Etude et conception analogique d'architectures d'acquisition acoustique très faible consommation pour applications mobiles

Anthony Baltolu

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

SPÉCIALITÉ ÉLECTRONIQUE

Par Anthony BALTOLU

ETUDE ET DESIGN ANALOGIQUE D'ARCHITECTURES D'ACQUISITION ACOUSTIQUE TRES FAIBLE CONSOMMATION POUR APPLICATIONS MOBILES

Sous la direction de : Jean-Baptiste BEGUERET
(co-directeur : Dominique DALLET)

Soutenue le 14/12/2018

Membres du jury :

M. BEGUERET Jean-Baptiste	Professeur de l'université de Bordeaux	Directeur
M. DALLET Dominique	Professeur de Bordeaux INP	Co-directeur
M. MALOBERTI Franco	Professeur de l'université de Pavia	Rapporteur
M. BARTHELEMY Hervé	Professeur de l'université Sud Toulon Var	Rapporteur
M. CHALET Frédéric	Ingénieur chez NXP Semiconductors	Examinateur
M. TARIS Thierry	Professeur de Bordeaux INP	Président du jury

Titre : Etude et conception analogique d'architectures d'acquisition acoustique très faible consommation pour applications mobiles

Résumé : Les récentes avancées technologiques des microphones de type microsystème électromécanique (MEMS) leurs permettent une utilisation sur une large gamme d'amplitudes sonores. Leur niveau de bruit ayant baissé, il devient possible de capter des sons provenant d'une distance plus lointaine, tandis que l'augmentation de leur pression acoustique maximale leur permet de ne pas saturer dans un environnement très bruyant de type concert ou évènement sportif. Ainsi le système électronique de conversion analogique-numérique connecté au microphone devient l'élément limitant les performances du système d'acquisition acoustique. Un besoin de nouvelles architectures de conversion analogique-numérique ayant une plage dynamique augmentée se fait donc ressentir. Par ailleurs, ces microphones étant de plus en plus utilisés dans des systèmes fonctionnant sur batterie, la contrainte de limitation de la consommation devient importante.

Dans la bande de fréquences audio, les convertisseurs analogiques-numériques de type sigma-delta sont les plus aptes à obtenir une grande résolution combinée à une faible consommation. Ils sont divisés en deux grandes familles: ceux à temps discret utilisant principalement des circuits à capacités commutées, et ceux à temps continu utilisant des circuits classiques. Cette thèse se concentre sur l'étude et la conception de chacun des deux types de convertisseurs sigma delta, en insistant sur la faible consommation, le faible coût de production (surface occupée) et la robustesse du circuit, cela en vue d'une production de masse pour équipements portables.

La conception d'un convertisseur analogique numérique de type sigma-delta à temps discret a été réalisé, ce dernier atteignant un rapport signal sur bruit de 100 décibels sur une bande de 24kHz, pour une puissance consommée de seulement 480µW. Pour limiter la consommation, de nouveaux amplificateurs à base d'inverseurs sont utilisés, et dont la robustesse contre les variations du procédé de fabrication ou de la température a été améliorée. Les spécifications ont été définies grâce au développement d'un modèle de haut-niveau précis, ce qui permet d'éviter le surdimensionnement tout en atteignant les performances voulues. Enfin, un grand ratio de suréchantillonnage a été choisi afin de réduire l'espace utilisé par les capacités commutées, minimisant le coût de fabrication.

Après une étude théorique de l'équivalence entre les modulateurs sigma-delta à temps discret et à temps continu, ainsi que des spécificités propres aux modulateurs à temps continu, une réalisation de ces derniers a été effectuée. Celui-ci atteint un rapport signal sur bruit de 95 décibels sur une bande de fréquence de 24kHz, tout en consommant 142µW. Pour réduire la consommation ainsi que l'espace utilisé, un filtre de boucle du second-ordre a été réalisé avec un seul amplificateur, et le quantificateur fait aussi office d'intégrateur grâce à l'utilisation d'une structure d'oscillateurs contrôlés en tension. Ce quantificateur à base d'oscillateurs est réalisé par des cellules numériques, réduisant la consommation et l'espace utilisé, mais est hautement non-linéaire. Cette non-linéarité a été prise en compte par des choix architecturaux afin de ne pas réduire les performances finales du modulateur.

Mots-clés : modulateur, sigma-delta, convertisseur analogique-numérique, audio, faible consommation, capacités commutées, temps continu, temps discret, quantificateur-intégrateur

Title: Study and analog design of low-power acoustic acquisition systems for mobile applications

Abstract: The recent technological advances in microelectromechanical system (MEMS) microphones allow them to be used on a large sound amplitude range. Due to their lower noise level, it becomes possible to capture sound from a faraway distance, while their increased acoustic overload point gives them the ability to capture sound without saturation in a loud environment like a concert or a sport event. Thus, the electronic analog / digital conversion system connected to the microphone becomes the limiting element of the acoustic acquisition system performance. There is then a need for a new analog / digital conversion architecture which has an increased dynamic range. Furthermore, since more and more of these microphones are used in battery-powered devices, the power consumption limitation constraint becomes of high importance.

In the audio frequency band, the sigma-delta analog / digital converters are the ones most able to provide a high dynamic range combined to a limited power consumption. They are split in two families: the discrete-time ones using switched-capacitors circuits and the continuous-time ones using more classical structures. This thesis concentrates on the study and the design of both of these two types of sigma-delta converters, with an emphasis on the low-power consumption, the low production cost (area occupied) and the circuit robustness, in sight of a mass production for portable devices.

A discrete-time sigma-delta modulator design has been made, the latter reaching a signal to noise ratio of 100dB on a 24kHz frequency bandwidth, for a power consumption of only 480µW. To limit the power consumption, new inverter-based amplifiers are used, with an improved robustness against the variations of the fabrication process or the temperature. Amplifier specifications are obtained thanks to an accurate high-level model developed, which allows to avoid over-design while ensuring that the wanted performances are reached. Finally, a large oversampling ratio has been used to reduce the switched-capacitors area, lowering the modulator cost.

After a theoretical study of the equivalence between discrete-time and continuous-time modulators, and of continuous-time modulators specificities, a design of the latter has been made too. It reaches a signal to noise ratio of 95dB on a 24kHz bandwidth, while consuming 142µW. To reduce the power consumption and the occupied area, a second-order loop filter is implemented using a single amplifier, and the quantizer uses a VCO-based structure that provides inherently an integrating stage. The VCO-based quantizer is made using digital cells, lowering the consumption and area, but is highly non-linear. This non-linearity has been handled by architectural choices to not influence the final modulator performances.

Keywords: sigma-delta, modulator, analog digital converter, audio, low-power, switched-capacitors, discrete-time, continuous-time, VCO-based, quantizer, low-area

Laboratoire de l'Intégration du Matériau au Système – UMR 5218 IMS

351 Cours de la libération, 33405 Talence cedex, France

Résumé français

I – Brève introduction aux convertisseurs sigma-delta

Les convertisseurs analogique-numérique de type sigma-delta combinent deux concepts pour atteindre une haute résolution. Le premier est le suréchantillonnage : en échantillonnant M fois plus vite que la fréquence d'échantillonnage nécessaire donnée par la théorie de Shannon-Nyquist, l'erreur de quantification est distribuée sur une plus grande plage spectrale, diminuant alors sa puissance dans la largeur de bande du signal F_B. Le second concept est la mise en forme du bruit de quantification : grâce au suréchantillonnage, le système travaille sur une plage spectrale beaucoup plus grande que nécessaire, il devient alors possible de « donner une forme » au bruit de quantification, qui consiste à diminuer sa puissance dans la bande du signal F_B et l'augmenter hors bande, la puissance totale restant inchangée. Ces deux concepts combinés, illustrés par la figure I-1, sont assez puissants pour obtenir une grande résolution même en utilisant un quantificateur simple bit.

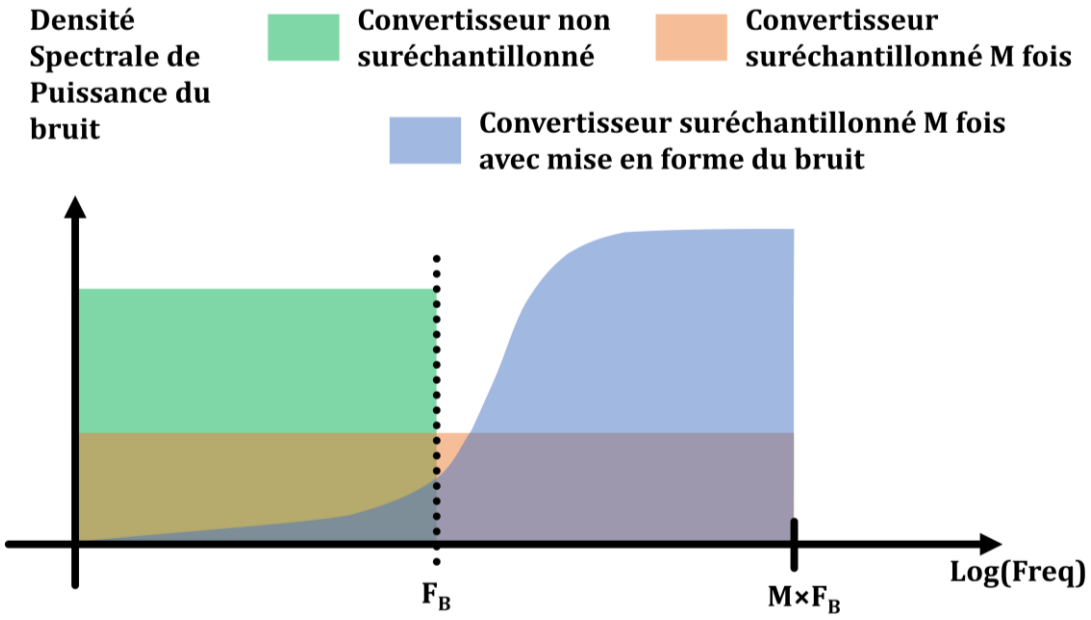


Figure I-1 Effets du suréchantillonnage et du concept de mise en forme sur le bruit de quantification restant dans la bande du signal F_B

Le schéma bloc d'un modulateur sigma-delta à temps discret est donné en figure I-2, avec son modèle linéaire. C'est un système hybride analogique-numérique rebouclé, composé d'un échantillonneur en entrée à une fréquence F_E, d'un filtre qui permet la mise en forme du bruit de quantification, d'un quantificateur et d'un convertisseur numérique-analogique pour former la boucle de retour. Le quantificateur étant un élément non-linéaire il est modélisé par l'addition d'une erreur de quantification. La sortie du modulateur $Y(z)$ est fonction du signal d'entrée $U(z)$, de l'erreur de quantification $E(z)$ ainsi que des fonctions de transfert du filtre pour le signal d'entrée et le signal issu de la boucle de retour, respectivement $H_0(z)$ et $H_1(z)$. La relation est donnée dans l'équation suivante :

$$Y(z) = \frac{H_0(z)}{1 + H_1(z)}\, U(z) + \frac{1}{1 + H_1(z)}\, E(z) \qquad \text{(I-1)}$$

Deux fonctions de transfert remarquables peuvent être dérivées : la fonction de transfert entre le signal d'entrée et la sortie, nommée $STF(z)$, et la fonction de transfert entre le bruit de quantification et la sortie, nommée $NTF(z)$:

$$STF(z) = \frac{H_0(z)}{1 + H_1(z)}, \qquad NTF(z) = \frac{1}{1 + H_1(z)} \qquad \text{(I-2)}$$

Pour mettre en forme correctement le bruit de quantification comme montré en figure I-1, la fonction de transfert $NTF(z)$ doit être proche de zéro pour les fréquences basses, ce qui requiert un fort gain à ces fréquences pour le filtre $H_1(z)$. Concernant la fonction de transfert $STF(z)$, elle ne doit pas changer le signal à basse fréquences et ainsi avoir un gain unitaire, ce qui impose un fort gain à la fois à $H_0(z)$ et à $H_1(z)$. Ce besoin de fort gain à basse fréquence explique pourquoi les filtres des modulateurs font appel à des intégrateurs.

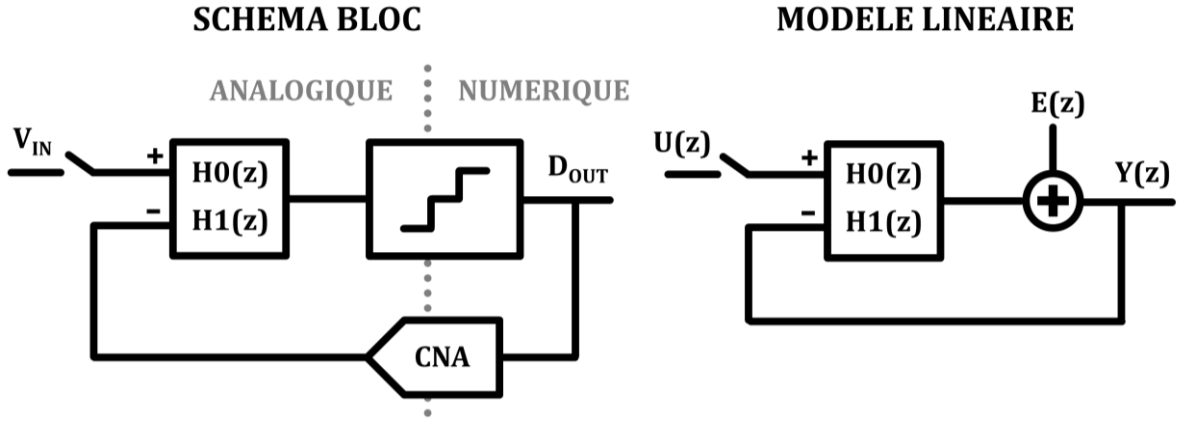


Figure I-2 Schéma bloc et modèle linéaire d'un modulateur sigma-delta

Le nombre d'intégrateurs mis en cascade détermine l'ordre du filtre de mise en forme, un ordre élevé rejetant plus fortement le bruit de quantification hors de la bande de fréquences du signal utile. L'ordre du filtre est ainsi le deuxième paramètre définissant la résolution du modulateur avec le ratio de suréchantillonnage. Cependant un ordre de filtre supérieur à 2 peut entrainer une instabilité du modulateur car une phase de -180° peut être atteinte alors que le filtre fournit toujours du gain dans le système rebouclé, saturant l'entrée du quantificateur. C'est pourquoi les modulateurs d'un ordre supérieur à 2 utilisent une NTF dont le gain à hautes fréquences est limité, de manière à ne jamais saturer l'entrée du quantificateur, perdant un petit peu de performance dans leur propriété de mise en forme du bruit quantification. Le troisième et dernier paramètre influant sur la résolution obtenue par le modulateur est le nombre de bits utilisés dans le quantificateur, qui augmente la résolution de 6dB par bit ajouté, comme dans le cadre de convertisseurs non suréchantillonnés. Le graphique I-3 montre le rapport signal sur bruit de quantification (SQNR) maximal atteignable par un modulateur en fonction de son ratio de suréchantillonnage (OSR) pour différents ordres du filtre de boucle et différents nombres de bits dans le quantificateur. Ainsi l'on peut voir qu'augmenter l'ordre du modulateur permet d'atteindre de meilleures résolutions, seulement à fort OSR pour les modulateurs d'ordre élevé à cause des problèmes de stabilité qui limitent l'efficacité de la mise en forme du bruit. Augmenter le nombre de bits du quantificateur aide à atteindre une forte résolution à faible

OSR, grâce à leur bruit de quantification intrinsèquement inférieur, ce qui relâche aussi le problème de stabilité des modulateurs d'ordre élevé.

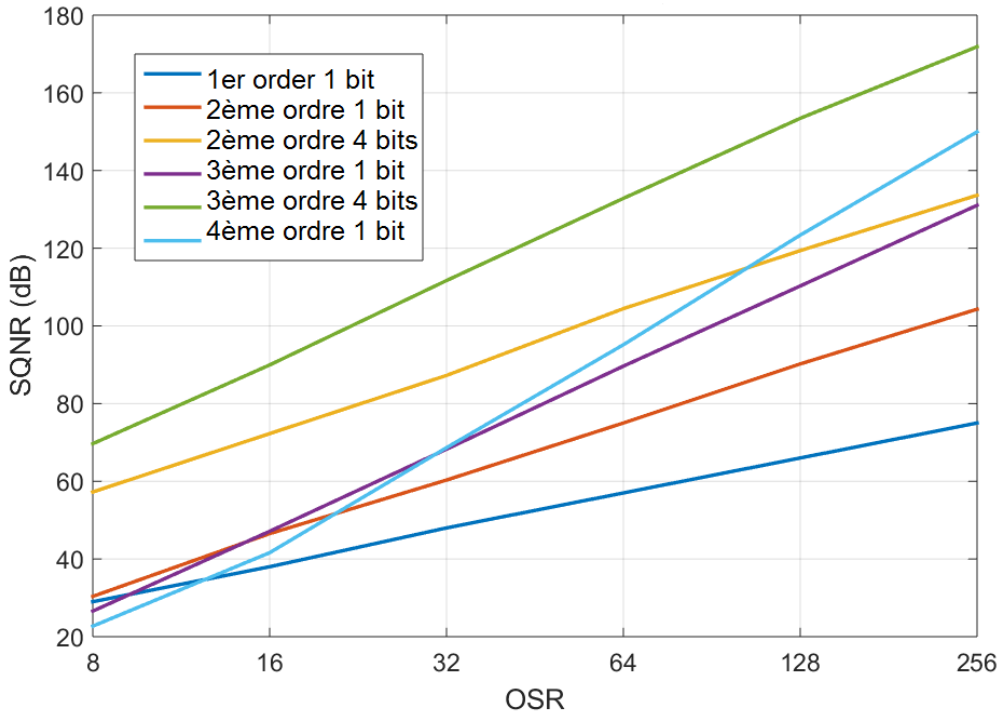


Figure I-3 Rapport signal sur bruit de quantification maximal atteignable pour un modulateur en fonction de son ratio d'échantillonnage, son ordre et du nombre de bits de son quantificateur

Le schéma d'un intégrateur à temps discret réalisé à l'aide de capacités commutées est donné en figure I-4 durant sa phase d'intégration, où A_0 est son gain DC, C_S la capacité d'échantillonnage, C_I la capacité d'intégration, C_O la capacité de charge et C_P la capacité parasite sur le nœud V_A. En considérant un intégrateur parfait (gain A_0 infini, capacités parasites et de charge C_P et C_O nulles), sa fonction de transfert est la suivante :

$$H(z) = \frac{C_S}{C_I}\frac{z^{-1}}{1-z^{-1}} \tag{I-3}$$

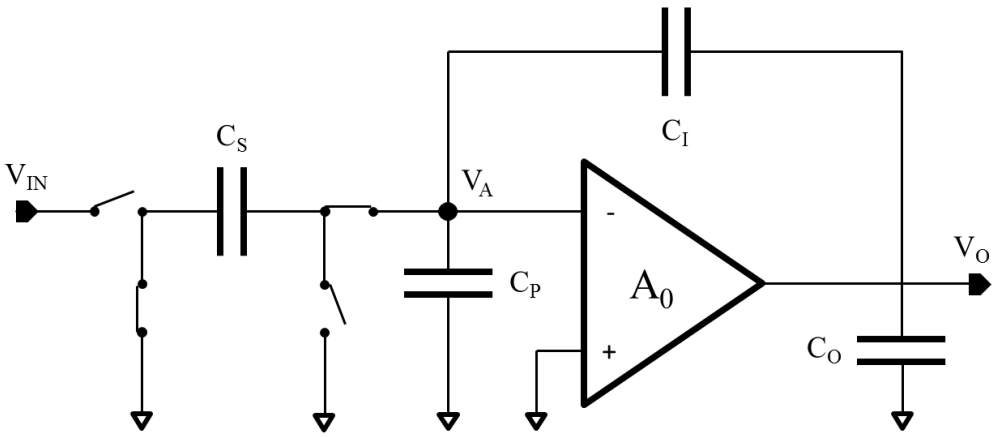


Figure I-4 Schéma d'un intégrateur à temps discret réalisé à l'aide de capacités commutées

Un intégrateur analogique n'étant jamais idéal, plusieurs effets doivent être pris en compte durant sa phase d'intégration pour s'assurer que ceux-ci ne viennent pas nuire aux performances du modulateur, et ainsi définir des spécifications pour les intégrateurs. Un modèle de reproduction précise de la réponse d'un intégrateur pendant sa phase d'intégration a été élaboré, permettant de s'assurer du bon fonctionnement du modulateur sans sur-dimensionner l'amplificateur, afin de réduire sa consommation au maximum. La figure I-5 montre la réponse d'un intégrateur analogique durant la phase d'intégration.

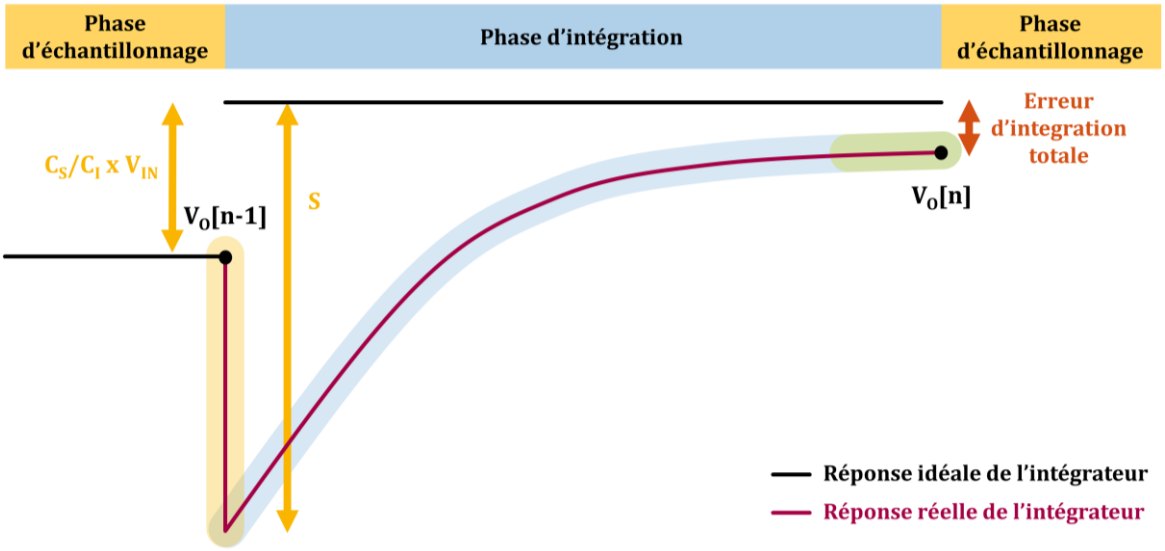


Figure I-5 Réponse de la tension de sortie de l'intégration durant la phase d'intégration

Au démarrage de celle-ci on connecte la capacité d'échantillonnage au reste de l'intégrateur et, avant même que l'amplificateur ne réagisse, une redistribution des charges dans toutes les capacités de l'intégrateur se produit. Cet effet perturbe le nœud V_O qui part, dans un premier temps, dans le sens opposé à son résultat final, augmentant la charge à délivrer par l'amplificateur durant la période d'intégration comme le montre la partie jaune du nœud V_O sur la figure I-5. La marche d'intégration n'est plus la tension échantillonnée multipliée par le rapport des capacités d'échantillonnage et d'intégration, mais une valeur notée S définie par l'équation suivante :

$$S = \frac{C_S}{C_I} \times V_{IN} \times \left(1 + \frac{C_I}{C_S + C_P + C_O \times \left(1 + \frac{C_S + C_P}{C_I}\right)} \right) \quad \text{(I-4)}$$

Ensuite l'amplificateur réagit et commence à délivrer du courant pour ramener le nœud V_O vers sa valeur finale (partie bleu de V_O sur la figure I-5), et sa réponse peut être de trois types : linéaire, dans ce cas le produit gain bande définit la réponse et l'erreur d'intégration à un instant T ; partiellement linéaire, dans ce cas l'amplificateur est limité par son courant de polarisation et fournit un courant constant jusqu'à ce qu'un comportement linéaire soit possible ; et enfin le cas où l'amplificateur a un courant de polarisation trop faible et est limité par celui-ci durant toute la période d'intégration. L'erreur d'intégration finale dépend alors de la marche d'intégration S, du produit gain bande de l'amplificateur (noté GBW) ainsi que de son slew rate (noté SR). Ce dernier est la valeur maximale de variation du nœud V_O par unité de temps, qui dépend du courant de polarisation de l'amplificateur. Selon les trois types de réponse, l'erreur

en fin de phase d'intégration ε est définie par les équations ci-dessous, où T_{SLEW} représente la durée pendant laquelle la réponse de l'amplificateur est limitée par son courant de polarisation. Un T_{SLEW} négatif ou nul équivaut à une réponse linéaire de l'amplificateur et l'erreur dépend d'une exponentielle négative, pouvant devenir négligeable avec un grand produit gain bande. Un T_{SLEW} inférieur à la phase d'intégration (équivalent à la moitié d'une période d'horloge T) produira une erreur d'intégration dépendante d'une exponentielle négative, pouvant être minimisée par un produit gain bande élevé. Enfin, si le T_{SLEW} dépasse une demie période d'horloge, l'erreur est linéaire et impacte grandement les performances du modulateur.

$$\varepsilon = \begin{cases} -S \times e^{-2\pi \times GBW \times \frac{T}{2}}, & T_{SLEW} \leq 0 \\ -(S - SR \times T_{SLEW}) \times e^{-2\pi GBW \times \left(\frac{T}{2} - T_{SLEW}\right)}, & 0 < T_{SLEW} < \dfrac{T}{2} \\ -\left(S - SR \times \dfrac{T}{2}\right), & T_{SLEW} \geq \dfrac{T}{2} \end{cases} \quad \text{(I-5)}$$

$$T_{SLEW} = \frac{S}{SR} - \frac{1}{2\pi \times GBW} \quad \text{(I-6)}$$

Enfin, après le produit gain bande et le slew rate, une dernière spécification peut entrainer une erreur d'intégration : le gain DC de l'amplificateur. Ce dernier détermine la valeur finale vers laquelle la réponse de l'amplificateur tend si le produit gain bande et le slew rate sont assez élevés pour entrainer une erreur négligeable. La loi de conservation des charges entre les phases d'échantillonnage et d'intégration permet de définir la fonction de transfert de l'intégrateur en tenant compte du gain fini :

$$H_{REELLE}(z) = \frac{C_S}{C_I} \times \frac{\dfrac{A_0}{1 + A_0 + (C_S + C_P)/C_I}\, z^{-1}}{1 - \dfrac{1 + A_0 + C_P/C_I}{1 + A_0 + (C_S + C_P)/C_I}\, z^{-1}} \quad \text{(I-7)}$$

Une erreur de magnitude est introduite, qui peut être compensée en ajustant le ratio des capacités, mais aussi une erreur de phase traduisant le fait qu'une partie de la charge intégrée est perdue à chaque cycle, plus difficile à compenser et obligeant le gain DC à avoir une valeur élevée.

Ce modèle permet ainsi de définir des spécifications précises pour les amplificateurs qui composent les intégrateurs. Ces dernières ne dépendent que de la valeur de composants passifs (valeurs de capacités) et de spécifications du modulateur (fréquence d'horloge par exemple). Ainsi les performances du modulateur peuvent être assurées sans sur-dimensionner les intégrateurs, bénéficiant à la consommation. La figure I-6 illustre la précision du modèle où la réponse de l'intégrateur est comparée à celle d'un circuit électronique ayant les mêmes spécifications et valeurs de capacités. La seule différence entre les deux réponses est l'effet de redistribution des charges qui est plus lissée sur un circuit électronique dû à la résistance non-nulle des commutateurs.

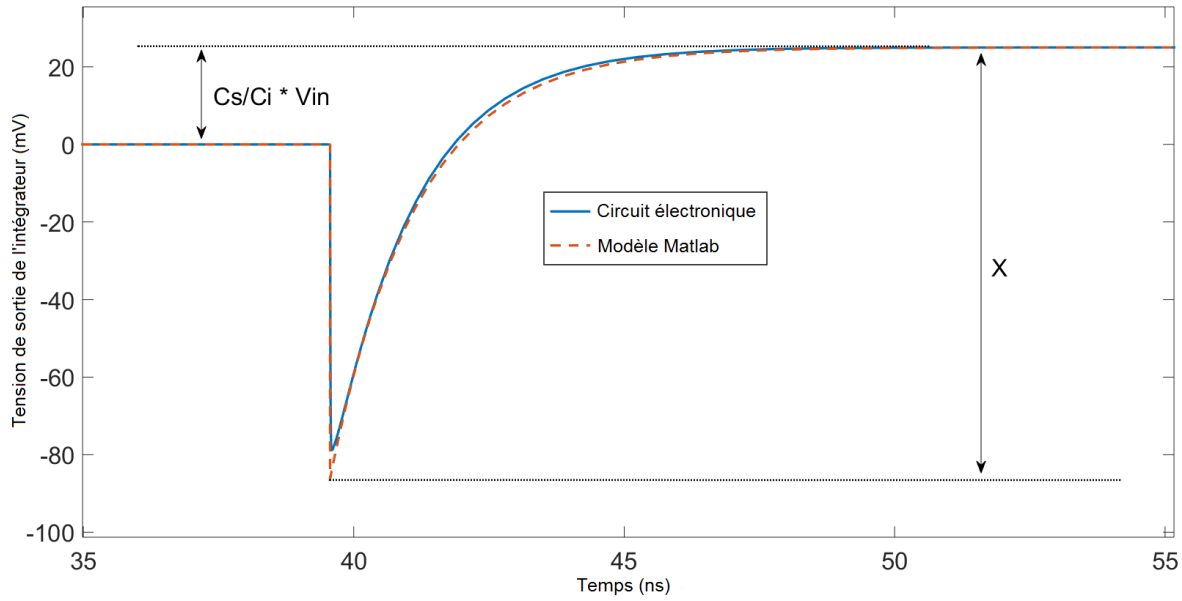


Figure I-6 Comparaison de la réponse du modèle et d'un circuit électronique

II – Design d'un modulateur à temps discret atteignant 103dB de plage dynamique

Dans le cadre de notre contexte lié à l'acquisition audio, le modulateur doit atteindre un rapport signal sur bruit de 100dB avec une consommation la plus faible possible tout comme une faible occupation d'espace. Les spécifications utiles pour définir l'architecture du modulateur sont données en table II-1.

Table II-1 Spécifications du modulateur utilisées pour définir son architecture

Spécification	Valeur	Unité
Tension de référence	1.5	V
Largeur de bande du signal (F_B)	24	kHz
Consommation (P)	480	µW
Niveau de bruit ramené à l'entrée	7.5	µV
Plage dynamique (D)	103	dB
Rapport signal sur bruit	100	dB
Rapport signal sur bruit et distorsion	>95	dB
Rapport signal sur bruit de quantification	110	dB
$FOM_S = D + 10\log_{10}(F_B/P)$	180	dB

L'architecture choisie est un modulateur d'ordre 2, utilisant un quantificateur à trois niveaux (1,5 bits) et un ratio de suréchantillonnage de 256, donnant un rapport signal sur bruit de quantification de 110dB. Un quantificateur de 1,5 bits est utilisé pour atteindre la résolution voulue sans augmenter l'ordre du modulateur, et cela pour plusieurs raisons : il n'y a pas de problème de stabilité particulier à prendre en compte, et il permet de réduire l'amplitude du

signal de sortie des intégrateurs, comme le fait un quantificateur multi-bits. Contrairement à ce dernier, un quantificateur sur trois niveaux simplifie largement l'algorithme de gestion des éléments du CNA de retour, lui permettant de conserver sa linéarité avec un faible surplus de consommation. Un grand ratio de suréchantillonnage est utilisé, afin de réduire l'espace occupé par les capacités qui représentent généralement la majorité de l'espace d'un modulateur à temps discret. Le modèle linéaire du modulateur est donné en figure II-1 :

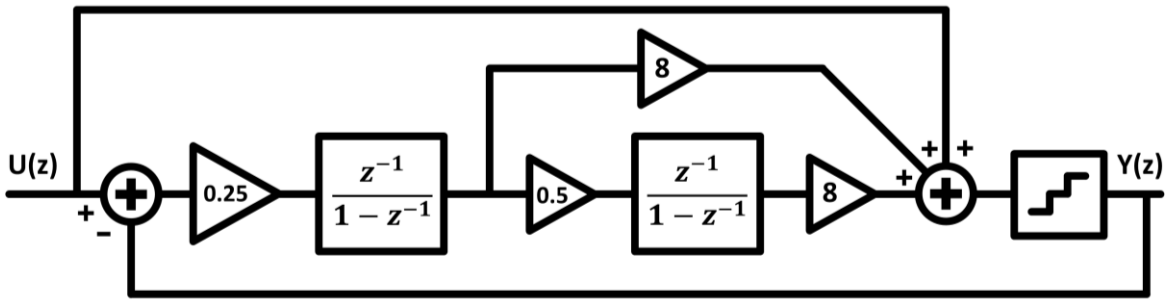


Figure II-1 Modèle linéaire du modulateur à temps discret développé

La contrainte de bruit permet de définir la valeur des capacités d'échantillonnage de chaque étage d'intégration, celles-ci sont de 2,3pF pour le premier intégrateur et de 150fF pour le second. Les spécifications dynamiques de chaque intégrateur sont ensuite définies grâce au modèle présenté en section I.

Les amplificateurs de classe AB sont particulièrement adaptés pour les circuits à capacités commutées de faible consommation. Ils présentent un courant de polarisation dynamique en fonction du signal d'entrée, ainsi ils procurent un fort slew rate en début de phase d'intégration, puis retournent à un courant de polarisation réduit le reste du temps. Parmi les amplificateurs de classe AB, ceux à base d'inverseurs ont la plus faible consommation grâce à leur unique branche de courant, les rendant particulièrement attractifs pour l'application souhaitée. Cependant, les amplificateurs inverseurs sont sensibles aux variations de tension d'alimentation, de procédé de fabrication ou de température, ce qui rend le point de polarisation variable sur une large plage et leur consommation mal maitrisée. Pour remédier à ce problème tout en conservant les propriétés dynamiques (classe AB) des amplificateurs inverseurs, le circuit présenté en figure II-2 est utilisé. Un miroir de courant fournit un courant de polarisation en utilisant des transistors dans leur régime linéaire (faible tension drain-source) afin de conserver le fonctionnement en class AB. Cela s'explique de la façon suivante : lorsqu'une assez large tension est appliquée en entrée de l'amplificateur, comme c'est le cas lors du début de la phase d'intégration des circuits à capacités commutées, les inverseurs sont dans des états opposés. Les transistors MN2 et MP3 sont passants et linéaires tandis que les transistors MN3 et MP2 sont bloqués. Ainsi la tension du nœud à la source du transistor MP3 chute tandis que celle du nœud à la source du transistor MN2 monte. Cela augmente les tensions drain-source des transistors de polarisation linéaires MN5 et MP5 : ceux-ci fonctionnent comme une résistance, une plus grande tension correspondant à un plus fort courant. A la fin de la période d'intégration, les entrées retrouvent leur position d'origine autour de la moitié de la tension d'alimentation, et le système retrouve alors le courant défini par l'étage de polarisation. Une modification présentée en figure II-3 permet d'amplifier le fonctionnement en classe AB de cet amplificateur en mettant l'étage de polarisation lui aussi dynamique. Pour cela, l'idée est de détecter les variations des nœuds A et B qui se produisent en début de phase d'intégration, pour

augmenter le courant de polarisation jusqu'au retour à la normale. Cette fonction est expliquée côté NMOS mais le raisonnement est le même côté PMOS. Un transistor MN2 est connecté en diode ce qui lui permet sur sa tension de grille de monitorer les variations de tension de sa source. Ainsi lorsque sa tension de source augmente, sa tension de grille aussi, augmentant la tension de grille du transistor MN1 de l'étage de polarisation. Comme la même chose se produit du côté PMOS, un plus grand courant passe dans la branche MP1-MN1, augmentant le courant de l'étage de polarisation, ce qui se répercute ensuite sur l'étage d'amplification. En combinant ces deux techniques, l'amplificateur à base d'inverseurs réalisé permet d'avoir un fonctionnement en classe AB tout en ayant un point de polarisation contrôlé, permettant une architecture robuste, à faible consommation et performante.

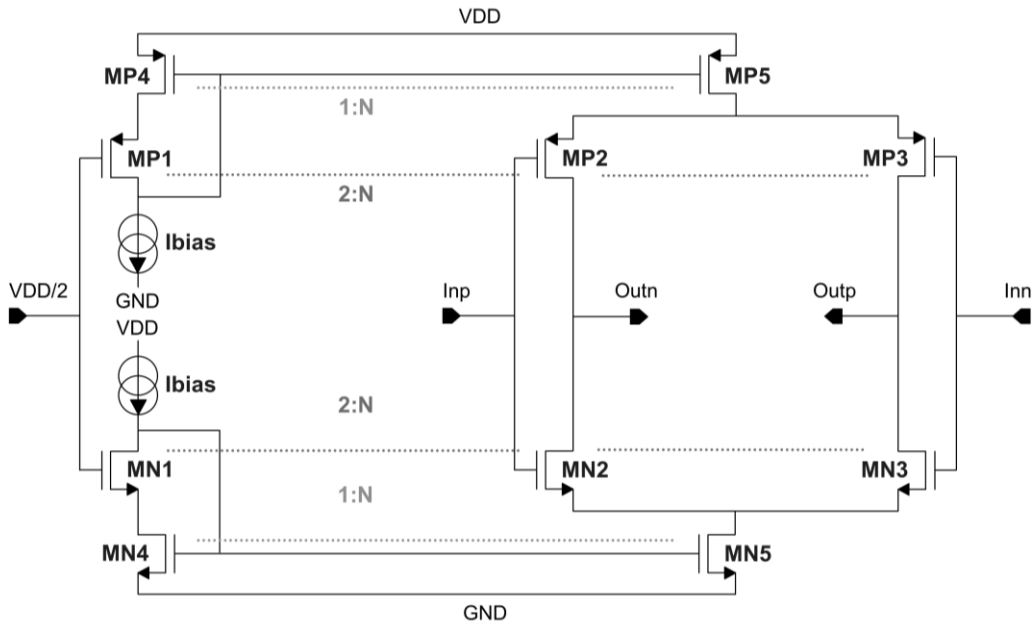


Figure II-2 Amplificateur à base d'inverseurs avec polarisation utilisant des transistors linéaires

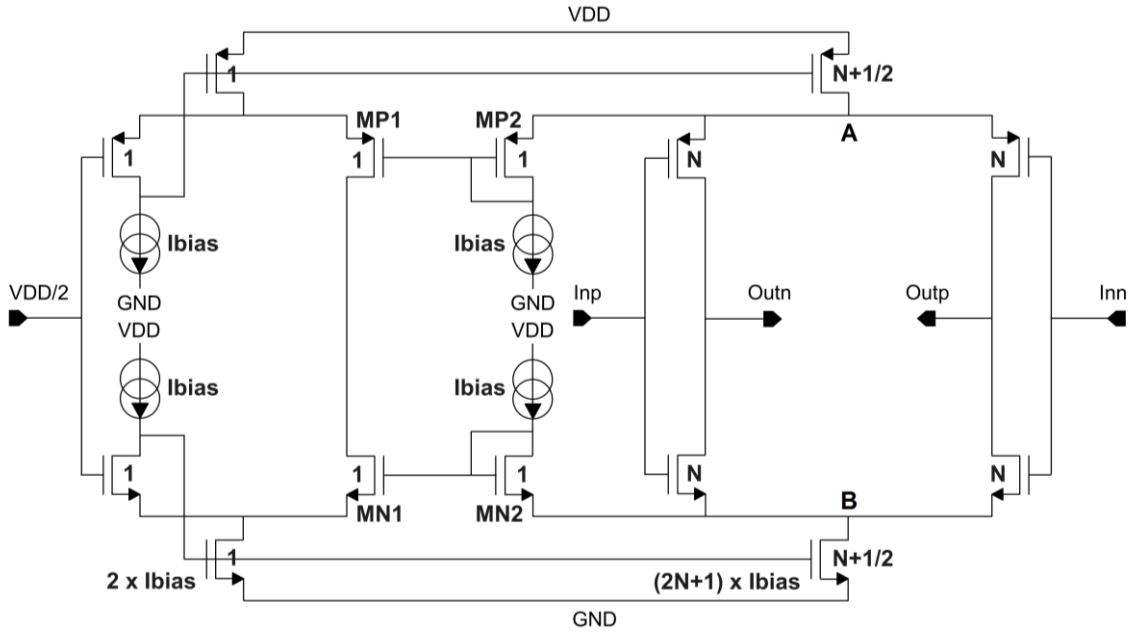


Figure II-3 Amplificateur à base d'inverseurs avec étage de polarisation dynamique

Pour minimiser l'espace occupé par les capacités ainsi que la consommation du modulateur, le CNA présent dans la boucle de retour du modulateur réutilise les capacités d'échantillonnage plutôt que d'en avoir des dédiées. Comme le CNA présente trois niveaux, les capacités d'échantillonnage sont coupées en deux pendant la phase d'intégration : elles sont connectées toutes deux à la tension de référence ou à la masse pour sortir les niveaux « 1 » ou « -1 », tandis que l'une est à la masse et l'autre à la tension de référence pour former le niveau intermédiaire « 0 ». Afin de conserver la linéarité du CNA, les capacités sont alternées à chaque fois que le niveau « 0 » doit être activé et la commande des commutateurs est fournie par un circuit logique simple. L'implémentation du CNA au travers de la capacité d'échantillonnage est présentée en figure II-4 (représentation non différentielle). Les signaux de commande du CNA à trois niveaux sont représentés par un bit de signe (état à « 0 » pour un signe positif, « 1 » pour un signe négatif) et un bit de donnée, correspondant à un code en complément à 2.

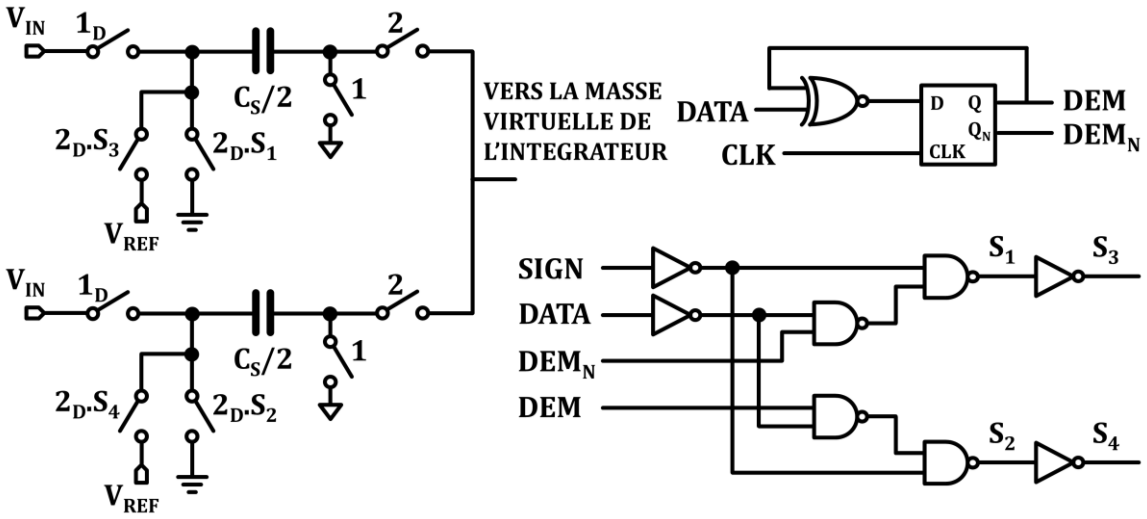


Figure II-4CNA à trois niveaux réutilisant la capacité d'échantillonnage

Le modulateur réalisé atteint 110dB de rapport signal sur bruit de quantification pour un signal d'entrée à -3dBFS sur une bande de fréquence de 24kHz. En tenant compte du bruit thermique, simulé à 7.5µVrms à l'entrée sur la même bande audio, le rapport signal sur bruit est de 100dB. Le spectre du signal de sortie du modulateur est donnée en figure II-5. Trois courbes sont affichées, celle d'un procédé de fabrication typique à 27°C, ainsi que les extremums : procédé fast à 125°C et procédé slow à -40°C, les trois atteignant 110dB de rapport signal sur bruit de quantification (SQNR). La distorsion est un peu plus élevée à 125°C mais toujours dans les spécifications, cela est dû à la résistance des commutateurs qui est non-linéaire en fonction du signal échantillonné. Ce modulateur atteint ces performances pour une consommation totale de 480µW, ce qui lui donne une figure de mérite FOM_S de 180dB, le plaçant ainsi parmi les meilleurs modulateurs audio à temps discret de la littérature. L'utilisation d'amplificateurs à base d'inverseurs a permis d'obtenir une faible consommation, tandis qu'un grand ratio de suréchantillonnage, conjugué à un CNA réutilisant les capacités d'échantillonnage permettent de réduire l'espace occupé (plus faible surface que l'état de l'art) et de produire à faible coût, répondant ainsi à la problématique de recherche.

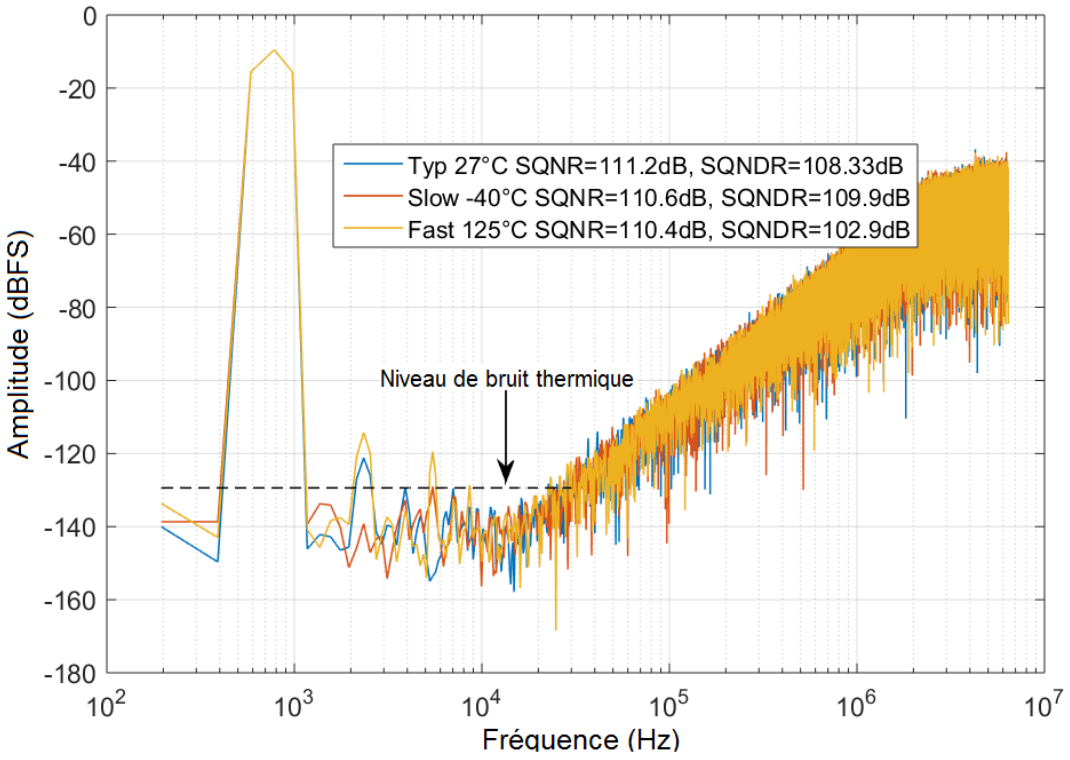


Figure II-5 Spectre de sortie du modulateur pour trois corners différents

Table II-2 Table des performances et comparaison avec l'état de l'art

	Ce Travail	**[16]**	**[17]**	**[20]**	**[21]**	**[22]**	**[33]**	**[34]**
Techno (nm)	140	180	180	180	160	350	130	90
Année	2017	2011	2012	2009	2016	2013	2012	2004
Ordre	2	4	3	3	3	4	4	3
F_B (kHz)	24	25	50	20	20	20	20	20
F_E (MHz)	12.288	1.6	3.2	4	11.29	2.4	2.5	4
OSR	256	32	32	100	282	60	62	100
Référence (V)	1.5	1	1.5	0.7	1.8	1.5	0.5	0.6
DR (dB)	103	93	94	85	109	92.6	85	88
SNR (dB)	100	92.5	-	84	106	-	82.4	85
SNDR (dB)	98	92	88.9	81	103	87.9	81.7	81
Conso (µW)	480	58	140	22	1120	140	35.2	140
Aire (mm^2)	0.084	-	0.49	0.715	0.16	0.207	0.57	0.18
FOMs (dB)	180	179.3	179.5	174.5	181.5	174	173	169.5

III – Théorie des modulateurs à temps continu et leurs spécificités

Cette section présente les connaissances théoriques nécessaires à la synthèse d'un modulateur sigma-delta à temps continu, ainsi que les spécificités de ces derniers par rapport à ceux à temps discret précédemment étudiés. Pour réaliser un modulateur à temps continu, on commence par

définir un modulateur à temps discret dont les performances satisfont les spécifications requises. Ensuite, en utilisant la méthode d'invariance de la réponse impulsionnelle, on synthétise la boucle de retour CNA – filtre de boucle - échantillonneur pour qu'elle ait la même réponse impulsionnelle que la boucle de retour du modulateur discret. Prenons l'exemple d'un modulateur à temps discret du troisième ordre ayant pour fonction de transfert du bruit de quantification $NTF_D(z)$ définie ci-dessous et dont on souhaite synthétiser l'équivalent continu. Travaillons premièrement avec un modulateur normalisé (période d'échantillonnage T_E de 1s).

$$NTF_D(z) = \frac{(1-z^{-1})^3}{1 - 2.2003z^{-1} + 1.6887z^{-2} - 0.4444z^{-3}} \quad \text{(III-1)}$$

La boucle de retour peut être décomposée comme la somme de trois chemins d'ordre d'intégration différent comme montré en figure III-1. Calculons par la méthode d'invariance de la réponse impulsionnelle l'équivalent discret de chacun de ces chemins, en commençant par celui du premier ordre. Le CNA a une réponse impulsionnelle de type échantillonneur-bloqueur idéal donnée par l'expression ci-dessous :

$$h_{CNA}(t) = \begin{cases} 1, & pour \ 0 \le t \le 1 \\ 0, & pour \quad t > 1 \end{cases} \quad \text{(III-2)}$$

Cette réponse est ensuite intégrée pour donner la réponse du chemin du premier ordre :

$$h_1(t) = \begin{cases} \int_0^t 1 \times dt\,, & for\ 0 \le t \le 1 \\ h_1(1) + \int_1^t 0 \times dt\,, & for\ t > 1 \end{cases} = \begin{cases} t, & pour\ 0 \le t \le 1 \\ 1, & pour\ t > 1 \end{cases} \quad \text{(III-3)}$$

On échantillonne ensuite cette dernière toutes les T_E secondes ($h_{1d}(n) = h_1(nT_E)$) :

$$h_{1d}(n) = \begin{cases} n, & for\ n \in \{0,1\} \\ 1, & for\ n > 1 \end{cases} = \begin{cases} 0, & for\ n = 0 \\ 1, & for\ n \ge 1 \end{cases} \quad \text{(III-4)}$$

Et l'on prend la transformée en z pour obtenir la fonction de transfert discrète équivalente :

$$H_1(z) = \sum_{n=0}^{\infty} h_{1d}(n) z^{-n} = \sum_{n=1}^{\infty} z^{-n} = \frac{z^{-1}}{1 - z^{-1}} \quad \text{(III-5)}$$

En intégrant $h_1(t)$ encore une et deux fois, on obtient respectivement $h_2(t)$ et $h_3(t)$ donnant la réponse impulsionnelle des deux autres chemins d'intégration. De manière analogue, en échantillonnant ces réponses et en prenant la transformée en z on obtient les fonctions de transfert discrètes équivalentes suivantes :

$$H_2(z) = \frac{0.5z^{-1} + 0.5z^{-2}}{(1-z^{-1})^2}, \qquad H_3(z) = \frac{\frac{1}{6}z^{-1} + \frac{4}{6}z^{-2} + \frac{1}{6}z^{-3}}{(1-z^{-1})^3} \quad \text{(III-6)}$$

Il ne manque plus qu'à additionner les transformées en z obtenues avec les coefficients de poids pour obtenir la réponse totale du chemin de retour :

$$H(z) = k_1 H_1(z) + k_2 H_2(z) + k_3 H_3(z) \quad \text{(III-7)}$$

Cela donne la fonction de transfert du bruit de quantification $NTF(z)$ suivante :

$$NTF(z) = \frac{(1-z^{-1})^3}{1 + \left(k_1 + \frac{k_2}{2} + \frac{k_3}{6} - 3\right)z^{-1} + \left(3 - 2k_1 + \frac{4k_3}{6}\right)z^{-2} + \left(k_1 - \frac{k_2}{2} + \frac{k_3}{6} - 1\right)z^{-3}} \quad \text{(III-8)}$$

Par identification avec $NTF_D(z)$ donnée en (III-1), on déduit alors la valeur des coefficients k_1, k_2 et k_3 qui sont respectivement 0,6703, 0,2441 et 0,0440. En utilisant ces coefficients, le modulateur à temps continu que l'on voulait synthétiser est équivalent à celui de référence à temps discret.

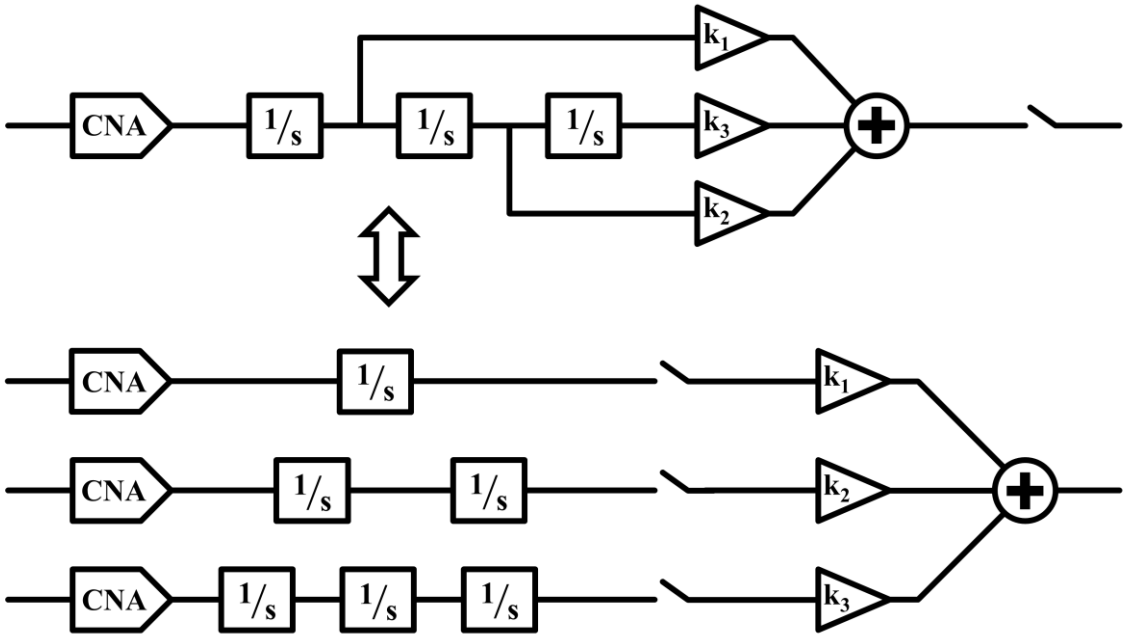


Figure III-1 Décomposition du CNA – filtre de boucle en chemins de différents ordres d'intégration

Les modulateurs à temps continu ont des intégrateurs à constante de temps fixe contrairement à ceux discrets dont la constante de temps est fonction de la fréquence d'horloge. Ainsi un modulateur à temps continu n'est adapté que pour une fréquence d'horloge particulière. Le modulateur normalisé présenté utilise une fréquence d'horloge de 1Hz, en changeant celle-ci par F_E, les intégrateurs auront un signal F_E plus petit à chaque échantillonnage. Pour retrouver le comportement correct des intégrateurs au cours d'une période d'échantillonnage, leur constante de temps doit alors augmenter et être égale à F_E.

Le principal avantage d'un modulateur sigma-delta à temps continu est sa propriété de filtre anti-repliement. En effet, le signal entrant est échantillonné en sortie du filtre de boucle, aussi ce dernier fait alors office de filtre anti-repliement car il est constitué d'intégrateurs qui ont une réponse fréquentielle passe-bas. La figure III-2 montre la réponse en fréquence idéale du filtre de boucle du modulateur de troisième ordre synthétisé auparavant, ainsi que celle du filtre de boucle utilisant des intégrateurs de gain DC fini et égal à 40dB. En considérant que ce modulateur a un taux de suréchantillonnage de 128, on peut voir que la performance minimum du filtre anti-repliement est donnée par les repliements à des fréquences proches de la bande passante du signal, on obtient dans notre cas 88dB de réjection. Pour des repliements plus basse fréquence, la rejection est meilleure car les intégrateurs ont plus de gain, cependant elle devient limitée lorsque l'on atteint les gains DC de chacun des intégrateurs à 112dB.

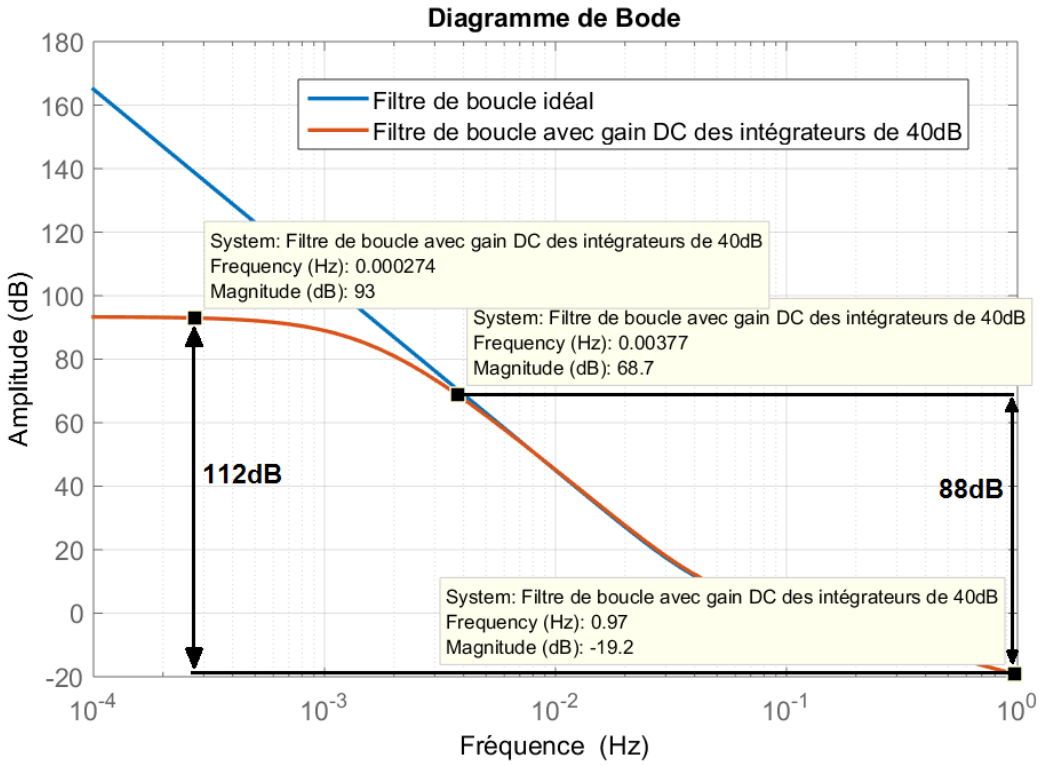


Figure III-2 Réponse en fréquence d'un filtre de boucle idéal et d'un filtre avec gain DC fini, ainsi que l'évaluation des performances anti-repliement du filtre

Les modulateurs à temps continu sont sensibles à l'exactitude de temps. Un délai dans la boucle de retour change sa réponse impulsionnelle : celle-ci ne sera plus équivalente, après transformée en z, à la fonction de transfert discrète que l'on voulait synthétiser et peut rendre le modulateur instable. Cela est illustré en figure III-3, où une compensation possible du délai est le rajout d'un chemin non intégrateur dans la boucle de retour, avec le coefficient k_0 adapté.

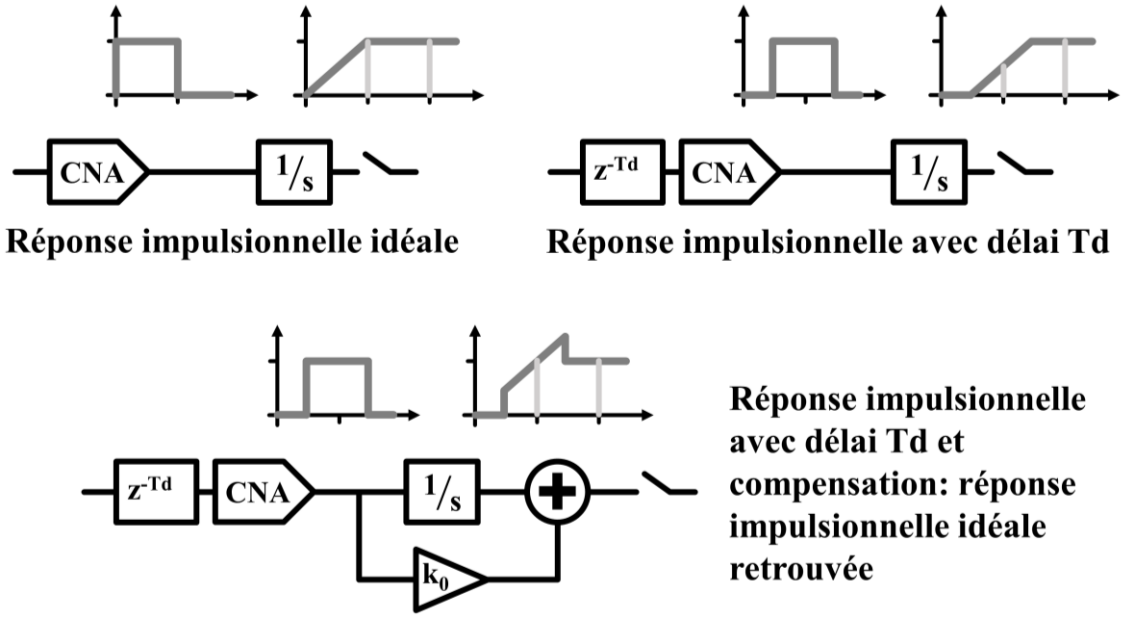


Figure III-3 Modification de la réponse impulsionnelle avec l'ajout d'un délai et compensation du délai pour retrouver une réponse impulsionnelle idéale

Un délai minimum est requis par les composants analogiques entre l'instant où l'on échantillonne le signal en entrée du quantificateur, et l'instant où le quantificateur donne une valeur qui fait changer le CNA de la boucle de retour. En considérant un délai d'une demi-période d'échantillonnage, on peut déterminer la valeur du coefficient k_0 de manière analogue à celle faite pour déterminer k_1, k_2 et k_3 ; en utilisant la méthode d'invariance de la réponse impulsionnelle. On se base sur le diagramme de la boucle de retour suivant :

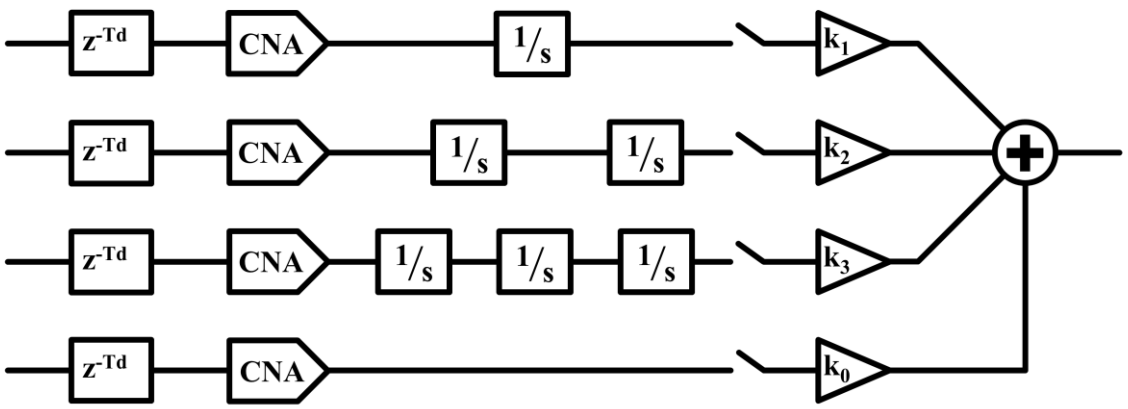


Figure III-4 Décomposition de la boucle de retour avec compensation du délai

Le CNA, retardé d'une demi-période, a la réponse impulsionnelle suivante :

$$h_{CNAD}(t) = \begin{cases} 0, & pour\ 0 \le t \le 0.5 \\ 1, & pour\ 0.5 \le t \le 1.5 \\ 0, & pour\ t > 1.5 \end{cases} \qquad \text{(III-9)}$$

De manière analogue à précédemment et en utilisant la réponse impulsionnelle retardée du CNA, on détermine alors les fonctions de transfert discrètes de chacun des chemins de la boucle de retour :

$$H_{0D}(z) = z^{-1}, \qquad H_{1D}(z) = \frac{0.5z^{-1} + 0.5z^{-2}}{(1 - z^{-1})}$$

$$H_{2D}(z) = \frac{0.125z^{-1} + 0.75z^{-2} + 0.125z^{-3}}{(1 - z^{-1})^2} \qquad \text{(III-10)}$$

$$H_{3D}(z) = \frac{\frac{1}{48}z^{-1} + \frac{23}{48}z^{-2} + \frac{23}{48}z^{-3} + \frac{1}{48}z^{-4}}{(1 - z^{-1})^3}$$

En prenant la somme pondérée, on obtient la fonction de transfert de la boucle de retour complète :

$$H_D(z) = k_0 H_{0D}(z) + k_1 H_{1D}(z) + k_2 H_{2D}(z) + k_3 H_{3D}(z) \qquad \text{(III-11)}$$

Ce qui conduit à la fonction de transfert du bruit de quantification suivante :

$$NTF_D = \frac{(1 - z^{-1})^3}{1 + \left(k_0 + \frac{k_1}{2} + \frac{k_2}{8} + \frac{k_3}{48} - 3\right)z^{-1} + \left(-3k_0 - \frac{k_1}{2} + \frac{5k_2}{8} + \frac{23k_3}{48} + 3\right)z^{-2} + \left(3k_0 - \frac{k_1}{2} - \frac{5k_2}{8} + \frac{23k_3}{48} - 1\right)z^{-3} + \left(-k_0 + \frac{k_1}{2} - \frac{k_2}{8} + \frac{k_3}{48}\right)z^{-4}} \qquad \text{(III-12)}$$

L'ajout d'un délai dans la boucle de retour augmente l'ordre de la NTF qui est maintenant du quatrième ordre. La compensation ajoute un degré de liberté supplémentaire au système lui

permettant alors d'annuler le terme d'ordre 4 de la NTF. Par identification avec (III-1), on obtient alors les valeurs de coefficients suivantes : $k_0 = 0{,}3666$, $k_1 = 0{,}7979$, $k_2 = 0{,}2661$ et $k_3 = 0{,}0440$. On remarque que les valeurs des coefficients k_1 et k_2 sont aussi impactées par l'ajout d'un délai de boucle. D'autres méthodes de compensation sont aussi présentées dans le corps de la thèse.

