

Point mémoire analogique basse tension basé sur l'effet GIDL

Nicolas Donckers, Olivier Butaye, Denis Flandre et Michel Verleysen
Laboratoire de microélectronique
Université catholique de Louvain
Place du Levant, 3
B-1348 Louvain-la-Neuve - Belgique

Email: {donckers,flandre,verleysen}@dice.ucl.ac.be

Résumé

La mémoire est un élément incontournable de tout circuit électronique; le développement des applications analogiques basse puissance et basse tension sera conditionné par la disponibilité de points mémoire présentant ces mêmes caractéristiques. Dans cet article, nous examinons les possibilités de réaliser un point mémoire analogique permanent en technologie SOI qui soit adapté aux circuits basse puissance, alimentés sous faible tension. Nous envisageons l'utilisation de transistors à grille flottante et montrons comment ces mémoires peuvent être programmées à l'aide de courants circulant à travers l'oxyde de grille. Nous utilisons l'effet GIDL afin de produire ce courant de grille ce qui, à notre connaissance, n'avait jamais été envisagé.

1. Introduction

La mémoire est un élément incontournable de tout circuit électronique. A l'heure actuelle, les mémoires digitales occupent la quasi totalité du marché parce que ce sont des circuits digitaux qui sont présents dans la majorité des applications.

Dans certaines situations, les circuits analogiques se sont pourtant révélés mieux adaptés. On citera par exemple la logique floue ou les réseaux de neurones. Plus récemment, l'idée d'utiliser des circuits analogiques pour réaliser des récepteurs itératifs dédiés au décodage de codes correcteurs d'erreurs s'est révélée très efficace [14]. Ces codes sont appelés à être utilisés dans les communications mobiles de troisième génération. Les circuits qui les manipulent seront alors soumis aux contraintes, de plus en plus sévères, de limitation de puissance consommée et de faible tension d'alimentation propres aux applications embarquées. Le développement de ces techniques est toutefois lié à la disponibilité de points mémoires compatibles avec ces contraintes.

Dans cet article, nous étudions la possibilité de réaliser des points mémoires analogiques adaptés aux applications embarquées. Le concept proposé se base sur l'injection de porteurs à travers une couche d'oxyde permettant de stocker, en un point électriquement isolé du reste du circuit, une quantité de charges fonction de la valeur à mémoriser.

L'utilisation de points mémoires analogiques permet d'éviter le recours aux convertisseurs A/D. Le choix de

mémoires permanentes évitera de devoir utiliser un circuit de rafraîchissement. Ces mémoires devront pouvoir être programmées électriquement, sans qu'une consommation importante de puissance ne soit nécessaire. Enfin, l'utilisation d'une technologie CMOS standard, permettant de réaliser un circuit à faible coût, sera également un argument à considérer.

La suite de cet article est structurée en 3 parties. Nous commençons par évoquer les architectures permettant d'obtenir des points mémoire flottants dans un circuit électronique ainsi que les phénomènes pouvant donner lieu à un courant à travers une couche d'oxyde. Dans la deuxième partie, le principe de notre architecture est présenté et validé par simulation. Enfin, nous terminons par des perspectives sur l'avenir de cette technique.

2. Points mémoire permanents

Il existe plusieurs possibilités permettant de réaliser un point mémoire analogique permanent. Celles reposant sur le principe d'une grille flottante sont les plus utilisées et conviennent bien à l'application que nous envisageons.

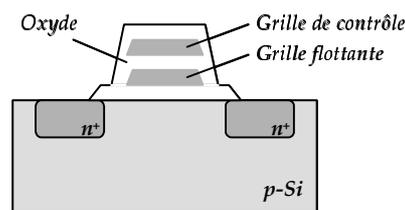


Figure 1. Transistor à grille flottante

2.1 Les transistors à grille flottante

Un transistor à grille flottante [5][6] (FGMOS – *floating gate MOS transistor*) comporte deux grilles (figure 1). La première, la grille de contrôle permet l'application d'une tension extérieure. La seconde, la grille flottante est électriquement isolée du reste du dispositif. Lorsqu'une tension suffisante est appliquée à la grille de contrôle, des charges peuvent traverser l'oxyde de grille et se retrouvent piégées sur la grille flottante. La tension de seuil du transistor est alors modifiée par la charge présente sur la grille.

La mise en œuvre de ce principe nécessite la réalisation de transistors à double grille (figure 1). L'un de nos objectifs étant de réaliser un circuit dans une technologie standard à faible coût, elle ne peut être retenue.

