

Conception optimale et réutilisable d'OTAs pour dispositifs médicaux implantables.

Fernando Silveira¹, Denis Flandre², Pablo Aguirre¹.

¹Instituto de Ingeniería Eléctrica
Universidad de la República
J. Herrera y Reissig 565
11300 Montevideo, Uruguay

²Laboratoire de microélectronique
Université catholique de Louvain
3, place du Levant
1348 Louvain-la-Neuve, Belgique

E-mail: silveira@fing.edu.uy

Résumé

Cet article analyse la conception de circuits analogiques orientée vers l'optimisation de la puissance et la réutilisation. Le premier aspect considéré est la minimisation de la consommation de puissance d'un amplificateur opérationnel de transconductance (OTA) pour un temps d'établissement total (portions linéaire plus rampe) donné. La réutilisation est illustrée dans la conception d'un amplificateur opérationnel rail-to-rail programmable par le courant de polarisation sur deux ordres de grandeur de fréquence de transition. L'outil de base est une méthodologie de conception qui traite le transistor MOS de façon unifiée dans tous les régimes de fonctionnement et permet une exploration aisée de l'espace de conception. Le cadre général qui fixe les spécifications pour les circuits analysés sont les applications micropuissance, particulièrement les pacemakers implantables.

1. Introduction. Les besoins des pacemakers implantables.

Les pacemakers constituent aujourd'hui le dispositif médical implantable actif le plus répandu. Un pacemaker cardiaque a pour but principal de rétablir le rythme normal d'un cœur malade qui présente une cadence de contraction trop lente. Ceci est réalisé en stimulant électriquement la contraction du cœur. En plus de cette fonction basique de stimulation, les pacemakers ajoutent plusieurs fonctions telles que la détection de la contraction spontanée du cœur, une liaison téléométrique pour recevoir des paramètres de configuration et transmettre des données d'état du dispositif et des statistiques d'opération; la capacité de mesurer la tension de la batterie et des moyens d'estimation de l'activité physique du patient (normalement en utilisant un accéléromètre) afin d'y adapter le rythme cardiaque. Les signaux analogiques traités dans ces systèmes concernent des bandes de fréquence allant jusqu'à 200Hz pour les signaux cardiaques et jusqu'à moins de 10Hz pour les signaux d'accélération nécessaires pour estimer l'activité physique.

Quelles sont les restrictions imposées, du point de vue de la consommation, par ces systèmes ? Deux aspects critiques doivent être considérés : la durée de vie et la taille physique du dispositif (liée à la taille de la batterie

nécessaire). La source d'alimentation des pacemakers est une batterie lithium-iode avec une tension nominale de 2.8V. Une valeur représentative de la capacité énergétique de la batterie, dans les systèmes actuels qui favorisent plus la réduction de la taille que l'augmentation de la durée, s'élève à 1A.h, équivalent à 114 μ A-année. En conséquence, pour viser une durée de dix ans, la consommation totale moyenne ne peut dépasser les 11.4 μ A. D'autre part, l'énergie requise afin de stimuler le cœur dépend de l'amplitude et de la périodicité des impulsions nécessaires pour le patient. Une valeur typique peut atteindre environ 3 μ A. Cette valeur est doublée s'il est nécessaire de stimuler les deux moitiés du cœur (ventricules et oreillettes) comme dans certaines pathologies traitées par des pacemakers dits bicaméraux. Ces valeurs montrent que du point de vue de la capacité de la batterie et comparé à l'énergie délivrée au cœur, la consommation totale du circuit intégré ne doit pas dépasser quelques μ A. Une partie considérable de cette consommation est liée aux circuits analogiques qui doivent rester actifs la plupart du temps, comme le circuit d'estimation de l'activité physique du patient et le circuit (ou circuits dans le cas des pacemakers bicaméraux) de détection du signal cardiaque. Ce cadre général, où les économies de fractions de μ A sont très significatives, sert de motivation au travail orienté vers la minimisation de la puissance consommée, présenté dans cet article. Par ailleurs, les fonctions analogiques d'amplification et filtrage appliquées dans les pacemakers sont représentatives des circuits nécessaires dans la plupart des systèmes biomédicaux en vue de réaliser la détection qualitative du comportement du système biologique et d'agir en conséquence.

