行政院國家科學委員會補助專題研究計畫 期末報告

※※※※※※※※※※※※※※※※※※※※※※

適用於通道特性長度低至十奈米左右之一維及二維奈米場效電晶體的非平衡通道背向散射模式架構

※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：■ 個別型計畫　　□整合型計畫

計畫編號：NSC 98-2221-E-009-164- MY3

執行期間： 98/08/01 ~ 101/07/31

計畫主持人：陳明哲

計畫參與人員：李建志，李韋漢，張立鳴，光心君，黃怡惠，葉婷銜，陳宛勵，李致葳

執行單位：**國立交通大學電子工程學系及電子研究所**

中　華　民　國　101　年　10　月　30　日

行政院國家科學委員會補助專題研究計畫期中精簡報告

適用於通道特性長度低至十奈米左右之一維及二維奈米場效電晶體的非平衡通道背向散射模式架構

Non-equilibrium Channel Backscattering Framework Suitable for 1-D and 2-D NanoFETs with Feature Size down to 10 nm and beyond

執行期限: 98/08/01 ~ 101/07/31

計畫編號: NSC 98-2221-E-009-164- MY3

主持人：陳明哲教授　　國立交通大學電子工程學系及電子研究所

一、中文摘要

本計畫為期三年，主要探討並實際產出更新的非平衡通道背向散射模式架構，使能適用於未來通道特性長度低至十奈米左右之一維及二維奈米場效電晶體，不管從元件物理或實驗數據分析觀點來看，預期能產生深遠的影響. 第一年將進行下列項目：(1) 建立非平衡下二維奈米場效電晶體通道背向散射模式架構，將以往長通道場效電晶體忽略掉的高階物理現象有系統的處理並容納之，如量子侷限，載子退化，速度高彈，源極通道介面瓶頸，非平衡源極通道障壁透納，長距庫倫交互作用等；(2) 修正或微調非平衡下二維奈米場效電晶體通道背向散射模式架構，運用購置的蒙地卡羅元件模擬器及自行製作並工業界提供的測試鍵分別模擬及量測在不同元件結構參數，不同材料物理參數，不同操作偏壓溫度條件下跌等元件背向散射係數及障壁入射速度，以使非平衡架構精確度至通道長度十奈米左右；以及 (3) 一維奈米線場效電晶體通道垂直方向量子模擬器程式撰寫及除錯，以為第二年研究準備. 第二年將進行下列項目：(1)建立非平衡下一維奈米場效電晶體通道背向散射模式架構，運用自行研發通道垂直方向量子模擬器程式，有系統的處理並容納高階物理現象如量子侷限，載子退化，速度高彈，源極通道介面瓶頸，非平衡源極通道障壁透納，長距庫倫交互作用等；(2) 運用購置的蒙地卡羅元件模擬器及自行製作並工業或學術界提供一維測試鍵以修正微調非平衡下一維奈米場效電晶體通道背向散射模式架構，使得一維非平衡架構精確度可至通道長度十奈米左右；以及(3) 繼續非平衡下二維奈米場效電晶體通道背向散射模式架構主題深入鑽研. 第三年將進行下列項目：(1) 繼續前兩年非平衡下一維及二維奈米場效電晶體通道背向散射模式架構主題之研究；(2) 繼續運用購置的蒙地卡羅元件模擬器作一深入探討；以及 (3) 理論及實驗雙重展示非平衡下一維及二維奈米場效電晶體通道背向散射模式架構之應用－統計擾動；自我加熱；以及電報雜訊等.

關鍵詞:

量子，奈米，金氧半場效電晶體，奈米線，散射，統計，擾動，雜訊

英文摘要

This is a three-year project proposal aimed at examining and creating an updated, non-equilibrium version of the channel backscattering framework suitable for 1-D and 2-D nanoFETs with the channel lengths down to 10 nm and beyond. This proposal is expected to have a profound impact, either from the aspect of device physics or the experimental data analysis. In the first year, we will have three items to be carried out: (i) construction of a non-equilibrium 2-D nanoFETs channel backscattering framework, where higher-orders physical effects, greatly ignored before in the case of long channel counterparts, will be systematically treated and therefore incorporated, such as quantum confinement, carrier degeneracy, velocity

overshoot, source-channel interface bottleneck, tunneling across the non-equilibrium source-channel barrier, long-range Coulomb interactions, etc; (ii) modification or refining of the non-equilibrium 2-D nanoFETs channel backscattering framework with the accuracy of 10-nm channel length and below, which will be done with the purchased Monte Carlo device simulators and fabricated device test-key to find out underlying backscattering coefficients and injection velocity over the barrier as a function of the device structural parameters, material physical parameters, operating biases and temperatures; and (iii) writing and debugging of a Schrodinger-Poisson equation self-consistent solver for the direction normal to the channel of 1-D nanowire FET and we can straightforwardly change some parameters as required for the next year proposal. In the second year, we will have three items to be carried out: (i) construction of a non-equilibrium 1-D nanoFETs channel backscattering framework, where higher-orders physical effects will be incorporated, such as quantum confinement, carrier degeneracy, velocity overshoot, source-channel interface bottleneck, tunneling across the non-equilibrium source-channel barrier, long-range Coulomb interactions, etc; (ii) modification or refining of the non-equilibrium 1-D nanoFETs channel backscattering framework with the accuracy of 10-nm channel length and below, which will be done with both the purchased Monte Carlo device simulators and the device test-key fabricated by our group and also provided by the industry and/or the academic colleagues; and (iii) in-depth study of the 2-D non-equilibrium nanoFETs backscattering issue. Finally, in the third year, three main items will be addressed: (i) continuing further investigation of 1-D/2-D non-equilibrium nanoFETs backscattering issues; (ii) more produced microscopic physical phenomena by executing the purchased Monte Carlo device simulators; and (iii) practical demonstrations of the 1-D/2-D non-equilibrium nanoFETs backscattering framework, both theoretically and experimentally, in terms of the statistical fluctuations, self-heating, and random telegraph signals.

Key Words:
Quantum, Nano, MOSFETs, Nanowire, Scatter, Statistical, Fluctuations, Noise

二、緣由與目的

- **本研究計畫領域歷史回顧**

1979 年, IBM Dr. Price 從蒙地卡羅模擬中首次觀察到通道靠近源極部分反射係數 rc 關鍵地決定了元件整體性能 (P. J. Price, Semicond. Semimetals. Vol.14, p.249, 1979; also refer to Landauer, IBM J. Res. Develop., p.223, 1957; McKelvey, etc. Phys. Rev. p.51, 1961). 但此重要的發現要 18 年後即1997年, 才被普度大學Lundstrom教授注意到並引用連結至同校Datta教授的平衡下背向散射係數公式 (Lundstrom, EDL, p. 361, 1997; Datta, Electronic Transport in Mesoscopic Structures, Cambridge, 1995), 此公式看似異常簡單, 但隱含重要物理現象, 使用時須小心. 在 2002 IEDM, 本人及共同研究者率先發表了溫度變化技巧以萃取背向散射係數 rc (M. J. Chen, et al. IEDM Tech. Dig., 2002, p. 39).

- **領域近況 – 應用面**

在 2004 IEDM, Stanford University Dr. Pop 等人(Professor Dutton group)引用我們的溫度變化技巧建立了 a new electro-thermal model 以分析 Self-Heating in Thin Film SOI and GOI FETs (E. Pop, et al. IEDM Tech. Dig., 2004, p. 411).

2005 Symp. VLSI Technology (p. 174), TSMC Dr. Lin 等人引用我們的溫度變化技巧實驗取得 rc and vinj in both uniaxial compressive stress p-FETs and uniaxial tensile stress n-FETs.

2006 IEDM (p.473), Singapore Dr. Liow 等人引用我們的溫度變化技巧實驗取得 rc and vinj in sub-30-nm FinFETs with SiC source/drain.

2007 IEDM (p.667), MIT Prof. Antoniadis 等人直接引用並改進 Prof. Lundstrom 原始有關 mobility 變化之關係式

使得不必用到溫度變化技巧而能在室溫下實驗決定 strain related rc and vinj.

2007 Symp. VLSI Technology (p. 128), French Dr. Barral 等人參考我們的溫度變化技巧從而提出了 a new compact model 使能在任何溫度下實驗決定 strain related rc and vinj.

2007 IEDM (p. 895), 北京大學 Prof. Huang 等人引起我們的溫度變化部分技巧用以分析 1-D silicon nanowire FETs.

- **領域近況 -- 批評面**

2003 ESSDERC (p.147), French Dr. Mouis and Dr. Barraud 從蒙地卡羅模擬中觀察到在一 near-ballistic transport 下 barrier 頂的背向速度分佈不為 Prof. Lundstrom 所認為的 Maxwellian 形態.

2004 SSE (p.1417), German Prof. Jungemann 從蒙地卡羅模擬中觀察到 non-equilibrium nature of vinj, 明顯與 Prof. Lundstrom 理論不符.

2005 TED (p.2727) and 2005 TED (p.2736), 義大利團隊 Dr. Palestri and Dr. Eminente 等人從蒙地卡羅模擬中觀察到 backscattered carriers 有部分來自 *l* 之外而非全出自 *l* 本身.

2006 IEDM (p.945), 義大利團隊 Dr. Palestri 等人從蒙地卡羅模擬得出 Carrier Degeneracy effect on mean-free-path for backscattering, 攻至 Prof. Lundstrom 理論脆弱之處.

2007 IEDM (p.105), 義大利團隊 Dr. Zilli 等人從蒙地卡羅模擬作一溫度效應之研究, 從而批判我們的溫度變化技巧.

2007 IEDM (p. 109), U. Massachusetts Prof. Fischetti 團隊從蒙地卡羅模擬, 提出一個重要觀念: 在 near-ballistic transport in channel 之下, source 被迫處在 non-equilibrium conditions, 間接指出了 Prof. Lundstrom 理論的另一弱點.

- **本研究計畫目的**

正如上述, Prof. Lundstrom 理論公式看似異常簡單, 但隱含重要物理現象, 使用時須萬分小心. 也因此 Prof. Lundstrom 理論易被人誤解, 雖則本人最近作了一些努力:

2008 TED (p.1265), 我們導出了 a new physically-based compact model for kT layer *l*.

接著 2008 Dec. TED (p.3594), 我們執行 3-D Bulk 蒙地卡羅模擬 (*L* down to 15 nm), 強力支持 Prof. Lundstrom 理論.