On peut électriquement isoler un point du reste du circuit en plaçant une capacité en série avec la grille du transistor. Le schéma équivalent obtenu est alors celui de la figure 2. La programmation du point mémoire nécessite dans ce cas le recours à un courant traversant la couche d'oxyde qui sépare le canal du transistor de la grille flottante. Plusieurs phénomènes permettent de l'obtenir.

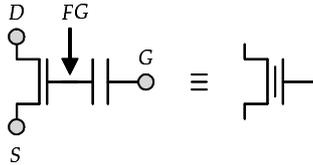


Figure 2. Modèle circuit

2.2 Courants de grille

Dans les circuits traditionnels, les courants traversant l'oxyde de grille sont considérés comme des courants de fuite qu'il s'agit de minimiser. Dans notre cas, ils constitueront l'effet utile. Traditionnellement, ces courants sont obtenus par effet Tunnel [8] ou par la technique des porteurs chauds [7][9]. Cette dernière consomme un courant de drain important parce qu'elle s'utilise lorsque le transistor est en régime saturé. L'utilisation du courant dit GIDL permet d'obtenir des courants de grille comparables lorsque le transistor est en régime sous-seuil (subbreakdown drain current) [1][8]. Les tensions nécessaires sont du même ordre de grandeur que celles permettant l'apparition de porteurs chauds.

2.2.1 Injection de charges par effet Tunnel.

L'injection de charges par effet Tunnel consiste à appliquer une différence de potentiel importante entre la grille et le canal d'un transistor MOS pour que les porteurs présents à la surface de celui-ci puissent traverser, par effet Tunnel, la barrière de potentiel. Ce phénomène n'est que partiellement bien adapté à l'application que nous envisageons. En effet, lors de son utilisation, une importante différence de potentiel est nécessaire entre la grille et le canal, ce qui n'est pas souhaitable. A l'opposé, la puissance consommée est faible, le courant de drain circulant dans le dispositif étant nul. En effet, dans cette configuration, la source et le drain du transistor sont court-circuités.

2.2.2 Effet GIDL

Cet effet apparaît dans la partie du drain recouverte par la grille du transistor (zone de recouvrement). Dans un transistor de type NMOS, lorsque la grille est portée à un potentiel négatif, une zone d'accumulation de trous apparaît à la surface du canal. Elle se comporte alors comme une zone plus fortement dopée que le substrat. Si la tension de grille devient fortement négative, la zone de recouvrement peut devenir déplétée, voire même passer en régime d'inversion (figure 3). Si le drain est amené à une tension suffisante, le champ électrique créé dans la zone de recouvrement peut être assez intense pour permettre aux

électrons, devenus porteurs minoritaires, de créer, par impact, des paires électron - trou.

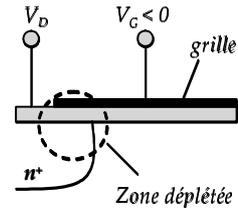


Figure 3. Effet GIDL

On constate alors l'apparition d'un double effet. D'une part, les porteurs créés par impact peuvent traverser l'oxyde de grille et donner lieu à un courant de grille. D'autre part, le champ peut devenir suffisamment intense pour créer un courant de drain par effet Tunnel à travers la zone déplétée recouverte par la grille [11][12]. Dans la technologie CMOS sur substrat massif, cet effet est à l'origine d'un courant dans le substrat. On cherche donc toujours à le minimiser.

3. Architecture du point mémoire

Le concept de point mémoire proposé est représenté à la figure 4. Un transistor de lecture M_L (monté en drain commun) est traversé par un courant fonction de la tension présente sur la grille flottante. Ce courant est comparé à la valeur I_{ref} que l'on souhaite mémoriser. Tant que les valeurs de ces courants ne sont pas suffisamment proches l'une de l'autre, le comparateur fournit une tension de sortie capable de modifier la charge présente sur la grille flottante par l'injection de charges sur cette grille à travers deux sources de courant commandées.