Du point de vue de la tension d'alimentation, les circuits doivent opérer de la valeur nominale de 2.8V jusqu'à 2V afin d'assurer le fonctionnement pendant temps suffisant entre la détection de la condition d'épuisement de la batterie et le remplacement du dispositif. Dans la suite de l'article, le cas le plus défavorable de 2V d'alimentation sera considéré.

Le deuxième aspect analysé dans cet article concerne la conception visant la réutilisation des circuits analogiques. Ce besoin est certainement partagé par la plupart des systèmes électroniques, sous la contrainte de réduction du temps de mise sur le marché de nouveaux produits. Dans notre travail, nous étudions la conception d'un amplificateur opérationnel programmable dans la

gamme de fréquences d'intérêt pour les systèmes biomédicaux et compatible avec la tension d'alimentation de ces systèmes.

Dans la section suivante nous rappelons très brièvement la méthodologie de conception utilisée. La section 3 est consacrée à la minimisation de la puissance pour un temps total d'établissement donné, tandis que l'aspect de réutilisation est examiné à la section 4. La section 5 résume les principales conclusions.

2. Méthodologie unifiée de l'inversion faible à l'inversion forte.

Les résultats présentés ci-dessous reposent sur la possibilité d'exploiter le transistor MOS dans toutes ses régions d'opération, sans limitations artificiellement introduites par le modèle du transistor ou la méthodologie de conception. De cette façon, l'espace complet de conception peut être exploré en vue de trouver le minimum de consommation du circuit dans une large plage de spécifications de fréquence, gain.... Cette démarche de conception se base sur l'application de la relation g_m/I_D comme variable clé, en la reliant à tous les aspects de performance des circuits analogiques et de dimensionnement des transistors qui peuvent être énoncés sous forme symbolique [1, 2]. Pour chacun des types de transistors disponibles dans une technologie donnée, la courbe g_m/I_D représentée en fonction du courant normalisé par rapport au facteur de forme, i.e. $I_D/(W/L)$, apparaît, en effet, au premier ordre, comme une caractéristique indépendante des inconnues du dimensionnement [1, 2]. Le transistor peut ainsi être modélisé soit par des mesures expérimentales de la courbe g_m/I_D , soit par des courbes issues de modèles analytiques continus comme le modèle EKV [3] ou le modèle ACM [4].

3. Puissance minimale pour un temps d'établissement total (linéaire plus rampe) donné.

La vitesse est bien connue pour être l'un des facteurs qui conditionne le plus la puissance consommée dans des systèmes électroniques. La vitesse requise est liée au temps d'établissement total dans de nombreuses applications des amplificateurs opérationnels (ampli ops.), telles que les convertisseurs A/D et D/A et les

filtres échantillonnés.

Deux périodes différentes déterminent le temps total d'établissement: la période de rampe ou "slewing" et la période d'établissement linéaire. La première est déterminée par le taux de rampe ou "slew-rate" de l'amplificateur résultant de la charge d'une capacité à courant constant, tandis que la deuxième est déterminée par les performances « petits signaux » en termes de la fréquence de transition et marge de phase. Alors que la modélisation du comportement d'établissement dans les ampli ops a retenu l'attention depuis longtemps ([5, 6, 7, 8, 9]), il n'existe presque aucun antécédent qui traite le sujet de manière systématique du point de vue de la synthèse.