但顯而易見地, 必須且緊迫性地, 現階段 Prof. Lundstrom 所一手建立的 Backscattering 架構需要進行根本的改進, 以外顯方式呈現出重要的物理現象: (i) Carrier degeneracy in 1-D/2-D confinement; and (ii) Non-equilibrium (scattering, heating, and self-consistent Poisson equation solving) in source, channel, and even the drain, as well as those of long-range Coulomb interactions from plasmons in heavily doped source, drain, and gate regions (as suggested by Prof. Fischetti, TED, p. 2116, 2007). 我們的溫度變化技巧也要一併回應. 可預見地, 更新後的非平衡下 Backscattering 架構及其相應溫度變化技巧將能有用於未來特性長度低至十奈米左右之一維及二維奈米場效電晶體, 不管從元件物理或實驗數據分析觀點來看, 預期能產生深遠的影響.

三、研究方法與成果

方法:

(1) 蒙地卡羅元件模擬軟體購置及應用

●國科會支持購置有蒙地卡羅功能的元件模擬軟體 AsiaPac Advanced TCAD Univ. Bundle (單一軟體, 可彈性使用, 無論是二維元件或一維奈米線).

●重現前面所述 Dr. Mouis, Dr. Barraud, Prof. Jungemann, Dr. Palestri, Dr. Eminente, Dr. Zilli, Prof. Fischetti 等人所發表的 2-D NanoFETs 蒙地卡羅模擬的微觀下非平衡現象.

●提供建立非平衡下 2-D NanoFETs Backscattering 架構所需要的微觀物理資訊.

(2) 建立非平衡下 2-D NanoFETs Backscattering 架構

●運用自行研發內含 rc 的 **TRP** (triangular potential approximation; 我們在 2000 IEDM (p. 679) 發表的論文就使用 **TRP**; 目前已校正，精確度與 Schrodinger-Poisson simulator **Shred** 一致; Available: https://www.nanohub.org) MOS simulator, 可求得 barrier 頂處 subbands and Fermi level 等 degenerate details, 進而得到 Qinv, Fermi-Dirac vinj, and the Fermi factor for both l and λ.

●將 velocity overshoot effect 植入我們已發表的 l compact model (M. J. Chen, et al., TED, p. 1265, 2008), 作法是在 high-field region near drain 解開 1-D Poisson equation 即得.

●Source non-equilibrium 與其 near-equilibrium 之間的 criterion 可以一 3-D source 與 2-D gas 以界面為主的系統運用機率統計方法決定，因之，2-D Backscattering 架構可以分解成 2 cases: source starvation and source near-equilibrium.

● 引用我們發表的 Esaki tunneling model (M. J. Chen, et al., EDL, p. 134, 1998) 加入 2-D Backscattering 架構 (Prof. Lundstrom 原只探討 thermal injection) 以處理 tunneling (field injection) across the source-channel barrier.

(3) 修正或微調非平衡下 2-D NanoFETs Backscattering 架構

●運用蒙地卡羅元件模擬軟體模擬在 2-D NanoFETs 不同元件結構參數，不同材料物理參數，不同操作偏壓溫度條件等，以逐步修正或微調前述非平衡下 2-D Backscattering 架構（特別能額外提供 DIBL 等重要參數).

●運用工業界提供的12吋晶圓級測試鍵量測在不同元件結構參數，不同製程參數，不同操作偏壓溫度條件等，以逐步修正或微調前述非平衡下 2-D Backscattering 架構. 量測方法將引用我們已發表者 (M. J. Chen, et al., TED, p. 1409, 2004).

●亦將大量引用文獻上公開發表的相關實驗或模擬數據以徹底修正或微調前述非平衡下 2-D Backscattering 架構.

●預期精確度至通道長度十奈米左右，可由元件蒙地卡羅模擬軟體決定之.

(4) 2-D Schrodinger and 2-D Poisson self-consistent 求解程式撰寫及除錯

● 普度大學所提供的 MOS（垂直於 channel 方向）1-D Schrodinger and 1-D Poisson simulator **Schred** 物理上不能使用於 1-D Nanowire FET 分析之用. 垂直於 1-D Nanowire channel 方向的 2-D Schrodinger- 2-D Poisson 且考慮了 rc 的 self-consistent 求解程式目前尚未在文獻公開發表，故時機上十分緊迫.

● 預期 2-D Schrodinger and 2-D Poisson self-consistent 求解程式可得到 1-D Nanowire FET source-channel potential barrier 頂處 subbands and Fermi level 等 degenerate details, 進而得出頂處 Qinv, Fermi-Dirac vinj, and the Fermi factor for both l and λ.

成果:

(1) 實驗成功分離出 Scattering by Plasmons in Gate Polysilicon 以及 Scattering by Plasmons in source/drain 因而使得非平衡下 2-D NanoFETs Backscattering 架構之建立成為可能.

(2) 並已實驗成功分離出 Scattering by Soft Optical Phonons in SiO_2 and High-k Dielectrics.

(3) 完成 2-D Backscattering 架構核心: self-consistent quantum simulator.

四、結論與討論

● 幾近 100% 完成預定改進之 Backscattering 架構，發表 IEEE T-ED 期刊論文七篇及 EDL 期刊論文二篇，另有多篇陸續投稿中。

● 本計畫畢業許智育博士，TSMC 2011 年研發替代役參與 10- and 7-nm FinFET R&D 團隊。

● 兩位博士班，李建志和李韋漢，2012 年 11 月將畢業，已獲 2012 年 TSMC 研發替代役，將分別加入 14-nm FinFET R&D 團隊和 10- and 7-nm FinFET R&D 團隊。

● 本計畫亦已畢業眾多碩士畢業生，在TSMC 服務，均參與 FinFET R&D 及 TCAD 團隊。

參考文獻:

1. C. Y. Hsu, C. C. Lee, Y. T. Lin, C. Y. Hsieh, and M. J. Chen, "Enhanced hole gate direct tunneling current in process-induced uniaxial compressive stress p-MOSFETs," *IEEE Trans. Electron Devices*, vol. 56, pp. 1667-1673, August 2009.

2. D. W. Lin, M. L. Cheng, S. W. Wang, C. C. Wu, and M. J. Chen, "A novel method of MOSFET series resistance extraction featuring constant mobility criteria and mobility universality," *IEEE Trans. Electron Devices*, vol. 57, pp. 890-897, April 2010.

3. D. W. Lin, C. L. Chen, M. J. Chen, and C. C. Wu, "An extreme surface proximity-push for embedded-SiGe in pMOSFETs featuring self-aligned silicon-reflow (SASR)," *IEEE Electron Device Letters*, vol. 31, pp. 924-926, Sept. 2010.

4. W. H. Lee and M. J. Chen, "Gate direct tunneling current in uniaxially compressive strained nMOSFETs: A sensitive measure of electron piezo effective mass," *IEEE Trans. Electron Devices*, vol. 58, pp. 39-45, Jan. 2011.

5. M. J. Chen, C. C. Lee, and K. H. Cheng, "Hole effective masses as a booster of self-consistent six-band k • p simulation in inversion layers of pMOSFETs," *IEEE Trans. Electron Devices*, vol. 58, pp. 931-937, April 2011.

6. C. Y. Hsu, H. G. Chang, and M. J. Chen, "A method of extracting metal-gate high-k material parameters featuring electron gate tunneling transition," *IEEE Trans. Electron Devices*, vol. 58, pp. 953-959, April 2011.

7. M. J. Chen, S. C. Chang, S. J. Kuang, C. C. Lee, W. H. Lee, K. H. Cheng, and Y. H. Zhan, "Temperature-dependent remote-Coulomb-limited electron mobility in n+ polysilicon ultrathin gate oxide nMOSFETs," *IEEE Trans. Electron Devices*, vol. 58, pp. 1038-1044, April 2011.

8. M. J. Chen, L. M. Chang, S. J. Kuang, C. W. Lee, S. H. Hsieh, C. A. Wang, S. C. Chang, and C. C. Lee, "Temperature-oriented mobility measurement and simulation to assess surface roughness in ultrathin-gate-oxide (1 nm) nMOSFETs and Its TEM evidence," *IEEE Trans. Electron Devices*, vol. 59, pp. 949-955, April 2012.

9. M. J. Chen and W. H. Lee, "Evidence for the fourfold-valley confinement electron piezo-effective-mass coefficient in Inversion layers of <110> uniaxial tensile strained (001) nMOSFETs," *IEEE Electron Device Letters*, vol. 33, pp. 755-757, June 2012.

10. Chien-Chih Lee, "A New In-House Fast Sophisticated Quantum Simulator for Silicon Hole Band Structure and Inversion-Layer Mobility under Three-Dimensional GPa-Level Stresses," *Ph.D. Dissertation*, National Chiao Tung University, November 2012.

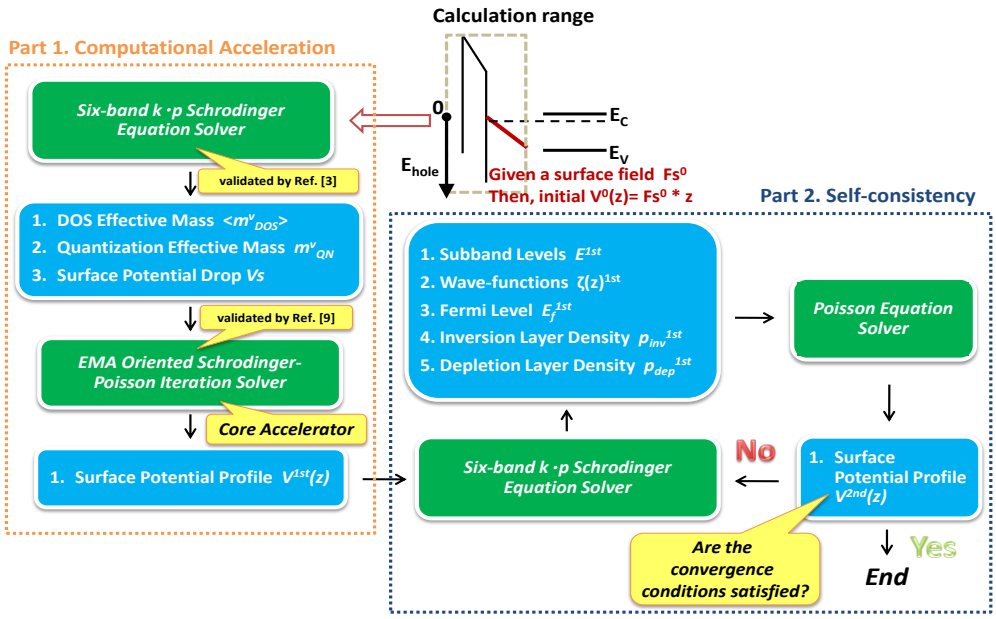
11. Wei-Han Lee, “Sophisticated Quantum Computation on n-type Inversion Layers: Strain, Subband, Mobility, and 3-D Structure,” *Ph.D. Dissertation*, National Chiao Tung University, November 2012.