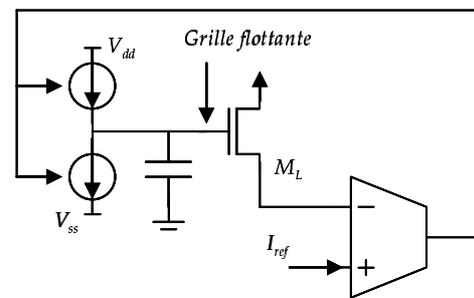


Figure 4. Schéma de principe du point mémoire

3.1 Comparateur de courant

Un schéma de principe du comparateur de courant utilisé est représenté à la figure 5. Les miroirs de courants utilisés ont été cascodés afin que le gain du comparateur soit suffisant. La résolution atteinte est de l'ordre du pourcent du courant de polarisation. Cette structure de miroir de courant est particulièrement adaptée à notre application. D'une part, la tension d'alimentation minimale garantissant son bon fonctionnement est limitée. D'autre part, l'impédance de sortie du circuit peut être arbitrairement augmentée en cascasant plus de deux étages de transistors [2].

On notera qu'étant donné la valeur des courants de grille disponibles (section 3.2), la programmation du point mémoire est un phénomène particulièrement lent. Le temps

de programmation est de l'ordre de la milliseconde. La bande passante du comparateur n'est donc aucunement un élément limitant. En revanche, le gain du comparateur est un élément important puisqu'il conditionne la précision avec laquelle on peut contrôler la charge présente sur la grille flottante.

Le gain de ce comparateur peut être aisément déduit:

$$\frac{V_{out}}{i_{ref} - i_{ML}} \approx g_m \cdot r_o^2$$

où g_m est la transconductance des transistors et r_o leur impédance de sortie.

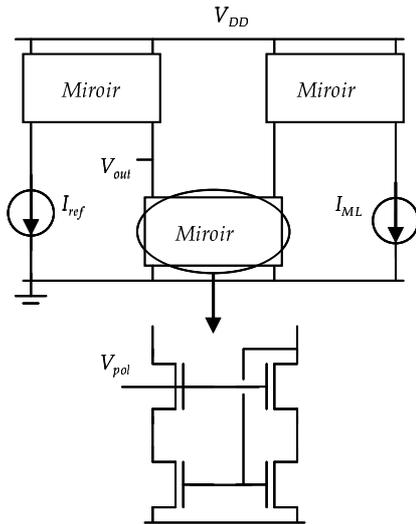


Figure 5. Schéma du comparateur de courant

La programmation du point mémoire utilise deux phénomènes. D'une part, un courant négatif (permettant de charger le nœud flottant) sera obtenu par effet Tunnel en utilisant un transistor de type N. D'autre part, l'effet GIDL à travers un transistor de type P permettra de diminuer la tension présente sur la grille flottante. L'architecture de notre point mémoire impose que ces deux phénomènes puissent être observés pour une même valeur de la tension de grille.

3.2 Mesures des courants de grille

La figure 6 représente les courants de grille mesurés sur un transistor de type N réalisé dans une technologie SOI 2 μm présentant une épaisseur d'oxyde de grille de 30 nm. Le courant obtenu est négatif, l'échelle verticale logarithmique ayant obligé à en modifier le signe.

On constate que l'on est bien en présence d'un phénomène d'injection de charges provoqué par l'effet Tunnel. Les tensions appliquées aux différents accès (voir figure 6) sont compatibles avec ce phénomène. De plus, la relation linéaire entre $\log(I_g)$ et V_g confirme le caractère exponentiel des relations présentées dans [8][10]. La valeur maximale du courant de drain pour l'ensemble des mesures de la figure 6 et de l'ordre du nano ampère. Cette constatation permet de confirmer la faible consommation de puissance du dispositif. Bien qu'une tension très élevée (environ 25 V) soit nécessaire pour voir apparaître l'effet Tunnel, ce dispositif est un candidat de choix pour réaliser le point mémoire nécessaire à notre application.

La figure 7 illustre les mesures effectuées sur le transistor de type P. On constate que l'on est face à une injection de porteurs à travers l'oxyde de type GIDL. Les courbes de mesures de courant de drain confirment que les courants de grille maximaux sont obtenus alors que le transistor est bloqué. Le courant de drain qui le traverse est, comme on s'y attendait (cfr. § 2.2.2), plus élevé: de l'ordre de la dizaine de micro ampère. De ce point de vue l'effet GIDL présente de moins bonnes caractéristiques. Par contre, on constate que les tensions nécessaires à l'apparition d'un courant Tunnel sont nettement supérieures à celles permettant d'observer un courant GIDL du même ordre de grandeur.