Un certain temps total d'établissement (qui est la quantité significative au niveau système) peut correspondre à différentes partitions entre l'établissement linéaire et le slew-rate. Cette partition est habituellement faite d'une façon plus ou moins arbitraire. Par exemple, pour les circuits à capacités commutées, [10] attribue 10% de la période totale au slew rate, tandis que [11] y assigne un quart. Pourtant, la distribution du temps total entre établissement linéaire et slew-rate a une influence déterminante sur la consommation, comme nous le montrerons ici. Nous allons, donc, aborder les questions suivantes:

Quelle est la meilleure combinaison, du point de vue de la consommation, entre slew-rate et produit gain-bande passante afin d'avoir un temps total d'établissement donné pour la réponse à un échelon d'une certaine amplitude ?

D'autre part, du point de vue de la méthodologie g_m/I_D , nous avons recherché en quoi la réponse à la question précédente influence la synthèse et plus particulièrement, comment cela se répercute sur le choix des relations g_m/I_D ou les niveaux d'inversion des transistors actifs de l'ampli.

Avant de considérer les idées de base du modèle utilisé pour le temps d'établissement, il faut apporter une précision quant au phénomène du slew-rate. L'origine du slew-rate est le chargement d'une capacité avec un courant limité. Ceci peut se produire soit en un nœud interne de l'amplificateur (par exemple, la sortie du premier étage dans un ampli Miller), soit au nœud de sortie, notamment dans le cas des amplis classe A, quand la combinaison de capacité externe et de limitation du courant de sortie est plus restrictive. Dans le premier cas, nous parlerons de slew-rate interne et de slew-rate

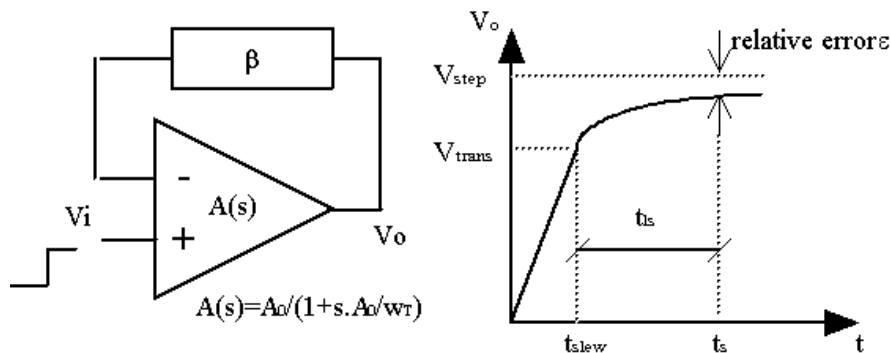


Fig. 1. Schéma pour la détermination du modèle orienté à la synthèse du temps d'établissement total.

externe dans le deuxième cas.

Le développement de notre modèle se base sur une modélisation de premier ordre de la réponse en fréquence de l'ampli, comme illustré à la Fig. 1. L'amplificateur est rétro-alimenté avec un facteur réel β . Les résultats de ce modèle par rapport au temps d'établissement linéaire sont corrigés a posteriori en ajoutant un facteur qui tient compte du fait qu'un système de deuxième ordre, plus représentatif d'un ampli. réel, nécessite un nombre plus grand de constantes de temps pour arriver à une certaine erreur d'établissement.

La transition entre le fonctionnement contrôlé par le slew-rate et celui en linéaire est calculée comme l'instant où la pente dVo/dt à la sortie de l'ampli., calculée en partant du modèle linéaire, devient égale au slew-rate.

Cette démarche amène à l'expression suivante pour le temps d'établissement total:

$$t_s = \tau \left(k_{corrsetl} \ln\left(\frac{1}{\varepsilon}\right) - 1 + \ln\left(\frac{\tau.SR}{V_{step}}\right) + \frac{V_{step}}{\tau.SR} \right)$$

où τ est la constante de temps du système en boucle fermée, égale à $1/(\beta.w_T)$, avec w_T la fréquence de transition de l'ampli. calculé en tenant compte de l'effet des pôles non dominants; ε est l'erreur relative considérée pour définir le temps d'établissement; $k_{corrsetl}$ est le rapport entre le nombre de constantes de temps nécessaires pour l'établissement dans un système de deuxième ordre et celui de premier ordre; V_{step} est l'amplitude de l'échelon à la sortie de l'ampli et SR le slew-rate, égal au minimum entre les slew-rate interne et externe.

