Bipolar Theory of MOS Field Effect Transistors and Experiments*, **

Chih-Tang Sah[1, 2, 3, †] and Bin B. Jie[3, †]

(1 *Chinese Academy of Sciences, Foreign Member, Beijing* 100864, *China*)
(2 *University of Florida, Gainesville, Florida* 32605, *USA*)
(3 *Peking University, Beijing* 100871, *China*)

Abstract: The bipolar theory of field effect transistor is introduced to replace the 55 year old classic unipolar theory invented by Shockley in 1952 in order to account for the characteristics observed in recent double gate nanometer silicon MOS field effect transistors. Two electron and two hole surface channels are simultaneously present in all channel current ranges. Output and transfer characteristics are computed over practical base and gate oxide thicknesses. The bipolar theory corroborates well with experimental data reported recently for FinFETs with metal/silicon and p/n junction source/drain contacts. Single device realization of CMOS inverter and SRAM memory circuit functions are recognized.

Key words: unipolar FET theory; bipolar FET theory; simultaneous hole and electron surface channels; volume channel; double gate; pure base
PACC: 7340Q **EEACC**: 2560S; 2560B
CLC number: TN386.1 **Document code**: A **Article ID**: 0253-4177(2007)10-1497-06

1 Introduction

The multiple gate (single, double and triple thus far proposed) thin-base short-channel metal-oxide-silicon (MOS) or insulated-gate (IG) field-effect transistor (FET), abbreviated as MOST or MOSFET, could be a best candidate to replace the single-gate semi-infinite-thick-bulk MOS transistor, in order to extend the downward scaling into the deep nanometer range[1]. These double or multiple gate, thin silicon-base MOS transistors are popularly known as the FinFET with rectangular thin-base semiconductor silicon and two or three isolated or connected gates, and as the WireFET with small-diameter cylindrical solid or hollow semiconductor silicon and wrapped around single gate or additional conductor core for concentric double gate. The mathematical theory of the intrinsic transistor (i.e. ideal geometry excluding parasitic effects) has employed the two classic assumptions made by Shockely in 1952[2], namely, one carrier (unipolar) and gradual channel (implies long and narrow-thin channel), which were introduced by Shockley to analyze the electrical current-voltage characteristics of the p/n junction-gate FETs (JGFET) he invented. These two classic Shockley assumptions have been followed in the subsequent 55 years (1952 to 2007) by all authors to give the baseline electrical (current-voltage) model of the JGFET and later the MOS transistor from which to derive their compact transistor electrical models[3, 4]. These compact models are necessary to design and analyze the electrical current-voltage characteristics of silicon MOSFET in order to speed up the computer-aided-design of multi-transistor (thousands to billions MOS transistor) integrated circuits using the circuit simulator SPICE[3, 4]. The channel-length modulation or shortening is one of the short channel effects excluded in Shockley gradual channel assumption. This has been modeled using the carrier depletion calculation, usually one dimensional, of the thickening of the drain p/n junction space charge layer, from applied reverse-voltage. It was first employed by Shockley in 2-dimenions for the JGFET[2] and later by all subsequent model developers for MOSFET[3~6]. For unipolar-bipolar device modeling, it is well-known that analytical solutions exist for electrical trans-

* Project supported by the CTSAH Associates (CTSA) which was founded by the late Linda Su-Nan Chang Sah
** The bipolar FET model was first presented by us in two Late News papers at the Workshop on Compact Modeling of the Nanotech Conference, in Santa Clara, California, on May 23, 2007.
† Corresponding author. Email: bb_jie@msn.com and tom_sah@msn.com
Received 4 September 2007, revised manuscript received 7 September 2007