Important Figures

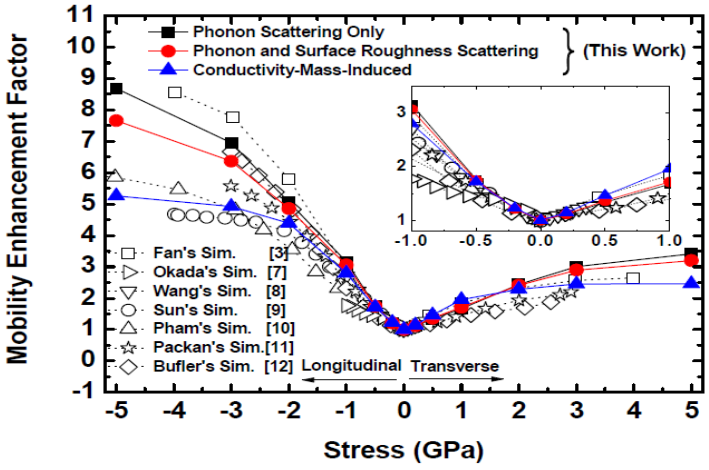
計畫至目前為止產生成果如下(包括已發表及尚在規劃發表者):

- **下世代元件量子力學模擬器 NEP**

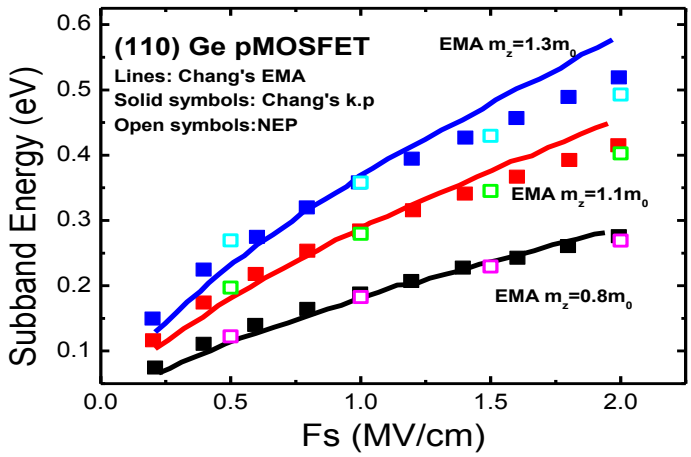
  NEP 是我們實驗室簡稱，英文全名為 Nano Electronics Physics，此模擬器架構流程如下:



此模擬器優異之處為可大幅度縮短 CPU 計算時間 更進一步整合自行建立的遷移率模擬器，整合後的模擬器已能精確計算不同應變及不同基座方位下電子及電洞之遷移率，此可從下面的電洞遷移率計算結果(尚未發表)得知:

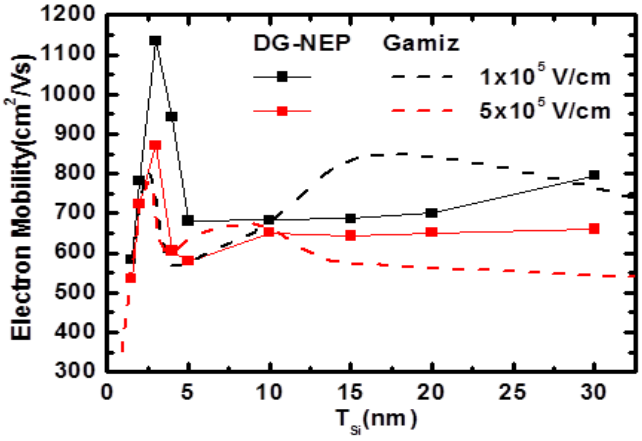


也能精確處理不同通道材料之計算，如下所示(尚未發表):



引用比較來源為 Shu-Tong Chang, in the title of "Subband structure and effective mass of relaxed Strained Ge (110) pMOSFETs" Solid-state physics, in January 2011.

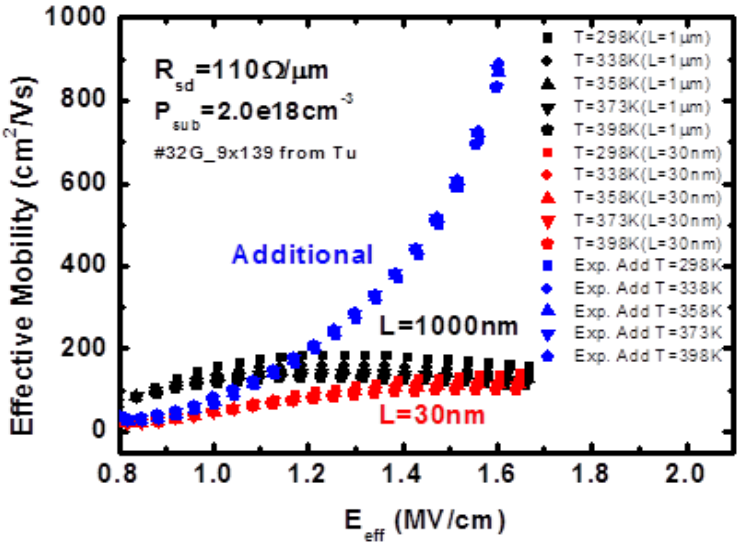
模擬器並已延伸至 Double-Gate 結構，即 FinFET 的 sidewall 結構，結果如下所示(尚未發表):



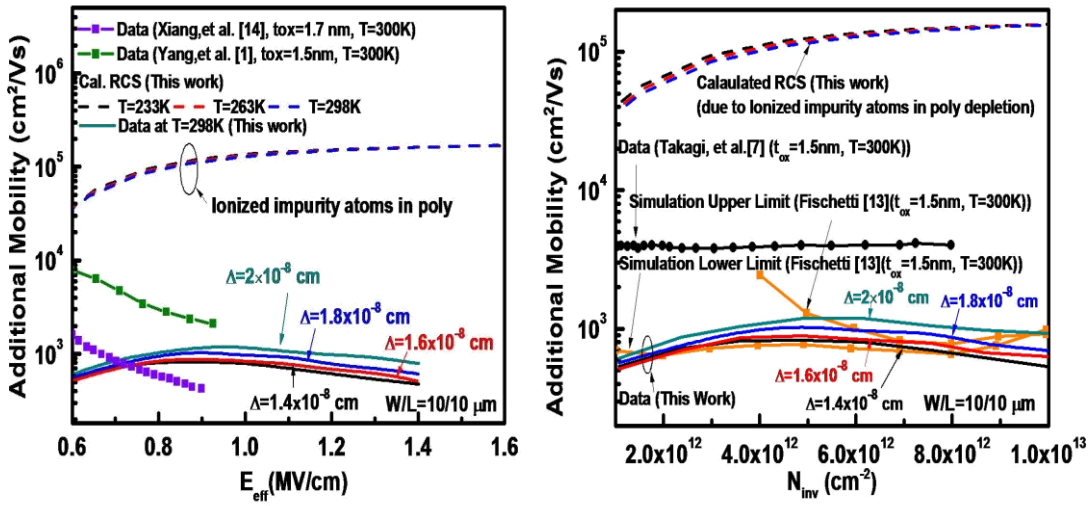
由此圖可見與 Monte Carlo simulation 比較趨勢十分符合，所見誤差係未考慮 wavefunction penetration 所致。此處 DG-NEP 為我們 NEP 模擬器 Double-Gate 版本。引用來源比較為 F. Gamiz and M. V. Fischetti, "Monte Carlo simulation of double-gate silicon-on-insulator inversion layers: The role of volume inversion, " *J. Appl. Phys.*, vol. **89**, no. 10, pp. 5478–5487, May 2001。

- **通道非平衡物理現象之實驗證據及相關模式建立**

我們已進行許久的通道長度從一微米至十幾奈米(已扣掉 overlap 部分)的實驗量測(包含 R_{sd})，且用到 TCAD 校正微調(主要 doping profile)，使得產出與實驗相符(from subthreshold to above-threshold and temperature dependencies)。萃取遷移率以及通道散射參數過程非常小心，我們也將一些不確定的因素給一些實驗誤差，最後我們世界首次獲取了通道非平衡下物理現象之實驗證據(source/drain plasmons) 。如下圖所示(尚未發表)，此現象在通道長度從一微米減至二十幾奈米時產生的效應就很明顯(此圖的 L 即為 gate length，扣掉 overlap 部分後就是 channel length 大約二十幾奈米):

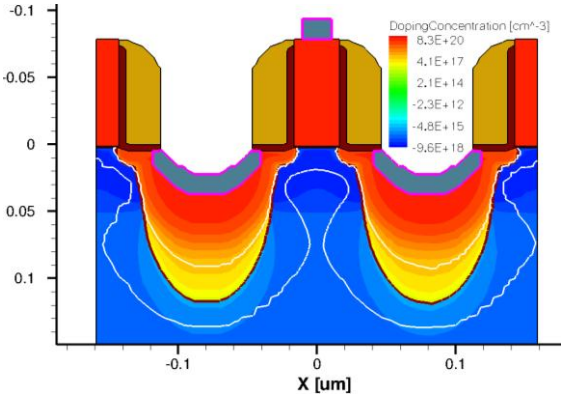


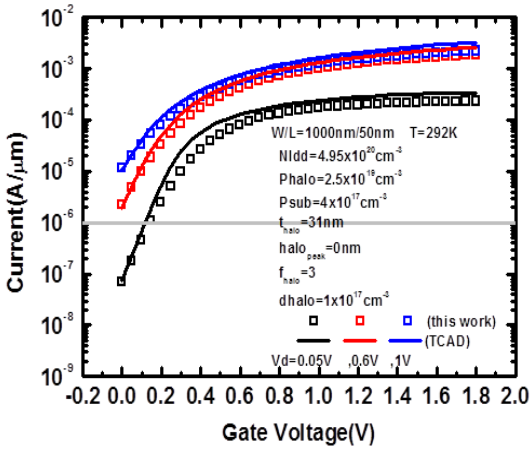
值得一提的是，我們也成功實驗萃取 interface plasmons in poly gate:



