AMS/RF Design for Reliability Methodology: a Reliable RF Front-end Design

Pietro Maris Ferreira

Thèse

présentée pour obtenir le grade de docteur

de l'Ecole Nationale Supérieure des Télécommunications

Spécialité : Electronique et Communications

Pietro MARIS FERREIRA

Méthodologie de conception AMS/RF pour la fiabilité : conception d'un frontal RF fiabilisé

Soutenance le 23 septembre 2011 devant le jury composé de

Fernando Silveira	**Rapporteurs**
Patrice Gamand	
Amara Amara	**Examinateurs**
Lírida Alves de Barros Naviner	
Hervé Petit	**Directeurs de thèse**
Jean-François Naviner	

My dear Flavia.

ACKNOWLEDGEMENTS

First of all, I would like to thank the sponsoring gave to my PhD research from Institut TELECOM. Also, STIC-AmSud had an important role in my work while funding the project NANORadio a cooperation of TELECOM ParisTech, Universidad de la Republica (Uruguay), Universidade Federal do Rio Grande do Sul (Brazil) and Universidade Federal do Rio de Janeiro (Brazil). Both funding made possible my dream to become a PhD in reality. Although I wished to personally thank each one of my colleagues at TELECOM ParisTech, I think that I cannot express my acknowledgements to them in this single page. Unfortunately, I may miss some of them.

I would like to express my love to my Lord who advised me and to my wife Flavia Tovo who had held up me. Both had guided me in this challenge of a PhD degree in France. Moreover, I would like to thank my family while seeding the valors during my background education. I also would like to express my pleasure to be advised by Jean-François Naviner and Hervé Petit. They have presented a special team play with complementary qualities and activities, while directing my PhD research.

In few words, I express my greetings to TELECOM ParisTech, the Department COMELEC and especially to the research group SIAM, as they welcome me with open heart. Inside such a group, I have shared a lot of experience with: Alban Gruget and Mariem Slimani, since the beginning; Fatima Ghanem, Germain Pham, and Hao Cai, after the first-half path. I realize that the discussions with Hussein Fakhoury had grown in great ideas; and the Arwa Ben Dhia's revision has improved my thesis and my writing skills. Furthermore, I am glad with the friendship found in all members of the group. I fell I am not able to enumerate all of them and highlight their participation in my achievement.

The inspiration in the end of this PhD life step and the beginning of my professional next step is found in George E. P. Box's thought. As it is: "Essentially, all models are wrong, but some are useful."

RÉSUMÉ ÉTENDU

INTRODUCTION

Le développement des technologies CMOS à l'échelle nanométrique a fait émerger de nombreux défis sur le rendement et la fiabilité des composants [1]. Les prochaines générations de circuits AMS et RF souffriront d'une augmentation du taux de défaillance durant le temps d'opération. La méthode classique de conception des circuits sur puce repose sur les compromis entre : surface, consommation de puissance et vitesse d'opération. L'optimum est le point où on obtient le maximum ou le minimum spécifié pour les caractéristiques du circuit.

L'augmentation de la variabilité et la diminution de la fiabilité des transistors a prouvé que ce point optimal est devenu une région étant donné que le rendement attendu se déplace sous l'effet du vieillissement. La combinaison de ces deux événements impose de nouveaux défis au concepteur par rapport à la méthode classique de conception des circuits sur puce. Le rendement est le rapport entre le nombre des puces qui sont conformes aux spécifications de conception et la totalité produite. Cette grandeur est mesurée au début du temps d'opération des circuits et juste après la production des puces, c'est-à-dire à $t = 0$. Pourtant, le rendement ne prend pas en compte l'évolution du taux de défaillance en fonction du temps. La fiabilité peut être définie comme la probabilité qu'un dispositif exécute une fonction exigée dans des conditions indiquées pendant une période indiquée [1].

Depuis l'avènement des téléphones mobiles, les architectures de frontal RF ont nécessité des innovations technologiques accrues dans un temps de commercialisation réduit. A côté de la recherche des architectures de frontal radio, les technologies des circuits intégrées sur puce ont suivi la célèbre Loi de Moore. Ces technologies ont permis d'obtenir les dimensions réduites, une faible consommation d'énergie et une augmentation de la vitesse. C'est pour cela que la surface, la consommation d'énergie et la vitesse sont devenues les éléments clés d'un compromis pour les circuits sur puce en général. Pour les circuits analogiques d'autres critères de performance sont définis, comme : le gain, le bruit et la linéarité. Le défi sera de trouver le compromis entre ces éléments.

Avec l'accroissement de la variabilité, la conception proche du point optimal est maintenant une région composée par un, deux ou trois σ selon le rendement souhaité. La solution la plus simple est de concevoir selon les caractéristiques maximales et minimales spécifiées en prenant le pire cas de ces caractéristiques selon la variabilité. Le coût imposé par ces marges mène à un circuit surdimensionné.

Les phénomènes de vieillissement des transistors MOS sont connus depuis les années 70, mais on sait que le circuit doit être soumis à un environnement qui va au delà de ses conditions normales de fonctionnement. Ce sujet avait disparu de l'état-de-l'art pendant quelques décennies puis est réapparu avec les circuits nanométriques. Pourquoi le vieillissement est-il devenu aujourd'hui si important? Les tendances du vieillissement sont très préoccupantes parce que la tension de seuil et la tension d'alimentation n'ont pas suivi le même rythme durant l'évolution des nœuds technologiques des circuits sur puce. Par conséquent, les conditions de fonctionnement auparavant très improbables qui pourraient contraindre le circuit sont devenues des conditions d'environnement courantes.

Dans ce scénario, nous sommes motivés à innover dans la conception d'un frontal RF en CMOS 65 nm (voir Figure 1). Nous proposons de nouvelles méthodologies d'analyse et de synthèse, qui comprennent la variabilité et le vieillissement en mettant en évidence les compromis présents parmi les critères de la conception. Le frontal RF a trois principaux blocs de construction: le BLIXER [2], re-

groupant un balun, un amplificateur large bande à faible bruit et un mélangeur I-Q ; l'oscillateur contrôlé numériquement (DCO) [3], et l'amplificateur de gain programmable (PGA) avec le filtre passe-bas [4].

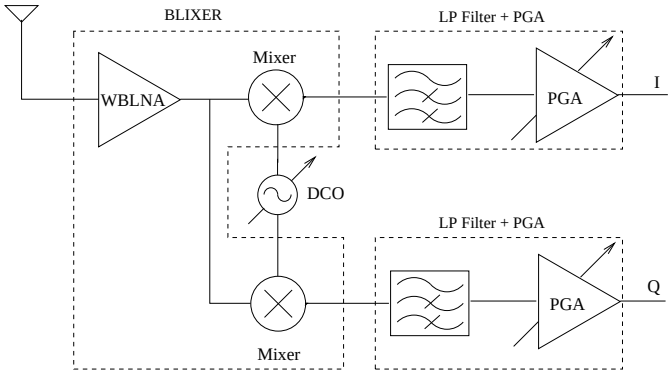


Figure 1: L'architecture du frontal radio pour les applications multistandards de communication sans fils : illustration [5].

L'objectif principal de ce travail est d'améliorer la conception de circuits du frontal RF basée sur la recherche des nouveaux compromis imposés par la variabilité du transistor et la dégradation par vieillissement. En proposant sur ces deux agents de variation des caractéristiques des circuits sur puce comme critères de conception, nous concevrons un frontal radio fiable pour une application RF multistandard. Ainsi, nous pouvons énumérer nos sous-objectifs qui compléteront cet objectif principal, comme

1. L'étude de la physique des phénomènes de vieillissement et des conditions de conception du circuit pour éviter le vieillissement et la variabilité des transistors

2. L'étude des sources de dégradation (vieillissement et variabilité) et ses tendances dans les technologies à l'échelle nanométrique

3. L'étude des méthodes de conception classique, comparant avec les besoins imposés par la variabilité et la fiabilité des composants

4. Une proposition de conception des circuits fiables pour le cas d'étude dans une approche montante (*bottom-up*)

5. Une proposition de conception de l'architecture fiable pour le cas d'étude dans une approche descendante (*top-down*)

6. Une proposition de généralisation de la méthode de synthèse qui relie les approchesb *bottom-up* et *top-down*

7. La comparaison des compromis imposés par le vieillissement et par la variabilité des composants pour la technologie CMOS 65 nm

LA VARIABILITÉ ET LA FIABILITÉ DANS LES TECHNOLOGIES AVANCÉES

La variation du processus d'intégration est observée comme la déviation de la valeur des paramètres des circuits sur puce. Ces variations sont dues à différentes causes et ont de nombreuses conséquences. La variabilité du circuit est l'ensemble des variations de paramètres du processus de fabrication, l'imperfection des masques d'intégration et les impacts sur la qualité des circuits entre les différents échantillons.

La variation de performance des circuits intégrés est liée à deux groupes de facteurs [6] :

- *Facteurs environnementaux* - Ils interviennent durant l'opération du circuit comme la source de l'alimentation, le couplage de bruit et la température. Ils sont appelés variations dynamiques et sont fortement dépendant des agents externes et du schéma du circuit.

- *Facteurs physiques* - Ils apparaissent durant la fabrication et ils changent la structure physique des circuits. Ils incluent les RDF (*Random Dopant Fluctuations*) et la LER (*Line Edge Roughness*).

Ainsi, les variations des paramètres physiques peuvent être classées en deux catégories : D2D (*Die-to-die*) et WID (*Within-die*) [7]. Les variations du type D2D sont des variations globales entre différentes puces. Les variations du type WID sont des variations internes à une puce et entre les différents composants qui forment les fonctions intégrées.

La conception de circuits RF plus fiables requiert la connaissance des phénomènes physiques qui réduisent la fiabilité. Les principaux phénomènes physiques à l'origine du vieillissement des dispositifs sont :

- *Hot Carrier Injection (HCI)* - un phénomène irréversible où une charge a assez d'énergie pour franchir une barrière de potentiel, (généralement du côté du drain, comme illustré par la Figure 2) en créant un défaut d'interface. Il a lieu quand V_{GD} est plus grand que zéro et V_{GS} est très élevée [1]. Il peut être évité en réduisant le temps pendant lequel les transistors sont en inversion forte, par le contrôle de V_{GD} et V_{GS} .

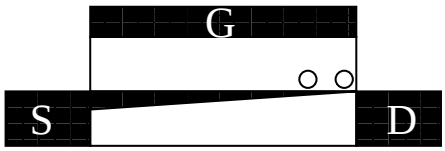


Figure 2: Phénomène HCI: illustration.

- *Negative Bias Temperature Instability (NBTI)* - un phénomène réversible de dérive des paramètres électriques du transistor sous une tension négative sur la grille et à hautes températures (illustré par la Figure 3). La récupération des paramètres a lieu quand les stress sont arrêtés. Cela implique qu'il est possible d'avoir moins d'impact si le temps de récupération est proportionnellement plus grand que le temps de stress.

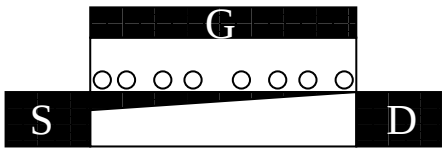


Figure 3: Phénomène NBTI: illustration.

- *Time Dependent Dielectric Breakdown (TDDB)* - un phénomène statistique de rupture continue du diélectrique (illustré par la Figure 4), qui par conséquent induit un manque de performance. La probabilité d'avoir le TDDB peut être réduite en contrôlant le courant de fuite par le biais de la réduction de la tension de grille et de la surface du transistor [8].

- *Electromigration (EM)* - un phénomène de transport de masse dans la couche de métal. L'EM peut être évitée par la réduction de la longueur et l'augmentation de la largeur des connexions. Aussi, les circuits qui n'utilisent pas de composants passifs (sujet à l'EM) sont plus robustes.

Il n'existe pas de modèle universellement accepté pour la dégradation de fiabilité due aux phénomènes physiques du vieillissement des composants. Nous présenterons les plus importants de l'état-de-l'art.

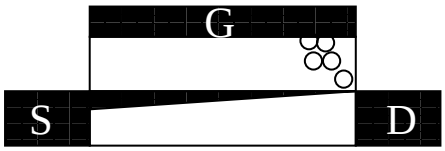


Figure 4: Phénomène TDDB: illustration.

LE MODÈLE DE RÉACTION ET DIFFUSION Le NBTI et le HCI ont des modèles similaires, car les paramètres électriques changent de la même façon après stress. Le modèle le plus accepté est le modèle dit de Réaction et Diffusion. Il est décrit en deux parties :

1. ***Réaction***: processus de rupture de liaison chimie Si-H dans un défaut d'interface oxyde/substrat et hydrogène libre;

2. ***Diffusion***: mouvement de l'hydrogène vers l'oxyde jusqu'à la grille.

Le processus de génération de défaut est décrit pour une loi de puissance du temps de stress, et la solution du modèle de Réaction et Diffusion est présenté en [9]. Le modèle de Réaction et Diffusion peut être illustré par la Figure 5

LE MODÈLE DE RUPTURE DE L'OXYDE La probabilité de défaillance pour un événement TDDB est bien décrite avec une distribution de Weibull [10]

$$F(x) = 1 - \exp\left(\left(-\frac{x}{\alpha}\right)^{\beta}\right); \tag{1}$$

où F est la probabilité cumulative de défaut, x peut être charge ou temps, α est la vie caractéristique pour 63% de probabilité de défaillance et β est le paramètre de forme de Weibull. Après rupture, le chemin de conduction est modélisé par une admittance de faible valeur [11]. Ceci modélise bien les courants de fuite, l'augmentation du bruit, la réduction de fréquence d'opération et la transconductance du transistor.

LE MODÈLE DE L'ELECTROMIGRATION La dégradation par l'EM est évaluée avec le temps moyen de défaillance (MTTF), qui est une méthode d'approximation statistique décrite par [12]

$$MTTF = AJ_e^n \exp\left(\frac{E_a}{kT}\right); \tag{2}$$

où A est une constante caractéristique de la couche de métal et du processus de fabrication, J_e est la densité de courant, E_a est l'énergie d'activation de l'EM, k est la constante de Boltzmann et T la température. n est dépendant du résidu de stress et de la densité de courant [13]. Après l'EM, la couche de métal perd sa caractéristique de haute conductance et est modélisée par une résistance parasite en série.

ETAT-DE-L'ART DES MÉTHODOLOGIES DE CONCEPTION

L'utilisation des technologies CMOS nanométriques entraîne d'importants défis imposés aux méthodologies de conception de circuits. Défis comprenant [14] :

- L'apparition de courants de fuite non négligeables et l'impact sur la consommation d'énergie;

- L'augmentation de la variabilité des paramètres du procédé technologique (V_{th}, le niveau de dopage, largeurs, longueurs, et autres) ;

- La réduction de la source d'alimentation sans que la tension de seuil suive ce rythme, ce qui réduit la dynamique du signal disponible pour les circuits AMS/RF ;
- L'apparition de nouveaux matériaux, comme les diélectriques à haut k, et de nouveaux dispositifs, comme FinFET et CNT-FET ;
- L'importance croissante des phénomènes de dégradation bien connus, comme HCI, NBTI et EM, et l'avènement de phénomènes de dégradation qui n'étaient pas observés avant, comme SDB et SM.

Ainsi, nous avons réalisé une étude de l'état-de-l'art des méthodes de conception classique, comparant les besoins imposés par la variabilité et la fiabilité des composants.

MÉTHODES CLASSIQUES DE CONCEPTION La conception des circuits intégrés peut être divisée en trois parties : la validation de l'architecture en utilisant des modèles comportementaux, la réalisation des schémas électriques, la synthèse des masques d'intégration. La plupart des méthodes de conception dans les différents niveaux hiérarchiques s'appuient sur de puissants outils d'optimisation numérique couplés à des outils d'estimation des caractéristiques, comme illustré dans la Figure 6. Dans l'ensemble de cette partie, nous discutons les avantages et les désavantages des approches *top-down* et *bottom-up*. Ensuite, nous présentons l'état-de-l'art des outils d'optimisation numérique et des outils d'estimation des performances.

LES MÉTHODES DE CONCEPTION POUR LA VARIABILITÉ La prise en compte de la variabilité dans les méthodes de conception est souvent réalisée par un changement de l'estimateur des caractéristiques des circuits. Les solutions les plus utilisées peuvent être distinguées entre la simulation des *corners* du circuit au pire cas et la simulation de Monte Carlo (illustré dans la Figure 7). Les plus grands désavantages d'une solution basée sur un simulateur sont le coût de calcul et le faible couplage avec l'outil d'optimisation. L'avantage est la précision qui peut être obtenue en utilisant des modèles physiques et des résultats de la caractérisation de circuit sur puce. Une autre solution est de changer l'évaluation dans l'optimisateur en utilisant aussi l'écart-type de la caractéristique et le rendement désiré. Cette solution réduit le coût de calcul et augmente le couplage entre l'estimateur et l'optimisateur au prix de l'utilisation d'un modèle coûteux et moins précis.

L'état-de-l'art ne présente pas un consensus dans le choix de la prise en compte de la variabilité en utilisant des outils commerciaux avec un seul outil de conception automatique. Depuis 2000, S. Nassif a souligné les besoins d'une conception qui prend en compte la variabilité [15]. Actuellement, G. Yu et P. Li ont proposé que la conception analogique soit optimisée non seulement pour les performances nominales, mais aussi pour la prise en compte de la variabilité, afin de maintenir un rendement raisonnable [16]. H. Onodera va plus loin en proposant que les méthodes et les outils de conception automatique doivent tolérer, atténuer, ou même exploiter la variabilité par des techniques appropriées à la conception avec la variabilité [17]. En 2009, V. Wang et al. ont introduit la formulation d'un modèle simplifié pour réduire la lacune entre les méthodes statistiques existantes et la conception de circuits [18].

LES MÉTHODES DE CONCEPTION POUR LA FIABILITÉ La prise en compte de la fiabilité dans les méthodes de conception est, à l'état-de-l'art, le développement de l'estimateur des caractéristiques des circuits vieillis (illustré dans la Figure 8). Ce concept a été ouvert avec les outils BERT (*Berkeley Reliability Tools*) et il est utilisé dans la modélisation des phénomènes de vieillissement [19], dans la simulation électrique des circuits vieillis [20, 21] et dans le calcul du temps de vie des composants [22].

L'avantage d'une approche de simulation électrique est la précision des résultats et le désavantage est le coût du calcul. Pourtant, la prise en compte de la fiabilité reste très peut utilisée dans les outils de simulation et les kits d'intégration commerciaux. Ainsi d'un côté, la recherche actuelle vise à améliorer les outils de simulation sans être liée aux données des kits d'intégration commerciaux, d'un autre côté

elle vise à améliorer les techniques de modélisation des phénomènes de vieillissement. Dans l'analyse de la fiabilité on peut distinguer : l'analyse pour un corner nominal, c'est-à-dire sans prendre en compte la variabilité ou en prenant un pire cas de variabilité ; et l'analyse en prenant en compte la variabilité, c'est-à-dire la fiabilité des points de simulation de type Monte Carlo.

CONCEPTION D'UN FRONTAL RADIO FIABLE DANS UNE APPROCHE *bottom-up*

BLIXER Nous avons proposé la conception d'un BLIXER fiable (schéma illustré dans la Figure 9). D'abord, nous avons caractérisé la variabilité et le vieillissement présent en CMOS 65 nm en utilisant les caractéristiques de courant de drain et de transconductance du transistor. Ensuite, nous avons proposé un modèle constructif pour estimer les caractéristiques du BLIXER et la variation des ses caractéristiques par rapport à la variation de courant de drain et de transconductance des transistors. Avec une analyse de sensibilité, nous avons identifié les transistors les plus sensibles et les impacts imposés au BLIXER en raison de la variation des caractéristiques des transistors. Avec ces informations, nous avons trouvé les conditions de polarisation et de dimensionnement qui réduisent les phénomènes de vieillissement. Les résultats obtenus sont illustrés parle Tableau 1 et les Figures 10 et 11.

Table 1: Résultats de simulation du BLIXER : performance typique.

Frequency (GHz)	1.0	2.4	5.0
Power (mW) Consumption	5.55	5.56	5.59
Differential Gain (dB)	13.5	13.5	14.3
NF_{max} (dB)	4.0	4.5	5.3
IP3 (dBm)	4.4	3.4	-0.73
S11 (dB)	-17.6	-15.4	- 13.3

Ces résultats de simulation du circuit typique sont cohérents avec les spécifications d'un frontal radio multistandard. Malgré la variabilité des processus d'intégration et la disparité des dimensions, nous avons observé que les caractéristiques du BLIXER sont conformes aux spécifications pour un rendement supérieur à 90 %. Par ailleurs, le vieillissement du BLIXER est négligeable selon la distribution de Poisson de la consommation de puissance ajusté avec 99,9 % de confiance [23].

DCO Nous avons proposé une comparaison entre deux DCO : l'un conçu sans les contraintes de vieillissement et variabilité et l'autre comme un circuit fiabilisé. Pour cela, nous avons choisi de concevoir un DCO pour les applications à 1 GHz avec le schéma proposé par [24] (illustré dans la Figure 12). Pour la conception classique du DCO, nous avons évalué le compromis entre bruit de phase, plage de fréquence pour le verrouillage et la consommation de puissance. Pour la conception du DCO fiabilisé, nous avons estimé la variation de ces caractéristiques par rapport à la variabilité et le vieillissement.

Nous avons développé une analyse de fiabilité du DCO qui nous donne les informations nécessaires pour la conception des circuits plus fiables. En concevant le DCO fiabilisé et non-fiabilisé, nous obtenons une réduction de la dégradation de la fréquence d'une valeur entre 15% et 30 %. Egalement, le DCO fiabilisé a un temps de vie cinq fois plus grand que le non-fiabilisé, si nous fixons la dégradation de la fréquence a un maximum de 2 %. Les inconvénients de la conception du DCO fiabilisé sont l'augmentation du bruit de phase et la réduction de la plage de fréquence disponible pour le verrouillage. Ceci pourrait être négligé dans la spécification d'un standard radio [25].

Les résultats du DCO sont illustrés pas les Figures 13(a), 13(b), 14(a), 14(b), 15(a), 15(b), 16, 17(a), 17(b), 18(a), 18(b).

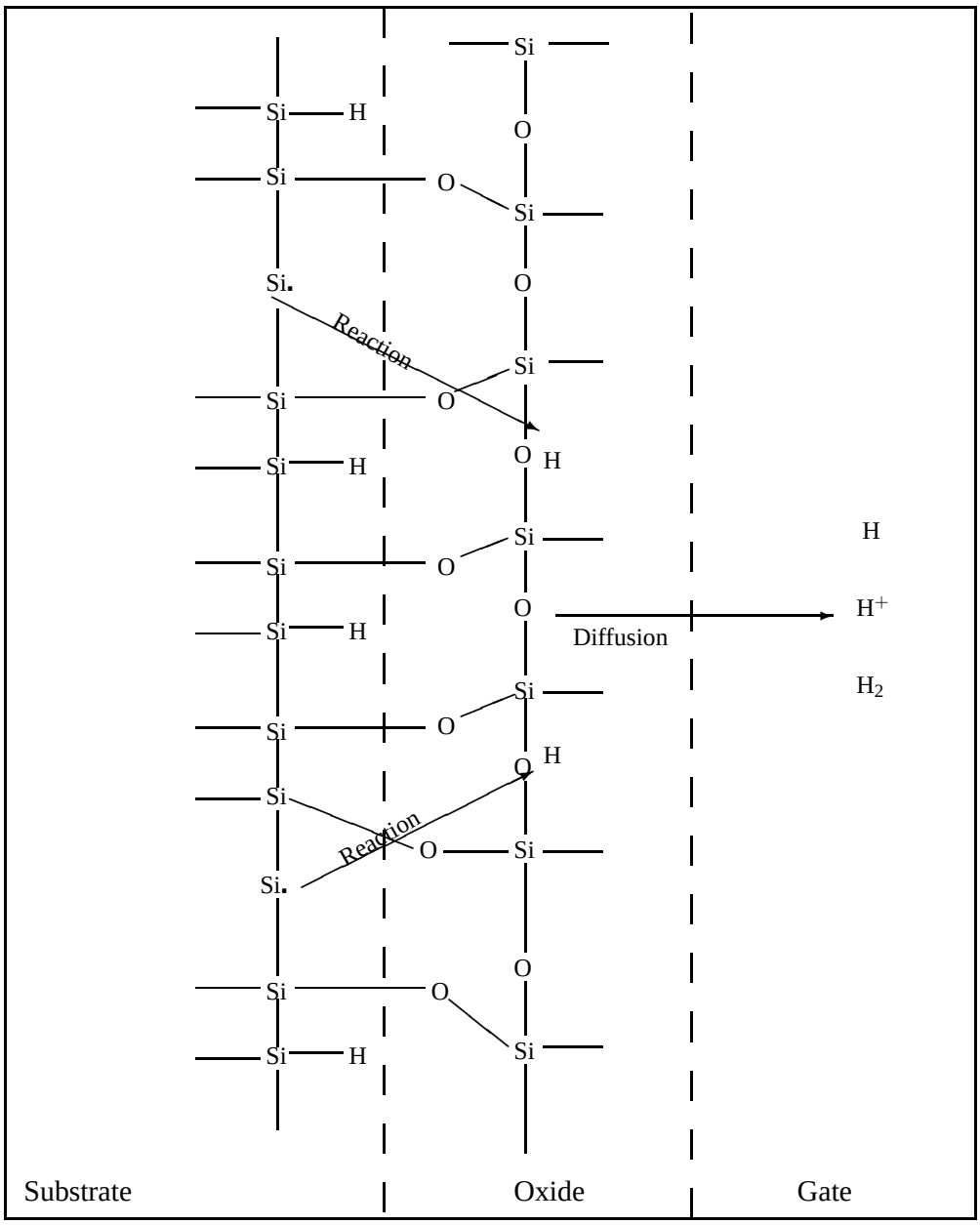


Figure 5: Le modèle de Réaction et Diffusion: illustration.

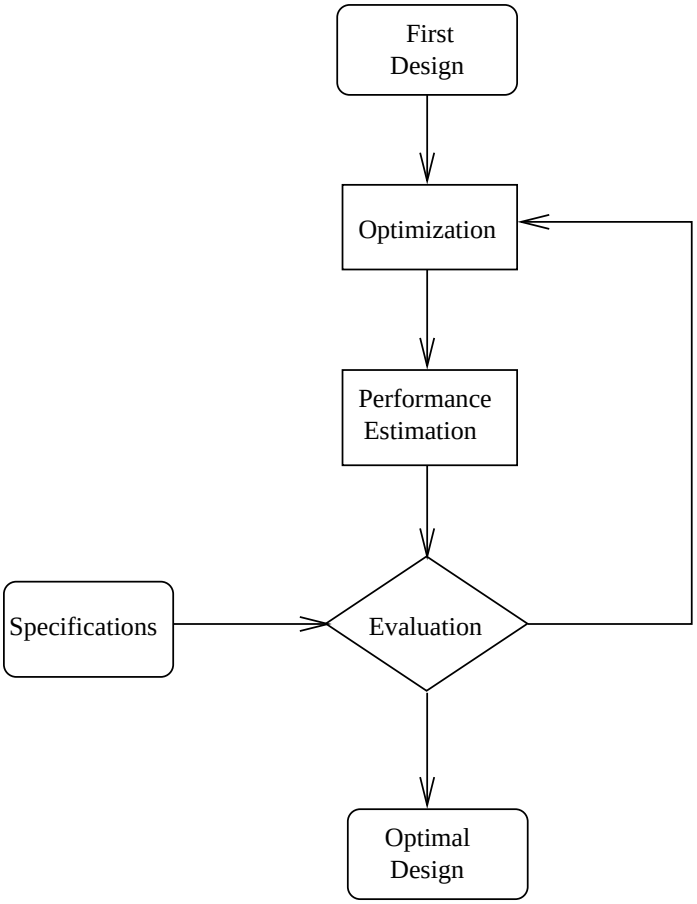


Figure 6: Flot de conception classique : illustration des étapes de conception.

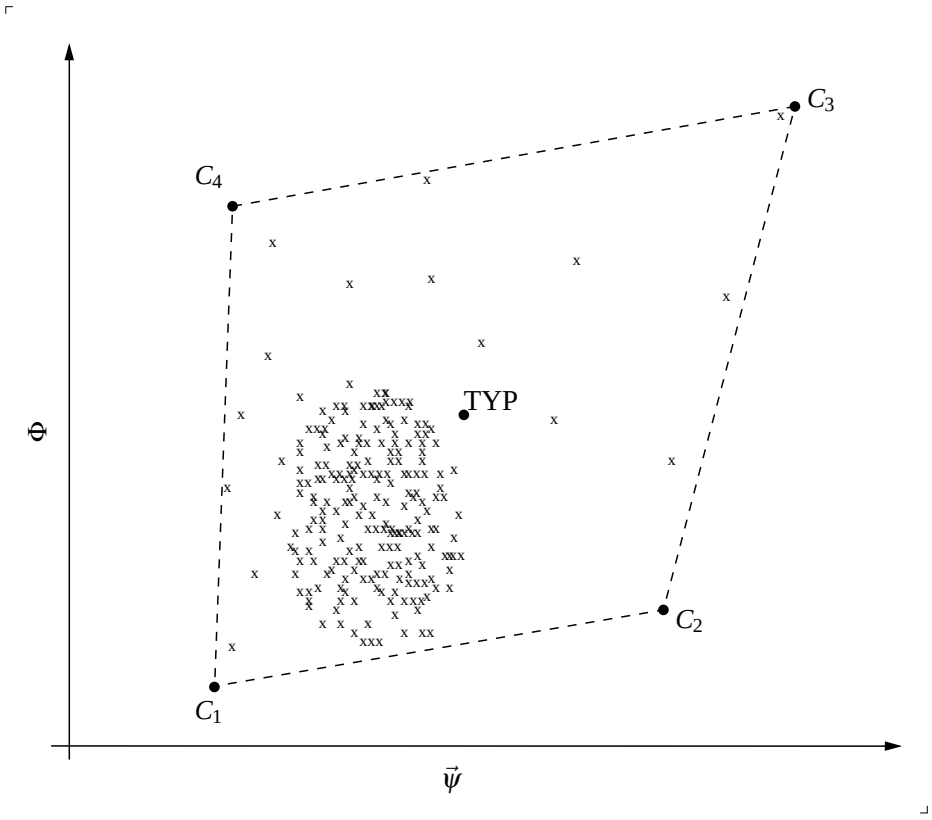


Figure 7: L'analyse des *corners* du circuit au pire cas et la simulation de Monte Carlo: illustration.