port in semiconductors and insulators with one carrier species, known as unipolar conductivity in photoconductor literature of the last century. Most if not all theoretical authors on compact MOS transistor models have followed the two classic Shockley assumptions to start their compact model development[3,4]. Such analytic solution using the two classic Shockley assumptions for FET was employed in 2000 by Yuan Taur[5] to mathematically analyze the double gate thin pure base MOS transistor, which has been followed by all recent authors as indicated by the review given by Adelmo Ortiz-Conde, Francisco J. Garcia-Sanchez, and Juin J. Liou in 2007[6]. However, Shockley's unipolar graduate channel theory, as employed by Taur and followers, cannot account for the recently observed experimental current-voltage characteristics of the nanometer double-gate thin pure base MOS transistors (FinFETs)[7], and in fact, not some of the current-voltage features in thin-film and SOI (semiconductor on insulator) field-effect transistors reported during the last four decades, including the 1971 Fu-Sah experimental demonstration and their bipolar theorectical analysis of the bistable memory transistor consisting of a pMOST and an nJGFET on a single physical silicon FET structure, i.e., a diffused n-silicon-well on p-silicon substrate[8]. In this report, we shall show that the unexplained current-voltage features of the recent nanometer double-gate short-channel MOS transistors or FinFETs are accounted for by the two carrier species (electrons and holes) in our new bipolar theory, absent in the previous unipolar MOSFET theories.

2 Bipolar Field-Effect Theory

To visualize and analyze the simultaneous appearance of the four surface channels from two carrier species (electron and hole species), we divide the transistor into an emitter section and a collector section along its length (y axis), following the bipolar-junction-transistor terminology. In the four surface channel bias mode of the double gate, with positive terminal voltages, these two sections are the electron-source-emitter and electron-drain-collector for the two electron channels induced by the positive voltage applied to the two gates relative to the source, with electrons flowing from the source to the drain, and simultaneously also the hole-drain-emitter and hole-source-collector for the two hole channels induced by the negative voltage applied to two gates relative to the drain, with holes flowing from the drain to the source. Similar four channels appear when the terminal voltages are negative, then the electron and hole current directions are reversed. The current-voltage characteristics are described by ten simple equations derived from the traditional MOS theory using (1) the correct boundary conditions and correct electric and electrochemical potential references and (2) the double integral electrochemical-potential-gradient driven channel current including both carrier species (electrons and holes).

Following the classical approach of decomposing the two-dimensional transistor in each of the two sections (emitter and collector), into two coupled one dimensional problems[3,9,10], and making the gradual channel assumption in the emitter section, the equations of the four surface channel currents and the drain terminal current are given by the following ten equations, in the electron source emitter and hole drain emitter.

In the electron source emitter ($0 \leqslant U_0 \leqslant U_S$):

$$U_{GB}-U_{FB}-U_S=\frac{C_D}{C_O}\sqrt{\exp(-U_N)(\exp U_S-\exp U_0)},\quad 0<V_{DS}\gtrless V_{GS} \tag{1}$$

$$2\int_{U_0}^{U_S}\frac{dU}{\sqrt{\exp(-U_N)(\exp U-\exp U_0)}}\equiv\chi_B=\frac{x_B}{L_D}\equiv X_B,\quad 0<V_{DS}\gtrless V_{GS} \tag{2}$$

$$I_{N3}=\frac{Z}{L}2D_n\frac{kT}{q}C_O\int_{U_{SB}}^{U_{DB}}(U_{GB}-U_{FB}-U_S)dU_N,\quad 0<V_{DS}\leqslant V_{GS} \tag{3}$$

$$I_{P4}=0\ \text{(This is the zeroth approximation.)},\quad 0<V_{DS}\leqslant V_{GS} \tag{4}$$

$$I_{N5}=\frac{Z}{L}2D_n\frac{kT}{q}C_O\int_{U_{SB}}^{U_{GB}-U_{FB}-1}(U_{GB}-U_{FB}-U_S)dU_N,\quad 0<V_{GS}\leqslant V_{DS} \tag{5}$$