D'autres effets temporels peuvent affecter la résolution du modulateur : le phénomène de gigue sur l'horloge entraine une incertitude sur le moment où le CNA va changer de valeur. Ainsi deux données consécutives du CNA peuvent ne pas avoir de durée exactement égale, entrainant une variabilité sur la durée de chaque valeur de sortie du CNA. L'erreur introduite est dépendante de plusieurs facteurs comme la gigue de l'horloge, le codage du CNA ou la fréquence de l'horloge F_E, venant ajouter du bruit au signal d'entrée et réduisant la résolution du modulateur. Prenons l'exemple d'un CNA utilisant un codage de non-retour-à-zéro, qui suit une séquence de données $y(n)$: son signal de sortie peut être décomposé comme la somme d'un signal idéal et d'un signal d'erreur comme présenté en figure III-5.

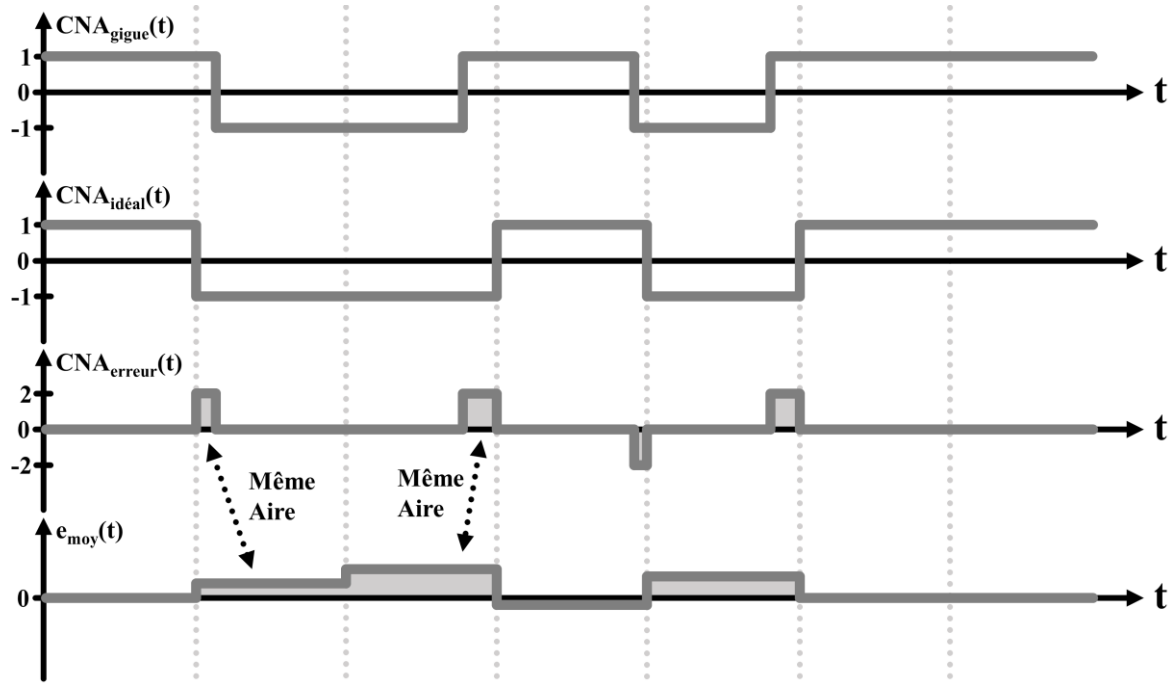


Figure III-5 Effet de gigue sur l'horloge du CNA introduisant une erreur et détermination de l'erreur moyennée sur une période

L'erreur introduite dépend de la différence entre nouvelle et précédente valeurs du CNA ainsi que de la variation de durée entre les deux valeurs, que l'on représentera par une séquence aléatoire $j(n)$. Moyennant cette erreur sur une période, on obtient la relation suivante :

$$e_{moy}(n) = \big(y(n) - y(n-1)\big) \times \frac{j(n)}{T_E} = \delta y(n) \times j(n) \times F_E \qquad \text{(III-13)}$$

La puissance du signal d'erreur injecté à l'entrée du modulateur dû à l'effet de gigue est :

$$\sigma_e^2 = \sigma_{\delta y}^2 \times \sigma_j^2 \times F_E^2 \qquad \text{(III-14)}$$

Ce dernier peut être modélisé comme un signal aléatoire de variance donnée par l'équation (III-14). Réduire la puissance du signal d'erreur pour une certaine variance de gigue donnée

peut être réalisé en diminuant la fréquence d'horloge du modulateur, mais cela revient à réduire sa résolution par la même occasion. Une autre solution consiste à réduire la variance du signal $\delta y(n)$ qui est la différence entre deux données consécutives de $y(n)$. Pour réaliser cela, un quantificateur et un CNA multi-bits doivent être utilisés dans le modulateur. Puisque $\delta y(n)$ est égal soit à zéro (quand deux données consécutives sont égales) soit à la taille d'un bit de poids faible, plus on utilise de niveaux de quantification plus la sensibilité du modulateur à l'effet de gigue diminue. D'autres implémentations du CNA sont plus sensibles à l'effet de gigue comme le codage retour-à-zéro, et d'autres moins sensibles comme un CNA à capacités commutées, c'est pourquoi un modulateur à temps continu est plus sensible qu'un modulateur à temps discret.

IV – Design d'un modulateur à temps continu ayant 16 bits de résolution

L'étude de l'état de l'art des modulateurs audio à temps continu montre que les concepts les moins consommateurs d'énergie combinent plusieurs fonctions dans un seul bloc de base : un double intégrateur n'utilisant qu'un seul amplificateur ou un quantificateur possédant une fonction d'intégrateur du premier ordre. Parmi ces quantificateurs, un type est utilisé dans des modulateurs à temps continu de fréquences plus élevées que l'audio : le quantificateur à base d'oscillateur contrôlé en tension. Ce dernier permet d'obtenir une quantification multi-bits, tout en y ajoutant certains bénéfices : une implémentation basée sur des cellules numériques (faible espace occupé et faible consommation), et opération intrinsèque d'intégration en utilisant la phase de l'oscillateur comme signal de sortie. De plus, grâce à leur implémentation numérique, la quantification peut se faire quasi-instantanément sur un front d'horloge, diminuant le délai de retour de boucle. Ce type de quantificateur permet de simplifier l'architecture du modulateur, mais apporte aussi certains défauts. La grande non-linéarité des oscillateurs contrôlés en tension limite leur application à des résolutions moyennes, c'est pourquoi ils sont principalement utilisés dans des modulateurs à temps continu de large bande de fréquence du signal où la résolution souhaitée est plus faible qu'en audio. Enfin, ces derniers sont sensibles aux variations de procédé de fabrication et de température, ce qui rend leur gain de conversion tension-fréquence variable et peut poser problème dans le cas d'un intégrateur où la constante de temps doit être fixe. Un modulateur à temps continu utilisant un quantificateur à base d'oscillateurs en tension est réalisé, atteignant 16 bits de résolution et dont les défauts du quantificateur ont été pris en compte par des choix architecturaux.

Commençons dans un premier temps par comprendre le fonctionnement d'un quantificateur à base d'oscillateurs contrôlés en tension. Celui-ci est expliqué à l'aide de la figure IV-1 présentant le schéma bloc et le modèle linéaire du quantificateur-intégrateur. Deux oscillateurs contrôlés en tension convertissent une différence de tension en une différence de fréquence à travers un gain K_{VCO}. La différence de fréquence fait augmenter ou diminuer continuellement la différence de phase entre les deux oscillateurs selon son signe. La position des phases de chaque oscillateur est échantillonnée à l'aide de bascules et un décodeur de phase quantifie la différence de phase avec un nombre prédéfini de valeurs de quantification équitablement réparties sur l'intervalle [0, π] ou [-π, π] selon les implémentations. La phase étant l'intégrale

de la fréquence multipliée par un facteur 2π, ce quantificateur réalise intrinsèquement un intégrateur parfait. Une boucle de retour est nécessaire pour s'assurer que la différence de phase reste dans l'intervalle de quantification choisi.

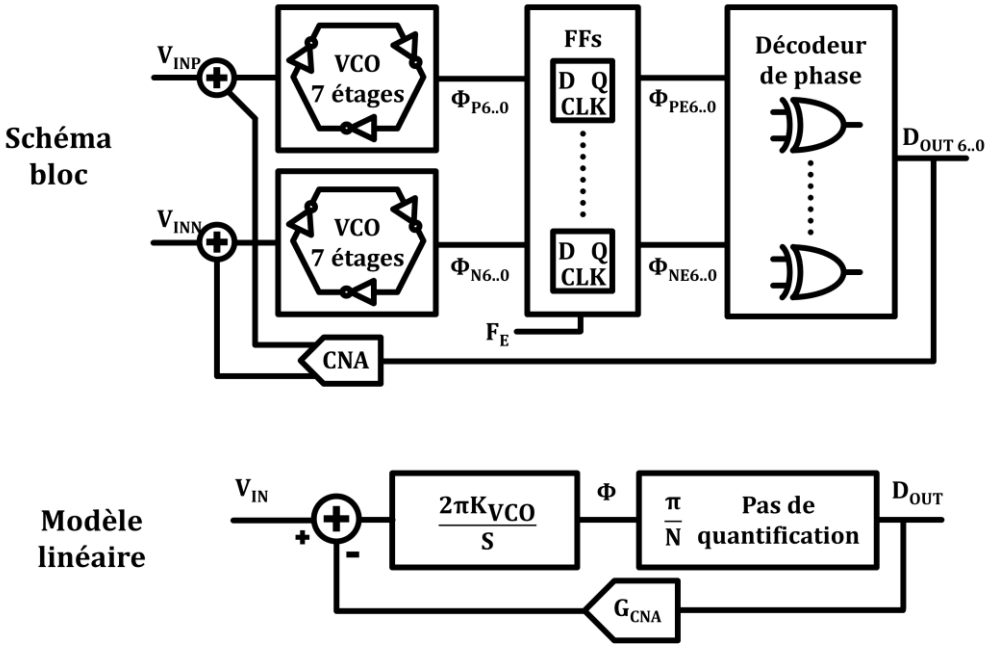


Figure IV-1 Schéma bloc et modèle linéaire d'un quantificateur-intégrateur utilisant des oscillateurs contrôlés en tension

La figure IV-2 explique le fonctionnement du décodeur de phase basé sur des portes logiques XOR en utilisant des oscillateurs à 7 phases. Les inverseurs entourés en rouge sont dans un état de transition : leurs grilles viennent de basculer mais leur sortie pas encore. Quand aucune tension différentielle n'est appliquée à l'entrée des oscillateurs, on se retrouve dans le cas a) où la phase des oscillateurs se propage à la même vitesse dans les deux oscillateurs. Toutes les phases sont égales, et la sortie du décodeur de phase est donc 0. Le cas b) est celui d'une tension différentielle positive : l'oscillateur VCO_P a une fréquence plus élevée que VCO_N, son inverseur en transition est en avance de deux éléments, comparé à VCO_N. Le décodeur de phase voit sa sortie avoir deux bits actifs sur sept, la différence de phase est alors de 2π/7. Si toutes les phases étaient dans un état opposé, la différence vaudrait alors π. Le cas c) est celui où la tension d'entrée est négative : l'oscillateur VCO_N est cette fois-ci plus rapide que VCO_P. Son avance est de 4 éléments, donnant alors une différence de phase de 4π/7 entre les deux oscillateurs. Le décodeur de phase à base de portes XOR n'est pas capable de définir le signe de la différence de phase : il en donne seulement la valeur absolue qu'il quantifie sur un intervalle [0, π] avec un pas de π/N où N représente le nombre de phases utilisées. Sa sortie est un code thermométrique de la différence de phase, donnant alors N+1 niveaux de quantification. Un circuit logique simple peut être ajouté en complément du décodeur de phase pour définir le signe de la différence de phase. Ainsi, l'intervalle de quantification passe de [0, π] à [-π, π] en conservant le même pas de quantification, ce qui permet d'obtenir avec le même système 2N+1 niveaux de quantification. Ce quantificateur, avec son architecture différentielle et un grand nombre de niveaux de quantification qui réduit la tension en entrée des oscillateurs, permet d'obtenir une linéarité de l'ordre de 80dB, et est intégré dans un modulateur à temps continu du troisième ordre.

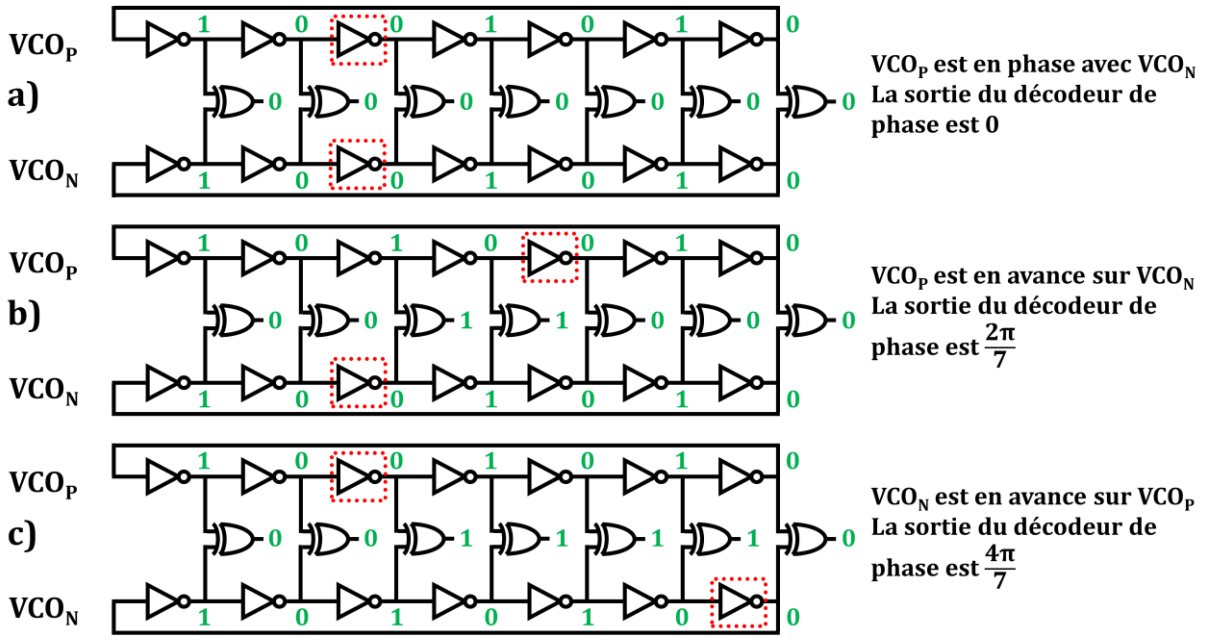


Figure IV-2 Principe de fonctionnement du décodeur de phase XOR a) oscillateurs en phase b) VCO_P en avance c) VCO_N en avance

Les spécifications du modulateur à temps continu sont présentées en table IV-1. Comme le quantificateur-intégrateur à base d'oscillateurs est sensible aux variations de procédé de fabrication et à la température, une marge de 20dB est prise entre le bruit de quantification et le bruit thermique. Ainsi, même si la constante de temps de l'intégrateur varie, la marge est assez grande pour assurer les performances du modulateur, sans avoir besoin de calibrer chaque circuit et de compenser ce dernier en température, ce qui ajouterait du coût au produit. Le modulateur réalisé (présenté en figure IV-3) est du troisième ordre, avec un taux de suréchantillonnage de 128 et un quantificateur de 4 bits donnant un rapport signal sur bruit de quantification de 115dB sur une bande de 24kHz. Une NTF à faible gain hors bande est utilisée pour assurer la stabilité du modulateur même en cas de large variation des constantes de temps des intégrateurs. Le filtre de boucle de second ordre rejette la non-linéarité restante du quantificateur-intégrateur afin d'obtenir une haute résolution du modulateur.

Table IV-1 Spécifications du modulateur utilisées pour définir son architecture

Spécification	Valeur	Unité
Tension de référence	1.5	V
Largeur de bande du signal (F_B)	24	kHz
Consommation (P)	150	µW
Niveau de bruit ramené à l'entrée	13	µV
Plage dynamique (D)	98	dB
Rapport signal sur bruit	95	dB
Rapport signal sur bruit et distorsion	>90	dB
Rapport signal sur bruit de quantification	115	dB
$FOM_S = D + 10\log_{10}(F_B/P)$	180	dB

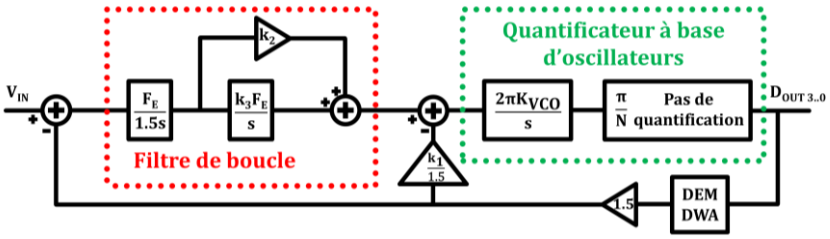


Figure IV-3 Modèle linéaire du modulateur à temps continu réalisé

Le filtre de boucle permet, par sa topologie « feedforward », une réalisation utilisant un seul amplificateur comme montré en figure IV-4. Ce dernier utilise un schéma à deux étages classique compensé par une capacité Miller, de façon à obtenir un grand gain DC pour la linéarité du modulateur et la réjection des repliements et des non-linéarités du quantificateur. La fonction de transfert du montage de la figure IV-4 est la suivante :

$$H_F(s) = \frac{-(1 + R_2 C_2 s)}{R_1(C_1 - C_2)s + R_1 R_2 C_1 C_2 s^2} \tag{IV-1}$$

En prenant $C_1 = C_2 = C$, on obtient :

$$H_F(s) = \frac{-\left(\frac{1}{R_1 R_2 C^2} + \frac{1}{R_1 C} s\right)}{s^2} \tag{IV-2}$$

Cette équation est maintenant de la même forme que celle du filtre de boucle présenté en figure IV-3 qui est :

$$H(s) = \frac{\frac{k_3 F_S^2}{1.5} + \frac{k_2 F_S}{1.5} s}{s^2} \tag{IV-3}$$

Comme la valeur de R_1 est définie par la contrainte de bruit, on peut ensuite définir R_2 et C par identification entre (IV-2) et (IV-3). Le quantificateur-intégrateur utilisé fournit en sortie un code thermométrique de la valeur ainsi qu'un bit de signe. Ces signaux sont particulièrement adaptés pour contrôler un CNA dont les éléments unitaires ont trois niveaux : « -1 », « 0 » et « 1 », réalisés à l'aide d'un duo résistances comme présenté en figure IV-4.

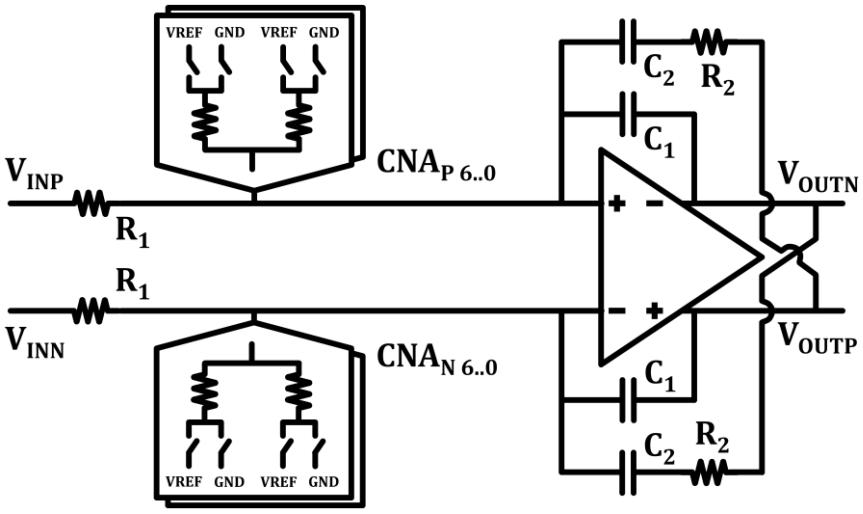


Figure IV-4 Filtre de boucle du second ordre réalisé avec un seul amplificateur

Le modulateur réalisé en technologie 140nm atteint un rapport signal sur bruit de quantification (SQNR) de 116dB en condition typique et de 113dB dans le pire cas sur une bande de 24kHz, comme montré en figure IV-5. Les choix architecturaux faits pour limiter la sensibilité aux variations de process et de température sont donc validés. Concernant la linéarité, l'autre point faible de l'utilisation d'un quantificateur à base d'oscillateurs, le troisième harmonique se situe sous le niveau de bruit thermique simulé, ce qui n'entraine donc pas de baisse de résolution du modulateur, validant là aussi les choix effectués. Ces performances sont atteintes pour une consommation de 142µW et sont résumées en table IV-2 où elles sont aussi comparées avec l'état de l'art.

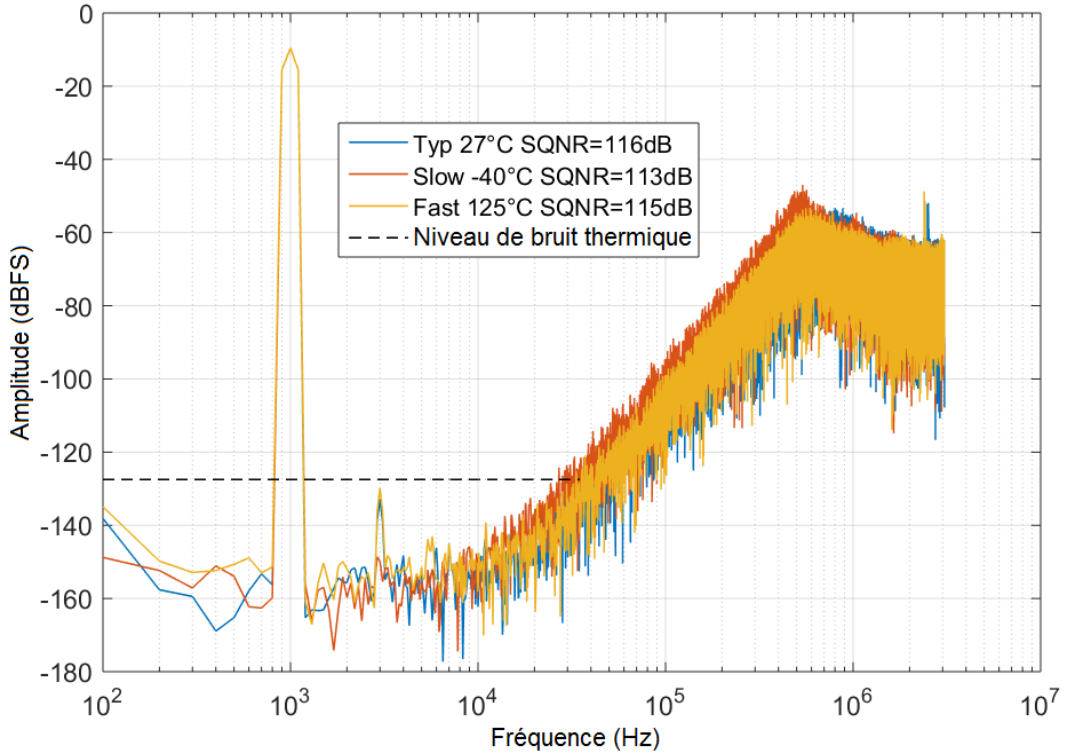


Figure IV-5 Spectre de sortie du modulateur pour trois corners différents

Table IV-2 Table des performances et comparaison avec l'état de l'art

	This work	**[50]**	**[51]**	**[53]**	**[39]**	**[45]**	**[63]**	**[64]**
Techno (nm)	140	180	180	65	160	28	40	65
Année	2018	2014	2016	2016	2016	2015	2011	2015
Ordre	3	3	3	3	3	2	2	3
F_B (kHz)	24	24	24	25	20	24	24	24
F_E (MHz)	6.144	6.144	6.144	6.4	3	24	6.5	3.072
OSR	128	128	128	128	75	500	135	64
Référence (V)	1.5	1.8	1.8	1	1.6	3.3/1	1.2	1.1
DR (dB)	98	103	103.6	103	103.1	100.6	102	88
SNR (dB)	95	98.9	99.3	100.1	-	100.6	-	-
SNDR (dB)	95	98.2	98.5	95.2	91.3	98.5	90	85
Conso (µW)	142	280	280	800	390	1130	500	121
Aire (mm^2)	0.232	1.25	1	0.256	0.21	0.022	0.05	0.6
FOMs (dB)	180.3	182.3	183	177.9	180	173.9	178.8	171

V - Conclusion

Pour répondre à la problématique d'un système d'acquisition audio à grande plage dynamique (d'environ 120dB), faible consommation et faible coût de production, une architecture composée d'un amplificateur à gain variable et d'un convertisseur analogique numérique à plage dynamique plus faible (de 100dB) est sélectionnée. La recherche s'est concentrée sur la conversion analogique numérique de faible consommation et faible espace occupé. Pour atteindre de hautes résolutions, deux convertisseurs sigma-delta (un à temps discret et l'autre à temps continu) ont été développés.

Le convertisseur sigma-delta à temps discret utilise des amplificateurs à base d'inverseurs pour obtenir une faible consommation. Un nouveau schéma d'amplificateur est présenté de manière à corriger les défauts de l'amplificateur inverseur et le rendre robuste aux variations de température et du procédé de fabrication, afin de contrôler sa consommation. Un taux de suréchantillonnage élevé est utilisé permettant de réduire la taille des capacités afin de réduire l'espace occupé. Dans le même esprit, le CNA de la boucle de retour est intégré aux capacités d'échantillonnage, bénéficiant à l'espace occupé et à la consommation de l'amplificateur en réduisant la capacité de charge. Ce modulateur atteint les performances de l'état de l'art, une figure de mérite FOM_S de 180dB, témoin d'une très bonne efficacité énergétique.

Le convertisseur sigma-delta à temps continu utilise un quantificateur-intégrateur à base d'oscillateurs contrôlés en tension. Les oscillateurs transforment la différence de tension en une différence de fréquence, et la phase instantanée (intégrale de la fréquence) est échantillonnée et quantifiée pour former l'opération d'intégration. L'avantage de ce type de quantificateur est l'augmentation de l'ordre du modulateur ainsi que son implémentation utilisant des cellules numériques simples, pour un faible espace occupé et une faible consommation. Un filtre de boucle du second ordre permet de rejeter les non-linéarités des oscillateurs, et est réalisé à l'aide d'un montage à amplificateur unique pour diminuer la consommation. Le modulateur est donc composé de seulement deux sous blocs, ce qui lui permet d'atteindre les performances de l'état de l'art (FOM_S de 180.3dB) tout en utilisant une implémentation pour moitié numérique, ce qui est bénéfique pour le transfert dans des technologies plus fines (réduction de l'espace occupé et de la consommation).

Les figures de mérite de 180dB atteintes par les modulateurs réalisés les placent parmi l'état de l'art dans le comparatif de convertisseurs analogiques/numériques de Boris Murmann [4] comme on peut le voir sur la figure V-1. Ce dernier a été modifié pour ne tenir compte que des modulateurs sigma-delta et la définition de la figure de mérite FOM_S utilise la formule originelle employant la plage dynamique plutôt que le rapport signal sur bruit et distorsion.

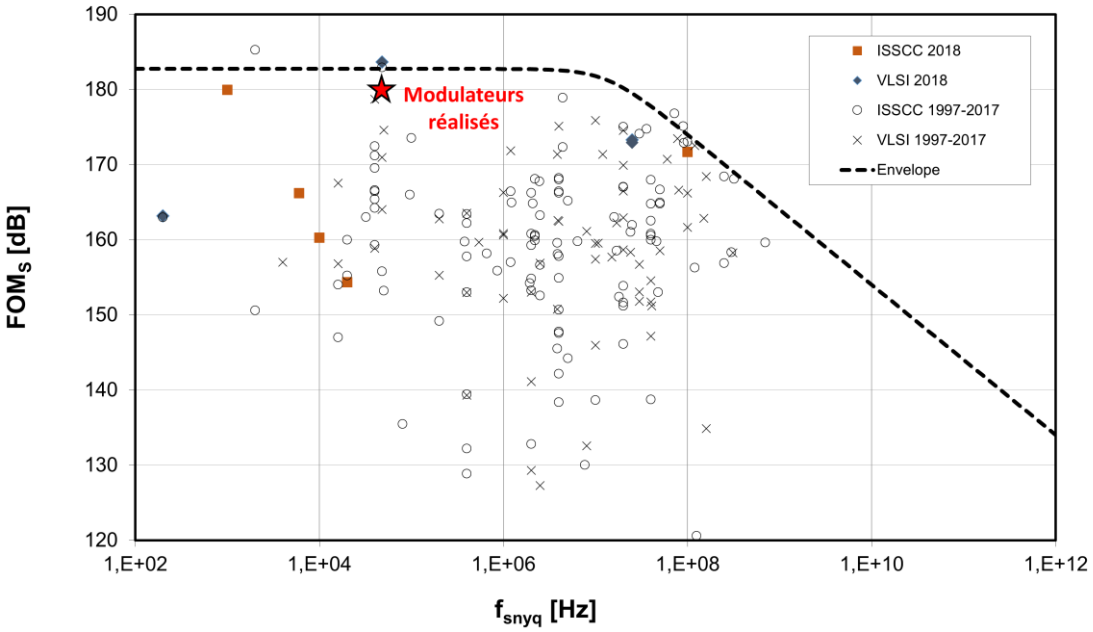


Figure V-1 Placement des modulateurs sur le comparatif des CAN de Murmann

Contents

List of figures

Chapter 1

Introduction

1.1 Motivation

The portable devices such as mobile and the new devices of the internet-of-things incorporate more and more voice applications. The software processing like voice recognition and authentication or microphone beamforming in case of a far-distance talk necessitates a low noise acquisition. In contrast, with high signal levels (wind noise, concert, sport event), the microphone can saturate and deteriorate the signal quality. To address these issues, microphone's research has made in the recent years a noticeable progress in their dynamic range (DR) by both reducing the microphone noise floor and increasing their acoustic overload point (AOP) as illustrated by figure 1.1.

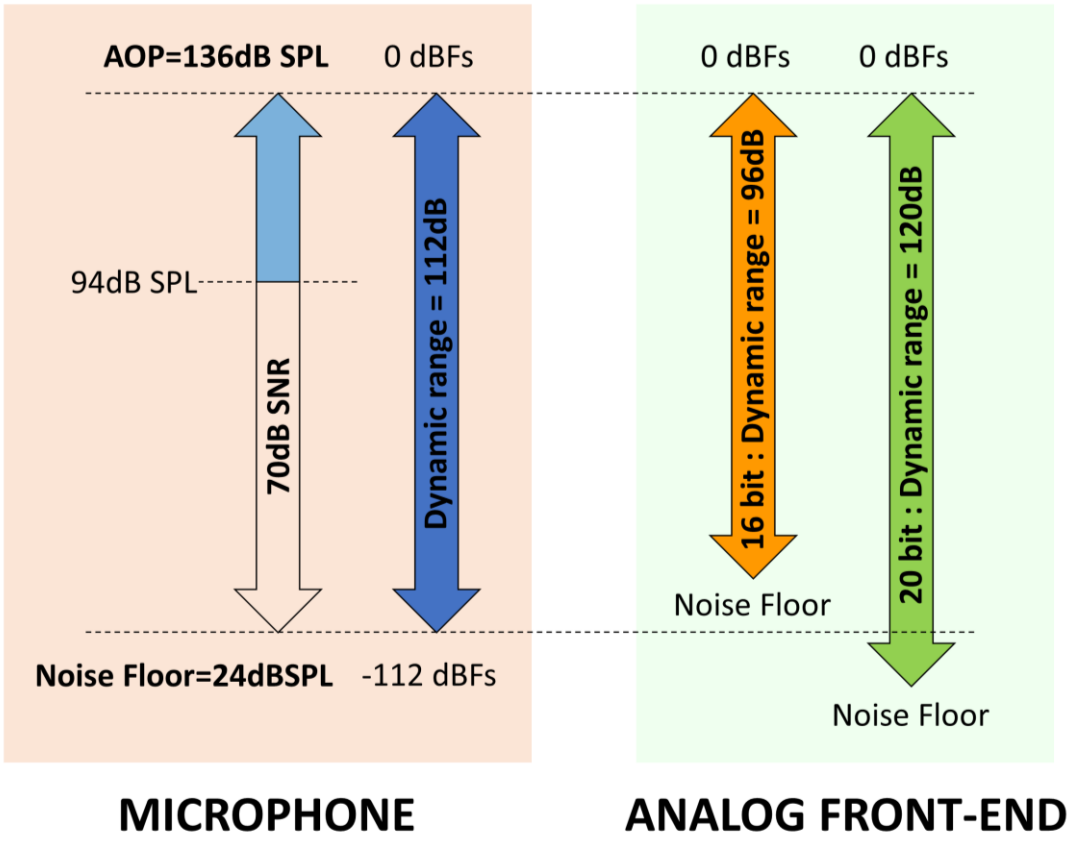


Figure 1.1 New generations microphones need extended dynamic range analog front-ends

This leads to a new generation of microphones with a dynamic range as high as 110dB which exceed by far the 16-bit resolution typically used by analog front-end acquisition systems. These front-ends need to not impact the microphone signal quality and then necessitate a large dynamic range of about 120dB. On the other hand, since they are employed on battery-powered devices, these analog front-ends should have lowest power consumption possible, as well as a low area occupation to facilitate their integration in constrained devices.

The DR requirement on the analog front-end can be decorrelated with other requirements like distortion: the microphone peak THD is generally comprised between 60 and 80dB [1, 2, 3], relaxing the acquisition system linearity requirement. The noise level can also be relaxed at high signal levels: the degradation between a 100dB and a 110dB SNR signal is quite hard to perceive and negligible, knowing also that the distortion at these signal levels will dominate over the noise. Then one can conclude that the analog front-end does not need a power-hungry 120dB DR analog/digital converter: a low-noise amplifier with a programmable gain from 0 to 20 dB and a 100dB DR ADC is then sufficient to cover the full microphone dynamic range without a perceptible signal degradation (as illustrated in figure 1.2) and with a reduced power consumption. This thesis will then concentrate on the research of power-efficient and low-cost data converters for a microphone analog front-end application. Targeting a high dynamic range of 100dB and a minimum THD of 90dB, sigma-delta modulators are the architecture of choice [4] for a good trade-off between resolution and power efficiency in the audio bandwidth. Both discrete-time and continuous-time solutions are possible and will be explored in this thesis, with a special attention made on the power consumption and the area occupation.

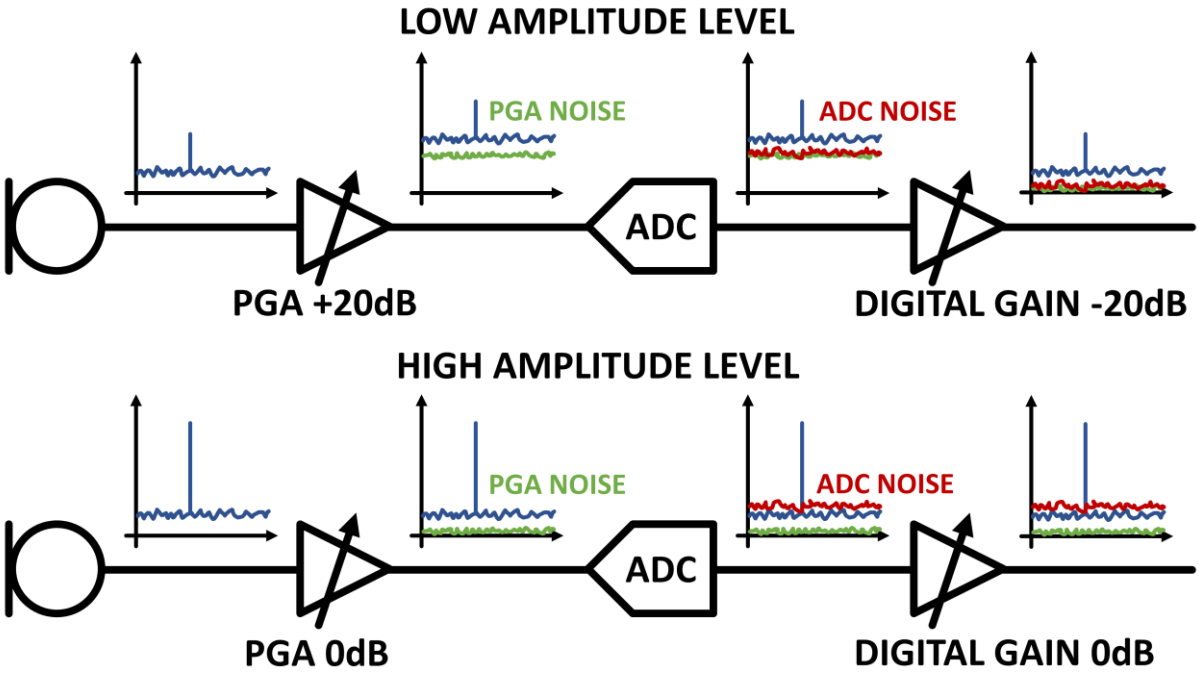


Figure 1.2 Programmable low-power analog front-end to cover microphone dynamic range

1.2 Contributions

The main contributions of this thesis are listed below:

- A high-level model of a discrete-time integrator that allows the designer to define accurately the amplifier specifications (DC gain, GBW product, Slew Rate) knowing only the sampling frequency and capacitors values (sampling, integration and parasitic ones)
- A novel inverter-based amplifier robust to Process, Voltage and Temperature variations and well suited for discrete-time (switched-capacitors) amplifiers or integrators
- The design of a high-resolution discrete-time modulator shows up some effects not present in lower resolution modulators that have to be handled
- A detailed calculation method of equivalence between continuous-time and discrete-time modulators, easing the understanding and design process and applicable to compensate the Excess Loop Delay
- A generic compensation DAC coefficients calculation method for continuous-time modulators using FIR feedback DACs
- The integration of a VCO-based quantizer in a high-resolution continuous-time modulator without impact on linearity and stability
- A modification of the Data-Weighted-Averaging algorithm to extend its use with tri-level DAC elements
- A highly insensitive to Inter-Symbol-Interference resistive DAC using tri-level DAC elements

1.3 Organization

Chapter 2 introduces the sigma-delta modulation principle which combines oversampling with noise-shaping. Discrete-time modulators theory in z-domain is presented together with linear models for several noise shaping orders, and the principal concepts (NTF, STF, feedback/feedforward architecture) are also introduced. In a second time, a behavioral model of a non-ideal discrete-time integrator is developed to help the designer to define quickly and accurately the integrators dynamic specifications at high-level like its DC gain, GBW product and Slew Rate (SR).

Chapter 3 starts by reviewing the state-of-the-art of audio discrete-time modulators and identifying their key points. Then a design of a low-area and 103dB Dynamic Range (DR) modulator is made for the targeted audio application, using low-power inverter-based amplifiers. These amplifiers are designed with a specific biasing stage which ensures PVT variations robustness while maintaining their class AB behavior for power efficiency. The complete modulator design is presented and post-layout simulation results are shown.

Chapter 4 concerns the continuous-time modulators theory and their specificities that are different from discrete-time ones. The equivalence with a discrete-time modulator and a coefficient calculation method is presented, based on the impulse-invariance transform. This method is generic and can be applied to any modulator architecture to synthesize a selected Noise Transfer Function (NTF) from z-domain into an equivalent continuous-time implementation. The principal advantages (anti-aliasing property) and drawbacks (Excess Loop Delay and jitter sensitivity) of continuous-time modulators are detailed. Theory for modeling and estimating the impact on modulator performance is developed, together with compensation methods. Special case of modulators using a FIR feedback DAC is also detailed, with a generic method to calculate the compensation DAC coefficients to restore the original NTF.

Chapter 5 reviews the state-of-the-art design of continuous-time modulators. VCO-based quantizers are more and more used in high-bandwidth modulators for their mostly digital nature (low area and power), and since they provide one integrating stage and a multibit quantization, while the relatively low linearity of VCOs is not limiting factor in medium-resolution converters. A continuous-time modulator design using a VCO-based quantizer is made, targeting high linearity for an audio bandwidth converter of 16-bit resolution. The VCO-based quantizer low linearity and PVT sensitivity is counterbalanced by architectural and design choices to have no impact on modulator resolution, linearity and stability.

Chapter 6 concludes the thesis and presents the perspectives of ameliorations for both discrete-time and continuous-time modulators in terms of power efficiency and area cost, which are the two main important specifications with the dynamic range.

Chapter 2

Discrete-time modulators and integrators non-idealities modeling

Sigma-delta Analog to Digital Converters (ADC) are the architecture of choice when targeting a good trade-off between high resolution, low signal bandwidth and low power consumption [4], while Nyquist-rate converters are generally limited in resolution by the matching of their elements [5]. To reach higher resolutions, sigma-delta modulators combine two concepts: oversampling and noise shaping. Shannon-Nyquist theory states that any signal can be sampled without loss of information with a minimum sampling frequency higher than twice the maximum signal frequency in the band of interest [6]. Oversampling is the process of sampling a signal with a frequency significantly higher than the Nyquist rate, and the ratio of the sampling frequency over the Nyquist rate is called oversampling ratio (OSR). Using a Nyquist ADC to convert a given signal will produce a certain amount of quantization noise on its output equal to $q^2/12$ where q is the quantization step (LSB size) [5, 7, 8]. The same ADC (same q), when oversampled, produces the same amount of quantization noise $q^2/12$ on its output but spreads it on a larger bandwidth, then the remaining amount of quantization noise in the signal bandwidth is lower due to a lower quantization noise pdf ($q^2/12/F_{SAMPLING}$). This is illustrated with the help of figure 2.1 where the integrated noise PSD of both converters up to half their sampling frequency are equals.

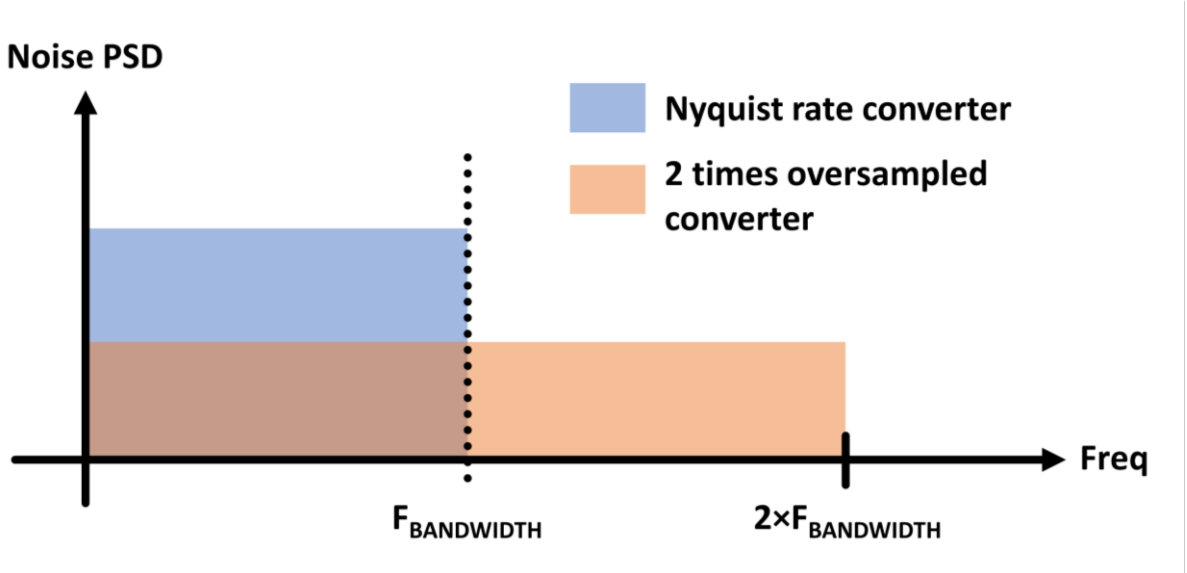


Figure 2.1 Effect of oversampling on remaining in-band noise power

The noise shaping consists in giving a determined shape to the quantization noise. In Nyquist rate ADCs this is useless since the total bandwidth is the signal bandwidth so it cannot influence the quantization noise value in the signal band. In oversampled converters however, the noise shaping allows to lower the quantization noise in the signal bandwidth to reject it out of band as shown on figure 2.2. Then, with a proper digital decimation filter, it is possible to filter out the out-of-band noise and to reduce the sampling frequency to go back to the Nyquist rate. The combination of oversampling and noise shaping can then extend the resolution of a fixed number of bits ADC. These combined concepts are powerful enough to obtain almost all practical resolutions even with a single-bit ADC.

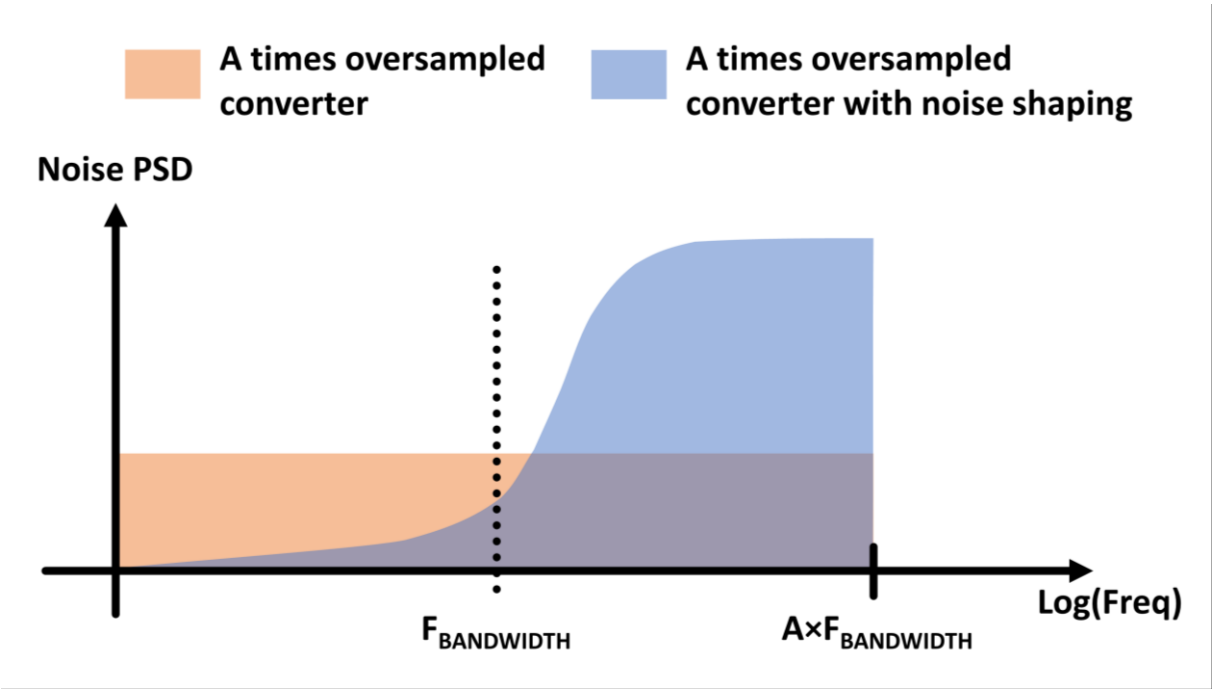


Figure 2.2 Effect of noise shaping on remaining in-band noise power

2.1 Discrete-time modulators theory

This part proposes a brief review of sigma-delta modulator theory, which is required to understand better the rest of the manuscript and to explain the different design choices made during this PhD work. More detailed theory explanations can be found in the references [5, 7, 8, 9]. A block diagram of a discrete-time sigma-delta modulator together with its linear model is presented on figure 2.3. It is a sampled hybrid analog/digital system composed of a loop filter, a quantizer and a DAC to close the feedback loop.

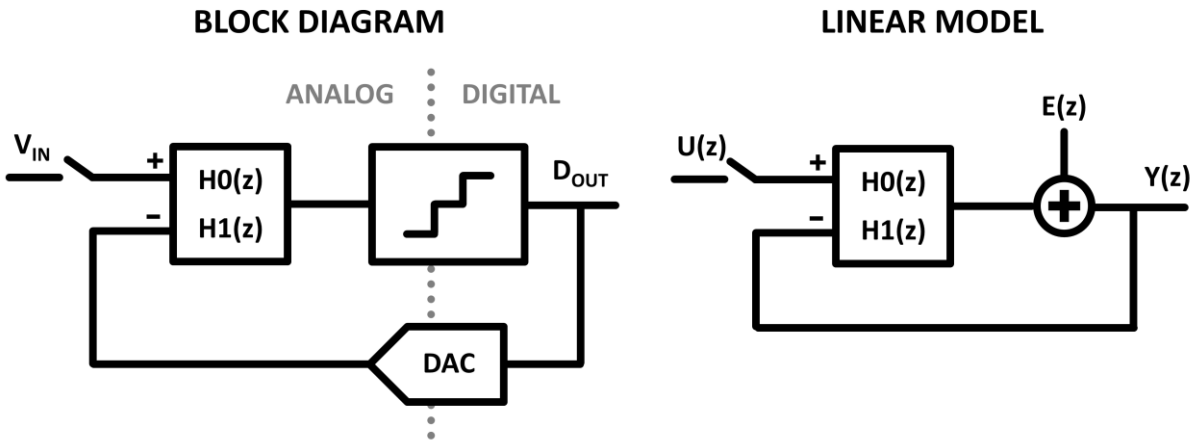


Figure 2.3 Generic sigma-delta modulator block diagram and linear model

The quantizer is the only nonlinear element and is modeled using an addition of an error signal representing the quantization error. The modulator output $Y(z)$ is function of its input $U(z)$, the quantization error $E(z)$ and the loop filter transfer functions $H_0(z)$ and $H_1(z)$, respectively for input signal and feedback signal; the relation is given in equation 2.1:

$$Y(z) = \frac{H_0(z)}{1 + H_1(z)}\, U(z) + \frac{1}{1 + H_1(z)}\, E(z) \tag{2.1}$$

From (2.1), two transfer functions can be derived: the signal transfer function $STF(z)$ and the noise transfer function $NTF(z)$ which are defined by:

$$STF(z) = \frac{H_0(z)}{1 + H_1(z)}, \qquad NTF(z) = \frac{1}{1 + H_1(z)} \tag{2.2}$$

$NTF(z)$ should be close to 0 in low frequencies to lower the quantization noise in the signal band, necessitating a high gain in $H_1(z)$. Concerning $STF(z)$, it should be close to one in signal band i.e. at low frequencies, then a high low-frequency gain is needed in both $H_0(z)$ and $H_1(z)$. That's why low-pass sigma-delta modulators employs integrators in their loop filter. The number of cascaded integrators in the loop filter determines the order of quantization noise shaping.

2.1.1 First-order modulator

The first order sigma-delta modulator is a case where its loop filter transfer functions $H_0(z)$ and $H_1(z)$ are equal (denoted $H(z)$) and are implemented by a single delayed integrator which transfer function is:

$$H(z) = \frac{z^{-1}}{1 - z^{-1}} \tag{2.3}$$

The corresponding $STF(z)$ and $NTF(z)$ are the following:

$$STF(z) = z^{-1}, \qquad NTF(z) = 1 - z^{-1} \tag{2.4}$$

The signal component of the modulator output is just a delayed version of the input signal, while the quantization noise is first-order differentiated. The high-pass filtering effect on the quantization noise is illustrated on figure 2.4. The first-order high-pass filter scales with the

sampling frequency: increasing the oversampling ratio will reduce the remaining in-band quantization noise as can be seen on figure 2.4. The resolution improvement as function of the oversampling ratio can be computed. Let's assume a white noise accounting for the quantization error with a power spectral density $S_e(f) = e_q$ (uniform distribution), this assumption is valid if the quantizer input changes for a sufficiently large amount from sample to sample [10]. The quantization noise at modulator output is shaped by the NTF, leading to a PSD given by:

$$S_q(f) = S_e(f) \times |NTF(f)|^2 = S_e(f) \times \left|1 - e^{-\frac{j2\pi f}{fs}}\right|^2 \tag{2.5}$$

The remaining in-band quantization noise power n_q is then given by:

$$n_q = \int_{-f_b}^{f_b} S_e(f) \times \left|1 - e^{-\frac{j2\pi f}{fs}}\right|^2 df = \int_{-f_b}^{f_b} e_q \times 4sin^2\left(\frac{\pi f}{f_s}\right) df \tag{2.6}$$

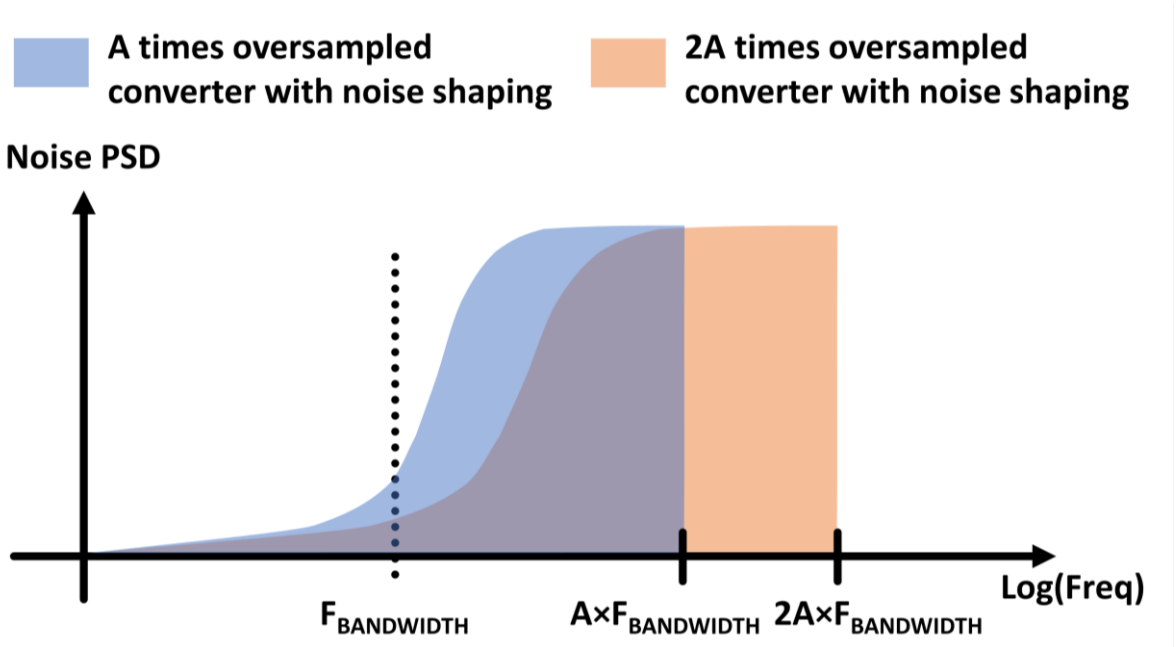


Figure 2.4 Effect of doubling the OSR on noise shaping response

Where f_b denotes the bandwidth frequency and f_s the sampling one. Considering a large oversampling ratio (OSR>>1), the sine function can be first-order approximated by sin(x) = x, leading to the following in-band noise power [8]:

$$n_q = \int_{-f_b}^{f_b} e_q \times 4\pi^2 \left(\frac{f}{f_s}\right)^2 d\left(\frac{f}{f_s}\right) = e_q \times \frac{\pi^2}{3(OSR)^3} \tag{2.7}$$

The in-band noise power is inversely proportional to the power of three of the oversampling ratio, meaning that each time the OSR is doubled a 9dB reduction of the quantization noise is obtained. The oversampling ratio is the first design parameter of a sigma delta modulator which trades speed with resolution.

Figure 2.5 shows the simulation results of an ideal first-order modulator with different oversampling ratios. The Signal to Quantization Noise Ratio (SQNR) given by spectral estimation is effectively improved by 9dB for doubling the OSR, in accordance with the theory. The quantization noise is first-order differentiated, meaning that it is following a 20dB per decade curve as illustrated in figure 2.5. Undesirable tones are also clearly visible, which are specific to first-order sigma-delta modulators [5, 7]: the input of the quantizer does not change sufficiently from sample to sample, then its position on the quantization range cannot be considered as totally random as assumed before, but signal-dependent. This effect is therefore amplified due to the use of a 1-bit quantizer since it has the maximum LSB size [8]. A correlation between the quantization error and the signal appears: the quantization noise cannot be still considered as white noise and this is visible at the modulator output. Adding a random dithering signal at the quantizer input can help to reduce the tones amplitude by decorrelation of quantizer input with the signal but it adds an additional noise to the quantization noise. The tones issue and the poor resolution vs sampling frequency (compared to higher-order modulators) are the reasons why in practice the first-order modulator is not used.

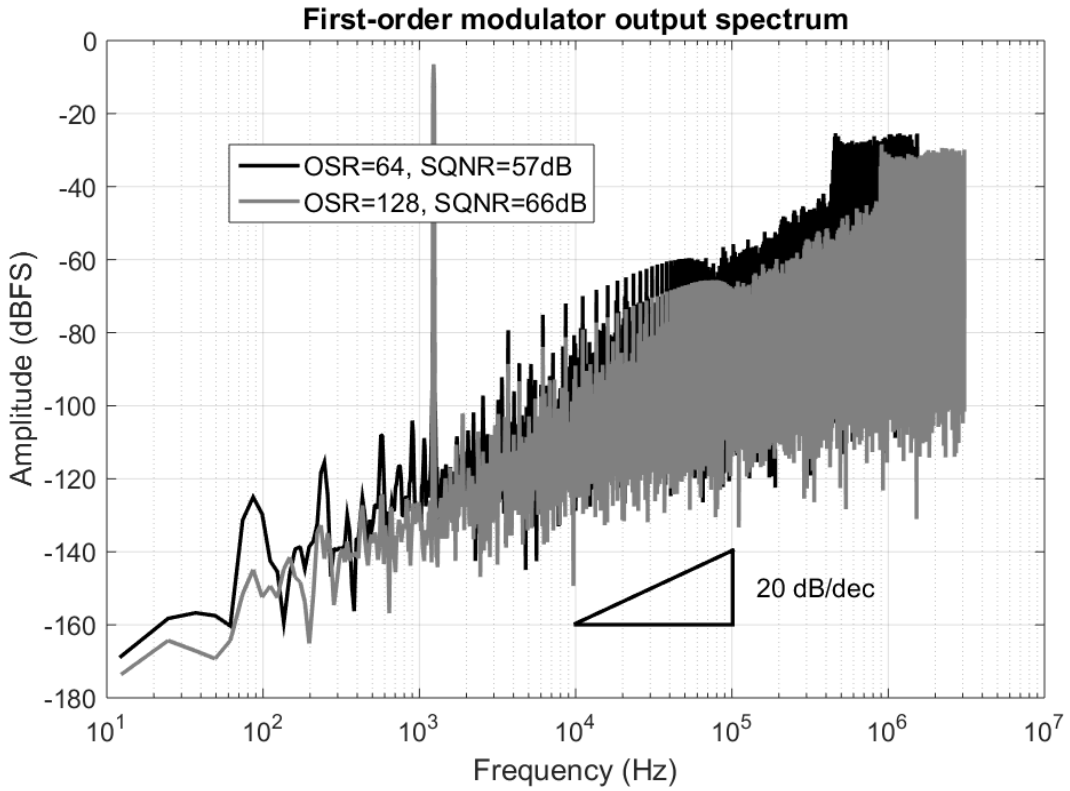


Figure 2.5 Output spectrums of audio (24kHz bandwidth) first-order modulators for different OSR values

2.1.2 Second-order modulator

Another way to improve the resolution of the modulator is to increase the noise shaping order. More quantization noise will be rejected out of band, reducing the OSR needed to obtain a certain resolution. To make a second order modulator, the loop filter is made of a cascade of two integrators to increase its gain at low frequencies. It can also be seen as a first-order modulator where the quantizer is replaced by a first-order sigma-delta modulator. Figure 2.6 shows the linearized block diagram of a second-order modulator.

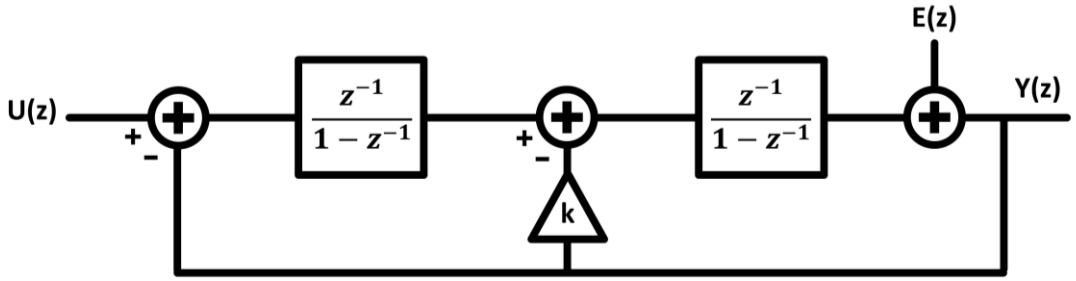


Figure 2.6 Second-order modulator linear model using a Cascade of Integrators with multiple FeedBack (CIFB) structure

The corresponding $STF(z)$ and $NTF(z)$ are:

$$STF(z) = \frac{z^{-2}}{1-(2-k)z^{-1}+(2-k)z^{-2}},\ NTF(z) = \frac{(1-z^{-1})^2}{1-(2-k)z^{-1}+(2-k)z^{-2}} \quad (2.8)$$

With $k = 2$, the STF becomes a two-sample delay and the NTF a pure second-order differentiator, with a quantization noise shaped by a 40dB per decade slope as visible on figure 2.7. The residual tones present on first-order modulator output are no more present on a second-order modulator, meaning that the quantizer input is made sufficiently random by the feedback loop. Here, with an oversampling ratio of 128, a SQNR of 90dB is reachable while the first-order reaches only 66dB for the same OSR, assuming a single-bit quantizer.