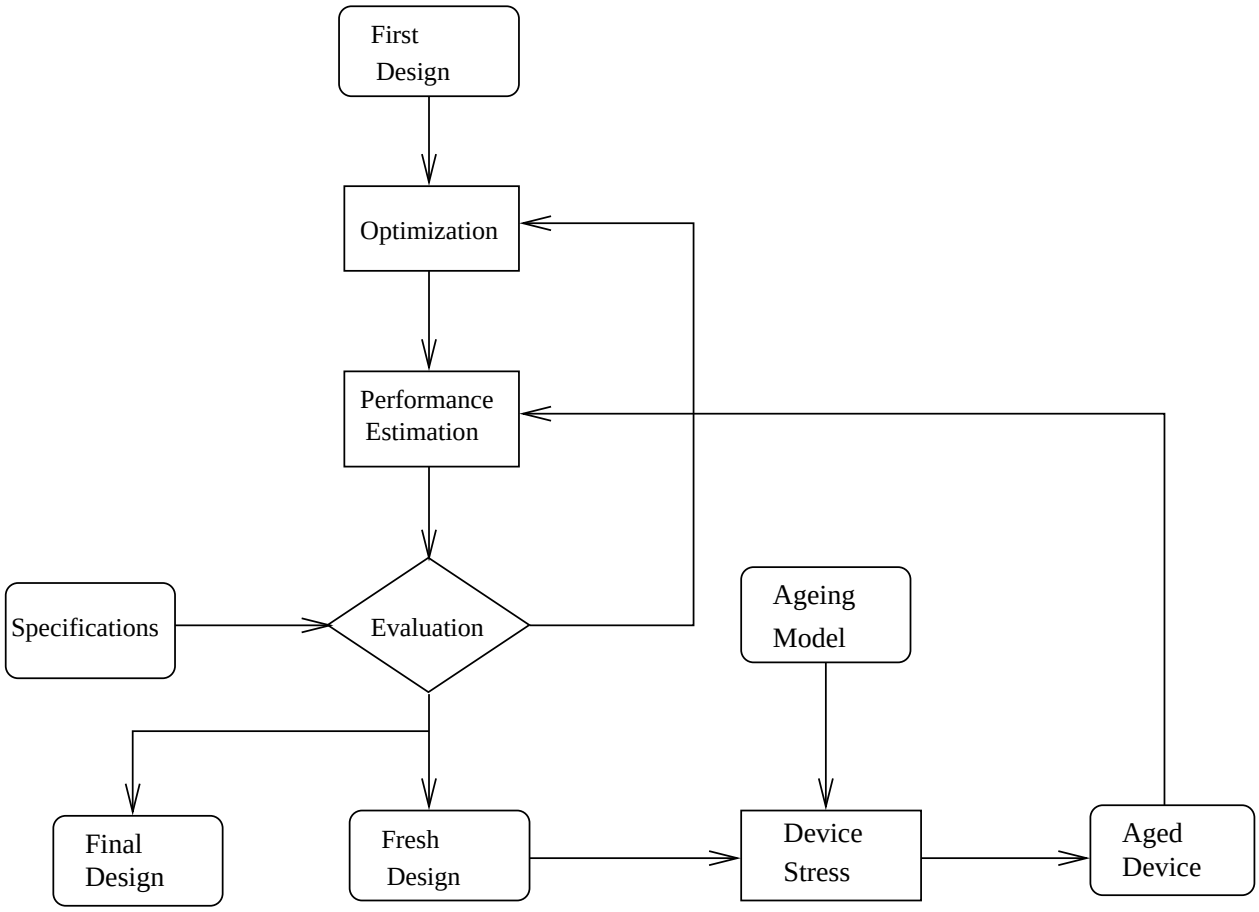


Figure 8: L'état-de-l'art des méthodes de conception avec l'analyse du vieillissement.

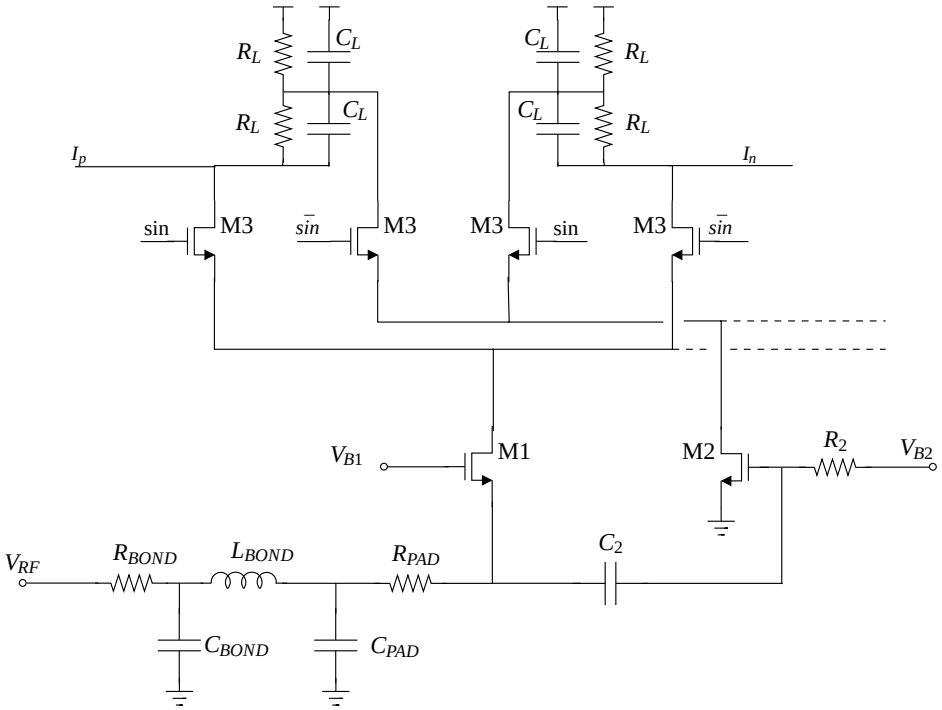
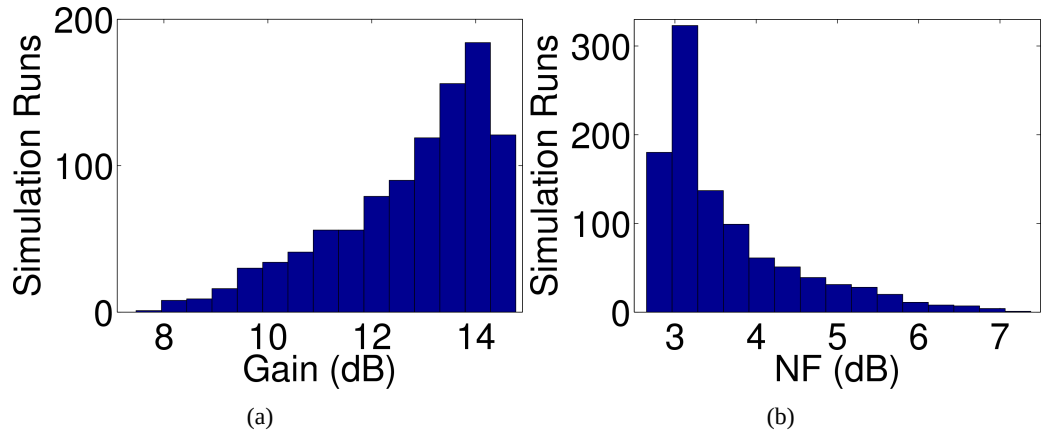
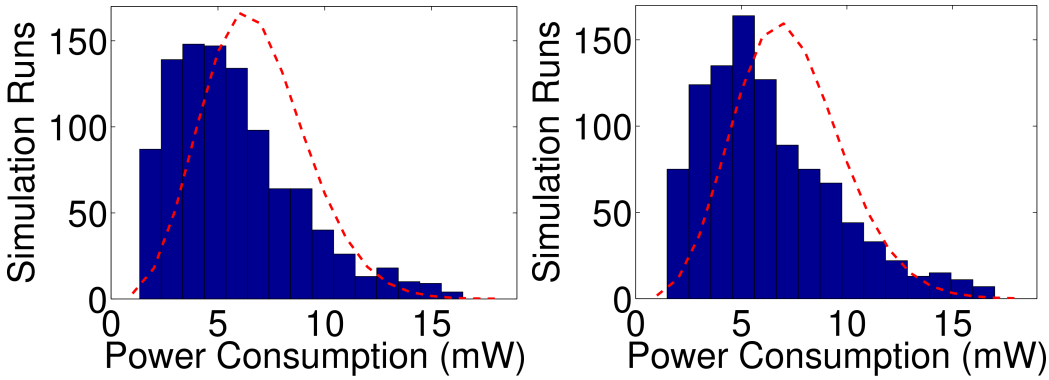


Figure 9: Schéma du BLIXER conçu [2].



(a) (b)

Figure 10: 1000 points de simulation de Monte Carlo du BLIXER typique pour (a) le Gain et (b) le NF_{max} avec 1 GHz de signal d'éntrée.



(a) (b)

Figure 11: 1000 points de simulation de Monte Carlo du BLIXER (a) typique et (b) après 30 ans de vieillissement pour la consommation de puissance en utilisant la bibliothèque de modèle obtenu.

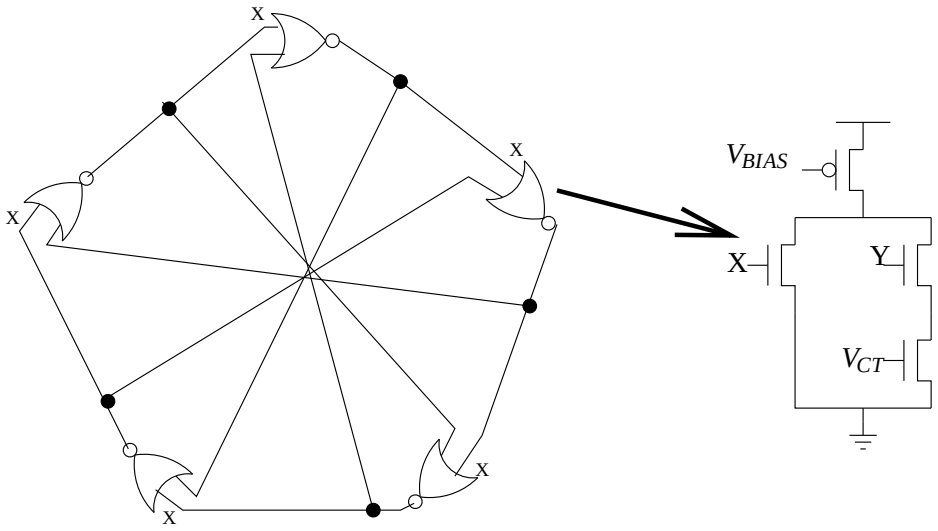
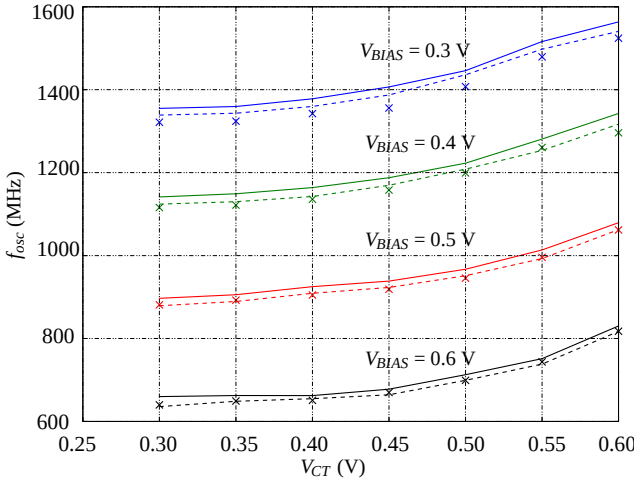
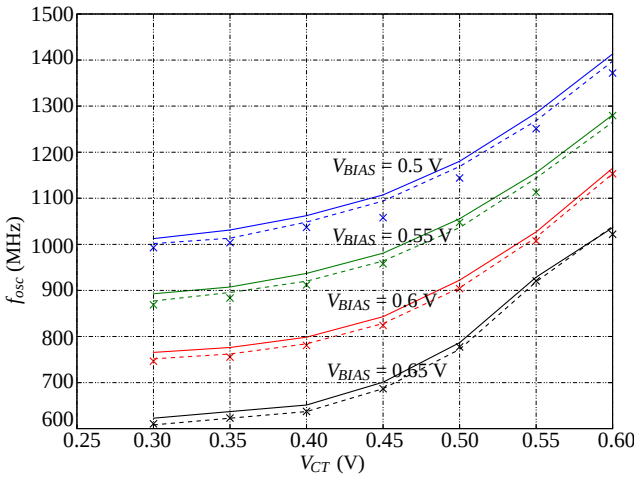


Figure 12: Schéma du DCO conçu [24].

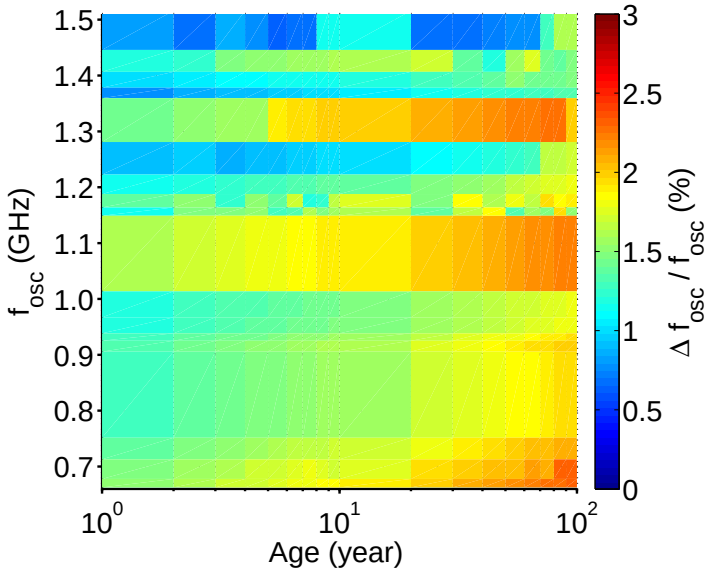


(a)

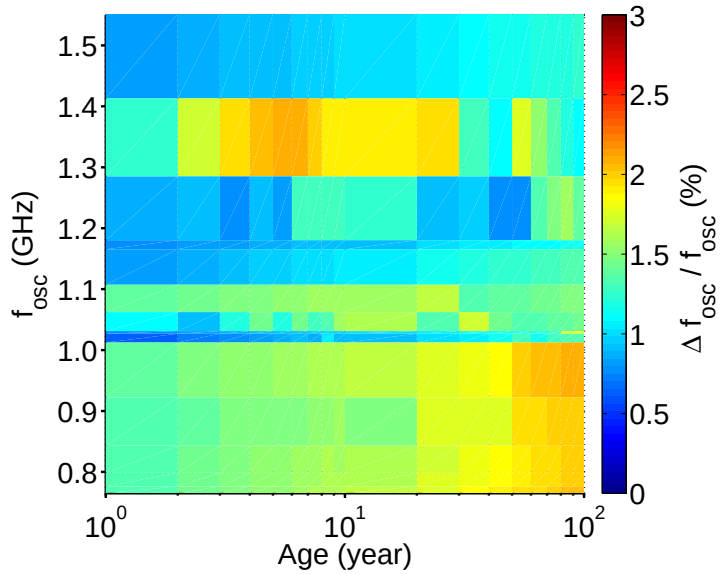


(b)

Figure 13: Fréquence d'oscillation du DCO (f_{OSC}) simulé à 27 ^{o}C pendant 30 ans de dégradation pour (a) le DCO et (b) le DCO fiable. Le résultat avant le stress est représenté par la ligne continue, après le stress par une ligne pointillée et le modèle de vieillissement marquée par x.

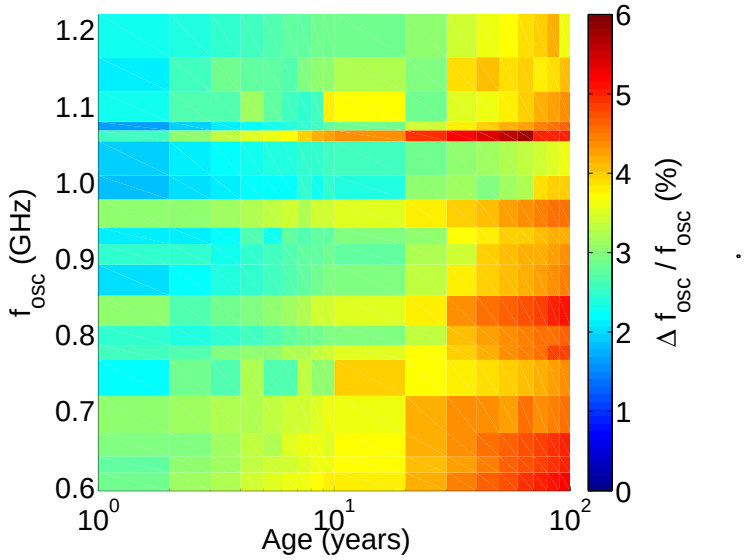


(a)

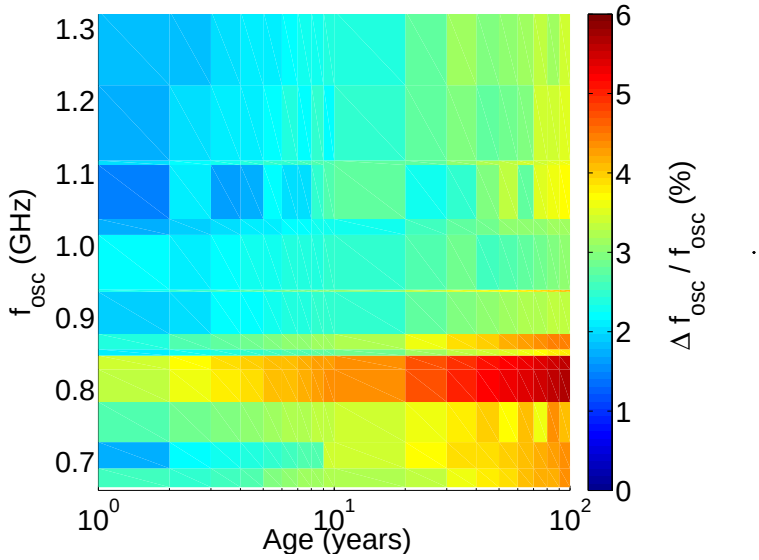


(b)

Figure 14: Analyse du $\Delta f_{osc}/f_{osc}$ à 27 ^{o}C pour (a) DCO et (b) fiable.



(a)



(b)

Figure 15: Analyse du $\Delta f_{osc}/f_{osc}$ à 150 °C pour (a) DCO et (b) DCO fiable.

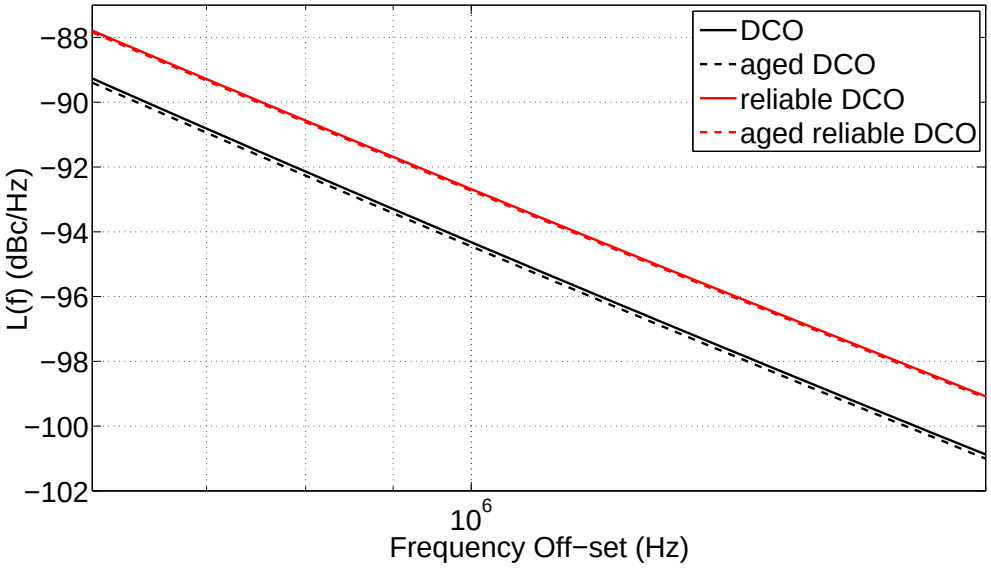
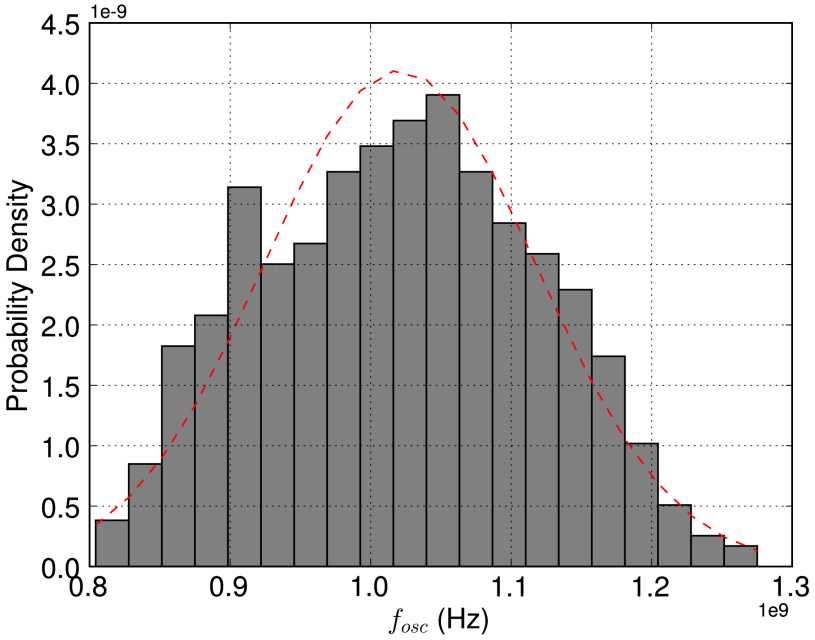
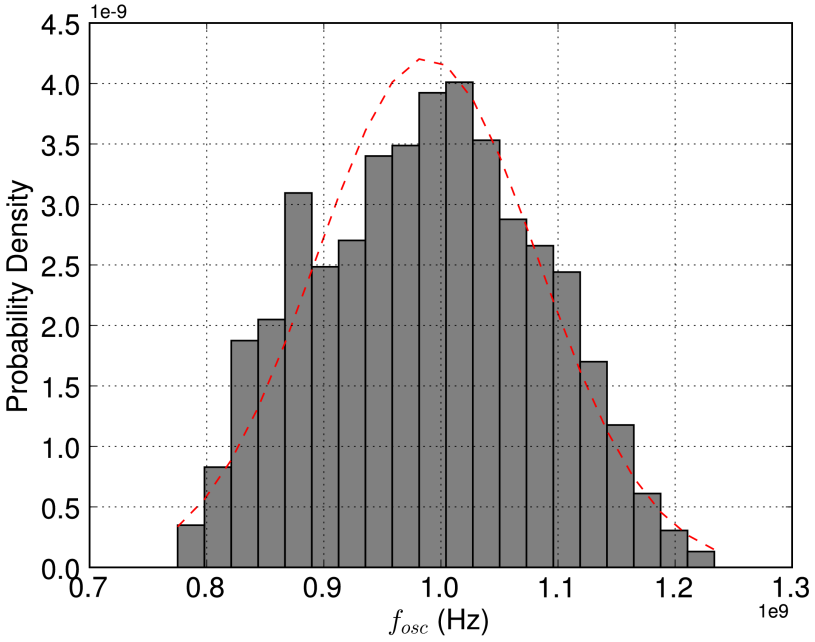


Figure 16: Simulation du bruit de phase du DCO et fiable pour $f_{osc} = 1$ GHz.

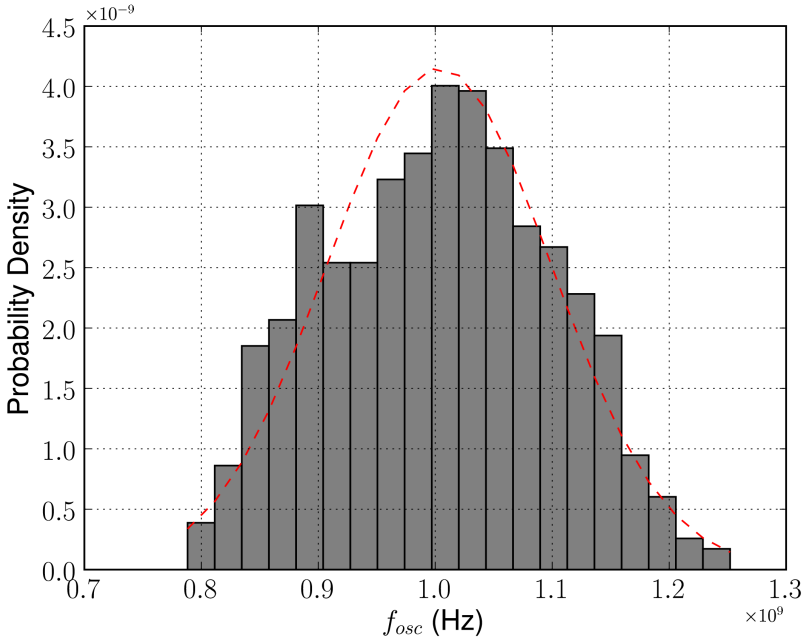


(a)

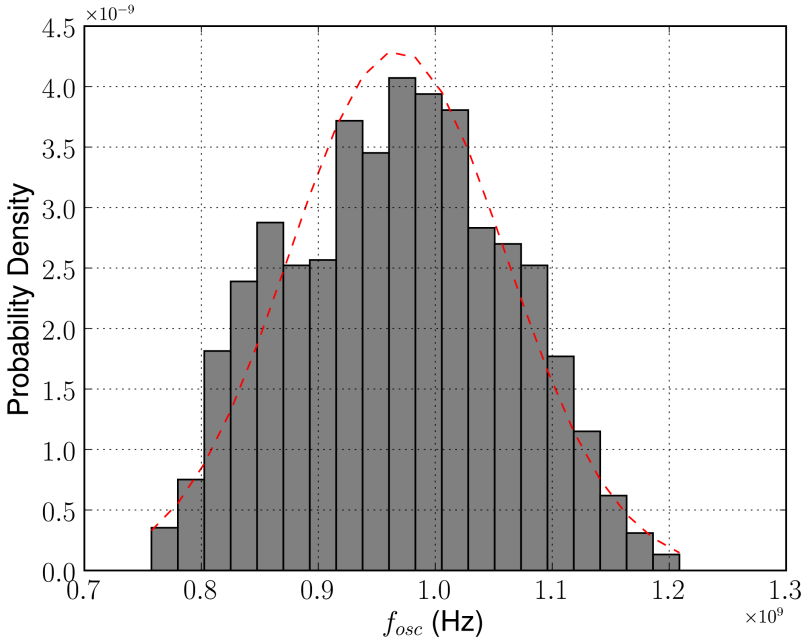


(b)

Figure 17: Variabilité du DCO, simulé stress avec $f_{osc} = 1$ GHz pour 1000 de simulation Monte Carlo. (a) DCO avec μ =1.020 GHz and σ =97.1 MHz ; et (b) DCO fiable avec μ =0.988 GHz and σ =94.7 MHz.



(a)



(b)

Figure 18: Variabilité du DCO, simulé avec le modèle de vieillissement de 10 ans de stress avec $f_{osc} = 1$ GHz pour 1000 de simulation Monte Carlo. (a) DCO avec μ =1.003 GHz et σ =95.9 MHz ; et (b) DCO fiable avec μ =0.967 GHz et σ =92.7 MHz.

PGA Dans cette partie, nous présentons pourquoi le circuit PGA (illustré dans les Figures 19(a) et 19(b)) est naturellement fiable en utilisant la méthode d'analyse et synthèse développée dans les cas précédents. La topologie la plus courante dans les applications multistandard est composée d'un filtre RC actif et d'un amplificateur [4]. L'analyse de la variation des caractéristiques du PGA montre que le choix des compromis dans une conception non-fiabilisée mène à un ensemble de caractéristiques de l'environnement qui dégrade très peu le transistor. Une autre caractéristique notable est la versatilité de la topologie en permettant la reconfiguration des paramètres du circuit.

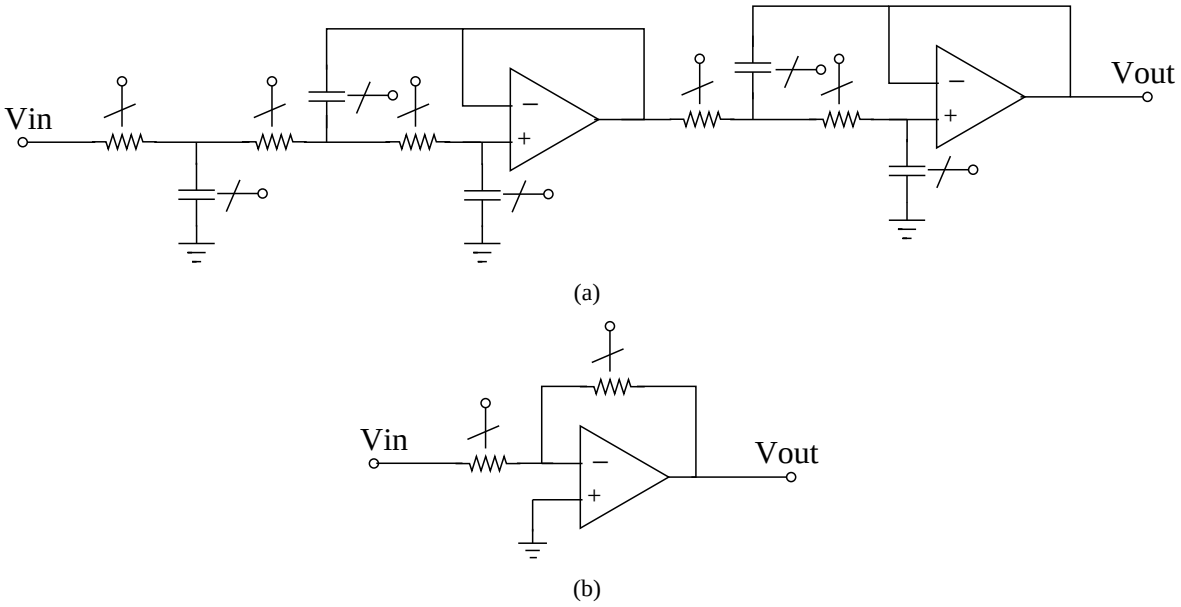


Figure 19: Schéma du PGA multi-bande : (a) filtre de bande base et (b) stage de gain programmable.

En ce qui concerne les contraintes de conception, la méthodologie de conception de circuits fiables est capable de montrer que le PGA est naturellement insensible au vieillissement comme présenté dans la conception du BLIXER. La fiabilité du PGA est ainsi directement contrôlée par la fiabilité des circuits de contrôle. Ces circuits sont pour la pluspart des circuits numériques dont la fiabilité n'est pas traitée dans ce travail.

