46, no. 9, pp. 1895–1900, Sep. 1999.
[10] N. R. Mohapatra, M. P. Desai, S. G. Narendra, and V. R. Rao, “Modeling of parasitic capacitances in deep submicrometer conventional and high-K dielectric MOS transistors,” *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 959–966, Apr. 2003.

Jyh-Chyurn Guo (M’06–SM’07) received the B.S.E.E. and M.S.E.E. degrees from National Tsing Hua University, Hsinchu, Taiwan, in 1982 and 1984, respectively, and the Ph.D. degree in electronics engineering from National Chiao Tung University (NCTU), Hsinchu, in 1994.

For more than 19 years, she was with the semiconductor industry, where her major focus was on device design and VLSI technology development. In 1984, she joined ERSO/ITRI, where she had been engaged in semiconductor integrated circuit technologies with a broad scope that covered high-voltage, high-power, submicron-project, and high-speed SRAM technologies, etc. From 1994 to 1998, she was with Macronix International Company, Ltd. and engaged in high-density as well as low-power Flash memory technology development. In 1998, she joined Vanguard International Semiconductor Corporation, where she was the Device Department Manager for advanced DRAM device technology development. In 2000, she joined Taiwan Semiconductor Manufacturing Company Ltd., where he was a Program Manager who was in charge of 100-nm logic CMOS FEOL, high-performance-analog, and RF CMOS technology development. In 2003, she joined the Department of Electronics Engineering, NCTU, as an Associate Professor, where she has been a Full Professor since 2008. She has authored or coauthored more than 50 technical papers and is the holder of 19 U.S. patents in her professional field. Her current research interests include RF CMOS and high-performance analog device design and modeling, nanoscale CMOS noise modeling and strain engineering effects, broadband and scalable-inductor modeling, novel nonvolatile memory technology, and device integration technologies for SOC.



Chih-Ting Yeh received the B.S. degree from the Department of Electrophysics and the M.S. degree from the Institute of Electronics, National Chiao Tung University, Hsinchu, Taiwan, in 2004 and 2006, respectively. For his M.S. thesis, he was focusing on 3-D capacitance model development for RF and analog circuit simulation.

Since 2007, he has been with the SoC Technology Center, Design Automation Technology Division, Department of Circuit Design, Industrial Technology Research Institute, Hsinchu, as an ESD Protection Design Engineer. His current research interests include on-chip ESD protection circuit design of CMOS integrated circuits.