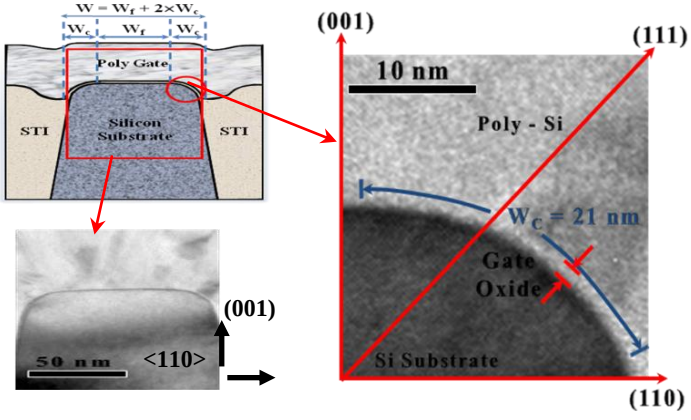
- **TCAD (Technology Computer-Aided Design)**

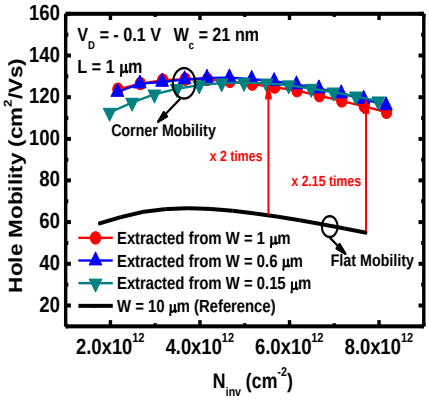
當進行上述曠時費久項目時，我們仍積極為進行下世代奈米元件之高等模擬作一準備。我們使用了 Sentaurus 軟體，已針對 extended planar bulk MOSFET 進行了 doping profiles 之 calibration，部份成果如下(尚未發表):





我們也進行 corner structure 的 TCAD 模擬， 有曲度的 corner structure 對未來奈米元件很重要，我們已針對此特殊結構作一深入量測及參數萃取。相應重點的實驗及數據，圖示於下:





附錄 （僅列兩篇發表論文）

Temperature-Oriented Mobility Measurement and Simulation to Assess Surface Roughness in Ultrathin-Gate-Oxide (~1 nm) nMOSFETs and Its TEM Evidence

Ming-Jer Chen, *Senior Member, IEEE*, Li-Ming Chang, Shin-Jiun Kuang, Chih-Wei Lee, Shang-Hsun Hsieh, Chi-An Wang, Sou-Chi Chang, and Chien-Chih Lee, *Student Member, IEEE*

***Abstract*—On a 1.27-nm gate-oxide nMOSFET, we make a comprehensive study of SiO_2/Si interface roughness by combining temperature-dependent electron mobility measurement, sophisticated mobility simulation, and high-resolution transmission electron microscopy (TEM) measurement. Mobility measurement and simulation adequately extract the correlation length λ and roughness rms height Δ of the sample, taking into account the Coulomb-drag-limited mobilities in the literature. The TEM measurement yields the apparent correlation length λ_m and roughness rms height Δ_m. It is found that the following hold: 1) $\lambda \approx \lambda_m$ for both the Gaussian and exponential models, validating the temperature-oriented extraction process; 2) the extracted Δ (~1.3 Å for the Gaussian model and 1.0 Å for the exponential one) is close to that (~1.2 Å) of Δ_m, all far less than the conventional values (~3 Å) in thick-gate-oxide case; and 3) the TEM 2-D projection correction coefficient Δ_m/Δ is approximately 1.0, which cannot be elucidated with the current thick-gate-oxide-based knowledge.**

***Index Terms*—Coulomb drag, gate oxide, interface plasmons, mobility, metal–oxide–semiconductor field-effect transistors (MOSFETs), scattering, surface roughness, transmission electron microscopy (TEM), universal mobility.**

I. Introduction

RANDOM roughness at the SiO_2/Si interface can critically affect the carrier transport in the inversion layers of MOSFETs. Thus, attempts to characterize the surface roughness parameters are essential. To facilitate the description of the surface roughness picture, two fundamentally distinct models were proposed [1]: Gaussian and exponential. Both models contain two elements, i.e., correlation length λ and roughness rms height Δ. To assess the underlying λ and Δ, numerical simulations were applied with the inversion-layer mobility data as inputs [1]–[4]. However, for MOSFETs having ultrathin gate oxides, such a scheme encountered difficulties due to the increasing importance of remote scatterers [5], [6]. Indeed, how to correctly distinguish between the surface roughness and remote scatterers has been a challenging issue [7]. To overcome the issue, we recently proposed a temperature-dependent extraction method [8]. As has been demonstrated [8] on 1.65-nm-gate-oxide nMOSFETs through the use of the Gaussian model, the merits of the method are summarized: 1) it can effectively distinguish the surface-roughness-limited mobility from the remote-scatterer-limited mobility; 2) it can accurately determine the surface roughness parameters; and 3) it can easily be conducted in a certain range around room temperature.

Alternatively, TEM measurements can be performed to provide the apparent correlation length λ_m and the apparent rms height Δ_m. To address the 2-D projection effect in the TEM measurement, a link to the aforementioned Δ was established in terms of correction coefficient Δ_m/Δ [1]. Furthermore, a mathematical transformation from digitized TEM surface roughness data, without directly accounting for λ and Δ, to mobility values was developed, thus producing a quantity associated with the projection correction [9]. However, these works [1], [9] were devoted to the thick-gate-oxide samples only.

To make a comprehensive surface roughness study and to advance the studies [1]–[9], in this paper, we integrate the aforementioned means, i.e., temperature-dependent mobility measurement, sophisticated mobility simulation accounting for both the Gaussian and exponential models, and TEM measurement. In addition, a thinner gate-oxide sample (1.27 nm), which enables fair citation of the simulated Coulomb-drag-limited mobilities (on 1-nm gate oxide) [5], is presented. The results are novel and might substantially improve current understanding of the surface roughness, particularly for the case of ultrathin gate oxides.

II. Sample and Electrical Characterization

The device under study was fabricated in a conventional manufacturing process. In this process, SiO_2 film was thermally grown on the (001) surface, followed by NO annealing. Corresponding process parameters can essentially be obtained by

Manuscript received October 24, 2011; accepted December 29, 2011. Date of publication February 3, 2012; date of current version March 23, 2012. This work was supported by the National Science Council of Taiwan under Contract NSC 98-2221-E-009-164-MY3. The review of this paper was arranged by Editor M. Ieong.

M.-J. Chen, L.-M. Chang, C.-W. Lee, S.-H. Hsieh, C.-A. Wang, and C.-C. Lee are with the Department of Electronics Engineering and Institute of Electronics, National Chiao-Tung University, Hsinchu 300, Taiwan (e-mail: chenmj@faculty.nctu.edu.tw).

S.-J. Kuang is with the Taiwan Semiconductor Manufacturing Company, Hsinchu 300, Taiwan.

S.-C. Chang is with the Department of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta, GA 30332-0250 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.2012.2182771

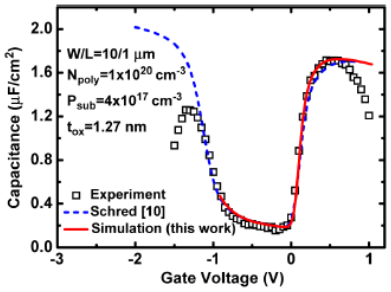


Fig. 1. Comparison of the measured (symbol) and simulated (lines) gate capacitance versus gate voltage. The lines came from the self-consistent Schrödinger and Poisson's equations solvers [10], [11].

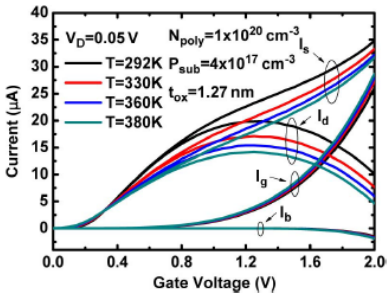


Fig. 2. Measured terminal currents at a drain voltage of 0.05 V versus gate voltage for four different temperatures.

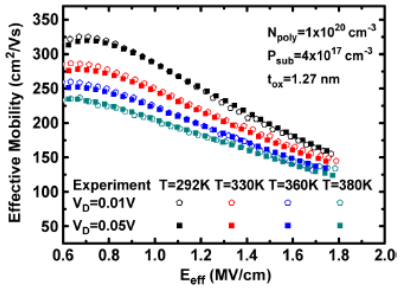


Fig. 3. Measured electron effective mobility at two drain voltages of 0.01 and 0.05 V versus vertical effective field for four temperatures.

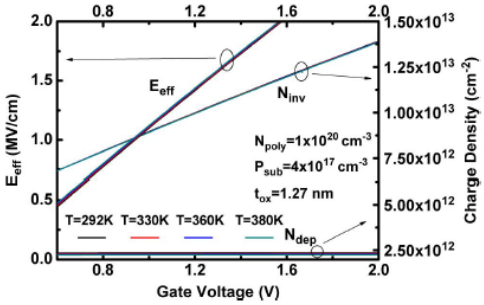


Fig. 4. Simulated vertical effective field, inversion-layer charge density, and substrate depletion charge density versus gate voltage with the temperature as a parameter.

fitting the measured gate capacitance versus the gate voltage (C_g–V_g), as shown in Fig. 1. This was realized with the use of a self-consistent Schrödinger and Poisson's equations solver. Two such solvers were cited: one named Schred [10] and the other in the previous work [11]. Obviously, the two sources [10] and [11] are consistent of each other in the data fitting. Although Cg-Cg data at high gate voltages were seriously distorted due to the prepared ultrathin gate oxide, where the direct tunneling current is profoundly large, the fitting was successfully done in the remaining regions, leading to n^+ polysilicon doping concentration $= 1 \times 10^{20}$ cm^{-3}, gate oxide (SiO_2) physical thickness $= 1.27$ nm, and p-type substrate doping concentration $= 4 \times 10^{17}$ cm^{-3}. As will be explained in detail later, the NO annealing used may have an impact on the SiO_2/Si interface formation but not the SiO_2 bulk one.

