

Oral presentation | 13 Semiconductors : 13.5 Semiconductor devices/ Interconnect/ Integration technologies

📅 Wed. Sep 20, 2023 9:00 AM - 12:00 PM JST | Wed. Sep 20, 2023 12:00 AM - 3:00 AM UTC | 🏢 A304 (KJ Hall)

[20a-A304-1~11] 13.5 Semiconductor devices/ Interconnect/ Integration technologies

Wen Hsin Chang(AIST)

9:00 AM - 9:15 AM JST | 12:00 AM - 12:15 AM UTC

[20a-A304-1] Additional High-Pressure Hydrogen Annealing Improving the Cryogenic Operation of Si (110)-oriented n-MOSFETs

○(D)Shunsuke Shitakata^{1,2}, Hiroshi Oka¹, Takumi Inaba¹, Shota Iizuka¹, Kimihiko Kato¹, Takahiro Mori¹ (1.AIST, 2.APPI, Keio)

9:15 AM - 9:30 AM JST | 12:15 AM - 12:30 AM UTC

[20a-A304-2] Condition to overestimate DIBL parameter in cryogenic operation of MOSFETs

○(M1)Yuika Kobayashi^{1,2}, Hidehiro Asai², Shota Iizuka², Junichi Hattori², Tsutomu Ikegami², Koichi Fukuda², Tetsuro Nikuni¹, Takahiro Mori² (1.Tokyo Univ. of Science, 2.AIST)

9:30 AM - 9:45 AM JST | 12:30 AM - 12:45 AM UTC

[20a-A304-3] Verification of a New Charge-Based Threshold Voltage Definition

○KIYOSHI TAKEUCHI¹, MASAHARU KOBAYASHI^{2,1}, TOSHIRO HIRAMOTO¹ (1.IIS, Univ. Tokyo, 2.d.lab, Univ. Tokyo)

9:45 AM - 10:00 AM JST | 12:45 AM - 1:00 AM UTC

[20a-A304-4] Effects of quantum confinement on electron velocity overshoot in Si nanosheet FETs

○Junichi Hattori¹, Koichi Fukuda¹, Tsutomu Ikegami¹, Yoshihiro Hayashi¹ (1.AIST)

10:00 AM - 10:15 AM JST | 1:00 AM - 1:15 AM UTC

[20a-A304-5] Integrate-and-Fire Operation by using “Dual-Gate PN-Body Tied SOI-FET”

○(M2)Haruki Yonezaki¹, Takayuki Mori¹, Jiro Ida¹ (1.Kanazawa Inst. of Tech.)

10:30 AM - 10:45 AM JST | 1:30 AM - 1:45 AM UTC

[20a-A304-6] Threshold Voltage Control by Al Diffusion for NMOSFET with Metal Gate Stack

○Yuya Omura¹, Shoji Aota¹, Tomoya Wada¹, Tomoyuki Funabasama¹, Kasumi Yasuda¹, Daisuke Watanabe¹, Hiroki Okamoto¹, Isamu Ito¹, Masato Koyama¹ (1.KIOXIA)

10:45 AM - 11:00 AM JST | 1:45 AM - 2:00 AM UTC

[20a-A304-7] Improvement of current drivability for metal S/D Ge n-MOSFET by introduction of recessed channel structure (III)

○Hajime Kuwazuru¹, Dong Wang², Keisuke Yamamoto² (1.IGSES, Kyushu Univ., 2.FES, Kyushu Univ.)

◆ English Presentation

11:00 AM - 11:15 AM JST | 2:00 AM - 2:15 AM UTC

[20a-A304-8] Performance Enhancement of Extremely-thin Body (111) Ge-on-Insulator nMOSFETs by Using Flipped Substrate Process

○Xueyang Han¹, Chia-Tsong Chen¹, Kei Sumita¹, Kasidit Toprasertpong¹, Mitsuru Takenaka¹, Shinichi Takagi¹ (1.Univ. Tokyo)

◆ Presentation by Applicant for JSAP Young Scientists Presentation Award

11:15 AM - 11:30 AM JST | 2:15 AM - 2:30 AM UTC

[20a-A304-9] A Nanosheet Oxide Semiconductor FET Using ALD InGaOx Channel for 3D Integrated Devices

○Kaito Hikake¹, Zhuo Li¹, Junxiang Hao¹, Chitra Pandey¹, Takuya Saraya¹, Toshiro Hiramoto¹, Takanori Takahashi², Mutsunori Uenuma², Yukiharu Uraoka², Masaharu Kobayashi^{1,3} (1.IIS, Univ. of Tokyo, 2.NAIST, 3.d.lab, Univ. of Tokyo)

11:30 AM - 11:45 AM JST | 2:30 AM - 2:45 AM UTC

[20a-A304-10] Design and analysis of an 8kB macro for a new ultralow-voltage retention SRAM

○Katsutoshi Ito¹, Yusaku Shiotsu¹, Satoshi Sugahara¹ (1.FIRST, Tokyo Inst. of Tech.)

11:45 AM - 12:00 PM JST | 2:45 AM - 3:00 AM UTC

[20a-A304-11] A binarized neural network accelerator macro with parallelized MAC units using ULVR-SRAM

○Yusaku Shiotsu¹, Satoshi Sugahara¹ (1.FIRST, Tokyo Inst. of Tech.)

追加高圧水素アニールによる(110)面 Si n-MOSFET の極低温特性改善

Additional High-Pressure Hydrogen Annealing Improving the Cryogenic Operation of

Si (110)-oriented n-MOSFETs

産総研¹, 慶大物情² ○(D) 下方 駿佑^{1,2}, 岡 博史¹, 稲葉 工¹, 飯塚 将太¹, 加藤 公彦¹, 森 貴洋¹

AIST¹, APPI, Keio², °S. Shitakata^{1,2}, H. Oka¹, T. Inaba¹, S. Iizuka¹, K. Kato¹, T. Mori¹

E-mail: s-shitakata@aist.go.jp

【背景】極低温下で集積回路を動作させるクライオ CMOS 技術は量子コンピュータの制御回路技術として期待されている。MOSFET の極低温動作にはシリコン/絶縁膜界面の品質が大きく影響することが明らかになっており[1, 2], 界面品質の向上は極低温動作 MOSFET の特性改善において重要である。今回, 界面品質の向上を狙い, 高圧水素アニール[3]を導入し, 極低温下動作 MOSFET への適用を検討した。

【実験】(110)面バルクウエハに MOSFET を作製, 400°C で 30 分水素アニールを行った。一部の試料には追加で 450°C, 30 分水素アニール (圧力 1 atm または 20 atm) も施した。作製した計 3 種の試料について低温での電気特性の評価を行った。

【結果と考察】Fig. 1 に 2.5 K での電流電圧特性を示す。追加水素アニールによりドレイン電流の増大が確認された。50 K でのチャージポンピングの結果から見積もられる界面トラップ密度は, 追加水素アニールなし(w/o), 1 atm, 20 atm それぞれで $5 \times 10^{10} \text{ cm}^{-2}$, $1 \times 10^{10} \text{ cm}^{-2}$, $7.4 \times 10^9 \text{ cm}^{-2}$ であり (Fig. 2), 追加の水素アニールによる界面トラップ密度の低減が確認され, その効果は高圧下の処理で大きく, 高圧水素アニールが界面準位密度低減に効果的であることを示している。Fig. 3 に 2.5 K での等価電圧ノイズ (S_{vg}) のドレイン電流依存を示す。追加水素アニールにより低電流時のノイズが低下, またその大小関係はチャージポンピングで得られた界面トラップ密度と相関することが確認された。

【まとめ】高圧水素アニールが(110)面 n-MOSFET の界面準位密度を低減させることを示し, 極低温下での電流電圧特性, ノイズ特性を改善させることを実験的に示した。

【謝辞】本研究は NEDO の委託業務 (JPNP16007) および文部科学省 Q-LEAP (JPMXS0118069228) の助成を受けて実施した研究の結果得られたものである。

【参考文献】[1] H. Oka *et al.*, Tech. Dig. Symp. VLSI Technol., TN2-2 (2020). [2] T. Inaba *et al.*, Tech. Dig. Symp. VLSI Technol., T13-4 (2023). [3] C. Diouf *et al.*, in Proc. Int. Conf. on Ultimate Integration on Silicon, 9 (2012).