Même si ce modèle correspond à une vision simplifiée, il propose un compromis entre précision et complexité approprié pour servir de guide à la démarche de synthèse. De plus, il présente deux avantages clés:

- * il est indépendant de l'architecture particulière de l'ampli.
- * il permet de traiter de façon unifiée les slew-rate interne et externe.

Ces deux caractéristiques sont absentes dans les modèles plus complexes orientés vers l'analyse mentionnés ci-dessus. Par ailleurs, la complexité additionnelle de ces modèles n'entraîne pas nécessairement un gain important en précision, puisque des phénomènes qui modifient le slew-rate comme le comportement non linéaire de l'étage de sortie ou la capacité parasite de la source de courant de la paire différentielle d'entrée ne sont pas pris en compte.

La précision atteinte par notre modèle est illustrée à la Table 1, où le temps d'établissement à 5% prédit par le modèle est comparé avec le temps mesuré pour deux OTAs avec sortie en classe AB. Pour chaque cas la valeur du facteur de correction $k_{corrsetl}$ correspondant a été dérivée de la valeur de la marge de phase PM, et vaut 1.3 pour l'ampli. en technologie silicium massif ("bulk") et 1.05 pour l'ampli. en technologie silicium sur isolant complètement déplété (FD SOI). Ce modèle a été appliqué à l'exploration de l'espace de conception d'un ampli. Miller avec compensation RC en technologie FD SOI 2 μ m [12]. Les variables d'exploration choisies sont

les relations g_m/I_D des étages d'entrée ($(g_m/I_D)_1$) et de sortie ($(g_m/I_D)_2$). Les courbes de niveau à consommation constante, dans le plan $((g_m/I_D)_2, (g_m/I_D)_1)$, sont illustrées à la Fig. 2, pour une synthèse visant un temps d'établissement total au 1% de 1 μ s avec rétro-action unitaire, un échelon de 0.25V, 67° de marge de phase et 10pF de charge.

Comme on peut le voir sur cette figure, on trouve un minimum de consommation de 7 μ A dans la région d'inversion modérée avec $(g_m/I_D)_1$ égal à 16 et $(g_m/I_D)_2$ égal à 23 V^{-1} . Le temps d'établissement simulé pour le circuit optimal s'écarte de moins de 20% de l'objectif de 1 μ s. Cette différence peut être ajustée en partant de l'optimum trouvé. L'optimum correspond à un partage du temps d'établissement de 80% pour la partie linéaire et 20% pour le slew-rate et une fréquence de transition de 0.75MHz. Par contre, imposer un critère, souvent appliqué, de partage du temps total en parties égales entre linéaire et slew-rate amène à une consommation de 15.6 μ A, soit plus du double du résultat optimal.

Techno, μ m	f_T (MHz)	PM (°)	SR (V/ μ s)	t_s mesuré (ns)	t_s calculé (ns)
Bulk, 0.8	9.0	57	3.2	157	169
FD SOI, 2	5.6	64	1.6	311	305

Table 1. Comparaison des résultats du modèle du temps d'établissement total avec mesures.