CONCEPTION D'UN FRONTAL RADIO FIABLE DANS UNE APPROCHE *top-down*

La conception d'un frontal radio fiable vis-à-vis des contraintes imposées par la variabilité et le vieillissement passe d'abord par une modélisation de l'architecture choisie. En connaissant les résultats obtenus dans l'approche *bottom-up*, nous pouvons réaliser une analyse des impacts de la défaillance que les blocs de construction peuvent causer aux caractéristiques spécifiées pour un frontal radio multistandard. Les résultats obtenus durant l'analyse de sensibilité sont suffisants pour proposer des stratégies de conception qui vont réduire les impacts aux caractéristiques du frontal radio. La fiabilisation du frontal radio se concrétise en deux solutions différentes qui sont compatibles avec les caractéristiques spécifiées pour un frontal radio multistandard. Ces deux solutions sont ainsi simulées en complétant le flot de conception du frontal radio. Ces résultats de simulation sont comparés au résultat de l'approche *bottom-up* pour conclure les compromis qui sont pris en compte dans chaque stratégie.

MODÉLISATION DU FRONTAL RADIO La modélisation du frontal radio est faite en deux parties : modèles comportemental et analytique. D'abord, nous avons proposé une modélisation comportementale en VerilogA pour pouvoir simuler l'architecture au niveau système. La plupart des blocs de construction et modèles comportementaux sont disponibles pour la mise en œuvre de l'architecture du frontal

radio [26]. Toutefois, l'ensemble des modèles ne dispose pas de caractéristiques nécessaires à la mise en œuvre de la conception des circuits fiabilisés. Ainsi, nous présentons ces détails complémentaires de mise en œuvre du modèle comportemental pour chaque bloc de construction. Les codes VerilogA sont disponibles en annexe.

Ensuite, nous avons développé un modèle constructif pour l'architecture choisie. Ce modèle analytique nous permettra d'évaluer les caractéristiques du circuit et les variations de ces caractéristiques en fonction des défaillances dans les blocs de construction. Le système des équations du modèle analytique obtenu est aussi suffisamment complet pour permettre l'optimisation des caractéristiques de l'architecture et des variations de ses caractéristiques.

ANALYSE DE DÉFAILLANCE DES BLOCS DE CONSTRUCTION Pour mettre en évidence l'analyse de sensibilité de l'approche top-down, nous avons généré des conditions de défaillance hypothétiques pour les blocs de construction. Ainsi, nous avons évalué l'impact que la défaillance peut causer sur les caractéristiques de l'architecture. Pour cela, nous avons décrit un tableau de tests qui couvrent tous les cas de défaillance des blocs de construction. Les résultats de l'analyse de défaillance des blocs de construction sont illustrés par le Tableau 2 et les Figure 20 et 21

Table 2: Tableau des cas de test de défaillance pour le frontal.

Cas	PGA défaillance	DCO défaillance	BLIXER défaillance
Pas de défaillance	non	non	non
teste 1	non	non	oui
teste 2	non	oui	non
teste 3	non	oui	oui
teste 4	oui	non	non
teste 5	oui	non	oui
teste 6	oui	oui	non
teste 7	oui	oui	oui

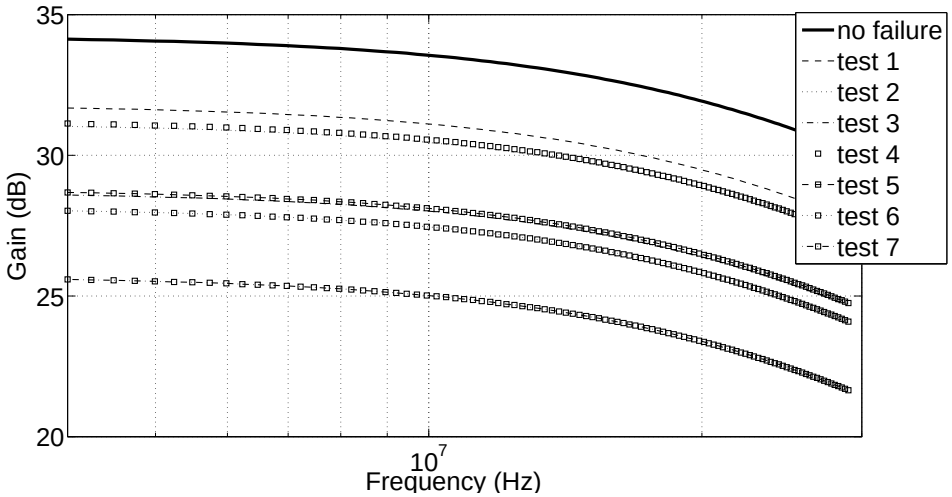


Figure 20: Résultats de simulation du Gain du frontal radio pour les cas de teste d'analyse de défaillance.

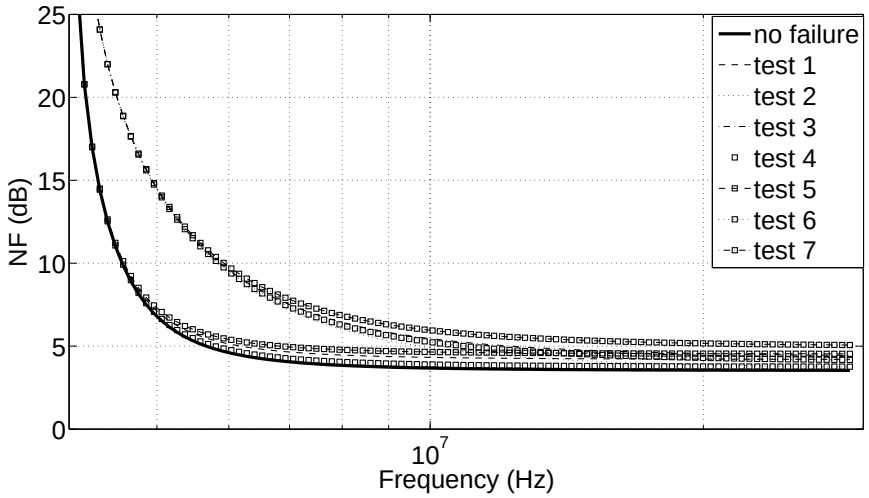


Figure 21: Résultats de simulation du NF du frontal radio pour les cas de teste d'analyse de défaillance.

Table 3: Résultats de simulation du IP3 du frontal radio pour les cas de teste d'analyse de défaillance.

Cas	IP3 (dBm)
Pas de défaillance	1.125
teste 1	2.127
teste 2	1.126
teste 3	2.127
teste 4	1.125
teste 5	2.127
teste 6	1.126
teste 7	2.127

Nous avons trouvé les éléments sensibles et les variations de performance les plus importantes. Nous avons identifié l'amplificateur à gain programmable (PGA) en tant que bloc de construction le plus sensible. La défaillance du circuit PGA est responsable de la dégradation du gain, du bruit, et de la linéarité de l'architecture. Dans les signaux basse fréquence, l'oscillateur à commande numérique (DCO) devient la principale source de dégradation de bruit. Par contre, c'est l'amplificateur de faible bruit (BLIXER) qui est la principale source de dégradation de bruit en haute fréquence [27].

CONCEPTION DU FRONTAL RADIO FIABILISÉ En connaissant les sensibilités des blocs de construction et les impacts des défaillances sur les caractéristiques de l'architecture du frontal radio, nous avons proposé une optimisation des caractéristiques de chaque bloc visant un meilleur compromis entre les caractéristiques, la variabilité et le vieillissement.

D'abord, nous présentons le résultat d'analyse de l'architecture conçue par la méthode classique en connaissant les caractéristiques des blocs de construction. Ensuite, nous discutons un ensemble de stratégies qui peuvent guider une conception fiabilisée du frontal radio. Visant l'équilibre des contraintes dans les compromis de la conception, nous présentons deux solutions possibles dans la prise de la décision pour un frontal radio fiabilisé.

SIMULATION DU FRONTAL RADIO Le résultat obtenu par la méthode classique est comparé aux deux solutions présentées, visant à démontrer les compromis qui ont été pris parmi les caractéristiques des blocs de construction. Nous présentons les résultats simulés pour le gain, le bruit et la linéarité du frontal radio. Ces résultats sont illustrés pas les Tableaux 4, 5 et 6 ; et Figures 22 et 23.

Table 4: Performances estimées pour les blocs des constructions du frontal radio.

BLIXER	PGA	DCO	Architecture
G = 14 dB	G = 20 dB	V_{LO} = 0.25 V	G = 34.1 dB
NF = 3.5 dB	NF = 10 dB	$L(1\ MHz)$ =	NF=
		- 120 dBc/Hz	6.87 dB@5 MHz
$f_{RF} = 5.0$ GHz		$f_{LO} = 5.004$ GHz	$f_{IF} = 4$ MHz
IP3 =1.1 dBm	IP3 =10.0 dBm		IP3 =1.12 dBm
R_{in} = 53 Ω			S_{11} = -15.0 dB

Table 5: Performances estimées pour les blocs des constructions du frontal radio fiabilisé avec la stratégie 1.

BLIXER	PGA	DCO	Architecture
G = 13.9 dB	G = 16.8 dB	V_{LO} = 0.24 V	G = 30.6 dB
NF = 3.51 dB	NF = 10 dB	$L(1\ MHz)$ =	NF=
		- 120 dBc/Hz	6.65 dB@5 MHz
$f_{RF} = 5.0$ GHz		$f_{LO} = 5.004$ GHz	$f_{IF} = 4$ MHz
IP3 =1.09 dBm	IP3 =10.05 dBm		IP3 =1.09 dBm
R_{in} = 61 Ω			S_{11} = -10.0 dB

Table 6: Performances estimées pour les blocs des constructions du frontal radio fiabilisé avec la stratégie 2.

BLIXER	PGA	DCO	Architecture
G = 13.9 dB	G = 16.8 dB	V_{LO} = 0.25 V	G = 30.8 dB
NF = 3.84 dB	NF = 10 dB	$L(1\ MHz)$ =	NF=
		- 119.5 dBc/Hz	7.26 dB@5 MHz
$f_{RF} = 5.0$ GHz		$f_{LO} = 5.004$ GHz	$f_{IF} = 4$ MHz
IP3 =0.16 dBm	IP3 =10.05 dBm		IP3 =0.16 dBm
R_{in} = 53 Ω			S_{11} = -14.8 dB

Avec les résultats de simulation du frontal radio, nous avons démontré la validité des discussions développées durant l'analyse de sensibilité. En plus, nous présentons de manière très claire les compromis présents possibles dans la prise de décision pour un frontal radio fiabilisé. Ainsi, les deux solutions de conception du frontal radio fiabilisé vont constituer des points d'équilibre entre l'optimum de la caractéristique spécifiée et les variations des caractéristiques dues à la variabilité et au vieillissement.

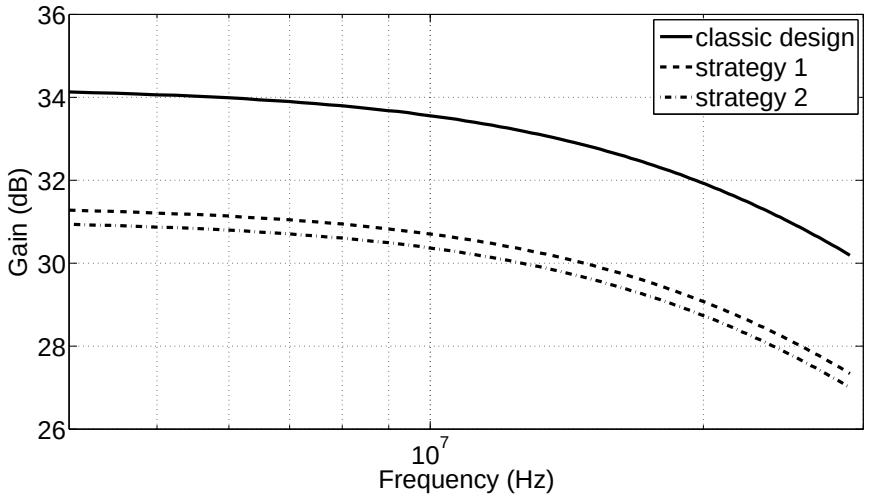


Figure 22: Simulation du gain du frontal radio, et les fiabilisées par les stratégies 1 et 2.

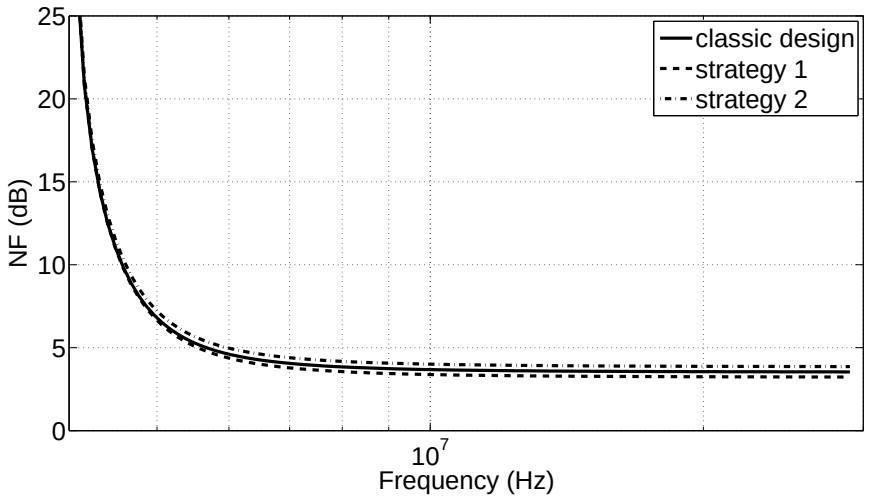


Figure 23: Simulation du NF du frontal radio, et les fiabilisées par les stratégies 1 et 2.

FLOT DE CONCEPTION DES CIRCUITS FIABLES

A partir des expériences obtenues dans la conception du frontal radio, nous avons rassemblé les étapes importantes du flot de conception des circuits fiables. Ainsi, nous proposons un nouveau flot de conception pour les circuits AMS/RF qui prend en compte les compromis de la variabilité et du vieillissement. Ensuite, nous pouvons démontrer comment ces compromis peuvent influencer les stratégies de conception.

Le flot classique de conception de circuits a besoin des spécifications minimum et maximum des performances. Souvent, les informations sur la technologie d'intégration et les blocs de construction sont disponibles. L'ensemble de ces informations vont constituer ce qu'on appelle l'espace de conception, c'est-à-dire la description des caractéristiques en fonction des choix de conception (par exemple dimensionnement et polarisation).

Prenant en compte les spécifications de performance et la caractérisation de l'espace de conception, une méthode de conception classique permet d'obtenir l'optimum pour les caractéristiques (ψ) des dis-

positifs de bas niveau à partir des caractéristiques (Φ) des dispositifs de haut niveau. Si l'optimisation prend déjà en compte la variabilité (comme en [18]) et que le stress de vieillissement n'est pas assez élevé en comparaison aux besoins de fiabilité, il est possible que ce circuit optimal satisfasse à l'évaluation de défaillance. Donc, nous avons un circuit optimal et fiable comme il est apparu dans les exemples de conception des circuits BLIXER et PGA. Pourtant, si la variation des caractéristiques du dispositif n'est pas nulle, cela signifie qu'il a souffert soit de variabilité, soit de vieillissement.

C'est clair qu'il est important de refaire l'étape d'optimisation, mais il faut apporter des éléments qui guideront l'optimiseur vers une solution aussi fiable. Donc, nous avons proposé l'analyse de sensibilité des dispositifs comme point de départ de la fiabilisation du circuit. Cette analyse est appliquée au même modèle utilisé dans l'étape d'optimisation et va conduire à une mesure de faiblesse des dispositifs. Avec l'expérience des concepteurs, ces chiffres vont constituer des poids pour le partage de la variation estimée dans l'étape précédente entre les variations autorisées aux caractéristiques des dispositifs de bas niveau.

Le partage des variations est un problème classique de prise de décision qui peut converger à un optimum ou présenter des points d'équilibre. Ainsi, nous discutons les stratégies possibles qui guideront la prise de décision. Dans la conception du DCO, nous avons présenté une stratégie assez simple guidée par les informations sur la variation de l'espace de conception. Cela signifie spécifier un nouveau jeu des caractéristiques pour atteindre les spécifications du circuit pour un certain rendement et fiabilité. Pourtant, ces stratégies peuvent être plus complexes et aboutir à un certain compromis pour relâcher la conception d'un bloc et rendre stricte la conception d'autres blocs, comme dans l'exemple présenté dans l'approche top-down.

Finalement, l'optimisation sera bouclée avec un nouveau jeu des caractéristiques à atteindre et un espace de conception réduit. Il est bien possible que l'optimisation ne soit pas capable d'obtenir le nouvel optimum qui va aboutir à la faisabilité du circuit avec ses spécifications de performance, rendement et fiabilité. Par contre, s'il existe un optimum il pourra être fiable ou pas, ce qui va aboutir au besoin de fiabilisation du circuit. La fiabilisation du circuit, que nous avons élaborée, constitue un nouveau flot de conception qui prend en compte la variabilité et le vieillissement. Ce flot de conception des circuits fiables est résumé dans la Figure 24.

Pour mettre en évidence la faisabilité des circuits dans les compromis existants parmi leurs spécifications de performance, rendement et fiabilité, nous démontrons ces compromis pour un oscillateur en anneau avec 13 portes inverseurs. L'oscillateur en anneau est un circuit souvent utilisé pour évaluer une technologie d'intégration. Faisant varier la tension d'alimentation, nous pouvons analyser le compromis entre consommation de puissance de chaque porte et sa vitesse par le délai de la porte.

Pourtant ces deux spécifications vont certainement changer avec la variabilité et le vieillissement des transistors. Pour cela, nous allons simuler ce circuit conçu en CMOS 65 nm sous un stress de 27^{o}C de température, jusqu'à une tension d'alimentation 10 % plus grande que la valeur spécifiée dans la technologie d'intégration et 30 ans de vieillissement. Nous mettons en évidence une variabilité proche de 30 % et un vieillissement proche de 1 % pour le point optimum du compromis entre consommation de puissance et vitesse. Donc, nous mettrons en évidence que le vieillissement sera souvent négligeable par rapport à la variabilité dans l'optimisation des compromis des performances en utilisant la technologie CMOS 65 nm sous un stress comparable aux conditions d'environnement courantes.

CONCLUSION

Ce travail de thèse a proposé un nouveau flot de conception des circuits fiables en s'appuyant sur la conception d'un frontal radio. Ainsi, nous arriverons à notre but principal qui était d'améliorer la conception de circuits du frontal RF. Durant ce travail, nous sommes à la recherche de nouveaux compromis imposés par la variabilité du transistor et dégradation par vieillissement. D'après ce nouveau compromis, nous avons proposé des stratégies de partage de la variation des caractéristiques des circuits entre les caractéristiques des blocs de construction.

Dans la recherche de l'état-de-l'art, nous avons étudié la physique des phénomènes de vieillissement

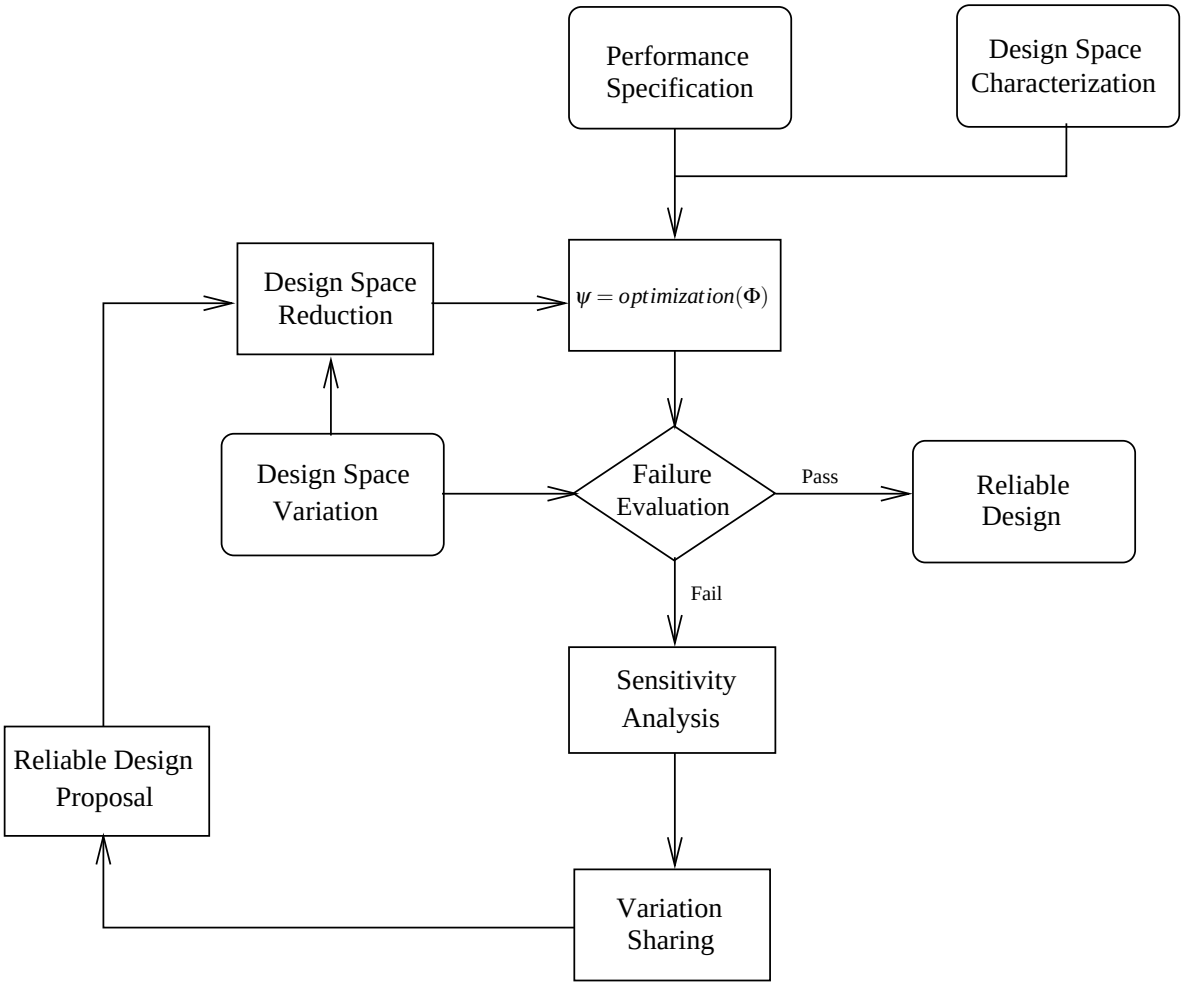


Figure 24: Flot de conception des circuits fiables: illustration des étapes de conception.

et des conditions de conception du circuit qui permettent d'éviter le vieillissement et la variabilité des transistors. Ensuite, nous avons étudié les sources de dégradation (vieillissement et variabilité) et ses tendances dans les technologies à l'échelle nanométrique. Enfin, nous avons présenté les méthodologies de conception classiques, comparant les besoins imposés par la variabilité et la fiabilité des composants.

Nous avons mis en œuvre des circuits fiables pour le cas d'étude du frontal radio dans une approche *bottom-up*. Par ailleurs, nous avons mis en œuvre l'architecture fiable pour le cas d'étude du frontal radio dans une approche *top-down*. Ainsi, nous avons pu lier les étapes de la conception *top-down* et *bottom-up* dans une méthode générale qui est la proposition d'un nouveau flot de conception des circuits fiables.

La conception des circuits fiables met en évidence le nouveau compromis entre les performances nominales désirées, les attentes de rendement et de temps de vie pour le circuit. Par la démonstration des compromis imposés par le vieillissement et la variabilité des composants en CMOS 65 nm, nous sommes capables de prédire les tendances dans les technologies à venir et mettre en évidence le besoin d'un flot de conception des circuits AMS/RF qui prend en compte les dégradations des performances pour le vieillissement et la variabilité.

ABSTRACT

In this work, we have been motivated to innovate in RF front-end design. New analysis and synthesis methodologies have been proposed including the variability and the ageing degradation in the center of the design trade-off. Moreover, the variability and the ageing degradation criteria have motivated us to propose changes in the classical design methodology with aim of a variability-aware and ageing-aware synthesis. Thus, the main objective of this work has been to improve the design of AMS/RF front-end circuits based on the investigation of a new trade-off imposed by transistor variability and ageing degradation. Aiming the proposition of both agents of characteristics variation as design criteria, we have designed a reliable RF front-end for a multi-standard application.

This state-of-the-art research has described, as best as we know, the sources and the trends of the device-performance variations. Variations will be separated as *static* or at time zero, and *dynamic* or time-dependent variation. The *static* variations will be represented by integration of process variability and devices mismatch. The *dynamic* variations will be represented by probabilistic events and device ageing. The knowledge of the physical mechanisms of variability and ageing has been essential to propose an improved design methodology aiming at more reliable devices. Furthermore, we have described, as best as we know, the state-of-the-art of design methodologies. The challenges and solutions for AMS/RF design in advanced CMOS technology have been presented. The classical design methodologies have been divided in three parts: architectural behavioral-model validation, electrical schematic implementation, layout synthesis. Most of variability-aware design methodologies have presented worst-corner and Monte Carlo simulation for design verification and performance estimation. However, some few works have been describing the variability-aware design methodologies using behavioral modeling and so improving the design optimization. Moreover, the reliability-aware design methodologies have been mostly concentrated on reliability analysis around a nominal corner. Few works have presented the combination of variability and ageing phenomena under a reliability analysis. Hence, we have identified an opportunity to innovate by proposing the reliability estimation in early design stages.

We have innovated by proposing AMS/RF circuit reliability improvements during the design of the multi-standard RF front-end using a bottom-up approach. First, the reliable-BLIXER design with a failure evaluation has been proposed in [23]. Next, the validation of a reliable-circuit synthesis method using a DCO design has been conducted. By designing a classical- and a reliable-DCO, we have published such analysis and discussions in [27] and [28]. Then, the PGA and the required elements to design a reliable-PGA have been analyzed. Furthermore, we have innovated by proposing architecture reliability improvements during the design of the RF front-end using a top-down approach. In this case, we have discussed the design of a reliable architecture for RF front-end and the variation sharing strategies to avoid an overdesign. We have published such analysis and discussions in [27]. Therefore, we have innovated linking top-down and bottom-up approaches in a general method which has been the proposition of a new AMS/RF design flow increasing the circuit reliability. The design of reliable circuits has highlighted a new trade-off among typical performance specification, the yield requirements and the circuit lifetime.

Therefore, our major objective has been successfully achieved; while improving the design of AMS/RF front-end circuits based on the investigation of new trade-offs imposed by transistor variability and ageing. Finally, we could point some research perspectives in: new analysis tools, new design models, and new synthesis methods; linking variability and ageing.

LIST OF SYMBOLS

AMS	Analog Mixed Signal
CMOS	Complementary Metal Oxide Semiconductor
CAD	Computer-Aided Design
DE	Differential Evolution
DoE	Design of Experiments
$\Delta\Phi$	Variation of Φ (defined by $\Phi_{worst} = \Phi_{typ} \pm \Delta\Phi$)
D2D	Die-to-die
EA	Evolutionary Algorithms
EDA	Electronic Design Automation
EM	Electromigration
GP	Geometric Program
HBD	Hard Oxide Breakdown
HCI	Hot Carrier Injection
IC	Integrated Circuit
IP	Intellectual Property
IP3	Third-order Input Intercept Point
LER	Line Edge Roughness
LHS	Latin Hypercube Sampling
MC	Monte Carlo
MOS	Metal Oxide Semiconductor
MTTF	Median Time to Failure
NBTI	Negative Bias Temperature Instability
NMOS	n type MOS transistor
OTA	Operational Transconductance Amplifier
Φ	a general circuit or system performance
ψ	a general circuit or system parameter
PA	Power Amplifier
PDF	Probability Density Function
PDK	Process Design Kit
PMOS	p type MOS transistor
QMC	Quasi Monte Carlo
RDF	Random Dopant Fluctuations
RF	Radio Frequency
RSM	Response Surface Model
SA	Simulated Annealing
SBD	Soft Oxide Breakdown
SiP	Systems in a Package
SM	Stress Migration
SoC	Systems on Chip
SOI	Silicon-On-Insulator

TDDB	Time Dependent Dielectric Breakdown
V_{DS}	Drain-to-source voltage
V_{GD}	Gate-to-drain voltage
V_{GS}	Gate-to-source voltage
V_{ov}	Overdrive voltage $V_{ov} = V_{GS} - V_{th}$
V_{SB}	Source-to-bulk voltage
V_t	Thermal voltage
V_{th}	Threshold voltage
VLSI	Very Large Scale Integration
WID	Within-die

CONTENTS

CHAPTER 1

INTRODUCTION

Sub-nanometer CMOS technology has emerged new yield and reliability challenges [1]. The next generation of analog-mixed signal (AMS) and radio frequency (RF) circuits may be touched by an increase in failure rate during all circuit operation-time. Classical design methodologies often look for the basic design criteria: die area, power consumption and speed; in terms of exploiting the technology limits. The optimum is the design point where we have the specified performance.

The increasing transistor variability has proved to be big enough to find a large number of chip samples with performance far away from the specification. This phenomenon is known as yield reduction. The yield can be defined as the ratio of chip samples which meet the design specifications and of all chip samples in a context of a complete production process. However, yield concept cannot measure the number of chip samples which still meet the design specifications in a context of continuous use under a known environment condition.

The transistor ageing is the agent of circuit characteristic changes under stressful environment condition during a period of time. The specified period of time is including a time-varying concept into circuit performance quality. If such a period of time is zero and so it is the complete production process moment; we measure the circuit yield which should be better than specified. If the circuit performance quality drops out of the specification, the time when it occurs is defined as the circuit lifetime. Combining stressful environment condition and the circuit lifetime, the reliability is defined as the ability of a circuit to conform to its specifications over a specified period of time and under specified conditions.

In order to evaluate the reliability of a circuit, we assume that a circuit is composed of n statistical identical and independent parts that were put into operation at time $t = 0$. The empirical reliability of a circuit can be defined according to

$$\hat{R}(t) = \frac{u(t)}{n}, \tag{1.1}$$

where $u(t)$ represents how many parts did not yet fail at time t. It can be noted that the behavior of $u(t)$ is a continuous decreasing step function. A direct application of the law of large numbers ($n \to \infty$) yields that $\hat{R}(t)$ converges to the reliability function $R(t)$ [29].