In the hole drain emitter ($U_S \leqslant U_0 \leqslant 0$):

$$U_{GB}-U_{FB}-U_S=-\frac{C_D}{C_O}\sqrt{\exp U_P[\exp(-U_S)-\exp(-U_0)]},\quad 0<V_{GS}\leqslant V_{DS} \tag{6}$$

$$-2\int_{U_0}^{U_S}\frac{dU}{\sqrt{\exp U_P[\exp(-U)-\exp(-U_0)]}}\equiv\chi_B=\frac{x_B}{L_D}\equiv X_B,\quad 0<V_{GS}\leqslant V_{DS} \tag{7}$$

$$I_{P8}=\frac{Z}{L}2D_p\frac{kT}{q}C_O\int_{U_{GB}-U_{FB}+1}^{U_{DB}}-(U_{GB}-U_{FB}-U_S)dU_P,$$

$$0 < V_{GS} \leqslant V_{DS} \quad (8)$$

The positive drain terminal current flowing into the drain contact is then given by

$$I_D = I_{N3} + I_{P4} \approx I_{N3}, \quad 0 < V_{DS} \leqslant V_{GS} \quad (9)$$

$$I_D = I_{N5} + I_{P8} \approx I_{N5sat} + I_{P8}, \quad 0 < V_{GS} \leqslant V_{DS} \quad (10)$$

The potentials and terminal voltages are all normalized to kT/q. The notations were defined in Ref. [3]. $L_D = (\varepsilon_S kT/2q^2 n_i)^{1/2}$ is the Debye length in the pure base ~26μm at room temperature with $n_i = 1 \times 10^{10}\,cm^{-3}$ and $\varepsilon_S = 11.7 \times 8.852 \times 10^{-14}$ F/cm for pure crystalline silicon. $C_D = \varepsilon_S / L_D$ and $C_O = \varepsilon_O / x_O$ are the Debye and oxide capacitance per unit area. $I_{N\#}$ and $I_{P\#}$ are the electron and hole surface channel currents given by the respective equation number (#). x_B is the silicon base thickness; W, the base width; and L, the base length. It should be noted that the current integrals in Eqs. (3), (5) and (8), were erroneously evaluated by compact model authors due to replacing U_N or U_P by $U_{NP} = U_N - U_P$[5,6] which arbitrarily sets U_N or U_P as the constant reference, including zero. This ten-equation solution is symmetrical in the polarity of the applied terminal voltages. Thus, in the negative terminal voltage range, these ten equations can be used by interchanging electron and hole, which will give the parabolic hole surface channel current in $V_{GS} \leqslant V_{DS} < 0$ and the constant-saturation hole and parabolic saturation electron surface channel currents in $V_{DS} \leqslant V_{GS} < 0$.

3 Computed Characteristics

The V-shaped solid curves with minimum in Fig. 1 show the computed transfer characteristics, I_D versus $V_{GS} = V_{GB} - V_{SB}$, with $V_{DS} = V_{DB} - V_{SB} =$ 0.1, 0.3, 0.5, 0.7, 1.0 and 2.0V, of a MOS transistor with two identical gates and pure base thickness $x_B = 30$nm, and the gate insulator electrical thickness $x_O = 1.5$nm. The three flatband voltages of the three terminals S, G, and D, are set to zero, or absorbed into the respective terminal voltages. Electron and hole mobilities are assumed equal and taken as $400\,cm^2/(V \cdot s)$. The electron I_{NS} (dots) and hole I_{PD} (circles) surface channel current components are also shown in Fig. 1. The salient features are the minimum of the total drain terminal current, and the positive-negative gate voltage