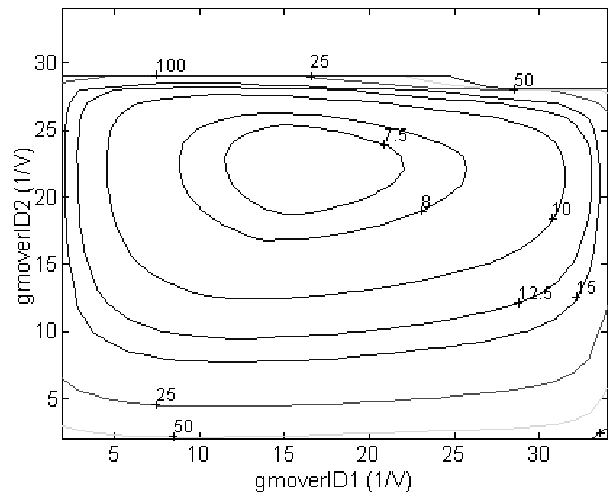


Fig. 2 Courbes de niveau pour consommation constant (présentée en μ A) dans le plan $((g_m/I_D)_2, (g_m/I_D)_1)$

Les Figs. 3, 4 et 5 montrent, dans le cas du même ampli, l'évolution du point de consommation optimale avec le temps d'établissement visé pour un échelon de 0.5V.

Un nouveau résultat intéressant ressort de la Fig. 5. Le point optimal, sauf pour de petites valeurs du temps d'établissement, vérifie que les slew-rate interne et externe sont égaux. Ceci peut être compris comme suit: dans la mesure où le slew-rate total est le minimum entre l'interne et l'externe, pour un slew-rate requis, avoir l'un plus grand que l'autre provoque seulement une augmentation de la consommation sans augmenter la vitesse. Pour des temps totaux réduits, il arrive que la

partie d'établissement contrôlé par le slew-rate soit très faible et donc, qu'il soit plus "profitable" du point de vue de la consommation d'augmenter le f_T , même si le slew-rate interne devient plus grand que l'externe. Des résultats similaires ont été observés pour la variation en fonction de l'amplitude de l'échelon, et aussi pour différentes valeurs de la capacité de charge.

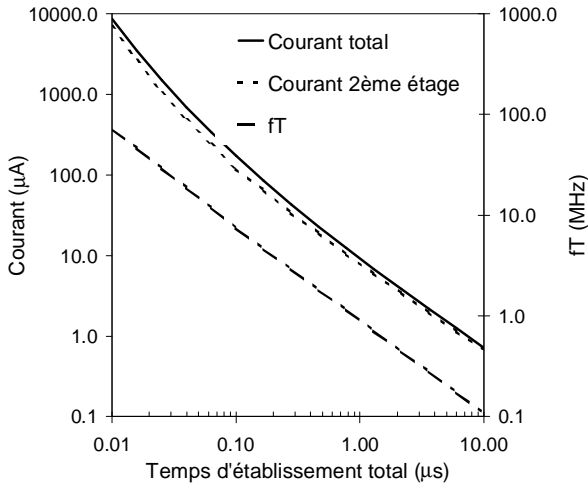


Fig. 3 Consommation minimale et fréquence de transition optimale en fonction du temps d'établissement total pour un échelon de 0.5V d'amplitude.

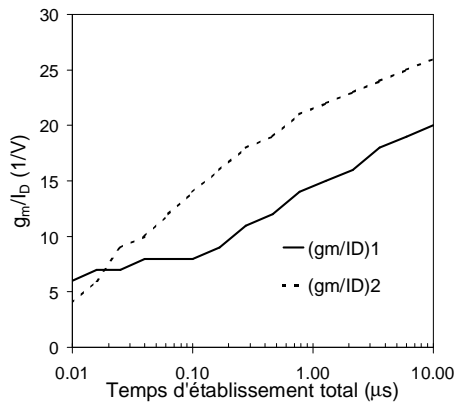


Fig. 4 $(g_m/I_D)_2$, $(g_m/I_D)_1$ pour consommation minimale en fonction du temps d'établissement total.