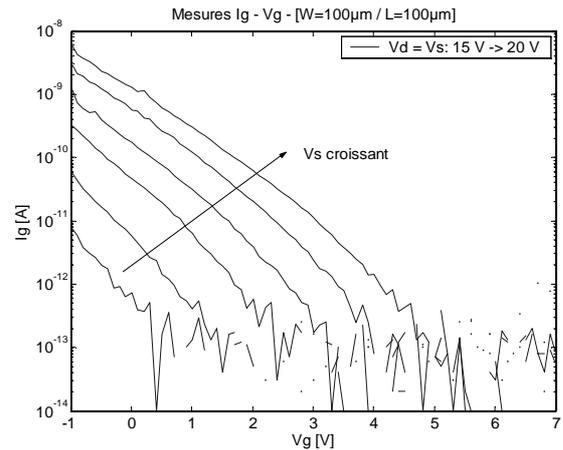


Figure 6. Mesures de courant de grille (NMOS)

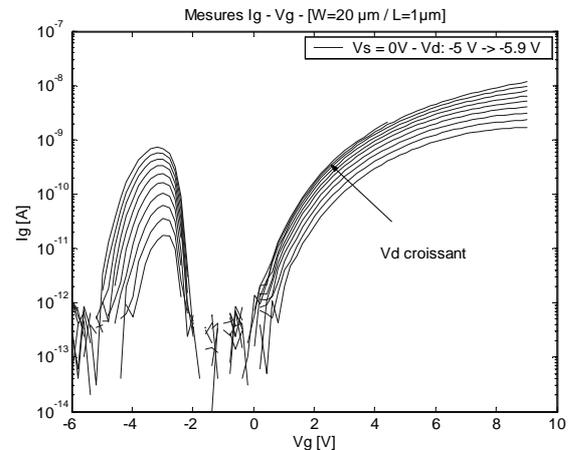


Figure 7. Mesures de courant de grille (PMOS)

3.3 Validation du concept

L'architecture a été validée par simulation sous ELDO pour une technologie SOI 2 μm . Les mesures représentées sur les graphes de la section précédente ont permis d'extraire des modèles approximant les courants de grille. Des sources de courant commandées ont ensuite été ajoutées entre le drain et la grille des transistors.

La tension sur la grille flottante se situe entre 3 V et 4 V. On constate toutefois que la précision avec laquelle la valeur stockée peut être contrôlée est limitée. Cette précision peut être évaluée en comparant le courant de référence avec celui sortant du transistor de lecture.

La résolution du comparateur est un élément important. Il est à noter que, ce dernier reposant sur l'utilisation de

miroirs de courant, les transistors doivent travailler en forte inversion afin d'améliorer l'appariement.

Enfin, l'utilisation d'une grille flottante fait inévitablement apparaître des effets d'injection de charges lors des commutations de la sortie du comparateur. Une capacité supplémentaire a été ajoutée au niveau du nœud flottant pour minimiser cet effet.

Le circuit nécessite une alimentation asymétrique afin de produire les 25 V nécessaires à l'apparition de l'effet Tunnel. Ces 25 V ne sont toutefois nécessaires qu'aux bornes d'un seul transistor. La consommation de ce dernier étant extrêmement réduite, la tension nécessaire peut être produite par un multiplieur de Dickson, également inclus dans notre simulation.

4. Perspectives

L'effet GIDL semble être particulièrement adapté à la réalisation de points mémoire analogiques. En effet, il permet d'obtenir des courants de grille sous des tensions bien moins élevées que celles nécessaires à faire apparaître l'effet Tunnel. L'effet GIDL s'obtient lorsque le transistor est bloqué. Par rapport à l'injection de charges au moyen de porteurs chauds, la puissance nécessaire est donc fortement réduite. Certains problèmes subsistent toutefois.

L'architecture proposée est mixte. Elle utilise à la fois l'effet Tunnel et le courant de type GIDL afin de pouvoir produire, vers la grille flottante, des courants tant positifs que négatifs. Nous n'avons pas pu mettre en évidence de courants de type GIDL dans des transistors NMOS présentant une épaisseur d'oxyde de 30 nm. Dans ces derniers, ce sont les trous qui devraient traverser l'oxyde de grille. Une énergie très importante est donc nécessaire. Il est possible de réduire la hauteur de cette barrière de potentiel en réduisant l'épaisseur de l'oxyde de grille. Des essais sur des transistors présentant une épaisseur d'oxyde de 2.5 nm se sont également révélés négatifs. L'utilisation de l'effet Tunnel semble donc incontournable pour extraire des charges de la grille flottante. On peut toutefois noter qu'une diminution de l'épaisseur de l'oxyde de grille profitera au courant de grille obtenu tant par effet Tunnel que par effet GIDL.