The ratio of channel width W to length L of the device is 1/1 μm. The channel length direction is along the $\langle 110 \rangle$ direction. We conducted I–V measurements at four temperatures (292, 330, 360, and 380 K). Measured I–V curves across different positions on wafer were found to be comparable with each other. This ensures the integrity of the presented sample. Fig. 2 shows the measured drain current I_d, source current I_s, gate current I_g, and bulk current I_b at drain voltage $V_D =$ 50 mV, plotted versus gate voltage with the temperature as a parameter. The effect of the huge gate tunneling current on the source and drain currents is evidently clear. Similar behaviors were also observed elsewhere [12]. In this situation, the correct mobility assessment should be formulated as [12]

$$\mu(V_g) = \frac{L}{W} \frac{(I_s(V_g) + I_d(V_g))}{2V_d} \frac{1}{qN_{\text{inv}}(V_g)}. \quad (1)$$

Resulting temperature-dependent mobilities are given in Fig. 3 and plotted versus vertical effective field E_{eff}. Here, E_{eff} followed the well-known expression

$$E_{\text{eff}} = \frac{q(0.5N_{\text{inv}} + N_{\text{dep}})}{\varepsilon_{\text{si}}} \quad (2)$$

where N_{inv} is the inversion-layer charge density, N_{dep} is the substrate depletion charge density, and ε_{si} is the silicon permittivity. With the aforementioned process parameters as inputs, the self-consistent Schrödinger and Poisson's equations solver [11] was executed to furnish N_{inv} and N_{dep}. Corresponding N_{inv}, N_{dep}, and E_{eff} are plotted in Fig. 4 versus gate voltage, with the temperature as a parameter. Moreover, we repeated the case of $V_D = 10$ mV and found that the change is little, as shown in Fig. 3. This ensures the quality of the presented mobility data, particularly for their temperature dependencies.

III. Simulation and Extraction

In the mobility simulation, we employed the self-consistent Schrödinger and Poisson's equations solver [11] to deliver

subbands and wavefunctions while computing the total mobility μ_{total}. Here, μ_{total} was limited to the high-E_{eff} region, where the microscopic scattering by acoustic and optical phonons in the channel region and by the SiO_2/Si surface roughness dominates. Two literature sources were quoted concerning the scattering rate formalisms: one for the phonon scattering rate [2] and the other for both the Gaussian and exponential surface roughness ones [13]. The material parameters used in the phonon scattering rate calculation were the same as in previous work [8]. These two surface roughness scattering rate formalisms are given as follows [13]:

$$\frac{1}{\tau_{\text{SR}}^{i}(E)} = \frac{m_{\text{dos}}^{i} e^2 E_{\text{eff}}^2 \Delta^2 \lambda^2}{2\hbar^3} \int_0^{2\pi} \exp\left(-\frac{q^2\lambda^2}{4}\right)(1-\cos\theta)\, d\theta \tag{3}$$

for the Gaussian model and

$$\frac{1}{\tau_{\text{SR}}^{i}(E)} = \frac{m_{\text{dos}}^{i} e^2 E_{\text{eff}}^2 \Delta^2 \lambda^2}{2\hbar^3} \int_0^{2\pi} \frac{1}{[1+(\lambda^2 q^2/2)]}(1-\cos\theta)\, d\theta \tag{4}$$

for the exponential model. In (3) and (4), τ_{SR}^{i} is the scattering rate of subband i, m_{dos}^{i} is the density-of-states effective mass of subband i, θ is the scattering angle, and $q^2 = 2k^2(1-\cos\theta)$, with $k^2 = 2m_{\text{dos}}^{i}(E-E_i)/2\hbar^2$. Simulated total mobility with different Δ's and different λ's were used to reproduce universal mobility data [14], [15] in a temperature range comparable with the measurement one in this work, as shown in Figs. 5 and 6 for the Gaussian and exponential models, respectively. The best fitting produces the solutions of Δ and λ, as shown in Fig. 7 for both models. Other values of Δ and λ only led to a poor fitting of the temperature dependencies and had been ruled out. It can be seen from Fig. 7 that Δ is considerably constant and is higher for the Gaussian model, λ has a broad range and is smaller in magnitude for the exponential model, and there is no overlap between the two models. Thus, in the subsequently analysis, λ will be fixed at the middle value, i.e., 14.9 and 23.2 Å for the Gaussian and exponential models, respectively. Accordingly, the values of Δ are 3.1 and 2.7 Å, respectively, which are close to the published values in the thick-gate-oxide case [1]–[4]. This dictates the applicability of the proposed temperature-dependent extraction process.

However, we found that, with the aforementioned Δ and λ values, simulated results cannot match all observed temperature dependencies of mobilities in a 1.27-nm sample. As shown in Figs. 8 and 9 for both models, simulated mobilities at 292, 330, and 360 K fall below data points, despite the apparent coincidence for the remaining temperature (380 K). Since this is physically unreasonable, the value of Δ must be reduced. The case of reducing Δ to lower values is depicted in these two figures.

The required reduction in Δ indicates the existence of the remote scatterers. In this sense, the additional mobility μ_{add} due to the remote scatterers, ionized substrate impurity atoms,

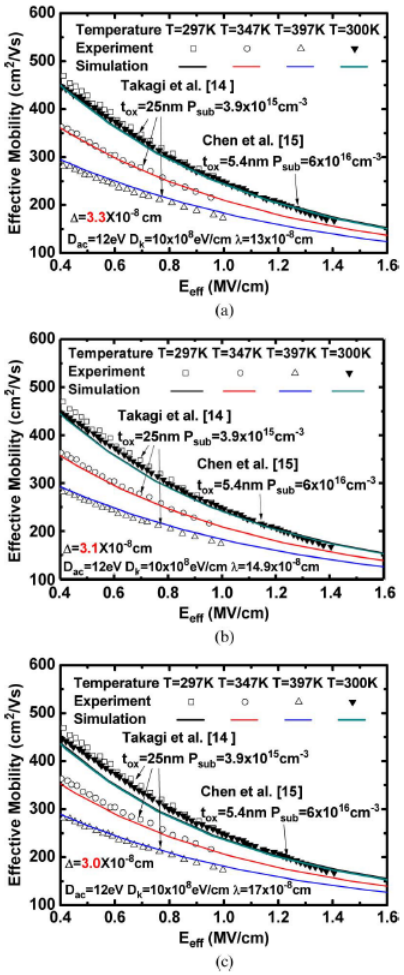


Fig. 5. Comparison of the experimental electron universal mobility curves for (open symbols) three temperatures [14] and (filled symbols) one temperature [15] with the (lines) simulated ones in this work for the Gaussian model. (a) $\Delta = 3.3$ Å, and $\lambda = 13$ Å. (b) $\Delta = 3.1$ Å, and $\lambda = 14.9$ Å. (c) $\Delta = 3.0$ Å, and $\lambda = 17$ Å. D_{ac} is the acoustic deformation potential. D_k is the deformation potential of the kth intervalley phonon.

and interface traps can be defined according to Matthiessen's rule, i.e.,

$$\frac{1}{\mu_{\text{eff}}} = \frac{1}{\mu_{\text{total}}} + \frac{1}{\mu_{\text{add}}} \tag{5}$$

where μ_{eff} is the measured mobility [i.e., (1)], and μ_{total} is the simulated total mobility that does not include the contributions by remote scatterers, substrate impurity atoms, or interface traps. Note that Matthiessen's rule can empirically apply as long as the undertaken E_{eff} or N_{inv} is high enough [16]. The effect of substrate impurity atoms or interface traps

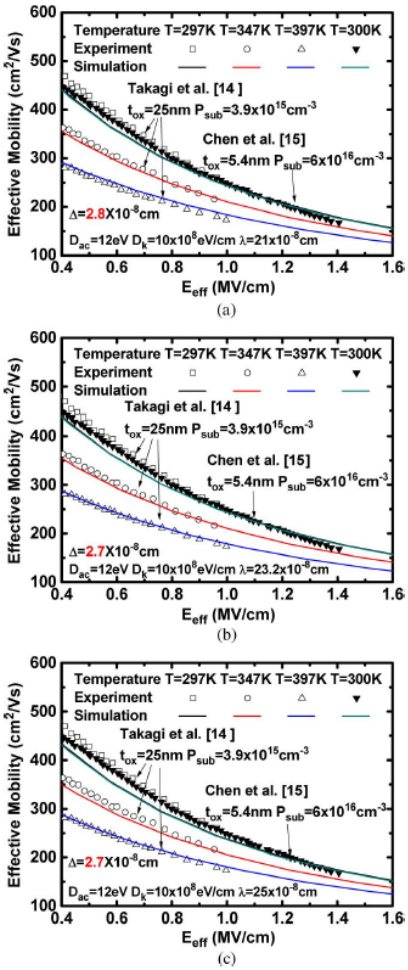


Fig. 6. Comparison of the experimental electron universal mobility curves for (open symbols) three temperatures [14] and (filled symbols) one temperature [15] with the (lines) simulated ones for the exponential model. (a) $\Delta = 2.8$ Å, and $\lambda = 21$ Å. (b) $\Delta = 2.7$ Å, and $\lambda = 23.2$ Å. (c) $\Delta = 2.7$ Å, and $\lambda = 25$ Å.

can also be suppressed in the high-E_{eff} region. In Fig. 10, we show the extracted μ_{add} at N_{inv} of 1×10^{13} cm^{-2}, with Δ as a parameter, along with the two separate curves for the simulated Coulomb-drag-limited mobilities [5]. These two Coulomb-drag-limited mobility curves represent two limiting conditions of the Landau-damping wave vector: one for "zero-temperature Landau damping wave vector" and the other for "damping at Thomas Fermi screening wave vector" [5]. According to Fischetti [5], the realistic Coulomb-drag-limited mobility is likely to be situated between the two limits. In this sense, the actual Δ of the presented sample should be approximately 1.3 and 1.0 Å for the Gaussian and exponential models, respectively.

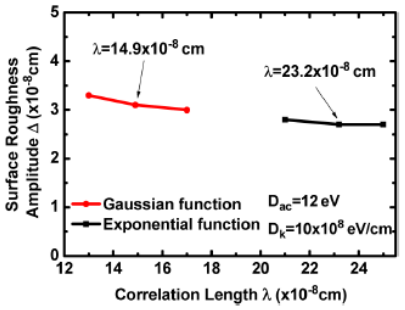


Fig. 7. Best-fitting results of Δ versus λ for both the Gaussian and exponential models.

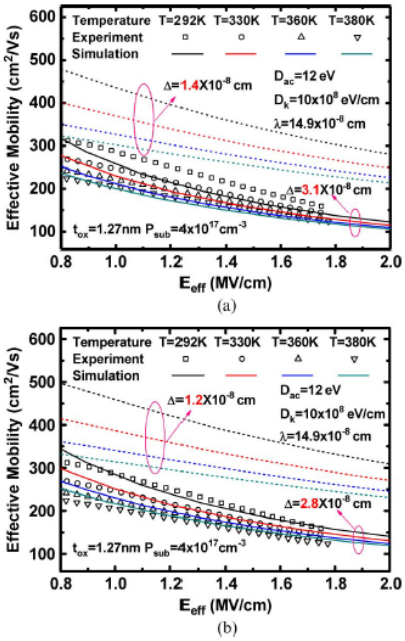


Fig. 8. Comparison of the (symbols) temperature-dependent electron effective mobility data with the simulated mobility curves using the Gaussian model for (a) Δ of (dashed line) 1.4 Å and (solid line) 3.1 Å, and (b) Δ of (dashed line) 1.2 Å and (solid line) 2.8 Å. $\lambda = 14.9$ Å.