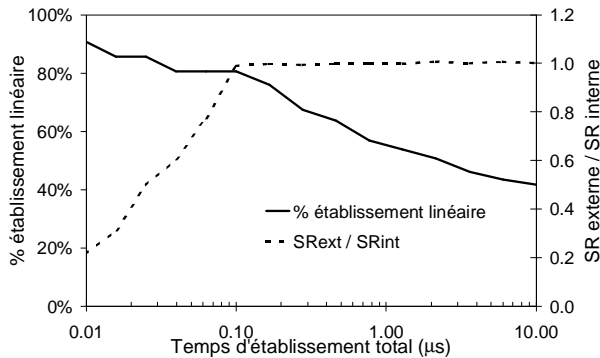


Fig. 5 Pourcentage du temps en établissement linéaire et relation entre slew-rate externe et interne qui donne le minimum de consommation en fonction du temps d'établissement total.

4. Amplificateur "rail-to-rail" programmable.

La faible dépendance de la tension grille-source et de la tension de saturation avec le courant quand le transistor MOS opère en inversion faible ou modérée permet d'envisager la programmation de la bande passante et du slew-rate d'un ampli. op. en utilisant le courant de polarisation, sans encourir de pertes importantes quant aux plages de mode commun d'entrée et à la dynamique de sortie.

Si nous ajoutons l'utilisation d'une architecture avec entrée rail-to-rail et sortie en classe AB, qui rendent l'ampli. approprié pour plusieurs conditions de tension d'entrée et charge de sortie, nous arrivons à une solution qui peut être reconfigurée pour une large gamme d'applications.

Ces deux caractéristiques ont été combinées grâce à l'application de deux techniques clés. D'une part, un principe de contrôle de g_m constant [13] pour l'étage d'entrée valable pour tous les régimes d'inversion. D'autre part, un étage de sortie en classe AB approprié pour la basse tension [14], qui est basée sur un effet de multiplication de la transconductance et permet de polariser les transistors en régime d'inversion faible.

La technique de g_m constant proposé par [13] est illustrée à la Fig. 6. Elle se base sur le contrôle du courant de polarisation d'une des paires différentielles d'entrée (la paire nMOS ici) au moyen d'une boucle d'asservissement qui maintient la transconductance totale égale à une transconductance de référence, réplique de la transconductance de la paire pMOS. À la Fig. 6, $T_{P,ref}$ et T_N sont des répliques du transconducteur pMOS d'entrée et T_P est une réplique du transconducteur nMOS d'entrée. Ces trois transconducteurs ont pour entrée une tension V petite pour assurer un fonctionnement linéaire, avec un niveau de mode commun dans la plage d'opération de mode commun des deux types de transconducteurs, nMOS et pMOS. $T_{P,ref}$ est polarisé avec le courant nominal de la paire d'entrée pMOS, tandis que T_P se polarise avec une copie du vrai courant de la paire d'entrée pMOS, lequel est affecté par le niveau de mode commun à l'entrée de l'ampli. Le courant de polarisation de T_N est contrôlé par la boucle de façon telle que la somme des courants de sortie de T_P et T_N soit égale au courant de sortie de $T_{P,ref}$ et donc, l'entrée étant la même pour les trois transconducteurs, la somme des transconductances de T_P et T_N sera égale à la transconductance de $T_{P,ref}$ laquelle est indépendante de la tension de mode commun d'entrée. La paire d'entrée nMOS est polarisée avec une copie du courant de polarisation de T_N , assurant de cette façon que la transconductance totale d'entrée soit égale à la transconductance de référence pour tous les niveaux de mode commun d'entrée.