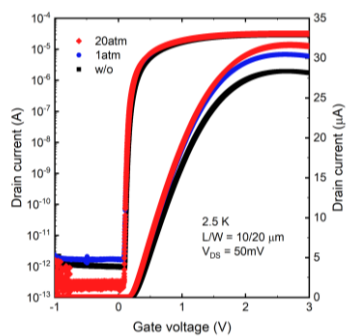


Fig. 1 Id-Vg characteristic at 2.5 K

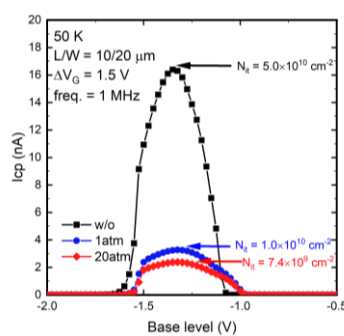


Fig.2 Icp as a function of base level at 50 K

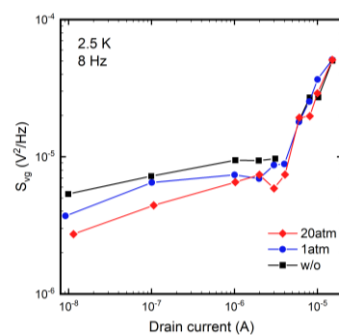


Fig. 3 Svg-Id at 2.5 K

MOSFET の極低温動作における DIBL パラメータ過剰見積の発生条件 Condition to overestimate DIBL parameter in cryogenic operation of MOSFETs

東京理科大¹, 産総研² ○(M1) 小林 唯華^{1,2}, 浅井 栄大², 飯塚 将太²,

服部 淳一², 池上 努², 福田 浩一², 二国 徹郎¹, 森 貴洋²

Tokyo Univ. of Science¹, AIST², ○Yuika Kobayashi^{1,2}, Hidehiro Asai², Shota Iizuka²,
Junichi Hattori², Tsutomu Ikegami², Koichi Fukuda², Tetsuro Nikuni¹, and Takahiro Mori²

E-mail: kobayashi-yuika@aist.go.jp

【背景】極低温環境における MOSFET の動作機構の解明は、量子コンピュータの制御用集積回路の実現にむけて重要である。回路設計において重要なパラメータの 1 つに DIBL (Drain Induced Barrier Lowering) がある。極低温下においては、拡散キャリアの減少とそれに伴うチャネル/ドレイン間の容量低下が発生するため、直感的には DIBL が抑制されると考えられる。一方で、近年極低温下において DIBL が増大するという報告が複数ある[1]。我々はこれまでに、ゲートに対するソース/ドレイン (S/D) のアンダーラップが存在する場合、低 V_d 条件下での V_t が異常を示し、DIBL パラメータが過剰に見積られ、DIBL が増大したように見えてしまう可能性を指摘した[2]。今回我々は、このような過剰見積が発生する条件を、TCAD シミュレーションを用いて詳細に検討したので、これを報告する。その原因は S/D アンダーラップ状況下での I_d - V_g 特性の異常な V_d 依存性にある。

【シミュレーション方法】産総研が独自開発する Impulse TCAD[3] を用いた。極低温特性計算のために、フェルミ・ディラック分布に対応した一般化されたアインシュタインの関係式[4]と、フリーズアウトを考慮したドーパントのイオン化モデル[5]を組み込んでいる。短チャネル pMOSFET を想定しており、計算に用いたデバイスのモデル図を Fig. 1 に示す。

【結果・考察】アンダーラップ長 L_{under} が 15 nm のときのデバイスについて、15 K における I_d - V_g 特性のシミュレーション結果を Fig. 2 に示す。 $|V_d|=0.2$ V 以下の低 $|V_d|$ 領域において、 $|V_d|$ の低下に伴う $|V_t|$ の急激な増加が見られる。アンダーラップのために発生するチャネル両端のポテンシャル障壁が、 $|V_d|$ の低下と共に発生、チャネルが切断されるためである[2]。一方、 $|V_d|=0.2$ V 以上の $|V_d|$ 領域においては、ソース・チャネル・ドレイン間が電氣的に接続された結果、通常の DIBL 現象による $|V_t|$ の低下がみられている。ここで、DIBL パラメータを線形領域での閾値電圧 $V_t(V_d=V_{d-\text{low}})$ と飽和領域での閾値電圧 $V_t(V_d=-1$ V) の差で定義する。このときの DIBL パラメータの $V_{d-\text{low}}$ 依存性を Fig. 3 に示す。15 K において、低 $|V_{d-\text{low}}|$ 領域での DIBL パラメータは、低 $|V_d|$ 領域における V_t の急激な低下を反映し、室温時よりも大きい値を見せている。このように、アンダーラップが存在している状況下で、 $V_{d-\text{low}}$ の設定によっては DIBL パラメータが過剰見積られ、DIBL が増大しているかのように見えることになる。

【謝辞】この成果の一部は、国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託業務(JPNP16007)の結果得られたものである。

【参考文献】[1] Y. Zhang *et al.*, IEEE Trans. Electron Devices **68**, 4267 (2021). [2] T. Inaba *et al.*, Appl. Phys. Express **15**, 084004 (2022). [3] T. Ikegami *et al.*, J. Comput. Electron. **18**, 534 (2019). [4] K. M. Van Vliet and A. H. Marshak, Phys. Status Solidi B **78**, 501 (1976). [5] P. P. Altermatt *et al.*, J. Appl. Phys. **100**, 113715 (2006).

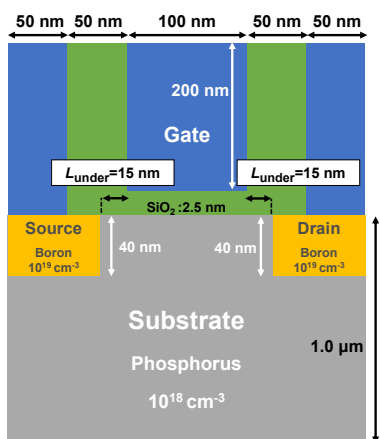


Fig. 1 Schematic of pMOS device structure.

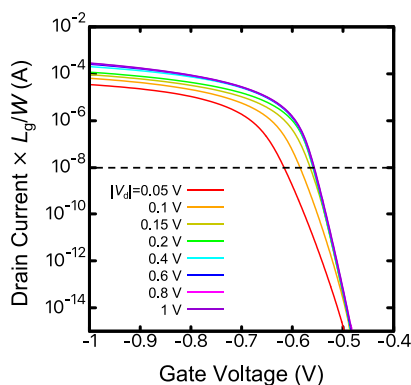


Fig. 2 $I_d \times L_g/W$ versus V_g characteristics at 15 K. V_t values were calculated by constant current method at 10 nA.

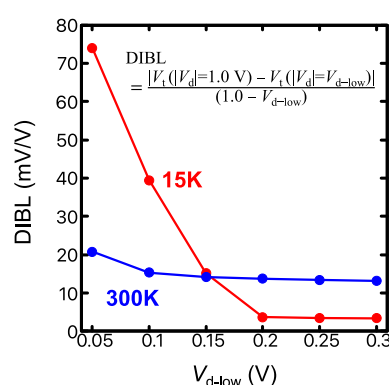


Fig. 3 $V_{d-\text{low}}$ dependence of DIBL parameter at 15 K and 300 K.

電荷密度に基づく新しい値定義の妥当性検証 Verification of a New Charge-Based Threshold Voltage Definition

東大生研¹, 東大 d.lab² *竹内潔¹, 小林正治^{2,1}, 平本俊郎¹
Univ. Tokyo, *Kiyoshi Takeuchi, Masaharu Kobayashi, and Toshiro Hiramoto
E-mail: takeuchi@nano.iis.u-tokyo.ac.jp

背景

MOSFET のモデリングにおいて、しきい値 (V_{TH}) をバンドが $2\psi_B$ 曲がった状態 (図1) と対応させて定義する方法が広く受け入れられている[1]。しかしこの $2\psi_B$ 定義には以下のような種々の問題点がある。

- (a) 反転状態をキャリアの面密度でなく体積密度で定義することの妥当性に疑問がある。
- (b) 基板濃度ゼロ ($\psi_B = 0$) の極限で $V_{TH} = V_{FB}$ となる不自然な基板濃度依存性を示す。
- (c) 基準となる基板を持たない FDSOI や FinFET など薄膜デバイスでそもそも定義ができない。厳密にはバルクデバイスでも基板濃度が深さ方向に不均一であれば同様である。
- (d) キャリア凍結が起こる極低温では反転層が全くできない状態がしきい値となる。

これらを解決するため筆者らは反転電荷の面密度に基づく新たな MOS ダイオードのしきい値定義を提案した[2]。本発表では新定義で用いた近似式の妥当性を量子閉じ込め効果を考慮しつつ議論する[3]。

新定義方法

ゲート電圧 V_G と反転電荷面密度 Q の間には近似的に以下の標準化した関係が成立する (図2参照)。

$$v' = Q' + \log(Q'), \quad Q' = \frac{Q}{m(k_B T/q)C_{OX}}, \quad v' = \frac{V_G}{m(k_B T/q)} - v_0. \quad (1)$$

ここで(1)の第一式右辺がゼロの状態をしきい値状態と定義する。すなわち、 Q がしきい電荷密度

$$Q_{TH} \equiv \xi m(k_B T/q)C_{OX}, \quad \xi \approx 0.56714 \quad (2)$$

と一致するときの V_G を V_{TH} とする。ここで ξ は $x + \log(x) = 0$ の解、 m は基板効果係数、 v_0 は定数である。この定義によれば V_{TH} は単純に Q 対 V_G 特性の形状から決定されることとなり、上記した従来の課題は解消できる。