The $\hat{\lambda}(t)$ is the empirical failure rate and it is given by

$$\begin{aligned} \hat{\lambda}(t) &= \frac{u(t) - u(t+\delta_t)}{u(t)\delta_t} \\ \hat{\lambda}(t) &= \frac{\hat{R}(t) - \hat{R}(t+\delta_t)}{\delta_t \hat{R}(t)}, \end{aligned} \tag{1.2}$$

converges to the failure rate expressed by

$$\lambda(t) = \frac{-\frac{dR(t)}{dt}}{R(t)} \tag{1.3}$$

for $n \to \infty$, $\delta_t \to 0$ and $n\delta_t \to 0$ [29]. Considering that at time $t = 0$, the circuit executes its functions

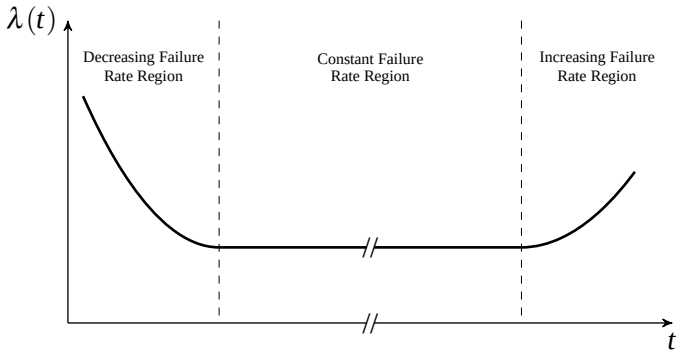


Figure 1.1: Bath Tub Curve representing the typical shape of the failure rate of a circuit composed of n statistical independent parts

perfectly, that means $R(0) = 1$. In this case, the reliability function can be defined as

$$R(t) = e^{-\int_0^t \lambda(x)dx}. \tag{1.4}$$

Analyzing the equation (1.4), it can be seen that the reliability function depends of the behavior of $\lambda(t)$. In fact, $\lambda(t)$ has a typical shape represented in Fig. 1.1. This curve is denominated *bathtub curve* due to its shape and it is described by three parts:

- *Decreasing Failure Rate* - the failures that occurs when the circuit is first introduced as a result of momentary weakness in materials or in the item's production process. At this point, the design should improve the circuit yield by the simulation of the integration process parameters variation.
- *Constant Failure Rate* - the period when $\lambda(t)$ can be approximated by a constant. It corresponds to the useful life of the circuit and it is reduced in advanced technologies when the sizes are shrunk and the ageing degradations are increased.
- *Increasing Failure Rate* - the end of the circuit's operation, the circuit lifetime is achieved due to wearout and ageing degradation. In sub-nanometer integrated circuits (IC), it begins earlier and is strongly dependent on the circuit and its operation conditions.

1.1 MOTIVATION

Since the advent of mobile phones, radio frequency (RF) front-end architectures were requiring increased technology innovations in a shortened time-to-market. At the beginning, the radio circuits are concerned in transmission of narrow signal bandwidth, a single service using a simple RF standard, and sometimes a double-talk communication was not required. In just few years, the humanity demands have changed a lot. Thus, the RF front-end architectures have been improved in the direction of wide-bandwidth, multi-standard, opportunistic radio and cognitive radio.

These improvements have been the base of new mobile generations. The called first-generation was an analog-based RF architecture for voice transmission only. The second-generation push forward the data rates to tenths of kilobits per second and it has innovated with the text message functionality. The third-generation has transformed the radio in a wireless multimedia center and so requiring the coexistence of different standards (like Bluetooth/UMTS/WCDMA) in a single or a multi-path front-end having data rates of some megabits per second. The fourth-generation is just beginning with the LTE standard, but the radio users expect GPS service, digital television, and WLAN Internet connection in a single wireless device which could also be able to do voice and text messages transmission.

Beside the RF front-end architecture research, integrated circuit (IC) technologies have been following the famous Moore's Law. IC technologies have achieved shrunk dimensions, lower power consumption, and increased limit-frequency. Area, power consumption and speed have become the basic trade-off for a circuit design mostly digital. Analog circuits have added gain, noise and linearity performance, forming a complex design trade-off challenge.

The IC technology node evolution has achieved the 100 nm barrier, moving inside of the molecular dimensions. Thus, the old drawback of circuit variability became the major agent of performance variation and the circuit yield a new trade-off challenge. With the advent of the important amount of variability the optimum-point design has become a region of one, two or more σ depending of the desired yield. The easiest way is to design the maximum and minimum circuit characteristics as the worst-case of such characteristics. The cost is imposing margins of circuit failure leading to an overdesigned circuit. The argument in opposition of a worst-case design is that such design does not take the statistical characteristic into account, but it is just a modified deterministic design.

CMOS ageing phenomena are known since the 70's decade, while circuits shall be stressed over its normal operation conditions to present ageing degradations. This subject has disappeared of the state-of-the-art for some decades to be back to the focus of research in the advent of the under 100 nm IC technologies. Why has ageing become so important in these dimensions? The ageing trends are very worrying because threshold voltage and voltage source did not keep pace during the dimension shrinking. Therefore, the very improbable operation condition which could stress the circuit is now the usual environment condition.

Most of ageing phenomena research is focused on ageing analysis, developing new estimation tools of circuit characteristics after some lifetime. Over 100 nm IC technologies, the circuit lifetime was near the infinity because the IC industry was sure that such a circuit would be replaced before its lifetime is achieved. Under 100 nm IC technologies, the IC industry has no more this guarantee because the circuit lifetime has become smaller than the before-expected time to replace the circuits. Therefore, ageing may change the statistical characteristics of the circuits during its lifetime. The easy and costly solution is using bigger margins and redundant circuits.

However, in spite of the advantages of the fourth-generation of mobile, the implementation of the future radios will require more and more shrunk IC technology dimensions, complex AMS circuits aiming an ADC just after the antenna, and multi-mode reconfigurable RF circuits. The future of telecommunications faces a number of design challenges which may be summarized as [5]:

- *Flexibility* in terms of wireless standards leads to multi-standard radios. Aiming the operation over a variety of different specifications, multi-standard radios need to be implemented with reconfigurable building blocks that can adapt to each standard specification and different signal conditions.

- *Integration* in terms of more and more functionalities in a single chip is leading to sub-nanometer IC technologies. Thus, the power supply is reduced and the limit-frequency is near the required speed. Moreover, the system performance may change during time pushed by the increasing variability and ageing degradation.

- *Optimization* in terms of single optimum is no longer the case. A multi-standard radio has multiples optima at least one for each standard specification. Furthermore, the integration challenges have transformed an optimum point in a region which may change during the circuit lifetime.

In this scenario, we are motivated to innovate in RF front-end design proposing new analysis and synthesis methodologies including the variability and the ageing degradation in the center of the design trade-off. Moreover, the variability and the ageing degradation criteria motivate us to propose changes in the classical design methodology aiming a variability-aware and ageing-aware synthesis.

1.2 Study Case Design

The multi-standard wireless applications require the most advanced IC technology [5], and thus we will study a design of an RF front-end in CMOS 65 nm. The wide-band specification and multiple frequency carriers will require a reconfigurable architecture. Although the many imposed challenges, one of the most advantageous architectures for such application is the direct conversion architecture. Moreover, such an architecture can easily be converted to the digital low-IF architecture if required. Hence, the RF front-end shown in Figure 1.2 was chosen [5].

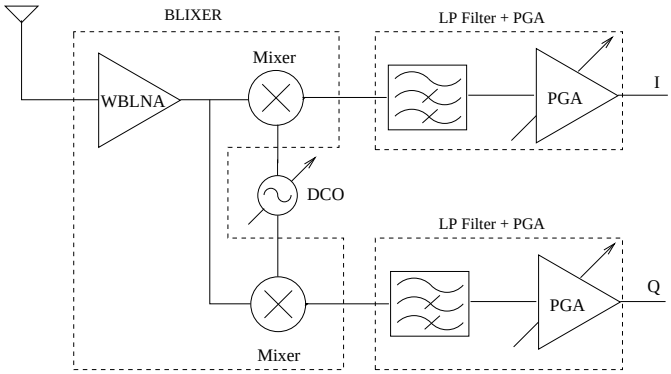


Figure 1.2: RF front-end architecture for multi-standard wireless applications: illustration [5].

The RF front-end has three main blocks: the BLIXER [2], aggregating a balun, a wide-band low noise amplifier, and an I-Q mixer; the digital controlled oscillator (DCO) [3], and the programmable gain amplifier (PGA) together with the low-pass filter [4]. The common functions of the direct conversion RF front-end were aggregated according to the available transistor level schematics.

The BLIXER, proposed in [2], avoids the area-consuming on-chip inductors delivering the signal at base-band by the I-Q mixer. Therefore, the area cost is lower with some trade-offs in noise performance and linearity. The receiver has often a single-ended RF-input which requires an external broadband balun and its accompanying losses. However, the BLIXER has two amplifier paths with signal phase opposition and common input noise which can be ideally canceled with the difference of the balanced outputs.

The DCO, proposed in [3], uses digital ring-gates interpolation avoiding inductors but guaranteeing a low phase noise performance. The trade-offs in the DCO is among limited resolution, design complexity, and power consumption. The advantage of the DCO is exploiting the digital domain with an all-digital phase locked-loops (ADPLL) and a time to digital converter (TDC), increasing the DCO reconfigurable capability. This characteristic is very probable in ageing stress environment, and it could mitigate the performance variations.

The PGA with the low-pass filter acts like a natural anti-aliasing filter and signal swing control for the following ADC. That is why a multi-mode, multi-band active-RC filter and tuning circuits are interesting for multi-standard applications [4]. The trade-offs are power and area consumption which allows the desired versatility and good linearity. The PGA versatility could also mitigate the performance variations.

The key performances of the RF front-end were defined in a context of multi-standard application requirements [5]. The RF front-end architecture specifications are summarized in Table 1.1.

1.3 Objective

The main objective of this work is to improve the design of AMS/RF front-end circuits based on the investigation of new trade-offs imposed by transistor variability and ageing degradation. Aiming at

Table 1.1: RF front-end architecture specifications of a multi-standard application [5].

Operational Frequency	1 GHz - 6 GHz
Bandwidth (max)	20 MHz
Gain	> 30 dB
NF	3.5 dB @ 1 GHz - 6 dB @ 6 GHz
IP3	> 0 dBm
S11	< -10 dB

the proposition of both agents of characteristics variation as design criteria, we will design a reliable RF front-end. Thus, we can enumerate our sub-objectives which will complete the main objective, as

1. Study the ageing phenomena physics and the circuit design conditions to avoid ageing degradation
2. Study variability and ageing degradation sources and trends in sub-nanometer IC technologies
3. Study the classical design methodology comparing to the variability-aware and reliability-aware needs
4. Propose the reliable-circuit design of the design study case in a bottom-up approach
5. Propose the reliable-architecture design of the design study case in a top-down approach
6. Propose the generalization of the synthesis method linking bottom-up and top-down approaches
7. Demonstrate the transistor variability and ageing degradation trade-off in CMOS 65 nm

1.4 ORGANIZATION

During this chapter, we presented the basic concepts of circuit reliability and the reliability definitions in CMOS technology. Introducing this work, we are motivated to innovate in RF front-end design proposing new analysis and synthesis methodologies, including the variability and the ageing degradation in the center of the design trade-offs. The chosen study case design is an RF front-end architecture. Hence, our main objective is to improve the design of AMS/RF front-end circuits based on the investigation of new trade-offs imposed by transistor variability and ageing degradation.

Chapter 2 describes, as best as we know, the sources and the trends of the device-performance variations. Variations will be separated as *static* or at time zero, and *dynamic* or time-dependent variation. The *static* variations will be represented by integration process variability and devices mismatch. The *dynamic* variations will be represented by probabilistic events and device ageing. At Chapter 2, we present the knowledge of the physical mechanisms of variability and ageing which is essential to propose improved design methodologies aiming at more reliable devices.

Chapter 3 describes, as best as we know, the state-of-the-art of design methodologies, the possible solutions and the not solved challenges for AMS/RF design in advanced CMOS technology. Thus, the classical design methodologies can be divided into three parts: architectural behavioral-model validation, electrical schematic implementation, layout synthesis. Most of the basic steps among the different synthesis levels rely on powerful numerical optimization tools coupled to performance estimation tools. Most of variability-aware design methodologies present the worst-corner and the Monte Carlo simulation for design verification and performance estimation. However, some few works are describing the variability-aware design methodologies using behavioral modeling and so improving the design optimization. Moreover, the reliability-aware design methodologies are mostly concentrated on reliability analysis around a nominal corner. Few works present the combination of variability and ageing phenomena under a reliability analysis. Hence, we identify the opportunity to innovate proposing the reliability estimation in early design stages.

Chapter 4 proposes the reliable-circuit design for the BLIXER, the DCO, and the PGA. First, we decided to represent variability and ageing degradations using simple transistor electrical characteristics as drain current, transconductance and threshold voltage. Using a CMOS 65 nm characterization, we are able to propose a nominal circuit design. Unfortunately, not all optimum circuits are reliable circuits. Then, we propose a failure analysis to early estimate how reliable the circuit is. Using the electrical characteristics variation to represent variability and ageing degradations, we can discuss some design strategies improving the circuit reliability in a bottom-up approach.

Chapter 5 proposes the reliable-architecture design for the RF front-end architecture using a top-down approach. Now, the challenges turn in characterizing the building blocks reliability and identifying system-level failure conditions. Thus, we develop failure analysis to early estimate how reliable the architecture is. Architecture design strategies not often lead to an optimum design, but more likely in some trade-off equilibrium. In such an equilibrium, the designer should suppose some trade-offs among the specified performance, variability and ageing degradation. In Chapter 5, we present some simulation results of two different strategies and the design trade-offs found.

Chapter 6 organizes the experiences gained in the design of the reliable RF front-end in a design flow. Thus, we propose a new design flow for AMS/RF circuits that takes into account the variability and ageing trade-offs. During Chapter 6, we develop a comprehensive discussion around the design strategies and how they lead to an optimum or equilibrium on variation sharing. After that, we present the variability and the ageing as new design criteria imposing new design challenges. Moreover, we complete the discussion presenting how that trade-off can influence the design strategies.

Chapter 7 concludes this work resuming how we achieve the main objective and the sub-objectives. Thus, we enumerate the main contributions of this work and our publications during this research. Finally, we can discuss future trends in CMOS technologies and highlight the needs imposed in an AMS/RF circuits design flow taking into account the ageing and variability performance degradation.

CHAPTER 2

VARIABILITY AND AGEING IN ADVANCED TECHNOLOGIES

2.1 INTRODUCTION

Technology scaling will continue imposing new criteria for IC design. The variability and ageing will be challenges to future RF devices, just as today challenges arise from area and power consumption. Die size, chip yields, and design productivity have thus far limited transistor integration in VLSI designs. However the market needs more functions, circuit reconfiguration, and multi-standard operation.

In advanced technologies, transistor leakage continues to increase. There are leakage avoidance, tolerance, and control techniques. However, as technology scales further, new challenges will emerge, such as variability, single-event upsets, and device ageing degradation. These problems will inevitably lead to inherent unreliability in components, posing serious design and test challenges. Even today, design methodologies attempt to consider variability and ageing issues, but both problems are not seen as the same trends. It copes with variability in transistor performance through careful design, as well as testing for different product quality binning. With continued technology scaling, the impact of these issues is becoming greater, and design for reliability techniques shall be required dealing effectively with such issues.

Random dopant fluctuations, negative bias temperature instability (NBTI) and hot carrier injection (HCI) cause variation in transistor parameters. Sub-wavelength lithography, which generates line edge roughness (LER) and thus variation in the device performance, will continue until extreme-ultraviolet technology becomes available. Electromigration (EM), which is a transport of mass in metals, causes LER and thus variation in the passive device performances. EM will increase with high current density stress in shrunk interconnections. The increasing power density increases heat flux, leading to greater demand on the power distribution system and increasing NBTI phenomenon. This greater demand provokes voltage variations, as well as hot spots on the die with increased leakage power consumption. Soft oxide breakdown (SBD) also increases the leakage power consumption, changes the voltage distribution leading the circuit to more HCI and finally to fail.

Thus, systems designs are facing static (integrated process variability) and dynamic (circuit ageing degradation) variations [30]. Most of design methodologies optimize performance, area, and power consumption. However, they ignore performance variation in the presence of variability and ageing. To include them, a complex-design optimization capability, taking into account all variation sources and their characteristics, is needed. Therefore, it is very important for the development of new design methodologies, studying techniques to avoid and to mitigate the variability and the ageing.

Process variability and device reliability have similar consequences. The first is partly responsible for device infant mortality and the second for device ageing failure. Moreover, the reliability constraints shall be included to the design techniques as the variability constraints are already included. Therefore, it is important to understand the physical phenomena (NBTI, HCI, EM and SBD) in the operation environment and the sources of variability. Both informations will help the proposition of a new design methodology including variability and ageing degradations.

2.2 SOURCES OF VARIABILITY

Process variation is defined as the deviation from designed values for a layout structure or circuit parameter. These variations may differ in their causes and consequences. Thus, the circuit variability is the set of the process parameters variations, misalignments of the integration masks, and the consequences in the circuit performance among different samples. The variability has long been a critical issue in integrated circuit (IC) design.

IC performance variations is caused by two groups of factors [6]:

- *Environmental factors* arise during the circuit operation, and include variations in power supply voltage, noise coupling among nets and temperature. They depend on time, schematic topology, and external agents. They are also called temporal (or dynamic) variations, and directly impact the circuit reliability. They have always been analyzed at the local parameters (or performance) variation.

- *Physical factors* arise during manufacturing and result in structural device and interconnect parameter variations. They include Random Dopant Fluctuations (RDF) and Line Edge Roughness (LER), causing connections, active and passive devices variability. Such variations are essentially permanent; they are also called spatial variations, and may reduce the parametric yield, and potentially introduce catastrophic yield loss. They have always been analyzed at the global parameters (or performance) variation. In the advanced technologies, the physical factors are also responsible for an increasing local parameters (or performance) variation. In fact, the ICs (like multi-core processors) are large enough that the old global sources of variation become local parameters variations.

Thus, the physical parameter variations can be classified into two categories: Die-to-die (D2D) and Within-die (WID) [7]. D2D are global variations, resulting from lot-to-lot, wafer-to-wafer, and a portion of the within-wafer variations. They affect all transistors and interconnections on a die equally. WID are local variations, consisting of random and systematic components inducing different electrical characteristics across a die.

In spite of the different factors and categories of sources of variability, we will deal with them as a variation of the device performance. At this point, we shall understand the sources of variability. Therefore, this knowledge consists in an important step in order to propose new design methodologies taking the variability into account.

2.2.1 DIE-TO-DIE (D2D) VARIATIONS

The D2D is the difference of some parameter values across supposed identical dies. Those dies are either fabricated on the same wafer, or different wafers, or come from different lots, resulting in the denominations lot-to-lot, wafer-to-wafer, and within-wafer variations.

In circuit design, the D2D is typically modeled with the same deviation with respect to the mean of such parameters across all devices or structures on any chip. The parameters often modeled are:

- the transistor model (e.g. BSIM4) parameters,

- wire width and length on a given layer connection, and

- resistivity, and permittivity on passive devices.

It is assumed that each contribution in the D2D variation is due to different physical and independent sources. Such assumption is usually sufficient to lump these contributions into a single effective D2D variation component with a unique mean and variance.

Thus, a parameter (or characteristic) distribution can be obtained by silicon measurements from a large number of randomly selected devices across chips on some wafers (or and lots). Then, the mean

and variance are estimated from the approximately normal distribution of these devices. In this approach, named the lumped statistics [6], the details of the physical sources of these variations are not considered. The combined set of underlying deterministic as well as random contributions is simply lumped into a combined random statistical description.

2.2.2 WITHIN-DIE (WID) VARIATIONS

The WID is the spatial parameter deviation within a single die. Such WID variation may have several sources depending on the physics of the manufacturing steps. In the designer's point of view, the concern is how a WID variation may impact on performance or parametric yield of the circuits. Moreover, the WID variation contributes to the loss of matched behavior between nearby structures. Individual MOS transistors or/and segments of signal lines may vary differently from designed or nominal values (e.g. RDF). They may also differ unintentionally from each other (e.g. LER).

Two sources of WID variations are particularly important:

- *Wafer-level* variations are small fluctuations across the spatial range of the die. Chemical and/or physical steps on IC layers deposition and/or polishing might introduce systematic variations across the die.

- *Layout dependencies* may create additional variations that are inherent to the component and its neighborhood. These variations are due to photolithographic interactions, plasma etch micro-loading, layer deposition and polishing, or other causes. The range of such perturbations can vary between microns to millimeter range. The line distortions in exposure and distortions in lens and other elements of the lithographic system are within the range of a micron or less. However, the film thickness variations arising in chemical-mechanical polishing (CMP) and chemical or physical vapor deposition (CVD or PVD) may occur in the millimeter range.

While such variations may be systematic in any given die, the set of these variations across different dies may have a random distribution. A random-WID parameter variation fluctuates randomly and independently from device-to-device. A systematic-WID parameter variation results from a repeatable and governing principle. In this case, the device-to-device correlation is empirically determined as a function of the distance between the devices. Although systematic-WID variations exhibit a correlated behavior, the profile of these variations can randomly change from D2D. From a designer's perspective, systematic-WID variations behave as continuous and smooth correlated random-WID variations [7].

2.3 VARIABILITY TRENDS

The ITRS report [31] has expected that the manufacturing variations will continue increasing relatively to their nominal values. Some of the technology parameters and their 3σ variations are summarized in Table 2.1. It is remarkable that the 3σ variability is becoming more significant in comparison to the nominal value with the technology shrinking. Furthermore, the WID variations were reported being significantly increasing, which brings the structure-mismatch as the most important concern.

Table 2.1: Technology process parameter (nominal/3σ variations): ITRS report trends.

Leff (nm)	Tox (nm)	W (μm)	H (μm)	ρ (mΩ)
250/80	5.0/0.40	0.80/0.20	1.2/0.30	45/10
180/60	4.5/0.36	0.65/0.17	1.0/0.30	50/12
120/45	4.0/0.39	0.50/0.14	0.9/0.27	55/15
100/40	3.5/0.42	0.40/0.12	0.8/0.27	60/19
70/33	3.0/0.48	0.30/0.10	0.7/0.25	75/25

The ITRS report [31] had shown even worst results for the V_{th} parameter, summarized in Table 2.2. At this point, the variability is achieving values comparable to the nominal value. These trends indicate that it will not be possible anymore to design an IC without a careful variability analysis and management.

Table 2.2: Technology process variability: ITRS report trends.

L (nm)	250	180	130	90	65	45	32	22
V_{th} (mV)	450	400	330	300	280	200	150	100
$\sigma_{V_{th}}$ (mV)	21	23	27	28	30	33	47	57

The state-of-the-art [15, 6, 32] has investigated the trends of the process-induced variations in technologies beyond 22 nm. Following the technology scaling trends, CMOS devices are expected to continue shrinking over the next two decades. While the semiconductor industry looks toward the 22nm technology node, some manufacturers are considering a transition from planar CMOS transistors to the three-dimensional (3D) FinFET device architecture. Beyond a 22nm node device, the thin width might be on the order of 10-15nm. As such devices approach the dimensions of the silicon lattice; they can no longer be described, designed, modeled, or interpreted as continuous semiconductor devices. In contrast, the process variability comparisons show that SOI FinFETs are likely to have superior matching characteristics [32]. Thin height and width are likely to be more easily controlled in the SOI process, while the bulk process faces significant manufacturing and process control challenges. Therefore, the beyond 22nm devices will need new analysis and synthesis tools, breaking the CMOS paradigm.

2.4 Ageing Physical Phenomena

The physical phenomena responsible for the most ageing degradation in active devices are

- Hot Carrier Injection (HCI),
- Negative Bias Temperature Instability (NBTI) and
- Time Dependent Dielectric Breakdown (TDDB).

On the other hand, the passive devices also have ageing degradation, mostly caused by

- Electromigration (EM).

The ageing physical phenomena degrade the circuit performance with crystalline structure defaults which changes the device parameters. The NBTI and the HCI are responsible for traps at substrate-oxide interface. These interface traps are dangling bonds no longer hydrogen passivated. The reaction and diffusion model is the most accepted model for interface trap creation in both physical phenomena. The microscopic model and the hydrogen role in present and future technologies are described; they are the basis of the reaction and diffusion model and also explain the TDDB event.

2.4.1 Hot Carrier Injection (HCI)

Hot carrier injection is the phenomenon in solid state devices or semiconductors where either an electron or a hole gains sufficient kinetic energy to overcome a potential barrier, becoming a *hot carrier*, and then migrates to a different area of the device. In electronic devices, this phenomenon occurs at the end of the drain junction of a transistor in saturation, as illustrated in Figure 2.1. These energetic carriers injected can get trapped or cause interface states to be generated, and then these defaults lead to threshold voltage shifts and transconductance degradation of MOS devices.

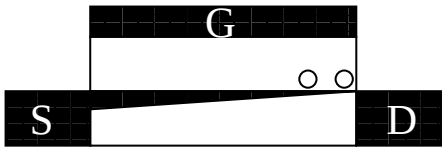


Figure 2.1: Interface traps in MOS structure under HCI degradation: illustration.

The first report of HCI measurements was done in 1970. The work of Hisashi Hara *et al.* [33] have reported a new instability found in p- and n-channel MOS transistors. The phenomenon occurs when a higher voltage in an excess of a breakdown voltage is applied to the drain electrode, so that the breakdown voltage drifts to a higher value and the drain current also increases. They highlighted that:

1. the semiconductor surface near the drain becomes p-like in the n-channel transistors and n-like in the p-channel transistors and thus the active channel length is shortened,

2. this is caused by charging of the gate oxide due to injection of electrons or holes generated during the drain avalanche breakdown, and

3. electron and hole injections are much affected by electric field across the oxide over the drain junction.

The research [1] has proven that the HCI in NMOS devices are more significant than in PMOS devices. The reason is to become *hot*, and enter the conduction band of the dielectric, an electron must gain a kinetic energy of 3.3 eV (for an SiO_2 dielectric), however for holes the valence band offset dictates, they must have a kinetic energy of 4.6 eV. Therefore, the holes are called much *cooler* than electrons (lower mobility), and suffer less HCI than electrons. However, the state-of-the-art has presented that PMOS devices suffer bigger consequences in the parameters degradation even the physical causes are smaller [34].

Later, for the injection of hot carriers into the dielectric, four distinguished injection mechanisms were studied [35]:

1. Channel hot-electron (CHE) injection - when the V_{GD} is approximately equal zero and the V_{GS} is very high, the hot carriers are attracted to the gate and gain enough energy from the electric field across the channel to surmount the Si/SiO_2 barrier at the drain end of the channel. It is normally identified by as a gate leakage current.

2. Drain avalanche hot-carrier (DAHC) injection - with very high V_{GS} and a high V_{GD}, the hot carriers enter in avalanche multiplication. Measurement is difficult as both carrier types are injected simultaneously as substrate and gate leakage currents.

3. Secondary generated hot-electron (SGHE) injection - under photo electron-hole pairs induced generation, the high field region near the drain injects both carriers causing the DAHC as a secondary effect.

4. Substrate hot-electron (SHE) injection - with a high positive or negative V_{SB} bias, then the carriers in the substrate are driven to the Si/SiO_2 interface which gain further kinetic energy in the surface depletion region. It is normally identified by a substrate leakage current.

Published works have demonstrated several mechanisms which can cause HCI, when the transistor is under stress condition. First, the carriers are accelerated by the strength of the electric field. Then, a carrier gains kinetic energy in the electric field, which happens only while it has a *room to run*. This *room* is essentially the mean-free path of the carrier. As mean-free path decreases with increasing temperature then low operating temperatures can be a problem.

Circuit designs, which use too high voltage coupled with small dielectric thickness, will create stronger field across the layer and increase the presence of hot carriers. Removal of the stress may anneal some of the interface traps, but these traps are only present at the drain junction of the transistor. As a result, this degradation cannot be recovered in most cases.

The developed models for HCI focused on the leakage gate or substrate current increase and the threshold voltage shift. The most accepted model for HCI degradation is the power law dependence on the stress time. In 2007, Wang *et al.* [36] have characterized the model of Reaction-Diffusion mechanism (R-D model) for HCI stress in 65 nm CMOS technology. Their work shows the threshold voltage degradation ΔV_{th} due to HCI and the mobility degradation expressed as a function of interface traps, as the same in negative bias temperature instability (NBTI). Normally, both phenomena have their device degradation modeled equally, but their environment stress conditions are different.

In order to avoid, or at least minimize hot carrier degradation, several device design modifications can be made. These are for example a bias control, larger channel length, double diffusion of source and drain, and graded-doped drain junctions to name a few. Then, these hot-carriers-related device instabilities have become a major reliability concern in modern MOS transistors. They are expected to get worse in future generation of devices. Therefore, the study of the fundamental physical processes that result in device parameter variation due to HCI is essential to provide guidelines avoiding such HCI degradation. It has been the subject of numerous studies and is one of the challenges in sub-nanometer MOS reliability.

2.4.2 Negative Bias Temperature Instability (NBTI)

The negative bias temperature instability is a physical phenomenon which generates positive charges and interface traps in Metal Oxide Semiconductor (MOS) structures under negative gate bias stress and mostly at elevated temperature, illustrated in Figure 2.2. At these conditions, the interaction of inversion layer with hydrogen-passivated Si atoms can break the Si-H bonds, creating an interface trap and one H atom, that can diffuse away from the interface (through the oxide) or anneal an existing trap [37]. The NBTI is mostly observed in PMOS transistors (p-type MOSFET), because in this case the positive oxide charge and positive interface charge are additive.

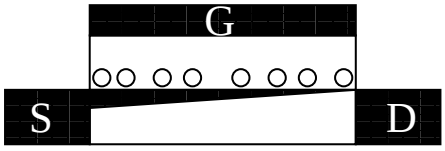


Figure 2.2: Interface traps in MOS structure under NBTI degradation: illustration.

A single microscopic model of NBTI is not fully established yet. One of the most accepted model is the reaction-diffusion model (R-D model), that attributes NBTI-induced degradation to a loss of passivated Si-H bonds at the oxide/Si interface and diffusion of hydrogen related species away from the interface. Mostly, the hydrogen release from the substrate/gate-oxide interface states and the hole trapping in the gate oxide are cited causes of NBTI [1].

The NBTI has already been reported 45 years ago, its effect has first been published by Miura and Matukura in 1966 [38]. The authors investigated a 300 nm SiO_2, thermally grown in a dry oxygen atmosphere, on an n-type silicon substrate. The metal contact was formed of aluminum. This MOS structure was stressed at a temperature of 300^o C at different gate voltages. The stress at each voltage was retained until the saturated degradation, which was then measured. Miura and Matukura proposed an electrochemical reaction under the influence of the strong electric field at the Si/SiO_2 interface. This reaction leads to positively charged oxygen vacancies in the SiO_2 film. As this mechanism proceeds at higher electric fields, it dominates the ion migration process which saturates at large bias [38].

Then, the NBTI was further characterized by researchers at Bell Laboratories, Fairchild Semiconductor, and RCA Laboratories. The effect was remarkable that an increase in positive charge under negative gate bias implicated a mechanism distinct from the migration of mobile ions [39]. The diffusion of hydrogen in poly-Si (poly silicon) gate, the dependence of the phenomenon on the oxide thickness and a variety of diffusing species (H-protio, H-deuterium or H_2), were observed. Also, the relaxation or recovery of the chemical bonds degradation were observed immediately after the stress voltage has been reduced, but not all degradation can be annealed and the remained interface traps are cumulative, reducing the circuit reliability.