IV. TEM Measurement and Analysis

In performing TEM measurements, the sample was 1000 nm long, and its cross-section thickness had a range of 20–60 nm. A TEM image was created, as shown in Fig. 11. In this picture, the labeled ⟨110⟩ direction is parallel to the SiO_2/Si interface, whereas the ⟨001⟩ direction is normal to the interface. Underlying λ_m and Δ_m were determined by following the work of Goodnick *et al.* [1]. First, digitalization of surface roughness was done by directly counting lattice points on the TEM photographs. The same sampling interval (1.92 Å) [1] was used. This led to the autocovariance function

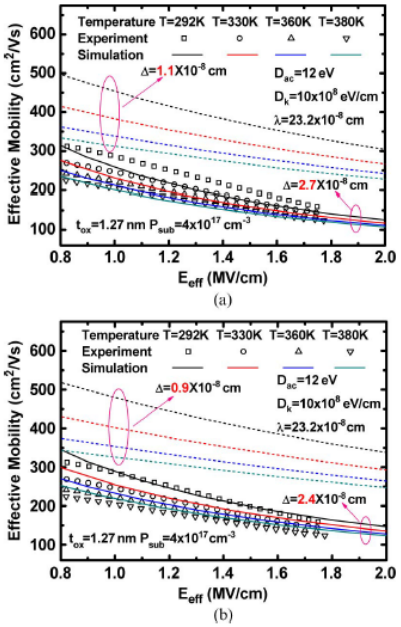


Fig. 9. Comparison of the (symbols) temperature-dependent electron effective mobility data with the simulated mobility curves using the exponential model for (a) Δ of (dashed line) 1.1 Å and (solid line) 2.7 Å and (b) Δ of (dashed line) 0.9 Å and (solid line) 2.4 Å. $\lambda = 23.2$ Å.

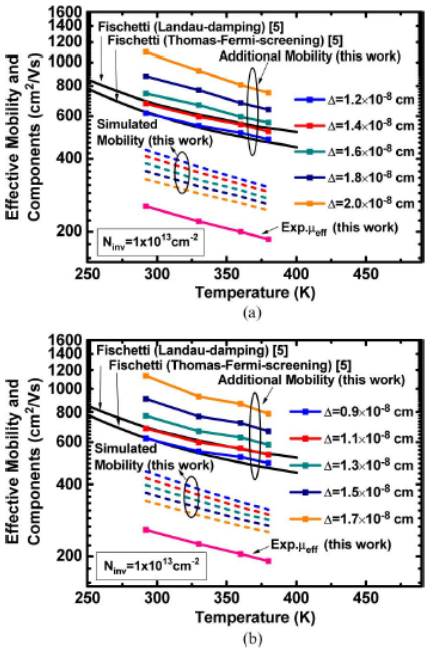


Fig. 10. Temperature dependence of the measured effective mobility, simulated mobility, and extracted additional mobility, all made at a fixed inversion-layer density of 1×10^{13} cm^{-2} to make a fair comparison with simulated interface-plasmons-limited mobility [5]. Simulated mobility and additional mobility are presented as a function of Δ. (a) Gaussian model. (b) exponential model.

in Fig. 12 as a function of the distance. In the figure, the squared root of the autocovariance at zero distance yields Δ_m of around 1.2 Å. Data fitting was carried out, producing $\lambda_m =$ 15.6 and 23.1 Å for the Gaussian and exponential models, respectively.

Straightforwardly, several key arguments can be drawn. First, the sample length effect [1] may be neglected due to a long sample used. Thus, the ratio of $\lambda_m/\lambda \approx 1.0$ is reached accordingly [1]. Strikingly, this is the fact because the extracted values of $\lambda_m = 15.6$ and 23.1 Å in this work separately are close to those (14.9 and 23.2 Å) of λ. One of the arguments can thereby be inferred: the sample length effect as cited in [1] can act to be corroborating evidence for the temperature-dependent extraction method. Second, the TEM 2-D projection correction coefficient Δ_m/Δ is approximately 1.0, valid for both the Gaussian and exponential models. However, the current thick-gate-oxide-based knowledge [1] cannot reasonably explain this because, for the TEM sample cross-section thickness range of 20–60 nm as in our work, the theoretical calculation pointed out [1] that the upper limit of Δ_m/Δ decreases sharply from 0.7 (see [1]). Even the Δ_m/Δ in this work is much higher than the published experimental Δ_m/Δ values of 0.50–0.71 [8]. We attribute such significant deviations to the sole use of the ultrathin gate oxide (1.27 nm) in this work. It is therefore suggested that further theoretical investigation of the TEM sample cross-section thickness effect is needed, particularly for the ultrathin-gate-oxide situation.

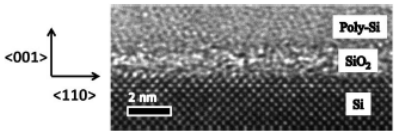


Fig. 11. TEM image of the sample.

Finally, we want to stress that the physical gate oxide thickness can be estimated from the TEM picture in Fig. 11, and it appears to be in proximity of 1.27 nm, which is the value of Cg-Vg fitting. The same gate oxide thickness had earlier been applied to p-type counterparts on the same wafer [17]. This strongly suggests that the presented SiO_2 bulk film was less nitrided during the NO annealing. Moreover, according to the literature [4], the use of the oxide nitridation process will produce a smoother Si/SiO_2 interface in terms of a reduced Δ. Extra traps may be created during the pure nitridation or mixed one (i.e., NO annealing process); however, their effect might be limited to the low-E_{eff} region, rather than the high-E_{eff} region where this work was focused on. Therefore, it is argued that the mixed NO annealing used may have an impact on the SiO_2/Si interface formation but not the SiO_2 bulk one.

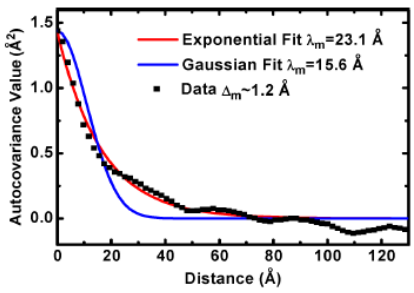


Fig. 12. Autocovariance function extracted from the TEM photograph as a function of distance. Both the Gaussian and exponential fits are shown.

V. Conclusion

Electron mobility of a 1.27-nm gate-oxide nMOSFET sample has been measured at various temperatures. Temperature-dependent numerical simulation has for the first time transformed existing universal mobility data into the solutions of surface roughness parameters and has been applied to the presented sample, taking into account the available Coulomb-drag-limited mobilities. The extracted surface roughness parameters have been correlated with the values from TEM measurements. As a consequence, novel results have been created and might substantially improve the current understanding of interface roughness, particularly for the case of ultrathin gate oxides.

References

[1] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface roughness at the Si(100) − SiO_2 interface," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 32, no. 12, pp. 8171–8186, Dec. 1985.

[2] S. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, "Comparative study of phonon-limited mobility of two-dimensional electrons in strained and unstrained Si metal-oxide-semiconductor field-effect transistors," *J. Appl. Phys.*, vol. 80, no. 3, pp. 1567–1577, Aug. 1996.

[3] G. Mazzoni, A. L. Lacaita, L. M. Perron, and A. Pirovano, "On surface roughness-limited mobility in highly doped n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 46, no. 7, pp. 1423–1428, Jul. 1999.

[4] A. Pirovano, A. L. Lacaita, G. Ghidini, and G. Tallarida, "On the correlation between surface roughness and inversion layer mobility in Si-MOSFETs," *IEEE Electron Device Lett.*, vol. 21, no. 1, pp. 34–36, Jan. 2000.

[5] M. V. Fischetti, "Long-range Coulomb interactions in small Si devices. Part II: Effective electron mobility in thin-oxide structures," *J. Appl. Phys.*, vol. 89, no. 2, pp. 1232–1250, Jan. 2001.

[6] F. Gámiz and M. V. Fischetti, "Remote Coulomb scattering in metal-oxide-semiconductor field effect transistors: Screening by electrons in the gate," *Appl. Phys. Lett.*, vol. 83, no. 23, pp. 4848–4850, Dec. 2003.

[7] M. V. Fischetti, T. P. ORegan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2116–2136, Sep. 2007.

[8] M. J. Chen, S. C. Chang, S. J. Kuang, C. C. Lee, W. H. Lee, K. H. Cheng, and Y. H. Zhan, "Temperature-dependent remote-Coulomb-limited electron mobility in n+-polysilicon ultrathin gate oxide nMOSFETs," *IEEE Trans. Electron Devices*, vol. 58, no. 4, pp. 1038–1044, Apr. 2011.

[9] Y. Zhao, H. Matsumoto, T. Sato, S. Koyama, M. Takenaka, and S. Takagi, "Comprehensive understanding of surface roughness limited mobility in unstrained- and strained-Si MOSFETs by novel characterization scheme of Si/SiO_2 interface roughness," in *VLSI Symp. Tech. Dig.*, 2009, pp. 18–19.

[10] Schred. [Online]. Available: http://nanohub.org/resources/schred

[11] M. J. Chen, C. C. Lee, and K. H. Cheng, "Hole effective masses as a booster of self-consistent six-band $k \cdot p$ simulation in inversion layers of pMOSFETs," *IEEE Trans. Electron Devices*, vol. 58, no. 4, pp. 931–937, Apr. 2011.

[12] S. Takagi and M. Takayanagi, "Experimental evidence of inversion-layer mobility lowering in ultrathin gate oxide metal-oxide-semiconductor field-effect-transistors with direct tunneling current," *Jpn. J. Appl. Phys.*, vol. 41, no. 4B, pp. 2348–2352, Apr. 2002, pt. 1.

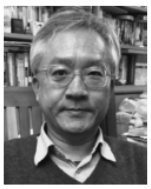
[13] S. Yamakawa, H. Ueno, K. Taniguchi, C. Hamaguchi, K. Miyatsuji, K. Masaki, and U. Ravaioli, "Study of interface roughness dependence of electron mobility in Si inversion layers using the Monte Carlo method," *J. Appl. Phys.*, vol. 79, no. 2, pp. 911–916, Jan. 1996.

[14] S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in Si MOSFETs: Part I—Effects of substrate impurity concentration," *IEEE Trans. Electron Devices*, vol. 41, no. 12, pp. 2357–2362, Dec. 1994.