Dans ce contexte, le multiplieur de tension de type Dickson utilise les diodes de type ULP présentées dans [3]. La très faible consommation du transistor dans lequel apparaît l'effet Tunnel rend l'utilisation de tels circuits possible. De plus, les courants de fuite des diodes ULP étant très faibles, la fréquence de l'horloge nécessaire au bon fonctionnement du multiplieur peut être également faible (typiquement 10 Hz), diminuant d'autant la consommation totale du circuit [13].

Le courant de drain créé dans le transistor soumis à l'effet GIDL est un effet indésirable que l'on doit tenter de minimiser. Comme suggéré dans [3], il est possible, dans une technologie SOI standard, de réaliser des transistors présentant différents dopages. Il sera intéressant de comparer les rapports I_g/I_d pour ces différents dispositifs.

5. Conclusion

Une architecture, utilisant un transistor à grille flottante, a été proposée pour réaliser un point mémoire analogique

permanent pouvant fonctionner dans une application basse tension. Elle s'est montrée performante à plusieurs égards.

D'une part, les différents blocs qui la composent sont tous bien adaptés à un fonctionnement sous faible tension d'alimentation.

D'autre part, nous avons mis en évidence que l'utilisation de l'effet GIDL pouvait être avantageuse pour injecter des charges à travers un oxyde de grille sans demander ni tension particulièrement élevée ni courant important.

Références

- [1] T. Brozek, V. Ramgopal Rao, A. Sridharan, J. Werking, Y. D. Chan, and C. Viswanathan. "Charge injection using gate-induced-drain-leakage current for characterization of plasma edge damage in cmos devices." *IEEE Transactions on Semiconductor Manufacturing*, 11(2):211–216, May 1998.
- [2] P. Crawley and G. Roberts. "High-swing MOS current mirror with arbitrarily high output resistance." *Electronics Letters*, 28(4):361–363, February 1992.
- [3] D. Levacq, C. Liber, V. Dessard, D. Flandre "Composite ULP diode fabrication, modeling and applications in multi-Vth FD SOI CMOS technology" *Solid State Electronics* 48 (2004), pp. 1017-1025.
- [4] John F. Dickson. "On-chip high-voltage generation in mmos integrated circuits using an improved voltage multiplier technique." *IEEE Journal of Solid-State Circuits*, sc-11(3):374–378, June 1976.
- [5] C. Diorio, S. Mahajan, P. Hasler, B. Minch, and C. Mead. "A high resolution nonvolatile analog memory cell." *Proceedings of the 1995 IEEE International Symposium on Circuits and Systems*, volume 3, pages 2233–2236., 1995.
- [6] R. Harrison, J. Bragg, P. Hasler, B. Minch, and S. Deweerth. "A cmos programmable analog memory-cell array using floating-gate circuits." *IEEE Transactions on circuits and systems-II : Analog and digital signal processing*, 48(1):4–11, January 2001.
- [7] T-C Ong, P-K Ko, and C. Hu. "Hot-carrier current modeling and device degradation in surface-channel p-mosfet's." *IEEE Transactions on Electron Devices*, 37(7):1658–1666, July 1990.
- [8] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand. "Leakage current mechanisms and leakage reduction techniques in deep submicrometer CMOS circuits." *Proceedings of the IEEE*, 91(2):305–327.
- [9] S. Tam, P-K Ko, and C. Hu. "Lucky-electron model of channel hole-electron injection in mosfet's." *IEEE Transactions on Electron Devices*, 31(9):1116–1125, September 1984.
- [10] L. Richard Carley. "Trimming analog circuits using floating-gate analog mos memory." *IEEE Journal of Solid-State Circuits*, 24(6):1569–1575, December 1989.
- [11] J. Chen, T. Y. Chan, I. C. Chen, P. K. Ko and Chemming Hu, "Subbreakdown drain leakage current in MOSFET" *IEEE Electron device letters*, 8 (11): 515-517, November 1987
- [12] A. Bouhdada, A. Touhami, S. Bakkali "New model of gate-induced drain current density in an NMOS transistor" *Microelectronics Journal* 29 (1998): 813-816
- [13] D. Levacq, C. Liber, V. Dessard and D. Flandre, "Ultra Low-Power design techniques using special SOI MOS diodes," 2003 IEEE International SOI Conference, Newport Beach, USA, September 29 - October 2nd, 2003.
- [14] V. Gaudet, G. Gulak "A 13.3 Mb/s 0.35 μm CMOS Analog Turbo Decoder IC With a Configurable Interleaver" *IEEE JSSC*, vol. 38, n° 11 – Novembre 2003.