The NBTI is typically seen as a threshold voltage shift after a negative bias stress has been applied to a PMOS transistor and later the degradation of channel carrier mobility as well as the driven current are observed [1]. Therefore, if the NBTI occurs, then the most important transistor parameters degraded are:

- decreasing transconductance g_m,
- decreasing linear drain current $I_{d,lin}$ and saturation current $I_{d,sat}$,
- decreasing channel mobility μ_{eff},
- decreasing subthreshold slope S,
- increasing off current I_{off}, and
- increasing absolute value of the threshold voltage V_{th}.

At DC operation, the NBTI rapidly shifts the threshold voltage. However, under AC operation condition, the recovery time may result in less severe shift in device parameters over a long term, if compared to the DC result. Moreover, at 50% duty cycle, the ΔV_{th} (threshold voltage shift) is less than one-half of its DC value [39]. A little or no frequency was noticed dependence up to 500 kHz, but then decreasing further above 2 MHz [39].

There are two important factors for accurate NBTI modeling:

- the physics of the degradation mechanisms have to be modeled as precisely as possible
- the experimental and measurement setup must lead to an exact description of the device state.

First, the physics of NBTI is very complex, and up to now it is well modeled by the R-D model. The model describes the trap generation and annealing as a power law dependence with the time of stress and the time of recovery. The most used parameter as a measure of degradation is ΔV_{th}. Wang *et al.* [36] have already characterized the R-D model for NBTI stress in 65 nm CMOS technology. Last, S. Mahapatra and M. A. Alam in 2008 [40] wrote an article which discusses the NBTI models and measured results, suggesting a common framework in all technologies for NBTI physical mechanism. The [40] highlights discrepancies on measured results, and points out ambiguities as a consequence of measurement artifacts. This conclusion was found with new measurements applying the on-the-fly measure technique developed by M. Denais et al. in 2004 [41] for ΔV_{th} characterization.

2.4.3 Time Dependent Dielectric Breakdown (TDDB)

One of the major factors presently assumed to be limiting the reliable lifetime of CMOS integrated circuits is the occurrence of the first failure (breakdown) of the gate dielectric. This phenomenon called time dependent dielectric breakdown (TDDB) and its physical mechanism is not yet fully understood [42]. The stochastic nature makes it difficult to determine the exact extrapolation of time-to-first event at operating conditions; and moreover, if the first event is sufficient for an observable circuit failure [8].

The gate dielectric layer has represented the critical component of metal-oxide-silicon field effect transistors (MOSFETs) since their adoption in 1960's. The dielectric or gate-oxide breakdown is manifested by a sudden loss of the layer's insulating properties. However, for older technologies it was less of the problem, where even more pessimistic reliability predictions satisfied the 10-year lifetime for 99.99% of circuits at operating conditions [11]. Nowadays, the probability of oxide breakdown has strongly increased with the increasing oxide field in downscaled MOSFETs technologies. The statistics of gate-oxide breakdown are described using the Weibull distribution [10]

$$F(x) = 1 - \exp\left(\left(-\frac{x}{\alpha}\right)^{\beta}\right); \tag{2.1}$$

where F is the cumulative failure probability, x can be either charge or time, α is characteristic life at 63% of failure probability and β is the Weibull slope.

For dielectric breakdown two scenarios are distinguished, extrinsic and intrinsic breakdown. Extrinsic breakdown is due to defects in the dielectric which can be introduced during different processing steps and it is out of the scope of this reliability study. The intrinsic breakdown is because of the nature of the dielectric itself and occurs at a certain electric field, defined by the dielectric strength. As the insulating layers are getting thinner and the technology is going to high-k insulator the probability of an external defect and therefore the probability of an extrinsic failure is decreasing. Hence, intrinsic failure is the most likely problem for future technologies.

The TDDB intrinsic failure was classified into two groups of failures depending on the magnitude of the post-breakdown conduction: hard breakdown (HBD) and soft breakdown (SBD) [43]. The HBD is considered a catastrophic failure of the device and consequently, of the entire circuit. HBD is the ancient gate-oxide breakdown, but in mid 1990's, the SBD was identified in ultra-thin oxide layers. The SBD is more likely to occur at operating conditions and has smaller effect on circuit operation. Because of SBD event, the reliable lifetime criterion has changed and SBD is a field of study in circuit reliability. After an SBD event, the devices and circuits can operate with a performance loss and the catastrophic nature was reexamined. Furthermore, there is some confusion in the literature over the precise characterization of breakdown modes: HBD and SBD. It may be caused by the lack of a precise definition of the terms and, for some experimental conditions, by the detection of one or the other breakdown mode may be difficult.

The SBD is the most interesting TDDB for this reliability study. Published works show that SBD occurs in a localized spot, and therefore the current after breakdown is independent of device area [44]. This spot is a number of aligned defaults into the dielectric, shown in Figure 2.3, which creates a conduction path between the channel and the gate. The size of this spot is estimated to be 10^{-14} to 10^{-12} cm^2. Several SBD events, occurring in different spots, may sometimes be seen in large-area devices prior to a HBD. In addition, the SBD current-voltage (I-V) characteristic between the gate and the channel path is independent of the oxide thickness, at least down to 3 nm, within a band of observed curves. In contrast to the HBD which shows a roughly linear (ohmic) I-V characteristic of resistance of about 10 kΩ between the gate and the channel. If the damage region remains localized and does not propagate, the SBD I-V characteristic between the gate and the channel is a power law of the position into the channel with an exponent of 3-6 [44].

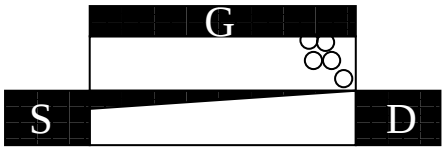


Figure 2.3: Aligned defaults in post-breakdown MOS structure: illustration.

2.4.4 ELECTROMIGRATION (EM)

Electrons passing through a conductor transfer some of their momentum to its atoms, and, at sufficiently high electron current densities (greater than 10^5 A/cm^2 [13]), these atoms may shift [12]. Electromigration is this physical phenomenon associated to the transport of mass in metals.

A metal crystalline structure can be defined as an orderly array of an aggregate of metal ions which are bound together by forces resulting from the ions sharing their valence electrons with the entire aggregate [45]. At any temperature other than absolute zero there is always a percentage of metal ions within the crystalline lattice that possess sufficient energy to escape from the potential wall which binds them in the lattice. The process of diffusion of the ions within their own lattice is termed *self diffusion* and is a random rearrangement of the individual ions which takes place under no concentration gradient or chemical potential, and therefore, it results in no net mass transport. However, under electromigration transport, the ion self diffusion is changed from a random process to a directional one by the presence of an electric field and charge carrier flow. This directional effect causes ions to migrate or diffuse downstream in terms of electron wind direction and vacancies move upstream [45]. Finally, the stressed conductor fails. The open in the film is caused by an accumulation of vacancies by a positive divergence in the ion flux.

The EM has been recognized as a potential semiconductor device wear out since J. R. Black works in 1970's. The most important circuit damage studied are [13]:

- the decrease of the electrical conductance,
- formation of open-circuit or close-circuit conditions,
- whiskers,
- thinning,
- localized heating,
- cracking of the passivated layer and
- the inter-level dielectrics.

There is no universally accepted model for EM failure mechanism, but the degradation is mostly evaluated by the median time to failure (MTTF). The MTTF is an approximate statistical method studied by J. R. Black [12], and described by

$$MTTF = AJ_e^n \exp\left(\frac{E_a}{kT}\right); \tag{2.2}$$

where A is a constant of the metal layer and process characteristics, J_e is the current density, k is Boltzmann's constant, T is the film temperature and E_a is the EM activation energy. The n is determined by Black as 2 [12], but it is proved that it may change depending on the residual stress and the current density conditions [13]. A range of values for E_a are reported, depending on the material characteristics, some of typical values are described at the Table 2.3

Table 2.3: EM activation energy values reported [13]

Metal Layer	E_a (eV)
Al	0.6 ± 0.1
Al + Cu 0.5%	0.95
Damascene Cu	0.94 ± 0.11

2.5 Ageing Trends

2.5.1 Hot Carrier Injection (HCI)

Future technologies advance in semiconductor manufacturing techniques and even increasing demand transistors with sizes close to their physical limits. However, it has not been possible to scale the supply voltage used to operate these devices proportionally due to factors such as compatibility with previous generation circuits, noise margin, power and delay requirements, non-scaling of threshold voltage, sub-threshold slope, and parasitic capacitance [1]. In aggressively scaled transistors, the consequent increase in internal electric fields comes with the additional benefit of increased carrier velocities, and also increased switching speed. The presence of large electric fields implies the presence of hot carriers. The carriers that have sufficiently high energies and momentum can get injected from the semiconductor into the surrounding dielectric films such as the gate and sidewall oxides as well as the buried oxide in the case of Silicon-On-Insulator (SOI) transistors [1].

2.5.2 Negative Bias Temperature Instability (NBTI)

In recent years, the NBTI is gaining much attention due to modern semiconductor technologies. Several effects conspire to bring NBTI to the attention of device and circuit designers. First, the operating voltage has not scaled as rapidly as gate oxide thickness, resulting in higher fields which enhance the NBTI [39]. Next, device threshold voltage scaling has not kept pace with operating voltage, which results in larger percentage degradation of drive current for the same ΔV_{th} [39]. Also, the NBTI is lower for thinner oxide, where sufficient tunneling current may flow to cause the generation of additional traps and interface states, the NBTI recovery is lower [39]. Finally, the addition of nitrogen into the gate dielectric used for gate leakage reduction and control of boron penetration, has had a side effect of increasing NBTI [39]. The following summarized aspects have been found to lead to increasing susceptibility to NBTI:

- higher oxide electric fields due to oxide scaling,
- higher temperatures due to higher power dissipation,
- replacement of buried p-channel MOSFETs with surface devices,
- the introduction of CMOS elevating the importance of p-channel MOSFETs, and
- nitride oxides with a higher permittivity.

Recent investigations show that NBTI is an important transistor reliability issue also in newest sub-nanometer MOS technologies. Measurements in high-k/metal-gate p-FETs [46, 47] show that NBTI induced ΔV_{th} is accompanied by an increase in the interface traps, similar to those observed for conventional p-FETs with oxide and poly-Si gates. Therefore, the NBTI is a circuit reliability constraint, being the threshold voltage shift or the drain current deviation as a criterion for device failure [37].

2.5.3 Time Dependent Dielectric Breakdown (TDDB)

Different technologies and circuits will have various degrees of sensitivity to oxide breakdown loss. Thus, more research is needed in order to develop a quantitative methodology for predicting the reliability of circuits after an SBD event. While the technology node is fast shrinking, the supply voltage is not keeping pace. Thus, the gate oxide field is increasing, resulting in a higher SBD probability event. Therefore, SBD degradation is becoming very critical as the technology node scales to 45nm and smaller.

In [48], the impact of voltage scaling on a 65nm and a 45nm SRAM bit cell was analyzed. V. Chandra and R. Aitken have evaluated the metric of the critical charge in the presence of gate oxide degradation [48]. They found that, as the oxide degrades, the value of the critical charge becomes a non-linear

function of the supply voltage. Furthermore, the values of critical charge also change with technology node scaling. The absolute values of the critical charge have decreased for both 65nm and 45nm nodes as the level of degradation increases. Moreover, the 45nm maximum critical charge is 30% smaller than the 65nm maximum critical charge, resulting on an increasing sensitivity to oxide breakdown.

The high-k technologies have presented higher breakdown immunity, as expected by the permittivity increasing. The physical properties of the Si/SiO_2 interface has been argued to be key to improve performance and reliability [10] in devices after a SBD event. As the technology node scales to sub-45nm, the effect of various reliability concerns becomes very crucial. The potential solution to TDDB is mitigating its consequences by oversized devices and redundancy. These features are already present in many systems, and they represent an increase of cost.

2.5.4 ELECTROMIGRATION (EM)

As ICs technologies continue to shrink the metal interconnections and interconnect current densities grow, EM will remain a concern. New technologies may reduce the EM impact, but performance requires increased interconnect reliability under condition of decreased metallization inherent reliability [13].

In advanced technologies, the interconnect Joule heating can strongly affect the maximum operating temperature of the metal interconnections and reduce the wire reliability by the combination of the EM with the Stress Migration (SM). The SM is a stress induced voiding in Cu dual damascene structure, first reported in 2002 [49]. SM is caused by the interaction between the thermomechanical stress in the interconnect systems and the diffusion of vacancies. Thus, the interaction between EM and SM is the interest of the state-of-the-art [50] studies with the advent of the dual-damascene Cu/low-k interconnects.

Unlike EM, SM may occur at the chip operating temperature, typically around 100-125^{o}C, where the thermal stresses are considerably higher. Moreover, EM failure time of the lower metal line could be strongly affected by the presence of the residual SM [50].

The interconnect Joule heating and the increasing of EM failure are already predicted in ITRS Report. These trends are summarized in Table 2.4. In the second line, the increasing interconnect Joule heating is presented by the metal temperature (T_m) at the maximum current density (Max. J_{rms}) projected in ITRS report [51]. In the third line, the required J_{rms} are presented in order to maintain the 105^{o}C metal temperature [50]. In both lines, the increasing EM and SM stress conditions with the continuous metal interconnection shrinking are clear. The consequences of EM and SM wearout are represented by the increase of the time delay (t_d) and the decrease of the Median Time to Failure (MTTF). In the fourth line, the increase of the time delay (Δt_d) are presented normalized by the t_d at 105^{o}C metal temperature [50]. In the fifth line, the decrease of the MTTF is presented by normalizing the technology MTTF values to the expected 130 nm MTTF. The MTTF decrease is estimated by the ITRS trends and the Black's equation (Eq. (2.2)).

Table 2.4: The increase of EM and SM failure, according to ITRS report trends [51].

L (nm)	130	90	65	45	22
T_m (^{o}C)	107	117	122	126	128
J_{rms} @ 105^{o}C (MA/cm^2)	0.79	0.82	1.0	1.2	1.6
$\Delta t_d/t_d$ (%)	1	5	9	9	10
MTTF	1.00	0.64	0.45	0.35	0.24

When the technology node is scaled down to the sub-45nm range, EM will remain a design and wearout issue in integrated circuits. The effect of SM on EM will become even more prominent in smaller vias, shrunk metal lines and the introduction of more porous low-k dielectric. The smaller vias imply in lower vacancy density, causing the metal line to reach the required critical tensile stress for void

nucleation. In addition, the introduction of more porous low-k dielectric lowers the critical nucleation stress. Therefore, this will reduce the MTTF, in turn, leads to a much earlier void nucleation at the interface, decreasing the IC's reliability.

2.6 Ageing Degradation Mechanisms

In this section, we will detail the physical and chemical mechanisms of HCI and NBTI degradation. Many physical and chemical models were proposed, however the most accepted are: Microscopic model and Reaction-Diffusion model. Both are able to well explain the physical and chemical origins of the transistor parameter degradation. On the other hand, we will not detail the physical and chemical mechanisms of TDDB and EM. We will see later (in Subsection 3.4.1) that both have not well adopted post-event models. Moreover, they are not taken into account in commercial simulation tools or well characterized in commercial design process kits.

2.6.1 Microscopic model

The hydrogen plays an important role in many technologically relevant processes of integration in silicon. The incorporation of hydrogen during the MOSFET fabrication process is associated to passivating silicon dangling bonds. These dangling bonds are found at surfaces, grain boundaries, interfaces, and in bulk silicon. They degrade the device quality, because they change the electrical characteristics of the transistors.

The hydrogen molecule (H_2) can easily form in most semiconductors. The binding energy is somewhat smaller than for H_2 in vacuum, but still large enough to make interstitial H_2 one of the more favorable configurations hydrogen can assume in the lattice. H_2 is the most common hydrogen species in semiconductors, because it has lower bond energy, lower migration barrier energy and higher chemical stability [52]. On the basis of the measurements, all of non-bonded hydrogen are present as H_2 molecules, located in centers of atomic dimensions inside the crystalline silicon lattice.

The occurrence of negative bias temperature instability (NBTI) and hot carrier injection (HCI) in semiconductor devices is directly connected with non passivated silicon dangling-bond and hydrogen-species diffusion through the crystalline silicon lattice. First studied in 1966, Miura and Yasuo Matukura [38] had reported an electrochemical reaction between holes and defects at the substrate/oxide interface, it was put forward as the physical and chemical mechanism for NBTI and HCI.

It is well established that the NBTI and the HCI phenomena are not dependent on current flow through the oxide, in contrast to time dependent dielectric breakdown (TDDB). The TDDB also provides reliability degradation. As oxides have become thinner, the increasing tunneling current leads to additional defect generation, with parameter degradation similar to the NBTI or the HCI.

The generation of interface traps is schematically expressed as

$$(\text{electrically inactive interface defect}) + \text{hole} \leftrightharpoons (\text{oxide positive charge}) + (\text{interface trap}) + X. \tag{2.3}$$

It is generally believed that the initially electrically inactive surface defect in equation (2.3) comprises a hydrogen passivated Si dangling bond (Si-H) and X is a hydrogen mobile specie (usually H_2) which diffuses through away from the interface. The interface trap is then supposed to be a silicon dangling bond ($Si^{\bullet}$) which results when H is removed from Si-H. The majority data for SiO_2/poly-Si, SiON/metal-gate and high-k/metal-gate FETs technologies are similar to those for conventional p-FETs [39, 47] expressed on the general scheme shown in equation (2.3).

2.6.1.1 SiO_2/poly-Si technology

The various proposals for reactants, in equation (2.3) are listed on Table 2.5 and they were presented in [39]. All the listed reactions, except the last, require a hole. However, attempts to prove this component have seen no variation in the rate of defect generation. The last reaction in Table 2.5 requires interstitial

atomic hydrogen. This reaction is known from radiation damage, plasma processing and hot electron experiments. The possible atomic hydrogen role has only recently been proposed. Note that the last three reactions do not place separate entities for the interface states and the oxide positive charge; in these reactions all of the positive charge is in the form of charged interface dangling bonds.

Table 2.5: Microscopic reaction models for SiO_2/poly-Si technology

Interface defect	X	Reaction
Si-H+Si-O-Si	Si-OH	Si-H+Si-O-Si+$h^+ \rightleftharpoons$Si$^{\cdot}$+Si^++Si-OH
Si-H+H_2O	H_3O^+	Si-H+H_2O+$h^+ \rightleftharpoons$Si$^{\cdot}$+H_3O^+
Si-H	H^+	Si-H+$h^+ \rightleftharpoons$Si$^{\cdot}$+H^+
Si-H	H_2	Si-H+$h^+ \rightleftharpoons$Si$^{\cdot +}$ + $1/2H_2$
Si-H	H	Si-H+$h^+ \rightleftharpoons$Si$^{\cdot}$+H
Si-H+H^+	H_2	Si-H+$H^+ \rightleftharpoons$Si$^{\cdot +}$+H_2

2.6.1.2 SiON/METAL-GATE TECHNOLOGY

In the 1980s, the use of nitrogen in gate oxides begins, as way of reducing boron diffusion from the gate into the channel and to improve low-field breakdown problems. In ultra-thin oxides ($\sim$3 nm or less), the silicon oxynitride (SiON) increases the dielectric constant, reduces the direct-tunneling leakage current and acts as a blocking barrier to impurities (similar to the silicon nitride passivating layer).

Although on relatively thick oxynitride ($\sim$35 nm), promising results for negative bias stress were found, the thinner oxides ($<$10 nm), used in current technology, with the addition of nitrogen to the gate oxide have presented serious adverse effect on NBTI. However, the recovery of interface traps is greater for SiON, if the stress is removed. One reason for this adverse effect is that the SiON has a lower activation energy ($\sim$0.1 eV) compared to SiO_2 ($\sim$0.2 eV) and also a shallower time dependence [39]. Other possible explanation is that oxynitride contains more water [39]. Also, electrical measurements have shown that the oxynitride exhibits a higher interface state density near the Si conduction band edge, whereas pure SiO_2 exhibits more states at mid-gap and close to the Si valence band edge [39].

Finally, a list of proposals for reactants are presented on Table 2.6, according to equation (2.3) [39].

Table 2.6: Microscopic reaction models for SiON/metal-gate technology

Interface defect	X	Reaction
Si-H+Si_2-N-Si-(N_xO_{3-x})	Si-NH-Si	Si-H+Si_2-N-Si+$h^+ \rightleftharpoons$Si$^{\cdot}$+Si^++Si-NH-Si or H^++Si_2-N-Si$\rightleftharpoons$$Si_2$-$NH^+$-Si
H_2O+Si-Si-N_3	H_2	Si-Si-N_3+H_2O+$h^+ \rightleftharpoons$Si-O-Si-NH^+-N_2+H then Si-H+H$\rightleftharpoons$Si$^{\cdot}$+H_2

2.6.1.3 HIGH-K/METAL-GATE TECHNOLOGY

In recent years, there has been a great effort to integrate high dielectric constant (high-k) materials and metal gates into FETs, in order to continue the scaling of FETs for future technologies. The hafnium oxide (HfO_2) is one of the leading high-k candidates and several different metals, such as W, Re, TiN, TaN; are being explored for replacing poly silicon gates. The NBTI is also a reliability issue in high-k/metal-gate p-FETs. The dielectric consists of a stack with an interfacial oxide plus HfO_2 layer, and,

under bias and temperature stress, measurements have shown an increase on threshold voltage and on interfacial trap density induced by NBTI.

Investigations have shown that NBTI in high-k/metal-gate p-FETs is similar to that in conventional SiO_2/poly-Si p-FETs at elevated temperature [39]. Also, HfO_2 and SiON p-FETs have comparable threshold variation at 10 years and high-k layers do not induce additional NBTI degradation [46, 47]. Earlier conclusions have appointed the interfacial oxide as the main cause of NBTI degradation and the high-k layers do neither reduce nor increase the device reliability.

2.6.2 Reaction-Diffusion model

One of the most accepted model is the reaction-diffusion model (R-D model), it was established by Jeppson and Svensson in 1977 [53]. In this model, the threshold voltage shift (ΔV_{th}) is caused by a strong increase of the density of Si-SiO_2 surface traps. On Jeppson and Svensson work, a detailed study of the increase of the number of surface traps in MOS structures after negative bias temperature stress (NBTI) at temperatures (25-125 ^{o}C) and fields (400-700 MV/m) was done. A power law dependence with the stress time as $t^{1/4}$ was characterized at low fields and at high fields it increases linearly with the stress time. The first physical model, proposed by Jeppson and Svensson, explains the surface-trap growth as being diffusion controlled at low fields and tunneling limited at high fields.

Since then, the model has been continuously refined and describes the degradation process (reaction) of the chemical bonds Si-H or Si-O, and the transport mechanism (diffusion) of the hydrogen species away from the interface. This generation process is described [9, 39] as

1. ***Reaction***: the degradation process as a reaction at the Si/SiO_2 interface generating an interface state (N_{it}) as well as releasing a mobile hydrogen related species (N_X) is described as

$$\frac{\partial N_{it}(t)}{\partial t} = \underbrace{k_F\left(N_0 - N_{it}(t)\right)}_{\text{trap generation}} - \underbrace{k_R N_{it}(t) N_X(t)}_{\text{trap annealing}}; \tag{2.4}$$

   where k_F is the interface-trap generation, k_R the annealing rate and at $x = 0$. The symbol N_0 denotes the initial number of electrically inactive Si-H bonds.

2. ***Diffusion***: the transport mechanism is driven by the gradient of the species density concentration, these species diffuse away from the interface toward the gate described by

$$\frac{\partial N_{it}(t)}{\partial t} = D_X \frac{\partial N_X(x,t)}{\partial x} + \frac{\delta}{2}\frac{\partial N_X(x,t)}{\partial t} - \frac{D_X E_{ox}}{V_t} N_X, \tag{2.5}$$

$$\frac{\partial N_X(x,t)}{\partial t} = D_X \frac{\partial^2 N_X(x,t)}{\partial x^2} - \frac{D_X E_{ox}}{V_t}\frac{\partial N_X(x,t)}{\partial x} \tag{2.6}$$

   at the dielectric, and

$$D_X \frac{\partial N_X(x,t)}{\partial x} = k_P N_X \tag{2.7}$$

   at oxide/poly-Si interface. Where D_X is the diffusivity of the hydrogen species, V_t is the thermal voltage, E_{ox} is the oxide electrical field and k_P is the surface recombination velocity at oxide/poly-Si interface.

In order to obtain an analytical solution, an infinitely thick oxide (i.e. $t_{ox} > \sqrt{4D_X t}$), the conservation rule

$$N_{it}(t) = \int N_X dx \tag{2.8}$$

and the case of a neutral diffusing species ($X = H$) are usually considered. This process can be illustrated in Figure 2.4, which shows the initially passivated dangling bond and the reaction of non passivating dangling bond at the Si/SiO_2 interface, as well as the hydrogen species diffusion.

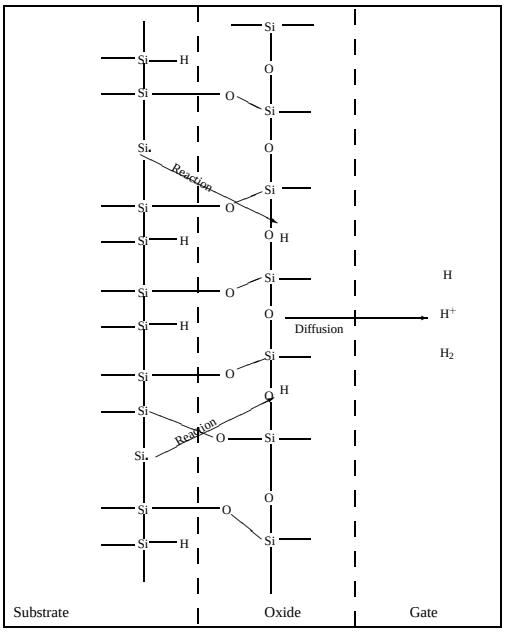


Figure 2.4: Reaction and Diffusion mechanism: illustration.

The general solution of the R-D model in stress phase anticipates five different regimes of time-dependent interface trap generation. They are illustrated in Figure 2.5, and described as

1. **Early phase**: N_{it} and N_H are small and, by equation (2.4), we found

$$\frac{dN_{it}}{dt} = k_F N_0 \Longrightarrow N_{it} = k_F N_0 t; \tag{2.9}$$

2. **Quasi-statical equilibrium**: $dN_{it}/dt \cong 0$, by equations (2.4) and (2.8), we found

$$k_F N_0 - k_R N_{it} N_H = 0 \Longrightarrow N_{it} = \sqrt{\frac{k_F N_0}{k_R}} t^0; \tag{2.10}$$

3. **Hydrogen diffusion through the oxide**: as analyzed by Jeppson and Svensson in 1977 [53], we found

$$N_{it} = 1.16 \sqrt{\frac{k_F N_0}{k_R}} D_H^{1/4} t^{1/4}; \tag{2.11}$$

4. **Hydrogen diffusion through the gate**: as analyzed by Alam and Mahapatra in 2005 [9], we found

$$N_{it} = \sqrt{\frac{1}{2\left(\frac{D_H}{k_P} - t_{ox}\right)}} \sqrt{\frac{k_F N_0}{k_R}} D_H^{1/2} t^{1/2}; \tag{2.12}$$

5. **Final equilibrium**: when all Si-H bonds are broken $dN_{it}/dt \cong 0$ and $N_{it} = N_0$.

For charged species, but again assuming infinite oxide thickness, the solution at hydrogen diffusion through the oxide phase changes to

$$N_{it} = 1.16 \sqrt{\frac{k_F N_0}{k_R}} D_X^{1/4} \frac{D_X E_{ox}}{V_t} \frac{1}{2\sqrt{D_X}} t^{1/2}; \tag{2.13}$$

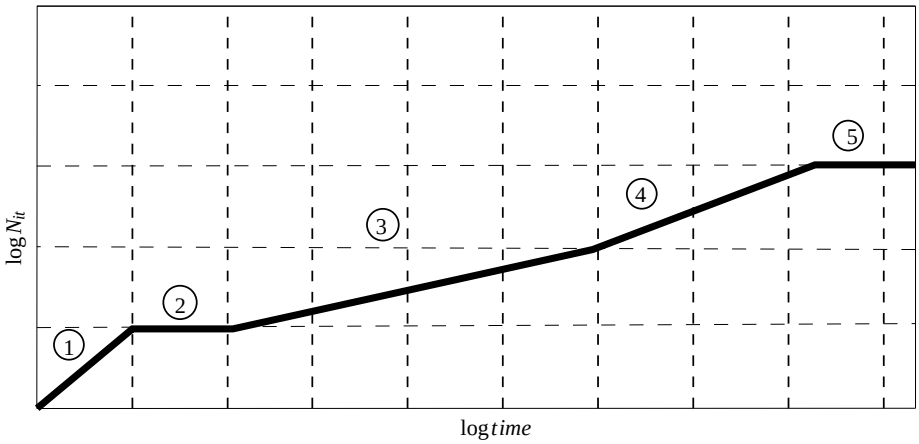


Figure 2.5: The five different regimes of time-dependent interface trap generation, as obtained from the reaction-diffusion model.

where D_X is the diffusivity of the charged hydrogen species. The larger time exponent was never measured, but the majority of the measured results present a time exponent very close to the time exponent at uncharged case. Therefore, the notion of a charged diffuser was rejected [39].

On the other hand, if the stress is removed (at t_0), the R-D model predicts the recovery of broken Si-H bonds with the number of traps generated ($N_{it}(t_0)$) and the density of interface hydrogen ($N_H(t_0)$). For a linear hydrogen profile,

$$N_{it}(t_0) = \frac{1}{2N_H(t_0)\sqrt{D_H t_0}} \tag{2.14}$$

and this traps are annealed at time $t + t_0$. By equation (2.4), with $k_F = 0$ (no more interface-trap generation), assuming that the original hydrogen front continues to diffuse, Alam and Mahapatra in 2005 [9] found

$$N_{it}(t + t_0) = N_{it}(t_0)\left(1 - \frac{\sqrt{\frac{\xi t}{t_0}}}{\sqrt{1 + \frac{t}{t_0}}}\right). \tag{2.15}$$

The interface trap species modeled by reaction and diffusion have as a consequence the variation of transistor parameters: threshold voltage, transconductance, saturation current. For example, the NBTI degradation is generally monitored through threshold voltage shift (ΔV_{th}). In this case, the transistor degradation mainly results from the loss of passivated Si-H bonds at the Si/SiO_2 interface and the ΔV_{th} is normally approximated by

$$\Delta V_{th} = \Delta V_{it} = \frac{qN_{it}}{C_{INV}}; \tag{2.16}$$

where q is the electron charge and C_{INV} is the inversion layer capacitance.