TCAD による新定義の検証

式(1)による近似の良さを調べるため TCAD[4]により1次元 Poisson/Schrödinger 方程式を解いて Q 対 V_G 特性を計算し、その結果に対して m 、 v_0 、 T_{OX} を可変パラメータとして(1)を合わせこんで一致度を調べた。図2にバルク nMOSFET についての結果例を示す。バルク FET では基板濃度 N_{SUB} が下がるほど、ダブルゲート FET (図示せず) では Si 膜厚が厚くなるほど近似が悪化することが分かった。しかしフィッティング範囲 $Q'_{LO} \leq Q' \leq Q'_{HI}$ に着目すると実特性は概ね式(1)をトレースすると見做せた。 $Q' = \xi$ (図2b縦破線)における理想特性と TCAD 特性とのズレをしきい値誤差と見做すと、図2中最悪の $N_{SUB} = 3e15 \text{ cm}^{-3}$ の場合でも誤差は 8mV である。

新定義の課題として、 Q_{TH} を決めるためにゲート絶縁膜容量 $C_{OX} = \epsilon_{OX}/T_{OX}$ 、従ってゲート絶縁膜厚 T_{OX} を知る必要があることが挙げられる。 T_{OX} は反転層膜厚を含み物理膜厚とは異なる。これは上記したフィッティングで決定できるが、これを毎回実施するのは実用上不便である。そこで T_{OX} として物理膜厚に一律 0.4nm を加算した値を利用した場合の V_{TH} 誤差を算出した。その結果、誤差に大きな劣化は見られなかった。

以上より、一定値の反転層厚さ補正を行うことで T_{OX} 決定の煩雑さを避けつつ、新定義により理想 Q 対 V_G 特性と良好に対応する V_{TH} を決定できると結論できる。

[1] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, 1981), 2nd ed., p.366. [2] K. Takeuchi et al., IEEE Trans. Electron Devices, 69, p. 942 (2022). [3] K. Takeuchi et al., 2023 *Silicon Nanoelectronics Workshop*, 7-2. [4] Sentaurus Device User Guide, ver. P-2019.03, Synopsys.

本研究は国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託業務(JPNP16007)の結果得られたものです。

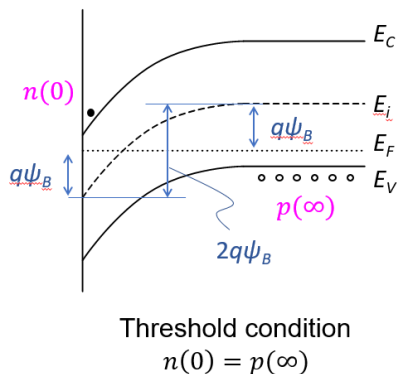


図1 $2\psi_B$ 定義によるしきい状態。

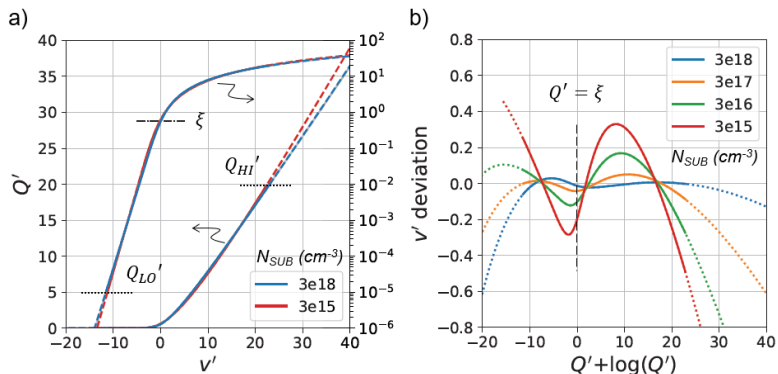


図2 TCAD と式(1)の比較、a) Q - v 特性、b) フィッティング誤差。

Si ナノシート電界効果トランジスタにおける 電子速度オーバーシュートに与える量子閉じ込めの影響

Effects of quantum confinement on electron velocity overshoot in Si nanosheet FETs

○服部 淳一, 福田 浩一, 池上 努, 林 喜宏 (産総研)

○Junichi Hattori, Koichi Fukuda, Tsutomu Ikegami, and Yoshihiro Hayashi (AIST)

E-mail: j.hattori@aist.go.jp

背景 ナノシートをチャネルにしてその表面をゲートで覆ったゲート・オール・アラウンド形ナノシート電界効果トランジスタ(NS GAAFET)は、次世代半導体集積回路のスイッチング素子として、現在、活発に研究開発が進められている [1]。NS GAAFET の想定されるゲート長は非常に短いため、キャリアが NS を走行中、強い電界で飽和速度を超えて加速される速度オーバーシュート(VO)が起こりえる。われわれは前回の講演でこの VO の影響を明らかにした [2]。他方、想定される NS の膜厚は一層薄いため、量子閉じ込めの影響も強く現れると考えられる。この量子閉じ込めが VO に与える影響は十分に理解されてはいない。そこで、本研究ではこれを調べる。

方法 Fig. 1 に示す n 型 Si NS GAAFET の静的挙動をデバイス・シミュレータ Impulse TCAD [3] でシミュレーションし、エネルギー輸送(ET)モデル [4] とドリフト拡散(DD)モデルに対する結果の差から VO の影響を評価した。次に、Impulse TCAD に密度勾配(DG)モデル [5] を実装して量子閉じ込めを考慮できるようにした上で同様のことを行い、VO の影響に違いが現れるか調べた。

結果 ゲート長 50 nm の NS GAAFET の伝達特性を Fig. 2 に示す。ET モデル、DD モデル、どちらも量子閉じ込めを考慮することで電流は減少する。これは、Fig. 3(a) に示すように、量子閉じ込めを考慮しない場合には電子が集中する絶縁膜との界面近傍で、電子密度が著しく減少するからである。また、ET モデルと DD モデルを比べると、前者の方が電流が多い。これは、Fig. 3(b) に示すように、電子の速度が飽和速度(約 10^7 cm/s)を超える VO が考慮されるからである。この電流増加を VO の影響とみなしてゲート長との関係を調べた結果を Fig. 4 に示す。図から、量子閉じ込めを考慮する方が VO の影響が強く現れると分かる。この直接の原因は、Fig. 3(b) から分かるように、量子閉じ込めによって電子が集中するナノシート中央部において電子速度がより高くなることにある。

結論 量子閉じ込めによって NS GAAFET における電子分布は大きく変わり、VO も影響を受ける。正確なシミュレーションに向けては、VO に加えて量子閉じ込めも考慮することが不可欠である。

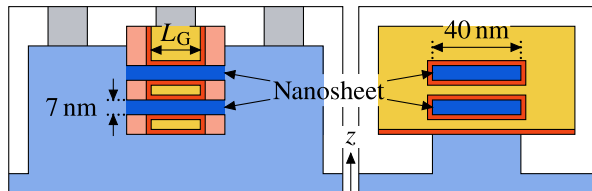


Fig. 1. Sectional views of Si nanosheet (NS) gate-all-around field-effect transistors (GAAFETs).

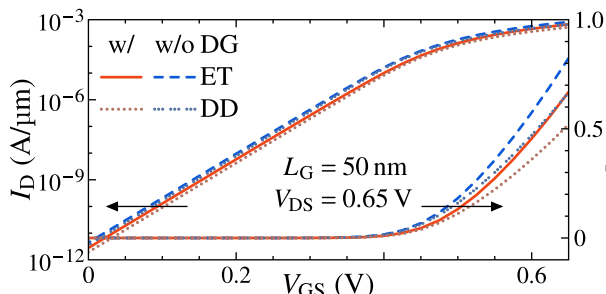


Fig. 2. Drain current I_D vs gate-to-source voltage V_{GS} characteristics of a NS GAAFET with a gate length L_G of 50 nm at a drain-to-source voltage V_{DS} of 0.65 V.

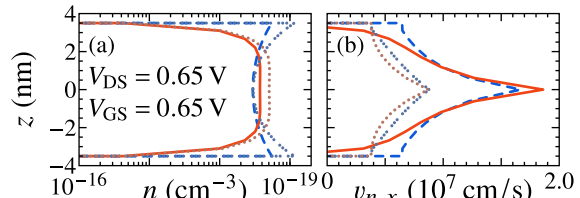


Fig. 3. Distributions of (a) electron density and (b) velocity along the z -axis at the center of the lower NS.