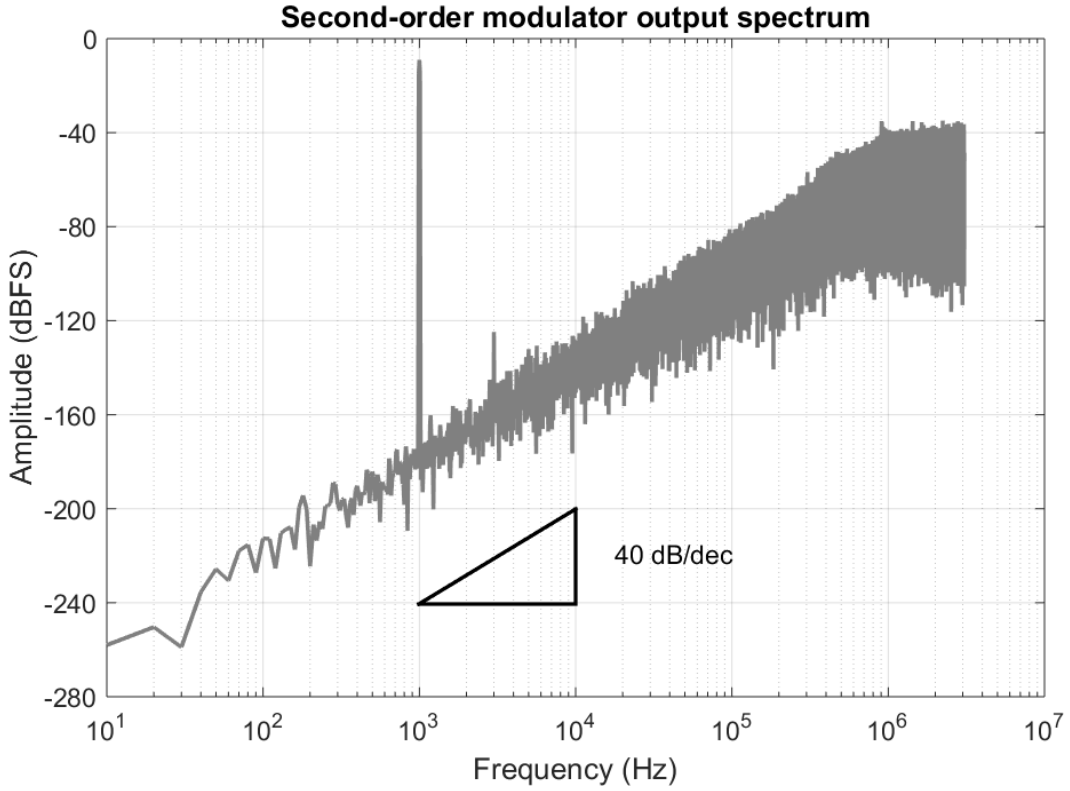


Figure 2.7 Second-order modulator output spectrum: SQNR=90dB over 24kHz for an OSR of 128

The architecture presented on figure 2.6 necessitates two feedback DACs and is generally called Cascade of Integrators with FeedBack (CIFB). Other possible realization is a Cascade of Integrators with FeedForward (CIFF) and is presented by the linear model of figure 2.8. Compared to the CIFB architecture, one feedback DAC is saved (only the main feedback DAC is needed), but a summation operation is required at the quantizer input.

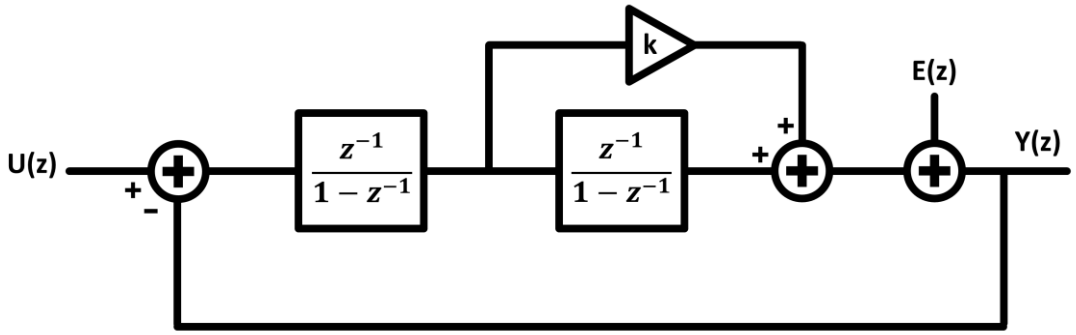


Figure 2.8 Second-order modulator linear model using a Cascade of Integrators FeedForward (CIFF) structure

The corresponding $STF(z)$ and $NTF(z)$ of this structure are given below:

$$STF(z) = \frac{kz^{-1}+(1-k)z^{-2}}{1-(2-k)z^{-1}+(2-k)z^{-2}}, \tag{2.9}$$

$$NTF(z) = \frac{(1-z^{-1})^2}{1-(2-k)z^{-1}+(2-k)z^{-2}} \tag{2.10}$$

With $k = 2$, the NTF remains unchanged compared to the CIFB structure and is still a pure double differentiator. The STF however is no more a number of delay cells but a $2z^{-1} - z^{-2}$ filter which magnitude response is given in the Figure 2.9.

In the signal bandwidth, the gain is flat and unity. At high frequencies (near half the sampling one), the gain can be close to 10dB: such frequencies can overload the quantizer input and make the modulator unstable but fortunately, in a discrete-time modulator, an anti-aliasing filter is needed before and then reduces the high frequency content of the input signal. The combination of the anti-aliasing filter and the STF frequency responses should have a maximum gain of 1 to ensure that these high frequencies cannot influence the modulator stability.

A popular variant of the CIFF structure is presented in figure 2.10. In this structure, a direct feedforward path from input signal to quantizer input is created: the loop filter processes now only the quantization error and the STF is equal to unity. The NTF remains unchanged and is still a second-order differentiator for $k = 2$. The advantage of this structure is that the loop filter processes only the quantization noise and not the signal component: lower linearity requirement is expected on the loop filter and the swings on the integrators are lowered too. This topology was presented in [11] and largely used. Other variants exist [12], lowering more the second integrator swing than the first.

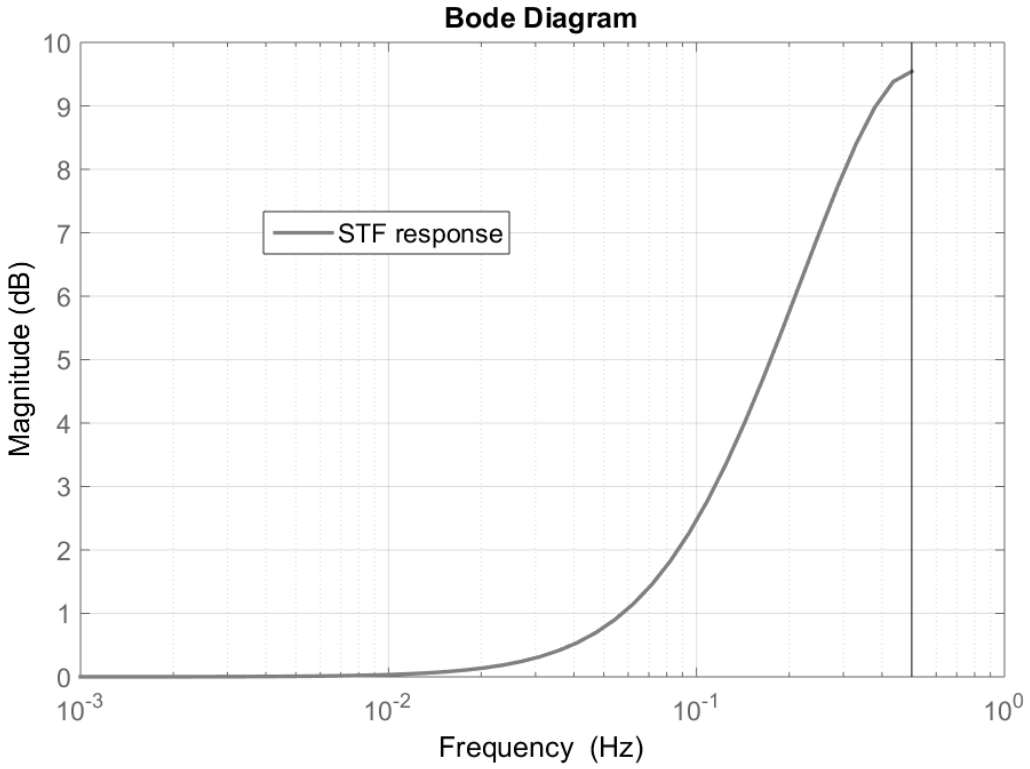


Figure 2.9 STF frequency response of a second-order modulator using a CIFF structure

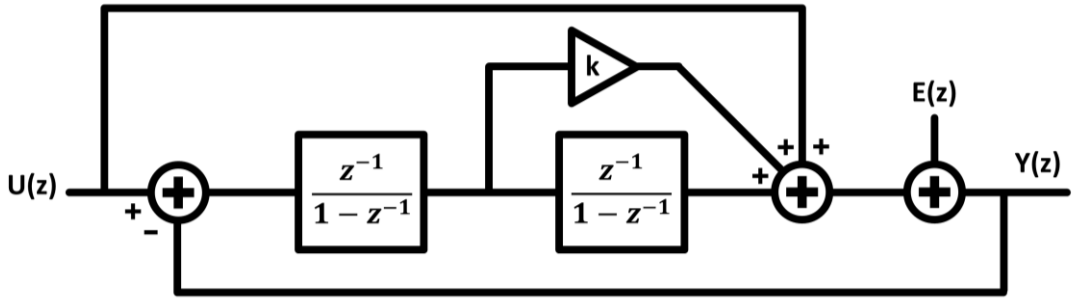


Figure 2.10 Second-order modulator using a CIFF structure with input feedforward

The integrator swing lowering property of feedforward architectures is illustrated with the help of the figure 2.11 where the swings of the two integrators are compared for the three presented structures using a single-bit quantizer. The CIFF structure (whatever input feedforward is used of not) reduces the first integrator swing by 40% compared to the CIFB. On the second integrator, the CIFF structure has also a lower swing, since part of the quantizer input signal is provided by the summation point and doesn't have to pass through the second integrator. This phenomenon is even amplified when adding the input feedforward: the signal component does not pass through the integrators, reducing further the second integrator signal swing.

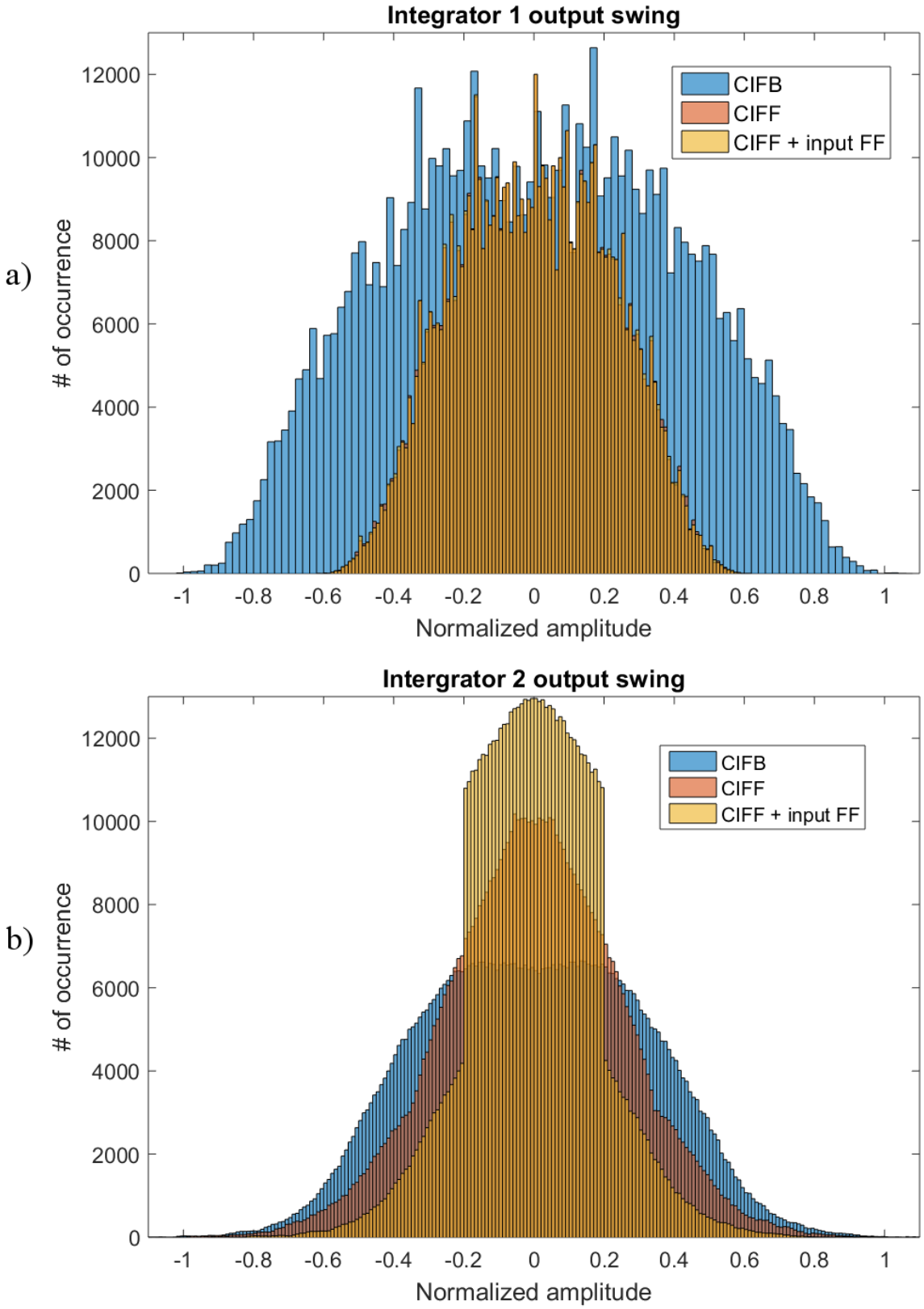


Figure 2.11 Comparison of integrators output swing for the three different structures using a single-bit quantizer a) first integrator, b) second integrator

2.1.3 Higher-order modulators

By cascading more integrators, it is possible to increase further the noise shaping order. Both CIFB and CIFF structures are possible to implement any modulator order [5]. Figure 2.12 shows the NTF frequency response in function of the ideal differentiation order. As can be seen, increasing the order by one increases the slope of the noise shaping by 20dB per decade, reducing further the remaining noise at low frequencies. Near half the sampling frequency however, the quantization noise magnitude is largely increased (up to 24dB for a 4th order modulator) which can lead to instability for modulator with orders N > 2.

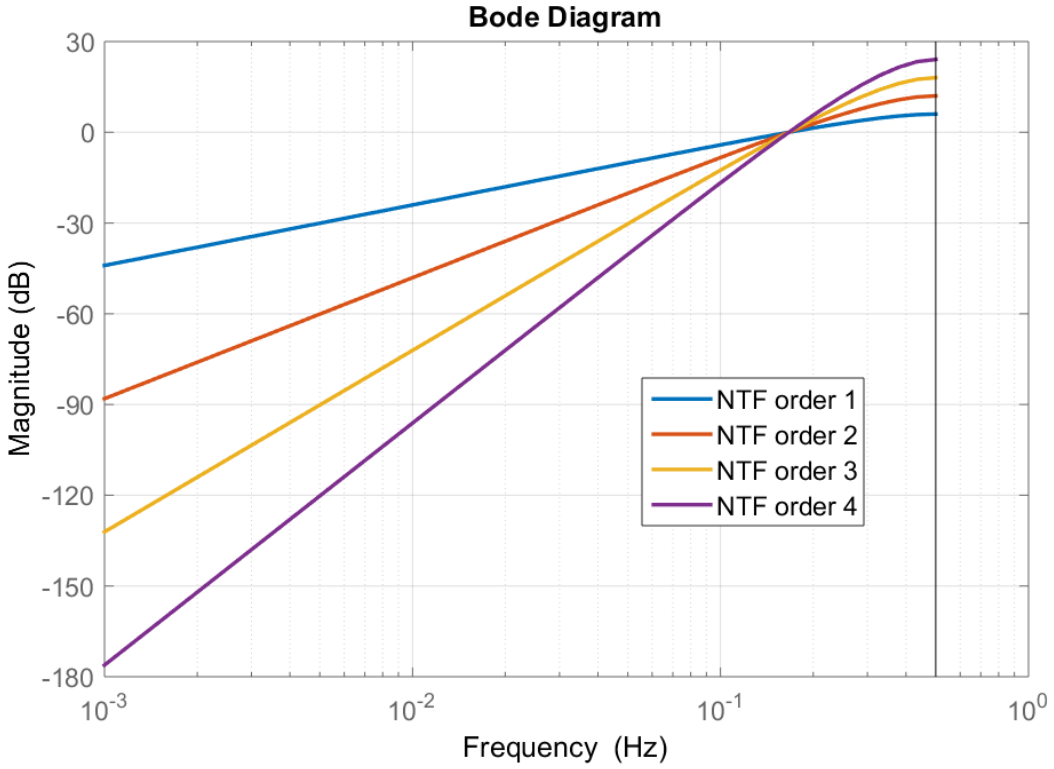


Figure 2.12 Comparison of NTF frequency responses for different orders

The modulator instability appears when the integrator internal states overload or more frequently, when the quantizer input overloads [5, 7, 8]. When the latter is a single-bit, contrarily to a multibit one, it has no clear defined gain and input range: it processes only the sign of the signal and so is a highly non-linear element, its gain varies from sample to sample depending the sample value and its input range is unlimited as illustrated by figure 2.13. For large signal, the quantizer gain is weak while the opposite occurs for small signals. When a multibit quantizer is used, its gain is defined and equals unity, and its input range is bounded as can be seen on figure 2.13. Combining equations (2.1) and (2.2), it is possible to express the quantizer input $V(z)$ as:

$$V(z) = Y(z) - E(z) = STF(z)U(z) + (NTF(z) - 1)E(z) \tag{2.11}$$

Where $U(z)$ and $Y(z)$ are the modulator input and output signals and $E(z)$ the quantization error. The quantizer input is then dependent on the input signal so it is more likely to overload at large input levels.

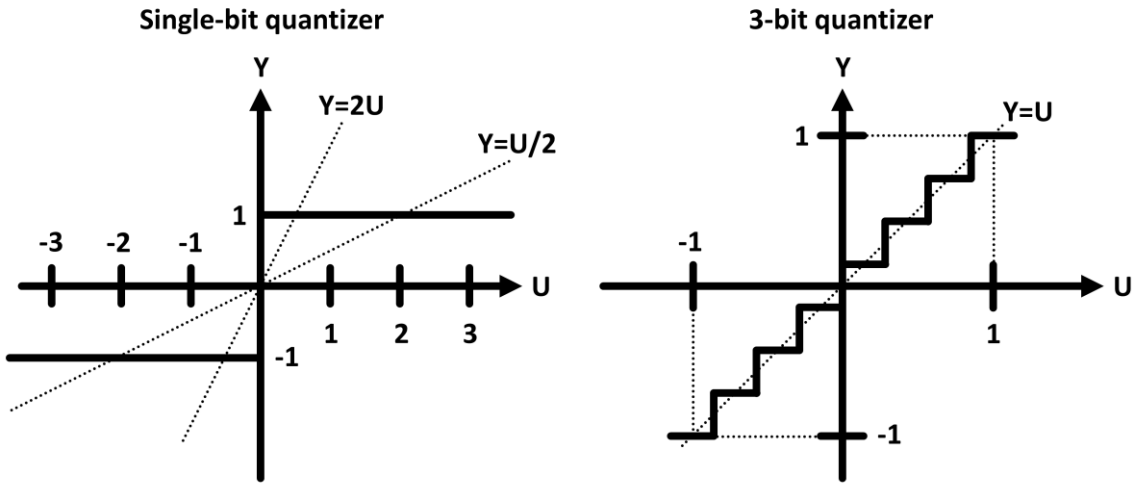


Figure 2.13 Single-bit quantizer with undefined gain and 3-bit quantizer with unity gain

The STF plays a role here because it can provide some gain at high frequencies (in feedforward structure) but this should not be a problem if the combination of anti-aliasing filter and STF frequency responses doesn't have a gain superior to unity as explained before. The quantizer input is also dependent of the quantization noise level, which is determined by the number of bits in the selected quantizer. A single-bit quantizer leads to the highest quantization error and then is more likely to drive the modulator into instability than a multibit quantizer. The NTF also plays a role here: as shown on figure 2.12, the large gain at high frequencies (near half the sampling one) increases the quantization amplitude which also could drive the modulator to instability. The idea to prevent this (given by Schreier et al. in [5, 7]) is to reduce the maximum gain of the NTF to an acceptable value preventing the quantizer overload, which must be checked by simulations with several input signals to ensure the modulator stability.

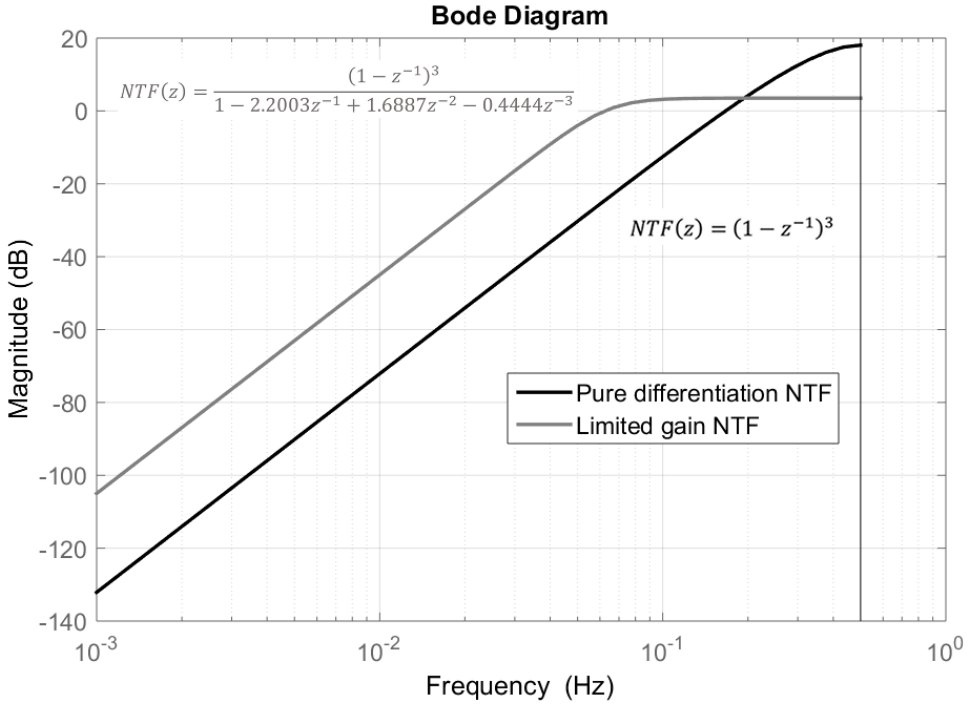


Figure 2.14 Third-order modulator NTFs frequency responses: ideal differentiation and limited gain

Figure 2.14 illustrates this concept: instead of having the poles of the NTF equal to zero (pure third-order differentiator), these poles are moved at a lower frequency than half the sampling one. In that case, the NTF gain at high frequency can be made flat and limited to 1.5. However, the noise shaping is not optimal and some loss is made compared to the ideal one, but this is the price to pay for stability. The amount of attenuation needed for stability depends on the maximum input level the modulator has to support compared to its DAC reference level and on the number of bits used in the quantizer.

Let's first work on a third-order modulator with a normalized full scale of 1 and a single-bit quantizer. The linearized model of this modulator is presented in figure 2.15 where the single-bit quantizer is modeled by a gain k (which can vary from sample to sample and is dependent of signal level and modulator coefficients) and an additional quantization noise. The corresponding $NTF(z)$ is the following:

$$NTF(z) = \frac{(1 - z^{-1})^3}{1 + (ka_1 - 3)z^{-1} + (3 + k(a_2 - 2a_1))z^{-2} + (k(a_1 + a_3 - a_2) - 1)z^{-3}} \tag{2.12}$$

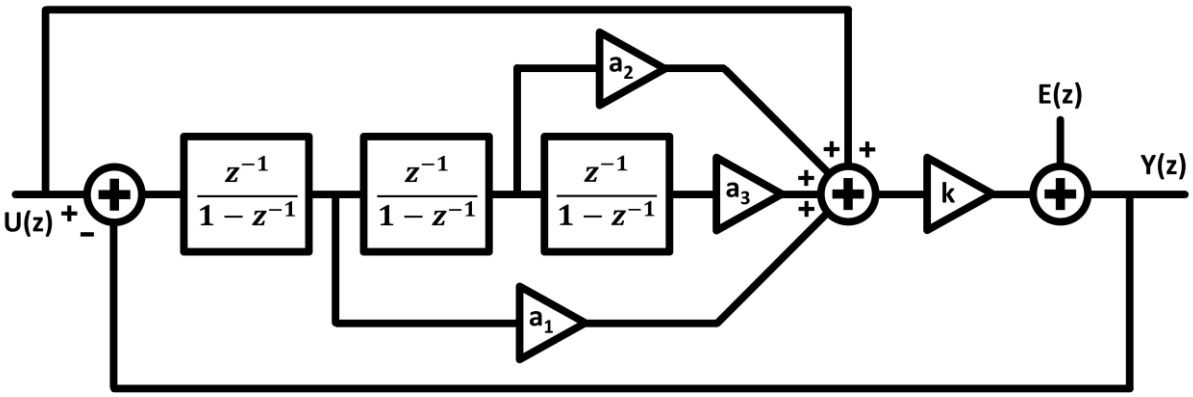


Figure 2.15 Third-order modulator linear model with undefined quantizer gain

For $a_1 = a_2 = 3$, $a_3 = 1$ and $k = 1$ (meaning a unity gain quantizer), the NTF is a pure third-order differentiator, with maximum noise shaping performance and high frequency gain. For $a_1 = 0.7997$, $a_2 = 0.2881$, $a_3 = 0.044$ and $k = 1$ the NTF has a limited high frequency gain of 1.5 enhancing the modulator stability, as plotted in figure 2.14.

The pole-zero plot of these two NTFs in function of the quantizer gain (from 0 to 1) is shown in figure 2.16. As already stated, the worse stability condition is for large input levels, when it is expected a large amplitude at quantizer input and then a low quantizer gain. For $k = 0$, the three poles of the NTFs starts at DC since in that case both NTFs equals unity. When k increases, one pole of each NTF goes along the real axis from dc to zero without impact on stability since they are always in the unit circle. However, there is a pair of complex conjugate poles for each NTF that describes a circle behavior with one part outside of the unit circle.

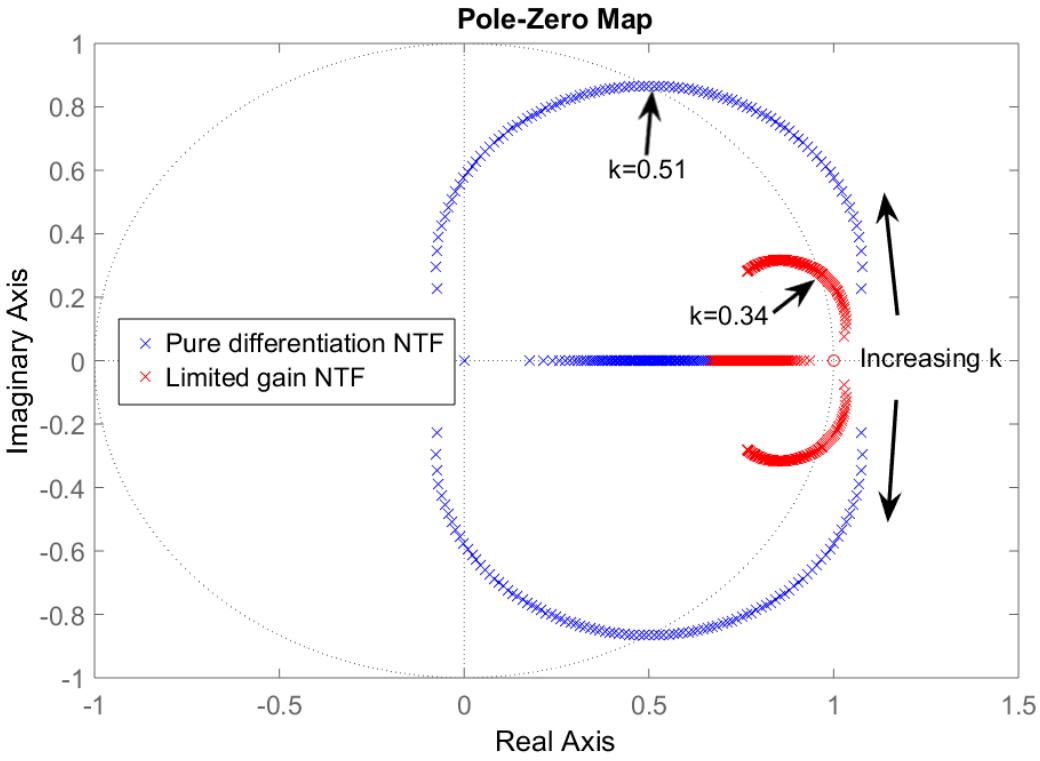


Figure 2.16 Pole-zero plot of third-order NTFs: pure differentiation and limited gain NTFs in function of quantizer gain k

The minimum value of k for which the NTFs complex conjugate poles enters the unit circle gives the condition for the modulator stability. Here, for the pure third-order differentiator NTF this value of k is 0.51 while for the limited gain NTF the value of k is 0.34. Since we work with a normalized modulator which quantizer outputs plus or minus unity, we can deduce that the maximum quantizer input can be about two for the ideal NTF (1/0.51) and about 3 for the limited gain NTF (1/0.34). The quantizer input is given by equation (2.11) and must be bounded by the $1/k$ values found before. Then, using (2.11) it comes the following relation on the maximum stable amplitude U_{max} which could be also find in close terms in [5] (assuming that STF and anti-aliasing filter combined frequency responses have a maximum gain of 1):

$$U_{max} < \frac{1}{k} - \|NTF(z) - 1\|_\infty \times \max\big(E(z)\big) \tag{2.13}$$

The maximum quantization error (considering the quantizer input is not overloaded) is half a LSB. For a single-bit quantizer outputting values of 1 or -1, the LSB is 2 leading to a maximum quantization error of 1. The infinite norm of $NTF(z) - 1$ is the maximum gain this function reaches over frequency and depends on the NTF as shown in figure 2.17.

Using the ideal NTF, the maximum gain of $NTF(z) - 1$ is 16.9dB (linear scale gain of 7) while the limited gain NTF gives a maximum gain of 6dB (linear scale gain of 2) for the $NTF(z) - 1$ function. Then, one can now compute the modulator theoretical maximum stable amplitude in function of the NTF implemented. This gives a result of -5 for the pure third-order differentiating NTF, meaning that the modulator is unstable whatever the input amplitude is. For the limited gain NTF the result is quite different and gives 0.94, meaning that the input signal can be as high as 94% of the modulator full scale.

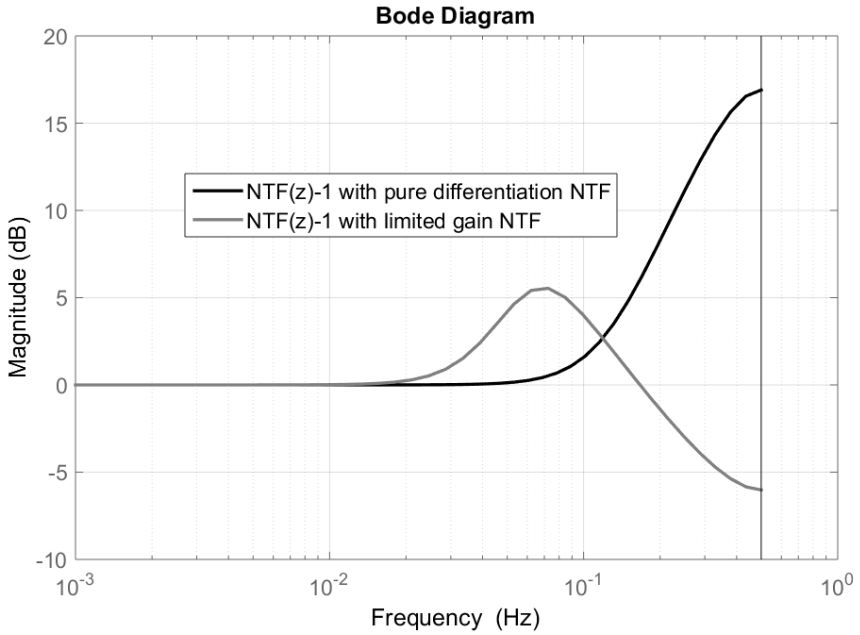


Figure 2.17 Frequency responses of $NTF(z)-1$ for both third-order NTFs: pure differentiation one and limited gain one

In practice however, the modulator maximum stable amplitude has been determined by simulations to 80% of the full scale, meaning that the assumption of a maximum quantization error of 0.5 LSB (linear quantizer) is not totally true for the single-bit quantizer, but still this method gives a useful information about modulator stability and an approximation of the maximum stable amplitude. In all cases, simulations with several input signals (dc, sinusoidal, square, noise) must be done to ensure the stability of a high-order modulator at a given amplitude level. A representative worst-case signal is a full swing square wave at the frequency where the $NTF(z)-1$ response reaches its maximum gain [5], but this occurs generally out-of-band and due to the anti-aliasing filter, this worst-case should not happen.

When using a multibit quantizer, this latter gain is well defined and equals unity: it doesn't change with the signal amplitude as does a single-bit one. Then the poles are in the unit circle for both NTFs but still we need to check that the quantizer is not overloaded. Taking the example of the 8-level quantizer presented in figure 2.13, the quantizer maximum acceptable input is one plus half a LSB leading to about 1.143 (1+1/7). Then, using equation (2.11) for the quantizer input amplitude, it can be found that the modulator maximum stable amplitude is:

$$U_{max} < 1 + 0.5LSB - \|NTF(z)-1\|_\infty \times \max\big(E(z)\big) \tag{2.14}$$

For the limited gain NTF, the modulator is stable for any input signal lower than 0.85, meaning that the modulator is stable with any signal up to 85% of its full scale. In practice, the maximum input level can be even higher since the limit given is accounting for the worst case: the error quantization is always at its maximum value and in the peak frequency of the $NTF(z)-1$ response. With simulation, the modulator keeps its stability with a 1kHz sinewave up to 95% of its full scale. For the pure third-order NTF, the modulator stability is guaranteed for input signals up to 0.143, meaning up to 14% of the full scale. With simulations, the quantizer input is not overloaded with an input signal up to 30% of the full scale. More generally, the relation (2.14) can be derived for a pure N-order differentiation NTF modulator with a M-bit quantizer to be [5]:

$$U_{max} < 1 + \frac{2}{2^M - 1} - \frac{2^N}{2^M - 1} \tag{2.15}$$

Using a number of bit M in the quantizer equals to N plus one ensures the stability of the modulator for at least input signals up to 50% of the full scale since (2.15) gives:

$$U_{max} < \frac{1}{2} < \frac{2^N + 1}{2^{N+1} - 1} \tag{2.16}$$

Hence, a multibit quantizer allows the use of aggressive NTFs and even the pure differentiation NTFs without stability issues while the single-bit quantization must use a limited gain NTF to ensure stability. Note that a multi-bit quantizer helps for stability concerns and increases the modulator resolution, but at the expense of a higher quantizer complexity and feedback DAC linearity. The latter must be ensured to not degrade the modulator resolution which can be done with Dynamic Element Matching (DEM) techniques [5, 7, 9] which noise shape the DAC non-linearity. In all cases the stability can be first order approximated for a modulator input signal level depending on the NTF and the number of bits used in the quantizer but only simulations can ensure the stability of the modulator for an input signal level. The resolution given by a modulator depends then on 3 variables: the noise shaping order N, the OSR (improving the modulator resolution by 3+6×N dB every time we double it), and the number of bits in the quantizer M (improving the modulator resolution by 6 dB per additional bit). Figure 2.18 shows the maximum reachable SQNR versus the OSR value for different modulator architectures. Single-bit architectures, which are prone to instability at modulator orders > 2, have their NTF limited enough in gain to be stable with a -3dBFS input. A fourth-order modulator provides resolution improvement only at high OSR compared to a third-order modulator: this is because its NTF gain is more limited to ensure stability at -3dBFS input level. Multi-bit architectures use pure differentiation NTFs even for the third-order without stability issue at -3dBFS input level. Clearly if a low OSR must be used, a high-order modulator with a multibit quantization is necessary to reach high resolution.

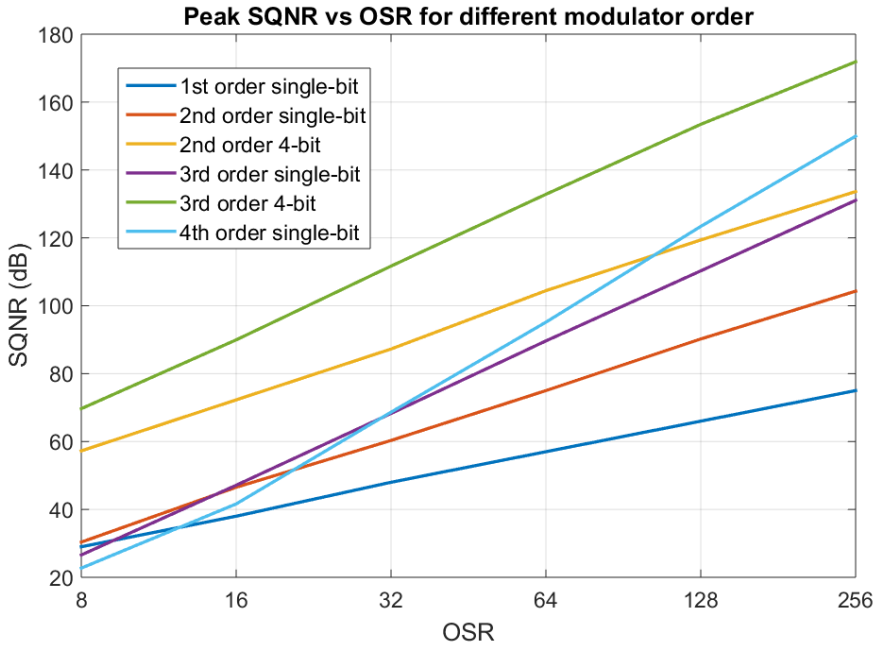


Figure 2.18 Maximum reachable SQNR for different modulators orders and quantizer bits in function of the oversampling ratio

2.2 Modeling integrator non-idealities

Once the architecture is fixed, the next design step is to define the specifications of the modulator sub-elements to ensure that overall modulator performances are not degraded when designed at transistor-level. The most sensitive sub-element of a sigma-delta modulator is its first integrator. In discrete-time modulators, the integrators are most often made by switched-capacitors circuits and although they are known and used for many years, there are no clear rules to get specifications values ensuring that the targeted modulator performances will be reached. That's why integrator non-idealities have to be modeled and inserted in high-level simulations with a relative high degree of accuracy. Finite DC gain can be modeled as a magnitude and a phase error on the integrator ideal transfer function, depending on DC gain value and capacitors ratios. Finite Gain-bandwidth product (GBW) and finite Slew Rate (SR) are much more complicated to model. Several attempts exist in literature [13, 14] but they involve transistor-level parameters which supposes a transistor level design already existing and increases the model complexity. A different approach is presented here, modeling these non-idealities at high-level: by taking into account parasitic capacitances, an accurate model of settling behavior of the integrator output voltage can be determined. Then, the integration error is expressed only as function of specifications (DC gain, GBW, SR) and capacitors values, and is added to each sample. Specifications values directly impact the modulator SNR or SNDR, helping the designer to accurately define requirements for integrators, while keeping the high-level advantage of simulation time. The switched capacitor integrator is presented in its single-ended representation in figure 2.19 during its integration phase.

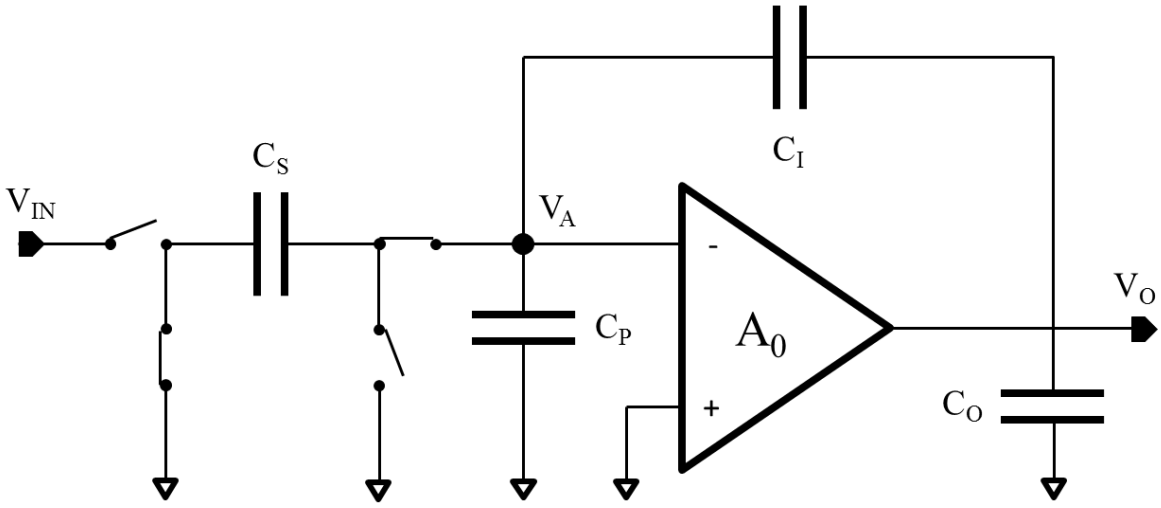


Figure 2.19 Switched-capacitor integrator during its integration phase

A_0 is the DC gain of amplifier, C_S the sampling capacitance, C_I the integration capacitance, C_P the parasitic capacitance at node V_A and C_O the load capacitance of the amplifier. C_P comprises the amplifier input capacitance (mainly C_{GS} of input transistors) and the parasitic capacitances connected to node V_A from sampling and integration capacitors. C_O is the smallest capacitor as it is mainly composed of the capacitances C_{DB} of the output transistors and eventually of a switched-capacitor common-mode feedback circuit. The next stage sampling capacitor is not considered in C_O since it is not connected during integration phase so it doesn't take part in the

amplifier load. Figure 2.20 shows the evolution of the integrator output voltage V_O during the integration period. At the starting time, a charge redistribution occurs in all the capacitors of the integrator. This effect, as it will be shown later, drives the integrator output voltage in the wrong direction, increasing the integration step X. Then the amplifier starts acting, and the output voltage response depends on its specifications (SR and GBW), which must be high enough to limit the integration error during the phase time. At the end of the integration phase, the limited DC gain of the amplifier defines the final integrator output value.

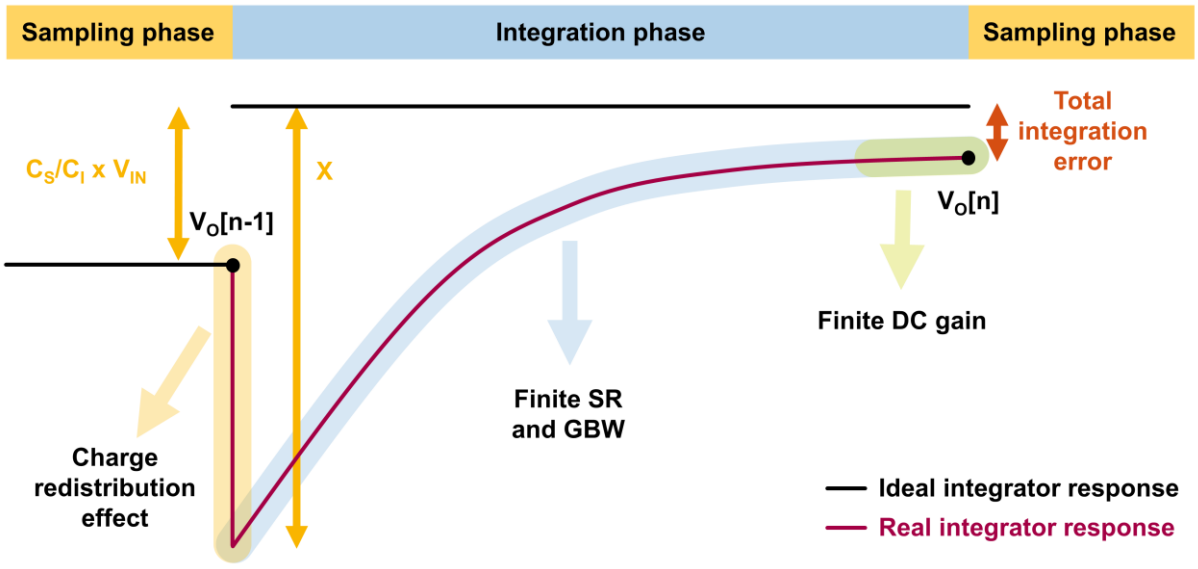


Figure 2.20 Amplifier output response during an integration phase

At the starting time of the integration phase, the sampling capacitor is connected to the internal virtual ground node V_A with a pre-charged voltage $-V_{IN}$. The effect of connecting suddenly this capacitor with a voltage $-V_{IN}$ on the integrator output voltage is the same as if this capacitor was already connected to the virtual ground node V_A with a zero voltage across it and if a step $-V_{IN}$ is applied on input as illustrated on figure 2.21:

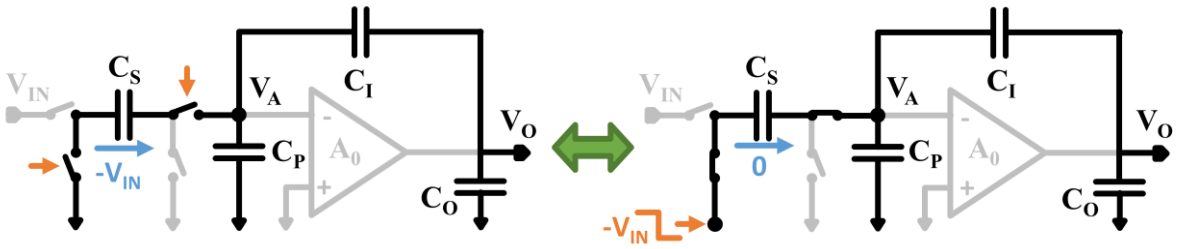


Figure 2.21 Capacitive charge redistribution effect occurring at the starting time of the integration phase

Voltages V_A and V_O are then determined by capacitive voltage division:

$$\Delta V_A = -V_{IN} \times \left(\frac{C_S}{C_S + C_P + \left(\frac{C_I \times C_O}{C_I + C_O} \right)} \right) \tag{2.17}$$

$$\Delta V_O = \Delta V_A \times \frac{C_I}{C_I + C_O} = -V_{IN} \times \left(\frac{C_S}{C_S + C_P + C_O \times \left(1 + \frac{C_S + C_P}{C_I}\right)} \right) \tag{2.18}$$

An interesting thing to note is the negative sign meaning that at the starting point, the integrator output voltage goes in the opposite direction than its final value as already discussed with figure 2.20. Then the total voltage step X used for exponential settling is no more only $C_S/C_I \times V_{IN}$ (figure 2.20): $|\Delta V_O|$ has to be taken into account to get accurate results leading to:

$$X = \frac{C_S}{C_I} \times V_{IN} \times \left(1 + \frac{C_I}{C_S + C_P + C_O \times \left(1 + \frac{C_S + C_P}{C_I}\right)} \right) \tag{2.19}$$

The charge redistribution effect clearly increases the step X, and often by more than 2 since the integration capacitor is usually larger than the sampling and parasitic ones. The increased step X directly impacts the dynamic requirements like GBW and Slew Rate, since the amplifier output has to settle in a fixed time to a larger voltage value. By taking into account this effect, the method used in [15] can be used to determine the settling integration error, dependent on GBW and SR values. This method has the advantages of reporting the settling error directly at the integrator output, avoiding the use of a complex integrator model with internal signals, and giving error values in function of the amplifier specifications, which helps during the design process. The response of the amplifier can be of 3 types during the integration phase: linear, partial slewing or totally slewing. According to [15], integration error ε is:

$$\varepsilon = \begin{cases} -X \times e^{-2\pi \times GBW \times \frac{T}{2}}, & T_{SLEW} \leq 0 \\ -(X - SR \times T_{SLEW}) \times e^{-2\pi GBW \times \left(\frac{T}{2} - T_{SLEW}\right)}, & 0 < T_{SLEW} < \frac{T}{2} \\ -\left(X - SR \times \frac{T}{2}\right), & T_{SLEW} \geq \frac{T}{2} \end{cases} \tag{2.20}$$

$$T_{SLEW} = \frac{X}{SR} - \frac{1}{2\pi \times GBW} \tag{2.21}$$

where T represents the period of one cycle sampling-integration, and T_{SLEW} the slewing duration. A negative or null T_{SLEW} means that the amplifier works in its linear region and is not current limited. A positive T_{SLEW} means that the amplifier response is limited by its maximum current (Slew Rate operation), and depending on the slewing duration the amplifier response is partially slewing and partially linear or completely slewing in the duration of the integration phase. Clearly SR must be high enough to avoid being in a totally slewing amplifier response: the error depends on X value and can be very large for low SR values (no exponential attenuation component in this expression). The partial slewing response shows that the requirements on the amplifier SR and GBW evolve in an opposite way: for a high SR, a lower GBW is required for a fixed error value. This is due to the fact that for a high SR, T_{SLEW} is reduced allowing more time for the linear settling, reducing the GBW constraint. Figure 2.22 compares the presented theoretical model including the capacitive redistribution effect with the real behavior of a switched-capacitor integrator under Cadence simulation at transistor level with the same GBW, SR and capacitor values. We can see in this practical realization that the settling step X is almost 4 times the $C_S/C_I \times V_{IN}$ step, which will increase the dynamic

requirements of the amplifier. The only difference between the two settling behaviors is the smoothed charge redistribution in real case due to the finite resistance of the switches. The proposed model takes into account the settling error in the integration phase but not in the sampling phase. However, in sampling phase the load is only the sampling capacitor of the next stage, which is usually well smaller than the sampling capacitor of current stage (due to different kT/C noise requirements); so in practice if the settling error of the integration phase doesn't lead to a modulator performance loss, the same applies to the settling error of sampling phase.

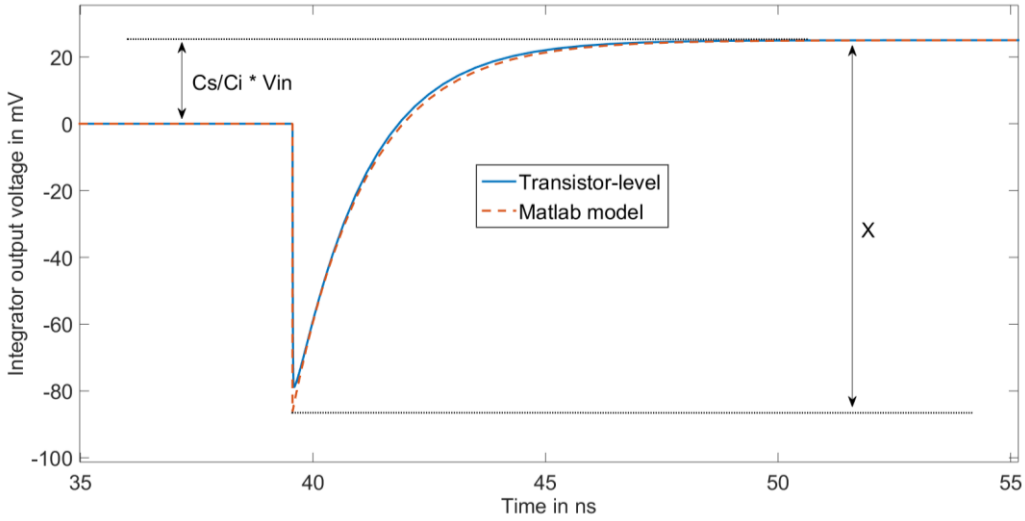


Figure 2.22 Comparison of integrator output voltage between transistor-level circuit and matched parameters model

The final integrator output value (or the value where it tends to go if there is a settling integration error due to finite SR and GBW) is defined by its DC gain and the capacitors values. The charge conservation law used on the integrator presented in figure 2.19 allows to determine the effect of the finite DC gain on the integrator transfer function $H(z)$ which is in ideal case (no parasitic capacitances C_P and C_O and infinite DC gain):

$$H(z) = \frac{C_S}{C_I}\frac{z^{-1}}{1-z^{-1}} \tag{2.22}$$

Applying the charge conservation law at the end of the sampling (2.23) and integration (2.24) phases and equating the two charges gives the following integrator transfer function $H_{REAL}(z)$:

$$Q_S = C_S V_{IN} z^{-1} + C_P \frac{V_O}{A_0} z^{-1} + C_I \left(1 + \frac{1}{A_0}\right) V_O z^{-1} \tag{2.23}$$

$$Q_I = (C_S + C_P)\frac{V_O}{A_0} + C_I \left(1 + \frac{1}{A_0}\right) V_O \tag{2.24}$$

$$H_{REAL}(z) = \frac{C_S}{C_I} \times \frac{\frac{A_0}{1 + A_0 + (C_S + C_P)/C_I} z^{-1}}{1 - \frac{1 + A_0 + C_P/C_I}{1 + A_0 + (C_S + C_P)/C_I} z^{-1}} \tag{2.25}$$

The finite DC gain has two effects on the integrator transfer function $H_{REAL}(z)$: a magnitude error, leading to integrate a step lower than the ideal, and a phase error, denoting a leakage effect. Both of them depends on A_0 and are reduced when DC gain is increased. The magnitude error is weak and can be easily compensated by changing C_S/C_I ratio. The phase error is much more important: it traduces the fact that a fraction of the charge in C_I (composing V_O) is lost at each cycle. The DC gain must be high enough to minimize the impact of this leakage. This leakage leads to a shift in the pole value which is no more 1 as in ideal case. We can note here that C_O doesn't play a role in finite DC gain effects: the system reaches its stable state at the end of each phase when $V_A = V_O/A_0$. V_A is then the driving node of the integrator and only capacitors connected to this node have an effect on the final solution. C_O only acts as an additional load and increases the settling time since the amplifier output current has to drive both C_I and C_O.

To validate the proposed model, a sigma-delta modulator has been designed at transistor level with amplifiers specifications deduced directly from MATLAB/SIMULINK behavioral simulations. The modulator uses a second-order fully feedforward architecture with a 1.5-bit quantizer and is presented on figure 2.23. It is designed for audio application, using a bandwidth of 24-kHz and an oversampling ratio of 256, leading to a 12.288MHz clock frequency. This modulator achieves 109.6 dB SQNR in ideal case, which will be the targeted performance to reach in amplifier specifications definition.

Figure 2.24 shows the implementation of the proposed integrator model in MATLAB/SIMULINK. The input signal is multiplied by C_S/C_I ratio and the finite DC gain magnitude error is applied. This gives the value to integrate. Then settling error due to finite GBW and SR is also added, depending on charge redistribution effect. To finish, the value is added to the preceding output voltage of the integrator, while taking into account the finite DC gain-phase error. This leads to a simple model representing accurately the internal effects in the integrator (as long as the capacitors values are exact) which will not degrade the simulation time. The model parameters are only specifications or capacitances values but there is no transistor or design parameter such as transconductance or bias current contrarily to [13, 14].

A set of behavioral simulations is then run with different values of DC gain, GBW and Slew Rate, for fixed capacitors values: C_S = 2.3pF, C_I = 9.2pF, C_P = 280fF, C_O = 120fF. Results are presented in figure 2.25 for DC gain and in figure 2.26 for GBW and SR effects. The minimum DC gain of integrators should be 40 dB; lower DC gain leads to an important pole shift of integrators transfer function, thus degrading the noise shaping performance.

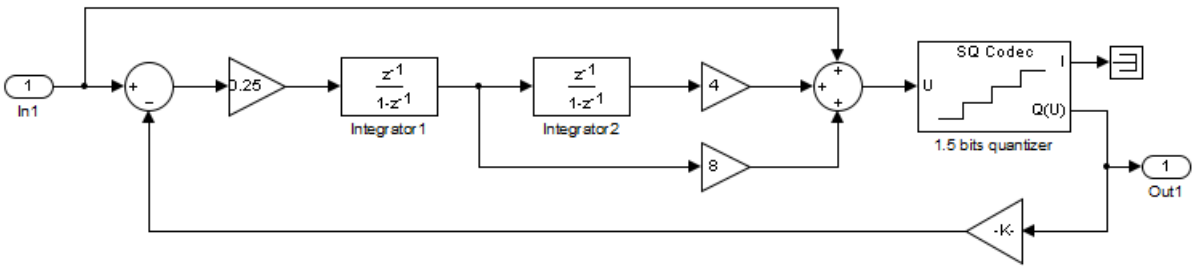


Figure 2.23 Linear SIMULINK model of the second-order modulator with 1.5-bit quantizer and an OSR of 256

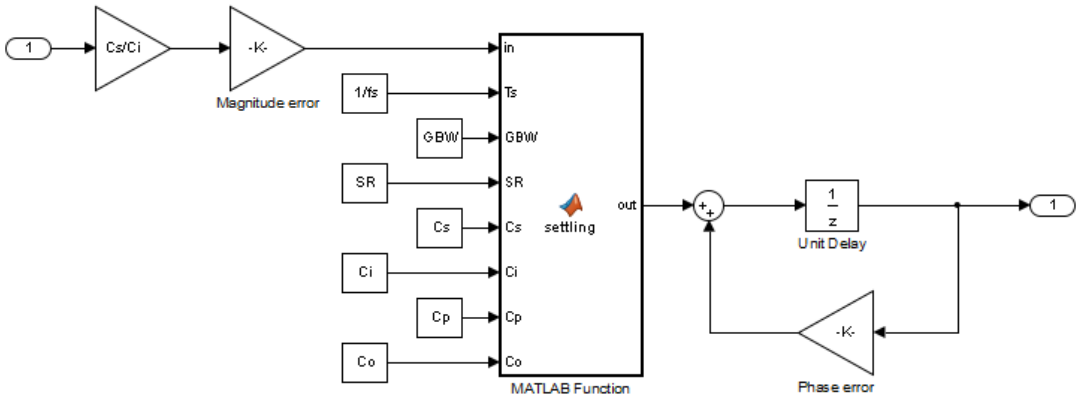


Figure 2.24 Model of a non-ideal integrator dependent on specifications (DC gain, GBW and SR) and capacitances values (C_S, C_I, C_P and C_O)

Figure 2.26 shows that the minimum GBW is 50 MHz which is about 4 times the clock frequency, but the targeted SQNR is in this case reached for a relatively high SR value. For lower SR values the GBW required is increased as expected from (2.20), with a non-linear relation as we can see in figure 2.27, leading to an optimal design point with both relatively low GBW and SR. Figure 2.27 shows also the impact of capacitive redistribution effect on specifications: the requirements curve to reach the targeted SQNR value is shifted to the right side when capacitive redistribution effect is taken into account. Slew Rate requirement is more impacted than GBW, this is due to the fact that the output voltage increases in slewing mode for most of the part of the settling step *X*. From figure 2.25, two DC gain values for modulator integrators are selected: 30 dB and 40 dB and implemented at transistor level to validate our model. Concerning the GBW and SR specifications, two couples of values are selected from figure 2.27: one follows the requirements including capacitive redistribution effect (design point A) and the other not (design point B), in order to validate our model accuracy. For sake of simplicity, in transistor design the bias current is kept constant leading to a constant slew rate of 120 V/us, and the design points A and B differ from GBW values by changing the width of input transistors pair: 70 MHz for the design point A and 50 MHz for the design point B.

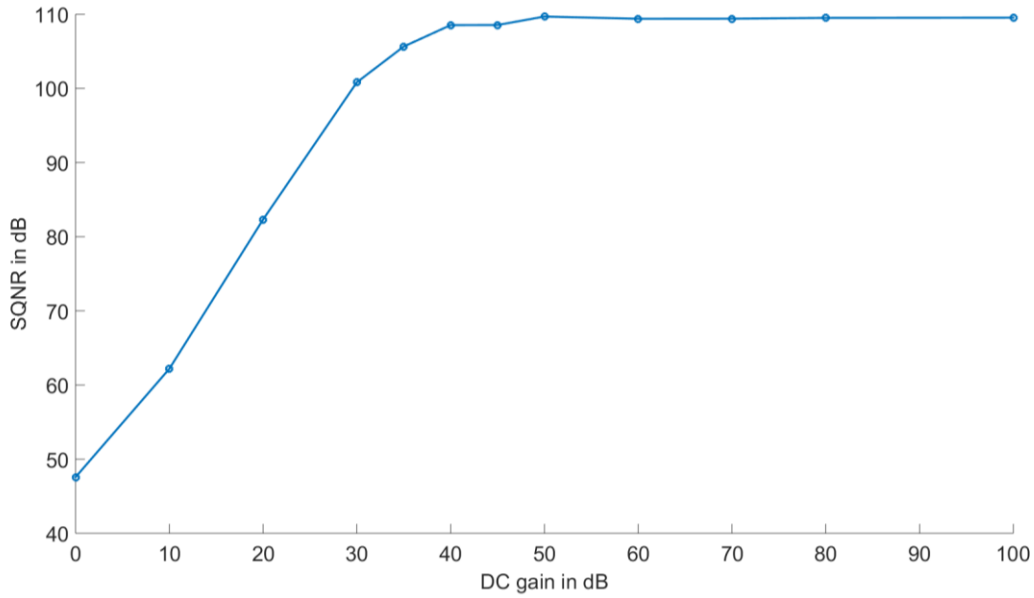


Figure 2.25 Simulated modulator SQNR in function of its integrators DC gain

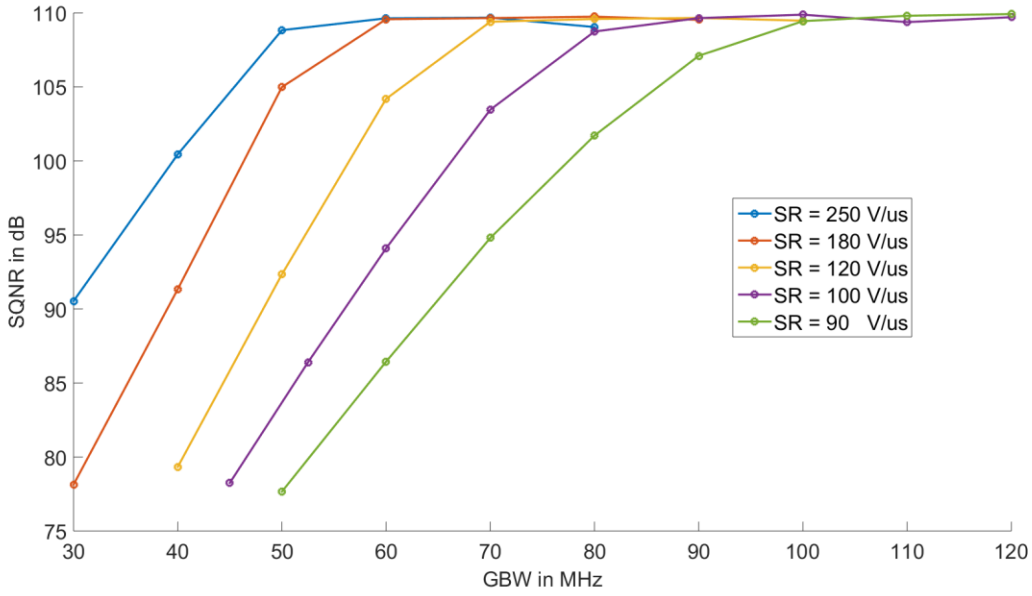


Figure 2.26 Simulated modulator SQNR in function of its integrators GBW and SR

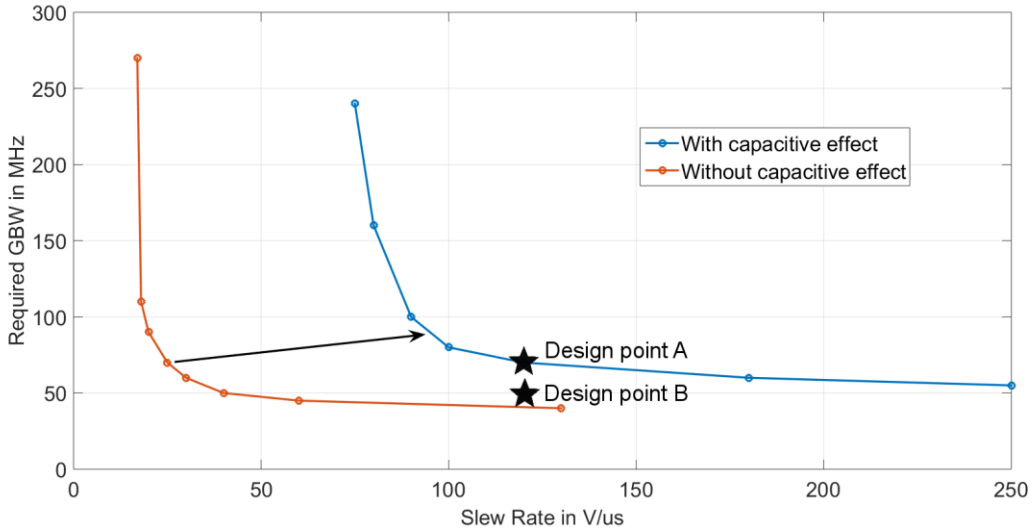


Figure 2.27 Minimum required GBW in function of SR value to not degrade modulator performance and effect of the capacitive charge redistribution effect on the specification requirements

To get representative results and not be influenced by switches related issues (charge injections, clock feedthrough), the modulator implementation uses ideal switches with a fixed on-resistance. The modulator output spectrums of figure 2.28 show the effect of varying the integrators DC gain value: the results are correlated with the model predictions presented in figure 2.25. For a 30 dB DC gain, the modulator reaches only 101.6 dB SQNR while 40 dB DC gain is sufficient to reach the desired SQNR target of 109.5 dB. Figure 2.28 clearly shows that noise shaping is degraded at low frequencies for a low DC gain, corresponding to a higher phase error in integrators transfer function. Figure 2.29 shows the modulator output spectrums of design points A and B. Design point A, taking into account the capacitive effects, reaches the targeted SQNR value, while the design point B shows 15 dB performance degradation, which

was predicted by model results in figure 2.26. These simulations confirm that the presented high-level integrator model provides sufficient accuracy to determine amplifier specifications without affecting the modulator performances.