Le schéma de l'ampli, sans le circuit de contrôle de g_m constant, est donné à la Fig. 7. Les sorties des paires différentielles d'entrée sont sommées par un étage cascode replié ("folded cascode"), qui transfère le signal d'entrée à l'étage de sortie en classe AB. L'étage cascode replié, qui présente une basse impédance à la sortie du

premier étage, et les caractéristiques de l'étage de sortie en classe AB, qui présente des chemins de basse impédance jusqu'à la sortie, permettent d'avoir un ampli. avec trois étages de gain mais avec une simple compensation par une seule capacité Miller. De plus, comme présenté par [14], l'étage de sortie a une transconductance totale équivalente égale à la transconductance du transistor M_a multipliée par $(1+k.m/h)$, avec k, m et h les gains des miroirs de courant, illustrés à la Fig. 7. Cette augmentation de la transconductance de l'étage de sortie, qui s'élève à une multiplication pour 35 pour cet ampli., permet d'avoir les fréquence de transition et marge de phase désirées avec une capacité de compensation et niveaux d'inversion inférieurs. Ceci réduit la consommation et favorise la programmation par le courant de polarisation.

Dans le schéma de la Fig. 7, VIBN est le signal de contrôle, provenant du circuit de g_m constant, employé pour régler la polarisation de la paire nMOS.

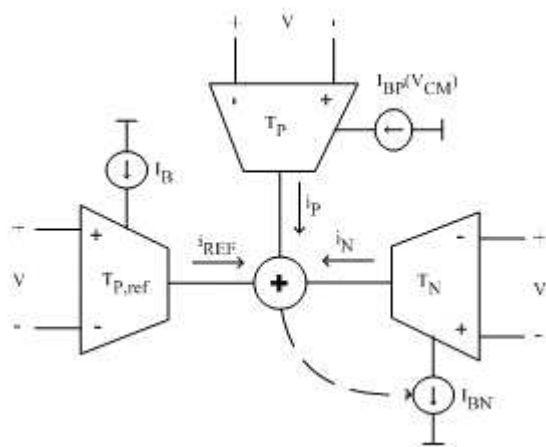


Fig. 6. Principe de fonctionnement du contrôle de g_m constante valable pour tous les régimes d'inversion proposé par [13] et appliqué ici à l'entrée rail-to-rail de l'ampli. programmable.

Cette architecture a été synthétisée suivant la méthodologie présentée à la section 3. La Fig. 8 présente des résultats de simulation où est illustrée la programmation de la bande passante en gardant la stabilité avec une capacité de charge de 50pF et pour différents niveaux de mode commun. La fréquence de transition peut être programmée entre 7.5kHz et 1.5MHz, de manière indépendante du niveau de mode commun d'entrée, avec une consommation totale de repos comprise entre 42nA et 27.5μA.

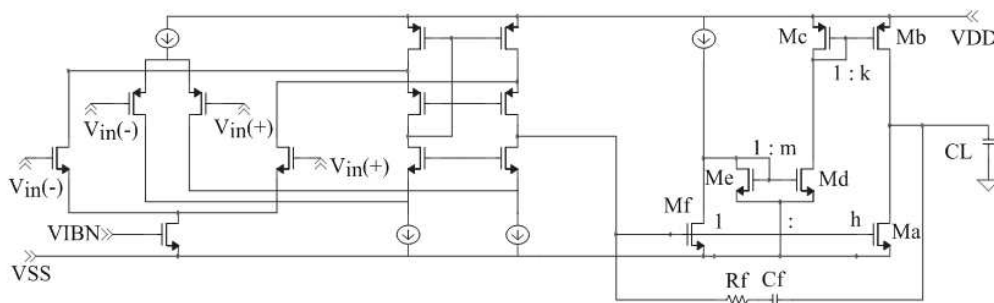


Fig. 7. Schéma de l'ampli. programmable. (Circuit de contrôle de g_m constant non représenté).

5. Conclusions.

Nous avons montré que l'application de la méthodologie g_m/I_D rend possible la conception de circuits optimaux et flexibles. D'une part, à partir d'un modèle simple du temps total d'établissement, approprié pour la synthèse, l'espace de conception d'un ampli. Miller avec compensation RC a été exploré et nous avons montré l'existence d'un minimum de consommation pour un temps total d'établissement donné, qui amène à une réduction considérable de la consommation par rapport à l'utilisation des règles traditionnelles pour le partage entre établissement linéaire et établissement en rampe. D'autre part, on a présenté la conception d'un ampli. op programmable pour des fréquences de transition entre 7.5kHz et 1.5MHz, en utilisant la démarche antérieure.