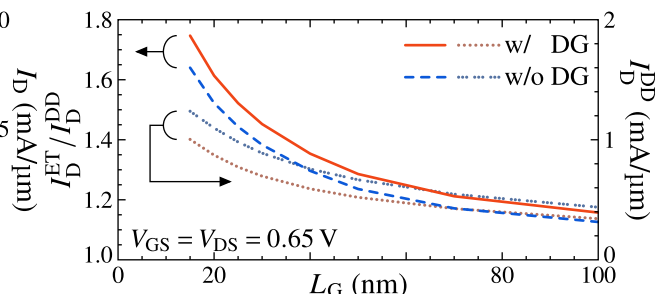


Fig. 4. Ratio of the I_D obtained from the energy transport (ET) model to that obtained from the drift-diffusion (DD) model plotted as a function of L_G .

謝辞 本成果は国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の助成事業(ポスト 5G 情報通信システム基盤強化研究開発事業, JPNP20017)の結果得られたものです。 **参考文献** [1] N. Loubet *et al.*, Proc. Symp. VLSI Technol., 2017, p. T230. [2] J. Hattori *et al.*, Ext. Abstr. JSAP Spring Meet., 2023, 17a-A403-2 [in Japanese]. [3] T. Ikegami *et al.*, J. Comput. Electron. **18**, 534 (2019). [4] W.-S. Choi *et al.*, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **13**, 899 (1994). [5] M. G. Ancona and H. F. Tiersten, Phys. Rev. B **35**, 7959 (1987).

“Dual-Gate PN-Body Tied SOI-FET”を利用した Integrate-and-Fire 動作の実証

Integrate-and-Fire Operation by using “Dual-Gate PN-Body Tied SOI-FET”

金沢工大, °(M2)米崎 晴貴, 森 貴之, 井田 次郎

Kanazawa Inst. of Tech., °Haruki Yonezaki, Takayuki Mori, Jiro Ida

E-mail: c6200787@st.kanazawa-it.ac.jp

はじめに： ニューロモルフィックハードウェア実現のため、スパイクングニューロンモデルの一つである Leaky Integrate-and-Fire (LIF) モデルを単体デバイスで実現する研究が行われている[1]. 我々の研究室でも LIF の内, IF 機能を”Dual-Gate PN-Body Tied SOI-FET DG PNBT SOI-FET)”を用いて単体デバイスで実現できることをシミュレーションで報告している[2]. 本研究では DG PNBT SOI-FET の実測で IF 動作が実現できるか初めて検証した.

実験結果： 図 1 に DG PNBT SOI-FET のデバイス構造を示す. 図 2 にパルス測定の電圧条件を示す. 図 3 に静特性を示す. 静特性では 2nd Gate 電圧を低くしていくことで急峻な立ち上がりを得られた. 図 4 に各 2nd Gate 電圧を一定にして, Body 電圧を 0V から 1V に立ち上げたときのターンオン特性を示す. 結果より, 2nd Gate 電圧を変化させることによって立ち上がる時間が変化している. また, 図 5 に示すように, 発火するためのパルス数が 2nd Gate 電圧によって変化している. $V_{g2} = 0.75V$ の場合は 1 パルス目で発火するが, $V_{g2} = 0.83V$ の場合は 8 パルス目で発火している. この結果より, 2nd Gate によって正孔の注入を制御でき, 発火するために必要なパルス数を調整可能であることが解る. これは, DG PNBT SOI-FET で 2nd Gate を重みとした IF 動作を実現することが可能であることを示す.

謝辞： 本研究は JSPS 科研費 JP21K14216 及び JST-CREST JPMJCR20Q1 の助成を受けたものである. また, 東京大学 VDEC 活動を通して, 日本シノプシス合同会社及びメンター・グラフィックス・ジャパン株式会社の協力で行われたものである.

参考文献： [1] S. Dutta et al., *Sci. Rep.*, Aug.2017. [2] T. Mori, and J. Ida, *JJAP*, vol. 60, no. SB, Feb.2021.

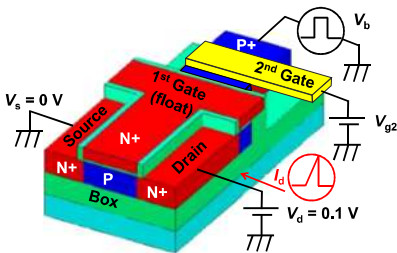


図 1 デバイス構造

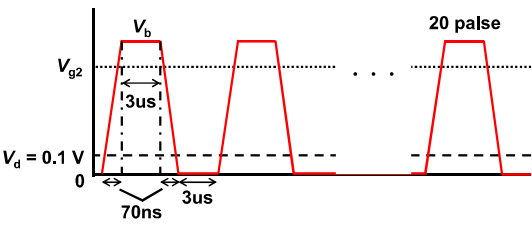


図 2 パルス測定の電圧条件

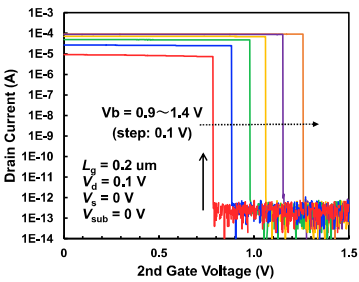


図 3 I_d - V_{g2} 特性

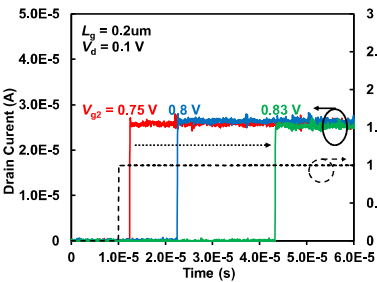


図 4 V_b 入力のターンオン特性

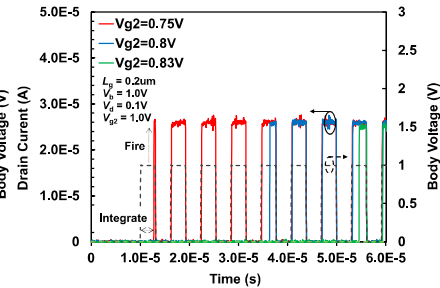


図 5 パルス測定結果

Al 拡散によるメタルゲートスタック構造のしきい値電圧最適化

Threshold Voltage Control by Al Diffusion for NMOSFET with Metal Gate Stack

キオクシア株式会社 メモリ技術研究所

○大村 祐弥、青田 正司、和田 智也、船迫 友之、安田 かすみ、渡邊 大輔、岡本 浩樹、伊藤 勇、小山 正人

Institute of Memory Technology Research & Development, KIOXIA Corporation

○Yuya Omura, Shoji Aota, Tomoya Wada, Tomoyuki Funabasama, Kasumi Yasuda,

Daisuke Watanabe, Hiroki Okamoto, Isamu Ito, Masato Koyama

E-mail: yuya1.omura@kioxia.com

近年、メモリ周辺回路の低消費電力、高速動作要求が高まり、従来の Poly-Si/SiON 構造トランジスタより高性能な Metal Gate/High-k(MG/HK)構造トランジスタの導入が必要とされている^[1]。また、メモリ周辺回路では、メモリセルの情報を誤りなく読み出すために、しきい値電圧(V_{th})などのデバイス特性を精度高く制御する必要がある^{[2][3]}。MG/HK 構造の NMOSFET では、複数金属からなる積層メタルゲート電極において、TiAl 層からゲート絶縁膜界面への Al 拡散によりメタルゲートの Work Function 制御を行うことが報告されている^[4]。本研究では、TiAl を含む積層メタルゲート電極の膜厚を系統的に変化させた時の V_{th} およびそのばらつきを評価することにより、NMOSFET の V_{th} 変化における Al の役割を考察した。

Fig. 1 に示すようなゲートスタック構造を有する NMOSFET を作成し、TiAl 膜厚を変化させたときの V_{th} およびそのばらつきを評価した。TiAl 膜厚を 2nm から 1nm 刻みで厚くするに従い、 V_{th} は負シフトし、膜厚 4nm 以上で所望の値に飽和する傾向が得られた(Fig.2)。SIMS 分析にて Barrier Metal/High-k 膜界面の Al 濃度を取得した結果、TiAl 膜厚が 2nm の場合には界面 Al は検出感度以下であったが、膜厚 3nm 以上では 1×10^{20} atoms/cm³ 台の界面 Al の存在が確認され、高 Al 濃度ほど V_{th} が負にシフトすることが分かった(Fig. 3)。講演当日は界面 Al による V_{th} 変化を考察するとともに、 V_{th} ばらつきの実験結果の解析と議論についても報告する。