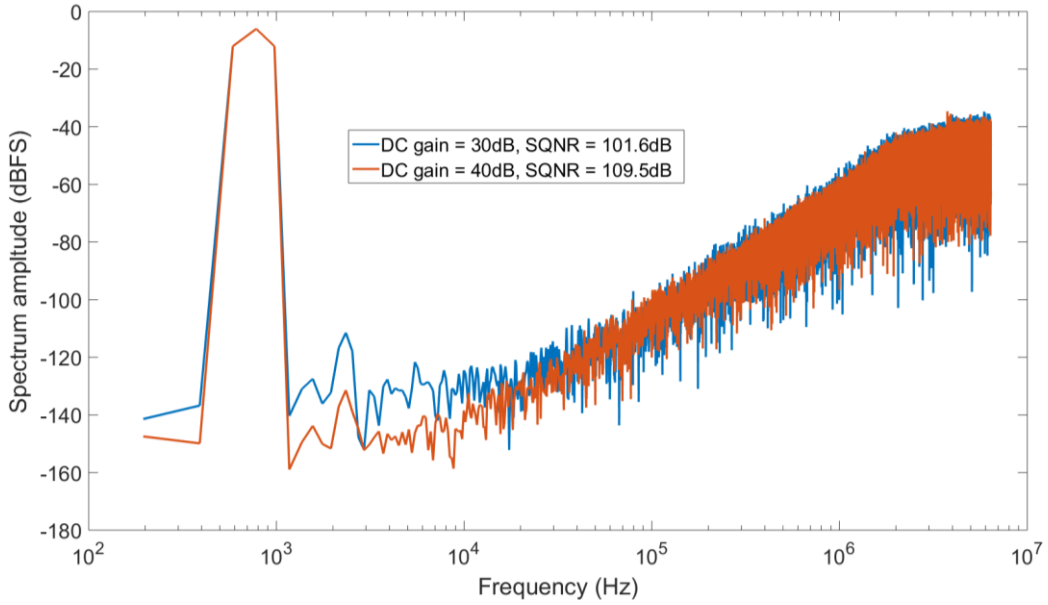


Figure 2.28 Transistor-level modulator output spectrums for different DC gain integrators

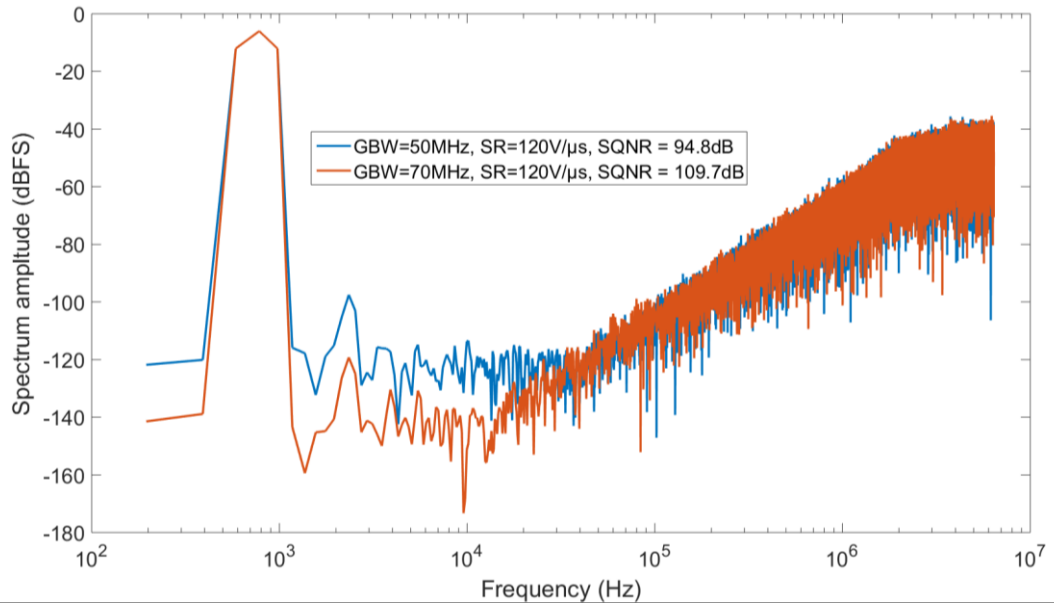


Figure 2.29 Transistor-level modulator output spectrums for different integrators (GBW,SR) couples

2.3 Conclusion

In this chapter, the basics concepts of sigma-delta modulation were introduced to better understand the rest of the manuscript. The main architectures have been presented with different modulator orders. Targeting a low-power modulator, the CIFF structure with input feedforward is the one of choice, since it has the lowest integrators swings, thus allowing area and power saving. For modulators order > 2, stability has been investigated in case of both single-bit and multi-bit quantizer. A high-order single-bit sigma-delta loop must have its noise shaping performance degraded to maintain stability, while a multi-bit one can use aggressive shaping, but losses its feedback DAC linearity and needs Dynamic Element Matching (DEM) techniques to keep modulator resolution. In a second part, a high-level model of a discrete-time integrator has been developed, to take into account the analog non-idealities and allowing the designer to define accurately integrators specifications such as DC gain, GBW or Slew Rate (SR). A high SR like a class-AB amplifier is suited, allowing to reduce the GBW specification and then the bias current, as it has been demonstrated, thus benefiting the power consumption. This model has been used in next chapter to design a 103dB Dynamic Range (DR) modulator and a good correlation has been achieved between high-level model and transistor-level simulations.

Chapter 3

A 103dB DR 2nd order discrete-time sigma-delta modulator

Let's first review several discrete-time modulator designs reaching the state-of-the-art performances. To compare ADCs performance, generally two figures-of-merit (FOMs) are used [4]: the Walden FOM denoted FOM_W and the Schreier FOM denoted FOM_S which definitions are given below.

$$FOM_W = \frac{Power}{2^{ENOB} \times F_{snyquist}} \tag{3.1}$$

$$FOM_S = DR + 10log_{10}\left(\frac{Bandwidth}{Power}\right) \tag{3.2}$$

The FOM_W illustrates the tradeoff power vs speed, which is good to compare ADCs where the speed is the dominant factor affecting the total consumption. But this FOM is not adequate to illustrate the tradeoff power vs resolution since in noise limited ADCs, improving the resolution by two requires a 4 times higher current consumption while at constant FOM_W, adding one bit of resolution traduces a doubled power only. That's why for noise limited ADCs (low-medium speed and medium-high resolution) the FOM_S is preferred: improving the dynamic range by 6dB requires 4 times more power to maintain a constant FOM, which traduces effectively a noise limited power vs resolution tradeoff. This FOM will then be used to compare designs and the state-of-the-art designs have nowadays a FOM_S of around 180dB in audio bandwidth [4].

Several advanced techniques can be used to reach this FOM_S level. In [16], Xu et al. use the switched-opamp technique to reduce the loop filter consumption by switching off the opamps during the sampling phases and switching on them only during the integration phase. Ideally the power consumption can be halved, but switching on/off the opamps requires an extra consumption to recover a normal opamp behavior at startup. This design achieves a 179.3dB FOM_S. In [17], Peña-Perez et al. employ the opamp sharing technique in combination with the noise enhancement presented in [18] to reach a third-order loop filter with a single amplifier. The amplifier is made by a conventional low power second-order OTA which is assisted by a dynamic Slew Rate enhancement circuit making it function as a class AB amplifier to reduce its power consumption [19]. This modulator reaches a FOM_S of 179.5dB. In [20], Chae and Han propose to replace conventional OTAs by inverters. Operating them in class AB or even better in class C amplifier, a high power efficiency is obtained compared to conventional OTAs, since also the number of current branches are reduced to one. This solution reaches a FOM_S of 174.5dB in 2009 and has inspired a large number of design and publications, which leads to the state-of-the-art design of discrete-time sigma-delta modulator presented in [21]. This modulator

uses the zoom concept: by mean of a coarse SAR ADC, the reference level of the feedback DAC of a third-order single-bit modulator is adjusted for each input sample. This allows to reduce the loop filter swings since it emulates a signal tracking feedback DAC, like a multibit quantizer can do. Combining it with low power inverter-based amplifiers, this design reaches a very high DR of 109dB which leads to a 181.5dB FOM_S.

3.1 Modulator specifications

The modulator, for the expected audio analog front-end application defined in the introduction, should have a peak SNR of 100dB. Let's reach this value for a -3dBFS peak input sinewave leading to a 103dB dynamic range requirement. The PGA placed before in the system will adapt the signal range of the microphone to the ADC full scale. The 140nm technology used has a nominal voltage of 1.8V, and from a system point of view, a Low-Drop-Out (LDO) regulator will be used to provide a low-noise reference level. This allows performances robustness even in case of noisy supply voltage. Then the modulator full scale cannot be higher than 1.5V to keep enough margin for the LDO to operate. The targeted FOM_S is 180dB, that gives for a 103dB DR and a 24kHz bandwidth a power budget of 480µW. To maximize the FOM_S, the power consumption should be minimized and then the input noise level should be as high as possible. Since the peak SNR is reached for a -3dBFS peak input sinewave, i.e. -6dBFS rms input sinewave level, and considering the 100dB SNR targeted, the rms total input noise level should be below -106dBFS. Using a 1.5V full scale, this leads to a targeted input-referred noise level of 7.5µVrms. Finally, to make the thermal noise the dominant modulator noise source, let's put a 110dB Signal to Quantization Noise Ratio (SQNR) requirement. The modulator specifications are summarized in table 3.1

Table 3.1 Modulator specifications used to define architecture

Specification	Value	Unit
Input full scale	1.5	V
Signal bandwidth	24	kHz
Power consumption	480	µW
Input-referred noise	7.5	µV
Dynamic Range (DR)	103	dB
Signal to Noise Ratio (SNR)	100	dB
Signal to Noise and Distortion Ratio (SNDR)	>95	dB
Signal to Quantization Noise Ratio (SQNR)	110	dB
FOMs	180	dB

The selected modulator architecture is a second-order modulator with a 1.5-bit quantizer and an OSR of 256, providing a SQNR of 110dB. It uses a CIFF structure with input feedforward for its relaxed loop filter linearity requirement and is presented on figure 3.1. A 1.5-bit quantizer is selected to reach the targeted SQNR instead of increasing the loop filter order or the OSR because it helps to reduce integrators output swings as a multibit quantizer does without a complicated DEM mechanism which could cost power [22].

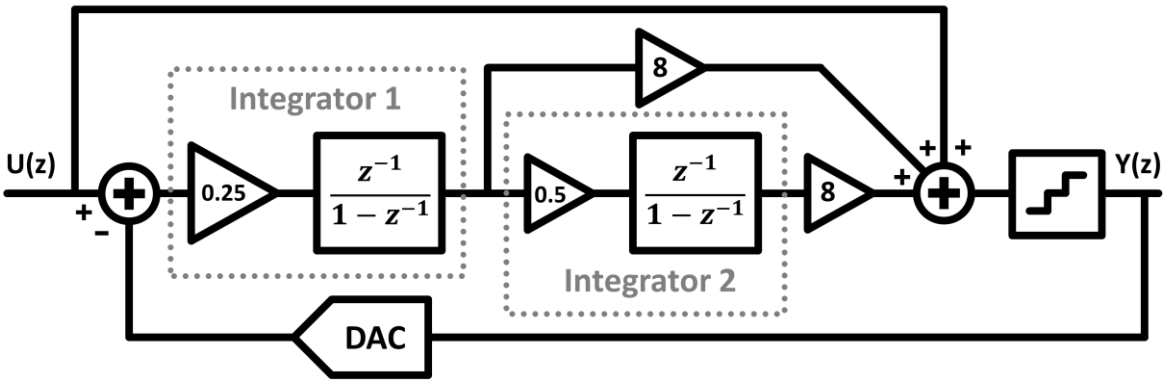


Figure 3.1 Second order 1.5-bit modulator linear model and coefficient scaling

A large OSR is chosen for the low area occupation it implies, which is important for a low-cost analog front-end application and the fact that it does not lead to power penalty, at least at first order: [21] proves that a high OSR can be used in a high power efficiency design. The drawback is the decimation filter which have to run at a high frequency too at least for its first stage. The justification is that for a giving noise budget, the input referred noise due to the sampling capacitors is proportional to:

$$V_{n,C_S} \propto \frac{kT}{OSR \times C_S} \tag{3.3}$$

Where k is the Boltzmann constant, T the temperature and C_S the sampling capacitance. Hence, increasing twice the OSR allows to halve the sampling capacitance, reducing the modulator area occupation. Concerning the power penalty, the required GBW product of a switched-capacitor integrator is dependent of its load capacitance. At first-order, it can be approximated by the sampling capacitance, since the effective load is the series of the integration and sampling capacitances giving the following relation:

$$GBW \cong \frac{gm}{2\pi \times C_S} \tag{3.4}$$

Hence, doubling the OSR requires twice the GBW in the integrators. Since the sampling capacitance is divided by 2, it comes without changing the gm value, i.e. the power consumption of the integrators.

3.1.1 First integrator sizing

Equation 3.3 can be extended to size, at first order, the sampling capacitances. Considering that (3.3) is the in-band noise power of a single sampling capacitance, a factor two is needed for a differential operation. Another factor of 2 is needed because during a clock period, two samplings on C_S capacitor occurs: one at the end of the sampling phase, and another one at the end of the integration phase, which directly impacts the noise transferred in the integration capacitor. This lead to the following:

$$V_{n,C_S} = \frac{4kT}{OSR \times C_S} \tag{3.5}$$

Allowing half the noise power budget to the sampling capacitances – the larger the value is the lower the occupied area will be – it comes:

$$C_S = \frac{4kT}{OSR \times \frac{(7.5e^{-6})^2}{2}} = 2.3\ pF \tag{3.6}$$

Since the ratio of C_S/C_I for the first integrator has been defined to 1/4 (figure 3.1) to limit the integrator output swing at 0.4V peak differential, it comes that C_I = 9.2pF. Using behavioral simulations, and considering 2µVrms noise for the DAC voltage reference and 5.3µVrms noise for the sampling capacitances, it can be deduced that the contribution of the first opamp to the total modulator input-referred noise level must not exceed 3µVrms to meet the required total input-referred noise of 7.5µVrms.

The non-ideal integrator model developed in chapter 2 is used to define the first integrator requirements. All curves given in section 2 of chapter 2 are for the first integrator of this modulator. It can then be concluded that the required DC gain is 40dB, and the optimal couple of GBW and SR value is 70MHz for the GBW and 120V/µs for the SR. The specifications deduced for the first integrator are summarized in table 3.2.

Table 3.2 First integrator specifications

Specification	Value	Unit
Sampling capacitance (C_S)	2.3	pF
Integration capacitance (C_I)	9.2	pF
DC gain	40	dB
Gain Bandwidth product (GBW)	70	MHz
Slew rate (SR)	120	V/µs
Input-referred opamp noise	3	µVrms
Output swing differential	0.4	V

3.1.2 Second integrator sizing

The input-referred noise of the second integrator is first-order noise shaped. Using behavioral simulations, the in-band noise can be as high as 40µVrms without impacting the total modulator noise level. Using (3.5) and allowing half of the noise contribution to the sampling capacitances, the latter size can be determined:

$$C_{S2} = \frac{4kT}{OSR \times \frac{(40e^{-6})^2}{2}} = 80\ fF \tag{3.7}$$

For matching considerations and in order to have enough differentiation between sampling and integration capacitances and parasitic ones, C_{S2} is chosen larger than the minimum required and is set to 150fF. As shown on figure 3.1, the second integrator gain is defined to 1/2 to limit its output swing to 0.3V differential; then the integration capacitor is C_{I2} = 300 fF. Considering a 150fF integration capacitance, it leads to a sampling input-referred noise of 20µVrms using (3.5). Since 40µVrms is allowed, the contribution of the opamp to the second integration input-referred noise can be up to 34µVrms.

The non-ideal integrator model in section 2 of chapter 2 is also used to define the second integrator specifications. The modulator SQNR evolution in function of the second integrator DC gain is shown on figure 3.2. As visible, a DC gain of 50dB is needed to ensure a 110dB

SQNR for the modulator. Figure 3.3 shows the required couple of GBW and SR values for the second integrator to meet the 110dB SQNR modulator requirement. The selected design point is a 50MHz GBW and 7V/µs SR. Second integrator specifications are summarized in table 3.3.

Table 3.3 Second integrator specifications

Specification	Value	Unit
Sampling capacitance (C_{S2})	150	fF
Integration capacitance (C_{I2})	300	fF
DC gain	50	dB
Gain Bandwidth product (GBW)	50	MHz
Slew rate (SR)	7	V/µs
Input-referred opamp noise	34	µVrms
Output swing differential	0.3	V

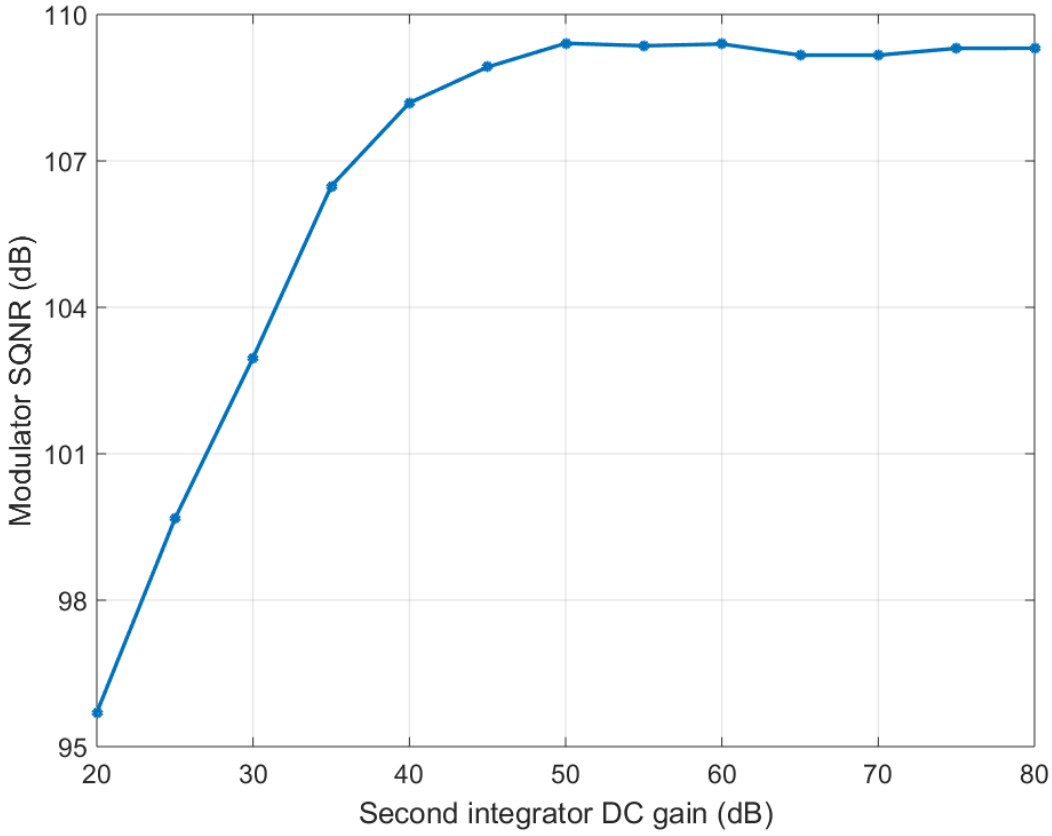


Figure 3.2 Modulator SQNR in function of second integrator DC gain

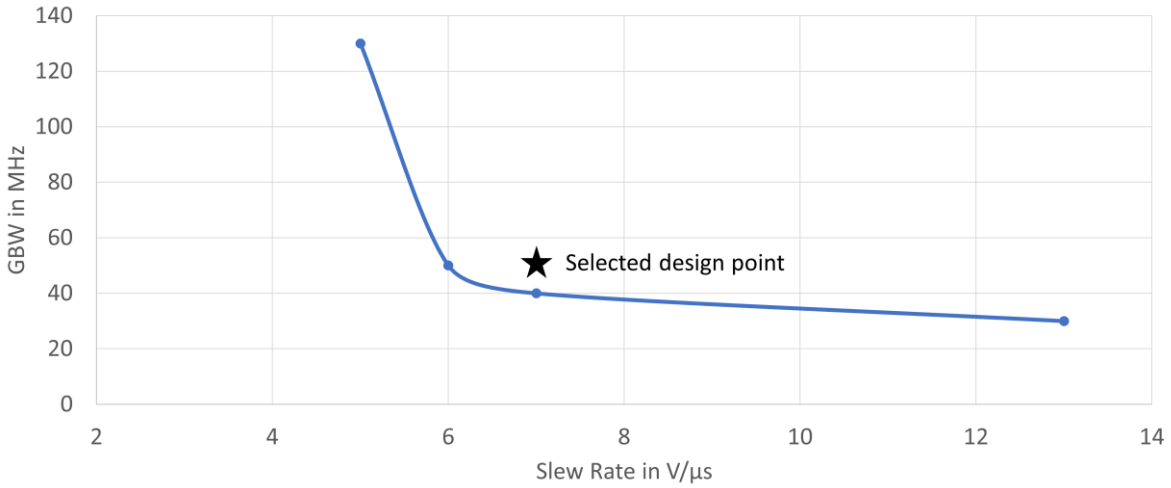


Figure 3.3 Minimum required GBW in function of SR value for second integrator to not degrade modulator performance

3.2 Integrators design

With their dynamic properties like doubled transconductance and class-AB behavior, inverters-based amplifiers are well-suited for switched-capacitor circuits as seen in [20, 21]. But inverters also have drawbacks that must be taken into account: their sensitivity to Process, Voltage and Temperature (PVT) makes their biasing point not stable, leading to an uncontrolled current consumption and common-mode voltage [20, 21, 23]. These aspects are important for applications where the low-power consumption is one of the key requirements and robustness is necessary for mass production. Moreover, the control of common-mode voltage is needed to interface with others building blocks of the analog front-end system.

Several attempts to circumvent these two issues can be found in literature. A voltage regulator can be used to control the supply voltage of the inverter, ensuring a constant current consumption [23]. But the supply voltage is often used as the reference voltage for the sigma-delta modulator, which can be problematic in sensing use-case since the absolute gain of the modulator can vary with the temperature for example. A dynamic biasing scheme can be adopted in switched-capacitor circuits: the sampling phase of the discrete-time integrator is used to bias the inverter in combination with an auto-zeroing technique [21]. The accuracy is however limited by mismatch and the auto-zeroing technique adds an extra capacitive load. Several structures have been presented to set the common-mode of pseudo-differential inverter-based amplifiers around half the supply voltage by using extra inverters [24, 25], but without a precise and reliable value. An interesting solution is to add biasing transistors operating in linear region [26]. In that case the current can be controlled and the common-mode too, without losing the class AB operation (this will be further developed in next section). But the selected self-biasing, which is the best solution for low-power, cannot totally control the sensitivity to PVT variations. This self-biased inverter-based amplifier [26] is presented in figure 3.4. The main idea is to use the inverter advantages (class AB operation, doubled transconductance, large dynamic swing, low-voltage capabilities) in an OTA architecture without inverter weaknesses

(control of bias current, sensitivity to PVT variations, regulation of common-mode). To do it, linear-biased transistors are added to set the biasing point of the amplifier.

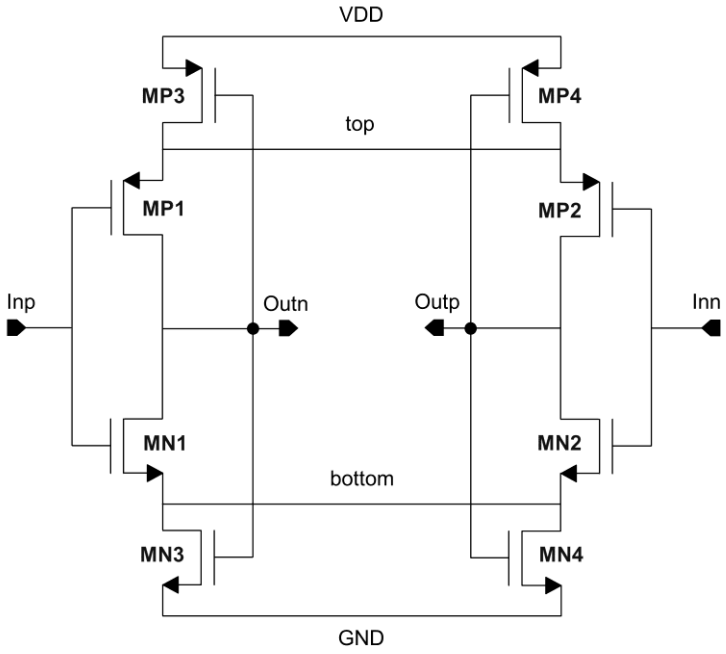


Figure 3.4 Self-biased inverter-based amplifier

The linear-biasing transistors combined to inverters allow a class-AB-like response when a large drive is applied on amplifier inputs. This is especially well suited for switched-capacitors circuits where large drive is present at the beginning of each phase. There are two reasons for a slew rate increase that will be explained in the case of an integrator application which is one of the most used common blocks in switched-capacitor circuits. First, large input glitches at the beginning of the integration phase (when integration switches close) generate large drive at the outputs of the inverters (nodes Outp and Outn). This drive increases the gate-source voltage of MN3,4 and MP3,4 (figure 3.4), which will provide more bias current. Second, still at the starting time of the integration phase, glitches on inputs make the inverters transistors in an opposite state: let's take MN1 and MP2 linear while MP1 and MN2 are off (figure 3.4). The linear ones (MN1 and MP2) will make the PMOS source node (called top on figure 3.4) to decrease and the NMOS source node (called bottom) to increase. This will increase the drain-source voltage of transistors MN3,4 and MP3,4; as they are in linear state, their current will also increase leading to a higher biasing current during the transient operation. Combination of these two effects allows the inverter-based amplifier to reach a peak current four times higher than its biasing current making this amplifier well suited for switched-capacitors circuits.

The self-biased amplifier is less sensitive to Process, Voltage or Temperature (PVT) variations than simple inverter but these variations still have a large impact. The biasing current varies from less than half to more than twice the typical value (as will be seen in simulation results) which not only affects the consumption but also the dynamic performances like the Gain-Bandwidth product (GBW) which follows the same variation range. The common-mode value is also affected with PVT variations, and can vary more than 100mV from its typical value of half the supply voltage. This large common-mode variation is not acceptable as input for the next inverter-based stages following the integrator and must be reduced by a dedicated solution.

The proposed solution (figure 3.5) keeps the linear-biased transistors (MP5, MN5) to control the inverters' biasing point (to keep the class AB response), and uses them in a current-mirror way (with MP4 and MN4) to control effectively the current consumption, lowering in the same time the spread of several specifications over PVT like the GBW product.

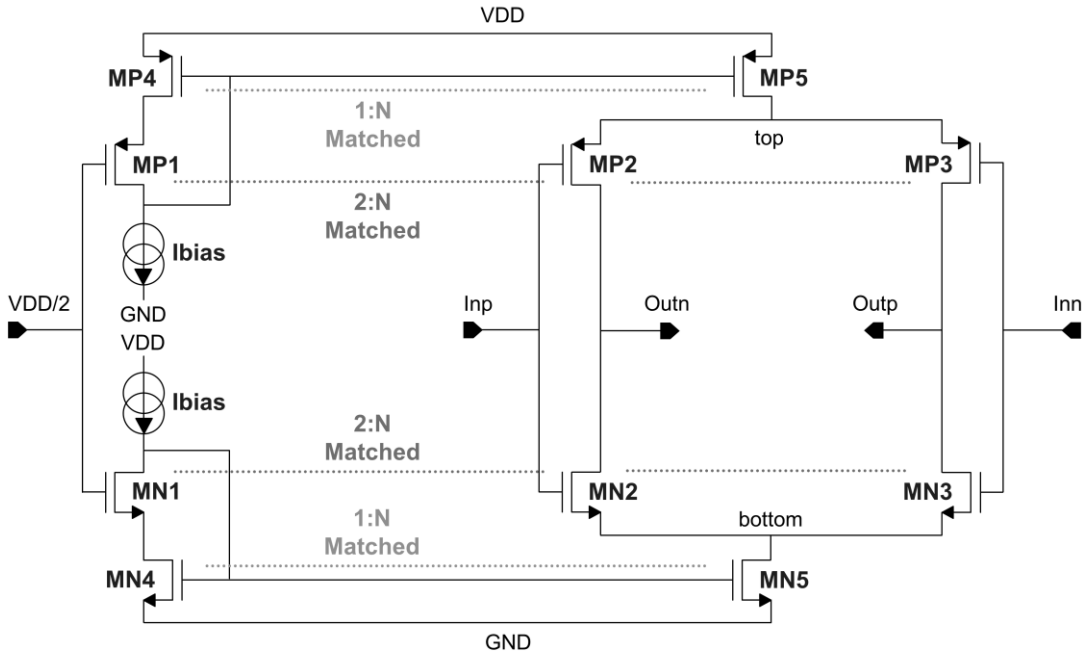


Figure 3.5 First biasing attempt using linear transistors to mirror current

To mirror accurately a current with linear-biased transistors, their drain-source voltages must be equal (not only their gate-source voltage); in that case the current mirror ratio depends only on transistor sizes. To ensure the same drain-source voltage, MN1 and MP1 transistors are added and act as source followers. Since they are matched with the inverters transistors, it should result that MN4-5 and MP4-5 share the same drain-source voltages. A fixed bias current is flowing through the source followers, providing the wanted biasing point for the amplifier. Nevertheless, the accuracy of this biasing solution is not sufficient: even if they are matched, MN1 and MP1 do not have the same drain-source voltage than MN2,3 and MP2,3. This results in a slight deviation in their gate-source voltage which makes the linear-biased transistors (MN4-5 and MP4-5) in an unmatched environment, their drain-source voltage is different and so the accuracy of the current copy is not ensured enough, especially over severe PVT conditions.

To reproduce more accurately the voltages of top and bottom nodes in the biasing stage, the complete structure of the amplifier is repeated with a ratio 1:10 in transistors sizes and the circuit is presented on figure 3.6. MN2 and MP2 form a matched inverter with the amplifier stage, it helps to reproduce in the biasing stage the same source voltage than bottom and top nodes in the amplifier stage. This lowers significantly the effect of MN1-MP1 drain-source voltage sensitivity and consequently reduces the spread of the biasing error over PVT conditions. The accuracy of this biasing solution is now mainly limited by the inverters mismatch between the biasing and amplifier stages.

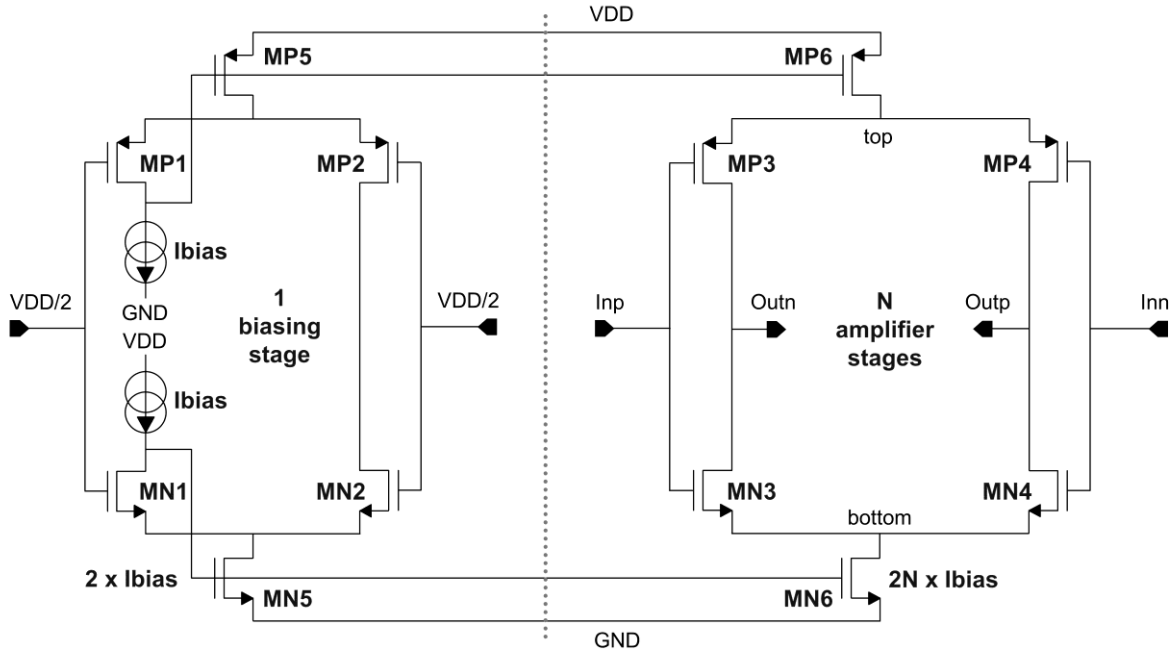


Figure 3.6 Proposed biasing solution using linear transistors to mirror current

The threshold voltage of both NMOS and PMOS transistors of the inverters is a key design parameter. It defines the gate-source voltage of the transistors and it allows to fix the biasing voltage of top and bottom nodes (figure 3.6). The closer these nodes are to the supply and ground lines, the larger the amplifier output swing will be. For this reason, low-leakage transistors (higher threshold voltage) are used in the inverters to maximize the output swing. Moreover, by determining top and bottom nodes which are about $V_{DD}/2 + |V_{thp}|$ and $V_{DD}/2 - V_{thn}$ respectively, the threshold voltage ensures the linear-bias of the current mirror transistors. This is crucial to benefit for the class-AB like response in switched-capacitor circuits as explained previously. Depending on the temperature and process variations, V_{th} of inverters has large variations (several hundreds of mV), and so varies top and bottom nodes (figure 3.6), making in worst case these nodes equal to respectively the supply and ground lines; and in that case the amplifier is not operating correctly anymore. To limit these variations, these threshold voltages are controlled by changing the bulk voltage since [27]:

$$V_{th} = V_{t0} + \gamma\left(\sqrt{|2\varphi_f + V_{source} - V_{bulk}|} - \sqrt{|2\varphi_f|}\right) \tag{3.8}$$

where V_{th} is the threshold voltage, V_{t0} the threshold voltage with a zero source-bulk voltage, γ the body effect parameter and $2\varphi_f$ the surface potential. If V_{t0} increases, the bulk voltage must increase to reduce or reverse the source-bulk voltage and then compensates the V_{th} variation (3.8). This is done with the circuit on figure 3.7. If the threshold voltage of NMOS transistor increases, lower current will flow through NMOS transistor and its drain node will increase. The inverse is done if threshold voltage decreases. The equivalent operation is done on upper side to provide PMOS bulk bias voltage. The bulk voltage regulation of the inverters allows a drain-source voltage of MN5-6 and MP5-6 (figure 3.6) comprised between 50 and 200mV for all process corners, for a temperature range of -40°C to 125°C and a supply voltage range of 1.4V to 1.6V, thus ensuring the linear operation. The process used is a triple-well technology,

keeping the substrate to the lowest voltage and then keeping the circuit out of latch-up risk: the bulk biasing is made only in separated wells containing the inverters' transistors. The power consumption of the circuit of figure 3.7 can be made negligible, by tuning the resistors values (increasing) and the transistors sizes (increasing length), it ends up with about 1µA current consumption in this branch.

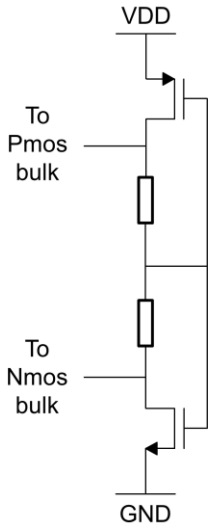


Figure 3.7 Bulk voltage generation circuit used to control the threshold voltage of inverters

The common-mode regulation of amplifiers using simple inverters requires complex feedback systems [24, 25]. It is composed of several other inverters leading to an increase of power consumption, and the common-mode value is not well defined, still sensitive to PVT variations. In our approach, the biasing stage provides two biasing voltages for the amplifier. One of them can be used in a low-power switched-capacitor circuit which compares the output common-mode level with a reference value and adapts the biasing voltage provided to the amplifier [28]. The circuit is presented on figure 3.8:

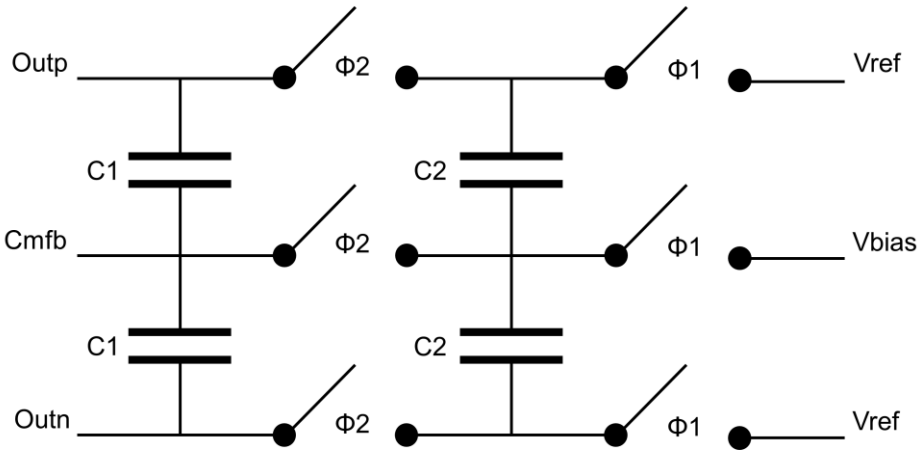


Figure 3.8 Switched-capacitor common-mode feedback circuit

On phase Φ_1, capacitors C_1 sense the output common-mode and C_2 are precharged with a voltage $V_{ref} - V_{bias}$, where V_{ref} is the wanted output common-mode value and V_{bias} is the

desired bias voltage. On phase Φ_2, C_{mfb} node is refreshed to adapt the bias voltage in function of the output common-mode level. The charge conservation law between phases Φ_1 and Φ_2 gives the following equation:

$$C_{mfb} = \frac{O_{utp} + O_{utn}}{2} - \frac{C_2 z^{-1/2}}{C_1 + C_2 - C_1 z^{-1/2}} \left(V_{ref} - V_{bias}\right) \tag{3.9}$$

Since V_{ref} and V_{bias} are two DC values, we can consider the limit $z \to 1$ and the equation can be rewritten as:

$$C_{mfb} = V_{bias} + \frac{O_{utp} + O_{utn}}{2} - V_{ref} \tag{3.10}$$

Then, C_{mfb} node is equal to V_{bias} node value which is added the difference between the common-mode level and its desired value. If the output common mode increases, so does the C_{mfb} node so any biasing point of the amplifier with a negative feedback action on the common mode value can then be used for the common-mode regulation with this switched-capacitor circuit. Both biasing voltages from the biasing stage to the amplifier stage provide a negative feedback on the output common-mode, we choose to place the regulation circuit in the NMOS current mirror for their lower parasitic capacitances compared to the PMOS side as shown on figure 3.9. The common-mode is then regulated with a better accuracy to $V_{DD}/2$ and with less sensitivity to PVT.

It has previously been explained how large input signals in switched-capacitors applications allow the self-biased amplifier to provide higher output current than its quiescent one. This is due to two factors: the variations of gate-source voltage of the linear-biased transistors which increase the output current and the variations of top and bottom nodes which also create higher current in linear-biasing transistors. With the new introduced biasing scheme, only the last of these two factors contributes to a slew rate increase.

To further extend the slew rate performance, the bias stage is made dynamic, with an increased bias current when amplifier needs. To make this, the main idea is to sense these variations of top and bottom nodes of the amplifier stage and use them to increase the current of the biasing stage. In that case, the biasing stage current will only increase when needed (in a transient operation), adding another increase of current in the amplifier stage.

The implementation of this function is presented in figure 3.9, and will be described for the NMOS side, but the same applies for the PMOS side. A diode-connected NMOS (MN2 on figure 3.9) senses the variations of bottom node and reports it on its gate voltage. Since this transistor is matched with the ones of the amplifier, its gate/drain node is biased at $V_{DD}/2$. This node is then used to drive the gate of the inverter present in biasing stage (MN1 on Figure 3.9): its biasing point is the same ($V_{DD}/2$), but in case of transient response the gate node will increase and a higher current will flow in the biasing stage, leading the current mirror to increase the current of the amplifier. Again, two current sources are needed to supply this transient-boosting-circuit, shown as ideal on figure 3.9.

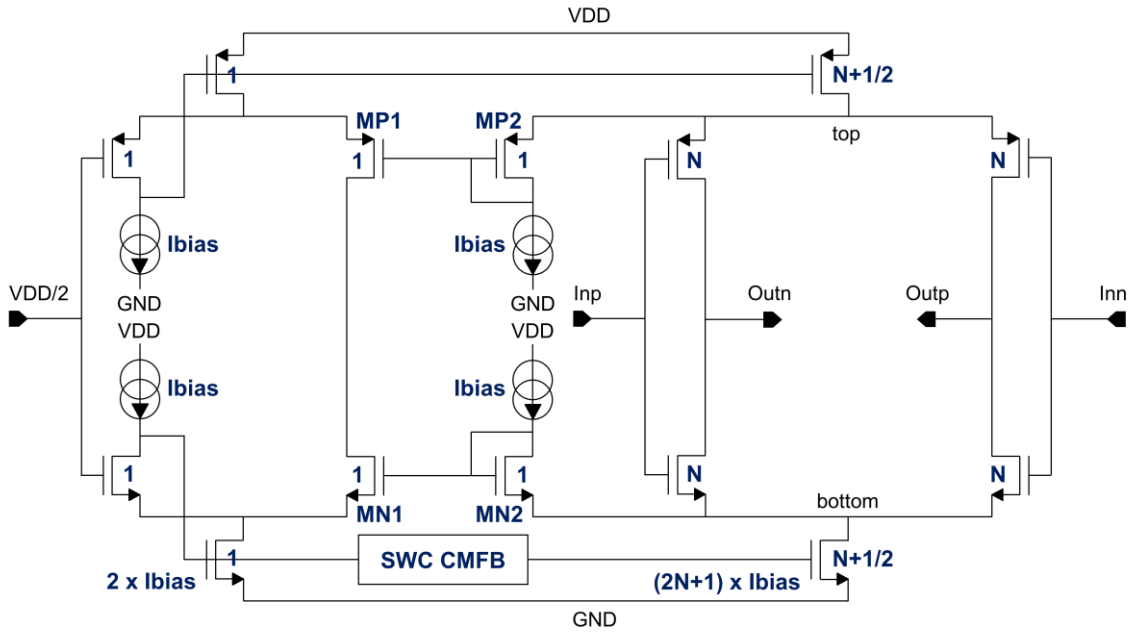


Figure 3.9 Proposed amplifier with slew rate enhancement circuit

Table 3.4 summarizes the error values of 100 points Monte Carlo simulations of process and mismatch variations for the self-biased and proposed amplifiers, with V_{DD} = 1.4V to 1.6V and temperature varying from -40°C to 125°C. The bias current of the differential amplifier stage is kept constant at 100µA between the two circuits. The amplifier consumption of circuit presented in figure 3.9 is 115µA including the biasing stage (10µA for biasing and 5µA for the slew rate enhancement circuit). A current biasing circuit is necessary to provide the current sources and consumes 15µA more, adding 1µA from the V_{th}-control circuit it leads to a 131µA total consumption for the amplifier. As we can see in Table 3.4, the spread over PVT and mismatch conditions has been largely reduced for the bias current, which was more than +100% in the self-biased circuit down to +/– 20% in our proposed solution, validating the developed biasing concept. The common-mode is now regulated to half the supply voltage with a maximum deviation less than 50mV while the self-biased inverter-based amplifier (figure 3.4) can have about 200mV deviation. Another approach to solve the PVT variations issue is the replica biasing: the current passing through a replica inverter is measured and the supply voltage of inverters is adjusted to have the desired bias current [29]. This solution can reach even better accuracy (error can be less than 10%) at the cost of a complex and power-hungry biasing system (1.7mW). In [29], this biasing system is used to provide reference signals to a massive number of parallel ADCs (1000) that's why its power consumption is negligible, but in the case of only one ADC the interest is more limited.

Table 3.4 Monte Carlo simulation results for self-biased and proposed inverter-based amplifiers

Parameter	Self-biased inverter			Proposed amplifier			[29]		
	min	typ	max	min	typ	max	min	typ	max
bias current (µA)	62	100	239	94	115	137	13.5	14.7	15.4
bias current error (%)	-38	0	139	-18.26	0.00	19.13	-8.16	0.00	4.76
common mode error (mV)	-185	0	99	-40	0	34	X	X	X

In figure 3.10, the differential output current of the amplifier is plotted in function of time for amplifier without (figure 3.6) and with (figure 3.9) the slew rate enhancement circuit. As we can see, circuit of figure 3.9 has a higher output current with a pronounced bounce during the first part of the integration phase, where the amplifier is slew rate limited. The slew rate enhancement circuit allows to reduce the settling time of the integrator output by 20% while costing only 5% of extra power consumption. This is an important benefit since the higher the slew rate is, the lower the required GBW product is for a fixed settling error as seen in section 2 of chapter 2. The reduction of GBW requirement allows to reduce the idle current of the amplifier and then reduces its power consumption.

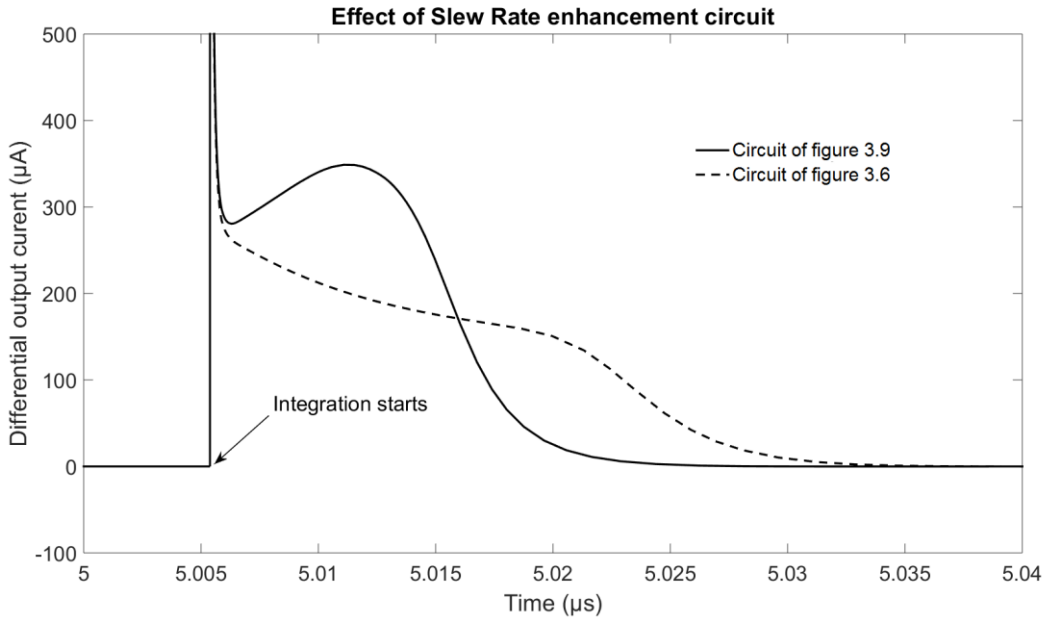


Figure 3.10 Comparison of output currents of figure 3.6 and figure 3.9 circuits

In figure 3.11, the NMOS bulk voltage generated by the circuit shown in figure 3.7 is plotted versus the temperature and process, the two most influent factors on the threshold voltage of inverters. As we can see, slow processes which have a higher threshold voltage are biased with a higher bulk voltage to compensate the threshold deviation from nominal value (as explained previously) and the opposite applies for fast processes. When temperature increases, the threshold voltage of transistors decreases. By decreasing the bulk voltage with the circuit of figure 3.7, this variation of threshold voltage can be reduced too. With this bulk regulation, the threshold voltage spread over PVT conditions of inverters goes down from about 300mV to 100mV. The bulk voltage can be as high as 500mV in worst case, and in that case the source voltage of NMOS transistor is about 100mV, leading to a 400mV forward biasing on the diode which is still enough away from the about 600mV diode forward biasing voltage. To check this, simulations of bulk leakage current have been done for the NMOS and PMOS transistors over PVT conditions and result in maximum total leakage currents of 8.6nA on PMOS side and 3.4nA on NMOS side.

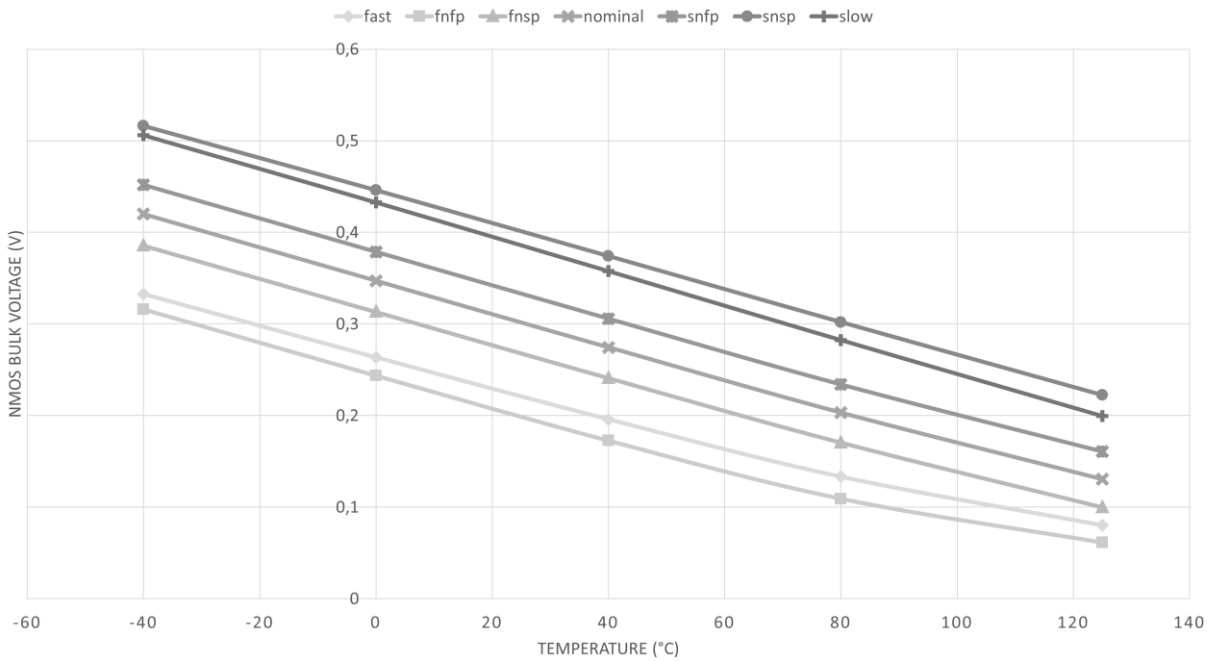


Figure 3.11 NMOS bulk voltage versus process and temperature generated by figure 3.7 circuit

Figure 3.12 shows the amplifier response during the integration phase over PVT conditions (all process corners, V_{DD} = 1.4V to 1.6V, Temperature = -40°C to 125°C) for A) output voltage and B) output current. The sampling frequency is 12.288MHz, and the end of the integration phase is not plotted to better zoom on the settling phase. Only the typical and the extremum curves are plotted on the graphs for sake of simplicity. As we can see, the fast process at 125°C with a 1.6V supply voltage have the longest settling time. This is due to the fact that the threshold voltage of inverter transistors for this process corner at this temperature is the lowest. In that case, the source node of inverter transistors (called top and bottom on figure 3.9) move away from supply/ground lines, leading to a higher drain-source voltage on linear-biased transistors (about 150mV). These transistors are now close to the linear/saturation limit region, limiting the peak current delivered to the amplifier. Oppositely, the slow corner at -40°C with 1.4V supply voltage has the shortest settling time. Since the linear-biased transistors have in that case the lowest drain-source voltage (about 50mV), the class-AB behavior due to transient spikes at the starting of the integration phase is even more amplified. These conclusions are also visible on figure 3.12 B) where the slow process reaches the highest output current. In all cases, the bump characteristic of the slew rate enhancement circuit is visible.

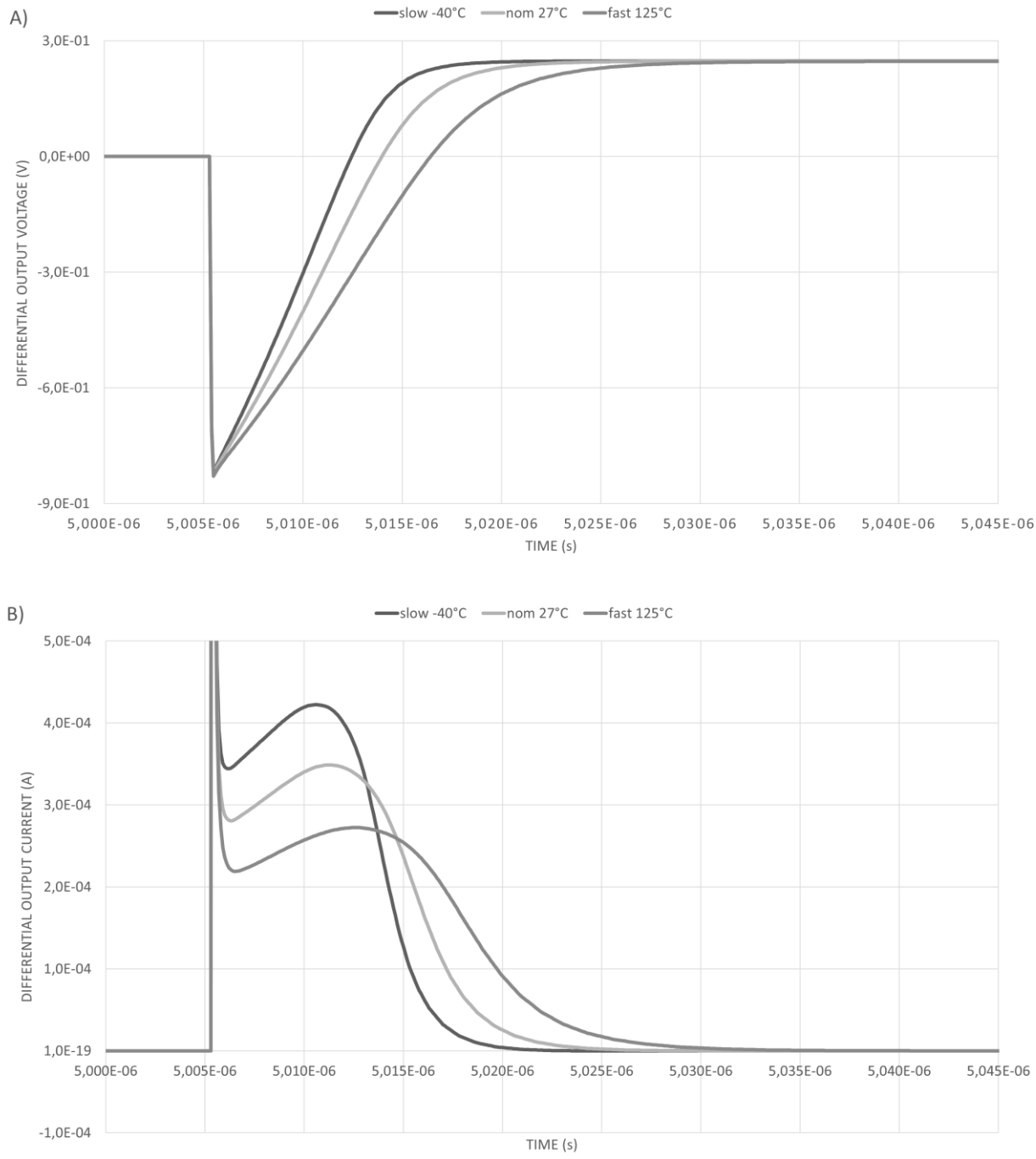


Figure 3.12 PVT simulation results during the integration phase of A) output voltage B) output current

Concerning the integrator noise, several solutions can be used to reduce or remove the flicker noise like the Correlated Double Sampling (CDS) technique or the chopper modulation [30]. The correlated double sampling is used in [20, 21, 23] following the structure presented by Nagaraj et al. in [31] and shown on figure 3.13. The principle is to sample the amplifier offset during the sampling phase into the compensation capacitor C_C and to remove it during the integration phase. In oversampled systems like sigma-delta modulators, the sampling frequency is quite high and the flicker noise value does not change much between the sampling and the

integration phases, that's why this technique allows also to remove a large part of the flicker noise. Another benefit shown in [31] is the integrator reduced finite gain sensitivity, allowing to use low-gain amplifiers such as inverters without impacting the modulator resolution.

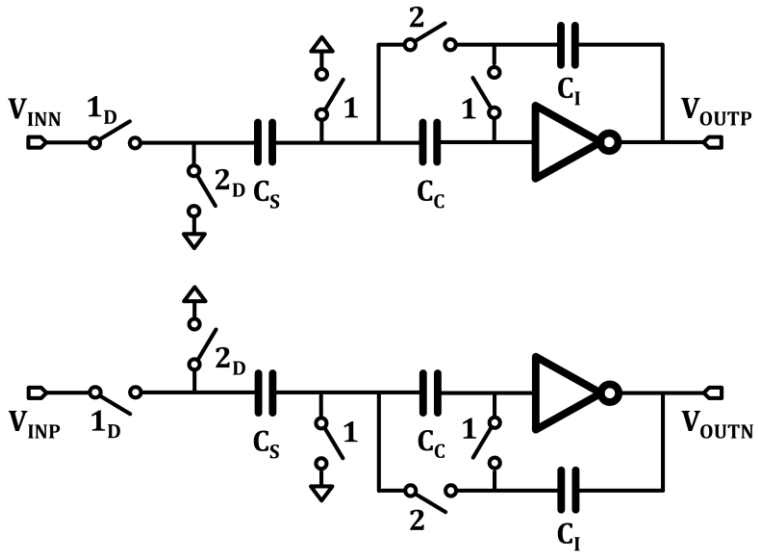


Figure 3.13 Nagaraj integrator using inverters as amplifier

Despite its apparent benefit and simplicity, this solution has also several drawbacks. First, the drain-gate capacitance of the inverter directly couples its input to its output, lowering its effective gain and thus degrading performance. In [21], inverters are cascoded to eliminate this coupling while in [23] cross coupled capacitances are added to reduce the effective parasitic capacitance. Second, the input capacitance of the inverters degrades the effectiveness of the flicker noise removal operation as it is demonstrated in [32]. The flicker noise is not completely eliminated and some white noise is folded in-band, that's why in this design CDS is not used.

The chopper modulation principle is explained with the help of figure 3.14. The baseband signal is first modulated to a frequency F_{CHOP} with a square-wave signal before going through the amplifier. The amplifier offset and flicker noise are then added in baseband without impacting the signal modulated at F_{CHOP} as visible on figure 3.14. At the amplifier output, signal is demodulated into baseband while in the same time, this operation modulates the amplifier offset and flicker noise around F_{CHOP}. This operation does not remove the offset and flicker noises but only modulate them in frequency so they don't appear in signal band.

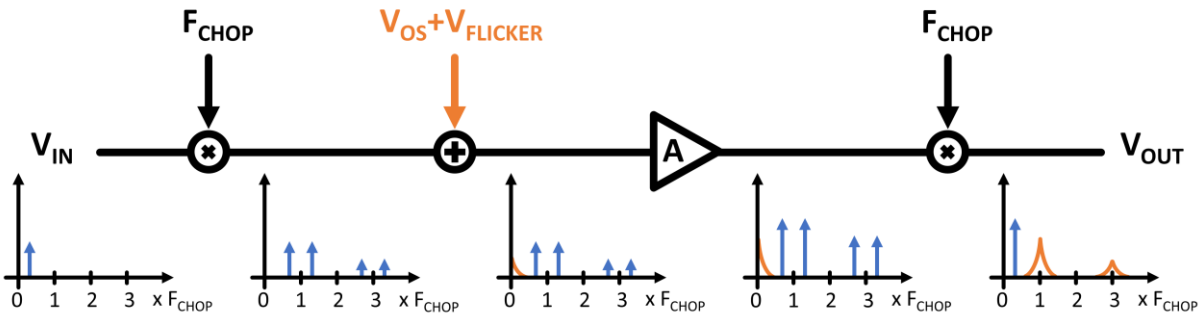


Figure 3.14 Chopper modulation principle

Using the inverter-based amplifier, chopper modulation correctly removes the low frequency noise and amplifier offset, but embedded in a sigma-delta loop degrades the modulator performances. The noise shaping seen is no more of second order but a first order. This is due to the parasitic capacitances C_{DG} of the inverter that directly connect its input to its output. This capacitor is in parallel of the integration capacitance C_I as visible on figure 3.15. When chopper phases alternate, the charge in C_I will be affected by the charge in C_{DG}, which was previously connected to the opposite integration capacitor and so charged to the opposite voltage value. The voltage in C_I then drops by a factor $\Delta = 2 \times C_{DG}/C_I$. This effect is further investigated by Leger et al. in [33] and is similar to a pole error in the integrator transfer function due to finite DC gain. To counteract this effect, a cascode topology is employed in the inverter-based amplifier, since sufficient voltage headroom is available regarding the low output swing needed (0.3V differential) and the supply voltage (1.5V).