Concluding, the R-D model successfully explains both the power-law dependence of interface trap generation and the mechanics of interface trap annealing. Furthermore, it offers no quantitative predictions regarding the field or temperature dependence. However, the k_F, k_R is field or temperature dependent, as predicted by Alam and Mahapatra in 2005 [9], the model covers the loss of passivated Si-H bonds and not the most microscopic reactions found in the literature [39]. Moreover, the precise power-exponents also make the R-D model difficult to fit the measured data, a wide variety of exponents is observed.

The first reason of the wide range for power exponents is due to the degradation recovery (under NBTI stress) that takes place during the delay between stopping the stress for measurement and the start of the next stress phase. The second is the assumption of constant time-independent diffusion coefficient for hydrogen diffusion in the oxide. Finally, the large variety of hydrogen species are difficult

to be analyzed separately, because the interfaces can be absorber or reflective for different species, and chemical reactions between them can change the hydrogen diffusion profile or velocity.

The [40] takes a survey of several works that have presented different stress conditions, measurement methods and gate-oxide process. At this point for the NBTI example, the threshold voltage shift is estimated as [40]

$$\Delta V_{th} = AP^{2/3} \exp\left(BE_{OX}\right) \left(D_H \exp\left(\frac{E_A}{kT}\right) t \right)^n ; \tag{2.17}$$

where A and B are function of N_0, P is the hole density at inversion layer, D_H is the diffusion coefficient of H_2, E_{OX} is the oxide field, E_A is activation energy necessary to create the dangling bond. k is Boltzmann's constant and T is the temperature.

2.7 CONCLUSION

In Chapter 2, we described, as best as we know, the sources and the trends of the device-characteristics variations. The variations were separated as *static* or at time zero, and *dynamic* or time dependent variation. The *static* variations are represented by integration process variability and devices mismatch. The *dynamic* variations are represented by probabilistic events and device ageing.

As the dimensions shrink in advanced technologies, the errors become more important, increasing the device failure rate. Moreover, the reduced lifetime has become a huge constraint for reliable devices with the increased performance lost. In this way, the knowledge of the physical mechanisms of variability and ageing is essential to propose improved design methodologies aiming more reliable devices.

CHAPTER 3

STATE-OF-THE-ART OF DESIGN METHODOLOGIES

3.1 INTRODUCTION

Technology scaling will continue to increase the number of functions per area and the complexity of the devices. Such scaling has allowed the design of huge Systems on Chip (SoC) and integrated Systems in a Package (SiP). Most of consumer-market applications demand an increasing performance for a low-cost full integrated solution in a shrunken time-to-market. Managing these criteria implies the use of electronic design automation (EDA) methodologies and tools, increasing the designer efficiency.

AMS/RF EDAs have changed the focus from the bottom-up design, transistor-level optimizations to behavioral system-level hierarchical design refinement flows. The systematic top-down design is only possible using computer-aided design (CAD) tools. However, EDA tool support for AMS/RF designs has strongly been lagging behind the digital design. The logic synthesis, standard-cell-based place and route, presented in digital design, are not possible for the application driven AMS/RF devices, mostly full-custom. As the AMS/RF IPs cannot be easily reused, the design methodology reuse and the design background are the only retained design experiences. Moreover, integrating AMS/RF and digital circuits on the same die creates additional problems of crosstalk and signal integrity, which require tool support for modeling and analysis.

The use of nanometer CMOS technologies brings significant challenges imposed to circuit design methodologies. These challenges include [14]:

- the advent of no longer negligible leakage currents and the impact of power consumption;
- the increasing variability of the technological process parameters (V_{th}, doping level, widths, lengths, and others);
- the reduction of the supply voltage without the pace of the V_{th}, dropping the headroom for AMS/RF circuits;
- the advent of novel materials, like high-k dielectrics, and novel devices, like FinFETs and CNT-FETs;
- the increasing importance of well known degradation phenomena, like HCI, NBTI, and EM; and the advent of degradation phenomena not observed before, such as SDB and SM.

Thus, the advent of new EDA methodologies and tools, including the nanometer CMOS technology challenges, is a need for AMS/RF design pace with the digital design. In this chapter, we show the state-of-the-art of design methodologies, the possible solutions and the not solved challenges for AMS/RF design in advanced CMOS technology.

3.2 CLASSICAL DESIGN METHODOLOGIES

IC's design has three parts: architectural behavioral-model validation, electrical schematic implementation, layout synthesis. Most of the basic techniques steps among the different synthesis levels rely

on powerful numerical optimization tools coupled to performance estimation tools, see Figure 3.1. Then, performance and specification are showdown by the evaluation of a cost function. If the optimal design is found then the designer can move down the synthesis level, in a top-down approach, or move up the synthesis level, in a bottom-up approach.

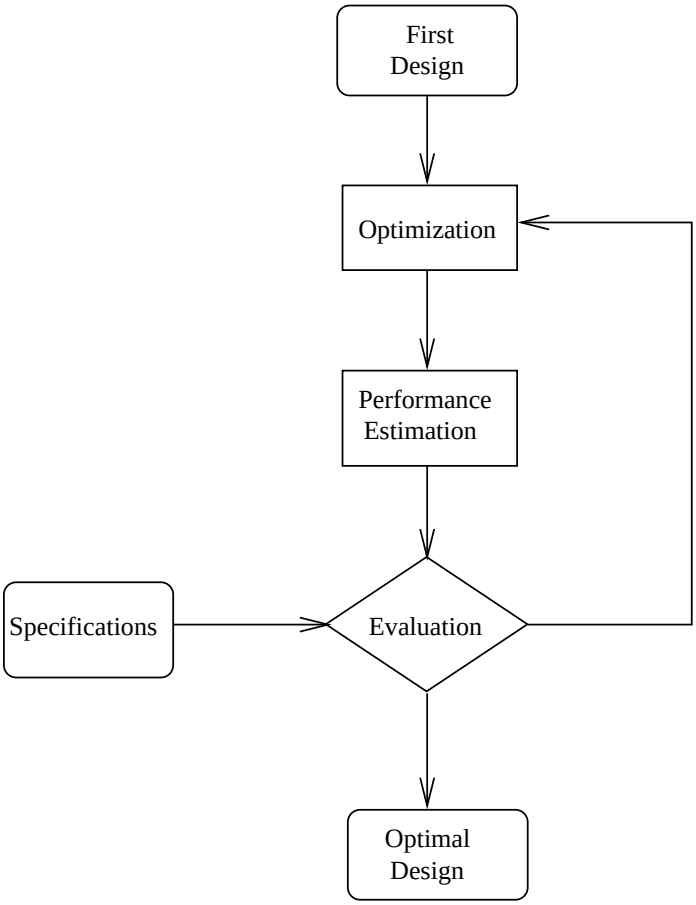


Figure 3.1: Classical design methodologies: the basic techniques steps.

The Figure 3.1 shows the basic flow of the most EDA methodologies and tools. Giving the electric-schematic synthesis level example, the sizes and biasing of all transistors as the values of needed passive components have to be determined in order to meet the circuit performance specification at some trade-off cost. First, the optimization tool determines such design values, using trade-off criteria. Then, the performance estimation tool, often a simulator, characterizes the optimal design candidate. Finally, the evaluator judges if the achieved performance is enough according to the specification. After some iterations and computational cost, the optimal design may be found. So that, the state-of-the-art EDA methodologies and tools have the objective of:

- increasing the performance estimation accuracy;
- reducing the optimization convergence time;
- controlling the computational cost as the design complexity increase;
- proposing new criteria trade-offs;
- reusing the design experience improving itself.

In Subsection 3.2.1, we will discuss the advantages and disadvantages of top-down and bottom-up approaches. Subsection 3.2.2 will describe the state-of-the-art optimization tools, and Subsection 3.2.3 will present the state-of-the-art performance estimation tools.

3.2.1 TOP-DOWN AND BOTTOM-UP EDA APPROACHES

A *top-down approach*, also known as step-wise design, is essentially the breaking down of a system to gain insight into its compositional sub-system. In a top-down point of view, the system is first formulated, specifying performances and parameters but not detailing them in any low-level building block. After, each building block is then detailed, sometimes in many additional lower-level devices, until the entire specification is reduced to base elements. A top-down model is often specified in blocks, making easier to manipulate the required mathematical behavioral. However, these blocks may fail to model and elucidate physical mechanisms or be detailed enough to realistically validate the behavioral model.

A *bottom-up approach* is the piecing together of low-level devices to give rise to high-level systems. Thus, such approach builds toward larger and more complex systems by starting at the low-level and maintaining precise control of physical structure. In a bottom-up approach, the individual base devices of the system are first designed in detail. These devices are then assembled together to form a larger building block, which then in turn are linked, sometimes in many levels, until the complete high-level system design. However, this strategy may result in a mass of confusedly interlaced devices and building blocks, developed in isolation and subject to local optimization as opposed to meeting a global architecture performance.

Starting from small device design building a product application is the most intuitive engineering approach. The bottom-up approach has high design space coverage in early phases, smaller time-to-market only a few number of IC functions, higher observability of parameter variations, and higher costs on design reuse. The top-down approach has a tactical but limited design space coverage, bigger time-to-market but higher number of IC functions adapted, masked parameter variations in the device performance, and lower costs on design reuse. And, the design reuse in the top-down approach is responsible on the speed-up of the time-to-market for new versions of the product.

Moreover, manage the design complexity of the AMS/RF systems integrated on a SoC or SiP requires designers to adopt more system-oriented EDA methodologies and tools. Such system-oriented EDA is only possible at the top-down approach, where the following phases can be identified: system specification, architectural design, circuit design, circuit layout and system layout assembly [14]. The advantages of adopting a top-down design methodology are:

- the possibility to perform system architectural evaluation and better optimization before a high cost circuit implementation;
- the anticipation and mitigation of risks and of problems interfacing building blocks, reducing the number of design iterations;
- the use of hybrid system simulation to fast explore critical parts in the early stage of the flow, combining different model levels like behavioral, schematic and extracted;
- the early development of test procedures in parallel with the block design;
- the better overall system solution in huge and complex SoC or SiP, achieving the required time-to-market.

The top-down solution requires however some investment from the design team, especially high-level modeling and setting up a sufficient model library for a target application. Such a high-level point of view may distance the designer of the physical implementation. In such a case, the feasibility or optimality of a solution may depend on the low-level details like matching limitations, circuit non-idealities, layout

effects and others. Therefore, the high-level models must include such effects to the extent possible, but it remains difficult in practice to anticipate or model everything accurately at higher levels.

At the end, there is no best design approach. A complex AMS/RF system integrated on a SoC or SiP shall be specified in top-down. However, the piecing together of low-level devices will require a bottom-up approach. Both approaches should be developed in parallel, linking architectural needs and low-level details. Therefore, accurate high-level models may be obtained optimizing the local low-level and the global high-level characteristics.

3.2.2 Optimization Tools

Synthesis can be carried out by the following two different approaches: knowledge- and optimization-based [54]. The basic idea of knowledge-based synthesis is to formulate design equations in such a way that given the performance characteristics, the design parameters can be calculated. In optimization-based synthesis, the problem is translated into function minimization problems that can be solved through numerical methods.

The optimization synthesis is based on the introduction of a performance evaluator within an iterative optimization loop. The system is often modeled with a performance function, capturing the behavior of a circuit topology. The performance estimation will be discussed in Subsection 3.2.3. Techniques for analog circuit optimization that appeared in literature can be broadly classified into two main categories: deterministic optimization algorithms and stochastic search algorithms.

The drawbacks of deterministic optimization algorithms are mainly in the following three aspects:

- requiring a good starting point,
- an unsatisfactory local minimum may be reached in many cases, and
- often requiring continuity and differentiability of the objective function.

The application of deterministic optimization can guarantee a local minimum like **Gradient Descent** algorithm or a global minimum like **Geometric Programming**.

Research efforts on stochastic search algorithms, especially **Simulated Annealing** and **Evolutionary Algorithm**, are been developed due to the ability and efficiency to find a satisfactory solution in a fast optimization convergence. However, this solution is often not the optimal neither the stochastic algorithm gives information about how good is the found solution.

Thereafter, we will describe the most implemented optimization algorithms that are:

- Gradient Descent,
- Geometric Programming,
- Simulated Annealing, and
- Evolutionary Algorithm.

The advantages and disadvantages of each one will be presented. And also, the state-of-the-art applications of the optimization algorithms will be highlighted.

3.2.2.1 Gradient Descent Optimization

The gradient descent search is an optimization algorithm of finding solutions at local minima that are known to be near the starting point of the optimization. Such an algorithm checks the slope of the cost function to take an optimization decision.

The gradient descent starts with the attribution of the cost function to be optimized. Also, the maximum number of iterations is specified the in first step of the algorithm. Then, the slope of the cost

function is calculated in the neighborhood of the starting point. If the slope is negative, it searches a new point in this direction in order to minimize the cost function. If the slope is positive, such a point is rejected. The algorithm continues to iterate until all specifications are satisfied or the total number of iterations exceeds the specified maximum number. Finally, it stops searching for the global optimum solution and simply reports a local minimum found as an optimum value [55].

Because of this simplicity, it is a very fast optimization algorithm. However, it has a serious drawback of easily becoming stuck at an undesirable local minimum. This case occurs when it is not sure how much near the starting point is to the optimum. The starting point choice increases in complexity with the cost function formulation and number of variables to be optimized. Thus, the odds of specifying such an initial condition approach zero ahead the complex real-world problems.

In 2008, J. Cong et al. [56] has unified a wide range of density smoothing techniques called global smoothing and presented a highly efficient method for computing the gradient of such smoothed densities used in several well-known analytical placement algorithm. This method reduces the complexity of the gradient computation used in analytical circuit placement optimization. In 2010, they [57] presented an analytical 3D placement method using a gradient projection method to solve the constrained optimization problem.

3.2.2.2 GEOMETRIC PROGRAMMING OPTIMIZATION

A geometric program (GP) is a type of mathematical optimization problem characterized by objective and constraint functions that have a special form. Recently developed solution methods can solve even large-scale GPs extremely efficient and reliable [58]. At the same time a number of practical problems, particularly in circuit design, have been found to be equivalent to (or well approximated by) GPs. Putting these two together, we get effective solutions for the practical problems.

A basic part of solving the GP is determining the problem feasibility. Such feasibility means to determine whether the design constraints are mutually consistent. Founding the feasible design space, the GP starts the research of the optimal point. If the problem is infeasible, there is certainly no optimal solution to the GP problem, since there is no point that satisfies all the constraints. To find the optimal point, a trade-off analysis of the constraints is applied to discover the effect on the optimal value of the problem [58]. This reflects the constraint variation idea that takes place in many practical problems. The analysis result is a point on the optimal trade-off design space. With the feasible region and the optimal trade-off of the design space, the GP rapidly converges to the optimal point by an objective criterion. Moreover, assuming the given criterion the trade-offs are immediately pointed-out. The given set of criteria indicates the optimal trade-off point for this formulation.

The basic approach in GP modeling is to attempt to express a practical problem, such as a circuit analysis or design problem, in GP format. One of the most disadvantages of GP is the required monomial and posynomial functions modeling the circuit performance, because such modeling requires a convex problem formulation and it is not often the case on circuit performance trade-offs. However, the bigger advantage is the great efficiency and the fast solution optimization. Moreover, the GP gives the feasibility of the design and the trade-offs to be paid in order to achieve the optimal.

In 2004, L. Vandenberghe [59] presented techniques for improving the accuracy of GP based analog circuit design optimization. He proposed a simple method to take the modeling error into account in GP optimization resulting in a robust design over the inherent errors in GP device models. In 2009, Y. Xu et al. [60] introduced a regular analog/RF IC using metal-mask configurability design methodology; named Optimization with Recourse of Analog Circuits including Layout Extraction (ORACLE). Such methodology is a combination of reuse and shared-use by formulating the synthesis problem as an optimization with recourse problem. Using a two-stage GP with recourse approach, ORACLE solved for both the globally optimal shared and application-specific variables.

3.2.2.3 SIMULATED ANNEALING OPTIMIZATION

The simulated annealing (SA) algorithm was introduced to overcome the limitations of the gradient descent optimization. The concept of simulated annealing optimization follows from the physical process of cooling of metal during manufacture for maximum strength and minimum defects.

In this algorithm, the metal heating process is implemented by picking a random number. And the metal cooling process, named annealing, is implemented by a gradient descent optimization. After a number of heating and annealing iterations, some cost function minima are found. Finally, the algorithm stops searching for the global optimum solution by simply reporting the minimum of local minima found as an optimum value. Thus, the global optimum is always found if we dispose of the enough number of iteration, and often tending to infinity.

Unlike gradient descent algorithms, simulated annealing provides a statistical hill climbing capability, and thus avoids, with high probability, being trapped in an undesired local minima. However, the hill climbing process needs to be controlled to save simulation time. In the simulation annealing technique, two control parameters on the hill climbing process are the slope of the cost function and the temperature ($temp$) coefficient [55]:

$$rand(0,1) < \exp\left(\frac{\Delta cost}{temp}\right), \tag{3.1}$$

where $\Delta cost$ is the variation of the cost function, always less than zero. So that, the right side of Equation (3.1) is always less than one. Thus, the condition of $temp$ being high indicates that the probability of hill climbing is high, and reducing the temperature of annealing the probability becomes lower. The second condition to a less likely hill climbing is a high $\Delta cost$, as a result the $\Delta cost$ reduction works as a gradient descent formulation.

In 1990, G. Gielen et al. [61] proposed an EDA tool based on a generalized formulation of the analog design problem. The EDA tool sizes all elements to satisfy the performance constraints, thereby optimizing a user-defined design objective. The global optimization method being applied on the analytic circuit models is SA. In 2001, R. Gupta et al. [62] had designed a 3-V 85-mW balanced fully integrated Class-C power amplifier (PA). The PA was optimized using a simulated-annealing-based custom CAD tool, integrated, and tested. The CAD tool described in [62] was designed to find such an optimum matching network for integrated PAs.

3.2.2.4 EVOLUTIONARY ALGORITHMS OPTIMIZATION

Evolutionary Algorithms (EA) are stochastic search and heuristic optimization methods whose mechanisms are analogous to biological evolution [63]. EAs differ from more traditional optimization techniques in that they involve a search from a population of solutions, not from a single point. Each EA's iteration involves a competitive selection that weeds out poor solutions. The solutions with high *fitness* are *recombined* with other solutions by swapping parts of a solution with another. Solutions are also *mutated* by making a small change to a single element of the solution. Recombination and mutation are used to generate new solutions that are biased towards regions of the space for which good solutions have already been seen.

A generic implementation of a single-population EA is presented in Algorithm 1. First, a random population is initialized. For a single population and a given generation, the objective fitness function is evaluated. While the desired fitness is not found neither the maximum number of generation is not achieved, EA performs the population evolution by natural genetic processes. These genetic processes are: selection, recombination, mutation, migration, locality and neighborhood. The genetic processes aim to improve the average fitness of the population by giving individuals of a higher fitness a higher probability to be recombined into the next generation. The mutation implements the stochastic nature of the EA optimization. The migration, the locality and the neighborhood are genetic processes applied among different populations in a multi-population EA context.

Algorithm 1 EA Pseudo-code implementation

```
initialize(population)
generation = 0
optimal = evaluate(population)
while (optimal ∉ optimalCriteria) and (generation ≠ maxGeneration) do
  naturalSelection(population,optimal)
  geneticRecombination(population)
  geneticMutation(population)
  generation += 1
  optimal = evaluate(population)
end while
return population
```

A single-population EA is powerful and performs well on a wide variety of problems. However, better results can be obtained by introducing multiple subpopulations. Every subpopulation evolves over a few generations isolated (like the single-population EA) before one or more individuals are exchanged between the subpopulation. The multi-population EA models the evolution of a species in a way more similar to nature than the single-population EA.

Several different types of evolutionary search methods were developed independently and the same advantages and disadvantages. These include:

- Genetic Programming, which evolve programs,
- Evolutionary Programming (EP), which focuses on optimizing continuous functions without recombination,
- Evolutionary Strategies (ES), which focuses on optimizing continuous functions with recombination, and
- Genetic Algorithms (GA), which focuses on optimizing general combinatorial problems.

In [64], GA was implemented to design analog circuits, where the classic GA was modified with respect to population size and the generation of valid circuits only. Also, simple modifications were introduced which reduce the computations required. In [54], a new EP algorithm, called competitive co-evolutionary differential evolution (CODE), was proposed to design analog ICs with practical user-defined specifications. The base of CODE is the combination of HSPICE and MATLAB. The CODE tool links circuit performances, evaluated through electrical HSPICE simulation, to the optimization system in the MATLAB environment, for a given circuit topology. In [65], an accurate method to determine the device sizes in an analog IC based on GAs was presented. To evaluate algorithm, a two-stage Operational Transconductance Amplifier (OTA) was designed in CMOS 0.18 μm process exemplifying both time and frequency domains optimization.

3.2.3 Performance Estimation Tools

An accurate synthesis begins with an accurate analysis. With the continuous downscaling of CMOS technology, to obtain a precise model and to control its parameters has become a highly challenging task. Most of options for circuit performance estimation consist of building a model of a circuit performance metric and then using this model to estimate the performance. However, one of the largest problems is to find systematic methods to create accurate behavioral or macro models. Moreover, such model shall be adapted to the needed tools to automate the synthesis process.

An accurate solution can be pointed out by a circuit simulation, which even today remains the cornerstone of many circuit designs. In general purpose, the SPICE-based simulator solves the nonlinear

system obtained from the netlist of the circuit. Such a circuit in turn is composed by modeled devices given by the process design kit. So, the problem turns to how good and available are the device models. AMS/RF device models are not often available in standard CMOS process. Moreover, the largest disadvantage of the circuit simulation is the time needed to complete the analysis. The simulation time grows exponentially up with the accuracy needs, and thus the accurate and fast analysis is not even feasible with circuit simulation.

Thereafter, we describe the solutions at the state-of-the-art for behavioral modeling and electric simulation. Foremost, these solutions do not solve the trade-off among accuracy, analysis time, and synthesis automation. They are able to point solutions to some problems adapting the trade-off to some application need.

3.2.3.1 Performance Behavioral Modeling

In order to systematically create analog behavioral models, the state-of-the-art of behavioral modeling can be distinguished [14]:

- **Fitting or Regression Approach** which is parameterized mathematical model and values parameters selection to best approximate the simulated circuit behavioral by minimizing the error between the simulation and model results. The disadvantages are finding a good starting meta-model and fit it with a small number of simulation points. However, once the model is fitted the performance can be directly estimated.

- **Constructive Approach** which is the physical inherent model by underlying circuit description. The advantage is the higher guarantee of the model fitting, but the model building includes not at all trivial simplifications from the electrical-simulated model.

- **Model Order Reduction Approach** which is a mathematical technique to generate a simplified model for a given circuit by direct analysis and manipulation of its detailed description by the eigenvalue and the eigenvector of the system matrix. Such models produce as output a similar state-space model, but with a state vector of lower dimensionality. These reduced-order models are very advantageous because they simulate much more efficiently, while approximating the exact response, with less system equations and computational efforts. However, reducing the system order is not often feasible, evident, or computationally less expensive than a direct simulation. The reason is that AMS/RF circuits are usually nonlinear, time-variant, and/or distributed parameters system.

In order to achieve a trade-off between accuracy and CPU time, efficient approaches based on meta-modeling techniques have been proposed. The meta-model can be distinguished:

- **Polynomial Function** which is a mathematical polynomial representation often obtained from a model order reduction in the form [66]

$$\Phi = \sum_{i=0}^{N} A_i \psi_i. \tag{3.2}$$

  In a single-objective optimization problem, x is one parameter and Φ is one performance. However, a multi-objective optimization problem will require a matrix representation. Such a model has an advantage to analyze the performance with simplicity, it is easy to hand design or EDA implementation. The disadvantage is the low accuracy for a small order model. The model is only accurate when process variations are sufficiently small which is hardly the case in nanometer technologies.

- **Posynomial Function** which is a mathematical posynomial representation often obtained from a model fitting or regression in the form [60]

$$\Phi = \sum_{k=1}^{t} c_k \prod_{i=1}^{N} \psi_i^{\alpha_{i,k}}, \tag{3.3}$$

  where $c_k \in \mathbb{Z}^{+*}$ and $\alpha_{i,k} \in \mathbb{R}$. Using a posynomial function, the advantageous GP optimization can be easily implemented and an exact solution achieved. However, the system can only be modeled by a posynomial function if it has a convex formulation, which is not often the case.

- **Symbolic Model** which is the generation of interpretable mathematical expressions from a set of templates, that relate the circuit performances to the design variables. Such templates are automatically chosen by the optimization tool from polynomial, posynomial and simple non-linear functions. Thus, its main weakness is that it is limited to linear and weakly nonlinear circuits [67]. The model strength comes from the automatically generation and fitting of the model in a single EA optimization tool.

- **Kriging Model** which is a framework to represent a circuit performance as a stochastic process $\Phi(\psi)$ that is cast into a regression model as

$$\Phi(\psi) = Z(\psi) + \vec{\beta}^T \vec{f}(\psi) \text{ [16]}, \tag{3.4}$$

  where $\vec{f}(\psi)$ is a vector of predefined regression functions and $\vec{\beta}^T$ is the vector of unknown regression coefficients. $Z(\psi)$ is a random process and used to capture the systematic departure of the performance from the global regression portion. The Kriging model enables robust regression of global trends of complex mapping between design parameters and resulting performances. Moreover, such a model provides an uncertainty level for each prediction in the form of mean-square-error [16]. Thus, model and model error are optimized simultaneously in a regression approach. The disadvantage is the mathematical complexity of the stochastic formulation of the model.

3.2.3.2 Performance Electrical Simulation

Advanced technologies and complex IC designs have turned AMS/ RF simulations in some days or even weeks of engineering tasks. As a result, the need for a fast, accurate and optimized simulation solution that incorporates analog, RF, and AMS needs has become more pressing and widespread. This is especially true in cutting-edge low-power, low-noise or high-frequency designs at 90nm and below. Thus, the resulting simulation performance bottleneck has increasingly strained IC designers' productivity and their products' time-to-market schedule [68].

Ever since the emergence of commercial simulation tools in the early 1980s, simulation performance has been the driving force for major simulation technology developments. Currently, the simulation tools are developed to specialized applications providing the fast and accurate results for a given circuit design. These simulators can be distinguished:

- **SPICE simulator** is based on a precise brute-force transient and small-signal analysis. The advantage is to be a general accurate circuit simulation, but it costs an exponential-increasing simulation time with the circuit complexity increase.

- **Steady-State simulator** is a simulation technology that takes advantage of the known periodic steady-state nature of time-variant designs, like RF and discrete-time circuits. Such simulator provides significant speed-up compared to SPICE analysis ensuring the required accuracy. However, the periodic circuit frequency shall be known a priori and the difference of frequency of circuit clocks shall not be small. The steady-state simulation time tends to the SPICE simulation, and even bigger, if the difference of frequency of circuit clocks decreases.

- **FastSPICE simulator** explores the hierarchical structure of full-chip design and takes a divide-and-conquer approach to ensure full-chip verification can be done in a reasonable timeframe. The advantageous speed-up SPICE simulation time is paid with a mandatory top-down approach. Moreover, the required analog small-signal analysis is not available which leaves a lack for a faster SPICE simulator.

- **AMS simulator** encourages a top-down design methodology that mixes much simpler behavioral models with full accuracy transistor blocks to speed-up transient analyses during both design and verification. The co-simulation approach takes advantage of different approaches to obtain fast and accurate result for each circuit block. However, combining the results in a full-chip simulation requires a common language to inter-communicate simulators. This disadvantage limits the AMS solutions to only few simulator tools, often not the faster neither more accurate ones.

The most advanced simulation engines have different approaches and achieve better results in specific applications. These features are presented below with the informations extracted from the white papers of the more important companies: Cadence, Mentor, Synopsys, and Magma. Although we are able to point the best engine for a specific application, the decisions are not only based on the technical capability. In most cases, the engine choice is driven by license cost, designers' experience, PDK model availability, and integration with other tools. Therefore, this thesis will present results based on Cadence Virtuoso and Mentor Graphics ELDO, because of the constraints of the integration with other tools and the PDK ageing model availability.

CADENCE VIRTUOSO ENGINE [68] using the simulator Spectre provides fast, accurate SPICE-level simulation for tough AMS/RF circuits. It is tightly integrated with the Virtuoso custom design platform and provides detailed transistor-level analysis in multiple domains. Its architecture allows low memory consumption and high-capacity analysis. The simulator features and benefits are presented below:

- Provides high-performance, high-capacity SPICE-level analog and RF simulation with out-of-the-box tuning for accuracy and convergence;

- Performs application-specific analysis of RF performance parameters (spectral response, gain compression, inter-modulation distortion, impedance matching, stability, isolation).

- Includes advanced statistical analysis (Smart, MonteCarlo, DCmatch) to help design companies improving the manufacturability and yield of ICs at advanced process nodes without sacrificing time-to-market;

- Ensures higher design quality using silicon-accurate, foundry-certified device models shared within Virtuoso Multi-Mode Simulation.

Cadence Virtuoso Multi-Mode [69] simulator combines industry-leading simulation engines like AMS Designer, Spectre, UltraSim, and device models. The objective is to deliver a complete design and verification solution. Such simulation engine meets the changing simulation needs of designers as they progress through the design cycle. The coverage solution is from architecture exploration to analog and RF block-level development and to final analog and mixed- signal full-chip verification.

Virtuoso Device Modeling (BSIMProPlus/RelProPlus) [70] automates HCI and NBTI stress and data collection through simultaneous control of various hardware measurements. The tool can stress multiple devices at the same time. After, the extracted AgeMos model makes it possible to have degradation models based on the age calculated after the fresh simulation. The Virtuoso UltraSim Full-chip Simulator can use these models directly. To run aged simulation with other commercial simulators, such as the Virtuoso Spectre Circuit Simulator, using a Cadence technology called RelXpert.

MENTOR GRAPHICS ELDO ENGINE [71, 72] is the core component of a comprehensive suite of AMS/RF tools. Moreover, Eldo offers a unique partitioning scheme allowing the use of different algorithms on differing portions of design. It allows the user a flexible control of simulation accuracy using a wide range of device model libraries, and gives a high accuracy yield in combination with high speed and high performance.

The following is a list of the major product features of Eldo:

- Eldo is the core technology allowing addressing RF simulation (Eldo RF) and mixed-signal (Questa ADMS, ADMS-ADiT);
- Simulation of very large circuits (up to around 300,000 transistors) in time and frequency domains;
- Advanced options such as pole-zero, enhanced Monte Carlo, DC mismatch, circuit optimization, statistical analysis (Design of Experiments), and S and Z-domain generalized transfer functions;
- Unique transient noise algorithm and reliability simulation (Eldo UDRM);
- Flexible netlist support, behavioral modeling with Verilog-A, and extensive device model libraries (MOS, bipolar, MESFET transistor models such as the BSIM3v3.x, BSIM 4, MM11, Mextram, HICUM, and PSP, and TSMC Model Interface);
- Integration into Cadence's Analog Artist environment (Artist Link).