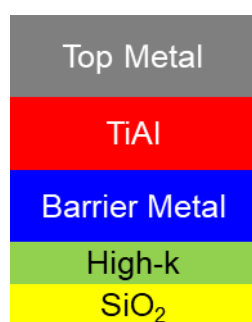


Fig.1: Schematic of Metal Gate Stack

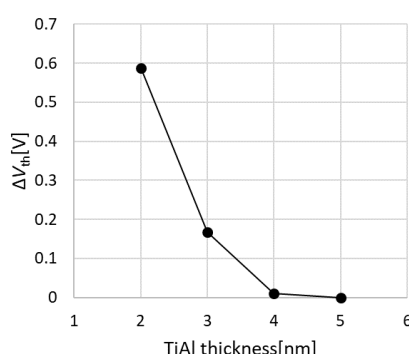


Fig.2: ΔV_{th} vs TiAl thickness

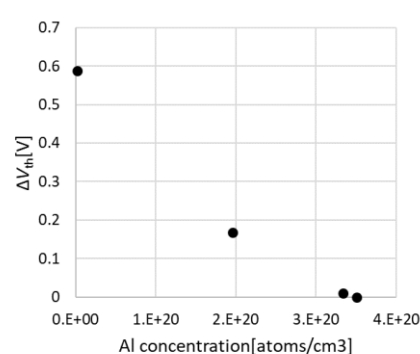


Fig. 3: ΔV_{th} vs Al concentration at interface between Barrier Metal and High-k

[1]E. Capogreco *et al.*, IEDM, 2022

[2]Sung Ho Jang *et al.*, IEDM, 2019

[3]Dongyeon Oh *et al.*, IEDM, 2020

[4]A. Veloso *et al.*, VLSI Tech. Dig. 2013

リセスチャネル化によるメタル S/D 型 Ge n-MOSFET の電流駆動力向上(III)

Improvement of current drivability for metal S/D Ge n-MOSFET by introduction of recessed channel structure (III)

九大・大学院総合理工学府/研究院, °(M2) 鎌 一, 王 冬, 山本 圭介

IGSES, Kyushu Univ., °Hajime Kuwazuru, Dong Wang, Keisuke Yamamoto

E-mail: kuwazuru.hajime.876@s.kyushu-u.ac.jp

1. はじめに

当グループはこれまでに、メタルソース・ドレイン (S/D) 型 Ge n-MOSFET に対して、チャネル部分を掘り込んだリセスチャネル構造を導入することで、電流駆動力の改善と寄生抵抗 R_p の低減に成功している。また、リセスチャネルの深さ d (Fig. 1) が浅いほど R_p が小さくなり、以後 d の増加に伴い R_p が漸増することを明らかにした [1]。これは、ゲートからの垂直方向電界がリセスチャネル構造における縦方向の反転層領域 ch_{recess} (Fig. 2, (a) Shape A) に印加されにくい形状であるためと考えられる。そこで今回、リセスチャネルの形状を、ゲートからの垂直方向電界を印加されやすいように傾斜領域を設けたデバイスを作製し、従来の形状と比較することでその影響を調査および考察した。

2. 試料作製

p 型 Ge (100) 基板の化学洗浄後、Ge に対し低電子障壁を示す TiN (60~70 nm) を全面にスパッタ堆積し、PMA (350 °C-10 min) 処理を施した。次に、CHF₃ ガスによる反応性イオンエッチング (RIE) により TiN を S/D 形状に加工した後そのまま自己整合的に Ge を掘り込み、リセスチャネル構造を形成した。この際、RIE の基板バイアス出力及びチャンバー内圧力をパラメータとしてエッチング部の形状を調整した (Fig. 2, (b) Shape B)。ゲートスタックとして Al/SiO₂/GeO_x (200/50/3 nm) 構造を ECR プラズマプロセスと真空蒸着によって形成した後、PDA (250 °C-30 min) 処理を施した。最後に、S/D 上にコンタクトホールを開口し、Al コンタクト電極を形成した。Fig. 1 に作製したデバイスの断面模式図を示す。MOSFET のチャネル長 L は 40-100 μm、チャネル幅 W は 390 μm、 d は 6.9~73.4 nm である。トランジスタの特性評価から各 d における R_p を算出した。

3. 結果

Fig. 3 に作製したデバイスの伝達特性 (ソース電流 (I_s) -ゲート電圧 (V_G)) を示す。いずれの形状においても非リセスチャネルデバイスに対して ON 電流の向上が確認された。Fig. 4 に寄生抵抗のチャネル深さ依存性を示す。従来の Shape A のデバイスでは、リセスチャネル化 (浅い d) によって R_p が急減した後、 d が深くなるにつれて R_p が漸増していたのに対し、新たに作製した Shape B のデバイスでは、 d が深くなっても低い R_p が保たれた。この原因は縦方向反転層 ch_{recess} とゲート直下反転層 ch の接触領域 $ch_{contact}$ にあると推察される。Shape A の場合、ゲートからの垂直方向電界が ch_{recess} に印加されにくい形状のため、 d の増加に伴い ch_{recess} 深部の領域が狭まり、 $ch_{contact}$ が狭まることで R_p が漸増すると考えられる。Shape B の場合、エッチング部底の傾斜領域が垂直方向電界を印加されやすい形状となっているため、 $ch_{contact}$ は狭まらず低い R_p が保たれたと考えられる。

以上よりリセスチャネル構造では、縦方向反転層 ch_{recess} の形成及びゲート直下反転層 ch との接触領域 $ch_{contact}$ を形成可能なリセスチャネル形状及びゲートスタック構造が有効であると考えられる。

謝辞 本研究は科研費・基盤研究 (S) (No. 19H05616) 及び RIEC 共同プロジェクト (No. R03/A02) の支援を受けた。

参考文献

[1] 那須 他, 第 83 回応用物理学会秋季学術講演会, 20a-A406-9 (2022).

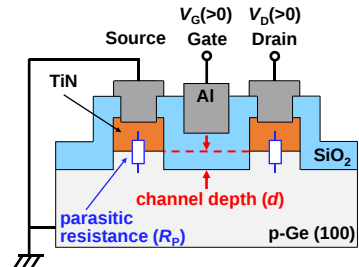


Fig. 1. Cross-sectional image of the metal S/D n-MOSFET.

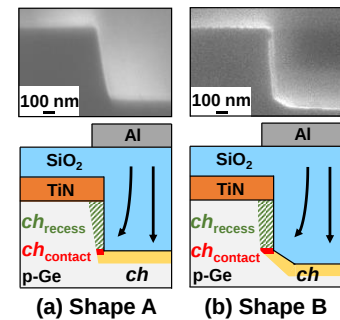


Fig. 2. Mesa shapes observed by SEM with different RIE conditions and cross-sectional image of channel.

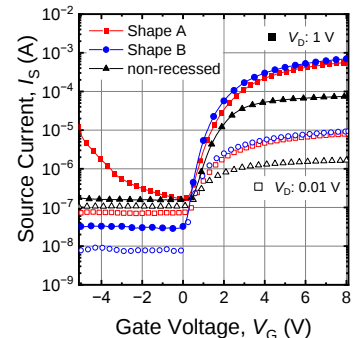


Fig. 3. I_s - V_G characteristics of the Ge n-MOSFETs.

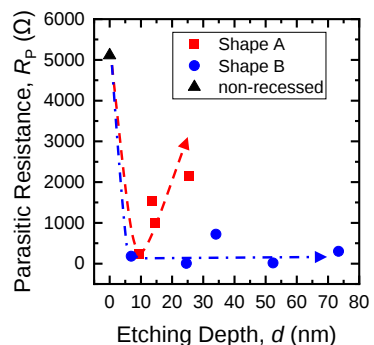


Fig. 4. Etching depth dependence of parasitic resistance.

Performance Enhancement of Extremely-thin Body (111) Ge-on-Insulator nMOSFETs by Using Flipped Substrate Process

[○]X. Han, C.-T. Chen, K. Sumita, K. Toprasertpong, M. Takenaka, and S. Takagi
The University of Tokyo, Faculty of Engineering. E-mail: hanxy@mosfet.t.u-tokyo.ac.jp

1. Introduction Due to the aggressive device scaling, Si logic CMOS is approaching its physical limit. In order to achieve higher integration density, Si FinFET is expected to be replaced by stacked GAA nanosheet [1] and 3D sequential CFET [2], which allow us to integrate hetero channel orientations and materials. (111) Ge is highlighted as one of the optimum choices for future nMOSFETs, thanks to its low in-plane effective mass and high quantization effective mass (m_z) [3], which is especially suitable for extremely-thin body (ETB) channel because of suppression of surface roughness scattering [4]. However, ETB (111) Ge-on-Insulator (GOI) nMOSFET, fabricated by the Smart Cut method [5], suffered severe mobility degradation for the channel thickness less than 30nm [6]. Since the previous work has proved that the post-annealing at 550-600 °C is applicable to the Ge/GeO_x/Al₂O₃ structure without deteriorating the interface quality [7], the quality of Smart Cut GOI, which has been reported to have a defective Ge layer near back interface [8], can be an origin of this severe mobility degradation. In this work, an effective flipped GOI method to employ second-time layer transfer after the smart cut, is proposed to avoid this intrinsic problem of Smart Cut GOI. As a result, flipped (111) GOI nMOSFETs exhibit a clear performance enhancement over a wide range of the channel thickness (<100 nm) compared to the Smart Cut ones, indicating that a superior ETB (111) GOI quality can be achieved by the flipped method.