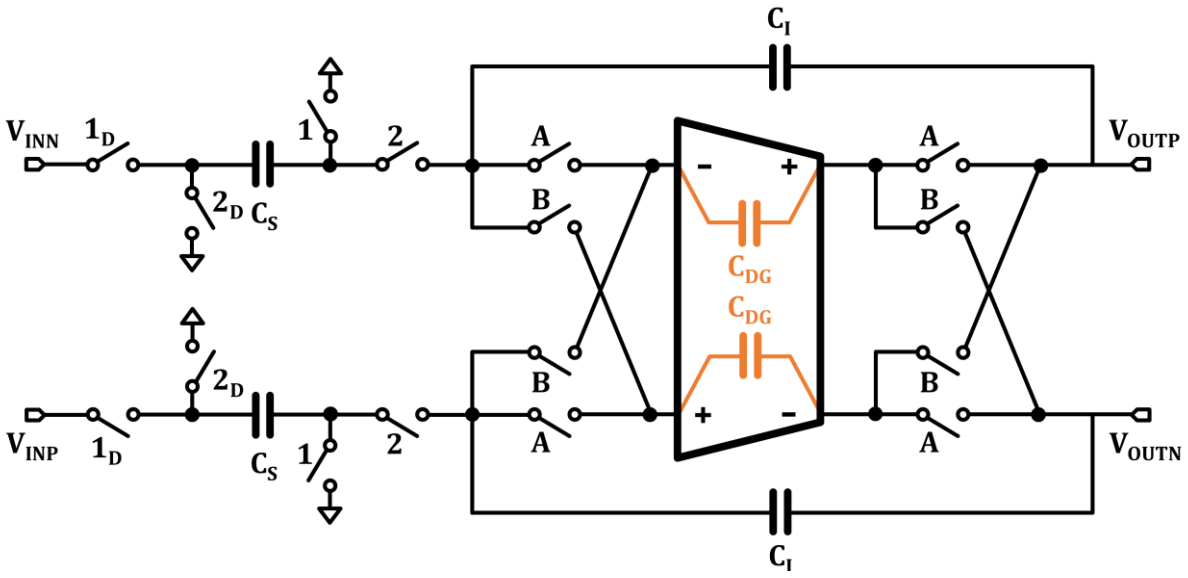


Figure 3.15 Switched-capacitor integrator with chopper modulation and parasitic capacitance C_{DG}

The main feedback DAC does not use dedicated capacitors, it instead reuses the sampling ones, which are no more connected to the common-mode level on the integration phase but on the DAC reference value, according to the data as illustrated on figure 3.16 (single-ended version).

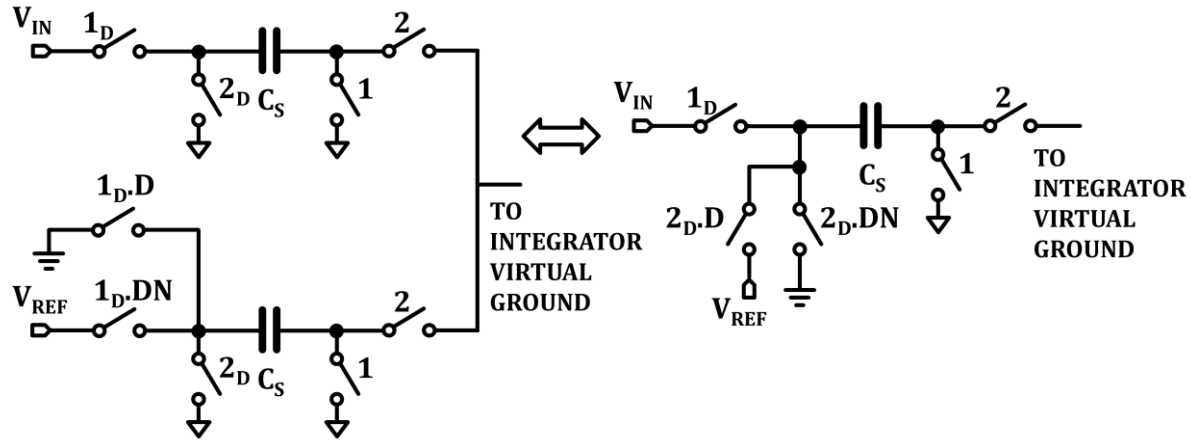


Figure 3.16 Main feedback DAC reusing the sampling capacitance concept

The advantage of this implementation is the lower input-referred noise since half of the capacitance is present compared to a dedicated DAC solution, saving area in the same time. The input sampling and virtual ground node switches are made by CMOS transmission gates, low-resistive enough thanks to the 1.5V supply voltage used. Special attention is made for the input switches size where a trade-off between low on-resistance and parasitic capacitance is made. Delayed clocks are used to avoid charge injections and are denoted with a subscript $_D$ like 1_D for the delayed sampling phase. No bootstrapped switches are needed, saving power and area. DAC connection switches use simple PMOS (for V_{REF}) or NMOS (for ground) transistors.

As said before, the quantizer is a 1.5-bit version, which can output three levels: -1, 0, or 1. To interface with the decimator, the quantizer output is coded using a 2's complement representation: code 11 for data -1, code 00 for data 0 and code 01 for data 1. The feedback DAC must use two capacitors to output the ground, common-mode and V_{REF} levels (corresponding respectively to quantizer levels -1, 0 and 1). Since the feedback DAC reuses the sampling capacitance, the latter is split in two as shown on figure 3.17 (single-ended). During the sampling phase, both capacitances are connected to the input signal. On the integration phase, if the feedback DAC must output data -1 or 1, both capacitances are connected to ground or V_{REF} and work as a normal behavior. But when the feedback must output data 0, one capacitance is connected to ground while the other is connected to V_{REF}. To avoid DAC nonlinearity issue, the data 0 alternates the capacitances connections to ground and V_{REF} as in [34]. The DAC driver decoding logic and DEM is also shown on figure 3.17 and its truth table in table 3.5. The 2's complement output of the quantizer is a 2-bit signal with the MSB representing the sign (0 for positive, 1 for negative) and the LSB the data value, this convention is used in table 3.5 and figure 3.17. The DEM signal alternates each time the data is 0 and is simply generated using a flip flop memory element and a NXOR gate. The switches control signals follow the table 3.5 and are synchronized with the integration phase clock (denoted 2 on figure 3.17) to drive the DAC.

Table 3.5 Truth table of tri-level DAC switches logic

DEM	SIGN	DATA	S_1	S_2	S_3	S_4
X	**0**	**1**	**0**	**0**	**1**	**1**
X	**1**	**1**	**1**	**1**	**0**	**0**
0	**0**	**0**	**1**	**0**	**0**	**1**
1	**0**	**0**	**0**	**1**	**1**	**0**
X	**1**	**0**	**X**	**X**	**X**	**X**

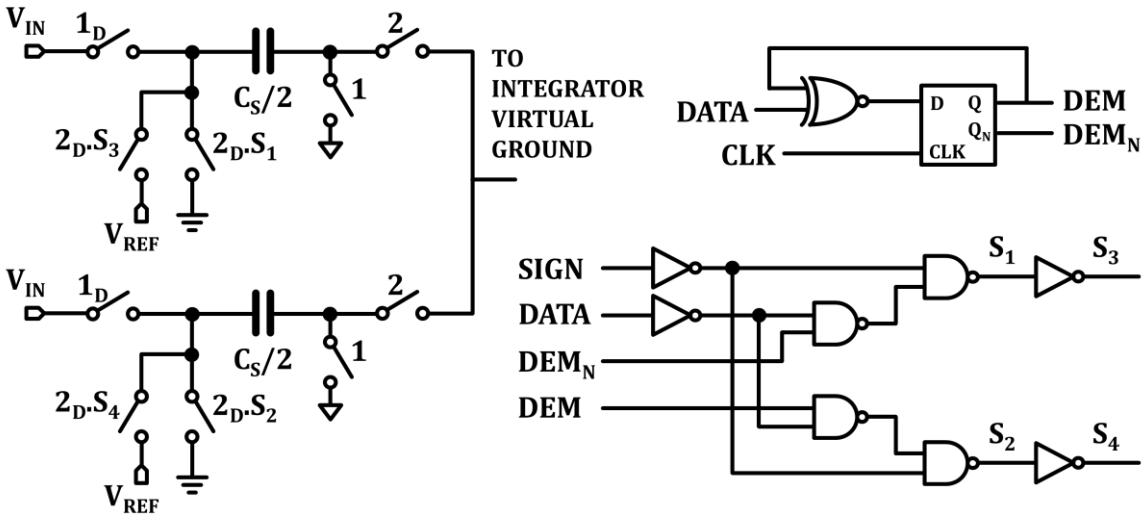


Figure 3.17 Split sampling capacitor to make a tri-level DAC during the integration phase and switches control signals generation (single-ended version shown)

The second integrator uses a scale-down version of the inverter-based amplifier present in the first integrator. A simpler amplifier is allowed due to relaxed requirements: no need of chopper modulation to remove the flicker noise (noise of second integrator is first-order shaped), and then no need to use cascode transistors. The current consumption also is divided by a factor 5 compared to the first integrator, thanks to the low capacitance values. No DAC is present on this stage and switches only follow the clock sampling and integration phases.

3.3 Summation circuit, quantizer and clocks generator design

Due to the feedforward architecture, the modulator needs a summation element in front of the quantizer. This summation is done in a passive way for a low-power solution, with switched-capacitors as presented on figure 3.18 in a single-ended version (design is differential). On the sampling phase, denoted 1 for the switches, the weighted capacitances (weights are derived from figure 3.1) are sampling the input signal, the first integrator output and the second integrator output. Then the summation phase, denoted 2 for the switches, directly connects all the capacitances in parallel and provides the output voltage almost instantly by charge redistribution effect. The kT/C noise of this switched-capacitor circuit is second-order shaped by the sigma-delta loop and has then very low requirements. Using low unitary capacitances C_U of 50fF is then sufficient and cap value is more defined to be robust against parasitic capacitances than to fulfill noise requirement. The output voltage at the end of the summation phase can be determined with the charge conversation law:

$$Q_1 = C_U V_{IN} + 8C_U V_{INTEG1} + 8C_U V_{INTEG2}, \qquad Q_2 = 17C_U V_{OUT} \tag{3.11}$$

$$Q_1 = Q_2 \underset{implies}{\Longrightarrow} V_{OUT} = \frac{V_{IN} + 8V_{INTEG1} + 8V_{INTEG2}}{17} \tag{3.12}$$

Equation (3.12) shows that the passive summation attenuates the signal by a factor corresponding of the total number of unit capacitances used in the summation circuit. Since the quantizer used is not single-bit with an undefined gain but a 1.5-bit with a unity gain, the quantizer gain must be adapted to compensate for the summation circuit attenuation and should then be 17.

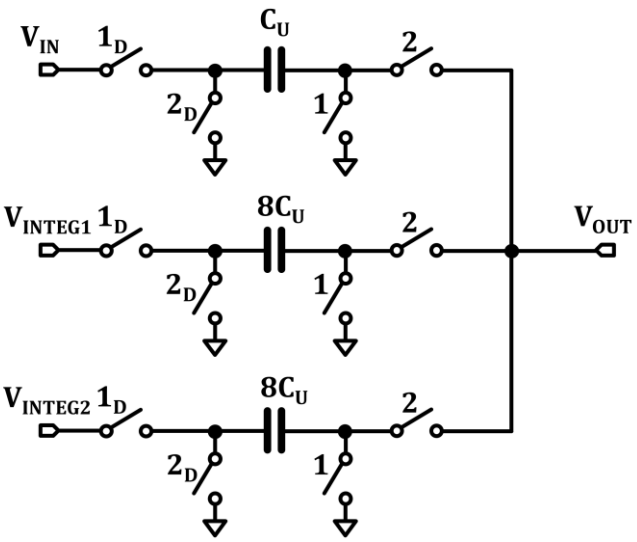


Figure 3.18 Passive switched-capacitor summation circuit

This can be done by scaling the quantizer references: if the signal is divided by 17, the references must be divided by the same factor to recover the normal behavior. Let's first determine the reference levels for the 1.5-bit quantizer with a unity gain. This is shown on figure 3.19 A) where a normalized (reference = 1, unity gain) 1.5-bit quantizer is illustrated. Two reference values are useful: -0.5 provides the limit between the output codes -1 and 0 while 0.5 is the limit between the output codes 0 and 1. Scaling this quantizer for a reference level of 1.5V as used in this modulator leads to multiply these reference levels by a 1.5 factor, giving respectively -0.75 and 0.75 for these limits as visible on figure 3.19 B). Now that the quantizer reference levels are defined, the last step is to divide them by a factor 17 to keep the same signal range than the summation circuit. The final reference levels must then be -44mV between output codes -1 and 0 and 44mV between output codes 0 and 1.

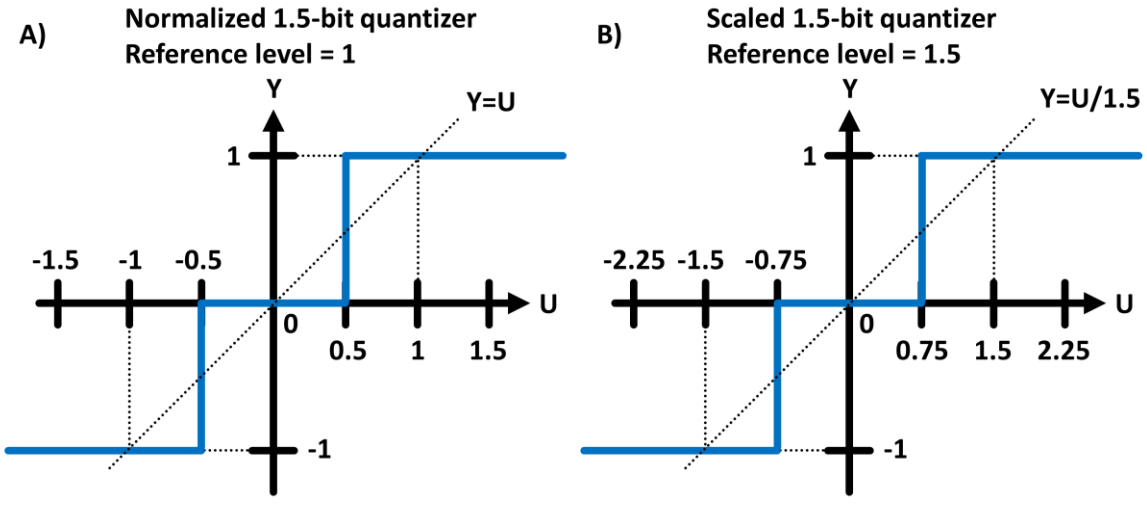


Figure 3.19 1.5-bit quantizer A) normalized B) scaled for 1.5V reference

The reference generation circuit must then provide a differential signal of $V_{REF}/(2 \times 17) =$ 44mV and depending of the polarity used, gives 44mV or -44mV reference level. This is done with the circuit of figure 3.20 where a resistor divider provides a 44mV differential signal centered around the common-mode of half the V_{REF} and differential capacitors are sampling the reference voltage.

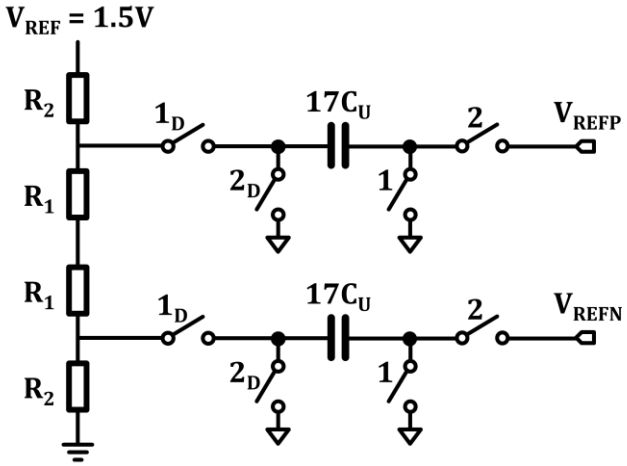


Figure 3.20 1.5-bit quantizer reference level generation

Capacitors have the same values than the summation circuit ones, to minimize mismatch between signal path and reference path. The resistors ratio is determined by the following relation:

$$V_{OUT} = \frac{R_1}{R_1 + R_2} V_{REF} = \frac{1}{34} \xrightarrow{implies} R_2 = 33R_1 \tag{3.13}$$

The complete summation circuit and 1.5-bit quantizer block diagram is given in figure 3.21. The dual difference amplifiers are making the difference $V_{IN1} - V_{IN2}$ and act also as pre-amplifiers for the clocked comparator, operating in open loop. Auto-zeroing capacitors C_{AZ} are used to get rid of the preamplifiers offset and then minimize quantizer errors. As said in chapter 3.2, the output of the quantizer should use a 2's complement coding with the MSB representing the sign and the LSB the data value. To simplify the decoding logic, the comparators outputs are not thermometer coded but use the following: Q_1 compares if the summation circuit output is greater than the 0 to 1 edge of the quantizer (44mV) while Q_0 compares if the summation circuit output is lower than the -1 to 0 edge of the quantizer (-44mV). Then, to obtain the 2's complement coding, the MSB D_1 is directly derived from Q_0 and the LSB D_0 is a simple logic OR between Q_0 and Q_1. This is summarized in table 3.6.

Table 3.6 Quantizer logic truth table

Quantizer input (= summation output V_{SUM})	**Output value**	**Comparator code**		**Quantizer code**	
		Q_1	**Q_0**	**D_1**	**D_0**
44mV < V_{SUM}	1	1	0	0	1
-44mV < V_{SUM} < 44mV	0	0	0	0	0
V_{SUM} < -44mV	-1	0	1	1	1

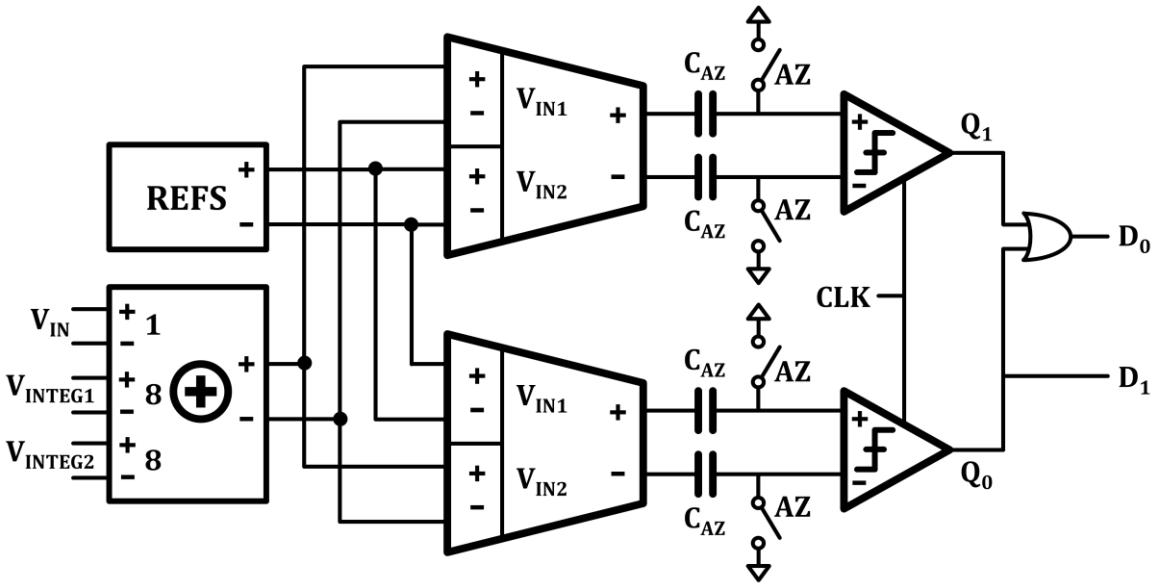


Figure 3.21 1.5-bit quantizer and summation circuit block diagram

The schematic of the dual difference amplifiers is given in figure 3.22. It is based on the self-biased amplifier presented in figure 3.4 where the input inverters are doubled and cross connected to amplify the inverters current difference. Since the requirements of the quantizer is relaxed by the second-order sigma-delta loop and the open-loop configuration of the pre-amplifiers using auto-zeroing, there is no need to control precisely the output common-mode that's why simpler self-biased amplifier structure is preferred over the loop filter amplifier. The second reason is the low biasing current used that is not needed to be controlled precisely since its variations do not influence much the total modulator consumption. The comparator is a classic clocked realization as in [23, 34] followed by a clocked SR latch to avoid data change when the comparator is reset. Its schematic is shown on figure 3.23.

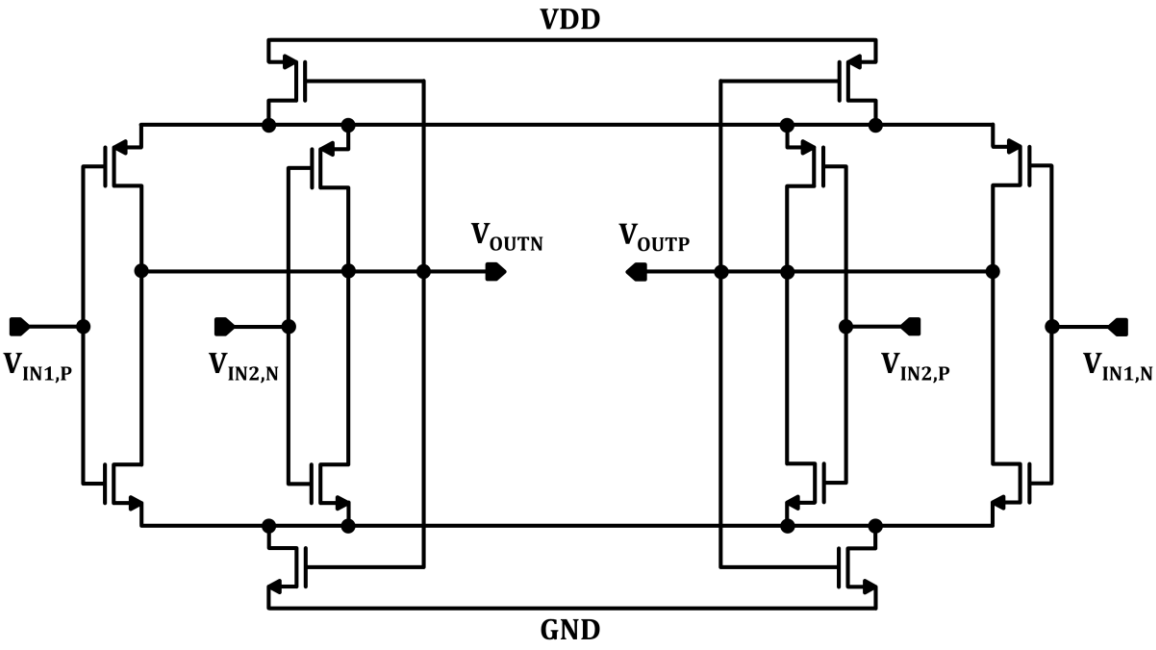


Figure 3.22 Schematic of the dual difference pre-amplifiers used in 1.5-bit quantizer

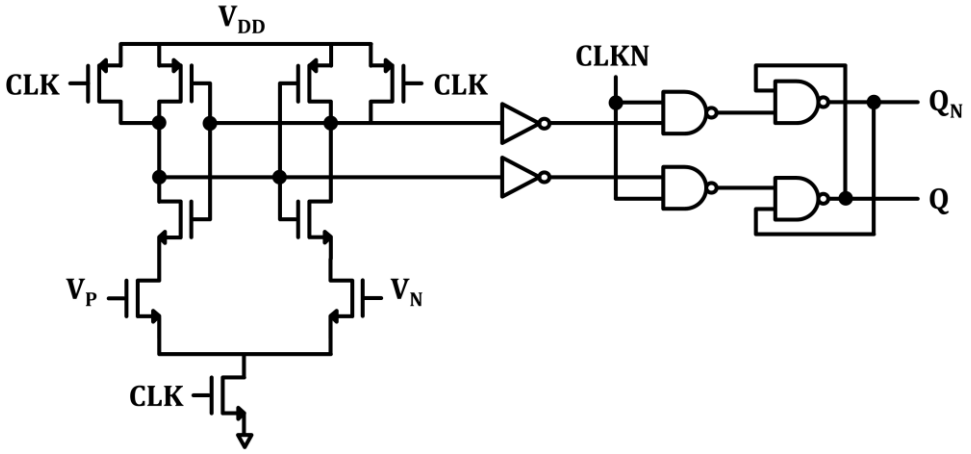


Figure 3.23 Schematic of the clocked comparator used in 1.5-bit quantizer

The clock phases generator schematic is presented in figure 3.24 with its associated timing diagram. It uses a cross-coupled NAND structure with several delays to ensure non-overlapping phases. Φ_1 is the sampling phase and Φ_{1D} is its corresponding delayed phase to avoid the switch charge injection issue [16, 34, 35]. The same applies for the integration phase Φ_2. Φ_1 and Φ_2 are not taken in a symmetric point of the structure: this is done to have a minimum delay for the Φ_2 to Φ_1 transition while the other transition (Φ_1 to Φ_2) has a larger delay to chop the amplifier during the non-overlapping time. The chopper runs at $F_S/2$ and uses a second non-overlapping structure to generate its phases Φ_A and Φ_B. The chopper transitions occur during a sampling to integration transition; if it was occurring during an integration to sampling transition, the charge integrated on integration capacitors could change due to the chopper transition and the next stage would sample a wrong value. These clock phases directly refer to the ones present in the first integrator (figure 3.15) and the other presented schematics.

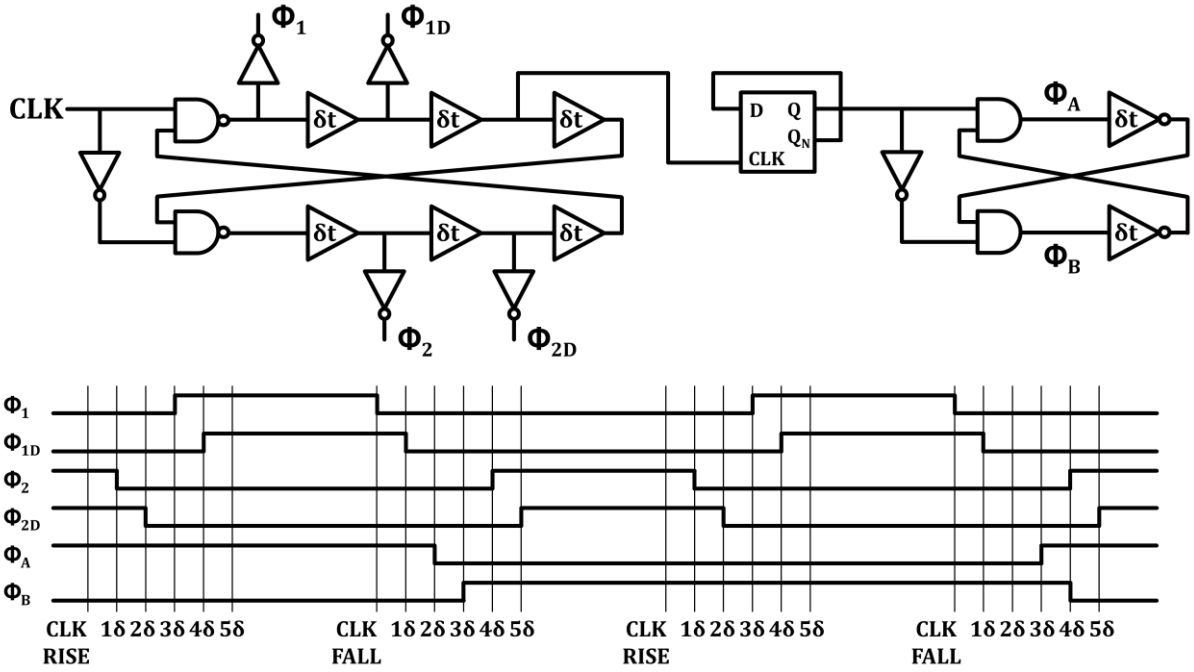


Figure 3.24 Schematic of the clock phases generator and its associated timing diagram

3.4 Simulation results

Complete transistor-level simulations of the modulator are made to check its performances. PVT simulations results for a -3dBFS input signal are visible on figure 3.25 where the typical modulator output spectrum is shown together with the two extremum corners: the slow at -40°C and the fast at 125°C. The expected thermal noise level, corresponding to the specification of 7.5µVrms integrated on a 24kHz bandwidth is also shown. In typical conditions, the modulator reaches a Signal-to-Quantization-Noise-Ratio (SQNR) of 111.2dB on 24kHz bandwidth while its Signal-to-Quantization-Noise-and-Distortion-Ratio (SQNDR) is 108.3dB; in accordance with the specifications. The second-order noise shaping is clearly visible on the spectrum, as well as a flat noise floor at low frequencies: this was expected due to the finite amplifiers GBW product and SR, the same behavior was shown by the high-level model in figure 2.29. The harmonic distortion was not predicted by the model and is coming mainly from the non-linear on-resistance of the CMOS switches, but is kept at a moderate level and better than the specification of 95dB SNDR. The slow corner at -40°C shows similar performances as the typical one, while the fast a 125°C has a slightly degraded harmonic performance, reducing its SQNDR to about 103dB, which is still in specification.

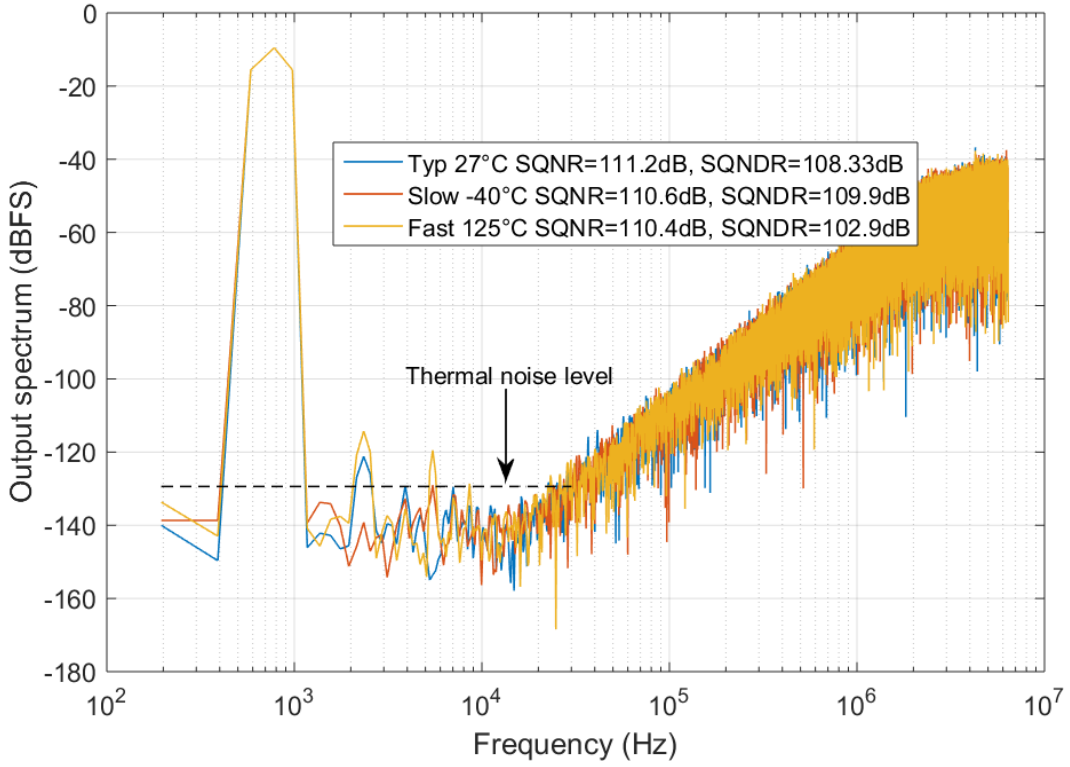


Figure 3.25 Modulator output spectrums over PVT conditions: typical, slow at -40°C and fast at 125°C

Mismatch simulations are also run to check that the DWA algorithm implemented works correctly and is sufficient. Results are shown on figure 3.26 where the modulator reaches the typical performances (111dB SQNR and 110dB SQNDR) when DWA is active while performances are highly reduced without DWA, to 93.7dB SQNR and 85.9dB SQNDR due to DAC non-linearity.

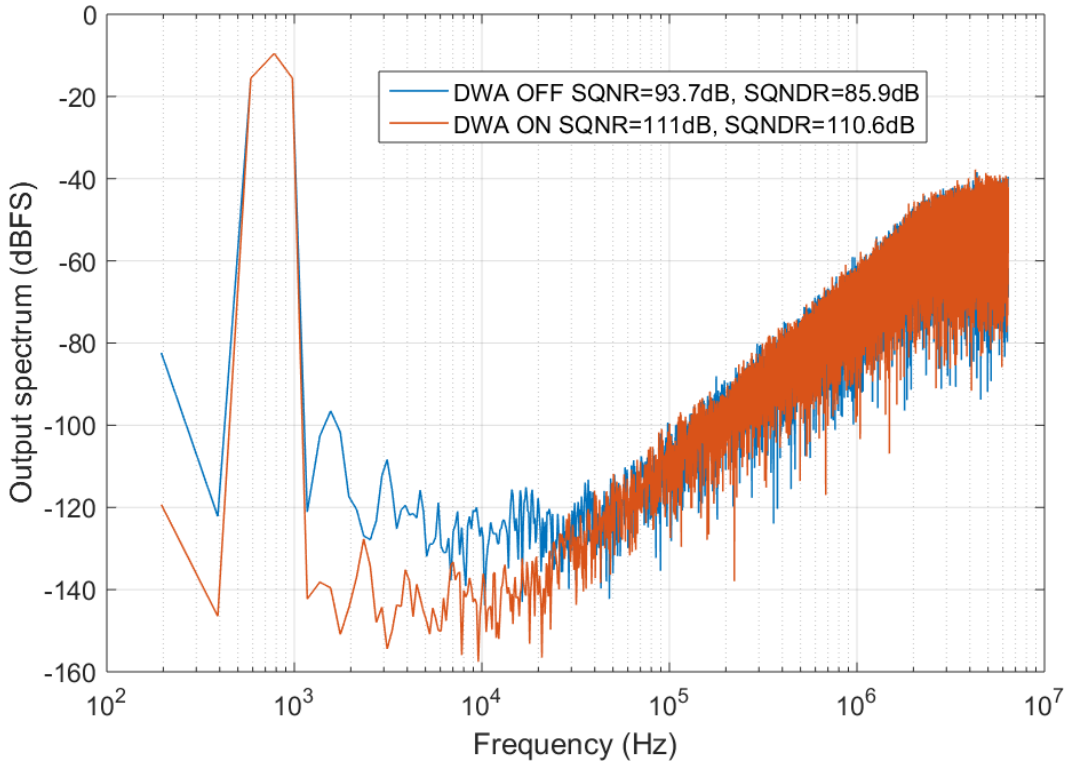


Figure 3.26 Modulator output spectrums under mismatch conditions with and without the implemented DWA algorithm

Corners simulation are also run with a low signal level (-60dBFS), checking that no tones appear when signal energy is dominated by the quantization noise as visible on figure 3.27.

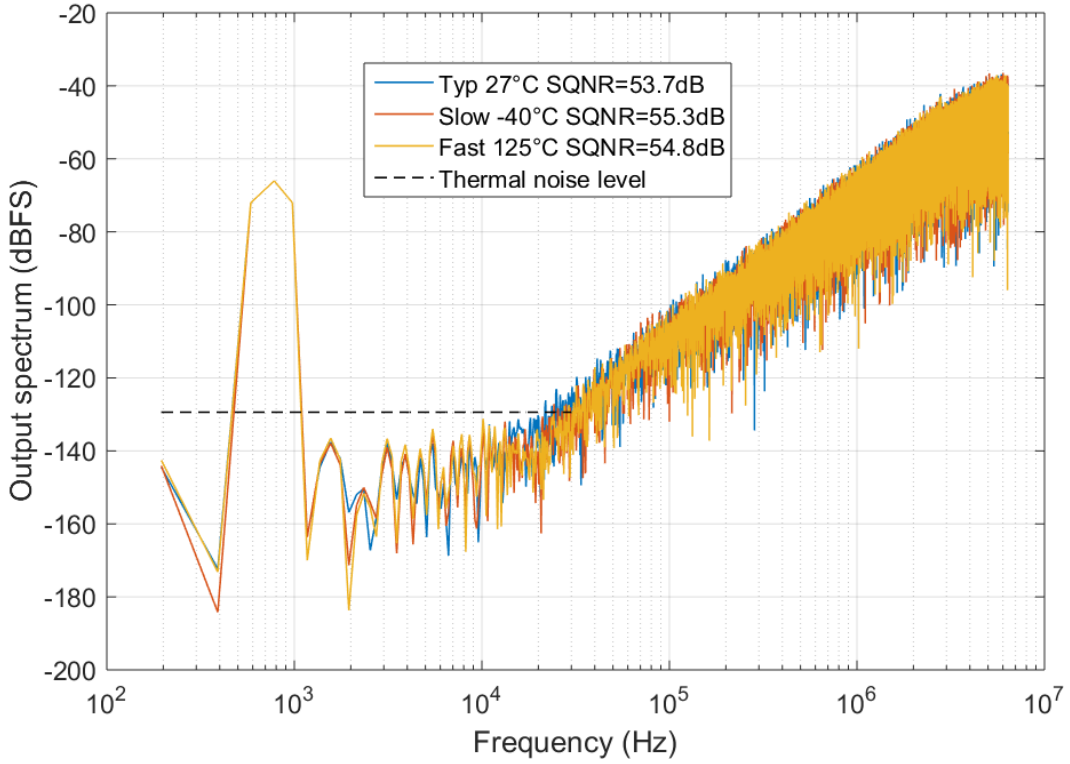


Figure 3.27 Modulator output spectrums over PVT conditions for a -60dBFS input signal

Figure 3.28 shows the power consumption split of the 480µW designed modulator power consumption. The most power-hungry element is the first integrator, which consumes almost half of the total power consumption, followed by the digital logic for the switch drivers and clock phases generation. The latter power consumption can be expected to be reduced using a lower node technology than used in this thesis (140nm).

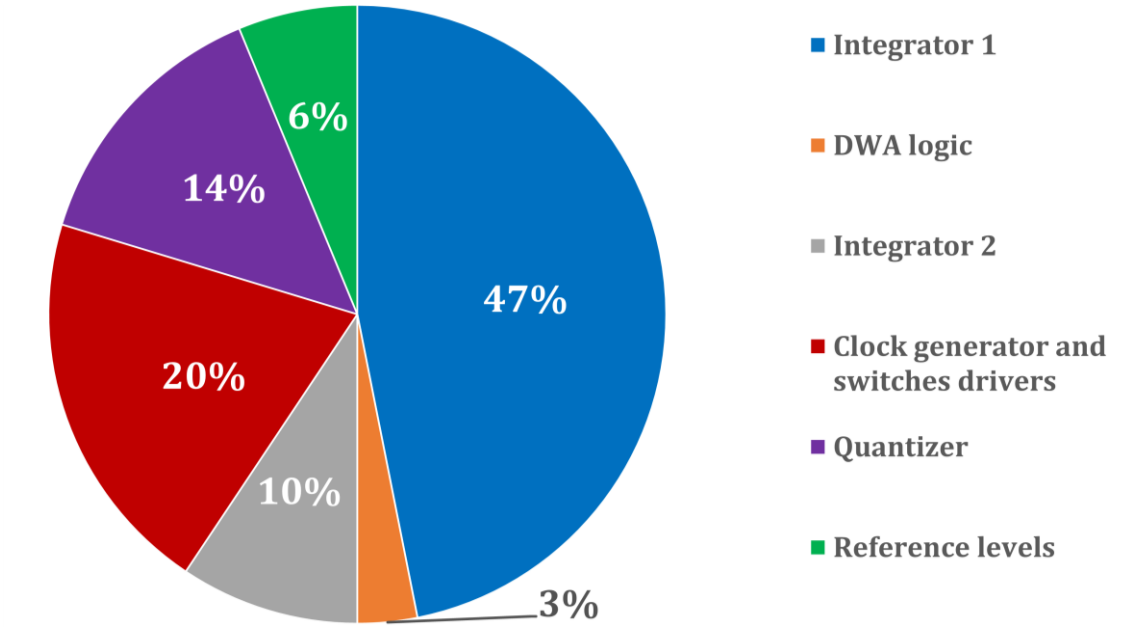


Figure 3.28 Modulator power consumption split

The modulator total input-referred noise is 7.5µVrms on a 24kHz bandwidth, as defined by the specification in beginning of this chapter. The noise power split is given in figure 3.29 where about 80% is coming from the first stage, 10% for the quantization noise and 7% from the DAC reference voltage (which is coming from a low-noise 2µVrms regulator). As expected, the second integrator noise account is negligible, since it is first-order shaped and capacitors values does not have the minimum size due to matching requirements.

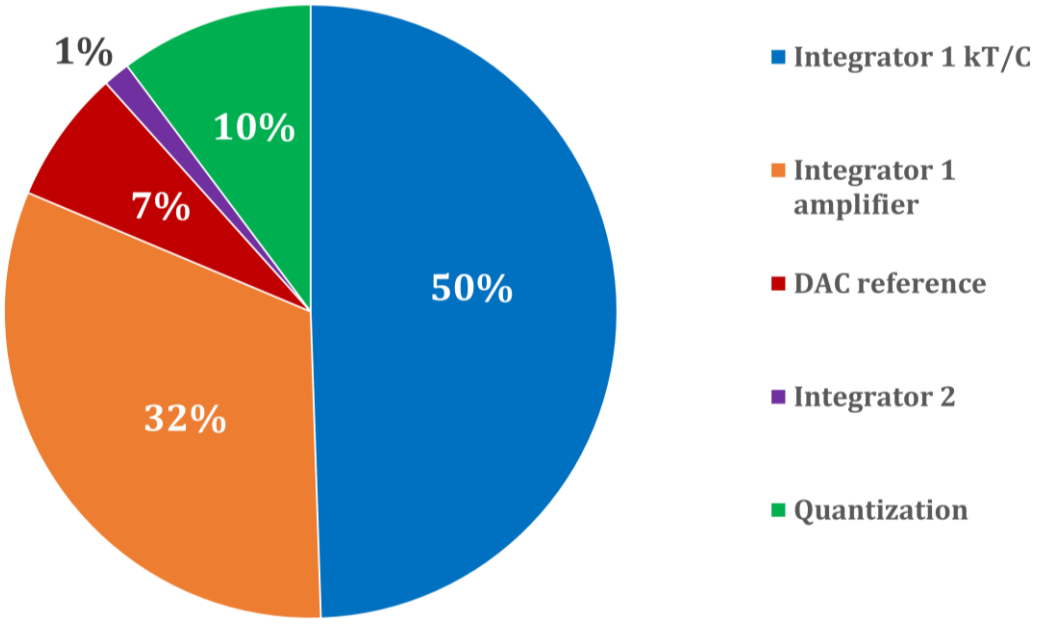


Figure 3.29 Modulator input-referred noise power split

The modulator designed in 140cm CMOS technology occupies an area of 0.084mm² and its layout is shown on figure 3.30.

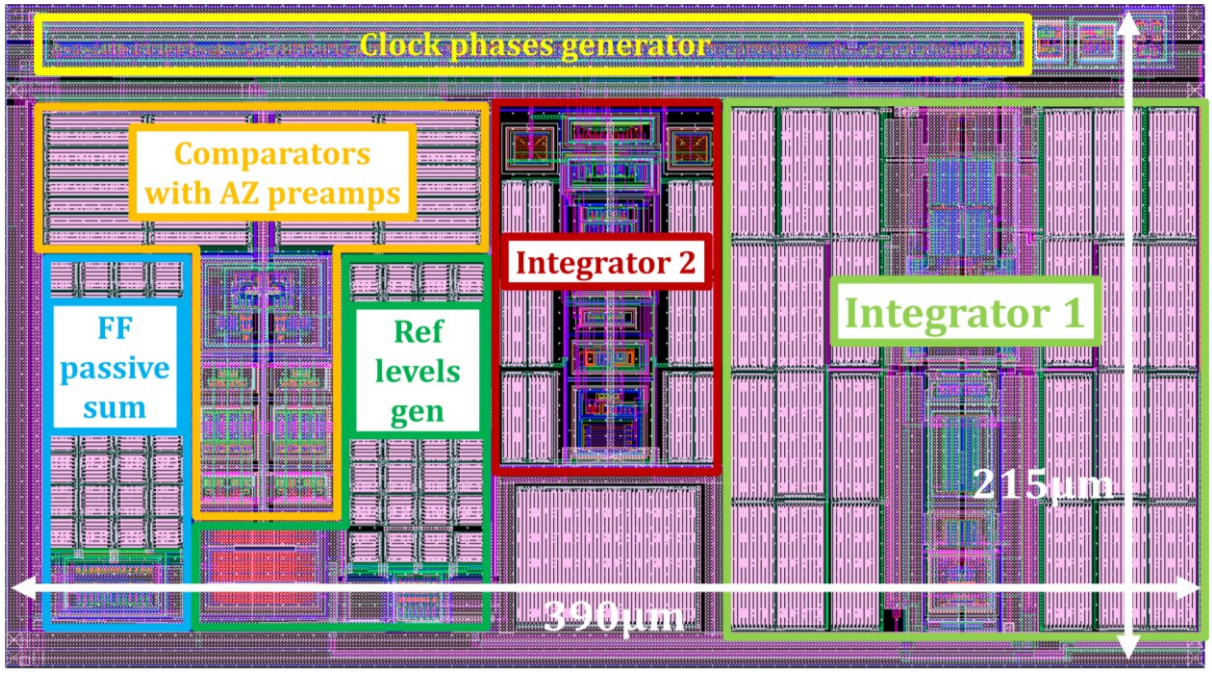


Figure 3.30 Layout of the designed modulator

The designed modulator performances are summarized in table 3.7. They are compared with other discrete-time audio sigma-delta modulators reported in literature, like the state-of-the-art and other interesting designs precedingly cited. As [21], this design shows that a high OSR can be used in a power-and-area efficient design, benefiting from a simpler quantizer and DEM circuit (if even needed), and low size capacitors. The modulator reaches a 180dB FOMs placing it among the state-of-the-art while occupying only half of the other designs lowest area.

Table 3.7 Performance summary and comparison with literature

	This work	**[16]**	**[17]**	**[20]**	**[21]**	**[23]**	**[34]**	**[35]**
TECH (nm)	140	180	180	180	160	350	130	90
YEAR	2017	2011	2012	2009	2016	2013	2012	2004
ORDER	2	4	3	3	3	4	4	3
BW (kHz)	24	25	50	20	20	20	20	20
FS (MHz)	12.288	1.6	3.2	4	11.29	2.4	2.5	4
OSR	256	32	32	100	282	60	62	100
Supply (V)	1.5	1	1.5	0.7	1.8	1.5	0.5	0.6
DR (dB)	103	93	94	85	109	92.6	85	88
SNR (dB)	100	92.5	-	84	106	-	82.4	85
SNDR (dB)	98	92	88.9	81	103	87.9	81.7	81
POWER (µW)	480	58	140	22	1120	140	35.2	140
AREA (mm²)	0.084	-	0.49	0.715	0.16	0.207	0.57	0.18
FOMs (dB)	180	179.3	179.5	174.5	181.5	174	173	169.5

The figure 3.31 shows the modulator placement on the ADC survey of Boris Murmann [4], where the FOMS has been changed to use DR instead of SNDR (original definition), and where only sigma-delta modulators are shown.

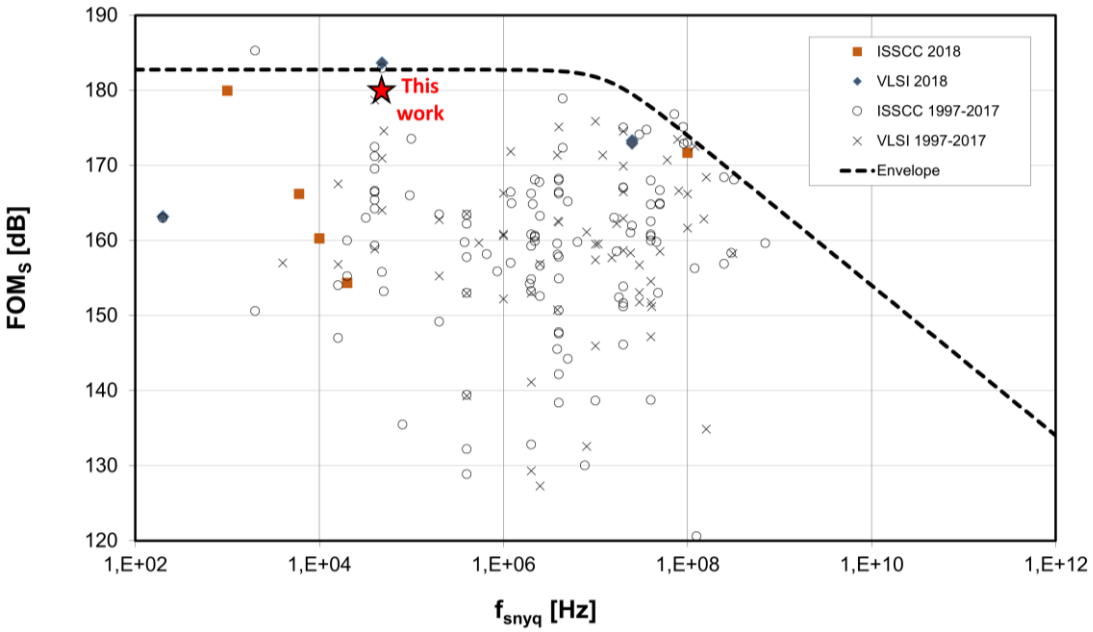


Figure 3.31 Placement of the discrete-time modulator on Murmann's ADC survey

3.5 Conclusion

In this chapter, a brief review of the state-of-the-art discrete-time sigma-delta modulator designs has been made, showing that inverter-based amplifiers are well suited for low-power designs due to their class AB behavior. Moreover, for future technology nodes, inverter-based amplifiers are also suitable due to voltage headroom limitations. However, inverters are highly sensitive to PVT making their biasing point unstable and not reliable for mass production. A new inverter-based amplifier is then introduced, with linear biasing transistor, which avoid limiting their peak current while controlling the biasing current and common-mode level. A modulator design using these amplifiers is then made, which reaches a 180dB FOMS, placing it among the state-of-the-art designs. To get a low-cost design for the targeted audio application (microphone analog front-end), a high OSR is used to reduce the capacitances, which represents the larger part of the modulator area. A 1.5-bit quantizer is used for its resolution increase and lower integrator swings it provides, while needing a simple and negligible power DEM mechanism. Then, a second-order architecture employing a feedforward topology is sufficient to reach a high DR of 103dB.

Chapter 4

Continuous-time modulators: theory and specificities

This chapter focuses on continuous-time sigma-delta modulators theory and their key points that make them differ from their discrete-time counterparts. Their key strengths are an easier input driving (generally resistive), an inherent anti-aliasing filtering capability and a lower required GBW product for the amplifiers (in order of the sampling frequency compared to about five times the sampling frequency in discrete-time realizations). These advantages allow to reduce the systems complexity and power (no need of an anti-aliasing filter, no need of a driving buffer, and a reduced loop filter power), or to enable the use of sigma-delta modulators at higher speed. But continuous-time modulators have also some specific drawbacks: they are sensitive to timing errors like timing offset (delay) or timing uncertainty (jitter) which can reduce the resolution or lead to instability. Their integrator time constants are also sensitive to process and temperature variations: they are defined by a RC product or a gm/C ratio and changes in their values can lead to instability. Contrarily, in a discrete-time modulator the integrators time constants are well defined (they depend on the sampling frequency and a capacitor ratio) and are sensitive only to capacitor mismatch [9,36].

Numerous papers have been published on the electrical design of continuous-time modulators [37, 38, 39, 40], some others concentrate on the theory [36, 41] and equivalence with the discrete-time ones but few of them combine both to provide a complete design methodology. This chapter presents a high-level design flow of a continuous-time sigma-delta modulator. It contains a coefficient calculation procedure to design a continuous-time modulator equivalent to a discrete-time one with the desired Noise Transfer Function (NTF), and how to handle the continuous-time specificities and drawbacks that are not present in discrete-time modulators. A high-level design example of a third-order audio continuous-time modulator is made all along to illustrate the developed theory with detailed calculations. It is a low-pass sigma-delta modulator but the procedure is similar for band-pass modulators. The modulator runs at 6.144MHz (with an OSR of 128), and reaches a Signal to Quantization Noise Ratio (SQNR) of 103dB on 24kHz band. The design procedure starts by defining a stable NTF in the z-domain as it can be done for designing a discrete-time modulator. In the studied example the desired $NTF_D(z)$ has a relatively low out-of-band gain of 1.5 which ensures modulator stability, and has been defined using the well-known Schreier theory [5]:

$$NTF_D(z) = \frac{(1 - z^{-1})^3}{1 - 2.2003z^{-1} + 1.6887z^{-2} - 0.4444z^{-3}} \tag{4.1}$$

The goal of the procedure developed is to define the continuous-time modulator coefficients to make it equivalent to the ideal discrete-time modulator implementing $NTF_D(z)$ as shown in figure 4.1. The modulator architecture is a generic cascade of integrators with a feedforward summation, but the procedure can be applied to any customized modulator architecture. The rest of this chapter concentrates on continuous-time modulators specificities which make them differ from discrete-time ones and specific techniques to compensate their drawbacks.

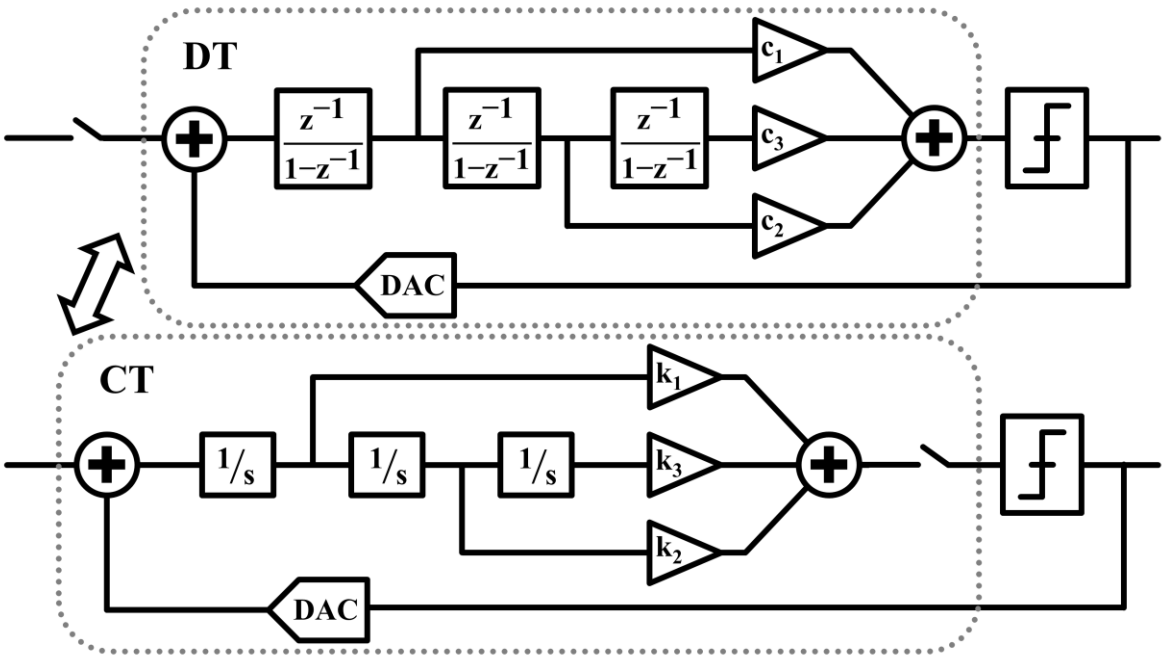
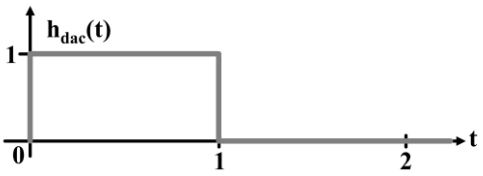


Figure 4.1 Equivalent discrete-time and continuous-time implementations of modulators with NTF_D

4.1 Continuous-time to discrete-time equivalence

A continuous-time sigma-delta modulator is a hybrid continuous-discrete time system where the continuous-time signal path (DAC + loop filter), which is sampled on output, must behave as its equivalent discrete-time one. To make this happen, the Impulse-Invariance Transform (IIT) is used [36,41]: the z-transform of the continuous-time signal path impulse response should be equal to the discrete-time DAC plus the loop filter z-transform. In that case, both continuous-time and discrete-time paths are equivalent, and the modulator can be designed using the standard discrete-time sigma-delta theory.

Let's do it on the third-order modulator presented before which has normalized sampling frequency (1Hz) and reference level (1V). The impulse response of the DAC and the loop filter must be determined. In this example a NRZ DAC is used but the procedure is the same with other DAC pulses; it begins by determining the DAC impulse response. The NRZ DAC behaves as a zero-order hold, its impulse response $h_{dac}(t)$ is shown on figure 4.2 and expressed in (4.2). The loop filter output can be decomposed in the weighted sum of 3 branches using the linearity property of the addition operation [42]: one with a single integrator, one with two integrators and one with a third-order integration function as it can be seen on figure 4.3.



$$h_{dac}(t) = \begin{cases} 1, & for\ 0 \le t \le 1 \\ 0, & for\ t > 1 \end{cases} \tag{4.2}$$

Figure 4.2 Impulse response of a NRZ DAC with 1Hz sampling frequency

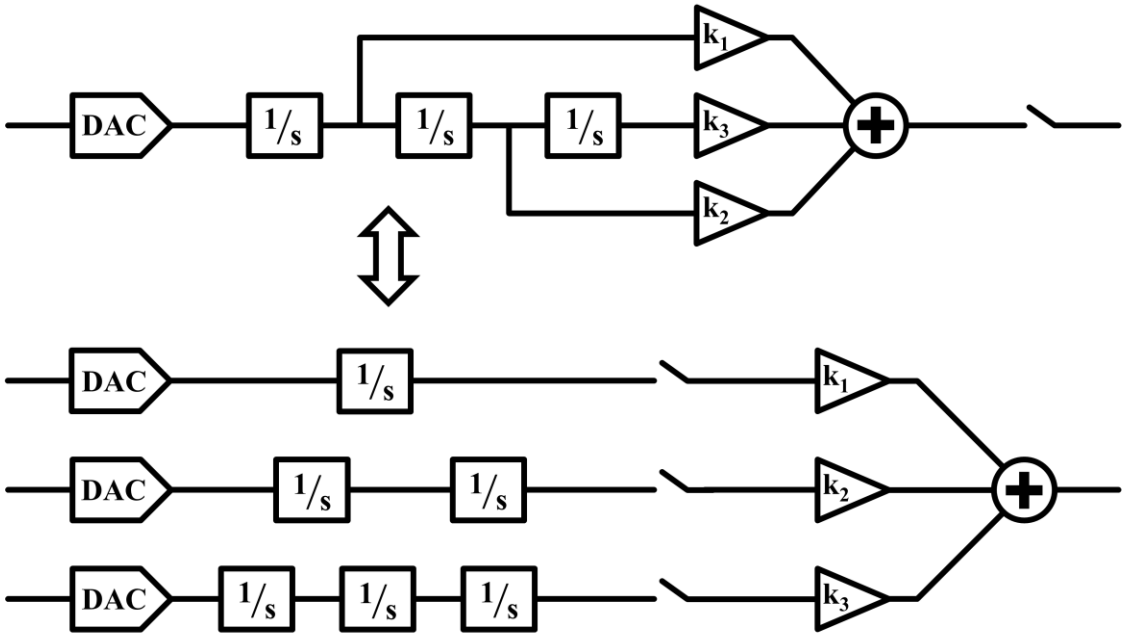


Figure 4.3 Decomposition of DAC + loop filter in different integrating order paths

Let's evaluate the impulse-invariance transform of each branch separately and add them up to get the final discrete-time equivalent transfer function of the loop filter. Starting with the first-order integrating branch, its impulse response is equal to the integration of the DAC impulse response. Then, its impulse response $h_1(t)$ can be defined as follows:

$$h_1(t) = \begin{cases} \int_0^t 1 \times dt\,, & for\ 0 \le t \le 1 \\ h_1(1) + \int_1^t 0 \times dt\,, & for\ t > 1 \end{cases} = \begin{cases} t, & for\ 0 \le t \le 1 \\ 1, & for\ t > 1 \end{cases} \tag{4.3}$$

The impulse response is then sampled every $T_S = 1$ second, leading to the corresponding $h_{1d}(n) = h_1(nT_S)$ discrete impulse response:

$$h_{1d}(n) = \begin{cases} n, & for\ n \in \{0,1\} \\ 1, & for\ n > 1 \end{cases} = \begin{cases} 0, & for\ n = 0 \\ 1, & for\ n \ge 1 \end{cases} \tag{4.4}$$

Last step, the z-transform of $h_{1d}(n)$ is taken to get $H_1(z)$:

$$H_1(z) = \sum_{n=0}^{\infty} h_{1d}(n) z^{-n} = \sum_{n=1}^{\infty} z^{-n} = \frac{z^{-1}}{1 - z^{-1}} \tag{4.5}$$

To continue, the impulse response of the second-order integrating branch is now evaluated, by integrating the impulse response of the first-order branch, then $h_2(t)$ is defined as:

$$h_2(t) = \begin{cases} \int_0^t t \times dt, & for\ 0 \le t \le 1 \\ h_2(1) + \int_1^t 1 \times dt, & for\ t > 1 \end{cases} = \begin{cases} \frac{t^2}{2}, & for\ 0 \le t \le 1 \\ t - 0.5, & for\ t > 1 \end{cases} \tag{4.6}$$

Then, by sampling $h_2(t)$ it comes $h_{2d}(n) = h_2(nT_S)$:

$$h_{2d}(n) = \begin{cases} \frac{n^2}{2}, & for\ n \in \{0,1\} \\ n - 0.5, & for\ n > 1 \end{cases} = \begin{cases} 0, & for\ n = 0 \\ n - 0.5, & for\ n \ge 1 \end{cases} \tag{4.7}$$

Taking the z-transform, it comes $H_2(z)$:

$$H_2(z) = \sum_{n=0}^{\infty} h_{2d}(n) z^{-n} = \sum_{n=1}^{\infty} n z^{-n} - 0.5 \sum_{n=1}^{\infty} z^{-n} = \frac{z^{-1}}{(1 - z^{-1})^2} - 0.5 \frac{z^{-1}}{1 - z^{-1}} = \frac{0.5z^{-1} + 0.5z^{-2}}{(1 - z^{-1})^2} \tag{4.8}$$

Repeating a last time this process to get the z-transform of the third-order integrating branch, where the impulse response is the integration of the second-order impulse response:

$$h_3(t) = \begin{cases} \int_0^t \frac{t^2}{2} \times dt, & for\ 0 \le t \le 1 \\ h_3(1) + \int_1^t (t - 0.5) \times dt, & for\ t > 1 \end{cases} = \begin{cases} \frac{t^3}{6}, & for\ 0 \le t \le 1 \\ \frac{t^2}{2} - 0.5t + \frac{1}{6}, & for\ t > 1 \end{cases} \tag{4.9}$$

Taking the discrete-time impulse response $h_{3d}(n) = h_3(nT_S)$:

$$h_{3d}(n) = \begin{cases} \frac{n^3}{6}, & for\ n \in \{0,1\} \\ \frac{n^2}{2} - 0.5n + \frac{1}{6}, & for\ n > 1 \end{cases} = \begin{cases} 0, & for\ n = 0 \\ \frac{n^2}{2} - 0.5n + \frac{1}{6}, & for\ n \ge 1 \end{cases} \tag{4.10}$$

It leads the z-transform $H_3(z)$:

$$H_3(z) = \sum_{n=0}^{\infty} h_{3d}(n) z^{-n} = \frac{1}{2} \sum_{n=1}^{\infty} n^2 z^{-n} - \frac{1}{2} \sum_{n=1}^{\infty} n z^{-n} + \frac{1}{6} \sum_{n=1}^{\infty} z^{-n} \tag{4.11}$$

$$H_3(z) = \frac{1}{2} \frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3} - \frac{1}{2} \frac{z^{-1}}{(1 - z^{-1})^2} + \frac{1}{6} \frac{z^{-1}}{(1 - z^{-1})} = \frac{\frac{1}{6}z^{-1} + \frac{4}{6}z^{-2} + \frac{1}{6}z^{-3}}{(1 - z^{-1})^3} \tag{4.12}$$

The discrete-time equivalent transfer function of each DAC plus loop filter branch is now defined, and taking the weighted sum gives the complete DAC and loop filter discrete-time equivalent transfer function:

$$H(z) = k_1 H_1(z) + k_2 H_2(z) + k_3 H_3(z) \tag{4.13}$$

$$H(z) = \frac{\left(k_1 + \frac{k_2}{2} + \frac{k_3}{6}\right) z^{-1} + \left(-2k_1 + \frac{4k_3}{6}\right) z^{-2} + \left(k_1 - \frac{k_2}{2} + \frac{k_3}{6}\right) z^{-3}}{(1 - z^{-1})^3} \tag{4.14}$$

Deducing the equivalent noise transfer function $NTF(z)$:

$$NTF(z) = \frac{(1 - z^{-1})^3}{1 + \left(k_1 + \frac{k_2}{2} + \frac{k_3}{6} - 3\right) z^{-1} + \left(3 - 2k_1 + \frac{4k_3}{6}\right) z^{-2} + \left(k_1 - \frac{k_2}{2} + \frac{k_3}{6} - 1\right) z^{-3}} \tag{4.15}$$

By identification with the desired discrete-time noise transfer function $NTF_D(z)$ presented in introduction, k_1, k_2 and k_3 are deduced by solving the following equation system:

$$\begin{cases} k_1 + \frac{k_2}{2} + \frac{k_3}{6} = 3 - 2.2003 \\ -2k_1 + \frac{4k_3}{6} = 1.6887 - 3 \\ k_1 - \frac{k_2}{2} + \frac{k_3}{6} = 1 - 0.4444 \end{cases} \overset{gives}{\Longrightarrow} \begin{cases} k_1 = 0.6703 \\ k_2 = 0.2441 \\ k_3 = 0.0440 \end{cases} \tag{4.16}$$

Theoretically, the resulting continuous-time modulator has the same noise transfer function than the discrete-time ideal modulator. The same methodology can be applied with higher-order modulator and/or different DAC pulse shape to determine continuous-time modulator coefficients. Let's now compare the behavior of the continuous-time modulator with its equivalent discrete-time one. As shown on figure 4.4, the noise shaping is undistinguishable between the two modulators, confirming that the coefficient calculation is correct and lead to the same NTF.