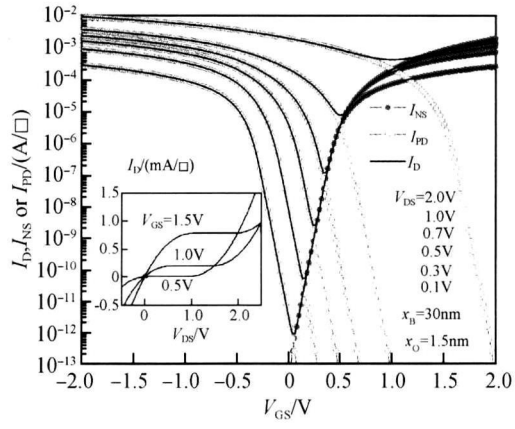

Fig. 1 Drain source voltage V_{DS} dependence of the DC transfer current voltage characteristics of an nFinFET with metal/silicon drain and source contacts. Total drain terminal current, I_D (solid line); Electron I_{NS} (dots) and hole I_{PS} (circles) channel current components of the drain terminal current. Inset: I_D-V_{DS} showing 4 surface channel currents, 2 hole and 2 electron each from two gates.

symmetry centered at the minimum due to the parallel gate voltage shift of the hole parabolic saturation current, I_{PD}, by the applied V_{DS}. The V-shaped transfer characteristics and the minima were observed in recent FinFETs reported by IMEC-Assignees[7] but misinterpreted as GIDL-like (Gate Induced Drain Leakage due to tunneling through reverse-biased drain p/n junction).

The inset in Fig. 1 shows the output characteristics I_D-V_{DS} in the presence of both electron channel in the source emitter section, I_{NS} (constant saturation current), and hole channel in the drain emitter section, I_{PD} (parabolic saturation current). I_{PD} has not been recognized by recent authors and seemed to be absent in their data, which could have been easily discarded by them as bad devices during first silicon or first-few silicon's. Hints of I_{PD} from the parabolic rising $I_D > I_{DSsat}$ when $V_{DS} > V_{DSsat}$ were present in both low-voltage SOI (silicon on insulator) and thin-film FETs in the past decades, interpreted as interband tunneling generation of electron-hole pairs which could be important only at high V_{DS} and V_{GD} in highly impure base/drain junctions. This parabolic hole surface-channel current in the saturation range would give large output conductance.

Figure 2 shows the computed base thickness dependence of the I_D-V_{GS} transfer characteristics at

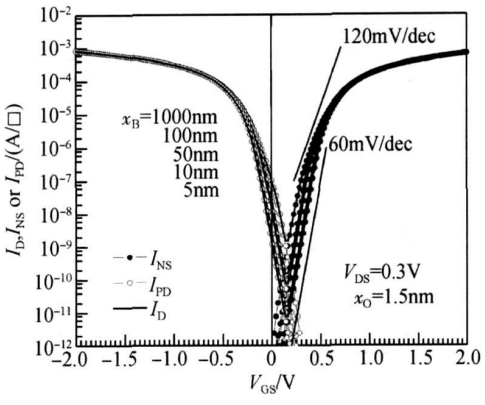

Fig. 2 Base thickness dependence of the DC transfer current voltage characteristics of nFinFETs with metal/silicon (pure base) drain and source contacts. Total drain terminal current I_D (solid line); Electron I_{NS} (dots) and hole (circles) channel current components of the drain terminal current. Note subthreshold voltage swing from 120 to 60mV/decade.

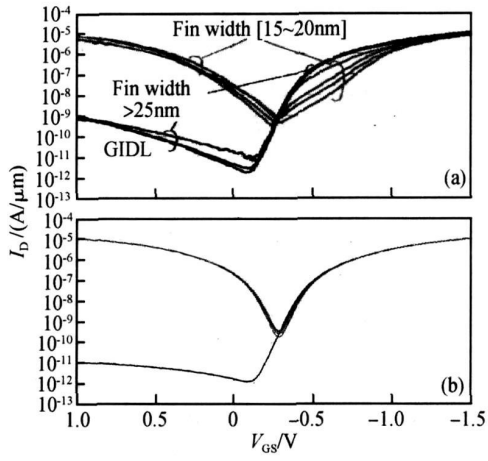

Fig. 3 Corroboration of experiments (a) of pFinFETs from Fig. 10 of IMEC + Assignees[7], with our theory (b) for two kinds of drain and source contacts: the ohmic metal/silicon (pure base thicknesses of x_B = 10, 15, 20nm and infinite interface recombination velocity) which gives the V shaped curves, and the n+/p (p base thickness x_B = 30nm and P_{IM} = ~ 10^{16} cm^{-3}) which gives the ∫ shaped curves