2. Experiments Fig. 1 shows the schematic fabrication flow of the flipped GOI using a Smart Cut GOI substrate. By this upside-down flipping, a Ge region with higher crystal quality is bonded to another Si/SiO₂ handling wafer. The Ge bonding interface is treated by Plasma Oxidation (P.O.) and ALD Al₂O₃ to minimize the interface states. After removing the top Si substrate and the damaged Ge layer, a flipped GOI with high Ge crystal quality can be obtained. Fig. 2 shows the schematic and fabrication flow of a recessed GOI nMOSFET, which has 100-nm-thick raised source/drain (S/D) to form high donor concentration S/D with low contact resistance. The thickness of thinned channels (0-100 nm) is well controlled by Digital Etching (D.E.) [9]. The top gate stack is formed by P.O. GeO_x and ALD Al₂O₃ with W/Al metal gate followed by PMA.

3. Results and Discussions The I_d - V_g and I_d - V_d characteristics of the Smart-Cut and flipped (111) GOI recessed nMOSFET are compared under similar channel thicknesses in Fig. 3. Here, the channel thickness is estimated by the minimum I_d which is expected to decrease monotonically with thinning the channel thickness. Compared to the Smart Cut GOI nMOSFET, the flipped one shows higher on-current over a wide range of the GOI thicknesses. The 8.1-time higher on-current at V_g of 2 V is observed at ~6-nm-thick channels. In addition, the maximum I_d at V_d of 0.5V and V_g - V_{th} of 0.4 V of the flipped GOI nMOSFET is enhanced by 5 times under the channel thickness of ~13 nm against the Smart Cut one. The minimum subthreshold swing (SS_{min}) of nMOSFETs with the two GOI substrates are compared in Fig. 4. The flipped GOI nMOSFET shows the lower SS_{min} , especially at ETB channel, which is lower than 100 mV/dec, indicating the better interface quality that the Smart Cut GOI, thanks to the flipped process. Fig. 5 summarizes the

field-effect mobility (μ_{FE}) of the flipped and Smart Cut GOI nMOSFETs. The μ_{FE} value of 70 cm²/Vs is obtained at the ~2.4 nm-thick flipped (111) GOI channel.

4. Conclusion We have demonstrated an effective way to improve the ETB (111) GOI quality by the flipped process. The performance of ETB (111) GOI nMOSFETs, fabricated by the proposed process, has been enhanced with higher on-current and electron mobility.

Acknowledgments This work was supported by JSPS KAKENHI Grant Number 22H00208.

References [1] N. Loubet et al., VLSI 2017, p. T230. [2] T. -Z. Hong et al., IEDM 2020, p. 15.5.1. [3] S. Takagi et al., Jpn. J. Appl. Phys. **50**, 010110 (2011). [4] K. Sumita et al., IEEE Trans. Electron Devices **69**, 2115 (2022). [5] C.-M. Lim et al., IEEE Electron Device Lett. **41**, 985 (2020) [6] C.-M. Lim et al., IEEE J. Electron Devices Soc. **9**, 612 (2021). [7] X. Han et al., Jpn. J. Appl. Phys. **62**, SC1089 (2023). [8] K. Yamamoto, et al., ECS Trans. **93**, 73 (2019) [9] X. Yu et al., IEDM 2015, p. 20.

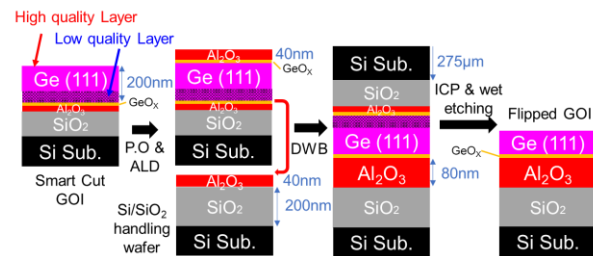


Fig. 1. Schematic and fabrication flow of flipped GOI

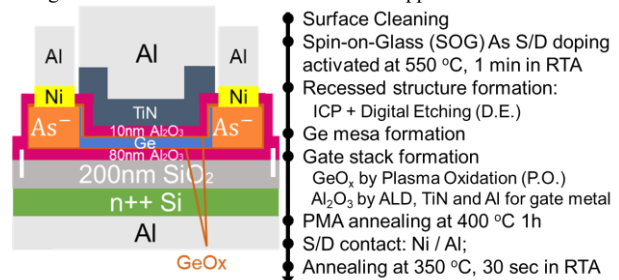


Fig. 2. Schematic and fabrication flow of recessed GOI nMOSFET

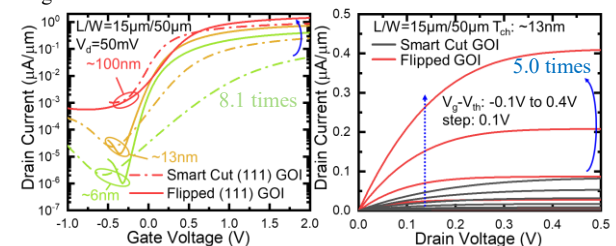


Fig. 3. I_d - V_g and I_d - V_d characteristics of Smart Cut and Flipped (111) GOI nMOSFETs

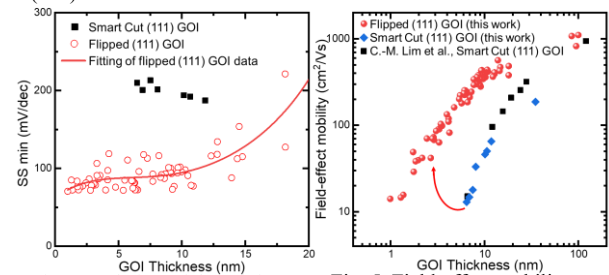


Fig. 4. Relation of the minimum subthreshold swing (SS_{min}) as a function of the estimated channel thickness

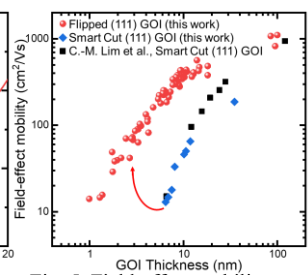


Fig. 5. Field-effect mobility μ_{FE} of the flipped and Smart Cut (111) GOI nMOSFETs

3次元集積デバイス応用に向けた原子層堆積法による

InGaOx チャネルナノシートトランジスタ

○日掛 凱斗¹, 李 卓¹, 郝 俊翔¹, パンディ チトラ¹, 更屋 拓哉¹, 平本 俊郎¹,高橋 崇典², 上沼 睦典², 浦岡 行治², 小林 正治^{1,3}東大生研¹, 奈良先端大², 東大 d.lab³

Email: hikake@nano.iis.u-tokyo.ac.jp

1. 背景と目的: アモルファスIGZO に代表されるアモルファス酸化物半導体は、透明で、比較的高い移動度を持ち、スパッタ成膜により室温で大面積に成膜出来るという優れた特性を活かして[1]、フラットパネルディスプレイ (FPD) 向けデバイスとして広く用いられている。近年では、それらの優れた特性をLSIデバイスに応用する流れが強まっており、BEOLへの3次元モノリシック集積[2-7]や、3次元構造メモリデバイスへの応用[8,9]が期待されている。しかし、従来よりFPDで用いられてきたスパッタ成膜では、高アスペクト比を持つ構造にコンフォーマルに成膜することが難しいため、原子層堆積法 (ALD) を用いて成膜する必要がある[10]。私たちはALDによるプロセスの制御性と電気特性、そしてLSI応用に向けた熱耐性の面から、InGaOx (IGO) に注目している[11]。本研究の目的は、ALD IGOチャネルFETのデバイス特性を系統的に評価し、ノーマリオフ動作、高移動度、高信頼性を実現するためのインテグレーションプロセスを構築することである。

2. デバイス試作方法: Fig.1(a) に作製したデバイスのプロセスフローを示す。Fig.1(b)とFig.1(c)にはそれぞれデバイスの断面図と断面TEM像、EDX元素分布を示す。シングルゲート(SG) FETはバックゲート構造であり、ダブルゲートナノシート(DG NS) FETはバックゲートとトップゲートを有する。

3. 結果と考察: SG IGO FETの熱安定性を確認するために、IGOをALDで成膜しパターンニングした後のアニール温度を変化させた際の I_d - V_g 特性を示す (Fig. 2)。作製した SG IGO FET は HfO_2 をゲート絶縁膜に用いたときは 500°C まで、 SiO_2 をゲート絶縁膜に用いたときは 600°C まで安定な電気特性が維持できることが分かった。Fig. 3 は、SG IGO FET について、In と Ga の組成比を In:Ga=2:1 から 1:1 まで変化させ、さらに IGO の膜厚を 3nm から 10nm まで変化させたときの、移動度、しきい値電圧、ストレス印加しきい値電圧シフトの関係を示すグラフである。移動度、しきい値電圧、しきい値電圧シフトの間に明確なトレードオフの関係があることが示された。このトレードオフを解消するために、DG NS IGO FET を提案し試作した。Fig. 4 と Fig. 5 に SG IGO FET と DG NS IGO FET の I_d - V_g 特性と実効移動度を示す。DG 構造にすることで、ノーマリオフ動作、SG と比較して 2.6 倍のドライバ電流、1.2 倍の移動度の向上を実証した。Fig. 6 には、SG と DG NS FET の正電圧ストレスおよび負電圧ストレス印加時のしきい値電圧シフトを示す。DG プロセスにおける表面パッシベーションプロセスと DG 動作による電界の減少により DG NS FET のしきい値電圧シフトは SG FET のしきい値電圧シフトと比較して飛躍的に向上した。

4. まとめ: ALD法で成膜したInGaOxをチャネルに用いたFETの特性を系統的に調査し、移動度、電気特性、信頼性の間のトレードオフ関係を明らかにした。また、DG NS IGO FETを作製することにより、ノーマリオフ、高移動度、高信頼性を両立するデバイス動作を実証した。

Acknowledgement: This work was supported by JSPS KAKENHI (21H04549), TSMC Advanced Semiconductor Research Project.