SYNOPSYS HSPICE ENGINE [73] uses a precision parallel technology which is a multi-core transient simulation extension to HSPICE for both pre- and post-layout of complex analog circuits such as PLLs, ADCs, DACs, SERDES, and other full mixed-signal circuits. HSPICE addresses the traditional bottleneck in accelerating SPICE on multi-core CPUs with new algorithms that enable a larger percentage of the simulation to be parallelized, with no compromise in golden HSPICE accuracy.

Additionally, efficient memory management allows simulation of post-layout circuits larger than 10 million elements. The HSPICE solution includes enhanced convergence algorithms, advanced analog analysis features and foundry-qualified support for process design kits (PDKs) that extend HSPICE gold-standard accuracy to the verification of complex analog and mixed-signal circuits. With HSPICE 2010, design teams can accelerate verification of their analog circuits across process variation corners, and minimize the risk of missing project timelines [73]. The Synopsys HSPICE features can be distinguished:

- Process and Interconnect Variation including models both device and interconnect variation;
- Variation Block using powerful and flexible mechanism for defining process variation effects;
- AC & DC Match using efficient statistical simulation for local parameter mismatch effects;
- Smart Monte Carlo, the all-purpose statistical simulation that runs several times faster than traditional Monte Carlo techniques;
- MOSRA device reliability analysis simulating HCI and NBTI device ageing effects;
- PLL and VCO simulation including the most accurate and fastest RF simulator.

MAGMA TITAN ENGINE [74] is well-known solution for leading the pack with regard to its digital IC design, analysis, implementation, and verification solutions. It's true to say, however, that (until now) most engineers and industry observers do not tend to think of Magma in the context of analog, full-custom, and/or mixed-signal designs.

For the last few years, Magma has been developing technology and expertise. The result is Titan, a truly unified, automated, full-chip, mixed-signal design solution. There is no question about the power and value of this technology in the minds of users who have had the opportunity to work with Titan.

The Magma Titan FineSim features can be distinguished:

- Interface for full-chip circuit simulation, offering SPICE-level accuracy for the analog portions of the design and Fast SPICE-level accuracy for the digital side of the design;
- Fast Monte Carlo is a revolutionary new alternative to traditional Monte Carlo analysis, making it possible to achieve much more accurate statistical analysis in a fraction of the time required by traditional methods;
- Especial steady-state simulation for RF circuits and ageing (HCI, NBTI) analysis are not available, or the specific tools are not found in this state-of-the-art research.

3.3 VARIABILITY-AWARE DESIGN METHODOLOGIES

In advanced technologies, the increase of process variation significantly impacts circuit yield. The impact of variability on the design of large ICs is becoming especially severe. Thus, the classical EDA methodology and tool is not able to solve the design problems under 100 nm technologies nodes.

In general, the first change in the classical design is the performance estimator. Most of the simulator engines have variability simulation included. The approaches can be distinguished between Corner-Based and Monte Carlos simulation. The biggest disadvantage of the simulator-based in performance estimation is the weak feedback for the optimizer tool. Thus, the optimizer shall be based on an optimal search without the information of the statistical characteristics of the performance. Moreover, the computational design cost can be too much expensive.

The second possible change in the classical design is the evaluator judge and the feedback of the optimizer. The evaluator shall judge the performance mean and its variance or standard-deviation as well. The feedback advantage is obtained with the model building cost, which is higher including the information of the statistical characteristics of the performance.

At the state-of-the-art, a single EDA consensus joining variability-aware design and commercial CADs was not found. Since 2000, S. Nassif pointed out the design for variability needs [15]. Nowadays, G. Yu and P. Li proposed that the analog designs must be optimized not only for nominal performance but also for robustness in order to maintain a reasonable yield [16]. H. Onodera goes further proposing that the EDA methodologies and tools shall tolerate, mitigate, or even exploit the variability by proper design techniques [17]. In 2009, V. Wang et al. introduced a simplified model bridging the gap between existing statistical methods and circuit design [18]. In next subsections, we address the details of each state-of-the-art answer to EDA lacks in the increasing variability of the advanced technologies.

3.3.1 WORST CORNER-BASED

In traditional variability-aware design methodologies, the variability has been considered in the ICs process by guard-banding, using a corner-based approach. Figure 3.2 illustrates all performance (Φ) samples in function of the parameters vector ($\vec{\psi}$). We see the guard-band linking the important corners of Φ. From all C_i corners, there is a corner named *Typical* situated at the center of the guard-band and some worst corner C_w, where Φ drops out of the specification. The corner-based method consists in identifying these important corners and designing the circuit to achieve the specification, even in the worst case.

The worst case corner models are generated by offsetting the selected parameters by a fixed number of its standard deviation ($\sigma\psi$). This selection guarantees a minimum circuit yield. The *Typical* corner is generated from the measured data on a single golden wafer of the centerline process. Other four corners often selected are the two analog and the two digital worst case scenarios. Thus, the worst-case corner models provide designers a capability to predict the assertions of the typical design and the pessimistic scenarios.

The state-of-the-art worst-case methods take the statistical distribution of the process parameters into consideration and evaluate the worst-case scenario based on the probability density function (Figure 3.2).

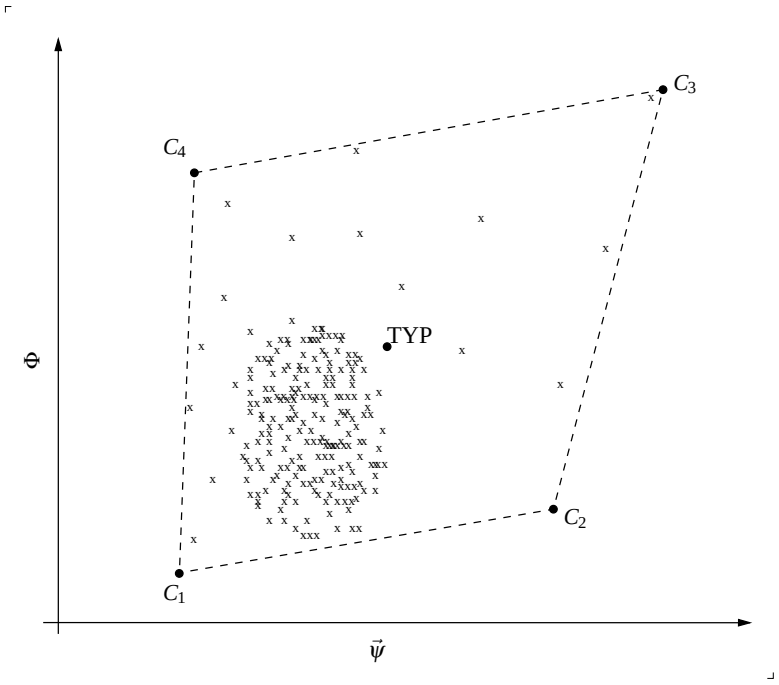


Figure 3.2: Corner-based variability analysis: illustration.

In [75], L. Silva et al. proposed a set of efficient branch-and-bound-based techniques for automating the computation of the exact worst-timing corners of combinational and sequential circuit. This research work pointed out difficulties to find the worst corner in a specific application. However, the corner-based approach easily leads to overly pessimistic worst-case design which will cost power and area consumption.

3.3.2 Monte Carlo Simulation

The Monte Carlo (MC) simulation avoids the pessimistic worst-case, performing a direct yield estimation of a given circuit. MC is an exhaustive method that never fails, and in some cases may be the only available option. It consists of several trials, each of which is a full-scale circuit simulation. On every simulation, each process parameter is sampled from its distribution, and then the circuit performance is estimated. The procedure is repeated over thousands of trials, and the performance distribution is derived from the statistics of the estimated performance among the trials. With a sufficient large number of trials, the performance distribution can be predicted with a measurable confidence.

MC-based methodologies are inherently accurate as they involve only measure-fitted models. This advantage leads to an accurate optimum for a given circuit yield. Based on a fully random choice of the samples, the number of employed samples is crucial to achieve a representative statistic. However, the simulation time directly depends on the number of samples, leading to a loss of efficiency for a large number of samples.

In order to reduce the number of MC runs, the state-of-the-art has proposed variance reduction techniques [6]. The idea is reducing intentionally the accuracy of the method speeding up the MC simulation. The most important methods derived from this idea are presented [76].

- **Latin Hypercube Sampling (LHS)** method samples each dimension (process parameter) by partitioning its domain into equi-probable sub-ranges. Hence it improves the uniformity of the samples in one-dimensional projection.

- **Quasi Monte Carlo (QMC)** utilizes low-discrepancy sequences to provide uniformity in higher

than 1-D projections. However, the convergence rate of the QMC method is dependent on the problem dimension, and it is found to be only asymptotically superior to MC.

Another disadvantage on a MC-based method is the weak feedback for an optimum search. Therefore, the design optimization easily leads to exhaustive random search increasing the design-cycle cost which consists of several analysis and redesign iterations.

3.3.3 Behavioral Modeling

The circuit performance metrics are generally non-linear functions. The performance's probability density function (PDF) cannot easily be computed. The aim of the behavioral modeling approach is to estimate an interval that encloses the circuit-performance variability. Thus, the circuit-performance distribution is estimated avoiding the MC exhaustive simulation. Moreover, this approach has a better optimization feedback since the estimation accuracy can be imposed as a trade-off.

Most noteworthy variability analyses are the response surface methods and models (RSM). RSMs first generate a model of the circuit performances as a function of the process parameters around the nominal design point. Then they use this model to estimate the performance's PDF. In this way, the state-of-the-art of the variability-aware design methodologies proposes different meta-model to estimate the mean and the standard deviation of the circuit performance distribution. Thus, the optimization problem only doubles the number of the objective equations giving to the desired performance a statistical characteristic.

The crucial issue in RSM lies in creating an accurate model with minimal computational cost. The required minimum number of simulations can turn the behavioral modeling in a prohibitive approach, and it is one of the bigger disadvantages. A proposed solution reducing the minimum number of simulations is the design of experiments (DoE). According to the type of meta-model, the DoE theory can indicate an efficient sampling strategy [77].

In [16], G. Yu and P. Li demonstrated how an efficient iterative search based optimization can be developed for extracting Pareto analog performance models in nominal case. The proposed methodology uses Kriging model advantage representing the parameter as a stochastic process. Thus, they applied an iterative search based optimization to efficiently seek optimal performance trade-offs under yield constraints in a high-dimensional design parameter and process variation spaces. Their experiments showed that the proposed method can efficiently extract desired optimal performance trade-offs by avoiding excessive time consuming MC simulations and such a method also leads to well-controlled accuracy.

In [60], Y. Xu et al. proposed a robust GP optimization capturing the variability by an ellipsoidal uncertainty expression. Formulating the optimization with ellipsoidal uncertainty, the statistical distribution information of the circuit variability can be included. They demonstrate that competitive performance can be achieved with guaranteed yield.

In [18], V. Wang et al. proposed an accurate posynomial model for circuit variability that provides insight into the sources of variability in a circuit design point of view. Modeling random process variability onto designer-controlled variables (size and bias), it will allow a variability-aware GP optimization. The work trends are using this model for circuit optimization problems and noise analysis for more complex analog blocks. Such a model has a great potential to find more methods of evaluating circuit performance variability without the use of computationally expensive MC methods.

3.4 Reliability-Aware Design Methodologies

Integrated circuit reliability simulation is not a new concept and a number of reliability models and simulation methodologies have been developed during the past decades [78, 79, 13]. Most state-of-the-art reliability simulation methods try to emulate the degradation process of aged devices based on the physical failure mechanisms and are based in Berkeley Reliability Tools (BERT) concept. As the technology advances under 90nm, the sources of ageing increase and the circuit lifetime shrinks. Since

2003 [79], reliability-aware design methodologies have been proposed. Most of studies are interested in model [19] and simulate the ageing stress [20, 21] and the circuit lifetime [22]. Thus, they proposed a new design flow (Figure 3.3) including the ageing analysis into the classic design (Figure 3.1).

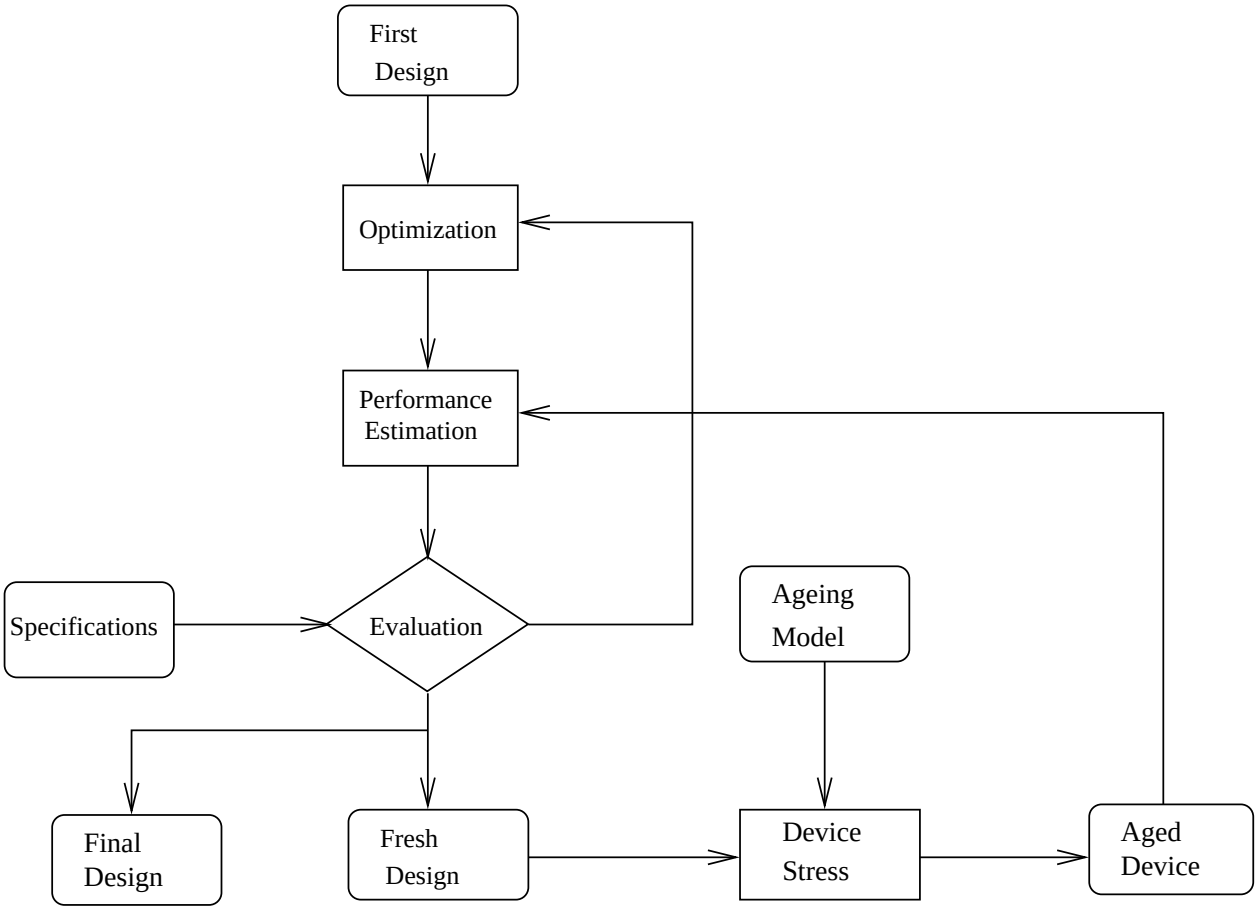


Figure 3.3: The state-of-the-art of design in reliability methodology: ageing stress flow.

In the Figure 3.3, we observe that after few optimization cycles a fresh design is found. Using the ageing model from the PDK, the devices inside the design are stressed leading to aged devices. The aged design has its performance estimated and evaluated, facing the specifications in an external feedback. If the aged design achieves the desired performance, we find the final design. However, if the aged design does not meet the specification, the design shall be optimized. The optimization step has not enough information to increase the circuit reliability, so that it shall implement some exhaustive search and probably do not converge in a reasonable time.

The reliability-aware design flow presented in Figure 3.3 has a complex feedback and it is often a tricky algorithm for an EDA tool implementation. The problem is always when the optimization stops: optimal design or reliable design? What's more, the trade-off between optimal and reliable is not clear. Therefore, the designer cannot decide which is preferred.

Reformulating the reliability-aware design flow at the variability-aware knowledge, the proposed improvements change the performance estimator. In order to take reliability and variability into account, the estimator should evaluate the fresh performance and the performance variation due to variability and ageing. This solution shares the characteristics of a simulation approach.

The advantage of this methodology is the accuracy of the results if the model is available and the stress conditions are known. However, the imprecision in describing the stress condition leads to wrong conclusions or overdesign. Moreover, the simulation approach leads to a random optimum search ahead

the weak feedback inside the not clear trade-offs. Such reliability simulation approach can be distinguished: Nominal Reliability Analysis, and Variability-Aware Reliability Analysis [80]. Both will be detailed in next subsections.

3.4.1 Nominal Reliability Analysis

The analog reliability analysis is a method to calculate the ageing of a circuit with fixed design parameters, like an IC process corner. Most ageing degradations have a dependence on the transistor bias (presented in Section 2.4). First proposed in 1993 [78], Berkeley Reliability Tools (BERT) simulates changes in a circuit performance due to NBTI and HCI. It also evaluates the circuit failure probability due to TDDB and EM. This tool was developed with the objective to provide circuit reliability simulation with little overhead to the design process.

The first attempt to evaluate the ageing degradation is analyzing the circuit ageing due to its DC operating point. The advantage is the fast simulation convergence; the disadvantage is the inherent small accuracy. Furthermore, the stress conditions depend on the circuit operating cycle and so vary with the time. In order to obtain accurate information about the time-varying stress at every circuit node, a transient or a steady-state simulation is required. The most accepted ageing analysis is based in BERT principles and is detailed in Algorithm 2. This algorithm has a linear complexity O(n) with respect to the number of transistors in the circuit. Currently, the ageing analysis is implemented in some commercial solutions [80] and the ageing model is available in few PDKs.

Algorithm 2 Ageing degradation analysis.

```
Require: freshNetlist, stressPeriod
  nodeWaveforms = transientSimulation(freshNetlist,stressPeriod)
  agedNetlist = freshNetlist
  for all transistor ∈ freshNetlist do
    agedTransistor = ageingAnalysis(transistor,nodeWaveforms)
    agedNetlist = update(agedNetlist,agedTransistor)
  end for
  return agedNetlist
```

According to the phenomenon, the analysis runs a specific model characterizing the ageing degradation into the transistor parameters. There are a lot of different models described in the state-of-the-art. Some of them are implemented for specific applications in dedicated simulators. They often uses simplified equations and need a specific model fitting. Others are described in a behavioral language including the transistor model and the ageing degradation description.

What's more, the commercial tools have their own models and only few foundries offers the ageing characterization as an option in the process design kit (PDK) description. Therefore, there is no standard ageing model, and they are not often available in the desired tool. The common characteristics among the models are summarized below for each ageing phenomenon.

3.4.1.1 HCI Simulation

The model estimates the NMOS or PMOS HCI ageing by the *Age* parameter defined as:

- NMOS

$$Age = \int_0^T \frac{I_{DS}}{W H_n} \left(\frac{I_{sub}}{I_{DS}} \right)^{m_n} dt, \tag{3.5}$$

- PMOS

$$Age = \int_0^T \frac{1}{H_p} \left(\frac{I_g}{W} \right)^{m_p} dt. \tag{3.6}$$

The T is the period of time of stress and W is the device width. I_{sub}, I_{DS} and I_g are the substrate, the drain and the gate currents, respectively. H and m are technology dependent parameters. Afterwards, the amount of degradation suffered by the MOSFETs is monitored in each model parameter (P) (for example the threshold voltage, the channel mobility, etc.) and the degradation dynamics can be written as

$$\Delta P(t) = K(Age \cdot t)^n. \tag{3.7}$$

So, such model equations use computationally simple and BSIM compatible reliability models to predict the degradation. Experimental evidence confirms that device degradation may be represented as a function of *Age* over a wide range of stress conditions and for channel lengths down to the submicron region [78].

3.4.1.2 NBTI Simulation

The behavior of the NBTI degradation is monitored by the threshold voltage shift (ΔV_{th}), represented in Figure 3.4. The ΔV_{th} is induced by both permanent damage and recoverable degradation. As soon as the stress is stopped, the recoverable degradation decreases in time. This degradation is activated by temperature (T) and a high negative field and hence is dependent on the bias (V_{GS}).

The permanent damage is modeled using the equation

$$\Delta V_{thp} = A_p \exp(\gamma_p V_{GS}) \exp\left(-\frac{E_{ap}}{kT}\right) t_{stress}^n. \tag{3.8}$$

The recoverable degradation is modeled by

$$\Delta V_{thr} = A_r \exp(\gamma_r V_{GS}) \exp\left(-\frac{E_{ar}}{kT}\right) \ln\left(1 + \frac{t_{stress}^n}{\tau_r}\right). \tag{3.9}$$

The recovered degradation is modeled by

$$\Delta V_{thrr} = A_{rr} \exp(\gamma_{rr} V_{GSrecover}) \exp\left(-\frac{E_{arr}}{kT}\right) \ln\left(1 + \frac{t_{recover}/t_{stress}}{\tau_{rr}}\right) \Delta V_{thr}. \tag{3.10}$$

In all ΔV_{th} equations (3.8), (3.9) and (3.10) [81], A is the technology dependent pre-factor, γ is the voltage accelerator factor, E_a is the activation energy in each case, τ is the physical time constant, k is Boltzmann's constant and T is temperature. Finally, the total threshold voltage shift is monitored by the equation

$$\Delta V_{th} = \Delta V_{thp} + \Delta V_{thr} - \Delta V_{thrr}. \tag{3.11}$$

This model is simple and BSIM compatible, but, to avoid any ΔV_{th} underestimation, it is important to use the on-the-fly stress and measurement technique [41] to fit the NBTI degradation model.

3.4.1.3 TDDB Simulation

Most of the simulators do not simulate post-breakdown transistor electrical behavior. They often evaluate the transistor gate electric field against the maximum technology allowed field, treating it as a hard breakdown (HBD) rather than a soft breakdown (SBD). HBD assumes a circuit catastrophic failure. In reality, the post-SBD circuit continues working with a performance loss. Thus, the HBD is of no interest in this work, and the design for reliability shall reduce the HBD probability. And often, the design-kit has the relied information about the mean time to failure and guidelines to reduce the HBD probability.

The statistics of SBD are described by the Weibull distribution (Eq. (2.1)). However, this equation is used only to estimate the first SBD event. To estimate the post-SBD, the electrical model presented in Figure 3.5 [11] is usually employed, where the SB_d, SB_s are the SBD low conductive paths modeled by a $\approx$10 kΩ resistance [82]. Despite the continuous breakdown characteristics, the second event cannot be easily described and it is often considered as a HBD.

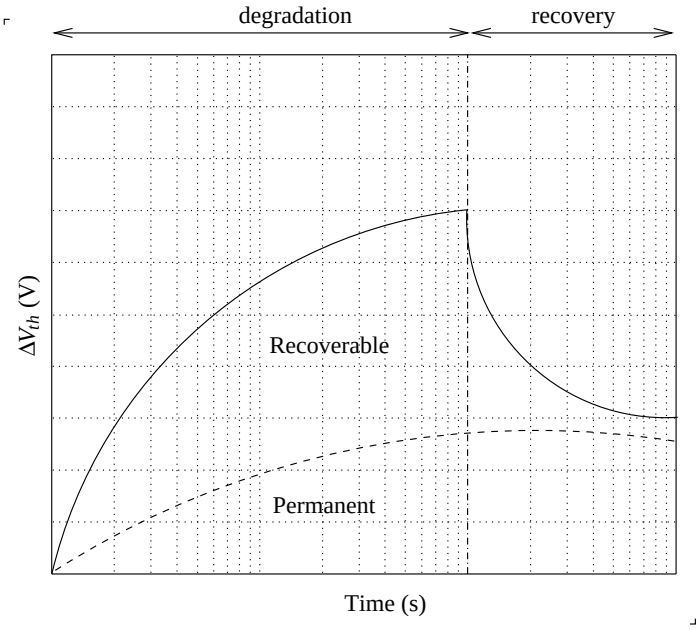


Figure 3.4: Behavior monitor of the NBTI permanent and recoverable damage: illustration [81].

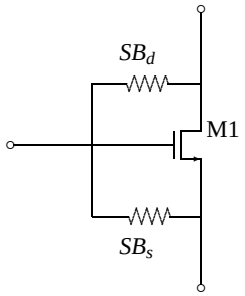


Figure 3.5: Behavior of the post-SBD degradation: NMOS illustration (PMOS is the analogous).

3.4.1.4 EM Simulation

The EM is often checked at the circuit layout level, because the EM involves all metal line connections and their current densities. In fact, there is no early EM estimation and in most cases it does not represent a constraint. However, RF circuits are particular cases because they often use inductors and capacitors as a tuned circuit, and in most cases they have larger currents passing through the inductors. These passive components are designed and software optimized, but the inherent low quality factor becomes worse without a robust design, because of the EM reliability degradation [83].

Many layout techniques may be used to reduce the EM, but the post-EM behavior is not often treated. In order to estimate the post-EM performance, it is suggested to model it as a quality factor reduction by increasing the inductor series resistance (R_s), see Figure 3.6. This idea does not increase the simulation effort, as the passive components optimizers return a lumped model or an equivalent S-parameters model which include the quality factor characteristic.

3.4.2 Variability-Aware Reliability Analysis

In fact, the ageing degradation is not totally independent of the IC process variability. Before ageing, the circuit suffers from the process variability that is a statistic variation. Then, the stress environment and

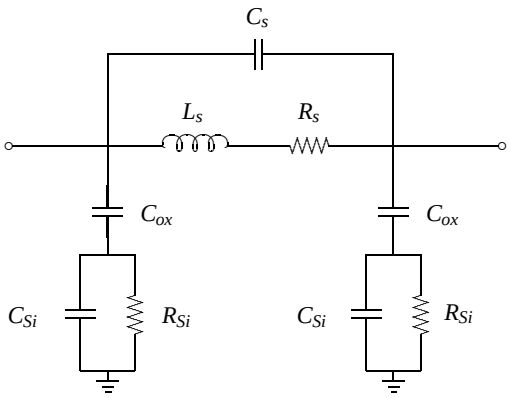


Figure 3.6: Inductor lumped model: illustration.

the stress time are the agents of a deterministic variation. Although both are combined in this physical happening order, the nominal reliability analysis does not take such an order into account.

In order to fill the nominal reliability analysis lack, the variability simulation, like MC simulation, can be run after the degradation with the aged netlist. The hypothesis behind such argument is that statistic and deterministic variations are naturally independent phenomena. Accurate results are achieved if the ageing degradation is small enough, and that is the case in 90 nm and 65 nm CMOS technologies aiming a lifetime less than 50 years. The huge advantage of this method is combining PDK accurate models and commercial tools in a single EDA methodology. The disadvantage is that the hypothesis cannot be true, moreover under 32 nm CMOS.

The solution is to break the nominal reliability analysis paradigm and propose a variability-aware reliability analysis. However, there is no commercial tool implementing this solution. In fact, random-sampling techniques, like MC simulation, are accurate, but very CPU intensive. Then implementing the ageing degradation, the EDA algorithm will have a complexity bigger than $O(n^2)$, which can rapidly achieve a not feasible number of simulation points.

One solution for this problem is proposed in 2010 by E. Maricau and G. Gielen. In [80], they demonstrated a deterministic, variability-aware reliability modeling and simulation method. In their work, a DoE with a quasi-linear complexity algorithm is used to build an RSM of the time-dependent circuit behavior. The simulation speed improvement is up to several orders of magnitude with good simulation accuracy. Another solution is proposed in 2010 by X. Pan and H. Graeb. In [84], they proposed an efficient method to predict the lifetime yield of analog circuits considering the joint effects of manufacturing process variations and parameter lifetime degradations. However, the analysis limitations are:

- a simplified ageing degradation model, far away from the commercial models;
- the model fitting availability for the commercial PDKs;
- a dedicated simulation until disconnected to the commercial tools.

3.4.3 Conclusion

Regarding the state-of-the-art, there is no most accepted solution. The reliability-aware remains in the analysis only and mostly simulation based approach. The clear disadvantage is the weak feedback to the optimizer algorithm. The reason is that the ageing degradation models are under research and the state-of-the-art efforts are concentrated in accurate modeling and simulating.

This scenario leaves an opportunity to innovate proposing a reliability-aware design methodology. The research needs resides in increasing the optimization feedback by design equations to estimate the

ageing degradation in early stages. These new equations should be integrated in a multi-objective optimization algorithm. Similarly, it is done in the variability-aware design methodology as proposed in [16, 60, 18]. The huge challenge is proposing an accurate design method to estimate the ageing degradation in early stages and use this information to find the optimum design space. The proposed method shall have in mind not only the advanced CMOS technologies, but also the future IC technologies when circuit reliability even worse is expected.

CHAPTER 4

RELIABLE RF FRONT-END: BOTTOM-UP DESIGN

In this chapter, we will present three applied examples of reliable-circuit design. Each example is one of the building blocks described in the study case design in order to propose a reliable RF front-end. Such building blocks will be designed in transistor level using CMOS 65nm technology.

The first block of an RF front-end is the low-noise amplifier. Aggregating balun, wide-band low-noise amplifier, and I-Q mixer; the BLIXER was designed following a classical design flow step-by-step. First of all, we collected the performance specification from the top-down design and the design space characterization from the CMOS 65 nm technology. Then, we obtained the BLIXER model equations using a constructive approach and the gradient for all design performances. Next, we obtained a BLIXER design solution using a classical design optimization.

In this step, the bias voltage obtained imposes the transistor in moderate inversion. This solution reduces the power consumption in required minimum gain and maximum noise. After achieving the optimal design we decided to evaluate the BLIXER failure for both variability and ageing. We found that such a solution imposes less circuit stress which implies in less ageing. The trade-off presented in design strategies will be discussed in Chapter 6. Now, we discuss such a trade-off inside the BLIXER design example presenting an increasing variability compromise.

In order to obtain a BLIXER failure evaluation, we developed a sensitivity analysis and a variation sharing. Thus, the BLIXER failure and design space will be discussed aiming a reliable-BLIXER design. Later, the transistor-level BLIXER was simulated and its performances estimated using a typical-corner electric simulation, a nominal reliability simulation and a Monte Carlo simulation. We analyzed these three results concluding that the ageing is negligible in comparison to the variability. We published these analyses and discussions in [23].