V_{DS} = 0.3V, and its electron and hole surface channel components, I_{NS} and I_{PD}, with x_B = 5, 10, 50, 100, 1000nm, and gate insulator electrical thickness x_O = 1.5nm. A salient feature is the subthreshold swing that slackens to 120mV/decade before it steepens to 60mV/decade when the silicon base thickens to x_B = 1000nm = 1μm which is still thin compared with the ~ 26μm pure base Debye length. The steep and ideal slope at the 60mV/decade voltage swing in the deep subthreshold range, illustrates rapid transformation from electron-hole equality bipolar conduction at the I_{Dmin} to unipolar conduction as the gate voltage is increased or decreased from the minimum, for examples, the I_{NS}/I_{PS} ratio changes from 1 to 100 or 0.01 for a change of $|V_{GS} - V_{GSmin}|$ of ±60mV, and by $10^{\pm 4}$ for ± 120mV, giving substantial on/off ratio in the CMOS inverter mode.

This variation of the subthreshold slope has been observed in all recent experimental thin-base FinFETs and WireFETs, although in some cases, rather subtly. An independent theoretical confirmation of the slope change was obtained by us recently from the transfer characteristics I_D-V_{GS}, computed using the 'exact' Jie-Sah integration equations[10] for the single gate, long and semi-infinite wide channel, bulk MOS transistor, by decreasing the base impurity concentration towards the pure base, from $P_{IM} = 10^{18}$ to 10^9 $cm^{-3} \ll n_i = 10^{10}$ cm^{-3}. The dependence on oxide thickness is not illustrated in this paper due to page number limit of short report, but also because it shows just a constant shift of the current level (log) and gate voltage (linear) with little shape change in I_{NS}-V_{GS} and I_{PD}-V_{GS} as predicted by the ten equations.

4 Corroboration with Experiments

Figure 3 (a) shows the experimental transfer characteristics, I_D-V_{GS}, of two groups of pFinFETs reported in 2005 IMEC + Assignees[7]. The V-shaped curves are from transistors with metal/silicon contacts to the drain and source ends of the thin base which were over-alloying-heated that usually gives ohmic contacts (traditionally represented by infinite interface recombination velocity[9]) while the ∫ shaped curves are transistors with p+ contacts to the drain and source ends of the thin base without over-heating. These experimental data are consistent with the present theory in two groups of curves shown in Fig. 3 (b). These experimental data triggered our investigation that led to our discovery of the bipolar conduction in FET which predicts the occurrence of four simultaneous surface channel channels (two hole and two electron channels) in double-gate thin-base transistors.

Another corroboration of experiments with theory is shown in Figs. 4 (a) and (b). Figure 4

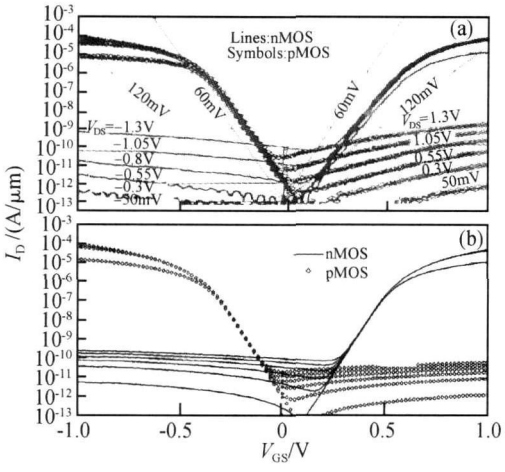

Fig. 4　Corroboration of experimental transfer characteristics (a) of nFinFET and pFinFET from Fig. 4 of IMEC + Assignees[7], with our theory (b) for n+/p (nMOST) and p+/n (pMOST) drain and source contacts