References: [1] K. Nomura, et al., NATURE, VOL 432, (2004). [2] H. Ye et al., IEDM, 613 (2020), [3] J. Wu et al., IEEE TED, 68, 12, 6617 (2021), [4] U. Chand et al., VLSI Symp., T10-2 (2021), [5] Z. Lin, et al., IEDM, 386 (2021), [6] M. Endo et al., IEDM, 134 (2022), [7] S. Subhechha et al., VLSI Symp., 292 (2022), [8] Z. Li et al., IEEE EDL, 43, 8, 1227 (2022), [9] K. Huang, et al., VLSI Symp., 296 (2022), [10] J. Sheng et al., J. Vac. Sci. Technol. A, 36, 6, 060801 (2018), [11] H. J. Yang et al., ACS AMI, 12, 47, 52937-52951 (2020).

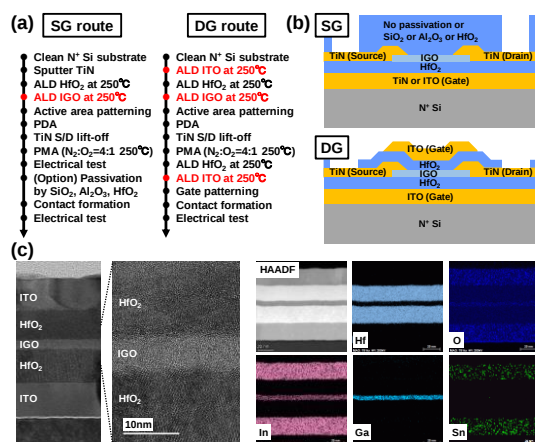


Fig. 1 (a) Device fabrication process flow of fabricated devices and (b) cross section schematics. SG FETs are fabricated with bottom gate, while DG FETs are fabricated with top and bottom gates. IGO and ITO are deposited by ALD process with O_3 . (c) Cross sectional TEM images and EDX mapping of the DG IGO FET. Each layer is conformally deposited and sharp interfaces are confirmed.

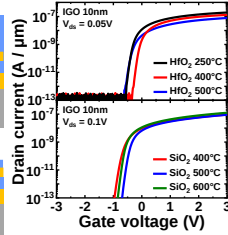


Fig. 2 Measured I_d - V_g curves of SG IGO FETs with HfO_2 and SiO_2 gate oxide. IGO was annealed at different temp.

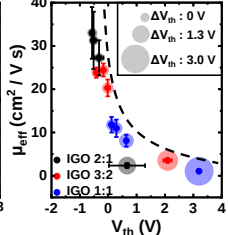


Fig. 3 Trade-off of μ_{eff} , V_{th} , and V_{th} shift in SG IGO FETs with different Ga concentration and IGO thickness.

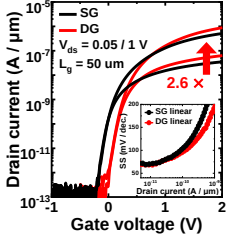


Fig. 4 Measured I_d - V_g of the DG NS IGO FET compared to the SG IGO FET. The inset is I_d -SS.

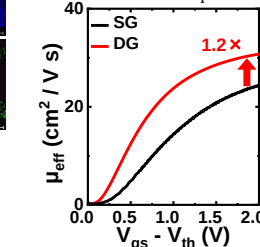


Fig. 5 μ_{eff} characteristics of the fabricated SG and DG IGO FETs.

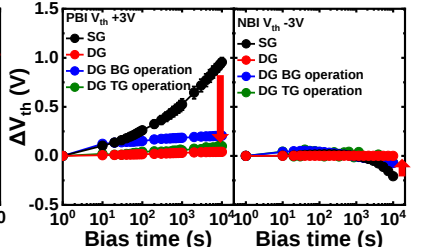


Fig. 6 PBI & NBI characteristics of the fabricated SG and DG IGO FETs.

新型超低電圧リテンションSRAM (ULVR-SRAM)マクロの設計と性能解析

Design and analysis of an 8kB macro for a new ultralow-voltage retention SRAM

東工大・未来研 °伊藤克俊, 塩津勇作, 菅原聡
°K. Ito, Y. Shiotsu, and S. Sugahara, *FIRST, Tokyo Inst. of Tech.*
E-mail: ito.k.bp@m.titech.ac.jp

【はじめに】我々は超低電圧でデータ保持可能な超低電圧リテンションSRAM (ULVR-SRAM)を提案し、その性能解析と応用の検討を進めている[1]. ULVR-SRAMは通常電圧を用いた高性能SRAM動作, エネルギー最小点(EMP)となる0.4Vを用いた動作時電力の削減, 0.2Vの超低電圧リテンション(ULVR)による実質的なパワーゲーティング(PG)を用いた待機時電力の削減が可能である. また, ULVR-SRAMを応用したニューラルネットワークアクセラレータでは, EMP動作によるエネルギー効率(TOPS/W)の最大化も実現できる[2]. しかし, ULVR-SRAMは10個のトランジスタ(Tr)で構成されることから, Tr数(セル面積)の削減が重要な課題となる. 最近, 我々は8個のTrで構成される新型のULVR-SRAMを提案した[3]. このセルでは従来型のULVR-SRAMよりさらにノイズマージン(NM)を拡大できることから, さらなる性能向上が可能である. 前回は, 新型ULVR-SRAMセルの設計とセルレベルの回路性能について報告した. 今回の発表では, 新型ULVR-SRAMの8kBマクロの設計とその性能について報告する.

【回路構成】図1に新型ULVRセルおよびその駆動制御系の構成を示す. 仮想電源電圧(V_{DD}), フィードバックTr (FBT)のバイアス V_{FB} , ワード線電圧 V_{WL} はそれぞれパワースイッチ(PS), コントロールスイッチ(CS), ワード線スイッチ(WS)により制御する. 通常SRAM動作では V_{DD} に V_{DDH} ($=1.2V$), V_{FB} に V_{FBH} ($=0.3V$)を用いる. スタンバイ(SB)では V_{WL} を V_{DDH} とし, ロードパスTr (LPT)のリークを用いて記憶を保持する. READ/WRITE時は V_{WL} を0Vとし, LPTによって読み出し, 書き込みを行う. ULVR時には V_{DD} を V_{DDL} ($=0.2V$), V_{WL} を V_{WLL} ($=0.2V$), V_{FB} を V_{FBL} ($=0.2V$)とし, このモードで生じる本セル特有の強いフィードバックの効果により高いNMを確保できる. WS_1 は V_{WLL} からデコーダ側に生じるリーク電流を遮断するために用いる. 図2に新型ULVR-SRAMマクロ(M[8T^{ULVR}])のレイアウトを示す. セルと周辺回路は従来型ULVR-SRAMマクロ(M[10T^{ULVR}])を参考にして配置した[1]. 今回用いたロジックデザインルールの制約とM[10T^{ULVR}]の周辺回路を流用することを考慮して, セルおよびアレイサイズをM[10T^{ULVR}]とほぼ同じとした. 今後, メモリデザインルールを用いれば, Tr数の削減率と同程度にまでセル・アレイサイズを削減できる可能性がある. セルおよび周辺回路はHSPICEを用いて設計し, マクロの動作と性能は寄生抵抗・容量を抽出してから, 高速SPICE (FineSim)による大規模シミュレーションを用いて解析した. デバイスには65nm CMOSプロセスのLPモデルを用いた.