[15] K. Chen, C. Hu, P. Fang, M. R. Lin, and D. L. Wollensen, "Predicting CMOS speed with gate oxide and voltage scaling and interconnect loading effects," *IEEE Trans. Electron Devices*, vol. 44, no. 11, pp. 1951–1957, Nov. 1997.

[16] F. Stern, "Calculated temperature dependence of mobility in Si inversion layers," *Phys. Rev. Lett.*, vol. 44, no. 22, pp. 1469–1472, Jun. 1980.

[17] C. Y. Hsu, C. C. Lee, Y. T. Lin, C. Y. Hsieh, and M. J. Chen, "Enhanced hole gate direct tunneling current in process-induced uniaxial compressive stress p-MOSFETs," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1667–1673, Aug. 2009.



Ming-Jer Chen (S'78–M'79–SM'98) received the B.S. degree in electrical engineering (with highest honors) from National Cheng Kung University, Tainan, Taiwan, in 1977 and the M.S. and Ph.D. degrees in electronics engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1979 and 1985, respectively.

In 1985, he was with the Department of Electronics Engineering, NCTU and, in 1993, became a Full Professor. From 1987 to 1992, he was a Consultant with Taiwan Semiconductor Manufacturing Company (TSMC), where he led a team from NCTU and the Electronics Research and Service Organization/Industrial Technology Research Institute to build up a series of process windows and design rules. From 2000 to 2001, he was a Visiting Professor with the Department of Electrical Engineering and the Center for Integrated Systems, Stanford University, Stanford, CA. He is the holder of eight U.S. patents and six Taiwanese patents in the field of the high-precision analog capacitors, 1-T memory cell, dynamic threshold MOS, electrostatic discharge protection, and Flash memory. He has graduated 17 Ph.D. students and more than 100 M.S. students. His current research interests include device physics, trap physics, and reliability physics, all in the area of nanoelectronics.

Dr. Chen is a member of the Phi Tau Phi.

Li-Ming Chang was born in Taipei, Taiwan, in 1981. He received the B.S. degree in electronics engineering, in 2003, from National Chiao Tung University, Hsinchu, Taiwan, where he is currently working toward the Ph.D. degree in the Department of Electronics Engineering and Institute of Electronics.

His current research interests include carrier transport phenomena and characterization of nanoscale devices.

Evidence for a Very Small Tunneling Effective Mass ($0.03m_0$) in MOSFET High-k (HfSiON) Gate Dielectrics

Ming-Jer Chen, *Senior Member, IEEE*, and Chih-Yu Hsu, *Student Member, IEEE*

***Abstract*—We have recently conducted experimental and modeling tasks on TaC/HfSiON/SiON n-type MOSFETs, leading to an effective mass of $0.03m_0$ for 2-D electrons tunneling in high-k HfSiON dielectrics. In this letter, we present extra evidence obtained from complementary MOSFETs undergoing the same TaC/HfSiON/SiON processing, which shows that such a very small tunneling effective mass is existent not only for 3-D electrons but also for 2-D holes. This new finding is very important because it can substantially enhance the current understanding of gate tunneling leakage suppression in metal-gate high-k MOSFETs.**

***Index Terms*—Effective mass, effective oxide thickness (EOT), HfO_2, HfSiON, high-k, metal gate, MOSFETs, tunneling.**

I. Introduction

HIGH-k GATE dielectrics are currently largely employed in advanced MOSFET manufacturing. Thus, understanding the fundamental properties of high-k dielectrics is crucial. Relative to conventional SiO_2 and SiON counterparts, high-k dielectrics feature two fundamentally distinct properties: a narrower energy bandgap and a lower optical phonon energy [1]. Concerning electrons or holes tunneling in high-k dielectrics, their tunneling effective masses should, in principle, differ fundamentally from those of gate oxide. More recently, we have conducted experimental and modeling tasks on TaC/HfSiON/SiON n-MOSFETs and found that 2-D electrons in a HfSiON dielectric have a tunneling effective mass of around $0.03m_0$ [2]. This value is quite unusual because it is far below that of gate oxide and is the smallest of high-k dielectrics to date. On the other hand, a countertrend with increasing effective oxide thickness (EOT) was experimentally observed [3]–[5]: HfO_2 gate tunneling leakage with respect to the SiO_2 one does not decrease as intuitively expected. To elucidate this, an intermixing action between a high-k dielectric and an interfacial layer was proposed [3], [4]; however, tunneling effective masses as responsible origins were not mentioned there.

Manuscript received October 26, 2011; accepted December 28, 2011. Date of publication January 31, 2012; date of current version March 23, 2012. This work was supported by the National Science Council of Taiwan under Contract NSC 98-2221-E-009-164-MY3. The review of this letter was arranged by Editor J. Cai.

The authors are with the Department of Electronics Engineering and the Institute of Electronics, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: chenmj@faculty.nctu.edu.tw; cyhsu.ee94g@g2.nctu.edu.tw).

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

Digital Object Identifier 10.1109/LED.2011.2182599

The aim of this letter is to provide extra evidence to confirm the existence of a very small tunneling effective mass and make it one of the fundamental properties of the high-k HfSiON dielectric. In a sense, the current understanding of the observed gate tunneling leakage suppression [3]–[5] is able to be significantly enhanced.

II. Experiment

TaC/HfSiON/SiON n- and p-MOSFETs were fabricated in a manufacturing process detailed elsewhere [6]. A TEM picture of the underlying TaC/HfSiON/SiON/Si system is shown in Fig. 1. The corresponding energy band diagram is together plotted for a p-MOSFET in flatband condition. All relevant material and process parameters are labeled in Fig. 1, along with the corresponding values. By performing a quantum mechanical numerical fitting of gate capacitance C_g–V_g measured from the p-MOS device in inversion, we obtained TaC work function = 4.48 eV, EOT = 1.5 nm, and n-type substrate doping concentration = 1×10^{17} cm^{-3}. Evidently, the p-MOS gate stack is slightly larger than the n-MOS one (1.4 nm) [2]. We attributed this to the different nitrogen concentrations encountered. To meet the same EOT (1.5 nm), ε_{IL} in the interface layer (IL) was changed to $6.2\varepsilon_0$. The corresponding φ_{ILC} and φ_{ILV} were 2.54 and 3.06 eV, respectively [7].

The carrier separation method in inversion condition was employed. The measured terminal currents are shown in Figs. 2 and 3 for n- and p-MOSFETs, respectively. Fig. 2 reveals the following: 1) The source/drain current $I_{S/D}$ dominates the gate current I_g due to 2-D electron tunneling, and 2) owing to 3-D valence-band electron tunneling to the gate, separated holes flow down the substrate and constitute the substrate current I_b. In the inset of the figure, the carrier separation measurement setup is shown. In Fig. 3, one can see that I_g comprises two distinct components: 1) $I_{S/D}$ due to hole tunneling from the inversion layer and 2) I_b due to 3-D electron tunneling from the metal side. In Fig. 3, we inserted experimental C_g–V_g for p-MOSFET in inversion, along with the aforementioned curve fitting. The corresponding energy band diagrams and tunneling paths are shown in Fig. 4.

III. Calculation and Fitting

A quantum gate tunneling simulator [2], [8] was used. Given the known material and process parameters ($m_k^* = 0.03m_0$,

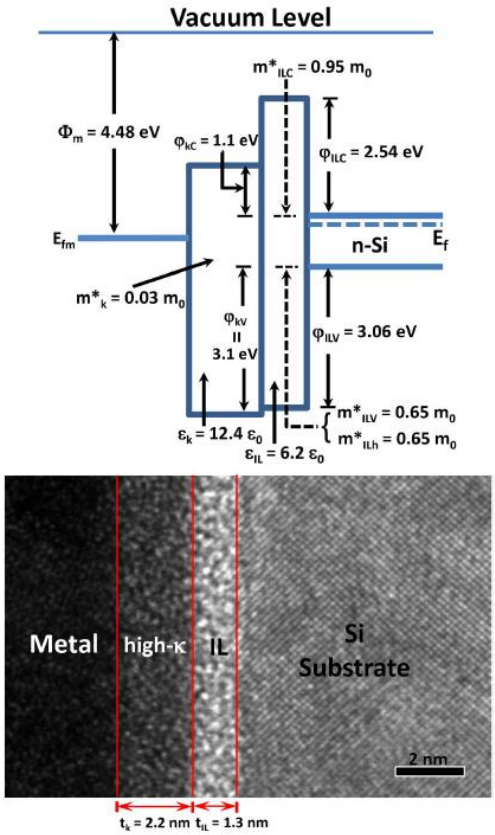


Fig. 1. Schematic of the energy band diagram in a metal-gate/high-k/IL/Si system for a p-MOSFET in flatband condition, along with a high-resolution TEM picture. The relevant material and process parameters are as follows: 1) the TaC metal work function Φ_m; 2) for the HfSiON high-k part, its physical thickness t_k, permittivity ε_k, tunneling effective mass m_k^*, conduction-band offset $\varphi_{\rm kC}$, and valence-band offset $\varphi_{\rm kV}$; and 3) for the SiON IL part, its physical thickness $t_{\rm IL}$, permittivity $\varepsilon_{\rm IL}$, conduction-band electron tunneling effective mass $m_{\rm ILC}^*$, valence-band electron tunneling effective mass $m_{\rm ILV}^*$, hole tunneling effective mass $m_{\rm ILh}^*$, conduction-band offset $\varphi_{\rm ILC}$, and valence-band offset $\varphi_{\rm ILV}$.

$\varphi_{\rm kC} = 1.1$ eV, $\varepsilon_k = 12.4\varepsilon_0$, $m_{\rm ILC}^* = 0.95m_0$, $\varepsilon_{\rm IL} = 7\varepsilon_0$, and $\varphi_{\rm ILC} = 2.36$ eV) [2], the calculated $I_{S/D}$ of the n-MOSFET and, hence, its $d\ln(I_g)/dV_g$ are shown in Fig. 2. The opposite tunneling case, namely the I_b of the p-MOSFET in inversion, should encounter the same tunneling parameters. To testify to this, we quoted an existing formula [9], and thereby, the underlying I_b can be written as

$$I_b = \frac{4\pi q m_M^*}{h_3} \int_0^{E_{\max}} E T_{\rm WKB}(E)\, dE \qquad (1)$$

where m_M^* $(= 1.0m_0)$ is the metal electron mass; E is the allowed electron energy, as shown in Fig. 4 for the p-MOSFET,

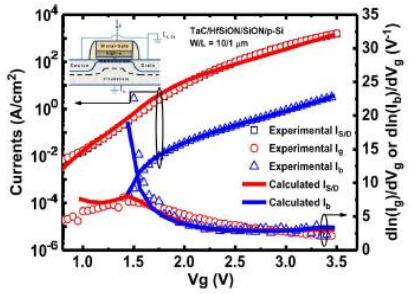


Fig. 2. (Symbols) Experimental I_g, $I_{S/D}$, and I_b, as well as the corresponding $d\ln(I_g)/dV_g$ and $d\ln(I_b)/dV_g$, plotted versus V_g for n-MOSFET in inversion. The (lines) calculated results are given. For $I_{S/D}$ calculation, $\varphi_{\rm kC} = 1.1$ eV, $\varphi_{\rm ILC} = 2.36$ eV, $m_k^* = 0.03m_0$, $m_{\rm ILC}^* = 0.95m_0$, and $\varepsilon_{\rm IL} = 7\varepsilon_0$. For I_b calculation, $m_k^* = 0.03m_0$, and $m_{\rm ILV}^* = 0.65m_0$. The inset schematically shows the current separation measurement.