(a) shows the experimental transfer characteristics of an nFinFET (lines) and a pFinFET (symbols) with respectively n+ and p+ drain and source contacts reported in 2005 IMEC + Assignees[7]. Note the subthreshold swing starts from 120mV/decade steepening towards the ideal theoretical minimum of 60mV/decade. Note also the nearly equal V_{DS} shift of the V_{GS} shift of the supposedly I_D off or accumulation range of the I_D-V_{GS}, $\Delta V_{GS} \cong \pm \Delta V_{DS} \cong \pm 0.25$V respectively for the nFinFET and pFinFET. These are telltales of bipolar conduction which the IMEC authors[7] attributed to GIDL-'like' and other recent investigators, GIDL, which is untenable (for interband tunneling) at such low dopant impurity concentrations, even in rather short physical channels. Figure 4 (b) is computed using the ten equations which demonstrate all the features of the experimental data observed in the pFinFET and nFinFET shown in Fig. 4 (a). The spread of the experimental data in Fig. 4 (a) reflects no doubt the well-known traditional technology variations of geometry and especially concentration of residual impurity in the supposedly pure base in these early-technology first-silicon transistors.

5　Concluding Summary

The new bipolar theory of the MOS field effect transistor is presented to replace the 55 year old classic unipolar theory of the field effect transistor, invented by Shockley in 1952[2]. The transfer and output current-voltage characteristics are computed and illustrated. They corroborate very well with that observed recently on nanometer double gate thin pure base silicon MOS transistors known as FinFETs.

Acknowledgments　The authors thank many veteran colleagues for encouragements and technical supports. They are Dr. Colin McAndrew of Free Scale Semiconductor Corporation (USA), Dr. Jin Cai, Dr. Tak H. Ning, Dr. Lewis M. Terman and Dr. Hwa-Nien Yu of IBM (USA), and six Professors (alphabetically): Dr. Carlos Galup-Montoro and Dr. Marcio Cherem Schneider of Federal University of Santa Catarina (Brazil), Dr. Gennady Gildenblatt of Arizona State University (USA), Dr. Chenming Hu of University of California Berkeley (USA), Dr. Mitiko Miura-Mattausch of Hiroshima University (Japan), and Dr. Xing Zhou of Nangyang Technological University (Singapore). We also thank four professors who have invited us to present the Bipolar FET model at their regional IEEE EDS Mini-Colloquium and ICCDCS: Dr. Marcel D. Profirescu of University Politehnica of Bucharest (Romania) on October 14, 2007, Dr. Adelmo Ortiz-Conde and Dr. Francisco J. Garcia Sanchez of Universidad Simon Bolivar (Venezuela), and Dr. Juin J. Liou of University of Central Florida (USA) on April 27 – 30, 2008.

References

[1] Chenming Hu, "From CMOS to Nanotechnology," Keynote, 17th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference, May 22 – 24 2006, Boston, Massachusetts.

[2] William Shockley, "A Unipolar 'Field Effect' Transistor," Proceedings of the IRE, 40(11), 1365 – 1376, November 1952. For experimental reduction to practice, see George C. Dacey and Ian M. Ross, "Unipolar 'Field Effect' Transistor," Proceedings of the IRE, 41(8), 970 – 979, August, 1953.

[3] Chih-Tang Sah, "A history of MOS Transistor Compact Modeling," Keynote. Technical Proceedings, Workshop on Compact Modeling (WCM), 349 – 390. Editors: Xing Zhou, Matthew Laudon and Bart Romanowicz. The NSTI Nanotechnology Conference and Trade Show, May 8 – 12, 2005. Nano Science and Technology Institute, Cambridge, MA 02139, USA. #PCP05040394. On line: http://www.nsti.org/publ/Nanotech2005WCM/1429.pdf

[4] J. Watts, C. McAndrew, C. Enz, C. Galup-Montoro, G. Gildenblat, C. Hu, R. van Langevelde, M. Miura-Mat

tausch, R. Rios and C.-T. Sah, "Advanced Compact Models for MOSFETs," pp. 3–12 of reference [3].