【解析結果】新型ULVRセルは全動作・全プロセスコーナーで130 mV以上のNMを確保できる. また, ローカルばらつきを考慮しても全動作で6 σ の故障率を満足できる. $V_{DD} = 1.2 V$ では動作周波数500 MHzでマクロ上のどの空間コーナー/プロセスコーナーでも正常に動作した. SRAM動作時の V_{DD} はNMを考慮して0.6Vまで低減可能であった. このとき, 動作時電力は90%程度削減できる. 低電圧動作用に特化したセル設計を用いることで, さらに低い V_{DD} でもSRAM動作が可能となり, 動作時電力もさらに削減できる. 図3にM[10T^{ULVR}], M[8T^{ULVR}], および6T-SRAMマクロ(M[6T])の待機時電力の解析結果を示す. M[8T^{ULVR}]は0.2VのULVRモードにより自身のSBモードから89%の待機時電力削減が可能である. また, M[10T^{ULVR}]と比較して, SB, ULVRモードともに待機時電力をさらに20%程度削減できる. 新型ULVR-SRAMはTr数の削減, ULVRにおけるNMの向上, そして待機時電力のさらなる削減が可能である.

【謝辞】シミュレーションは東京大学大規模集積システム設計教育センター(VDEC)を通しシノプシス株式会社の協力で行われたものである.
【参考文献】[1] H. Yoshida *et al.*, IEEE OJCAS 2, pp. 520-533, 2021. [2] Y. Shiotsu *et al.*, IEEE JXCDC 8, pp. 134-144, 2022. [3] 伊藤他, 第83回応用物理学会秋季学術講演会, 21a-C105-5, 2022.

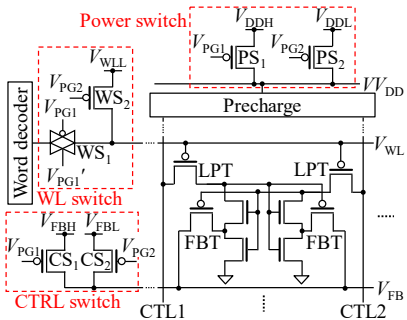


図1. 新型ULVRセルの構成

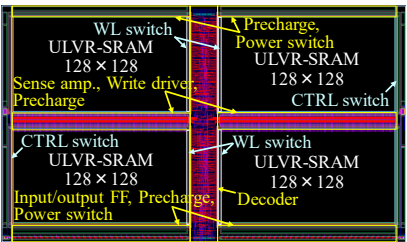


図2. 新型ULVRマクロのレイアウト

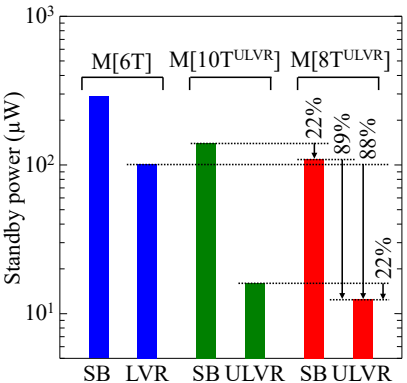


図3. 各マクロの待機時電力

並列化MACユニットを有するULVR-SRAMを用いたBNNアクセラレータマクロ

A binarized neural network accelerator macro with parallelized MAC units using ULVR-SRAM

東工大・未来研 °塩津勇作, 菅原聡

°Y. Shiotsu, and S. Sugahara, *FIRST, Tokyo Inst. of Tech.*

E-mail: y.shiotsu@isl.titech.ac.jp

【はじめに】 将来のスマート社会ではモバイルエッジ階層のコンピューティングシステムにAI技術を導入することが期待されており、これまで以上にエネルギー効率の高いニューラルネットワーク(NN)アクセラレータが要求される。Processing-in-memory (PIM)型のハードウェアはNNアクセラレータの高性能化に有効である。特に、SRAMを用いたPIM型NNアクセラレータは、現状のロジックCMOS技術で実装が可能なため、応用上重要である。このようなNNアクセラレータの性能向上には、エネルギー最小点(EMP)となる駆動電圧(V_{EMP})を用いた推論動作が極めて効果的であるが、この実現は容易ではない。最近、我々は従来のCMOS技術で実装され、 V_{EMP} でのSRAM動作と超低電圧(V_{UL})を用いたパワーゲーティングが可能なULVR-SRAM [1]を用いて、PIM型2値化NNアクセラレータ(BNA)マクロを開発した[2]。このBNAマクロを複数個用いることで任意のサイズ・形状のネットワークを構成できる。特に、BNAマクロに複数の積和(MAC)演算ユニットを導入し、EMP動作を用いた推論動作を並列化することで、エネルギー効率(TOPS/W)が高く、処理性能(TOPS)の高いアクセラレータを構成できる[2]。本発表では、ULVR-SRAMを用いた8並列MACユニットを有するBNAマクロの設計と性能について議論する。

【8並列BNAマクロ】 本BNAマクロでは、XNORベースのMAC演算ユニットを用いた[2]。MAC演算結果に整数バイアスを加えた結果の最上位ビット(MSB)を用いてアクティベーション(ACTV)を判定する[2]。マルチポートセルを用いることなく配線の工夫のみで重みデータと整数バイアスの読み出しの並列化を行い、複数のMACユニットを実装することでマクロ内のMAC演算の並列化を実装した。図1に開発した8並列MACユニット内蔵BNAマクロのレイアウトを示す。このマクロでは256bの入力ベクトルに対するMAC演算を8並列で同時に実行できる。設計したマクロは、1並列のマクロ[2]と比較して、増設した読み出し系(センスアンプ、FFなど)およびMACユニットの分だけ全体の面積が増大したが、SRAMアレイ部は1並列のマクロとほぼ同じ面積で実装することができた。

【解析結果】 設計したマクロの性能は、寄生抵抗・容量を抽出して高速SPICE (FineSim)を用いた大規模シミュレーションにより解析した。デバイスには65nm CMOSプロセスのLPモデルを用いた。図2に8並列MACユニットを有するBNAマクロにおける動作周波数 f 、平均電力 P_{avg} 、1サイクル当たりの消費エネルギー E に関する仮想電源電圧 V_{DD} 依存性を示す。 E は $V_{DD}=0.4V$ で最小となっている($V_{EMP}=0.4V$)。このとき、 f は450MHzから25MHzに低下するが、 P_{avg} を99%削減できる。また、推論のエネルギー効率 η は V_{EMP} で最大となり、104TOPS/Wと極めて高い値を実現できる。図3に8並列MACユニットを有するBNAマクロのEMP動作時の性能を示す。同図には1並列MACユニットを有するマクロ[2]の結果も示す。1並列MACユニットのBNAマクロをEMP (0.4V)で動作させると、通常電圧動作(1.2V)と比較して、 f は減少するが、 P_{avg} はさらに大幅に削減できる。この結果、 η は1.2V動作の~10TOPS/WからEMP動作の65TOPS/Wに大幅に改善する。8並列MACユニットのBNAマクロのEMP動作では、1並列の場合と比べて f がわずかに減少するが、 P_{avg} は周辺回路の効率化の効果によりMACユニットの並列数ほどには増加しない。したがって、8並列MACユニットのマクロでは η はさらに改善され104TOPS/Wまで増大する。次に、8並列MACユニットのBNAマクロを16個用いて構成される1024ノードの全結合層[2]の評価を行った。この全結合層全体のMAC演算の並列数は128となる。この構成では、高いエネルギー効率を維持したまま(~100TOPS/W)、処理性能を1.5TOPSまで向上できる。したがって、並列化MACユニットを有するBNAマクロを用いることで、高性能のBNAを構成できる。

【謝辞】シミュレーションは東京大学大規模集積システム設計教育センター(VDEC)を通しシノプシス株式会社の協力で行われたものである。

【参考文献】 [1] H. Yoshida *et al.*, IEEE OJCS 2, 2021, p. 520. [2] Y. Shiotsu *et al.*, IEEE JXDC 8, 2022, p. 134.

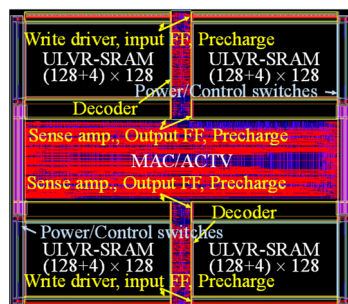


図 1. 8 並列 BNA マクロのレイアウト

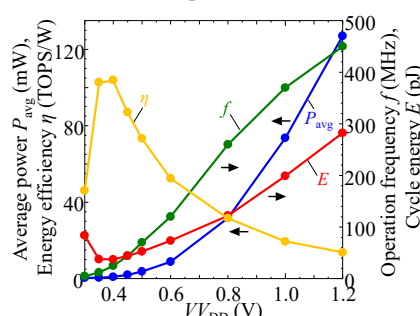


図 2. P_{avg} , f , E , η の V_{DD} 依存性

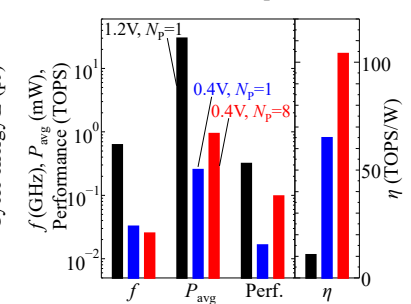


図 3. 1 並列と 8 並列 BNA マクロの性能