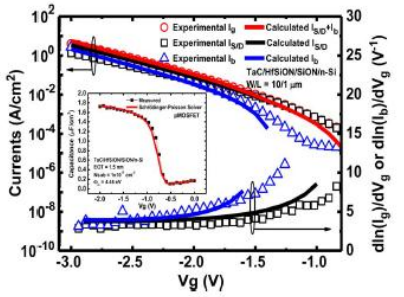


Fig. 3. (Symbols) Experimental I_g, $I_{S/D}$, and I_b, as well as the corresponding $d\ln(I_{S/D})/dV_g$ and $d\ln(I_b)/dV_g$, plotted versus V_g for p-MOSFET in inversion. The (lines) calculated results are given. For I_b calculation, $\varphi_{\rm kC} = 1.1$ eV, $\varphi_{\rm ILC} = 2.54$ eV, $m_k^* = 0.03m_0$, and $m_{\rm ILC}^* = 0.95m_0$. For $I_{S/D}$ calculation, $\varphi_{\rm kV} = 3.1$ eV, $\varphi_{\rm ILV} = 3.06$ eV, $m_k^* = 0.03m_0$, and $m_{\rm ILh}^* = 0.65m_0$. The inset shows a comparison of the (symbol) experimental and (line) simulated C_g versus V_g for TaC/HfSiON/SiON-gate-stack p-MOSFET in inversion.

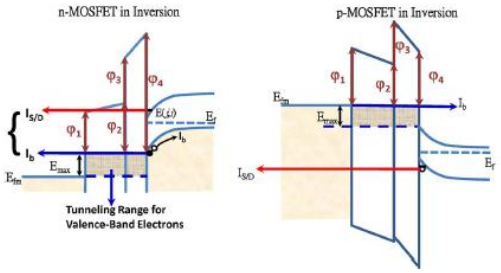


Fig. 4. Schematics of the energy band diagrams and tunneling paths for n- and p-MOSFETs.

for tunneling between metal Fermi level and conduction-band energy at the Si/IL interface; and $E_{\max}$ is the corresponding energy difference. The WKB transmission probability $T_{\rm WKB}$ in (1) can make use of existing analytic expressions (i.e., [2, eq. (2)]) as long as the tunneling criteria (i.e., φ_1, φ_2, φ_3, and

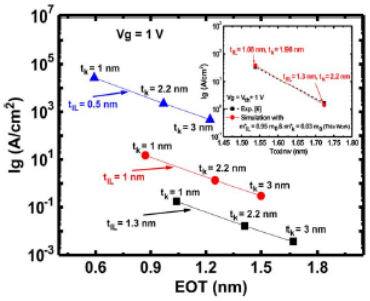


Fig. 5. Simulated I_g due to electron tunneling from the inversion layer versus EOT with $t_{\rm IL}$ as a parameter. The corresponding t_k values are labeled. The material parameters used in the simulation are the same as those in Fig. 2. The inset shows a comparison of the simulated I_g with experimental values [6] in the same TaC/HfSiON/SiON process, plotted versus electrical gate oxide thickness in inversion. $V_{\rm th}$ is the threshold voltage.

φ_4 therein) are modified according to the energy band diagram in Fig. 4. Strikingly, the resulting I_b appears to match the experimental data well, as shown in Fig. 3. This was achieved without changing any parameters.

Next, the values of $m_k^* = 0.03m_0$ and $m_{\rm ILV}^* = 0.65m_0$ were drawn from a fitting of the experimental I_b of the n-MOSFET in inversion. This was done by using (1) but with the following changes: m_M^* was replaced by a valence-band electron effective mass of $0.65m_0$ [9], $E_{\rm max}$ was redefined as the difference between metal Fermi level and silicon valence-band edge, and the corresponding criteria (φ_1, φ_2, φ_3, and φ_4) for $T_{\rm WKB}$ were altered, in accordance with the energy band diagram in Fig. 4. The fitting quality is fairly good, as shown in Fig. 2.

Physically speaking, $m_{\rm ILh}^*$ in IL should be equal or close to $m_{\rm ILV}^*$. In this work, we made $m_{\rm ILh}^* = 0.65m_0$. To calculate the hole tunneling component $I_{S/D}$ of the p-MOSFET, a hole tunneling simulator [8] was utilized. $T_{\rm WKB}$ can be easily modified accordingly. Then, a comparison of the calculated $I_{S/D}$ with the experimental one led to $\varphi_{\rm kV} = 3.1$ eV. As shown in Fig. 3, good fitting holds again.

IV. Discussion

To see the individual effects of varying t_k and $t_{\rm IL}$, we show in Fig. 5 the simulated I_g due to electron tunneling from the inversion layer, plotted versus EOT for three $t_{\rm IL}$ values. The simulation points are also labeled with corresponding t_k. In addition, the simulated I_g was found to match existing data in the same manufacturing process [6], as shown in the inset of the figure for two different combinations of t_k and $t_{\rm IL}$.

From Fig. 5, we can see the following: 1) the gate leakage increases with decreasing EOT, in agreement with [3]–[5], and 2) an increase in $t_{\rm IL}$ can suppress I_g more significantly than t_k. We also show in Fig. 5 that reducing $t_{\rm IL}$ will seriously increase I_g until it is intolerably high. This seems to be inconsistent with recent experiments [5]: I_g through HfO_2 is tolerable even for the case of $t_{\rm IL}$ approaching zero. However, one of the fundamental differences should be kept in mind: HfO_2 permittivity is higher than that of HfSiON, and as a consequence of maintaining the same EOT, the HfO_2 dielectric is much thicker. Indeed, this is the fact since a fair comparison of experimental I_g between HfSiON and HfO_2 has been published in the literature [6].

As corroborated earlier, a fundamental very small tunneling effective mass is existent. In this sense, a tunneling effective mass point of view is able to enhance current understanding of the observed I_g suppression [3], [4]. First, the HfO_2 dielectric is featured by a very small m_k^* while the interfacial layer can have a much higher $m_{\rm IL}^*$. Next, as stated in [3] and [4], extra annealing treatments increase $t_{\rm IL}$ while simultaneously making more hafnium atoms appear in the regrown interfacial layer. Thus, the corresponding $m_{\rm IL}^*$ is likely to be lowered according to this work. Consequently, the gate leakage suppression ability relative to the SiO_2 gate oxide is degraded, as experimentally observed [3], [4].

V. Conclusion

Characterization and modeling of gate tunneling components of TaC/HfSiON/SiON complementary MOSFETs in inversion have been carried out. A fundamental tunneling effective mass featuring a very small value has been corroborated in the HfSiON dielectric. Current understanding of gate leakage suppression has been substantially enhanced.

References

[1] M. V. Fischetti, T. P. O'Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, and Y. Zhang, "Theoretical study of some physical aspects of electronic transport in nMOSFETs at the 10-nm gate-length," *IEEE Trans. Electron Devices*, vol. 54, no. 9, pp. 2116–2136, Sep. 2007.

[2] C. Y. Hsu, H. G. Chang, and M. J. Chen, "A method of extracting metal-gate high-k material parameters featuring electron gate tunneling current transition," *IEEE Trans. Electron Devices*, vol. 58, no. 4, pp. 953–959, Apr. 2011.

[3] A. Callegari, P. Jamison, E. Cartier, S. Zafar, E. Gusev, V. Narayanan, C. D'Emic, D. Lacey, F. M. Feely, R. Jammy, M. Gribelyuk, J. Shepard, W. Andreoni, A. Curioni, and C. Pignedoli, "Interface engineering for enhanced electron mobilities in W/HfO_2 gate stacks," in *IEDM Tech. Dig.*, Dec. 2004, pp. 825–828.

[4] A. Callegari, P. Jamison, D. Neumayer, F. M. Feely, J. Shepard, W. Andreoni, A. Curioni, and C. Pignedoli, "Electron mobility dependence on annealing temperature of W/HfO_2 gate stacks: The role of the interfacial layer," *J. Appl. Phys.*, vol. 99, no. 2, pp. 023709-1–023709-5, Jan. 2006.

[5] T. Ando, M. M. Frank, K. Choi, C. Choi, J. Bruley, M. Hopstaken, M. Copel, E. Cartier, A. Kerber, A. Callegari, D. Lacey, S. Brown, Q. Yang, and V. Narayanan, "Understanding mobility mechanisms in extremely scaled HfO_2 (EOT 0.42 nm) using remote interfacial layer scavenging technique and Vt-tuning dipoles with gate-first process," in *IEDM Tech. Dig.*, Dec. 2009, pp. 423–426.

[6] Y. T. Hou, F. Y. Yen, P. F. Hsu, V. S. Chang, P. S. Lim, C. L. Hung, L. G. Yao, J. C. Jiang, H. J. Lin, Y. Jin, S. M. Jang, H. J. Tao, S. C. Chen, and M. S. Liang, "High performance tantalum carbide metal gate stack for nMOSFET application," in *IEDM Tech. Dig.*, Dec. 2005, pp. 31–34.

[7] X. Guo and T. P. Ma, "Tunneling leakage current in oxynitride: Dependence on oxygen/nitrogen content," *IEEE Electron Device Lett.*, vol. 19, no. 6, pp. 207–209, Jun. 1998.

[8] C. Y. Hsu, C. C. Lee, Y. T. Lin, C. Y. Hsieh, and M. J. Chen, "Enhanced hole gate direct tunneling current in process-induced uniaxial compressive stress p-MOSFETs," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1667–1673, Aug. 2009.

[9] J. W. Yang, J. G. Fossum, G. O. Workman, and C. L. Huang, "A physical model for gate-to-body tunneling current and its effects on floating-body PD/SOI CMOS devices and circuits," *Solid-State Electron.*, vol. 48, no. 2, pp. 259–270, Feb. 2004.