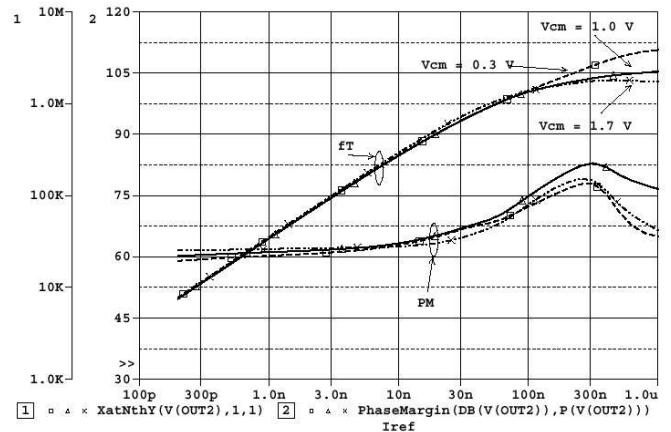


Fig. 8 Programmation de la fréquence de transition (f_T) au moyen du courant de polarisation en gardant la marge de phase pour différents niveaux de mode commun d'entrée.

Références

1. F. Silveira, D. Flandre, P. Jespers, "A gm/ID Based Methodology for the Design of CMOS Analog Circuits and its Application to the Synthesis of a Silicon-on-Insulator Micropower OTA", *IEEE JSSC*, Sept. 1996.
2. D. Flandre, "Méthodologie gm/ID: un chaînon entre l'analyse symbolique et la synthèse de circuits analogiques basse puissance", TAISA 2002.
3. C. Enz, F. Krummenacher and E. Vittoz, "An Analytical MOS Transistor Model Valid in All Regions of Operation and Dedicated to Low-Voltage and Low-Current Applications", *AICSP*, No. 8, 1995.
4. A.Cunha; M.Schneider, C.Galup-Monoro, "An MOS transistor model for analog circuit design", *IEEE JSSC*, Oct. 1998.
5. B. Kamath, R. Meyer and P. Gray, "Relationship between frequency response and settling time of operational amplifiers", *IEEE JSSC*, Dec.1974.
6. C. Chuang, "Analysis of the settling behavior of an operational amplifier," *IEEE JSSC*, Feb. 1982.
7. C. Turchetti, G. Masetti, "A Macromodel for Integrated All-MOS Operational Amplifiers", *IEEE JSSC*, Aug. 1983.
8. I. Lin and J. Nevin, "A modified time-domain model for nonlinear analysis of an operational amplifier," *IEEE JSSC*, June 1986.
9. F. Wang, R. Harjani, "An Improved Model for the Slewing Behavior of Opamps", *IEEE T. CAS II*, Oct. 1995.
10. R. Gregorian, G. Temes, *Analog MOS Integrated Circuits for Signal Processing*, Johnv Wiley & Sons, New York, 1986, p. 497.
11. R. Taylor, "Switched Capacitor Filters" in *Analog Circuit Design*, Kluwer Academic Publishers, Dordrecht, 1994.
12. F. Silveira, "Low-power analog IC design and optimization in Bulk and SOI CMOS technologies in view of application to pacemakers", Thèse de doctorat, UCL, 2002.
13. J. Duque-Carrillo et al., "Robust and universal constant gm circuit technique," *Electronics Letters*, no. 9, April 2002.
14. F.Silveira, D.Flandre, "Analysis and design of a family of low-power class AB operational amplifiers" Proc. XIII SBCCI, Sep. 2000, IEEE Computer Press.