[5] Yuan Taur, "An analytical solution to a double gate MOSFET with undoped body," IEEE Electron Device Letters, 21(5), 245–248, May 2000.

[6] A. Ortiz-Conde, F. J. Garcia Sanchez, J. Muci, S. Malobabic, J. Liou, "A review of core compact models for undoped double gate SOI MOSFETs," IEEE Transactions on Electron Devices, 54(1), 131–139, January 2007.

[7] T. Hoffmann, G. Doornbos, I. Ferain, N. Collaert, P. Zimmerman, M. Goodwin, R. Rooyakers, A. Kottantharayil, Y. Yim, A. Dixit, K. De Meyer, M. Jurczak, and S. Biesemans, "GIDL(gate induced drain leakage) and parasitic Schottky barrier leakage elimination in aggressively scaled HfO_2/TiN FinFET devices," IEDM Technical Digest, 743–746, December 2005.

[8] This bipolar FET, a CMOS transistor, with dual field effect transistors in one physical transistor structure, was first conceived and theorized by one of us in 1970. It consists of the p-surface inversion channel MOS field effect transistor (pMOST) and an n-type volume channel junction gate field effect transistor (pJGFET) in one physical FET. This was experimentally demonstrated in H. S. Fu and C. T. Sah, "A distributed gate bistable MOS transistor," Solid State Electronics, 14(9), 799–804, September 1971.

[9] Chih-Tang Sah, *Fundamentals of Solid State Electronics*, Singapore: World Scientific Publishing Company, 1991. See Chapter 6 for the elements of the partition method to solve the 2-dimensional MOSFET by two 1-dimensional devices. See sections 580 to 583 on pages 500 to 509 for the complete mathematical solutions of currents and their limits in two terminal devices such as p/n junctions and metal/semiconductor barriers that are central parts of short (electrically and physically) channel field effect transistors.

[10] Bin B. Jie and Chih-Tang Sah, "Accuracy of surface potential based long wide channel thick base MOS transistor models," IEEE Transactions on Electron Devices, 54(8), 1897–1909, August 2007.

MOS 场引晶体管双极理论和实验*, **, ***

薛支唐[1,2,3,†] 揭斌斌[3,†]

(1 中国科学院外籍院士, 北京 100864)
(2 佛罗里达大学, 佛罗里达州, Gainesville FL32605, 美国)
(3 北京大学, 北京 100871)

摘要: 提出场引晶体管双极理论. 替代已 55 年久, 1952 Shockley 发明单极理论. 解释近来双栅纳米硅 MOS 晶体管实验特性——两条电子和两条空穴表面沟道, 同时并存. 理算晶体管输出特性和转移特性, 包括实用硅基及栅氧化层厚度. 理算比较最近报道实验, 利用硅 FinFET, 含(金属/硅)和(p/n)结, 源和漏接触. 实验支持双极理论. 建议采用单管, 实现 CMOS 倒相电路和 SRAM 存储电路.

关键词: 单场引 FET 理论; 双场引 FET 理论; MOSFET; 同时存在空穴电子表面沟道和体积沟道; 双栅; 纯基 FET 理论

PACC: 7340Q **EEACC**: 2560S; 2560B

中图分类号: TN386.1 文献标识码: A 文章编号: 0253-4177(2007)10-1497-06

* CTSAH Associates (CTSA) 资助项目(CTSA 由已故的薛夫人张淑南女士创建)

** 本篇内容来自我们的两篇最新进展报告. 这些报告于 2007 年 5 月 23 日在美国加州 Santa Clara 举行的 Workshop on Compact Modeling 上宣读过.

*** 薛支唐写成此摘要基于揭斌斌的现代语初稿. 感谢潘胜和北京大学原物理系教师赵立群和潘桂明的修改建议.

† 通信作者. Email: bb_jie@msn.com, tom_sah@msn.com

2007-09-04 收到, 2007-09-07 定稿