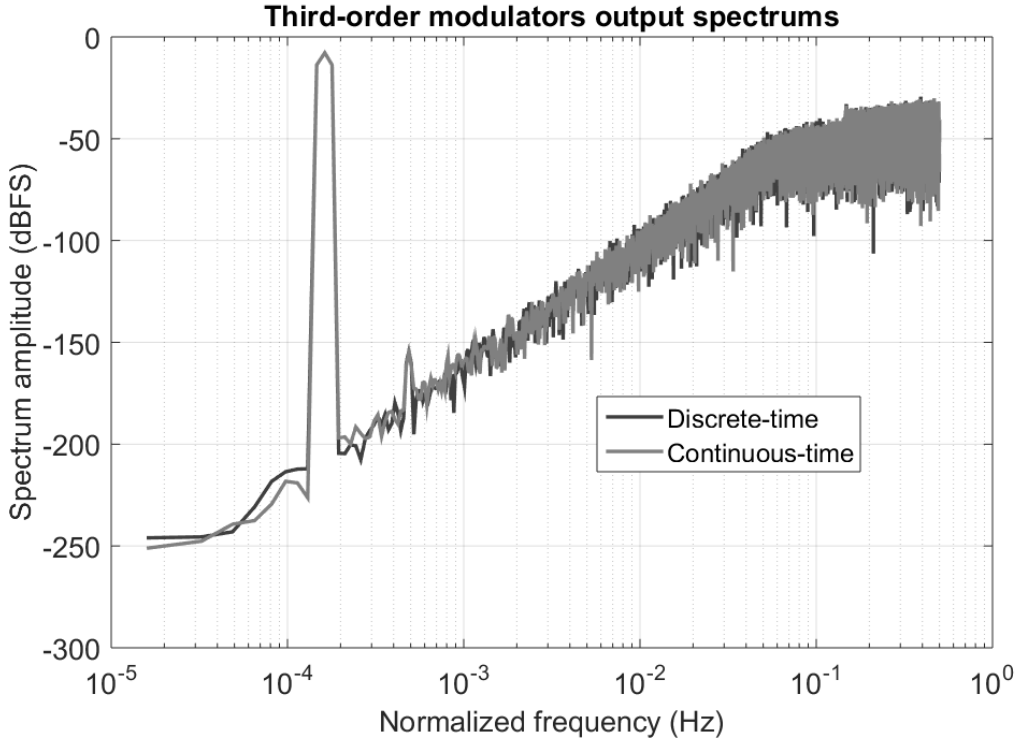


Figure 4.4 Output spectrums of ideal discrete-time modulator and synthesized continuous-time one

4.2 Integrators time constant

The modulator is still normalized and runs at 1Hz. Contrarily to a discrete-time modulator where the integrator time constants scale according to the sampling period [5], in a continuous-time modulator the integrator time constants are defined by a RC product and are then fixed. Hence, scaling the modulator in function of the desired sampling frequency is necessary. Sampling the loop filter output at the rate F_S implies that integrators have to output the same value in a time divided by F_S, so they need a F_S scaling factor to restore the desired behavior. Then, $1/s$ becomes F_S/s and the implementation leads to a $1/F_S$ integrators RC time constant [36]. A continuous-time modulator is adapted to only one F_S: if the sampling frequency is reduced, the integrator outputs will have too much gain which can lead to instability (overloading of quantizer input range). This is shown on figure 4.5 where a large increase in the out of band gain appears when the continuous-time modulator runs at $F_S/2$. The magnitude is plotted in linear and not in dB to better emphasize the difference. Contrarily, a discrete-time modulator will have a decreased resolution but will keep the same out of band gain, keeping its stability. If the sampling frequency is increased, integrators swings will be lower which is good for the stability but the resolution doesn't grow as fast as a discrete-time does when increasing the oversampling ratio: the noise shaping will remain the same as shown on figure 4.5 when the continuous-time modulator runs at $2F_S$; only the quantization noise level will be lower since it will be spread on a higher bandwidth. In contrast, a discrete-time modulator will have a better noise shaping behavior scaled by the sampling frequency. This is one drawback of a continuous-time modulator versus a discrete-time one: the modularity given by its resolution scaled with the sampling frequency.

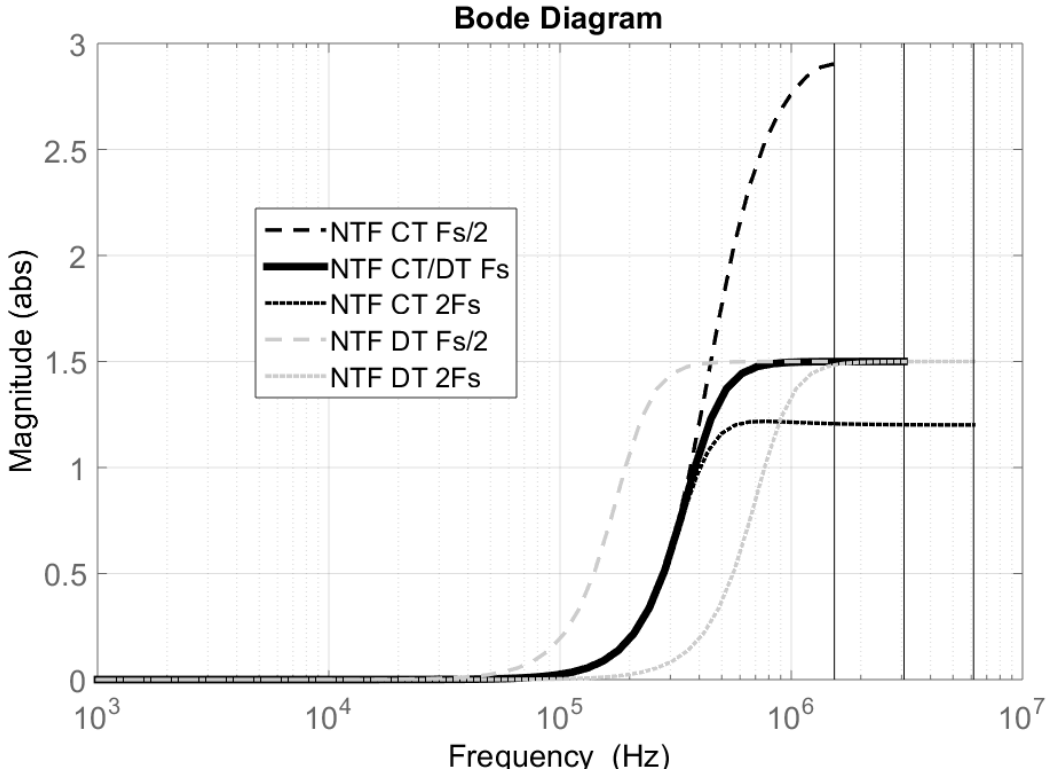


Figure 4.5 Effect of the clock frequency variation on both continuous-time and discrete-time NTFs

In a discrete-time modulator, the integrators time-constants are precise since they depend on the clock frequency accuracy and on a capacitor ratio, sensitive only to the capacitors mismatch [9]. However, in a continuous-time modulator, integrator time constants are defined by a RC product (or gm/C ratio). This product is not accurate and is mainly sensitive to Process, Voltage, and Temperature (PVT) variations. Figure 4.6 a) shows the impact on the noise transfer function of a +/- 30% variation of the integrators time constant. The magnitude is still plotted in linear and not in dB to better emphasize the difference. Decreasing the time constant will result in a better noise shaping and an increased resolution but the out of band gain is also increased, which will affect the modulator stability. Contrarily, when the time constant increases the out of band gain is reduced, which is good for loop stability, but the noise shaping is not as efficient as the ideal NTF, more noise is present on lower frequency and affects the resolution. This reflects on the simulated modulator performance as it can be seen on figure 4.6 b). An increase of the integrators time constant leads to a higher maximum stable amplitude combined with a lower achievable resolution, and the contrary happens for a reduction of integrators time constant. To handle these time constants variations, either a low out of band gain NTF can be selected, which will provide sufficient margin for stability even with 30% variation, or the RC time constant have to be made tunable, allowing a more aggressive NTF to be used.

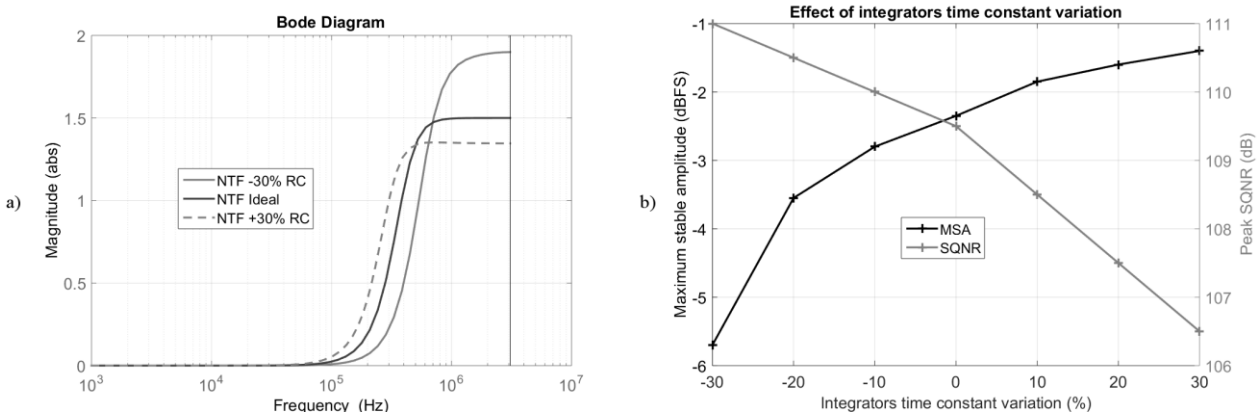


Figure 4.6 Effect of integrators time constant variation on a) NTF b) maximum stable amplitude (MSA) and reachable SQNR

4.3 Finite DC gain and anti-aliasing property

Like in discrete-time modulators, the integrators DC gain is an important parameter in continuous-time modulators [9, 36, 41]. Since the loop filter is composed of integrators, the finite amplifiers DC gain limits the achievable loop filter gain at low frequencies. This impacts the noise transfer function as it can be seen on figure 4.7 a). In this plot, a comparison between an ideal NTF and its degraded version where the three integrators have a DC gain of 40dB is presented. In low frequencies, there is no more noise shaping and the NTF response tend to be flat. To not impact the achievable SQNR, the zeros in the NTF introduced by finite DC gain must be placed at a lower frequency than the signal bandwidth. The finite DC gain value also increases the level of distortion: as shown on figure 4.7 b), where the rejection of the third harmonic (HD3) is 30dB lower with a DC gain of 40dB than with a 60dB DC gain. This is not

much visible on SQNR since there is only a 3dB difference between these two DC gain values, but in case of a multiple frequencies input signal, with distortion and intermodulation several tones can appear in signal bandwidth and degrade the achievable SQNR if the harmonic rejection is not sufficient.

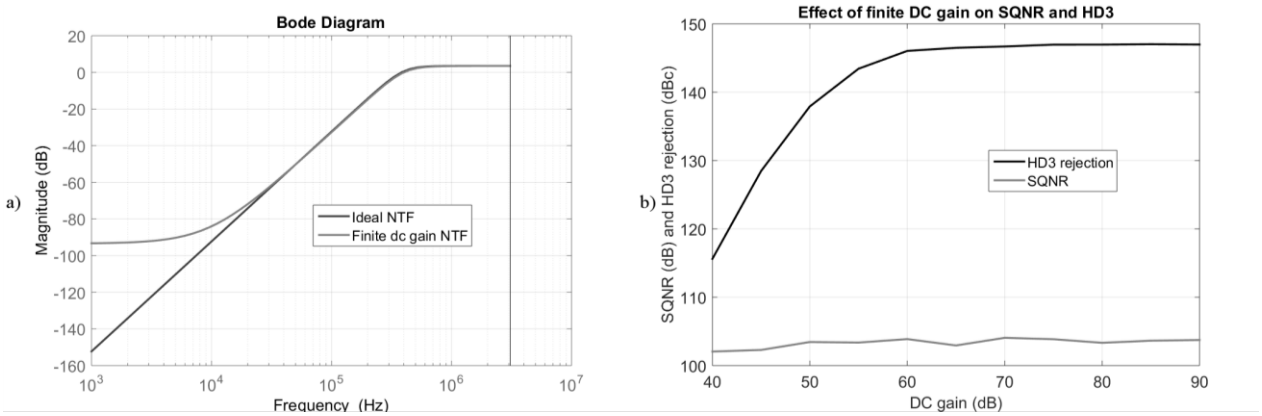


Figure 4.7 Effect of the integrators finite DC gain on a) NTF b) SQNR and harmonic performance

One important property of continuous-time modulators and one of its principal advantages over discrete-time ones is their inherent anti-aliasing filter [5, 36, 41]. In a continuous-time modulator, the input signal is not sampled at the modulator input but at the loop filter output. Then the loop filter, constituted of integrators, forms a low-pass filtering of the input signal before it is sampled by the quantizer as can be seen on the figure 4.8 a). Since the finite DC gain of the integrators affects the loop filter response and then the noise transfer function, it also affects the alias rejection property of the continuous-time modulator. The alias rejection can be evaluated by taking, on the loop filter frequency response, the amplitude attenuation between the signal frequency F_{signal} and the $F_S - F_{signal}$ frequency. The alias rejection is then frequency dependent as shown on figure 4.8 a), and has its worst value at the signal band edge. In the example modulator, the alias rejection is 91.3dB at 20kHz, while it goes up to 112.6dB for frequencies below 1kHz. The 40dB DC gain of the three integrators limits the achievable alias rejection at low frequencies, but at the band edge the alias rejection is limited by the third-order integrating shape. If higher alias rejection is required near the signal band edge, increasing the modulator order would then be the way to go.

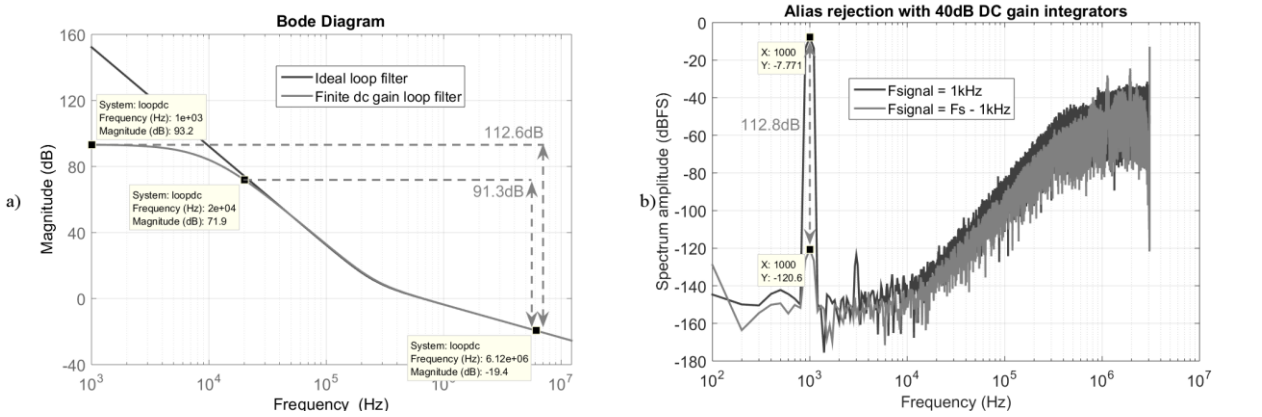


Figure 4.8 Alias rejection determined by a) loop filter frequency response b) modulator simulation

This alias rejection behavior is checked by running successively the modulator with a -3dBFS input signal at respectively 1kHz and F_S –1kHz. The output spectrums are plotted on figure 4.8 b) where an alias rejection of 112.8dB is visible too, correlating with the estimation of alias rejection made with the loop filter frequency response.

4.4 Excess loop delay: effect and compensation

In a continuous-time modulator, the timing of the feedback signal is important as it defines the feedback coefficients. If some delay is inserted, the impulse response of the feedback DAC and the loop filter will be different (delayed) and the amplitude sampled lower than expected as it can be seen on the simple example of figure 4.9. This can lead to instability since the discrete-time equivalent transfer function will not be equal to the ideal one when applying the impulse-invariance transform. The delay inserted is called Excess Loop Delay (ELD) and corresponds to the non-zero delay between the time when the loop filter output is sampled and the time when the feedback DAC switches and the signal propagates into the loop filter to the quantizer input [9, 36, 41]. For the example modulator, the finite gain-bandwidth of the amplifiers has a negligible effect since F_s is quite low and ELD is mainly determined by the quantizer regeneration time and the feedback DAC time to switch.

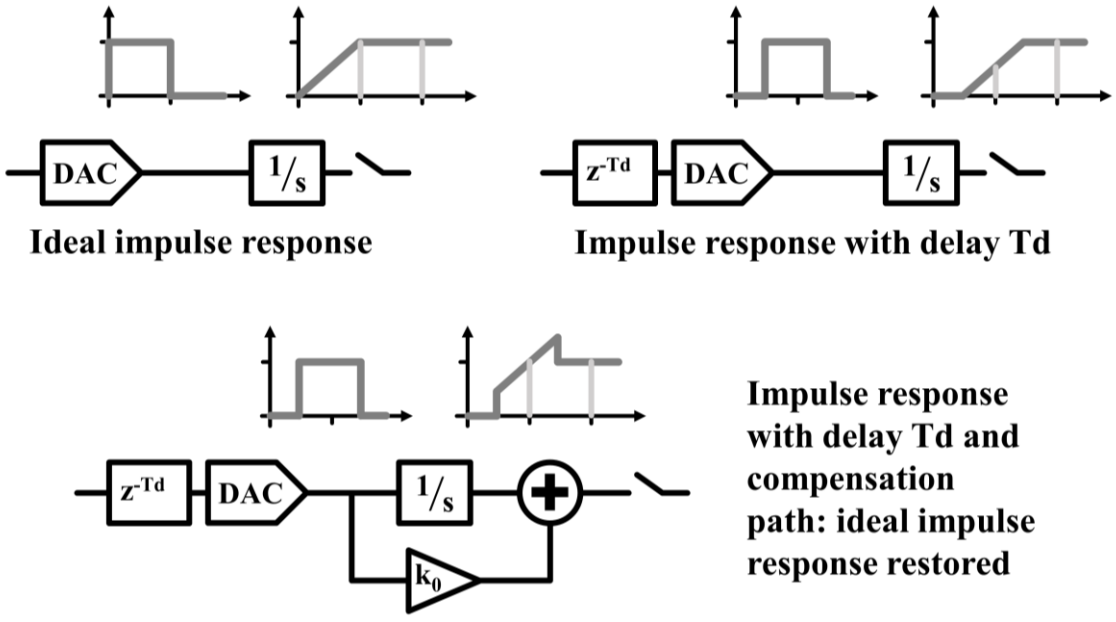


Figure 4.9 Effect of ELD on a first-order loop filter impulse response and the compensation method

Figure 4.10 a) shows the evolution of the NTF with an increasing ELD value in percentage of the sampling period. As it can be seen, the NTF exhibits a peaking proportional to the ELD value, and a maximum out-of-band gain of 2 is rapidly reached. In consequence, modulator stability is not ensured as it can be seen on figure 4.10 b) where the SQNR is slightly degraded with an increased ELD and brutally drops when instability occurs (30% ELD). One way to compensate for ELD is to use a direct path around the quantizer [43]: this is shown on figure 4.9. In that case, an amount k_0 of the feedback signal is added at the quantizer input to

compensate for the loop filter missing amplitude and the ideal noise transfer function is then recovered. The coefficient k_0 is determined in the same way than k_1, k_2 and k_3 have been defined in section 4.1 using the Impulse-Invariance Transform (IIT).

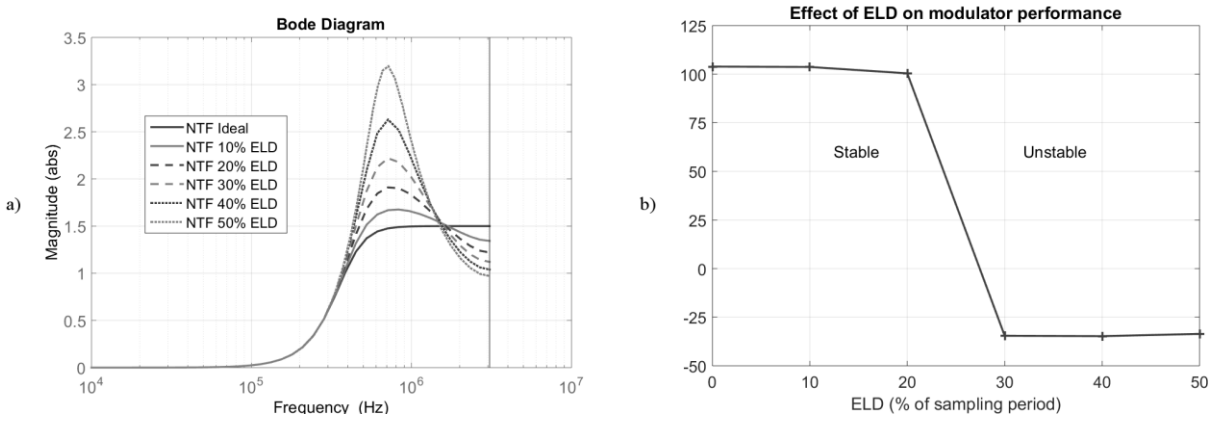


Figure 4.10 Effect of ELD on a) 3rd order NTF b) maximum reachable SQNR

Let's assume an ELD of $0.5T_S$ and work with a normalized sampling frequency of 1 second to better understand the process. Based on the diagram of figure 4.11, the values of all the coefficients must be recalculated considering ELD, since the DAC impulse response will be modified in the following way (still assuming a NRZ DAC type):

$$h_{dac}(t) = \begin{cases} 1, & for\ 0 \le t \le 1 \\ 0, & for\ t > 1 \end{cases} \xrightarrow{becomes} h_{dacELD}(t) = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ 1, & for\ 0.5 \le t \le 1.5 \\ 0, & for\ t > 1.5 \end{cases} \tag{4.17}$$

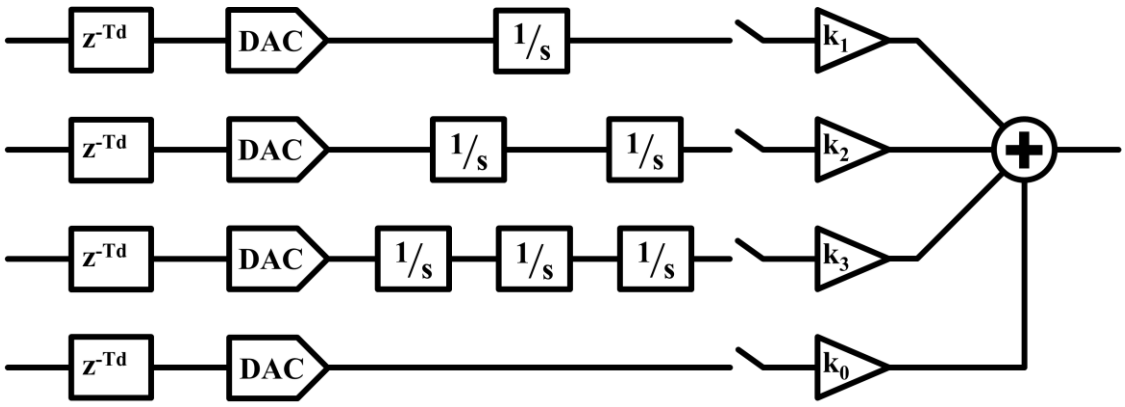


Figure 4.11 Equivalent diagram of DAC + loop filter in presence of ELD and its compensation path

Since the transfer function of the k_0 path is directly the DAC transfer function, its discrete-time equivalent $H_{0ELD}(z)$ using Impulse-Invariance Transform (IIT) turns out to be:

$$H_{0ELD}(z) = \sum_{n=0}^{\infty} h_{dacELD}(n) z^{-n} = z^{-1} \tag{4.18}$$

Using the DAC impulse response in presence of ELD, the first-order, the second-order and the third-order integrating paths transfer functions are computed using the IIT (calculations are presented in appendix):

$$H_{1ELD}(z) = \frac{0.5z^{-1} + 0.5z^{-2}}{(1 - z^{-1})} \tag{4.19}$$

$$H_{2ELD}(z) = \frac{0.125z^{-1} + 0.75z^{-2} + 0.125z^{-3}}{(1 - z^{-1})^2} \tag{4.20}$$

$$H_{3ELD}(z) = \frac{\frac{1}{48}z^{-1} + \frac{23}{48}z^{-2} + \frac{23}{48}z^{-3} + \frac{1}{48}z^{-4}}{(1 - z^{-1})^3} \tag{4.21}$$

The equivalent DACs plus loop filter transfer function is then:

$$H_{ELD}(z) = k_0 H_{0ELD}(z) + k_1 H_{1ELD}(z) + k_2 H_{2ELD}(z) + k_3 H_{3ELD}(z) \tag{4.22}$$

$$H_{ELD}(z) = \frac{\left(k_0 + \frac{k_1}{2} + \frac{k_2}{8} + \frac{k_3}{48}\right)z^{-1} + \left(-3k_0 - \frac{k_1}{2} + \frac{5k_2}{8} + \frac{23k_3}{48}\right)z^{-2} + \left(3k_0 - \frac{k_1}{2} - \frac{5k_2}{8} + \frac{23k_3}{48}\right)z^{-3} + \left(-k_0 + \frac{k_1}{2} - \frac{k_2}{8} + \frac{k_3}{48}\right)z^{-4}}{(1 - z^{-1})^3} \tag{4.23}$$

As shown in (4.23), ELD increases the order of the discrete-time equivalent DACs and loop filter transfer function. This impacts the noise transfer function which becomes:

$$NTF_{ELD} = \frac{(1-z^{-1})^3}{1+\left(k_0+\frac{k_1}{2}+\frac{k_2}{8}+\frac{k_3}{48}-3\right)z^{-1}+\left(-3k_0-\frac{k_1}{2}+\frac{5k_2}{8}+\frac{23k_3}{48}+3\right)z^{-2}+\left(3k_0-\frac{k_1}{2}-\frac{5k_2}{8}+\frac{23k_3}{48}-1\right)z^{-3}+\left(-k_0+\frac{k_1}{2}-\frac{k_2}{8}+\frac{k_3}{48}\right)z^{-4}} \tag{4.24}$$

The noise transfer function has now one more pole as it can be seen by the z^{-4} term which must be nulled to recover the expected behavior. By identification with the ideal discrete-time transfer function $NTF_D(z)$, it comes the following linear equation system and solving it gives the values of coefficients:

$$\begin{cases} k_0 + \frac{k_1}{2} + \frac{k_2}{8} + \frac{k_3}{48} = 3 - 2.2003 \\ -3k_0 - \frac{k_1}{2} + \frac{5k_2}{8} + \frac{23k_3}{48} = 1.6887 - 3 \\ 3k_0 - \frac{k_1}{2} - \frac{5k_2}{8} + \frac{23k_3}{48} = 1 - 0.4444 \\ -k_0 + \frac{k_1}{2} - \frac{k_2}{8} + \frac{k_3}{48} = 0 \end{cases} \xrightarrow[gives]{\Longrightarrow} \begin{cases} k_0 = 0.3666 \\ k_1 = 0.7979 \\ k_2 = 0.2661 \\ k_3 = 0.0440 \end{cases} \tag{4.25}$$

The ELD compensation consists of a $k_0 \times z^{-1}$ parallel path with the loop filter which translates in the noise transfer function as a k_0 coefficient present up to the z^{-4} term, helping to null it. The implementation of this compensation path requires one additional DAC path around the quantizer to create the coefficient k_0 and changes the other coefficients values like k_1 and k_2. Since these last ones are increased, higher integrators swings are expected: ELD and compensation must then be considered before scaling the integrators swings to take this effect into account. The modulator is now scaled for a fixed ELD of $0.5T_S$, allowing half the sampling period to the quantizer, the eventual DEM logic (in case of multibit modulator) and the DAC driver to evaluate and switch to the new feedback DACs values.

Thanks to the selected feedforward modulator topology, the k_0 path can directly be added to the summation node at the quantizer input. But in case of a hybrid or feedback topology, an extra

adder is needed to implement the compensation. This extra adder can be avoided by first differentiate the ELD compensation path and add it on the last integrator input [44] as shown on figure 4.12, with a coefficient k_0'.

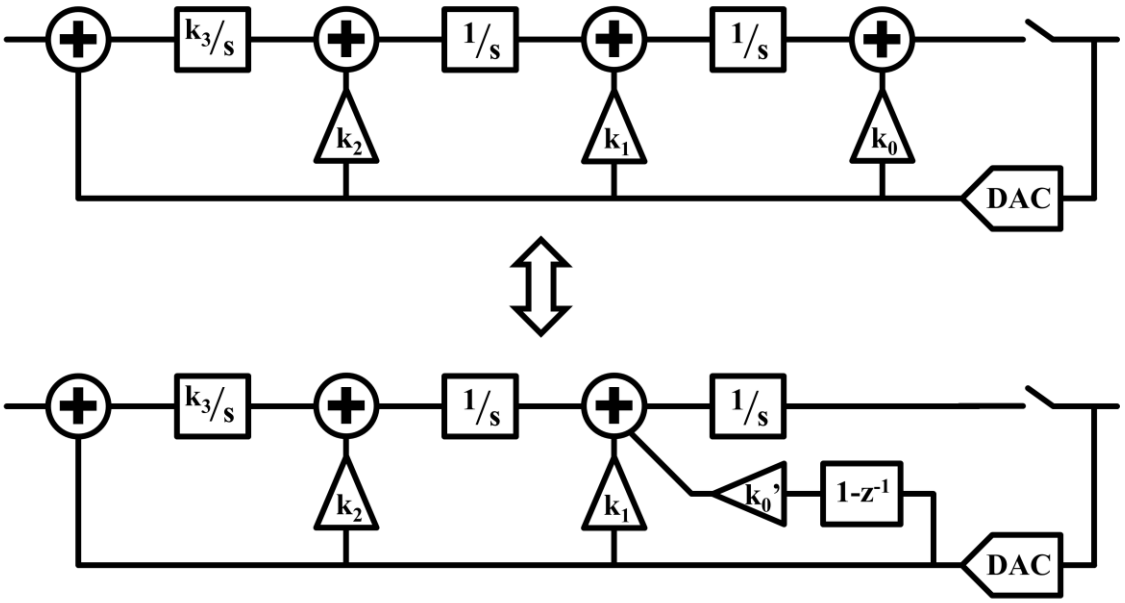


Figure 4.12 Alternative ELD compensation method to avoid an extra summation node at loop filter output

Let's now compute the value of this k_0' coefficient. To effectively compensate the ELD effect and restore the original noise transfer function, the k_0' compensation path should be equivalent to the $k_0 \times z^{-1}$ path: in that case, the z^{-4} term in the noise transfer function can be eliminated. Let's first attempt with a NRZ DAC for the k_0' path, the equivalent discrete-time transfer function is then:

$$H_{COMP} = k_0' \times (1 - z^{-1}) \times H_{1ELD}(z) = k_0' \times (0.5z^{-1} + 0.5z^{-2}) \tag{4.26}$$

As shown in (4.26), the compensation path not only creates a z^{-1} term as desired, but also a z^{-2} term which is unwanted and would introduce a z^{-5} term in the noise transfer function. This is because a NRZ DAC is used for compensation: it is not only active on one half of the sampling period but also on one half of the next sampling period too, creating a z^{-2} term. To avoid it, the compensation DAC must be active on the current sampling period only; its DAC is consequently not of NRZ type anymore. In the example case, this DAC can be active between $0.5T_S$ (ELD time) and T_S, leading to the following impulse response:

$$h_{dacCOMP}(t) = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ 1, & for\ 0.5 \le t \le 1 \\ 0, & for\ t > 1 \end{cases} \tag{4.27}$$

By integration, the impulse response of this DAC and the integrator path $h_{1COMP}(t)$ is:

$$h_{1COMP}(t) = \begin{cases} \int_0^t 0 \times dt, & for\ 0 \le t \le 0.5 \\ h_{1COMP}(0.5) + \int_{0.5}^t 1 \times dt, & for\ 0.5 \le t \le 1 \\ h_{1COMP}(1) + \int_1^t 0 \times dt, & for\ t > 1 \end{cases} = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ t - 0.5, & for\ 0.5 \le t \le 1 \\ 0.5, & for\ t > 1 \end{cases} \quad (4.28)$$

Sampling the impulse response gives $h_{1COMP}(n) = \begin{cases} 0, & for\ n = 0 \\ 0.5, & for\ n \ge 1 \end{cases}$ and the z-transform gives $H_{1COMP}(z)$:

$$H_{1COMP}(z) = \sum_{n=0}^{\infty} h_{1COMP}(n) z^{-n} = 0 + 0.5 \sum_{n=1}^{\infty} z^{-n} = \frac{0.5z^{-1}}{1 - z^{-1}} \quad (4.29)$$

Then, the total ELD compensation path $H_{COMP}(z)$ becomes:

$$H_{COMP}(z) = k_0' \times (1 - z^{-1}) \times H_{1COMP}(z) = k_0' \times (1 - z^{-1}) \times \frac{0.5z^{-1}}{(1 - z^{-1})} = k_0' \times 0.5z^{-1} \quad (4.30)$$

Equating $H_{COMP}(z)$ with the ideal compensation path $k_0 \times z^{-1}$ gives the k_0' value:

$$H_{COMP}(z) = k_0' \times 0.5z^{-1} = k_0 \times z^{-1} \overset{gives}{\Longrightarrow} k_0' = 2\ k_0 \quad (4.31)$$

The ELD compensation coefficient k_0' is now defined when the compensation is made at the last integrator input, and the noise transfer function is totally restored. Another method is to replace the last integrator by a proportional-integrating element to avoid an additional DAC [45, 46], at the expense of a peaking in STF which is generally not suitable since the main advantage of a feedback architecture is to have no peaking in the STF (versus feedforward). Digital ELD compensation is also possible but a special attention to not overload the quantizer input must be made otherwise the modulator stability is impacted [45, 47].

4.5 Clock jitter in continuous-time modulators

Timing accuracy in switching and sampling instants is a major concern in continuous-time modulators as illustrated by the previous section where the switching instant is delayed (ELD). Delay is not the single effect that can affect the sampling instants: the clock reference signal has a limited timing accuracy and its period value can weakly fluctuate between several periods: this effect is called clock jitter. The sampling period duration can be approximated (at first order) by a normal random value with a mean corresponding to its reference value and a variance corresponding to its associated jitter [9, 36, 41]. When sampling the loop filter output, the error introduced on the sampled voltage due to the timing uncertainty will be shaped by the loop filter in the same way the quantization noise is: then the clock jitter does not play a role here. However, on the feedback DAC switching instant the situation is different and the jitter changes the DAC duration on each sample. An error signal is then injected at the modulator input and will not be noise shaped by the loop filter. This error signal is dependent of several factors like the jitter variance, the DAC type or the sampling frequency.

Let's take an example of jitter effect on the NRZ DAC which follows an alternating sequence $y(n)$: its output signal can be decomposed as the sum of the ideal and error signals as done in figure 4.13.

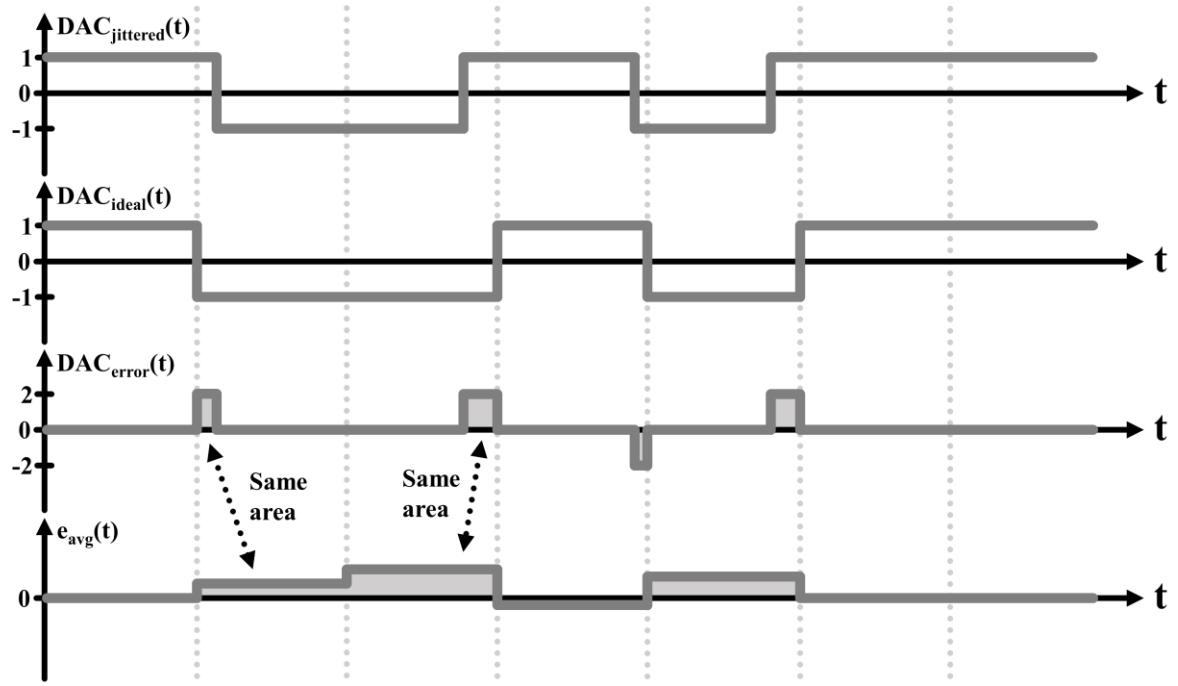


Figure 4.13 Jittered NRZ DAC waveform and determination of the one-period averaged DAC error

The error introduced depends on the difference between the DAC new value and its old value and the jitter value (symbolized by a random sequence $j(n)$); averaging it over a sampling period leads to the following equation:

$$e_{avg}(n) = \big(y(n) - y(n-1)\big) \times \frac{j(n)}{T_s} = \delta y(n) \times j(n) \times F_s \tag{4.32}$$

The power of the error signal can then be deduced together with its in-signal-bandwidth contribution:

$$\sigma_e^2 = \sigma_{\delta y}^2 \times \sigma_j^2 \times F_s^2 \overset{gives}{\Longrightarrow} \sigma_{e,inband}^2 = \frac{\sigma_e^2}{OSR} = \sigma_{\delta y}^2 \times \sigma_j^2 \times F_s \times 2B_w \tag{4.33}$$

The error signal injected at the modulator input due to jitter effect can be modeled as a random noise with a variance given in (4.33). The jitter variance is generally known as it is one main specification of clock references. To reduce the jitter sensitivity two options are available: either reduce the sampling frequency or reduce the variance of the $\delta y(n)$ signal. By reducing the sampling frequency, the jitter error will occur on a lower portion of the feedback DAC signal, making it more negligible compared to the DAC output $y(n)$ duration but this will also affect modulator resolution. The solution to reduce the variance of the $\delta y(n)$ signal is to use a multibit quantizer and feedback DAC in the modulator. Since $\delta y(n)$ is either equal to zero (when the output sequence $y(n)$ doesn't change) or equal to one LSB size (when $y(n)$ switches), increasing the number of DAC levels reduces the LSB size (see figure 4.14): it goes from 2 for a 2-level quantizer to 2/4 in case of a 5-level quantizer. The more the DAC levels are, the lower

the jitter sensitivity is. The use of a multibit quantizer and DAC to counter the jitter sensitivity has the same effect than in discrete-time modulators: the DAC non-linearity will mainly determinate the linearity of the modulator and discrete-time solutions like Data-Weighted-Averaging (DWA) to shape the introduced DAC error noise can be used [48].

Let's take a rms jitter value of 1% of the sampling period T_S and simulate the behavior of the third-order modulator with jittered feedback DAC using (4.32). The result (plotted on figure 4.15) shows a flat noise floor in the signal bandwidth (24kHz) due to the jitter effect, where the quantization noise is negligible. It gives a signal to jitter noise ratio SJNR of 51.96dB with a 2-level DAC which increases to 74.94dB by using a 15-level DAC. The jitter-induced noise attenuation is directly proportional to the LSB size reduction, since the 1/14 factor translates in a -23dB reduction, which is the amount of SJNR difference given by the simulation.

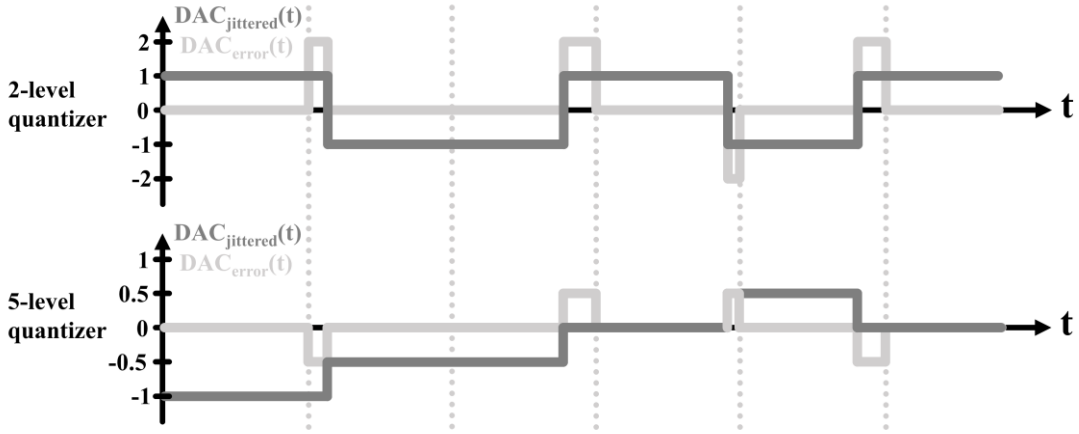


Figure 4.14 Influence of NRZ DAC LSB size on introduced jitter error amplitude

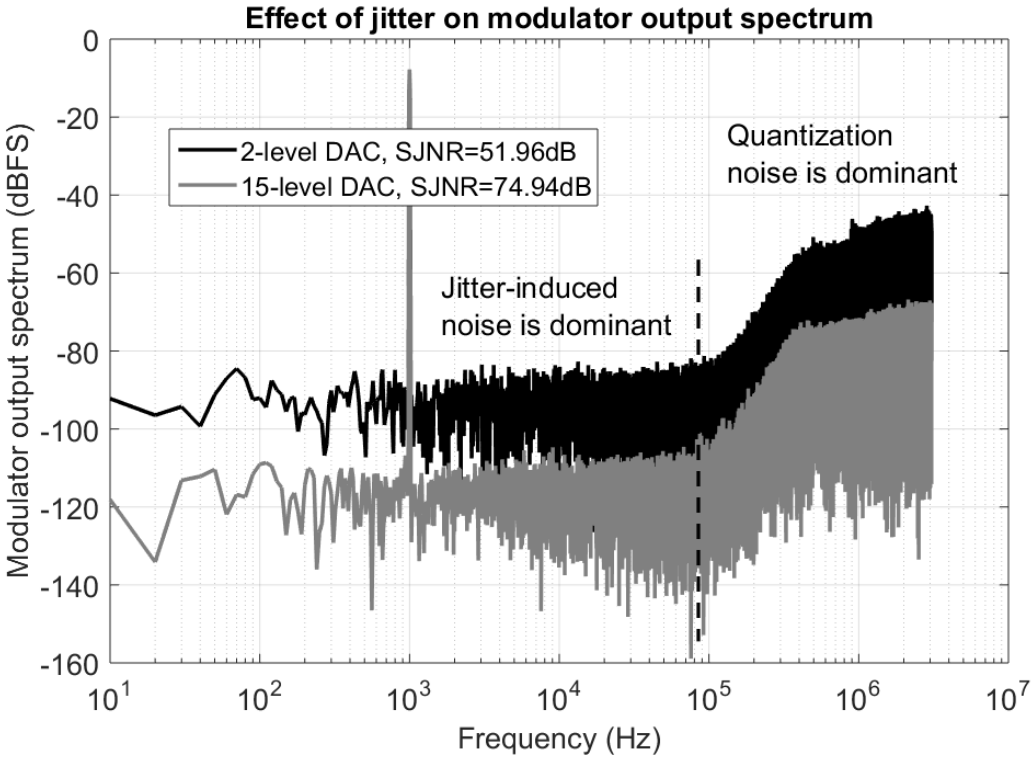


Figure 4.15 Modulator simulation results with jittered 2-level and 15-level NRZ DACs

Another benefit of a multibit modulator is its increased resolution. This extra resolution allows to reduce the OSR while maintaining a resolution equivalent to the single-bit modulator. This allows to combine both effects to reduce the jitter sensitivity: a lower sampling frequency and a higher number of DAC levels for the same resolution.

Let's now study the effect of jitter in case of a Return-to-Zero (RZ) DAC using a 50% duty-cycle. Its output can also be decomposed in an ideal response summed with an error signal, as shown on figure 4.16. The error introduced depends on the modulator output sequence $y(n)$ value, the jitter value at rising times (symbolized by $J_R(n)$) and the jitter value at the falling DAC edges (symbolized by $J_F(n)$). Averaging this error during the time of DAC activity (which is 50% of the sampling period) leads to the following error:

$$e_{avg}(n) = y(n) \times \big(J_R(n) + J_F(n)\big) \times \frac{1}{0.5T_s} = 2 \times y(n) \times \big(J_R(n) + J_F(n)\big) \times F_s \qquad (4.34)$$

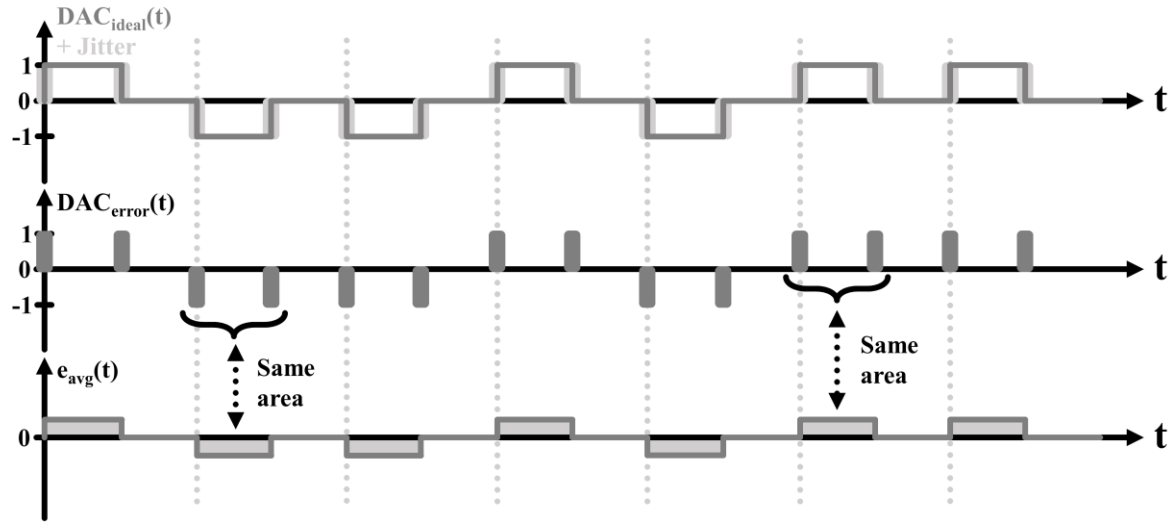


Figure 4.16 Jittered RZ DAC waveform and determination of the one-period averaged DAC error

The rise and falling jitter sequences are uncorrelated (white noise model) so their power sums up, and since they come from a single reference circuit, they have the same variance: $\sigma_J^2 = \sigma_{JR}^2 = \sigma_{JF}^2$. Then the error power can be expressed as:

$$\sigma_e^2 = 4 \times \sigma_y^2 \times 2\sigma_J^2 \times F_s^2 \overset{gives}{\Longrightarrow} \sigma_{e,inband}^2 = \frac{\sigma_e^2}{OSR} = 8 \times \sigma_y^2 \times \sigma_J^2 \times F_s \times 2B_w \qquad (4.35)$$

Reducing the jitter-induced noise power with a RZ DAC is more complicated, still the sampling frequency can be reduced at the expense of a reduced resolution. Using a multibit quantizer and feedback DAC does not help much: the error power depends on the output signal power and not on the power of the difference between two samples as in a NRZ DAC case. Then, increasing the number of DAC levels would only help at very low signal levels, since the DAC will switch from zero to a LSB value and goes back to zero at half the sampling period. But with a large input signal, the DAC will switch from zero to a large value and return to zero at half the sampling period as it can be seen on figure 4.17, keeping a large jitter error injected by the feedback DAC depending on signal amplitude.

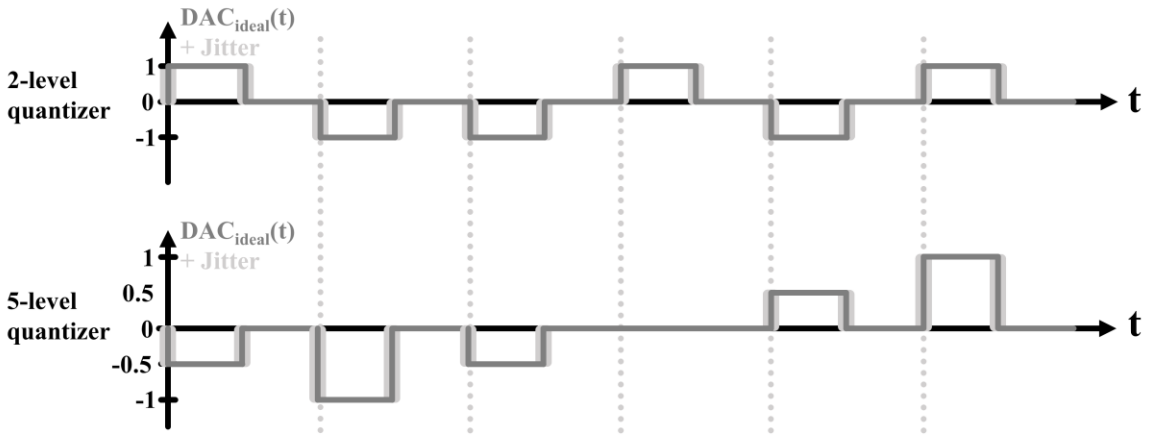


Figure 4.17 Influence of RZ DAC LSB size on introduced jitter error amplitude

To verify this behavior, simulations of the modulator with a RZ DAC have been done with jitter modeled using (4.34) and results are plotted on figure 4.18. The gain in SJNR with a multibit RZ DAC is not as important as with a NRZ DAC: the simulated SJNR for a rms jitter of 1% of the sampling period goes from 46.06dB with a 2-level DAC to 51.93dB for a 15-level DAC, correlating with the expectations deduced before.

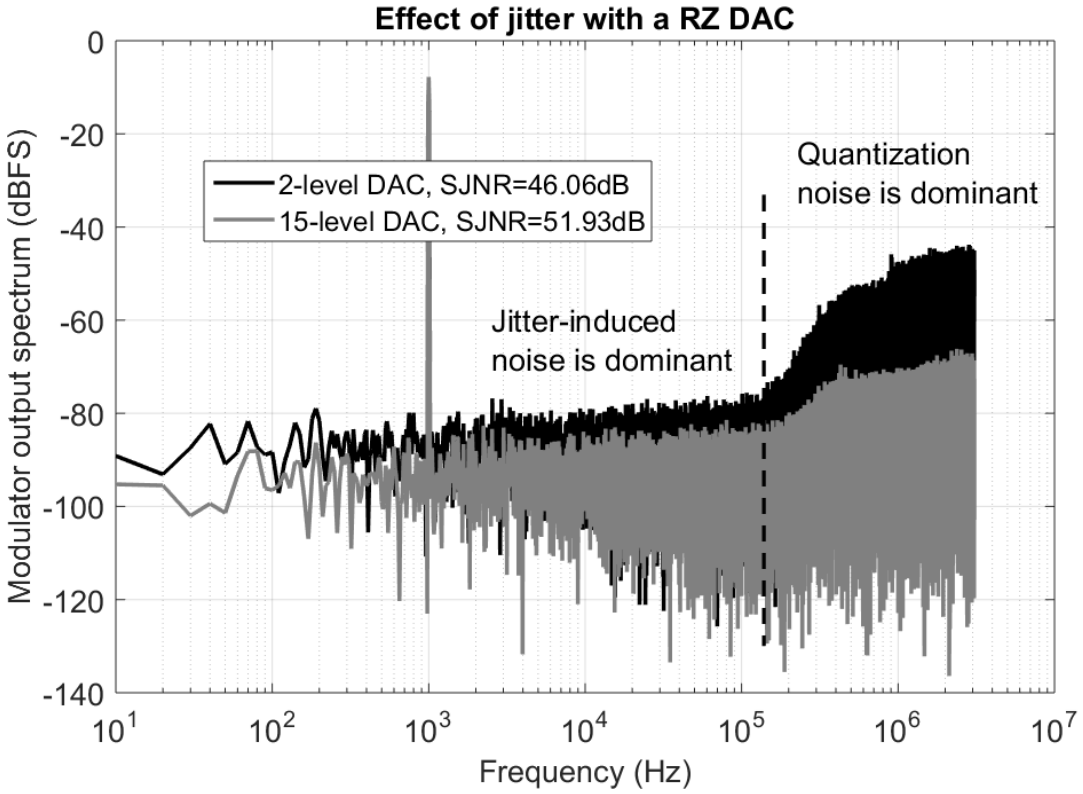


Figure 4.18 Modulator simulation results with jittered 2-level and 15-level RZ DACs

Others DAC output pulses can be used to reduce the jitter sensitivity like the Switched-Capacitor-Resistor (SCR) or the cosine shape (COS) as presented in figure 4.19. The SCR has one of the best jitter immunity [49, 50] and simple implementation but requires a higher integrator drive due to its high peak value, and can also degrade the anti-aliasing property of the modulator; that is the reason why it is not used in most of the cases since the jitter sensitivity

of a multibit NRZ DAC is low enough. The cosine shape has the best jitter immunity and a relatively low peak current compared to the SCR but requires a complex generation and the synchronization of the zero-level with the clock edge is quite difficult [51], and still have a higher peak current than a NRZ DAC.

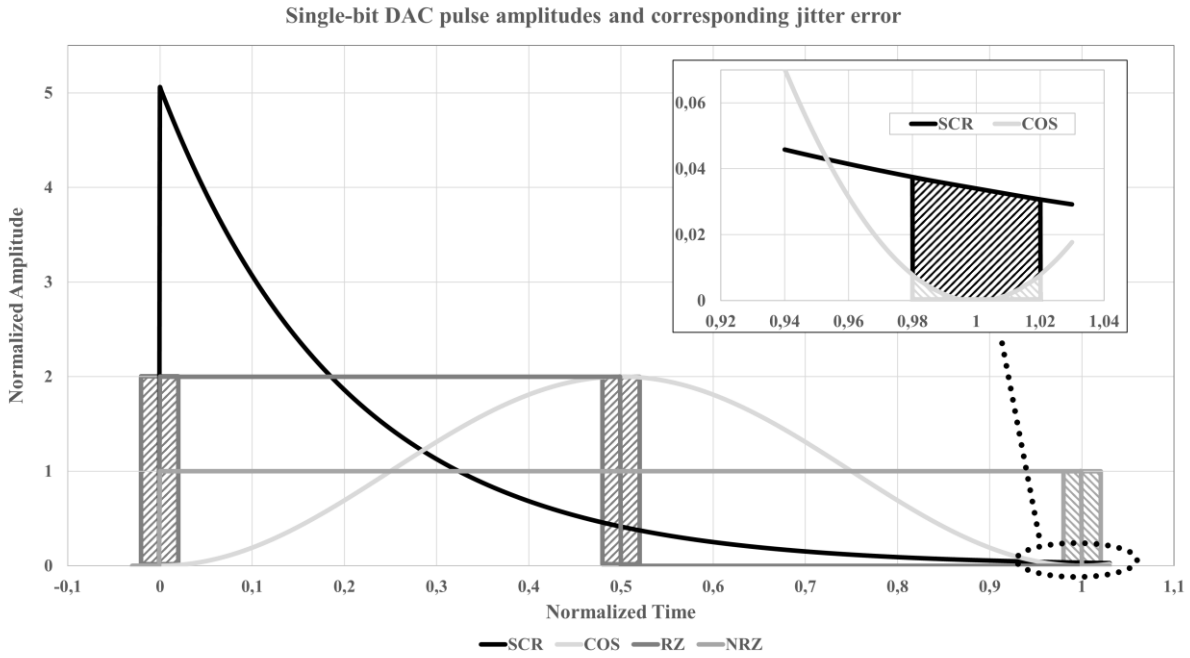


Figure 4.19 Comparison of different DAC waveforms and their corresponding introduced jitter error

4.6 FIR feedback DAC in CT modulators

Finite Impulse Response (FIR) feedback DAC can be used in single-bit quantizer modulators to add them some benefits of the multibit quantization: by filtering the quantization noise, the feedback path provides a signal tracking the input signal, on a multi-level way depending on the number of taps used in the FIR filter. Then, the error signal feeding the loop filter has a lower amplitude, reducing the integrator swings like a multibit quantization does, and also reducing the jitter sensitivity. A N-tap moving average FIR filter provides the same jitter robustness than a N level quantization, as shown by Pavan et al. in [52]. The DAC elements mismatch does not lead to a lower modulator resolution as in multibit quantization, it only moves slightly the frequency filter response which is not problematic. Then, FIR feedback DAC combines the advantages of a multibit quantization with the simplicity of a single-bit one (simple quantizer, no DEM mechanism needed). But it also has drawbacks: the FIR filter introduces in the feedback path delays of several clock periods (depending on the number of taps chosen). These delays have the same effect as ELD, modifying the equivalent discrete-time NTF and a compensation path must be added to recover the original NTF. Secondly, the single-bit quantization provides lower modulator resolution than a multibit one, but this can be compensated with an OSR increase. This section describes the compensation filter coefficients calculation to restore the NTF to the desired one, using a 4-tap FIR feedback DAC with equal

coefficients (moving-average) in a third-order modulator. The feedback path can be represented as done in figure 4.20:

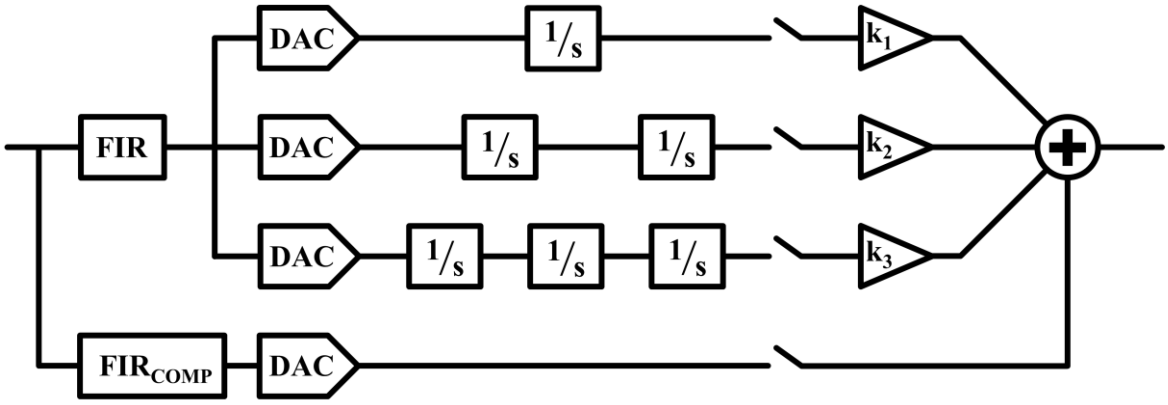


Figure 4.20 Modulator feedback path using a FIR DAC and its compensation

The FIR filter transfer function is cascaded with the first-order feedback path $H_1(z)$, the second-order $H_2(z)$ and the third-order $H_3(z)$, which transfer functions using IIT are calculated in section 4.1 and reminded below:

$$H_1(z) = \frac{z^{-1}}{1 - z^{-1}}, \qquad H_2(z) = \frac{0.5z^{-1} + 0.5z^{-2}}{(1 - z^{-1})^2}, \qquad H_3(z) = \frac{\frac{1}{6}z^{-1} + \frac{4}{6}z^{-2} + \frac{1}{6}z^{-3}}{(1 - z^{-1})^3} \tag{4.36}$$

The moving-average FIR filter and the compensation FIR_{COMP} filter transfer functions are given below:

$$FIR(z) = \frac{1}{4}(1 + z^{-1} + z^{-2} + z^{-3}), \qquad FIR_{COMP}(z) = c_0 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} \tag{4.37}$$

The complete feedback path $H(z)$ is then:

$$H(z) = FIR(z) \times \left(k_1 H_1(z) + k_2 H_2(z) + k_3 H_3(z)\right) + FIR_{COMP}(z) \tag{4.38}$$

Using expressions of (4.36) and (4.37) in (4.38) gives the following loop filter $H(z)$:

$$H(z) = \frac{c_0 + \left(\frac{k_1}{4} + \frac{k_2}{8} + \frac{k_3}{24} + c_1\right)z^{-1} + \left(\frac{-k_1}{4} + \frac{k_2}{8} + \frac{5k_3}{24} - 3c_1 + c_2\right)z^{-2} + \left(\frac{k_3}{4} + 3c_1 - 3c_2 + c_3\right)z^{-3}}{(1 - z^{-1})^3} + \frac{\left(\frac{k_3}{4} - c_1 + 3c_2 - 3c_3\right)z^{-4} + \left(\frac{-k_1}{4} - \frac{k_2}{8} + \frac{5k_3}{24} - c_2 + 3c_3\right)z^{-5} + \left(\frac{k_1}{4} - \frac{k_2}{8} + \frac{k_3}{24} - c_3\right)z^{-6}}{(1 - z^{-1})^3} \tag{4.39}$$

The ideal loop filter transfer response $H_D(z)$, which is related to the $NTF_D(z)$ given in (4.1) is:

$$H_D(z) = \frac{0.7997z^{-1} - 1.3113z^{-2} + 0.5556z^{-3}}{(1 - z^{-1})^3} \tag{4.40}$$

By identification between (4.39) and (4.40), it comes the following linear system of equations, that allows to find the modulator coefficients k_1, k_2 and k_3 and the FIR compensation filter coefficients c_0, c_1, c_2 and c_3:

$$\begin{cases} c_0 = 0 \\ \frac{k_1}{4} + \frac{k_2}{8} + \frac{k_3}{24} + c_1 = 0.7997 \\ \frac{-k_1}{4} + \frac{k_2}{8} + \frac{5k_3}{24} - 3c_1 + c_2 = -1.3113 \\ \frac{k_3}{4} + 3c_1 - 3c_2 + c_3 = 0.5556 \\ \frac{k_3}{4} - c_1 + 3c_2 - 3c_3 = 0 \\ \frac{-k_1}{4} - \frac{k_2}{8} + \frac{5k_3}{24} - c_2 + 3c_3 = 0 \\ \frac{k_1}{4} - \frac{k_2}{8} + \frac{k_3}{24} - c_3 = 0 \end{cases} \xrightarrow[gives]{\Longrightarrow} \begin{cases} k_1 = 1.0585 \\ k_2 = 0.3101 \\ k_3 = 0.044 \\ c_0 = 0 \\ c_1 = 0.4945 \\ c_2 = 0.3888 \\ c_3 = 0.2277 \end{cases} \quad (4.41)$$

Like the method presented in [52], this calculation method of the compensation filter coefficients can be used for any number of FIR filter taps, necessitating the compensation DAC to have the same number of taps than the main FIR DAC used to provide the system sufficiently degrees of freedom. Contrarily to [52], this method is valid for any type of modulator topology (here feedforward but can be feedback or hybrid as [52]), any modulator order and can be mixed with ELD compensation; making it generalized for the compensation of continuous-time modulators using FIR feedback DAC. The figure 4.21 presents the modulator output spectrum of the third-order modulator with a single-bit quantizer and DAC and with a single-bit quantizer associated with a 4-tap FIRDAC and its compensation, showing that the compensation effectively restores the modulator stability without performance degradation.

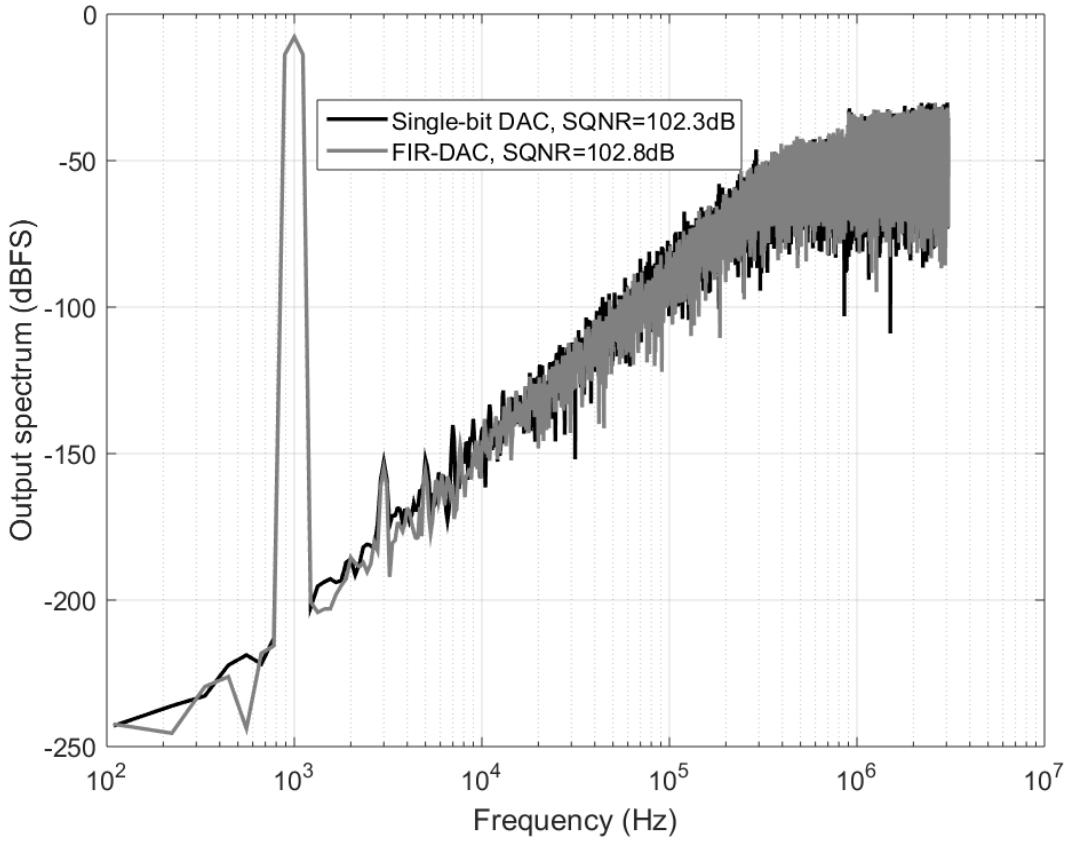


Figure 4.21 Single-bit third-order modulator output spectrums with and without FIRDAC

4.7 Conclusion

This chapter has introduced the basic concepts of the continuous-time modulator theory, starting in a first part by explaining the equivalence between continuous-time and discrete-time loop filters. Continuous-time coefficient synthesis method is presented, using the Impulse-Invariance-Transform (IIT) to make matching a continuous-time modulator with an ideal discrete-time one. Specificities of continuous-time modulators were in a second part presented: scaling for a fixed sampling frequency, impact of the integrator time constant variations, the inherent anti-aliasing property, and the effect of integrators finite DC gain on alias rejection performance. The timing non-idealities are crucial in continuous-time modulators and must be handled to reach the expected resolution. The delay between the sampling instant and the feedback DAC response, called excess loop delay, increase the NTF order and can lead to instability. This effect was studied, and the main compensation methods presented with calculations detailed on a third-order modulator example. The clock timing uncertainty, called generally clock jitter, was also investigated and several DAC pulse shapes were presented to illustrate their jitter tolerance. The NRZ DAC pulse shape with a multibit quantization is a good trade-off between jitter tolerance, robustness and anti-aliasing performance. To finish, theory on FIR feedback DAC is explored, which introduces delays of several periods in the feedback path, and a generic compensation DAC coefficient method is proposed.

Chapter 5

A 16-bit audio continuous-time sigma-delta modulator using a VCO-based quantizer

As it has been done for discrete-time modulators, let's first review several audio continuous-time designs reaching the state-of-the-art performances of around 180dB FOM_S. In [52], Pavan et al. reaches a 182dB FOM_S with the help of a FIR feedback DAC and low-power feedforward-compensated two-stage amplifiers. The FIR feedback allows the use of a simple single-bit quantizer with the benefits of a multi-bit architecture on the input stage: reduction of the integrators swings for a lower power consumption and lower sensitivity to clock jitter. No complex quantizer and DEM circuits are needed, saving power compared to a multi-bit architecture, but needs a higher OSR to reach the same resolution due to its single-bit quantizer. In [53], a version with chopper is presented, allowing to reduce the 1/f noise and the input stage amplifier area, while reaching an 183dB FOM_S, and an extended chopper and FIR feedback DAC theory is given in [54]. In [55], Leow et al. present third-order modulator using a second-order loop filter, a low-power 5-bit SAR quantizer and the noise coupling technique to add one order of noise shaping. Using a SAR quantizer, at the end of the quantization process, the quantization noise is stored on the SAR DAC capacitors and is used with a delay to be subtracted in the next sample quantization; the quantization noise is then filtered with a $1-z^{-1}$ transfer function, adding one more order of noise shaping. This solution reaches a FOM_S of 178dB. In [40], De Berti et al. propose a third-order modulator with a 4-bit flash quantizer reaching a high DR of 106dB. The modulator adapts dynamically its noise level to the signal range to achieve a larger DR. To make this, tri-level current-steering DAC unit elements are used: at low signal level, only one single DAC unit element is used, and the others are disconnected from the modulator input (state 0 of the tri-level DAC unit). When the input signal increases, the quantizer output (through decoding logic) activates the other DAC unit elements and the input-referred noise increases too. Using a single amplifier to implement a second-order resonator, which benefits the power consumption, this design reaches 180dB of FOM_S.

A new type of quantizers, based on VCOs, are introduced in high bandwidth continuous-time modulators [56, 57] to face their speed issue. For these high-speed continuous-time modulators benefiting from technology scaling, flash quantizers complexity increases as lower supply voltages must be used making them power-hungry while SAR quantizers become not fast enough. In the same time, frequency and time resolution reach new standards. For these reasons, replacing the voltage-based quantizer with a time-based one becomes more and more suitable [58, 59]. While keeping the multibit quantization advantages in a continuous-time sigma-delta modulator (reduced jitter sensitivity and internal swings, increased dynamic range), the VCO-

based quantization adds extra benefits: the ring-inverter-oscillator implementation using digital cells gains from technology scaling (area and power), an extra order of noise shaping can be obtained for free, and some implementations provide an intrinsic Dynamic Element Matching (DEM) function for the feedback DAC [60]. Moreover, due to their pseudo-digital signals, quantization can be made on a clock edge, avoiding the regeneration time of voltage-based comparators, thus reducing the modulator ELD. The VCO-based quantizer allows to significantly simplify the modulator architecture, making it well suitable for a low-power audio modulator. Nevertheless, a VCO-quantizer brings also its specific drawbacks: the linearity of a ring-oscillator based VCO is poor, its specifications (idle frequency, voltage-frequency coefficient) are Process, Voltage and Temperature (PVT) sensitive, and the interface with the rest of the loop filter requires a special attention. Due to their poor linearity, VCO-based quantizers are often used in high-speed continuous-time modulators which have less stringent requirement on linearity and resolution than audio ones. In this chapter, an implementation of a third-order continuous-time sigma-delta modulator for an audio application is proposed. Using a VCO-based quantizer, its linearity and PVT sensitivity will be under special attention to achieve a 16-bit resolution.

5.1 Principle of a VCO-based quantizer

Two types of VCO-based quantizer can be used depending on the output variable selected: the frequency or the phase position. Let's first study the properties of a frequency output with the help of figure 5.1. A ring-oscillator VCO converts the input voltage into a frequency variation through the gain K_{VCO}. Since the instantaneous value of the frequency is difficult to extract, the phases positions are read at the sampling rate F_S. The position of the VCO phases rotates at a speed proportional to the VCO instantaneous frequency: the phase is proportional to the integral of the frequency. Once sampled, phase positions are first-order differentiated digitally to recover a frequency output. The frequency is quantized on N levels, depending on the number of phases used (3 in figure 5.1). These quantizers are the simplest implementation: open-loop systems, they can reach very low-power values but suffer from the VCO non-linearity and have a poor harmonic performance. Since the quantization is made on the phase, the digital differentiation provides a first-order quantization noise shaping.

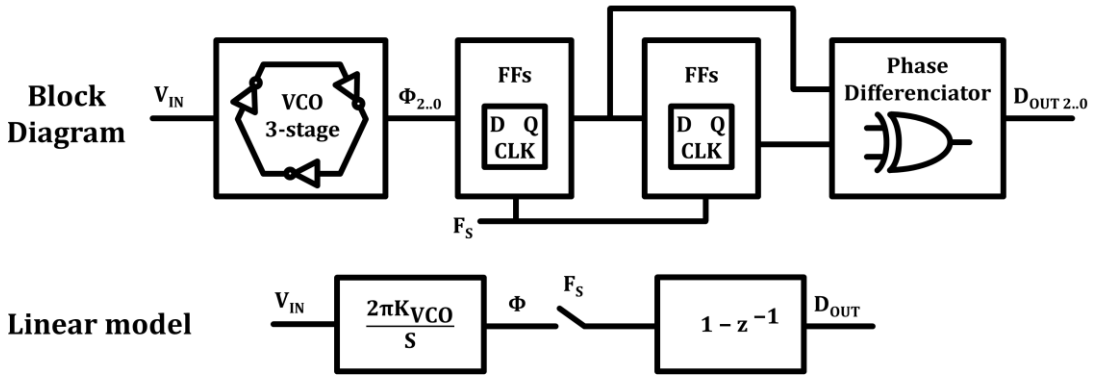


Figure 5.1 Frequency output VCO-based quantizer block diagram and linear model

Differential realization allows to suppress even-order harmonics [61], but still constrains these quantizers to high-speed and low-resolution converters. Embedding it in a continuous-time loop helps to reduce the VCO non-linearity effect due to the high loop gain provided by the feedback loop, and the main DAC benefits from a free Data-Weighted-Averaging (DWA) mechanism since the quantizer output is a first-order differentiation [56, 62] but the final linearity performance is not good enough to target 16-bit resolution.

The operation principle of a differential VCO-based quantizer with a phase output is presented with the help of figure 5.2. Two VCOs convert a differential input voltage into a frequency difference through a gain value K_{VCO}. Since the instantaneous value of the frequency is difficult to extract, the phase position is read at the sampling rate F_S. The advantage of working with the phase is the integration operation (integral of the frequency) which provides the quantizer a first order noise shaping. The phase difference is then quantized with a selected number of quantization values equally distributed on [0, π] or [-π, π] interval, depending on the implementation of the phase decoder. A negative feedback is needed to ensure that the VCO phase difference will be kept in the selected quantization range. The VCOs are generally made by inverter delay cells in a ring-oscillator configuration for its simplicity. Here 7-stage ring oscillators are represented, leading to 7 output phases for each VCO. A bank of flip flops samples the state of each phase of both VCOs at the end of the sampling period. A phase decoder composed of XOR cells compares each output phase of VCO_P with its corresponding output phase of VCO_N (Φ_{PS0} with Φ_{NS0}, Φ_{PS1} with Φ_{NS1} and so on); generating a thermometer output code of the phase difference between the two VCOs. Working with the phase necessitates a feedback operation, the latter reduces the swing at the VCOs input: the more the phases used, the lower the VCOs input swing will be, thus enhancing the harmonic performance. Differential operation also helps to provide common-mode rejection, even order harmonic suppression and double the input range. The differential operation allows to work on the phase difference between the two VCOs: their idle frequency can be minimized and does not need to be tuned accurately, which is a good point toward PVT robustness [60].

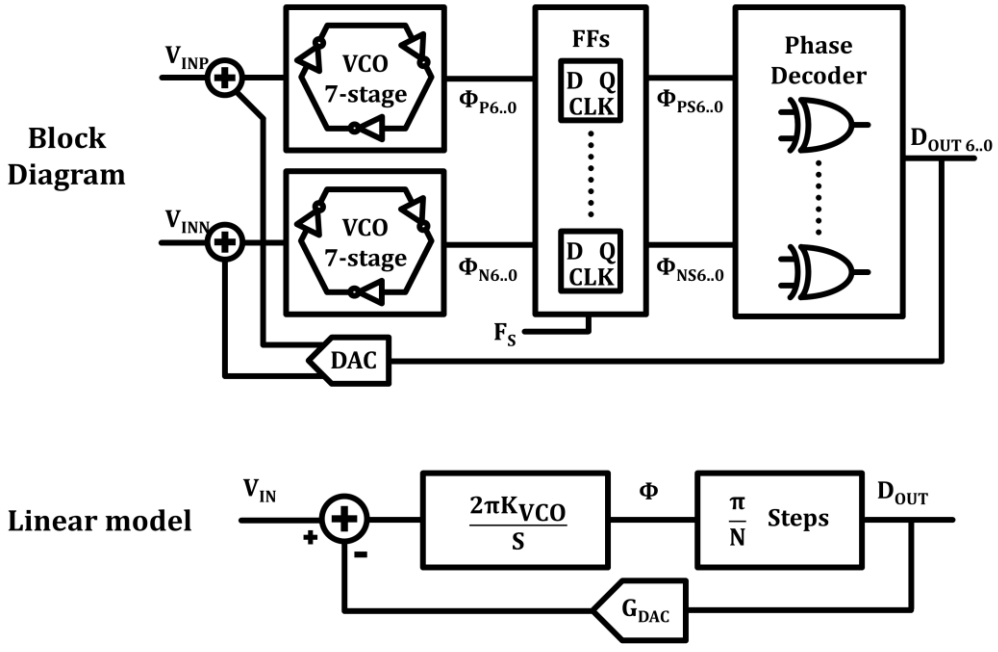


Figure 5.2 Block diagram of a differential VCO-based quantizer with a phase output

Figure 5.3 helps to describe how the XOR-based phase decoder works using 7-stage ring VCOs which are driven by their supply voltage. Inverters highlighted with red dotted boxes are the ones in transition state (their gates have just switched but not their output yet - reflecting the gate propagation time). Figure 5.3 a) shows the case where a zero differential input is applied. The two VCOs are in the same state: their 7 phases are equal, leading to a zero output word on the XOR-based phase detector (the phase difference is $\pi \times 0/7$). Figure 5.3 b) shows the case when VCO_P runs faster than VCO_N: its transition state is in advance by two inverters compared to VCO_N. The phase difference between the two VCOs is then $\pi \times 2/7$. If all the phases of the VCOs are in opposite state, the phase difference will be π. Figure 5.3 c) represents the case where VCO_N is leading VCO_P. Here the transition appears four delay cells before the one of VCO_P, leading to a phase difference of $\pi \times 4/7$. One thing to note is that the XOR-based phase decoder is not able to differentiate if VCO_P is faster than VCO_N or inversely. To conclude, the XOR-based phase decoder quantizes the absolute value of the phase difference in an interval $[0, \pi]$ with a step of π/N where N is the number of used phases. The output is directly a thermometer coding of the phase difference, leading to N+1 quantization levels. In [63], the XOR-based phase decoder is completed with some logic gates to detect the sign of the phase difference between the two VCOs. This minor modification extends the quantization range from $[0, \pi]$ to $[-\pi, \pi]$, while keeping the same number of stages in the VCOs. The quantization step of π/N remains unchanged, leading to 2N+1 quantization levels on the interval $[-\pi, \pi]$. This makes an efficient hardware use since with N delay cells in the VCOs, 2N+1 quantization levels can be generated, reducing further the VCOs input swing and then increasing the harmonic performance. The modulator presented in [63] (which is an amelioration of [60]) reaches up to 83dB SNDR which is the highest value for a VCO-based only modulator, making it very suitable to be included in a higher-order continuous-time sigma-delta loop.

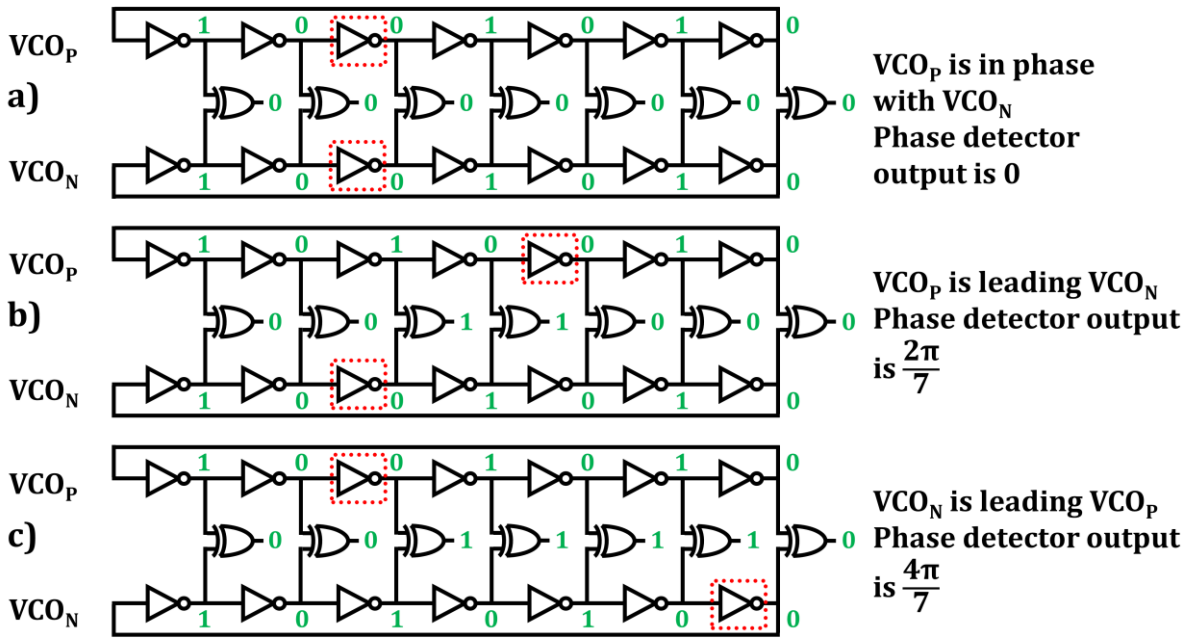


Figure 5.3 XOR-based phase decoding principle a) VCOs in phase b) VCO_P leading c) VCO_N leading

The sign extraction principle used in [63] is explained with the help of figure 5.4. In this example, a constant negative input is applied leading to a -2π/7 phase difference. Looking closer at the XOR-based phase decoder behavior, one can note the following: the rising edge in the phase decoder word indicates the element number where the VCOs enters in opposite phase while the falling edge of the phase decoder word indicates the element number where the VCOs returns in phase state. These two remarkable elements correspond to the inverters in transition in each VCO. Then, for a negative phase difference sign, the VCO_P transiting element coincides with a rising edge in the phase decoder output word, and the VCO_N transiting element coincides with a falling edge in the phase decoder output word; while the opposite occurs for a positive phase difference. A simple logic gates combination testing one of the two affirmations is sufficient to detect the sign of the phase difference.

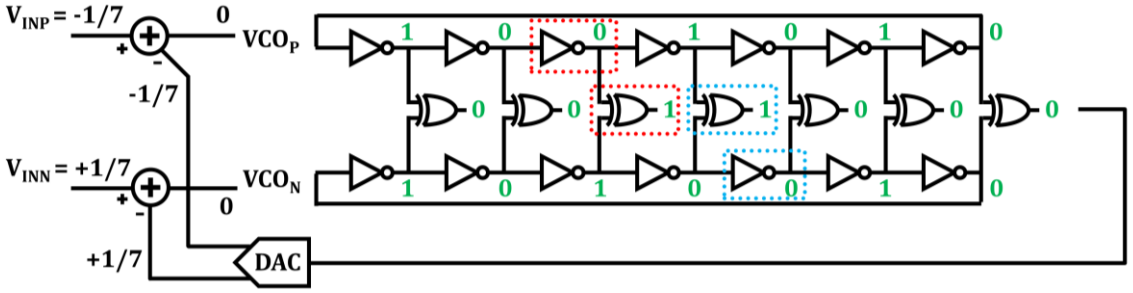


Figure 5.4 Phase difference sign extraction principle: inverters in transition state coincide with rising / falling edges in phase decoder word

Frequency output VCO-based quantizer provides an intrinsic DWA for a feedback DAC due to their digital differentiation which first-order noise-shapes the DAC mismatch errors. In [60,63], authors claim that an intrinsic DEM mechanism is also present using the XOR-based phase decoder and a differential VCOs architecture. To illustrate this DEM mechanism, let's still work with the figure 5.4. Assuming a constant differential input voltage of -2/7 times the voltage reference, the VCO_P frequency will decrease while the opposite occurs for VCO_N, until the phase decoder reaches -2π/7 on its output. In that case, the feedback DAC will null the differential input of the VCOs and they will return to their idle frequency. The system reaches a stable steady-state operation where both VCOs run at the same frequency, with a constant -2π/7 phase difference. The VCOs have the same but non-zero frequency: their phases still continue to propagate with a constant difference. The phase decoder keeps outputting -2π/7 in a thermometer coding, but a rotation in the elements in state 1 appears. Since each phase directly controls a DAC element unit, a rotative selection pattern of the DAC elements is made, following the VCOs idle frequency. DAC mismatch errors will then be modulated around VCOs idle frequency.

This behavior is acceptable in low / medium resolution converters but embedding this quantizer in a high-resolution sigma-delta loop shows the limitation of this DEM mechanism as illustrated in figure 5.5. The reason is that each DAC unit element is controlled by one phase signal of the quantizer. The latter can be seen as a PWM signal running at the VCO idle frequency and where quantized signal information is coded in the PWM modulation. The issue occurs when sampling this PWM signal: in-band noise and tones aliases appear. When no DAC mismatch are present, the sum of each phase signal aliases cancels since the phases signals are equally distributed on the unit circle and contain the same information. But with DAC mismatches, this cancellation does not occur anymore, making the aliases due to PWM sampling to appear into the signal

band. Then, the VCO-based quantizer proposed in [63] is still a good candidate to be embedded in a high-order sigma-delta loop, for its linearity performance and efficient hardware use, but will necessitate the use of an explicit DWA mechanism to noise-shape the DAC mismatch errors and reach the expected 16-bit resolution.

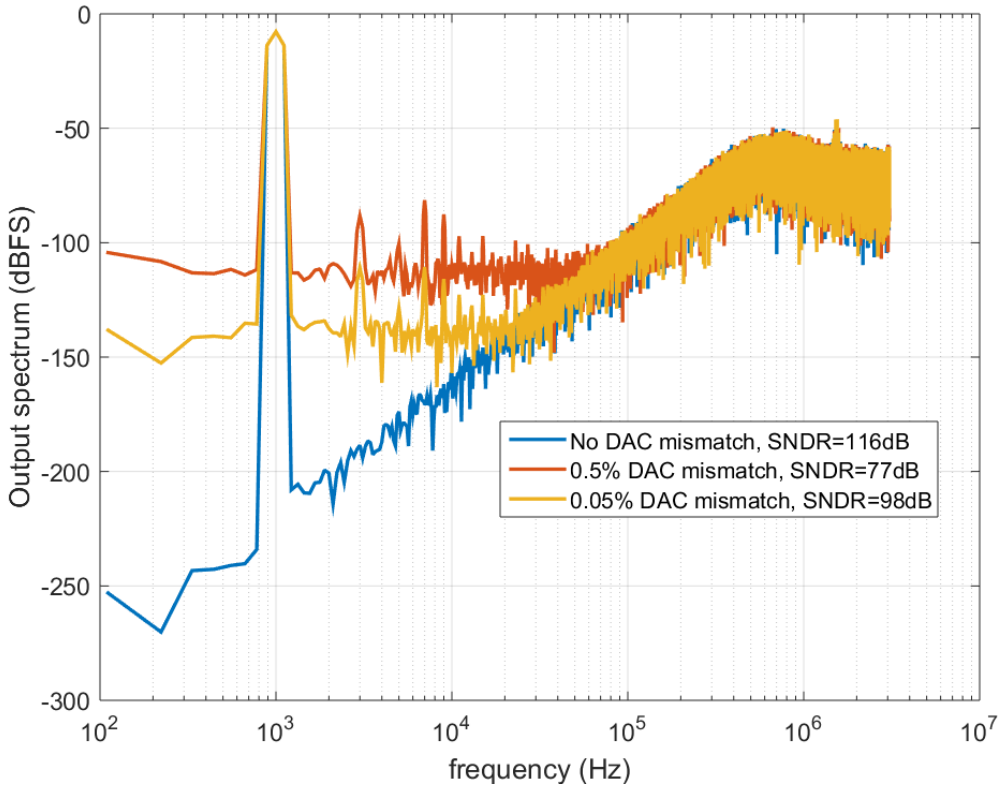


Figure 5.5 Limitation of the [63] intrinsic DEM mechanism when high resolution is targeted

5.2 Modulator specification and definition

For this continuous-time modulator, the dynamic range specification is 98dB, a little higher than 16-bit resolution (96dB). Let's reach this value for a -3dBFS peak input sinewave leading to an expected 95dB peak SNR. The PGA placed before in the system will adapt the signal range of the microphone to the ADC full scale. As for the discrete-time modulator, a low-noise LDO regulator will be used to provide a reference voltage of 1.5V. The targeted FOMs is 180dB, this gives for a 98dB DR and a 24kHz bandwidth a power budget of 150µW. In an analogue way as it has been done for the discrete-time modulator, the input-referred noise level is calculated to be 13µVrms. Finally, to make the thermal and flicker noises the dominant modulator noises contributors, let's have 20dB margin between these noise sources and the quantization noise leading to about 115dB SQNR requirement. Compared to the discrete-time modulator, larger margin is taken here due to the higher integrators time constants variability versus PVT, which can degrade the noise shaping performance as seen in chapter 4. This effect is further amplified with the use of a VCO-based integrator, highly sensitive to PVT. Since the targeted application must be low-cost, a robust architecture is preferable than tuning and calibration which will cost test time. The modulator specifications are summarized in table 5.1.

Table 5.1 Modulator specifications used to define architecture

Specification	Value	Unit
Input full scale	1.5	V
Signal bandwidth	24	kHz
Power consumption	150	µW
Input-referred noise	13	µV
Dynamic Range (DR)	98	dB
Signal to Noise Ratio (SNR)	95	dB
Signal to Noise and Distortion Ratio (SNDR)	>90	dB
Signal to Quantization Noise Ratio (SQNR)	115	dB
FOMs	180	dB

The selected modulator architecture is a third-order one in a feedforward loop topology as presented in figure 5.6 a). Direct feedforward path from input signal to quantizer input is not used since it will suppress the anti-aliasing property of the continuous-time modulator. A 4-bit (15 levels) VCO-based quantizer is used to reduce the internal swings (lower power consumption in the loop filter) and to increase the clock jitter immunity using a NRZ feedback DAC scheme. Since the VCO-based quantizer uses a phase output, it needs a feedback path at the VCO input: the modulator topology has to be changed from a feedforward one to a hybrid feedforward/feedback as illustrated in figure 5.6 b). The NTF remains unchanged while the STF has a lower peaking due to the presence of one feedback path [52]. The second order loop filter provides sufficient gain to noise-shape the remaining VCO non-linearity, since the VCO-based quantizer can reach up to 83dB SNDR itself. To have a robust architecture against PVT variations, which makes varying the integrator time constants (RC product for the loop filter, K_{VCO} gain for the last integrator), a low out-of-band gain NTF is selected: 1.5. This helps to ensure stability together with the 4-bit quantizer, and the OSR is fixed to 128 to obtain sufficient quantization noise shaping.

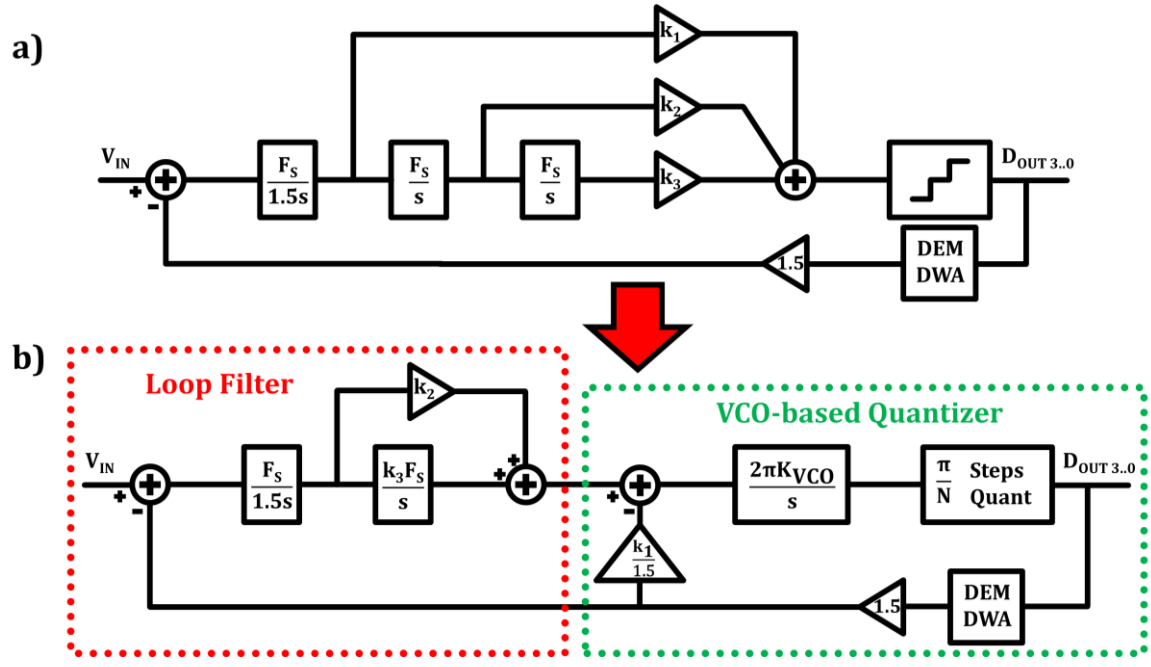


Figure 5.6 Continuous-time modulator linear model a) feedforward topology b) hybrid feedforward-feedback topology to insert a VCO-based quantizer

The k_1, k_2 and k_3 coefficients are determined by the procedure defined in chapter 4 that gives values of resp. 0.6703, 0.2442 and 0.044. The K_{VCO} value determines the time-constant of the last integrator (which should be F_S) and is defined as follows: since the phase quantizer has a [-π, π] range and not a [-1, 1] one, a π gain has to be inserted to scale the signal according to the quantization range leading to a $\pi \times F_S$ time constant needed for the last integrator. The voltage to phase gain provided by the VCO is $2\pi \times K_{VCO}$, thus providing already the needed π factor. It results that K_{VCO} should be equal to $F_S/2$ to provide the correct last integrator time constant.

With the help of a high-level model of the defined modulator one can check the architecture robustness against several non-idealities. With the selected modulator architecture, K_{VCO} can vary from $F_S/4$ to F_S without stability issue or performance degradation (minimum SQNR of 110dB), ensuring the modulator robustness against VCO PVT sensitivity.

The second-order loop filter in a feedforward configuration allows the use of a single-opamp loop filter as it will be detailed in the next section. This lowers the modulator power consumption, but the specifications of the opamp are more stringent like its DC gain, since the DC gain of two amplifiers is not cascaded. Thanks to the help of the high-level model, finite DC gain and GBW can be simulated to see their impact on the modulator performances. This leads to the figure 5.7 which shows that a low GBW amplifier (in order of the F_S or below) necessitate around 80dB DC gain to reach modulator performance target, while doubling the amplifier GBW reduces the DC gain requirement to 60dB. Keep doubling the GBW does not help much, since the required DC gain is somewhere around 55dB. As a trade-off power consumption versus reasonable DC gain, the design point of 60dB DC gain and $2F_S$ GBW is selected.

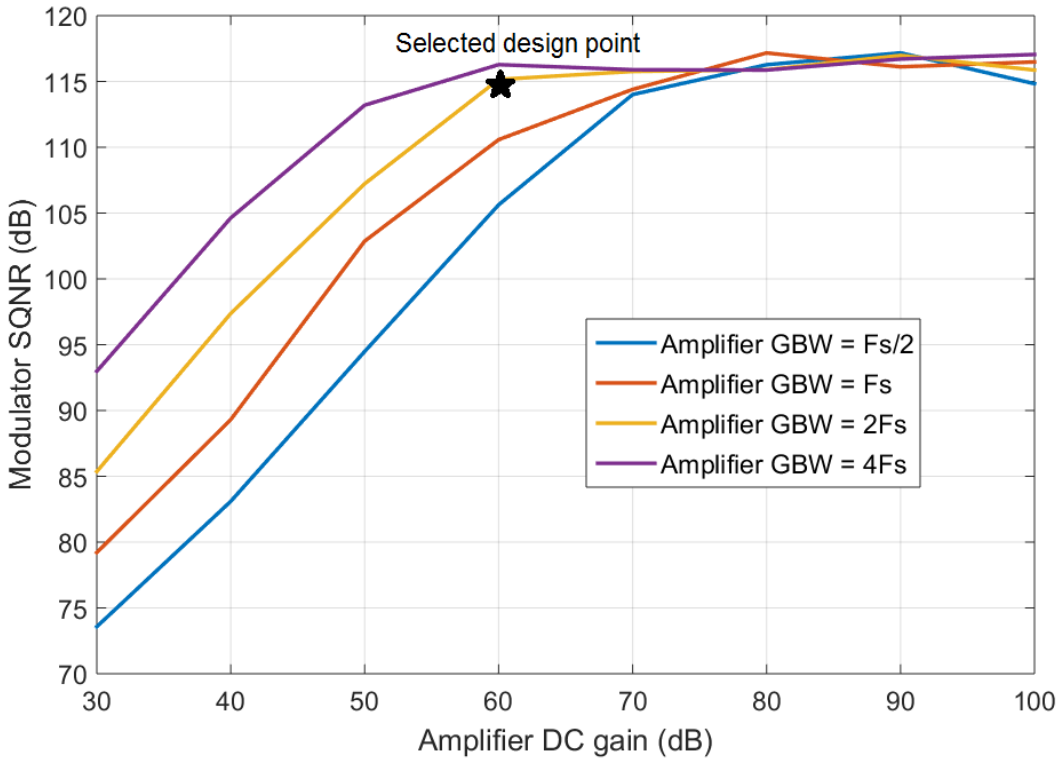


Figure 5.7 Modulator SQNR for different loop filter amplifier DC gain and GBW values

The alias rejection with finite DC gain is also determined by running simulations with a -3dBFS sinewave at 1kHz and F_S –1kHz and 20kHz and F_S –20kHz. Figure 5.8 shows the simulation result for the alias rejection at 1kHz, which is as high as 163dB. This high value (and increased rejection compared to the 3-rd order modulator of chapter 4) is explained by two factors: first, the VCO-based integrator is exploiting the mathematic relation of the phase position being directly proportional (with a 2π factor) to the frequency integral, thus providing a perfect integrator with infinite DC gain. Second, the hybrid feedforward-feedback architecture, that filters the signal before it is sampled, is at least a cascade of 2 integrators, while in a feedforward-only architecture, the filtering is only of first-order at high frequencies. In the same way, the alias rejection at the audio band edge (20kHz) is simulated to be 115dB, the alias rejection performance of the modulator is then high enough on the whole signal bandwidth to not impact the modulator performances.

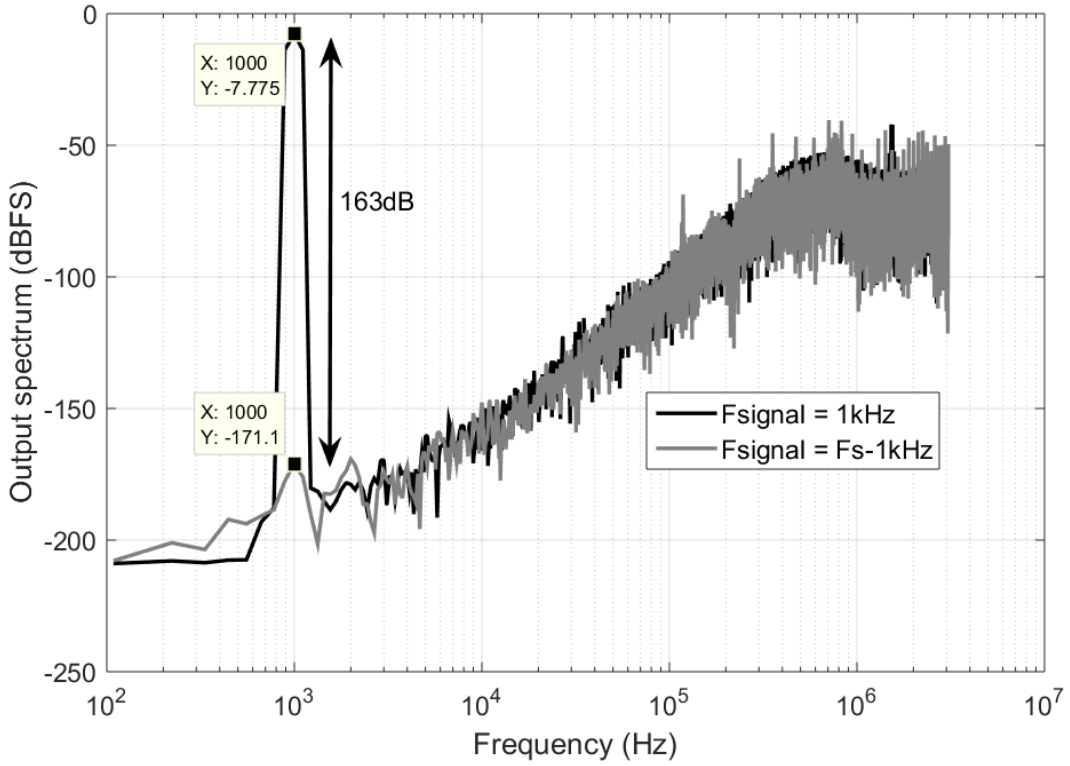


Figure 5.8 Simulated alias rejection a 1kHz for the hybrid feedforward-feedback topology

Concerning the ELD tolerance, simulations show that the modulator is able to support a delay up to 50% of the clock period without stability issue as shown on figure 5.9. Thanks to the VCO-based quantizer, which provides almost instantaneous data on a sampling edge, only the decoding and DWA logics have to operate before the feedback DAC can switch. Then, a low delay of 15-20% of the clock period can be used, and no need to be compensated, simplifying the modulator architecture. The jitter sensitivity is also investigated, the modulator architecture uses a 15-level feedback DAC with a NRZ switching scheme to reduce it, and simulations are then run to define the modulator jitter tolerance level. Adding the expected input-referred noise level of 13µVrms in the simulations, and using the jitter model developed in chapter 4, modulator SNDR is given as function of the clock jitter level in figure 5.10. As visible, jitter can be as high as 25ps rms before starting to degrade the modulator performances.

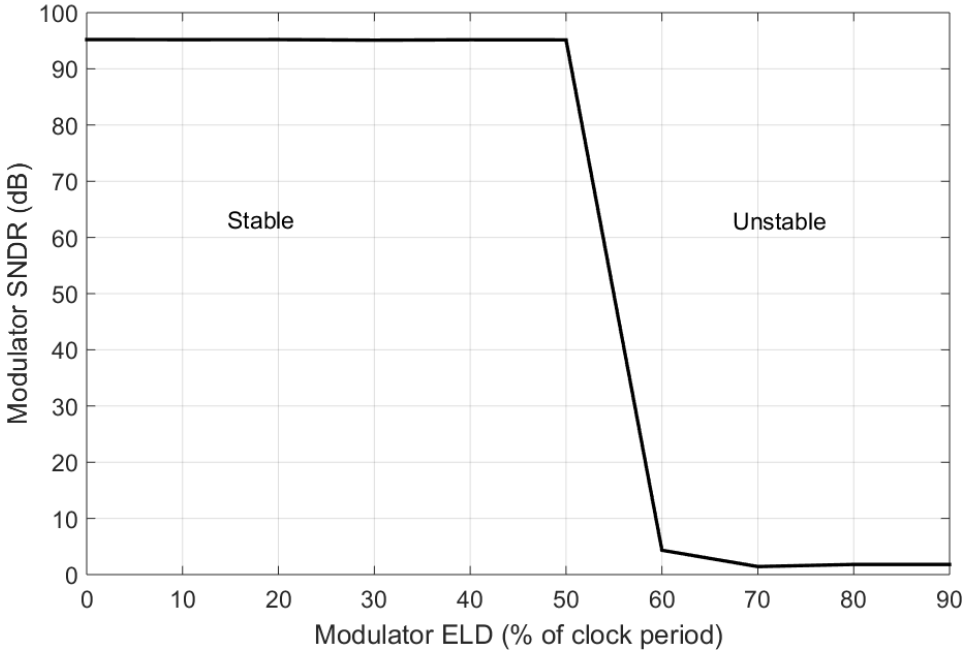


Figure 5.9 Modulator SNDR as function of the ELD in percentage of the clock period

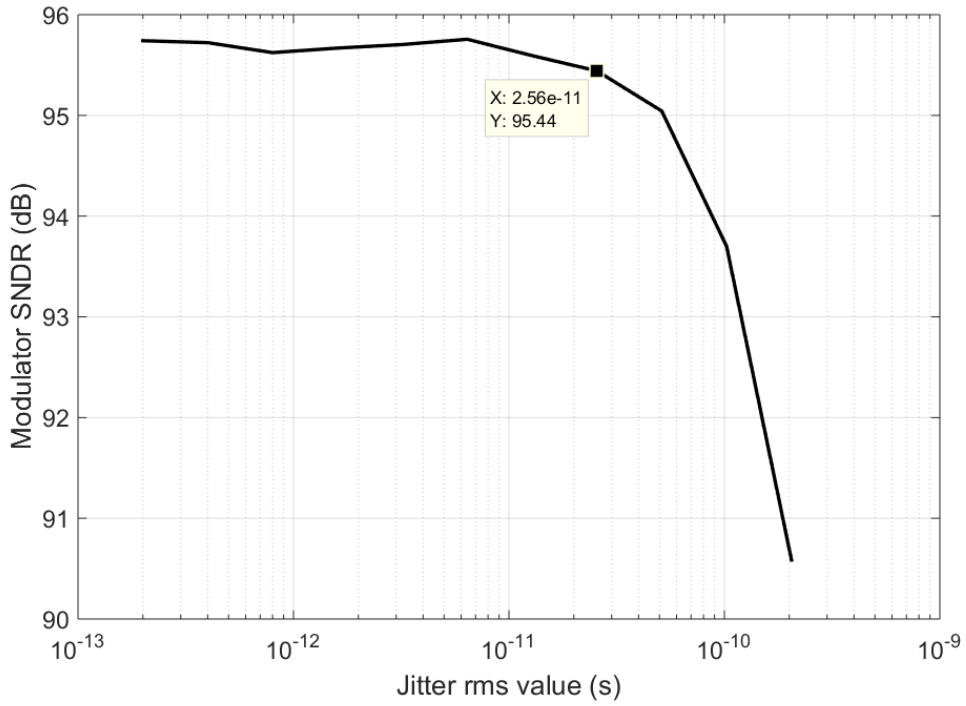


Figure 5.10 Modulator SNDR as function of the clock jitter rms value

The modulator will use the Data-Weighted-Averaging (DWA) algorithm to first-order noise-shape the DAC errors introduced by mismatch in its elements. Implementing this algorithm in the high-level model allows to define a matching requirement on the DAC elements before starting the design. Random DAC elements values are generated, using a normal distribution centered on the ideal element value and having a standard deviation σ representing the matching deviation. Simulation results are plotted on figure 5.11 where we can note that a σ higher than 0.5% starts to impact the modulator SNDR.

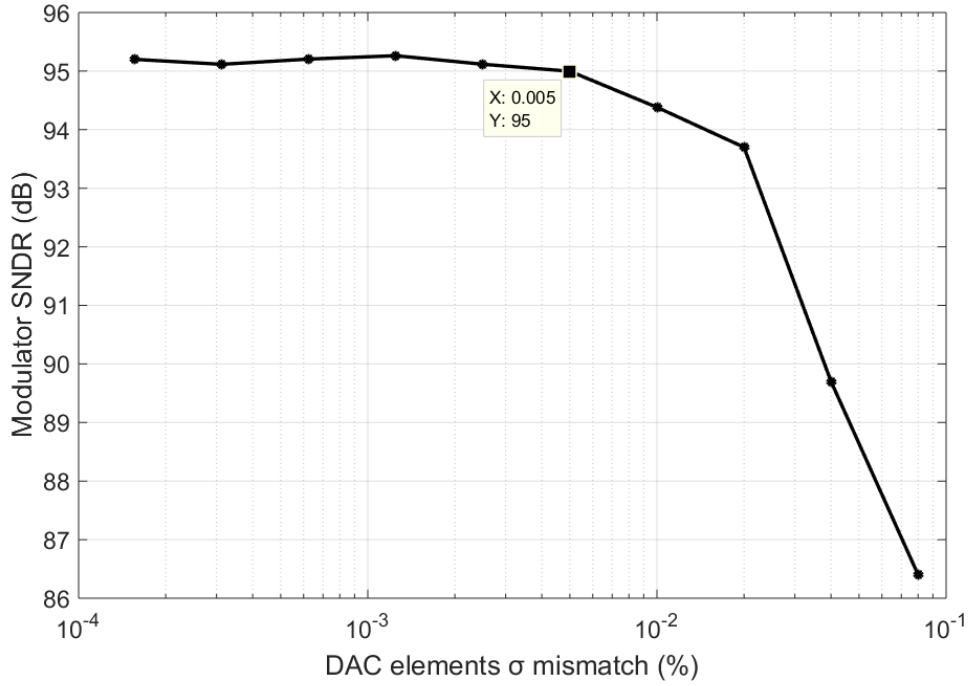


Figure 5.11 Modulator SNDR as function of the DAC elements standard deviation percentage

5.3 Modulator design

This section presents the modulator design which can be divided in two main parts: the loop filter and main feedback DAC in a first part, and the summation node and VCO-based quantizer in a second part, including also the digital logic for the DAC DEM mechanism.

5.3.1 Second-order loop filter and tri-level elements DAC design

The second-order loop filter allows by its feedforward topology an implementation with a single amplifier, which is benefiting the power consumption and employed also in [40]. A modified version of the RCX filter presented in [64] is used in this work, where the loop filter does not need a resonator function and then the high resistive path is suppressed, saving area. The implemented loop filter is shown on figure 5.12 and its transfer function $H_F(s)$ is:

$$H_F(s) = \frac{-(1 + R_2C_2s)}{R_1(C_1 - C_2)s + R_1R_2C_1C_2s^2} \tag{5.1}$$

By using $C_1 = C_2 = C$, $H_F(s)$ can be rewritten as:

$$H_F(s) = \frac{-\left(\frac{1}{R_1R_2C^2} + \frac{1}{R_1C}s\right)}{s^2} \tag{5.2}$$

Equation (5.2) is now in the same form as the equation $H(s)$ of the modeled loop filter presented in figure 5.6 b):

$$H(s) = \frac{\frac{k_3F_S^2}{1.5} + \frac{k_2F_S}{1.5}s}{s^2} \tag{5.3}$$

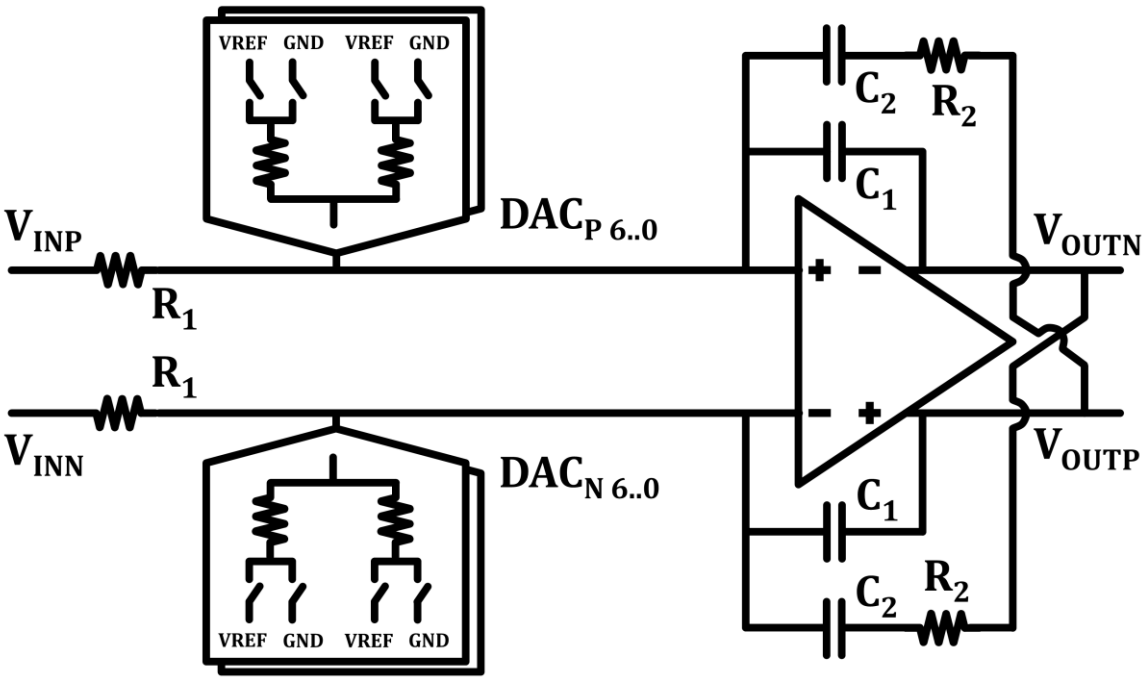


Figure 5.12 Single-opamp second-order loop filter with main feedback DAC using tri-level unit elements

Since the input-referred noise of the VCO-based quantizer and the second feedback DAC is second-order noise shaped, and the quantization noise is set 20dB below the expected SNR level, almost all of the input-referred noise specification can be attributed to the loop filter and the main feedback DAC. Considering a low-noise reference voltage regulator as used for the discrete-time modulator design (~2µVrms), its contribution can be neglected. The remaining contributors are the input resistors, the DAC resistors and the amplifier. Let's specify 12µVrms input-referred noise for this stage (1µVrms is kept as margin), and each of the cited contributors before account for 1/3 of this specification, it leads that:

$$R_1 = \frac{V_n^2}{3 \times 8 \times k \times T \times BW} = 60k\Omega \tag{5.4}$$

Where k is the Boltzmann constant, T the temperature in Kelvin (300K), and BW the signal bandwidth (24-kHz). Since R_1 is determined by noise constraint, C and R_2 can be determined by identification between (5.2) and (5.3). This gives C = 16.63pF and R_2 = 54.425kΩ. Due to the relatively high DC gain requirement (>60dB), a two-stage amplifier with miller compensation is used. The schematic of the amplifier and its common-mode feedback loop is presented on figure 5.13. The design is classic, using a PMOS input pair for its lower flicker noise, and source degeneration for the NMOS current sources to reduce their noise contribution. The loop filter input-referred noise (including input resistors) is 9.8µVrms, leaving 1/3 of the total allowed noise power (square of 12µVrms) for the resistive DAC. The designed amplifier reaches a 75dB DC gain with a 12.5MHz GBW product, satisfying the requirements made in the previous section while consuming 43µA current.

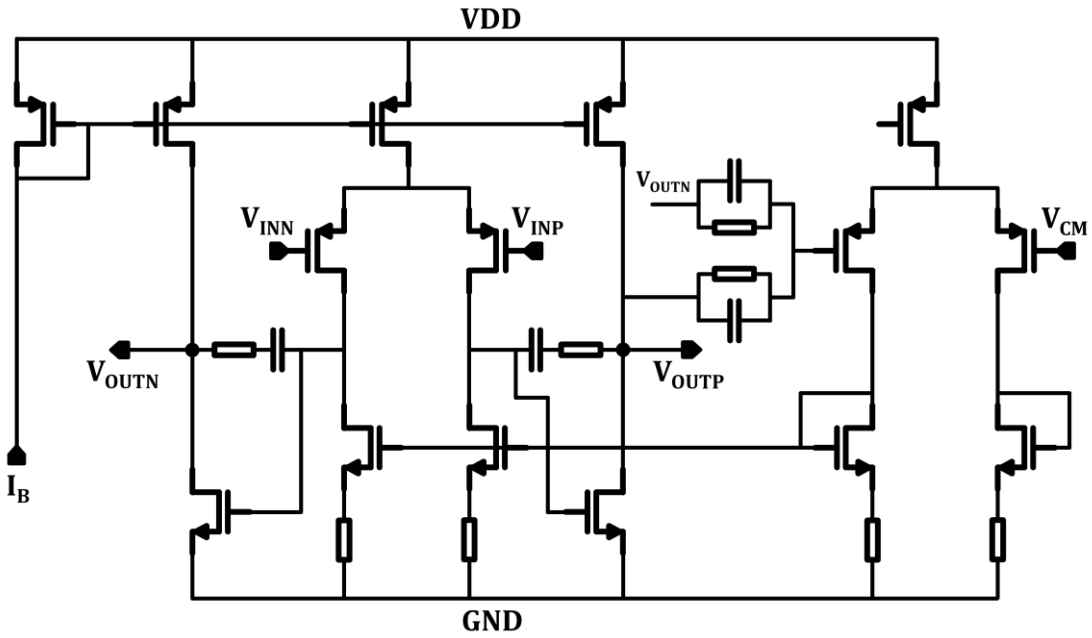


Figure 5.13 Loop filter amplifier schematic with common-mode feedback amplifier

The VCO-based quantizer used in this design outputs a thermometer code for the data and an additional bit for the sign, generating inherently the proper control signals for the use of a tri-level DAC. A NRZ DAC switching scheme is adopted to benefit from the multibit quantization to reduce the jitter sensitivity as shown in chapter 4. Each DAC control signal is then synchronized with flip flops clocked with a delayed clock. This delayed clock leaves enough time for the decoding and DEM logics to operate and has a quite low delay (20% of a clock period); thus, this design does not need to compensate ELD as shown previously by the model simulations. In a sigma-delta modulator, the performances of the main feedback DAC are critical and transient effects like Inter-Symbol Interference can limit the achievable resolution. A tri-level DAC increases the number of possible different transitions (0 to +1, +1 to 0, 0 to -1 and -1 to 0) and so the sensitivity to this issue. Instead of using a single resistor switched between 3 levels, the proposed solution (figure 5.14) consists of using two resistors in each unit element of the tri-level DAC. The latter are switched both to V_{REF} to output +1, one to V_{REF} and the other to GND to output 0 and both to GND to output -1. Then, each resistor is only switched between two levels reducing the number of different transitions. For each resistor, both switches (a PMOS connected to V_{REF} and a NMOS connected to GND) can use the same control signal: one is active while the other is off due to the use of a NRZ switching scheme. This implies that each flip-flop synchronizing the DAC switching instant will be loaded by the same elements (one NMOS and one PMOS switches). Thus, the switches form an inverter driving the resistor, leading to smooth and equal transitions and a reduced sensitivity to ISI. Since two resistors are used to generate the 0 level, a linearity issue can occur in case of mismatch. A local DEM in each unit element alternates the resistors each time the 0 level has to be outputted to avoid this issue, with the generation shown on figure 5.14 (same as used in discrete-time modulator). These tri-level DAC unit elements are similar to the discrete-time modulator tri-level DAC, except that they are not active on half of a clock period, but use a NRZ scheme and control signals are synchronized using flip-flops. The truth table for the control logic of a unit element DAC is given in table 5.2.

Table 5.2 Truth table of tri-level unit element DAC switches logic

DEM	SIGN	DATA	S_1	S_2	S_3	S_4
X	0	1	1	1	0	0
X	1	1	0	0	1	1
0	0	0	0	1	0	1
1	0	0	1	0	1	0
X	1	0	X	X	X	X

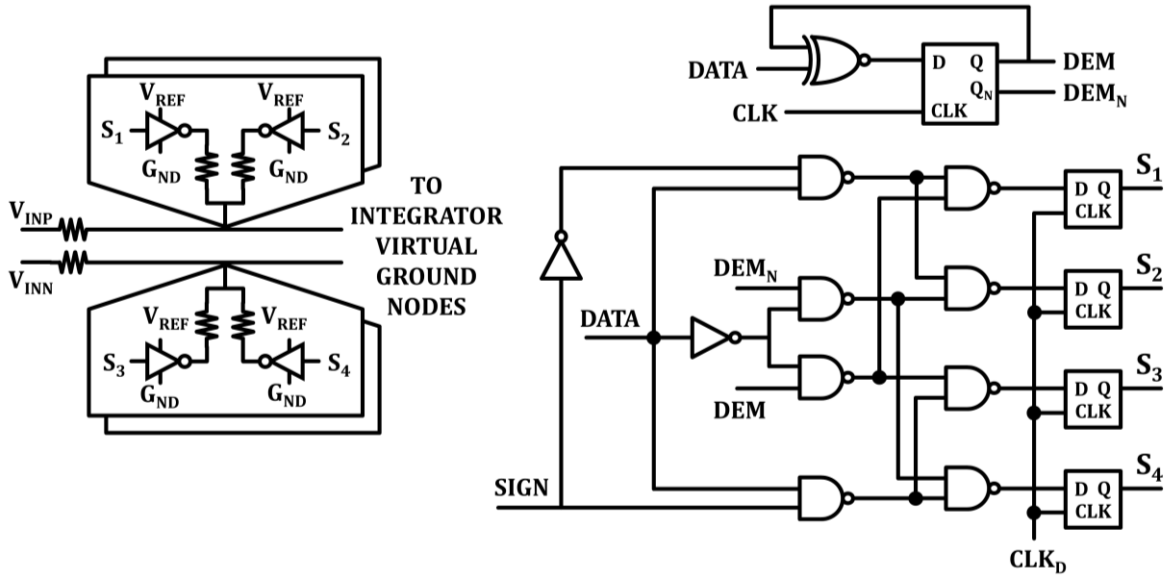


Figure 5.14 Tri-level unit elements of main feedback DAC and switches generation logic

5.3.2 VCO-based quantizer design and interface

The VCO-based quantizer uses the concepts presented in [63] to achieve a good linearity as well as an interesting power efficiency. Since the selected modulator architecture needs 15 quantization levels (N_Q), and knowing that $N_Q = 2N + 1$ where N is the number of phases used in the VCOs (developed in the preceding section), we can deduce that the number of stages in the VCOs to be used in our configuration is 7. The VCOs are implemented with inverter delay cells in a ring oscillator configuration as shown in figure 5.15 a). The ring oscillators are controlled in current, their biasing current sets the idle frequency of the oscillators F_0. To interface with the loop filter output, a V-to-I conversion has then to be done. Two possibilities are shown on figure 5.15 a). One uses resistors (left side) to inject an additional signal current to the bias current. This solution is the simplest but if V_P or V_N node has a different voltage than loop filter output common-mode one, an extra current will be injected or subtracted, changing F_0 and K_{VCO} of the ring oscillators. V_P and V_N voltages are mainly determined by the threshold voltages of the inverter and are sensitive to PVT (300mV to 600mV range). Moreover, when a differential signal is applied, oscillators frequencies follow VCOs inputs, changing also the V_P and V_N voltages, making the extra injected current signal dependent and lowering the linearity. A better solution to interface the loop filter is the use of a transconductor stage as shown on figure 5.15 a) on the right side. In that case, V_P and V_N variations due to PVT can be handled without any impact on injected current as long as the differential pair remains in saturation region.

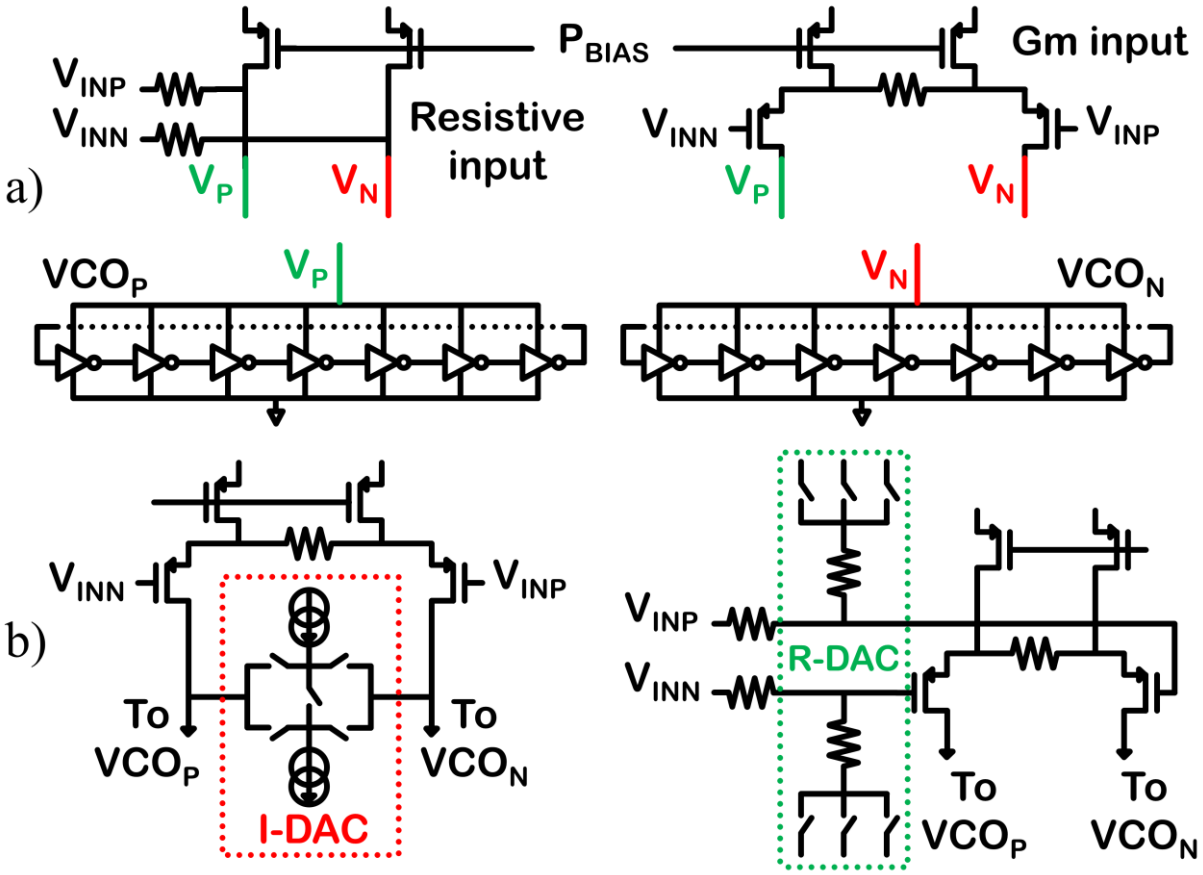


Figure 5.15 a) Possible interfaces between the loop filter and VCO-based quantizer b) Summation operation done either in current or voltage

Nevertheless, the gm solution comes also with its drawback: since V_{INP} and V_{INN} nodes are biased at half the supply voltage of 1.5V, only one half is left to handle the threshold voltage V_{thp} of the differential pair, the signal swing, and the drain-source voltage of the biasing source. Using large W/L ratio for the differential pair reduces the V_{thp} value while a large L for the biasing sources increases their output impedance and allows lower V_{DS} without degradation of the bias current value. But this is not sufficient and a signal swing reduction must be done. To do it, let's have a look at the interface between the loop filter and the VCO-based quantizer which implements a summation function with the second feedback DAC as it can be seen on figure 5.6 b). The summation can be done in current with a current DAC placed after the gm-stage adjusting the current feeding the VCOs, as visible on the left side of figure 5.15 b). In that case, the input signal swing of the transconductance stage is the loop filter output signal swing. Another approach is to use a passive summation before the gm-stage, lowering the gm input signal swing. This is shown on the right side of the figure 5.15 b), where the passive summation is made by resistors, subtracting the signal component to the loop filter output and lowering the differential signal swing from 600mV to 200mV, which is sufficient to solve the swing issue. These resistors create with the input capacitance of the gm-stage a parasitic pole which could disrupt the modulator stability. Behavioral simulations show that a pole higher than $1.5F_S$ has no impact on the modulator stability, linearity or noise shaping performance. Since in an audio application the F_S is quite low (here 6.144MHz for this modulator), a careful design meets this requirement. Using high-density poly resistors (for low area) and a proper sizing of gm-stage input transistors, the parasitic capacitance is made low enough to keep relatively high resistance

values and then to ensure a low-power consumption. Since the secondary DAC errors due to mismatch are second-order noise-shaped, there is no need here for a DWA algorithm and single resistors are switched between the ground, the common mode and the supply level to create a tri-level DAC unit element.

The idle frequency F_0 of the VCOs is determined by the number N of delay cells and by the transition time of one delay cell multiplied by 2 (two transitions needed for one period). This time corresponds to the charging or discharging time of the next stage gates. Hence, this time is proportional to several factors: the bias current I_{BIAS}, the capacitance on delay cell output node C_{EQ} (including the input capacitance of next delay cell), and the value of V_P and V_N node called V_S. Then F_0 is approximated by:

$$F_0 = \frac{I_{BIAS}}{2 \times N \times C_{EQ} \times V_S} \tag{5.5}$$

K_{VCO} is the gain factor for the voltage-to-frequency conversion. It can be expressed by the multiplication of the gm value and the derivative of F_0 versus the current, resulting to the equation below:

$$K_{VCO} = \frac{gm}{2 \times N \times C_{EQ} \times V_S} \tag{5.6}$$

Since K_{VCO} is defined by the time constant value of the last integrator it has then a relatively high value in the order of F_S. It becomes interesting from a power point of view to maximize the ratio K_{VCO}/F_0: it allows to reduce F_0 and so the biasing current while still reaching the desired value of K_{VCO}. The K_{VCO}/F_0 ratio can be reduced combining (5.5) and (5.6) to a gm/I_{BIAS} ratio. The gm value is bounded: gm times the maximum input voltage $V_{IN,MAX}$ cannot be larger than the biasing current I_{BIAS}. To ensure a proper work of the gm-stage with an acceptable distortion level (simulated with behavioral model), the following equation is used to define the gm value:

$$gm \times V_{IN,MAX} = \frac{I_{BIAS}}{\sqrt{2}} \tag{5.7}$$

Since the ratio K_{VCO}/F_0 is equal to gm/I_{BIAS} and using (5.7), one can rewrite the K_{VCO}/F_0 ratio as:

$$\frac{K_{VCO}}{F_0} = \frac{1}{\sqrt{2} \times V_{IN,MAX}} = 3.53 \tag{5.8}$$

Equation (5.8) shows that the reduction of the input signal swing by the passive summation from 600mV to 200mV allows to get a three times higher K_{VCO}/F_0 ratio, which allows to reduce the power consumption of the two VCOS and their gm-biasing stage to 6µA. Figure 5.16 shows the implementation of the VCOs, the sampling latches and the phase decoder. The VCOs inverter cells are made differential to improve PSRR with cross coupling outputs to synchronize the switching between inverters. A differential flip-flop samples each phase of the VCOs at every falling edge of the F_S clock. Thanks to the pseudo-digital nature of the phase signals, flip-flops are composed of a dynamic preamplifier operating on a clock falling edge with a positive feedback load to benefit from its high gain and a latch stage to lock the data value. A careful design of the dynamic preamplifier is necessary: it must sense correctly the data under mismatch

and PVT conditions with a relatively low input swing for a pseudo-digital signal. Since ring oscillators produce a signal between ground and V_P and V_N which are in the V_{th} order, a NMOS input pair is chosen: one transistor will be completely off while the other will be conducting or near its conducting limit, helping to drive the positive feedback load in the right direction even under mismatch conditions. This allows the transistors size reduction in the differential flip-flop to limit the power consumption during switching, since the design needs 14 sampling elements.

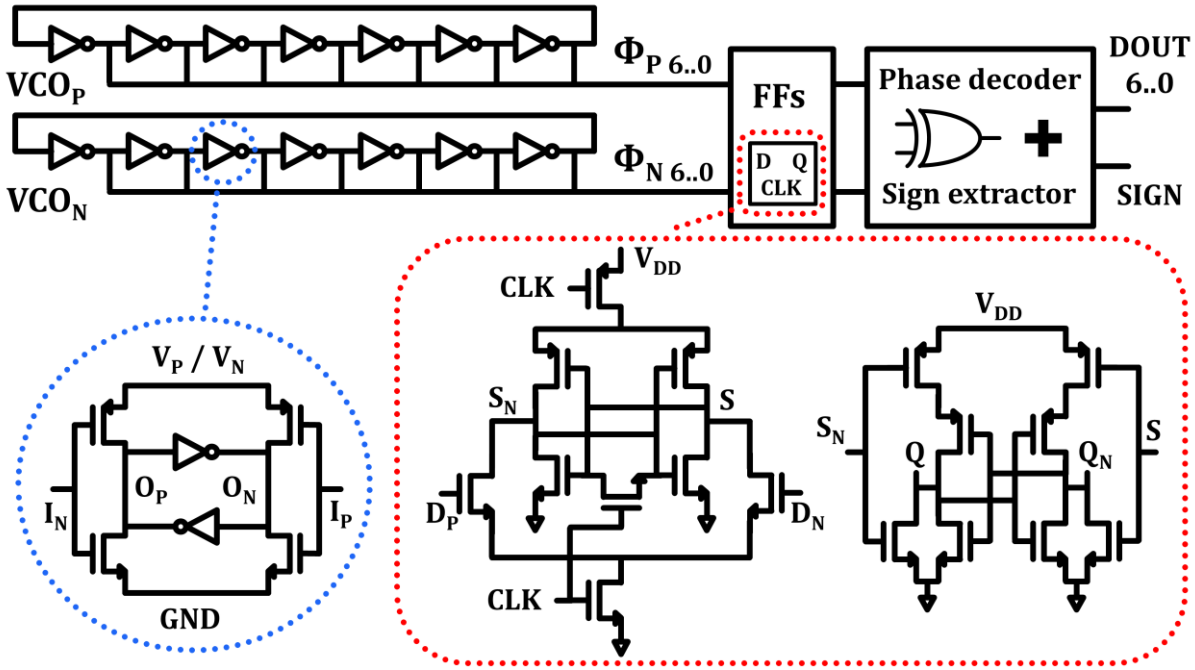


Figure 5.16 VCO delay cell implementation and custom flip-flops schematics

The phase decoder schematic is presented in figure 5.17. It uses XOR gates to compare phase vectors coming from VCO_P and VCO_N. This generates a quantized signal of the phase difference between the two VCOs as explained in section 5.1 and provide the Data vector. A logic circuit similar to [63] is used to determine the sign of the VCOs phase difference. As stated in section 5.1, a sufficient condition to detect a negative phase difference sign is the coincidence between a 0 to 1 edge in the XOR-based phase decoder output word and the transiting element of VCO_P. Detection of the transiting element of VCO_P is obtained by comparing if each phase is in the same state than the previous one. This is done algorithmically with a XORN gate and a left rotation operation as visible on figure 5.17. The detection of a 0 to 1 edge in the phase decoder output word is obtained similarly with a left rotation operation and a AND gate with one inverted input. This logic generates a remarkable vector which is a pointer of the first "1" in the rotative thermometer coded DATA vector. This logic is one-hot coded and will be reused in the DEM logic to suppress the rotation of the Data vector thermometer code. That's why the implementation of the sign extractor differs slightly from [63], because of the need of this pointer generation. To finish, AND gates compares the pointer vector and the transiting element vector (figure 5.17) and result is summed by an OR gate. An issue not handled in [63] occurs when the phase difference is equal to π. In that particular case, the Data vector is only constituted by "1" values. A 0 to 1 edge is then impossible to detect, making the feedback DAC to output $+V_{REF}$ instead of $-V_{REF}$ and resulting in an unstable modulator behavior. To correct this, when the Data vector is full of "1", an AND gate compares

if the preceding value of the phase difference sign was already negative, and the result is added in the OR sum (figure 5.17). This is possible in an oversampled system since when the output of a multibit modulator is near its full scale or minus its full scale, the signal sign will not change for the next sample.

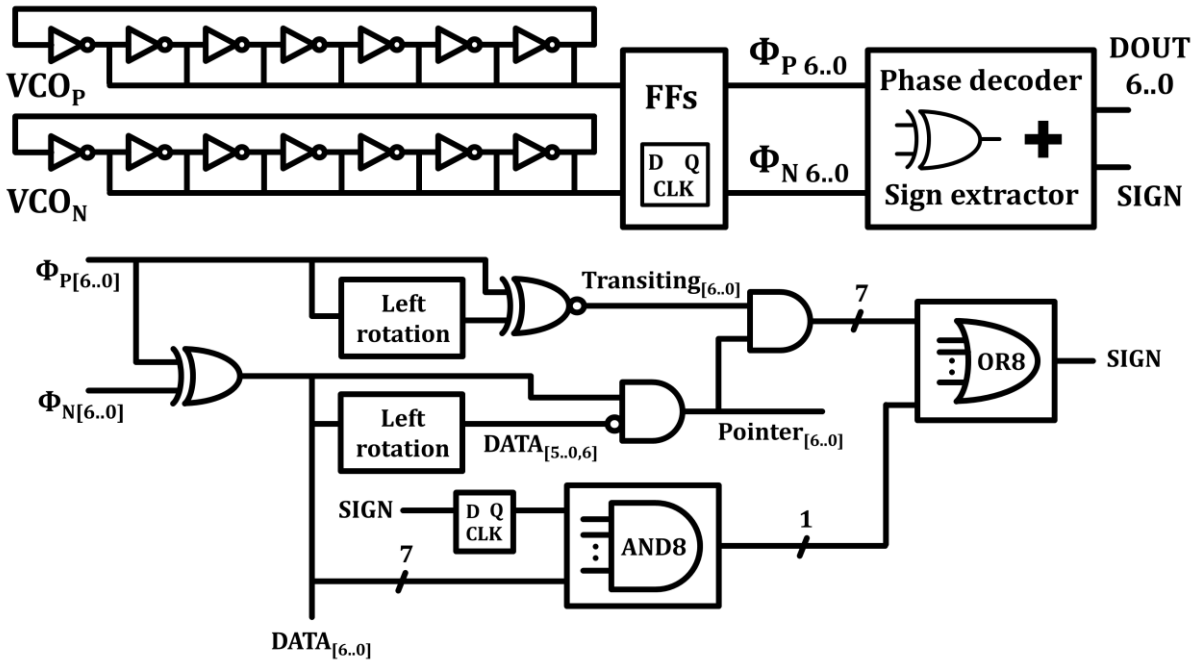


Figure 5.17 Schematic of phase decoder and sign extractor logic

As said before, a DWA algorithm to noise shape the DAC mismatch errors is needed to achieve the desired resolution. The proposed solution (figure 5.18) implements an explicit DWA algorithm for the tri-level DAC to benefit from its first-order mismatch error shaping property. The VCO-based quantizer outputs a rotating thermometric code: decoding the binary output necessary for the DWA pointer generation and to interface with the decimator is then quite complex due to the number of possible combinations. A first barrel shifter is used to restore the Data vector in a classical thermometric code. Its pointer generation is coming from an already existing logic in the sign extractor function, as shown in figure 5.17. The output of the first barrel shifter (DataThermometric vector) is then binary coded and sent to the decimator with the sign bit.

The sign of the data output in figure 5.18 alternates between each value to better illustrate the tri-level algorithm. Two pointers are used for the tri-level DAC, one for the +1 values and the other for the -1 values. A 7-modulo adder generates the pointer for the DWA barrel shifter by adding the data with the corresponding pointer old value depending on the sign value. A binary to one-hot decoder is needed to control correctly the barrel shifter. Then, as it can be seen on figure 5.18, DAC elements selected to output +1 (in green) rotate independently from DAC elements selected to output -1 (in orange), each one following a classic DWA algorithm. Monte Carlo simulations of modulator with DAC mismatch and DWA algorithm exhibit a worst case SQNR of 110dB (90dB without DWA), validating the proposed concept.

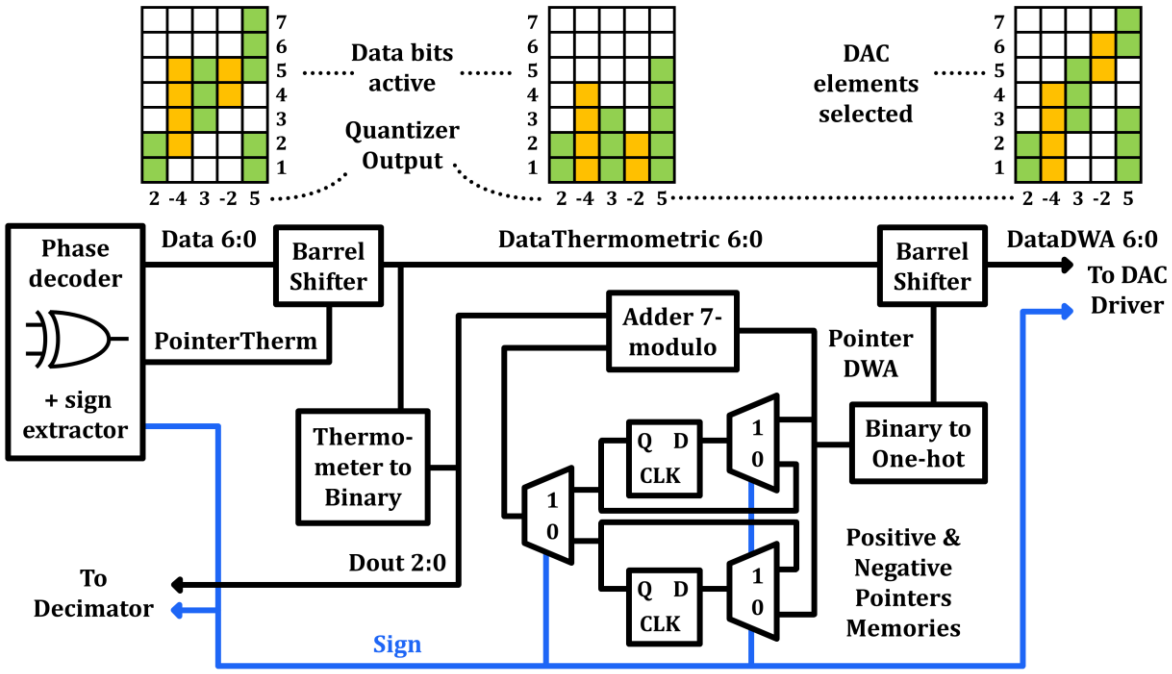


Figure 5.18 Block diagram of the tri-level DWA algorithm implemented

5.4 Simulation results

The proposed modulator has been designed in a 140nm standard CMOS technology under 1.5V supply voltage. Figure 5.19 shows its output spectrum (Hann-windowed 61440-pt FFT) with a -3dBFS input signal at 1kHz in typical conditions and with the two extremum corners results (slow -40°C and fast 125°C). The expected thermal noise level, corresponding to the specification of 13µVrms integrated on a 24kHz bandwidth is also shown. It reaches a SQNR of 116dB in typical conditions and 113dB SQNR in worst case on a 24kHz band. The VCO-based quantizer non-linearity has no impact since the third harmonic is below the simulated thermal noise level (figure 5.19), showing the proposed modulator architecture robustness. The VCOs idle frequency is also visible, with a tone near $F_S/2$ which varies a little over corners.

Mismatch simulations are also run to check that the tri-level DWA algorithm implemented works correctly and is sufficient. Its effectiveness is proven with a mean SQNR going from 92dB without DWA to 116dB with DWA as it can be seen on figure 5.20 spectrums

Corners simulation are also run with a low signal level (-60dBFS), checking that no tones appear when signal energy is dominated by the quantization noise as visible on figure 5.21.

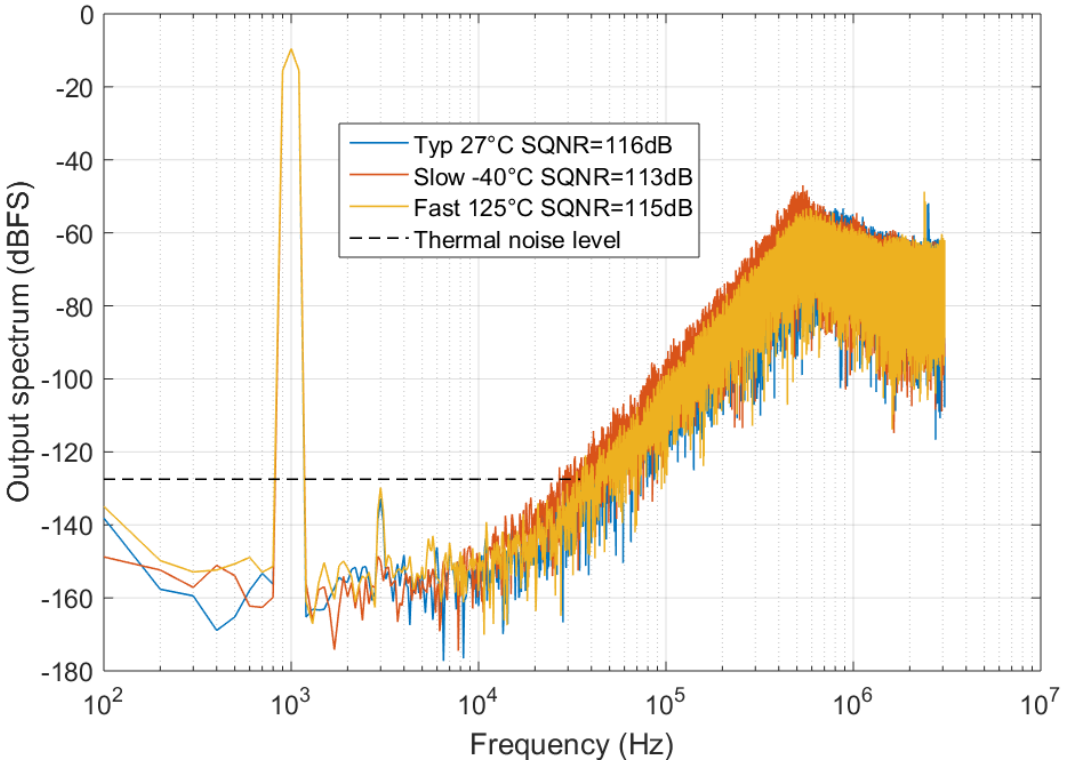


Figure 5.19 Modulator output spectrums over PVT conditions: typical, slow at -40°C and fast at 125°C

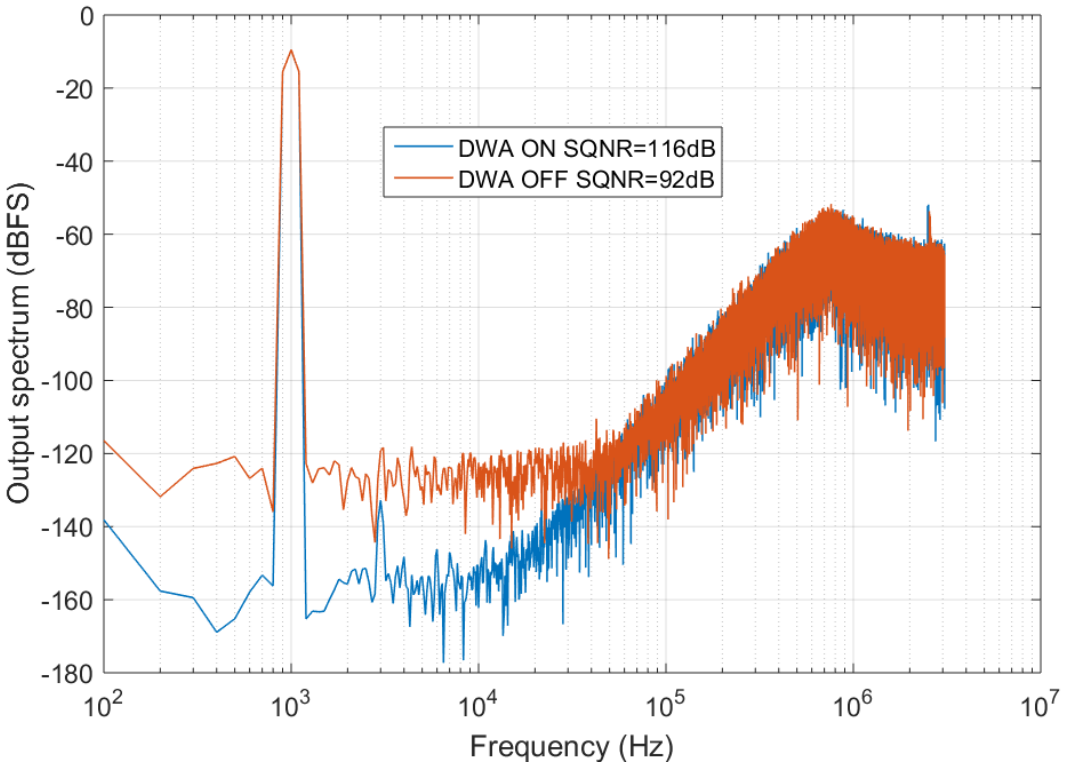


Figure 5.20 Modulator mismatch simulations with / without DWA output spectrums

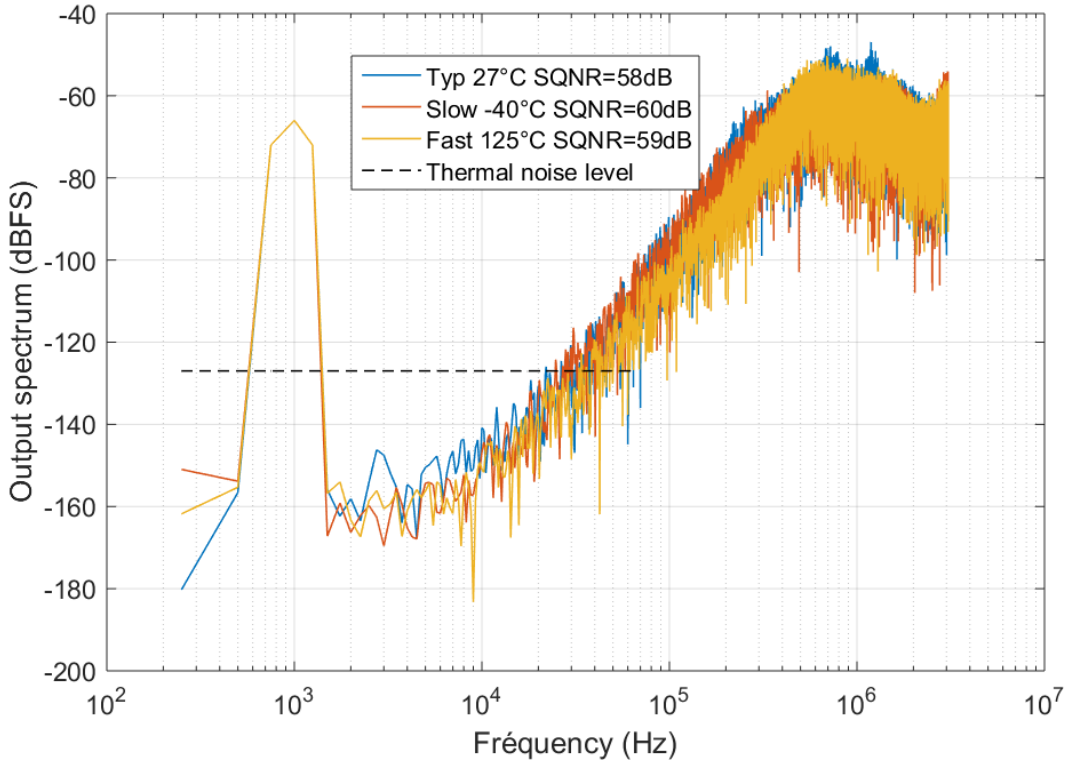


Figure 5.21 Modulator output spectrums over PVT conditions for a -60dBFS input signal

The total power consumption of the modulator is 142µW, and the power split is given in figure 5.22. The main contributor is the first stage, where the loop filter and the resistive DAC consumes 60% of the total power, followed by the digital logic of the phase quantizer and DEM algorithm. The latter represents almost one third of the total power consumption which is non-negligible, but with nanometer technologies further power improvement can be made on digital and VCO parts due to the lower consumption needed for switching.

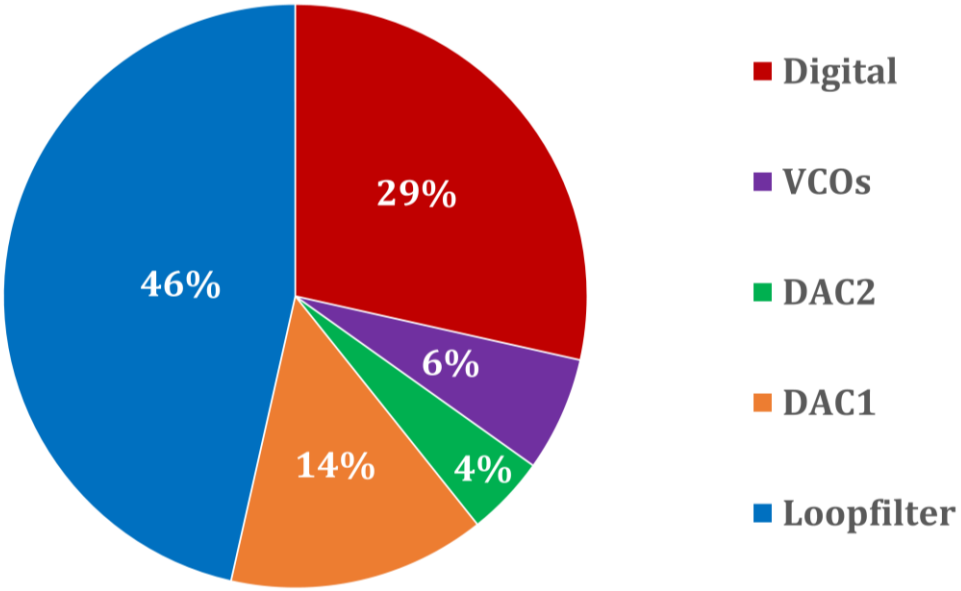


Figure 5.22 Modulator power consumption split

The modulator total input-referred noise is 13µVrms on a 24kHz bandwidth and its power split is given in figure 5.23. The input resistors and the main feedback DAC ones account for 60%

of the total noise power, while 36% is coming from the implemented two-stage amplifier. Thanks to the single-amplifier second-order loop filter, the VCOs noise and the second feedback DAC noise are second-order noise shaped and made negligible, together with the quantization noise which is about 20dB below the thermal noise level. Using behavioral simulations, the modulator can tolerate up to 25ps rms clock jitter without losing resolution.

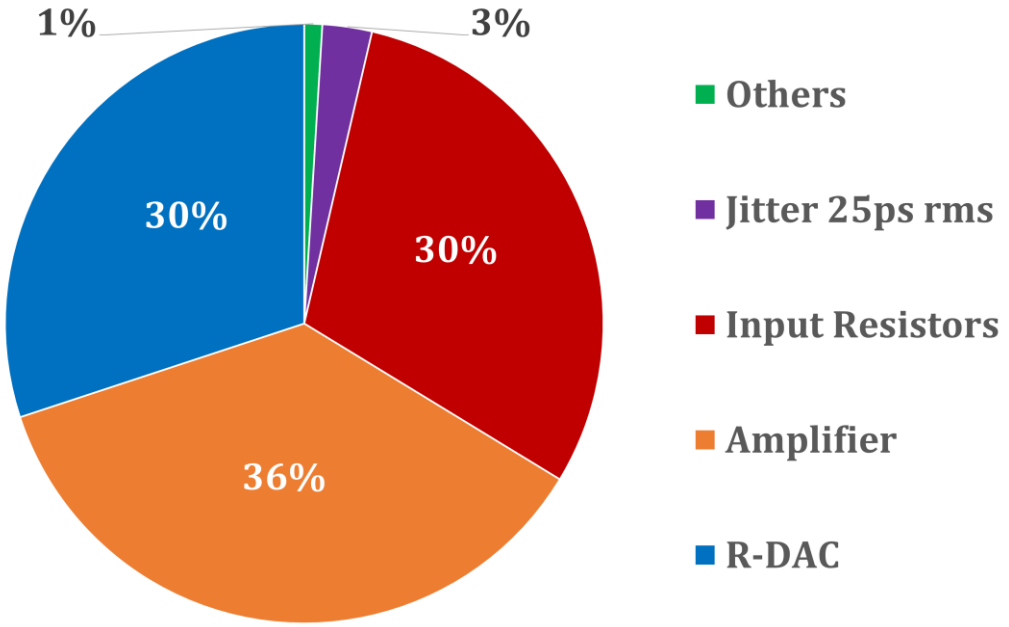


Figure 5.23 Continuous-time modulator input-referred noise power split

The modulator designed in 140nm CMOS technology occupies 0.232mm² and its layout is presented on the figure 5.24. The loop filter capacitances occupy most of the area, due to their metal fringe structure which ensures a very high linearity compared to the high density poly-well capacitor and allows then to meet the requirements at the price of area.

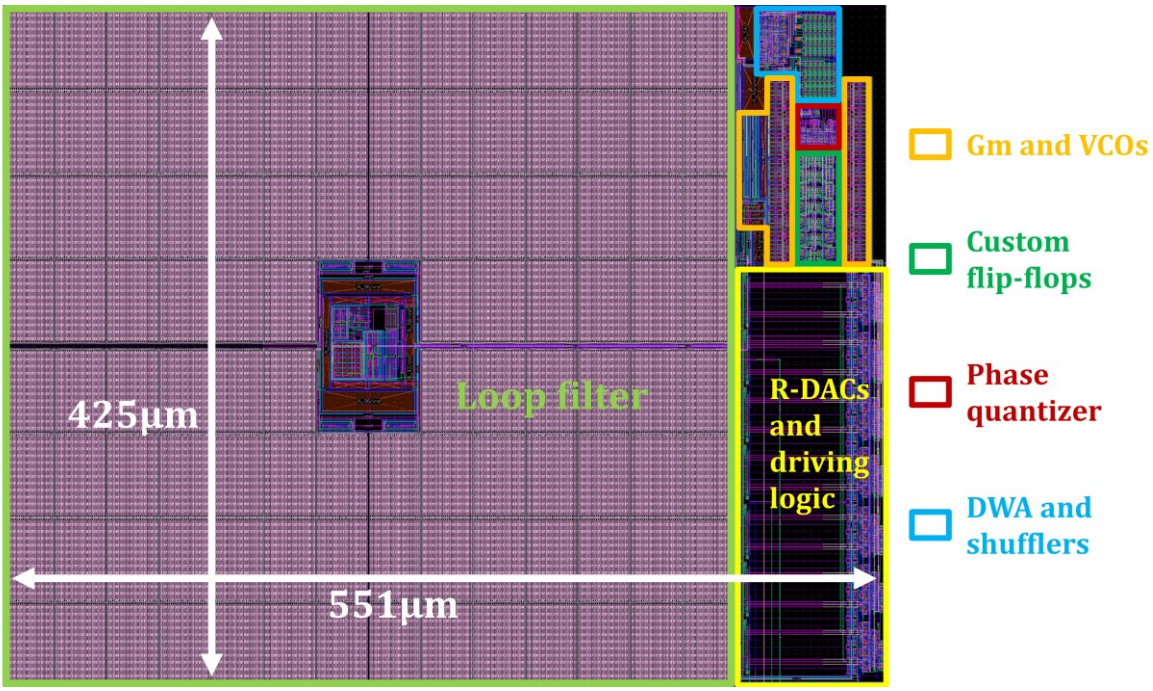


Figure 5.24 Layout of the designed continuous-time modulator

The designed modulator performances are summarized in table 5.3. They are compared with other continuous-time audio sigma-delta modulators reported in literature, like the state-of-the-art and other interesting designs previously cited. The proposed modulator reaches among the best state-of-the-art performances, demonstrating the usability of a VCO-based quantizer in a high resolution sigma-delta converter. The occupied area is among the lowest ones, excepting the low nanometer nodes of designs [47] and [65]. As seen on figure 5.24, a process technology with a better density-linearity capacitance compromise can reduce significantly the modulator area and will increase even more its architecture attractiveness from a cost point of view.

Table 5.3 Performance summary and comparison with literature

	This work	**[52]**	**[53]**	**[55]**	**[40]**	**[47]**	**[65]**	**[66]**
TECH (nm)	140	180	180	65	160	28	40	65
YEAR	2018	2014	2016	2016	2016	2015	2011	2015
ORDER	3	3	3	3	3	2	2	3
BW (kHz)	24	24	24	25	20	24	24	24
FS (MHz)	6.144	6.144	6.144	6.4	3	24	6.5	3.072
OSR	128	128	128	128	75	500	135	64
Supply (V)	1.5	1.8	1.8	1	1.6	3.3/1	1.2	1.1
DR (dB)	98	103	103.6	103	103.1	100.6	102	88
SNR (dB)	95	98.9	99.3	100.1	-	100.6	-	-
SNDR (dB)	95	98.2	98.5	95.2	91.3	98.5	90	85
POWER (µW)	142	280	280	800	390	1130	500	121
AREA (mm²)	0.232	1.25	1	0.256	0.21	0.022	0.05	0.6
FOM_S (dB)	180.3	182.3	183	177.9	180	173.9	178.8	171

The figure 5.25 shows the modulator among the ADC survey of Boris Murmann [4], where the FOM_S has been changed to use DR instead of SNDR (original definition), and where only sigma-delta modulators are shown.

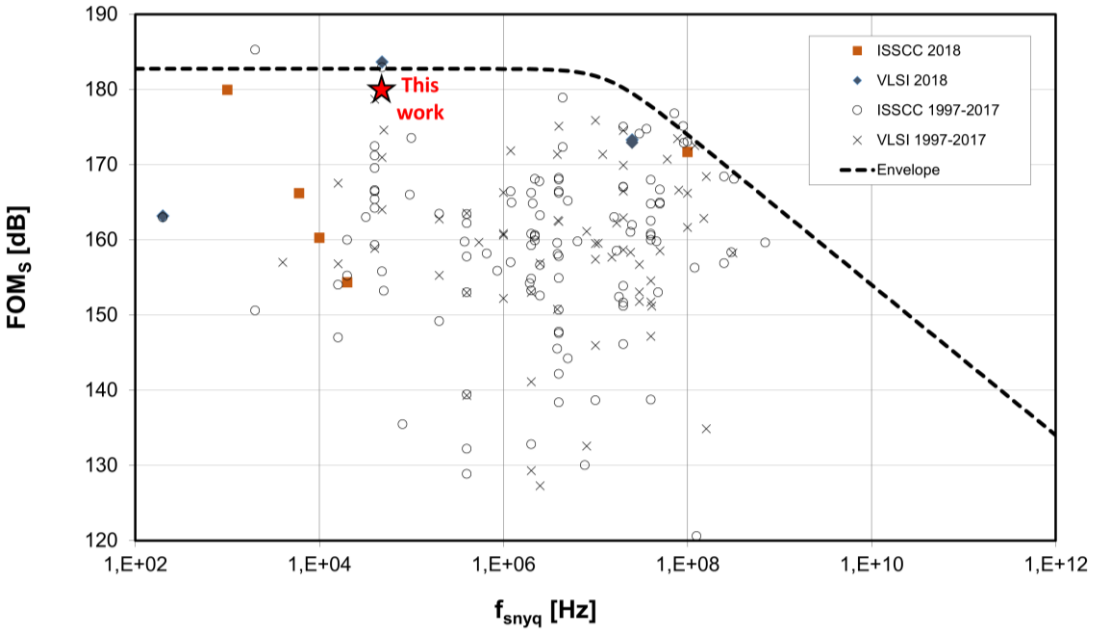


Figure 5.25 Placement of the continuous-time modulator on Murmann's ADC survey

5.5 Conclusion

In this chapter, a brief review of the state-of-the-art continuous-time sigma-delta modulator designs has been made, showing that low-power is obtained by using FIR or multibit feedback DAC, to reduce the loop filter swing and consumption, and by combining functions such as single-amplifier second-order resonator or integrating quantizer. Outside of the audio applications, a new type of integrating quantizer has emerged: the VCO-based ones. Due to their highly digital implementation, scaling in lower and future process nodes is ensured which is not the case for flash quantizers. Moreover, a VCO-based quantizer offers an extra benefit compared to a voltage-based multibit one: using a phase output, it provides one more integrating stage, simplifying the modulator architecture. But VCO-based quantizers are sensitive to PVT and are highly non-linear, which limits their application to low-to-medium resolution converters. A design of a 16-bit audio sigma-delta modulator using a VCO-based quantizer is proposed, reaching a 180.3dB FOM_S and placing it among the state-of-the-art. To limit the power consumption, a single-amplifier second-order loop filter is employed and an integrating quantizer too, dividing then the modulator into only two sub-circuits. To handle the PVT and linearity issues, several architectural choices have been made: use of differential VCOs, high OSR and low out-of-band gain for the NTF, second-order loop filter which noise-shapes the remaining non-linearity. With a lower process node, a substantial power consumption improvement can be expected since one third of the total is attributed to digital circuits.

Chapter 6

Conclusion

To answer the need of a high dynamic range audio analog front-end of the new generation microphones, an acquisition system composed of a programmable gain amplifier and an analog-digital converter is a suitable solution since the DR requirement is decorrelated from other requirements like peak SNR, allowing the use of a more power efficient ADC with a reduced dynamic range. Since microphones are more and more embedded in battery powered devices, the power consumption of the analog front-end is a major concern, as well as a low area occupation to facilitate the integration in small devices. This thesis concentrates on the design of audio sigma-delta modulators with an emphasis on low power and area for both discrete-time and continuous-time solutions.

The state of the art of discrete-time modulator design employs class AB amplifiers to achieve a high power efficiency. Among the class AB amplifiers, inverter-based ones show the best efficiency due to their single-current branch making them very suitable to make discrete-time integrators. But inverters are sensitive to PVT and not reliable for mass production: their current consumption cannot be accurately controlled. A new inverter-based amplifier solving this issue has been presented in chapter 3; by using biasing transistors operating in their linear region the current consumption can be controlled while the amplifier class AB operation is kept. A second-order modulator with a 1.5bit quantizer (3 levels) is then designed employing these amplifiers to reach a low power consumption. An accurate high-level model of a discrete-time integrator has been developed in chapter 2 to define amplifier specifications and to avoid over-design while ensuring the modulator performances. To save area, a high OSR of 256 is selected allowing to reduce the capacitance sizes since they generally represent the most of a discrete-time modulator area. In the same spirit, the input sampling capacitances are split to form a tri-level DAC during the integration phase which avoids the need of a dedicated DAC. Not only this reduces the area occupation but also lowers the load of the first integrator which benefits to the power consumption too. The final area occupation is 0.084mm² which is better than comparable performance designs and ensures the modulator low cost. This design reaches 100dB SNR and 103dB DR on a 24kHz bandwidth for a 480µW power consumption. It leads to an 180dB FOM_S placing it among the state-of-the-art audio discrete-time modulators, validating the design choices and demonstrating that a high OSR can be used in power-efficient converters.

The review of audio continuous-time modulators literature shows that the best power efficiency is reached when combining several functions into a single bloc element. For example, an integrating-quantizer allows to reduce the loop filter order by one while keeping the same noise shaping performance. Among these quantizers, the VCO-based ones are very attractive for our application: a differential input voltage is first converted into a frequency difference, then the

VCOs phase difference (integral of the frequency with a 2π factor) is sampled and quantized, providing a perfect integrating stage. Their highly digital implementation (inverters, flip-flops and logic gates) is an advantage for area and power furthermore in nanometer nodes. Due to the pseudo-digital nature of signals, quantization is done almost instantly on a clock edge, relaxing the ELD issue of continuous-time modulators. But the VCOs are highly non-linear (poor harmonic performance) which limits their applications in high-bandwidth medium-resolution converters, and are sensitive to PVT. These issues have been handled by architectural choices: the presented modulator embeds a 4-bit VCO-based quantizer in a third-order loop, which noise-shapes the VCOs non-linearity by a second-order loop filter. A low out-of-band gain NTF is selected to tolerate large integrator time constant variations, then the voltage-to-frequency gain of the VCOs (which sets the integration time constant) can vary with PVT without impact on the modulator stability. The loop filter-VCOs interface is made by a gm-stage which decouples the fixed loop filter output common-mode with the PVT-sensitive VCOs common-mode. The second-order loop filter allows, thanks to its feedforward topology, a single amplifier realization, saving one integrator and benefiting to the power consumption and the area occupation. This design has a peak SNR of 95dB on a 24kHz bandwidth, a DR of 98dB and a 142µW power consumption using a 0.232mm² area. Its 180.3dB FOM_S places the designed modulator among the state-of-the-art of continuous-time ones, showing that VCO-based quantizers can be used in high resolution converters without performance loss.

The principal advantages of a continuous-time solution over a discrete-time one are the resistive input which is easier to drive than a switched-capacitor one, and its inherent anti-aliasing filter property, which simplifies the programmable gain amplifier design. On the other hand, a discrete-time realization offers much more system modularity thanks to its clock frequency scaling: resolution and power scale with the clock frequency value without stability issues. This allows discrete-time modulators to support several audio sampling frequencies commonly used in commercial products like 16kHz, 32kHz, 44.1kHz and 48kHz.

The improvement perspective of the designed discrete-time modulator is to include a switched-amplifier behavior: the amplifier current is cut-off during the sampling phase which can reduce in theory all the amplifiers current consumption by 50%. But some fast biasing circuit must be added, lowering the potential benefit. On the continuous-time modulator side, a cascade of several VCOs stages in a low nanometer process node is suitable to further reduce the modulator power consumption and area occupation. But using VCOs in the first stage lowers the harmonic performance and a higher number of bits in the quantizer (6-8) must be considered to sufficiently reduce the signal swing at the VCOs inputs.

Publications

The work done in this thesis gave rise to several publications in journals or conference participations and are listed below:

- **« A design-oriented approach for modeling integrators non-idealities in discrete-time sigma-delta modulators »**

  Poster presentation at the IEEE ISCAS 2017 Conference and paper available in IEEE ISCAS 2017 conference proceedings

- **« A Robust Inverter-Based Amplifier versus PVT for Discrete-Time Integrators »**

  Paper published in the International Journal of Circuit Theory and Applications (IJCTA), Wiley, July 2018

- **« Continuous-time sigma-delta modulators: a tutorial on design procedure »**

  Paper submitted to the International Journal of Circuit Theory and Applications (IJCTA), Wiley, June 2018

- **« A 16-bit Audio Continuous-Time Sigma-Delta Modulator Using VCO-based Quantizer »**

  Lecture presentation at the IEEE NEWCAS 2018 conference and paper available in IEEE NEWCAS 2018 conference proceedings

References

[1] InvenSense ICS-40618 microphone datasheet

[2] Infineon IM69D130 microphone datasheet

[3] Knowles SPA1687LR5H-1 microphone datasheet

[4] B. Murmann, "ADC Performance Survey 1997-2018," [Online]. Available: http://web.stanford.edu/~murmann/adcsurvey.html

[5] R. Schreier, G. C. Themes, "Understanding delta-sigma data converters," USA: IEEE-Wiley, 2005

[6] C. E. Shannon, "Communication in the presence of noise," Proceedings of the IRE, vol. 37 no. 1, 1949

[7] S. R. Norsworthy, R. Schreier and G. C. Temes, "Delta-sigma data converters: theory, design and simulation," USA: IEEE-Wiley, 1997

[8] F. Maloberti, "Data converters," USA: Springer, 2007

[9] J. M. De la Rosa and R. Del Rio, "CMOS sigma-delta data converters : practical design guide," USA: Wiley, 2013

[10] A. Sripad and D. Snyder, "A necessary and sufficient condition for quantization error to be uniform and white," IEEE Transactions on Acoustic, Speech, and Signal Processing, vol.25 no.5, 1977

[11] J. Silva, U. Moon, J. Steensgaard and G. C. Temes, "Wideband low-distorsion delta-sigma ADC topology," IET Electronics Letters, vol. 37 no. 12, 2001

[12] Y. Matsuya and J. Yamada, "1V power supply, low-power consumption A/D conversion technique with swing-suppression noise shaping," IEEE Journal of Solid-State Circuits (JSSC), vol. 29 no. 12, 1994

[13] G. Suarez, M. Jimenez, and F.O. Fernandez, "Behavioral modeling methods for switched-capacitor ΣΔ modulators," IEEE Transactions on Circuits and Systems, 2007

[14] R. del Rio, F. Medeiro, B. Perez-Verdu and A. Rodriguez-Vazquez, "Reliable analysis of settling errors in sc integrators – application to the design of high-speed ΣΔ modulators," IEEE Int. Symp. on Circuits and Systems (ISCAS), 2000

[15] S. Brigati, F. Francesconi, P. Malcovati, D. Tonietto, A. Baschirotto and F. Maloberti, "Modeling sigma-delta modulator non-idealities in simulink," IEEE Int. Symp. on Circuits and Systems (ISCAS), 1999

[16] J. Xu, X. Wu, M. Zhao, R. Fan, H. Wang, X. Ma, and B. Liu, "Ultra low-FOM high-precision ΔΣ modulators with fully-clocked SO and zero static power quantizers," IEEE Custom Integrated Circuits Conference (CICC), 2011

[17] A. Peña-Perez, E. Bonizzoni and F. Maloberti, "A 88-dB DR, 84-dB SNDR very low-power single op-amp third-order ΣΔ modulator," IEEE Journal of Solid-State Circuits (JSSC), vol. 47 no. 9, 2012

[18] F. Ali and A. A. Hamoui, "Continuous-time ΔΣ modulators with noise-transfer-function enhancement," IEEE Int. Symp. on Circuits and Systems (ISCAS), 2008

[19] A. Peña-Perez, N. Kumar Y. B, E. Bonizzoni and F. Maloberti, "Slew-rate and gain enhancement in two stage operational amplifiers," IEEE Int. Symp. on Circuits and Systems (ISCAS), 2009

[20] Y. Chae and G. Han, "Low voltage, low power, inverter-based switched-capacitor delta-sigma modulator," IEEE Journal of Solid-State Circuits (JSSC), vol. 44 no. 2, 2009

[21] B. Gönen, F. Sebastiano, R. Quan, R. van Veldhoven and K. A. A. Makinwa, "A dynamic zoom ADC with 109-dB DR for audio applications," IEEE Journal of Solid-State Circuits (JSSC), vol. 52 no. 6, 2017

[22] R. H. M. Van Veldhoven and A. H. M. Van Roermund, "Robust sigma delta converters and their application in low-power highly-digitized flexible receivers," USA: Springer, 2011

[23] T. Christen, "A 15-bit 140-μW scalable-bandwidth inverter-based ΔΣ modulator for a MEMS microphone with digital output," IEEE Journal of Solid-State Circuits (JSSC), vol. 48 no. 7, 2013

[24] R. G. Vieru and R. Ghinea, "An ultra-low voltage sigma delta modulator with inverter based scalable amplifier," IEEE Int. Symp. on Electronics and Telecommunications (ISETC), 2012

[25] A. J. Ginés, A. Villegas, E. Peralías and A. Rueda, "Self-biased input common-mode generation for improving dynamic range and yield in inverter-based filters," IEEE Int. Conf. on Electronics, Circuits and Systems (ICECS), 2012

[26] M. Chae, J. Kim and W. Liu, "Fully-differential self-biased bio-potential amplifier," IET Electronics Letters, vol. 44 no. 24, 2008

[27] B. Razavi, "Design of analog CMOS integrated circuits," USA: McGraw-Hill, 2000

[28] O. Choksi and L. R. Carley, "Analysis of switched-capacitor common-mode feedback circuit," IEEE Trans. on Circuits and Systems II, vol. 50 no. 12, 2003

[29] J. Yeo, Y. Choi, J. Roh, G. Han, Y. Chae and S. Ham, "A current regulator for inverter-based massively column-parallel ΔΣ ADCs," IEEE Trans. on Circuits and Systems II, vol. 61 no. 4, 2014

[30] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of ap-amp imperfections: autozeroing, correlated double sampling, and chopper stabilization," Proceedings of the IEEE, vol. 84 no. 11, 1996

[31] K. Nagaraj, J. Vlach, T. R. Viswanathan and K. Singhal, "Switched-capacitor integrator with reduced sensitivity to amplifier gain," IET Electronics Letters, vol. 22 no. 21, 1986

[32] O. Oliaei, "Noise analysis of correlated double sampling sc integrators with a hold capacitor," IEEE Trans. on Circuits and Systems I, vol. 50 no. 9, 2003

[33] G. Léger, A. Gines, E. J. Peralías and A. Rueda, "On chopper effects in discrete-time ΣΔ modulators," IEEE Trans. on Circuits and Systems I: regular papers, vol. 57 no. 9, 2010

[34] Z. Yang, L. Yao and Y. Lian, "A 0.5-V 35-μW 85-dB DR double-sampled ΔΣ modulator for audio applications," IEEE Journal of Solid-State Circuits (JSSC), vol. 47 no. 3, 2012

[35] L. Yao, M. S. J. Steyaert, and W. Sansen, "A 1-V 140-μW 88-dB audio sigma-delta modulator in 90-nm CMOS," IEEE Journal of Solid-State Circuits (JSSC), vol. 39 no. 11, 2004

[36] M. Ortmanns, F. Gerfers, "Continuous-time sigma-delta A/D conversion fundamentals, performance limits and robust implementations," USA: Springer, 2006

[37] M. Tamaddon and M. Yavari, "A wideband time-based continuous-time sigma-delta modulator with 2nd order noise-coupling based on passive elements," Wiley Int. Journal of Circuit Theory and Applications (IJCTA), 2016

[38] L. Breems, M. Bolatkale, H. Brekelmans, S. Bajoria, J. Niehof et al. "A 2.2GHz continuous-time ΔΣ ADC with -102 dBc THD and 25 MHz bandwidth," IEEE Journal of Solid-State Circuits (JSSC), vol. 51 no. 8, 2016

[39] T. Kim, C. Han and N. Maghari, "A 7.2 mW 75.3 dB SNDR 10 MHz BW CT delta-sigma modulator using gm-C based noise shaped quantizer and digital integrator," IEEE Journal of Solid-State Circuits (JSSC), vol. 51 no. 8, 2016

[40] C. De Berti, P. Malcovati, L. Crespi and A. Baschirotto, "A 106 dB A-weighted DR low-power continuous-time ΣΔ modulator for MEMS microphones," IEEE Journal of Solid-State Circuits (JSSC), vol. 51 no. 7, 2016

[41] J. A. Cherry and W. M. Snelgrove, “Continuous-time delta-sigma modulators for high-speed A/D conversion: Theory, practice and fundamental performance limits,” USA: Kluwer, 2002

[42] O. Belotti, E. Bonizzoni and F. Maloberti, “Exact design of continuous-time sigma-delta modulators with multiple feedback DACs,” Springer Analog Integrated Circuits and Signal Processing, vol. 73 pp. 255-264, 2012

[43] Y. C. Chang, W. H. Chiu, C. C. Lin and T. H. Lin, “A 4MHz BW 69dB SNDR continuous-time delta-sigma modulator with reduced sensitivity to clock jitter,” IEEE Asian Solid-State Circuit Conference (ASSCC), 2011

[44] C. Chu, J. Anders, J. Becker and M. Ortmanns, “Finite GBW compensation technique for CT ΔΣ modulators with differentiator based ELD compensation,” IEEE New Circuits and Systems Conference (NEWCAS), 2015

[45] M. Keller, A. Buhmann, J. Sauerbrey, M. Ortmanns and Y. Manoli, “A comparative study on excess-loop-delay compensation techniques for continuous-time sigma-delta modulators. IEEE Trans. on Circuits and Systems I: regular papers, 2008

[46] M. Huang, D. Chen, Z. Wang, J. Guo et al. “A power-area efficient, 3-band, 2-RX MIMO, TD-LTE receiver with direct-coupled ADC,” Wiley Int. Journal of Circuit Theory and Applications (IJCTA), 2015

[47] T. C. Wang, Y. H. Lin and C. C. Liu, “A 0.022mm² 98.5dB SNDR hybrid audio ΔΣ modulator with digital ELD compensation in 28nm CMOS,” IEEE Journal of Solid-State Circuits (JSSC), vol. 50 no. 11, 2015

[48] C. Y. Ho, C. Liu, C. H. Lo, H. C. Tsai, T. C. Wang and Y. H. Lin, “A 4.5 mW CT self-coupled ΔΣ modulator with 2.2 MHz BW and 90.4 dB SNDR using residual ELD compensation,” IEEE Journal of Solid-State Circuits (JSSC), vol. 50 no. 12, 2015

[49] S. Tao, J. Garcia, S. Rodriguez and A. Rusu, “Analysis of exponentially decaying pulse shape DACs in continuous-time sigma-delta modulators,” IEEE Int. Conf. on Electronics, Circuits and Systems (ICECS), 2012

[50] J. G. Jo, J. Noh and C. Yoo, “A 20-MHZ bandwidth continuous-time sigma-delta modulator with jitter immunity improved full clock period SCR (FSCR) DAC and high-speed DWA,” IEEE Journal of Solid-State Circuits (JSSC), vol. 46 no. 11, 2011

[51] D. Feng, F. Maloberti, S. W. Sin, S. P. U, and R. P. Martins, “Jitter-resistant capacitor-based sine-shaped DAC for continuous-time sigma-delta modulators,” IEEE Int. Symp. on Circuits and Systems (ISCAS), 2014

[52] A. Sukumaran and S. Pavan, “Low power design techniques for single-bit audio continuous-time delta sigma ADCs using FIR feedback,” IEEE Journal of Solid-State Circuits (JSSC), vol. 49 no. 11, 2014

[53] S. Billa, A. Sukumaran and S. Pavan, “A 280µW 24kHz-BW 98.5dB-SNDR chopped single-bit CT ΔΣM achieving <10Hz 1/f noise corner without chopping artifacts,” IEEE Int. Solid-State Circuit Conference (ISSCC), 2016

[54] S. Billa, A. Sukumaran and S. Pavan, “Analysis and design of continuous-time delta-sigma converters incorporating chopping,” IEEE Journal of Solid-State Circuits (JSSC), vol. 52 no. 9, 2017

[55] Y. H. Leow, H. Tang, Z. C. Sun and L. Siek, “A 1 V 103 dB 3rd-order audio continuous-time ΔΣ ADC with enhanced noise shaping in 65 nm CMOS,” IEEE Journal of Solid-State Circuits (JSSC), vol. 51 no. 11, 2016

[56] M. Park and M. H. Perrott, “A 78 dB SNDR 87 mW 20 MHz bandwidth continuous-time ΔΣ ADC with VCO-based integrator and quantizer implemented in 0.13 µm CMOS,” IEEE Journal of Solid-State Circuits (JSSC), vol. 44 no. 12, 2009

[57] K. Reddy, S. Rao, R. Inti, B. Young, A. Elshazly, M. Talegaonkar and P. K. Hanumolu, “A 16-mW 78-dB SNDR 10-MHz BW CT ΔΣ ADC using residue-cancelling VCO-based quantizer,” IEEE Journal of Solid-State Circuits (JSSC), vol. 47 no. 12, 2012

[58] S-J. Huang, N. Egan, D. Kesharwani, F. Opteynde and M. Ashburn, “A 125MHz-BW 71.9dB-SNDR VCO-based CT ΔΣ ADC with segmented phase-domain ELD compensation in 16nm CMOS,” Int. Solid-State Circuit Conference (ISSCC), 2017

[59] T. Kim, C. Han and N. Maghari, “An 11.4mW 80.4dB-SNDR 15MHz-BW CT delta-sigma modulator using 6b double-noise-shaped quantizer,” Int. Solid-State Circuit Conference (ISSCC), 2017

[60] K. Lee, Y. Yoon and N. Sun, “A scaling-friendly low-power small-area ΔΣ ADC with VCO-based integrator and intrinsic mismatch shaping capability,” IEEE Journal on Emerging and Selected Topics in Circuits and Systems (JETCAS), vol. 5 no. 4, 2015

[61] C-C. Tu, Y-K. Wang and T-H. Lin, “A low-noise area-efficient chopped VCO-based CTDSM for sensor applications in 40-nm CMOS,” IEEE Journal of Solid-State Circuits (JSSC), vol. 52 no. 10, 2017

[62] M. Z. Straayer and M. H. Perrott, “A 12-bit, 10-MHz bandwidth, continuous-time ΣΔ ADC with a 5-bit, 950-MS/s VCO-based quantizer,” IEEE Journal of Solid-State Circuits (JSSC), vol. 43 no. 4, 2008

[63] S. Li, A. Mukherjee and N. Sun, “A 174.3-dB FoM VCO-based CT ΔΣ modulator with a fully-digital phase extended quantizer and tri-level resistor DAC in 130-nm CMOS,” IEEE Journal of Solid-State Circuits (JSSC), vol. 52 no. 7, 2017

[64] S. Zeller, C. Muenker and R. Weigel, “A X-coupled differential single-opamp resonator for low-power continuous-time ΣΔ ADCs,” IEEE Conf. on Ph.D Research in Microelectronics and Electronics (PRIME), 2011

[65] T-Y. Lo, “A 102dB dynamic range audio sigma-delta modulator in 40nm CMOS,” IEEE Asian Solid-State Circuit Conference (ASSCC), 2011

[66] I. Ahmed, J. Cherry, A. Hasan, A. Nafee, D. Halupka, Y. Allasasmeh and M. Snelgrove, “A low-power Gm-C based CT- ΔΣ audio-band ADC in 1.1V 65nm CMOS,” IEEE Symp. on VLSI Circuits (VLSI Circuits), 2015

Appendix

In this appendix are presented the discrete-time equivalent transfer functions calculations using IIT of the different figure 4.11 integrating paths in presence of ELD. The first-order integrating path integrates the DAC impulse response $h_{dacELD}(t)$, and its impulse response $h_{1ELD}(t)$ is:

$$h_{dacELD}(t) = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ 1, & for\ 0.5 \le t \le 1.5 \\ 0, & for\ t > 1.5 \end{cases}$$

$$h_{1ELD}(t) = \begin{cases} \int_0^t 0 \times dt, & for\ 0 \le t \le 0.5 \\ h_{1ELD}(0.5) + \int_{0.5}^t 1 \times dt, & for\ 0.5 \le t \le 1.5 \\ h_{1ELD}(1.5) + \int_{1.5}^t 0 \times dt, & for\ t > 1.5 \end{cases} = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ t - 0.5, & for\ 0.5 \le t \le 1.5 \\ 1, & for\ t > 1.5 \end{cases}$$

Sampling it gives $h_{1ELD}(n) = \begin{cases} 0, & for\ n = 0 \\ 0.5, & for\ n = 1 \\ 1, & for\ n \ge 2 \end{cases}$ and taking the z-transform yields $H_{1ELD}(z)$:

$$H_{1ELD}(z) = \sum_{n=0}^{\infty} h_{1ELD}(n) z^{-n} = 0 + 0.5z^{-1} + \sum_{n=2}^{\infty} z^{-n} = 0.5z^{-1} + \sum_{n=1}^{\infty} (z^{-n}) - z^{-1}$$

$$H_{1ELD}(z) = \frac{z^{-1}}{(1 - z^{-1})} - 0.5z^{-1} = \frac{0.5z^{-1} + 0.5z^{-2}}{(1 - z^{-1})}$$

To get the second-order integrating path impulse response $h_{2ELD}(t)$, $h_{1ELD}(t)$ is integrated:

$$h_{2ELD}(t) = \begin{cases} \int_0^t 0 \times dt, & for\ 0 \le t \le 0.5 \\ h_{2ELD}(0.5) + \int_{0.5}^t (t - 0.5) \times dt\,, & for\ 0.5 \le t \le 1.5 \\ h_{2ELD}(1.5) + \int_{1.5}^t 1 \times dt\,, & for\ t > 1.5 \end{cases} = \begin{cases} 0, & for\ 0 \le t \le 0.5 \\ \frac{t^2 - t + 0.25}{2}, & for\ 0.5 \le t \le 1.5 \\ t - 1, & for\ t > 1.5 \end{cases}$$

The discrete-time equivalent impulse response is $h_{2ELD}(n) = \begin{cases} 0, & for\ n = 0 \\ 0.125, & for\ n = 1 \\ n - 1, & for\ n \ge 2 \end{cases}$ leading to $H_{2ELD}(z)$:

$$H_{2ELD}(z) = \sum_{n=0}^{\infty} h_{2ELD}(n) z^{-n} = 0 + 0.125z^{-1} + \sum_{n=2}^{\infty} (n-1) z^{-n} = 0.125z^{-1} + \sum_{n=2}^{\infty} n z^{-n} - \sum_{n=2}^{\infty} z^{-n}$$

$$H_{2ELD}(z) = 0.125z^{-1} + \left(\sum_{n=1}^{\infty} (n z^{-n}) - z^{-1} \right) - \left(\sum_{n=1}^{\infty} (z^{-n}) - z^{-1} \right) = 0.125z^{-1} + \sum_{n=1}^{\infty} n z^{-n} - \sum_{n=1}^{\infty} z^{-n}$$

$$H_{2ELD}(z) = 0.125z^{-1} + \frac{z^{-1}}{(1 - z^{-1})^2} - \frac{z^{-1}}{1 - z^{-1}} = \frac{0.125z^{-1} + 0.75z^{-2} + 0.125z^{-3}}{(1 - z^{-1})^2}$$

Finally, deducing $h_{3ELD}(t)$ by repeating the integration operation on second-order branch impulse response:

$$h_{3ELD}(t)=\begin{cases}\int_0^t 0\times dt, & for\ 0\le t\le 0.5\\ h_{3ELD}(0.5)+\int_{0.5}^t \frac{t^2-t+0.25}{2}\times dt, & for\ 0.5\le t\le 1.5\\ h_{3ELD}(1.5)+\int_{1.5}^t (t-1)\times dt, & for\ t>1.5\end{cases}=\begin{cases}0, & for\ 0\le t\le 0.5\\ \frac{4t^3-6t^2+3t-0.5}{24}, & for\ 0.5\le t\le 1.5\\ \frac{3t^2-6t+3.25}{6}, & for\ t>1.5\end{cases}$$

Then sampling it gives $h_{3ELD}(n)=\begin{cases}0, & for\ n=0\\ \frac{1}{48}, & for\ n=1\\ \frac{3n^2-6n+3.25}{6}, & for\ n\ge 2\end{cases}$ and the z-transform gives $H_{3ELD}(z)$:

$$H_{3ELD}(z)=\sum_{n=0}^{\infty} h_{3ELD}(n)z^{-n}=0+\frac{1}{48}z^{-1}+\sum_{n=2}^{\infty}\frac{3n^2-6n+3.25}{6}z^{-n}$$

$$H_{3ELD}(z)=\frac{1}{48}z^{-1}+\frac{24}{48}\sum_{n=2}^{\infty}n^2z^{-n}-\frac{48}{48}\sum_{n=2}^{\infty}nz^{-n}+\frac{26}{48}\sum_{n=2}^{\infty}z^{-n}$$

$$H_{3ELD}(z)=\frac{1}{48}z^{-1}+\frac{24}{48}\left(\sum_{n=1}^{\infty}(n^2z^{-n})-z^{-1}\right)-\frac{48}{48}\left(\sum_{n=1}^{\infty}(nz^{-n})-z^{-1}\right)+\frac{26}{48}\left(\sum_{n=1}^{\infty}z^{-n}-z^{-1}\right)$$

$$H_{3ELD}(z)=\frac{-1}{48}z^{-1}+\frac{24}{48}\sum_{n=1}^{\infty}n^2z^{-n}-\frac{48}{48}\sum_{n=1}^{\infty}nz^{-n}+\frac{26}{48}\sum_{n=1}^{\infty}z^{-n}$$

$$H_{3ELD}(z)=\frac{-1}{48}z^{-1}+\frac{24}{48}\frac{z^{-1}(1+z^{-1})}{(1-z^{-1})^3}-\frac{48}{48}\frac{z^{-1}}{(1-z^{-1})^2}+\frac{26}{48}\frac{z^{-1}}{1-z^{-1}}$$

$$H_{3ELD}(z)=\frac{\frac{1}{48}z^{-1}+\frac{23}{48}z^{-2}+\frac{23}{48}z^{-3}+\frac{1}{48}z^{-4}}{(1-z^{-1})^3}$$