The second building block in the RF front-end is the digital controlled oscillator (DCO). The DCO design has followed two different points of view. First design is aiming a reliable-circuit synthesis and the result is a full analysis comparable to the BLIXER. We published these analysis and discussions in [28]. The second point of view is the answer to the question: how much has the reliability increased during the reliable-circuit synthesis? In order to answer this simple question, we redesigned the DCO using a classical design methodology. As a result, we compare the different points of view: if we choose a reliable-circuit and if we do not take the reliability into account during the design flow. We published theses analyses and discussions in [25].

The conclusion of the DCO design clarifies the advantages and the disadvantages to analyze the reliability in early stages and take measures to manage this criterion [25]. Designing a classical and a reliable DCO, we achieve a frequency degradation reduction by a value between 15 % and 30 %. And also, the reliable-DCO has a circuit lifetime five times longer than the classical-DCO, if we fix the maximum frequency range degradation at 2.0 %. Disadvantages of the method are the increase of the phase noise and the reduction of the frequency range.

The third building block in the RF front-end is the programmable-gain amplifier (PGA).The PGA design is mostly composed of reconfigurable passive components with some few transistor-level constraints in the amplifiers. These few transistor-level constraints are similar to the constraints found in the BLIXER design. Thus, we explain why the PGA is naturally reliable (in CMOS 65 nm) in terms of their

amplifiers. The sensitivity analysis and the PGA failure evaluation reveals that the amplifier-performance variation is not big enough to imply in a PGA failure. Therefore, we do not need to full design the PGA to be sure that the transistor ageing in PGA is negligible in comparison to the transistor variability. Finally, the PGA reliability is dependent on the reliability of the control circuits, mostly digital circuits. The reliability in digital circuits is out of the scope of this work.

4.1 RELIABLE-BLIXER DESIGN

4.1.1 BLIXER SCHEMATIC AND SPECIFICATION

The BLIXER topology is presented in Figure 4.1 and its specifications in Table 4.1. For the circuit design, we used the transistor-level characterization which will be presented in Section 6.2. This characterization consists in a I_{DS} and *gm* versus bias simulation for a fresh normalized NMOS (W = 1 μm and L = 60 nm) and the CMOS 65 nm integrated process ageing simulation.

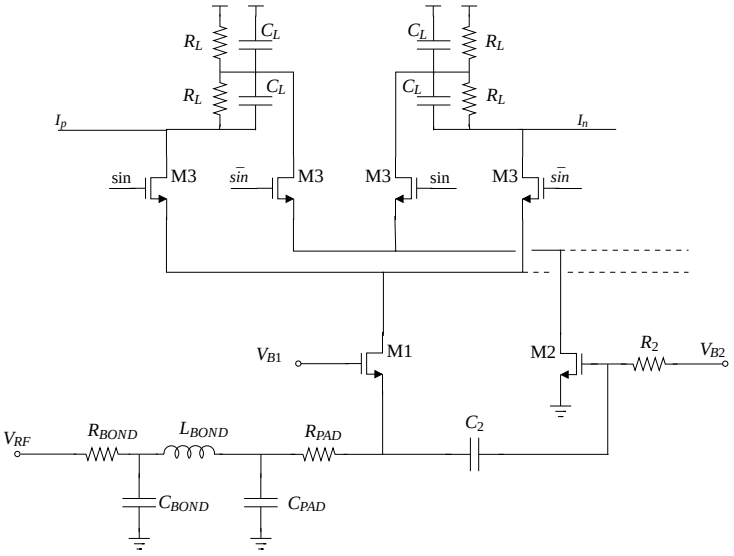


Figure 4.1: The BLIXER topology composed of a balun, a wide-band low noise amplifier (WBLNA), and a mixer connected: illustration of the I-path (the Q-path is not present as it is similar to the I-path) [2].

Table 4.1: The BLIXER specifications for wireless applications.

Operational Frequency	1 GHz - 5 GHz
Bandwidth	$>$ 100 MHz
Gain	$\geq$ 10 dB
NF	$\leq$ 5.0 dB
IP3	$\geq$ 0.0 dBm
S11	$\leq$ -10 dB

4.1.2 BLIXER MODEL EQUATION

POWER CONSUMPTION The power consumption is the first important circuit characteristic to be evaluated; it is described using the equation

$$\mathrm{P} = V_{DD}\left(I_{D1} + I_{D2}\right). \tag{4.1}$$

Then, we found the total derivative of such a characteristic as

$$\Delta\mathrm{P} \approx V_{DD}\left(\Delta I_{D1} + \Delta I_{D2}\right), \tag{4.2}$$

which leads to a decreasing power consumption with the I_{DS} ageing degradation. Thus, the circuit degradation does not lead to a worst characteristic as a power consumption decreasing is desirable. Although, the power consumption is an easy way to evaluate the circuit ageing degradation [28].

GAIN The BLIXER gain is defined as:

$$\mathrm{G} = \frac{\pi}{4}\left(2gm_1R_L + gm_2R_L\right), \tag{4.3}$$

where $\frac{\pi}{4}$ is the passive mixer conversion gain and we assume an ideal matching impedance. So, its total derivative is

$$\Delta\mathrm{G} \approx \frac{\pi}{4\sqrt{2}}\left(2\Delta gm_1R_L + \Delta gm_2R_L\right). \tag{4.4}$$

The BLIXER differential gain is $G_d = 2\mathrm{G}$. The gain degradation shall be estimated by the gm ageing degradation, shown in Figure 6.4(b) (in Chapter 6.2). The circuit bandwidth is not controlled by transistor parameter as it is

$$\mathrm{BW} = \frac{1}{2\pi R_L C_L}, \tag{4.5}$$

and it does not degrade with transistor ageing. The R_L resistance might only suffer EM degradation, but it will not be the case if the designer follows the PDK guide reducing the EM probability. Thus, only a gain variation could be interpreted as a BW variation. This will have no impact on RF front-end performance, because the base-band signal has a smaller bandwidth. Furthermore, the PGA will be later charged to control output signal swing and bandwidth when programmed to a specific standard.

NOISE FACTOR The noise factor ($F = 10^{NF/10}$) is defined as:

$$\mathrm{F} = 1 + \frac{\frac{\left((2gm_1R_{L1})^2 - (gm_2R_{L2})^2\right)}{gm_1R_s+1}\overline{e_{n1}^2} + (gm_2R_{L2})^2\,\overline{e_{n2}^2} + \overline{e_{nL1}^2} + \overline{e_{nL2}^2}}{\frac{\left((2gm_1R_{L1})^2 + (gm_2R_{L2})^2\right)}{gm_1R_s+1}\overline{e_{nS}^2}} + \frac{\pi}{4}, \tag{4.6}$$

where $\overline{e_{ni}^2}$ is the noise voltage of the component i and $\frac{\pi}{4}$ is the passive mixer conversion gain degrading the noise. This equation was obtained by the BLIXER noise analysis of the noise model presented in

Figure 4.2. The total derivative of the noise factor is

$$\Delta F \approx \tag{4.7}$$

$$\left(\frac{R_s\left(\overline{e_{nL1}^2}+\overline{e_{nL2}^2}+{R_{L2}}^2\overline{e_{n2}^2}{gm_2}^2+\frac{8kT\left({R_{L1}}^2{gm_1}^2-{R_{L2}}^2{gm_2}^2\right)}{3gm_1(R_sgm_1+1)}\right)}{\overline{e_{nS}^2}\left({R_{L1}}^2{gm_1}^2+{R_{L2}}^2{gm_2}^2\right)} \right. \tag{4.8}$$

$$-\frac{(R_sgm_1+1)\left(\frac{8kT\left({R_{L1}}^2{gm_1}^2-{R_{L2}}^2{gm_2}^2\right)}{3gm_1{}^2(R_sgm_1+1)}-\frac{16{R_{L1}}^2kT}{3(R_sgm_1+1)}+\frac{8R_skT\left({R_{L1}}^2{gm_1}^2-{R_{L2}}^2{gm_2}^2\right)}{3gm_1(R_sgm_1+1)^2}\right)}{\overline{e_{nS}^2}\left({R_{L1}}^2{gm_1}^2+{R_{L2}}^2{gm_2}^2\right)} \tag{4.9}$$

$$\left. -\frac{2{R_{L1}}^2gm_1\left(R_sgm_1+1\right)\left(\overline{e_{nL1}^2}+\overline{e_{nL2}^2}+{R_{L2}}^2\overline{e_{n2}^2}{gm_2}^2+\frac{8kT\left({R_{L1}}^2{gm_1}^2-{R_{L2}}^2{gm_2}^2\right)}{3gm_1(R_sgm_1+1)}\right)}{\overline{e_{nS}^2}\left({R_{L1}}^2{gm_1}^2+{R_{L2}}^2{gm_2}^2\right)^2} \right) \Delta gm_1 \tag{4.10}$$

$$\left(\frac{(R_sgm_1+1)\left(\frac{8{R_{L2}}^2\mathrm{kT}}{3}-\frac{2{R_{L2}}^2\overline{e_{n1}^2}gm_2}{R_sgm_1+1}\right)}{\overline{e_{nS}^2}\left({R_{L1}}^2{gm_1}^2+{R_{L2}}^2{gm_2}^2\right)} \right. \tag{4.11}$$

$$\left. -\frac{2{R_{L2}}^2gm_2\left(R_sgm_1+1\right)\left(\mathrm{enL1}+\overline{e_{nL2}^2}+\frac{\left({R_{L1}}^2{gm_1}^2-{R_{L2}}^2{gm_2}^2\right)}{R_sgm_1+1}+\frac{8{R_{L2}}^2gm_2\mathrm{kT}}{3}\right)}{\overline{e_{nS}^2}\left({R_{L1}}^2{gm_1}^2+{R_{L2}}^2{gm_2}^2\right)^2} \right) \Delta gm_2. \tag{4.12}$$

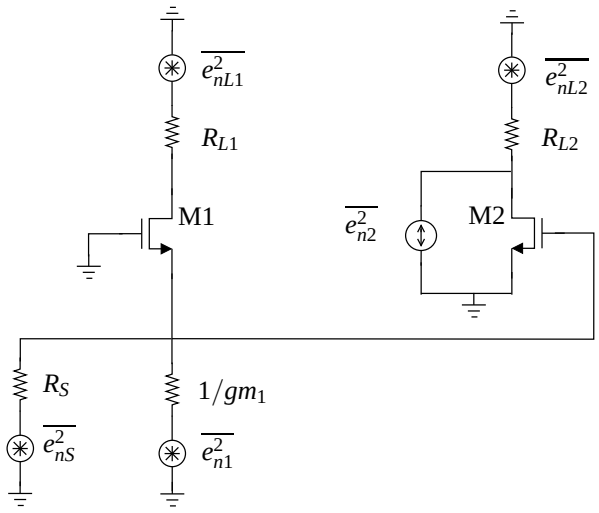


Figure 4.2: The BLIXER noise model: illustration.

INPUT IMPEDANCE MATCHING The BLIXER matching impedance is defined:

$$S_{11} = \frac{1-gm_1R_s}{1+gm_1R_s}, \tag{4.13}$$

where R_s is the source resistance and its total derivative is

$$\Delta S_{11} \approx \frac{-2R_s}{(1+gm_1R_s)^2}\Delta gm_1. \tag{4.14}$$

Thus, the S_{11} is a performance with a large sensitivity to the ageing degradation of the gm_1.

4.1.3 RELIABLE-BLIXER DESIGN

Using the performance specification and the initial design space, the optimization looks for the transistor parameters which lead to an optimum BLIXER performance. Thus, the required circuit bias is $V_{DS} = 0.4$ V and $V_{GS} = 0.44$ V for both M1 and M2 transistors. The optimized sizes are: $W_1 = 64.0\ \mu$m, $W_2 = W_3 = 200.0\ \mu$m, $L = 60$ nm. The passive components values are: $R_L = 150\ \Omega$, and $C_L = 3$ pF. The components of the bond wire and pad modeling are $R_{BOND} = 0.5\ \Omega$, $C_{BOND} = 200$ fF, $L_{BOND} = 1$ nH, $R_{PAD} = 0.2\ \Omega$, $C_{PAD} = 200$ fF, $R_2 = 12.0$ kΩ, and $C_2 = 30$ pF.

Regarding the design space (presented in Section 6.2), a normalized drain current of 17.4 μA, and a normalized transconductance of 0.1 mS are expected. Hence, the first performance estimation gives us: $P = 5.5$ mW, $G_d = 14.7$ dB, $NF = 3.9$ dB. The design space also indicates the ageing variation under 1% and the variability of 75% for the drain current and 55% for the transconductance. The bias choice has reduced the HCI stress, because the overdrive voltage is small and the NMOS transistors are in moderate inversion. However, the bias choice has increased the importance of the transistor variability which could be responsible for the BLIXER failure.

SENSITIVITY ANALYSIS The BLIXER sensitivity analysis has pointed the transistor M2 as the most sensitive transistor. In power consumption sensitivity analysis, we found

$$S_P^{I_{DS1}} = 0.24, \text{ and} \tag{4.15}$$
$$S_P^{I_{DS2}} = 0.75; \tag{4.16}$$
$$\tag{4.17}$$

which means that M2 consumes more and is more sensitive. In gain sensitivity analysis, we found

$$S_G^{gm_1} = 0.39, \text{ and} \tag{4.18}$$
$$S_G^{gm_2} = 0.60; \tag{4.19}$$
$$\tag{4.20}$$

which is a similar result for M2 as presented in power consumption sensitivity analysis. In noise sensitivity analysis, we found

$$S_F^{gm_1} = 0.83, \text{ and} \tag{4.21}$$
$$S_F^{gm_2} = -1.34; \tag{4.22}$$
$$\tag{4.23}$$

thus it is clear the M2 noise cancellation role in the sensitivity signal. The M2 parameters variation will decrease the noise cancellation.

FAILURE EVALUATION Considering only the nominal circuit characteristic, a classical synthesis will just optimize the circuit sizing. After that, the designer should simulate the circuit variability and ageing. Using a given specification of maximum ageing degradation of the transistor parameters and the total derivatives (Equations (4.4), (4.12) and (4.14)), we propose to early estimate the circuit characteristic variability and ageing.

Assuming the case of 1 % ageing degradation of the transistor parameters, we found the estimated characteristic degradation as: $\Delta P/P = 1$ %, $\Delta G/G = 1$ %, $\Delta S11/S11 = 0.7$ %, and $\Delta F/F = 0.5$ %. Such information shall be used on the design flow to improve the circuit against ageing. Moreover, it leads to how sensitive the chosen circuit schematics and technology are to ageing degradation. Hence, such failure evaluation proves that the designed BLIXER is reliable.

Assuming the case of 75% for the drain current and 55% for the transconductance variability, we found the estimated characteristic variation as: $\Delta P/P = 75$ %, $\Delta G/G = 55$ %, $\Delta S11/S11 = 39.2$ %, and

$\Delta F/F = 27.7$ %. Such information clarifies the importance of the transistor variability on the BLIXER failure. Therefore, the BLIXER performance and failure should be discussed in higher level evaluating the impacts of the characteristics variation on the yield requirements.

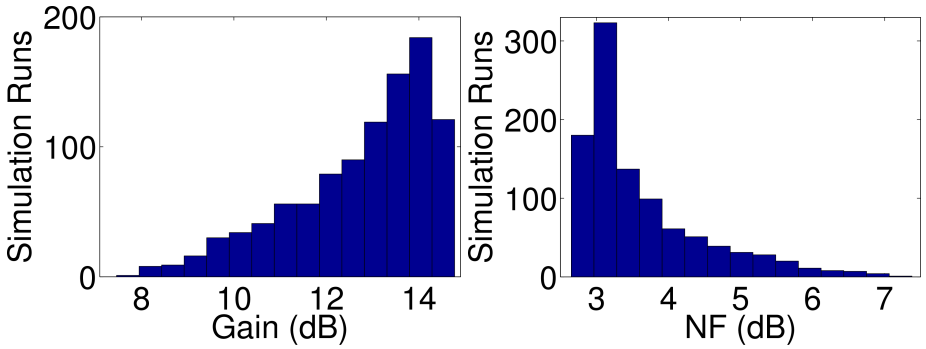
4.1.4 BLIXER Simulation Results

First, the BLIXER design example was simulated obtaining the nominal characteristics for RF-input-signal frequencies: 1.0 GHz, 2.4 GHz, and 5.0 GHz. The simulation results of the typical circuit are summarized in Table 4.2 and they are coherent with the RF front-end specification.

Table 4.2: The BLIXER simulation results: typical performance.

Frequency (GHz)	1.0	2.4	5.0
Power (mW) Consumption	5.55	5.56	5.59
Differential Gain (dB)	13.5	13.5	14.3
NF_{max} (dB)	4.0	4.5	5.3
IP3 (dBm)	4.4	3.4	-0.73
S11 (dB)	-17.6	-15.4	- 13.3

Subsequently, the BLIXER design example was simulated obtaining the variability of the Gain (Figure 4.3(a)) and NF_{max} (Figure 4.3(b)) characteristics. We present in Figure 4.3 the results of a 1000 points of Monte Carlo simulation of the fresh BLIXER for the 1.0 GHz of input-signal frequency. The simulated result shows that 96.4 % of the simulation runs have Gain $>$ 10.0 dB, and 92.1% of the simulation runs have $NF_{max} < 5.0$ dB. Therefore, both characteristics did not lead to the BLIXER drop out of the specifications despite the integrated process variation and mismatch.



(a) (b)

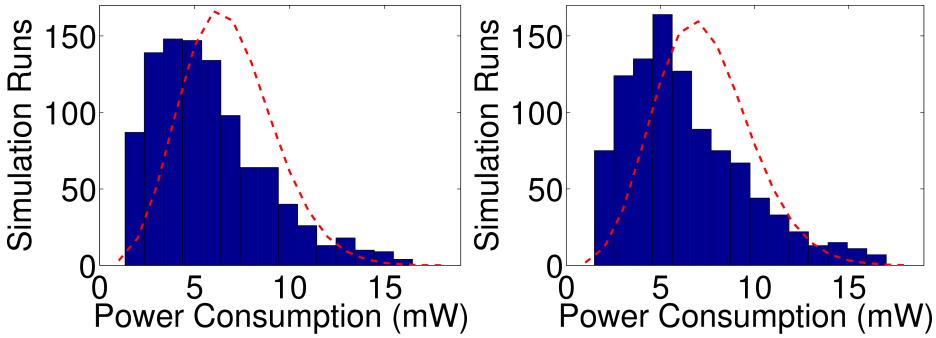
Figure 4.3: A 1000 points Monte Carlo simulation of the fresh BLIXER (a) for the Gain and (b) for the NF_{max} characteristics at 1.0 GHz of input-signal frequency.

In order to characterize the BLIXER ageing degradation, the power consumption (P) will be evaluated. We propose to evaluate the power consumption by 1000 points of Monte Carlo simulation of the fresh BLIXER and the 30 years old stressed BLIXER. Firstly, we obtained the model library for the 30 years aged BLIXER using the typical integrated process parameters. Next, we did a 1000 points Monte Carlo simulation of the fresh BLIXER (presented in Figure 4.4(a)) and of the 30 years aged BLIXER using the aged model library (presented in Figure 4.4(b)). The mode of the fresh BLIXER histogram

is 4.2 mW. The mode of the 30 years aged BLIXER histogram is 4.4 mW. Then, we fitted a Poisson distribution for 99.9% of confidence, described by the probability density function

$$f(P;\lambda) = \frac{\lambda^P e^{-\lambda}}{P!}, \tag{4.24}$$

and we found $\lambda_{fresh} = 5.8$ mW and $\lambda_{aged} = 6.3$ mW. As expected, the bias voltages chosen will lead to an ageing degradation much lesser than 1 % checking the Figures 6.4(a) and 6.4(b) (in Section 6.2). Therefore, the performance-ageing degradation of the BLIXER is negligible.



(a) (b)

Figure 4.4: A 1000 points Monte Carlo simulation (a) of the fresh BLIXER and (b) of the 30 years aged BLIXER power consumption using the obtained model library.

4.1.5 Reliable-BLIXER Design Conclusions

In this section, we proposed the reliable-BLIXER design using the failure evaluation. In the bottom-up approach the circuit design equations have been used to obtain an early estimation of the circuit characteristics ageing. Using the CMOS 65 nm transistor ageing characterization, the sensitive circuit bias condition was identified. The HCI can be avoided reducing the time in which the transistors are in strong inversion, controlling the V_{GD} and the V_{GS}. In order to increase the circuit reliability and also to achieve the desired performance, the required circuit bias is $V_{DS} = 0.4$ V and $V_{GS} = 0.44$ V.

The simulation results of the typical circuit are coherent with the RF front-end specification. Despite the integrated process variability and mismatch, we observed that 96.4 % of the simulation runs have Gain > 10.0 dB, and 92.1% of the simulation runs have $\text{NF}_{max} < 5.0$ dB, and they did not lead the BLIXER drop out of the specifications. Moreover, the BLIXER ageing degradation is negligible according to the fitted Poisson distribution of the power consumption for 99.9% of confidence.

4.2 Reliable-DCO Design

4.2.1 DCO Schematic and Specification

In order to illustrate the reliable-DCO design method, we chose to design a realistic DCO for RF applications near 1 GHz. The 5-3 NOR interpolative DCO [24] was chosen because it has a large signal at the transistor gates, resulting in HCI degradation. And also, we cannot control the PMOS transistor NBTI stress time, as this time is correlated with the oscillation frequency. The DCO, designed in CMOS 65 nm, has to generate a frequency clock between 600 MHz and 1.2 GHz. The Figure 4.5 shows the circuit schematic, where V_{CT} and V_{BIAS} are the control words converted into bias voltages for NMOS and PMOS respectively.

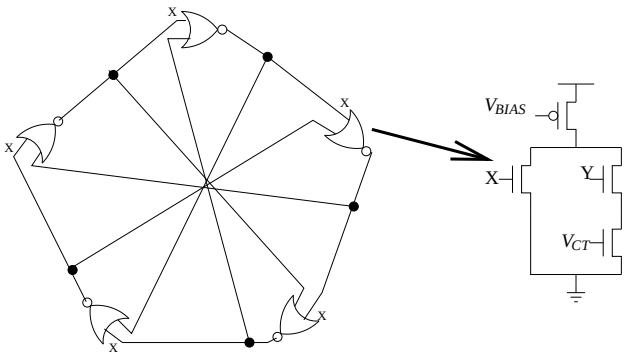


Figure 4.5: The 5-3 NOR interpolative DCO schematic [24].

4.2.2 DCO Failure Evaluation

We proposed a reliability analysis at the transistor level in [25]. The device behavior chosen to model the HCI and the NBTI degradations in this analysis is the transistor drain current. Thus, we assume that

$$I_{DSaged} = (1-\alpha)\, I_{DSfresh}, \tag{4.25}$$

where $\Delta I_{DS} = \alpha I_{DSfresh}$ and α is the percentage of I_{DS} degradation. α is evaluated between the fresh circuit simulation and the aged circuit simulation as [85]

$$\alpha = \frac{I_{DSaged} - I_{DSfresh}}{I_{DSfresh}}. \tag{4.26}$$

The [1] and [82] have shown similar models. In order to evaluate the circuit reliability, ΔI_{DS} will feed the transistor model shown in Figure 4.6. This transistor model will replace its corresponding transistor and indicate the sensitivity of the devices [85].

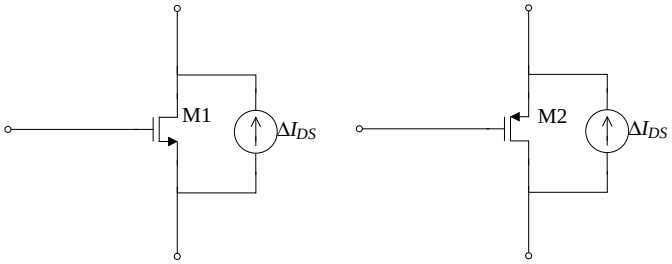
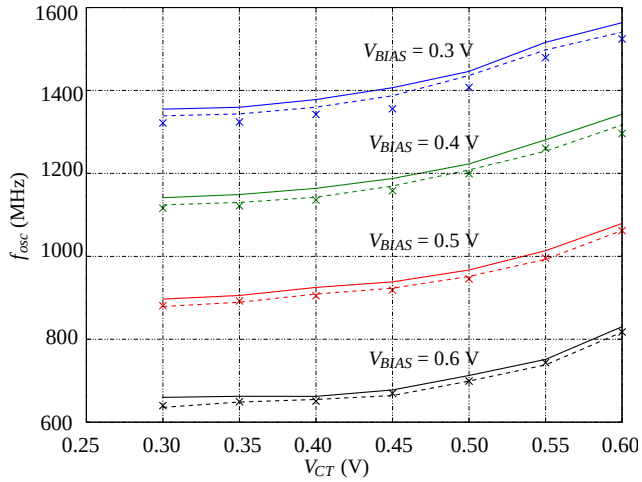


Figure 4.6: NMOS and PMOS transistors reliability model for analysis purposes ($\Delta I_{DS} = \alpha I_{DSfresh}$).

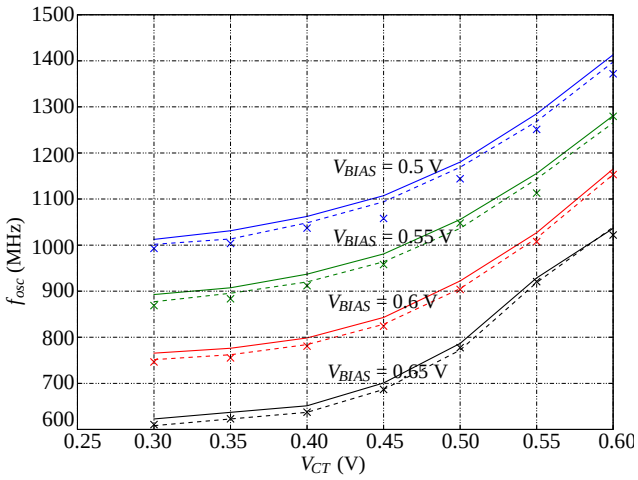
The most important step consists in describing the typical environment conditions, such as temperature, signal swing and circuit frequency. In order to obtain a consistent model, the circuit test bench should represent these stress conditions. Thus, the circuit degradation is measured by the chosen behavior of the transistor (I_{DS}) under HCI and NBTI degradations.

As earlier the designer can simulate the cell performance, as better it is to evaluate ΔI_{DS} degradation. Moreover, such degradation represents the behavior variation that the circuit should be robust to. Thus, a design methodology could predict how robust the circuit is and if a possible redesign can increase its reliability. In order to achieve better results, the analysis is iterated with the design in a new reliable-design flow (see Chapter 6).

APPLICATION OF THE ANALYSIS MODEL In order to demonstrate the reliability analysis method, we show the analysis of a DCO designed with the classical methodology and a reliable-DCO (design presented in Subsection 4.2.3). Figure 4.7(a) and 4.7(b) show both DCOs stressed for 30 years of ageing at 27 ^{o}C for all digital control words, converted into analog voltages. Actually, the full validation of the analysis model was presented in [28].



(a)



(b)

Figure 4.7: The DCO oscillation frequency (f_{OSC}) simulated at 27 ^{o}C during 30 years of degradation for (a) the classical and (b) reliable designed circuit. The result before the stress is represented by the solid line, after the stress by a dashed line and the aged model marked by x

4.2.3 RELIABLE-DCO DESIGN

First of all, we will not consider the variability of the integration process technology, discussing only ageing degradation. Thus, only the reliability could change the circuit performance and so the result will be the transistor-level considerations aiming a reliable-DCO. In this circuit, the minimum number of stages is 3 and the maximum (N_{max}) is 5, so we have 4 effective stages (N_{eff}) in the feedback loop. The oscillation frequency is

$$f_{osc} = \frac{1}{2N_{eff}t_d} \tag{4.27}$$

where t_d is the NOR cell delay. We found 80 ps $< t_d <$ 200 ps to be the cell delay, which can be described as

$$t_d = \frac{C_L}{gm}. \tag{4.28}$$

We assume that the delay is the same for all cells for both rise and fall transitions of the signal. Approximately,

$$gm = \frac{2I_{DS}}{V_{ov}}, \tag{4.29}$$

where V_{ov} is the overdrive voltage and I_{DS} is the drain to source current. Therefore, f_{osc} is a function of I_{DS}, as

$$f_{osc} = \frac{I_{DS}}{N_{eff} C_L V_{ov}}, \tag{4.30}$$

and the current degradation will represent a loss in the oscillation frequency. The oscillation frequency degradation can be obtained by deriving the equation (4.30). As presented in

$$\Delta f_{osc} = \frac{\Delta I_{DS}}{N_{eff} C_L V_{ov}}, \tag{4.31}$$

Δf_{osc} has a linear dependence with ΔI_{DS} transistor behavior chosen in the reliability analysis.

As the frequency, the maximum power consumption (P_{max}) is a function of the maximum current at the load. Thus, P_{max} , evaluated as

$$P_{max} = N_{max} V_{DD} I_{DS}, \tag{4.32}$$

is only limited by I_{DS} (assuming constant V_{DD}). Consequently, the maximum power consumption degradation is

$$\Delta P_{max} = N_{max} V_{DD} \Delta I_{DS}. \tag{4.33}$$

Another important performance parameter to be evaluated in oscillators is the phase noise ($L(f)$). The f is the frequency distance between the carrier frequency and the noise frequency. According to [86], the $L(f)$ is a function of the transistor bias current (I_{DS} in our design) and a constant γ_n which represents the transistor noise model, as

$$L(f) = 10 \log \frac{\gamma_n}{I_{DS}}. \tag{4.34}$$

Hence, the phase noise ageing degradation is

$$\Delta L(f) = 10 \log (1 - \alpha), \tag{4.35}$$

where α was defined in equation (4.26).

During the design optimization, we naturally found that the power consumption is reduced, if the designer chooses $V_{GS} = V_{BIAS} - V_{DD}$ as low as possible. In this bias condition, the reliability analysis shows lower NBTI degradation. Figure 4.8 illustrates the amount of NBTI degradation over years of degradation for each V_{GS} at 150 ^{o}C of temperature stress (see Section 6.2 for more details). Clearly, the circuit will present lower NBTI degradation for V_{GS} >-0.7 V and will consume less power. Aiming to cover the oscillation frequency span, the designer has to choose between accepting the degradation presented in Figure 4.8 or redesigning the DCO to be reliable. If the reliability and the power consumption constraints are chosen to be optimized, the designer will need an overdesigned DCO to cover from 600 MHz to 1.2 GHz for V_{BIAS} >0.5 V ($V_{DD} = 1.2$ V) . Such an optimization will be the design for reliability choice. On the other hand, the classical design methodology will choose not to overdesign the DCO, and hence the DCO will be exploited for 0.3 V $< V_{BIAS} <$0.6 V.

The classical design indicates that the PMOS width should be 3 times bigger than the NMOS width (W) for the DCO clock with an equal duty cycle. Increasing the reliability, we propose to use V_{BIAS} >0.5 V and to design the width as to W/2 for both NMOS in the Y path (marked with Y and V_{CT} in the transistor gate in Figure 4.5). Comparing the classical design and the reliable design options, the dimensions of the DCO are present in the Table 4.3.

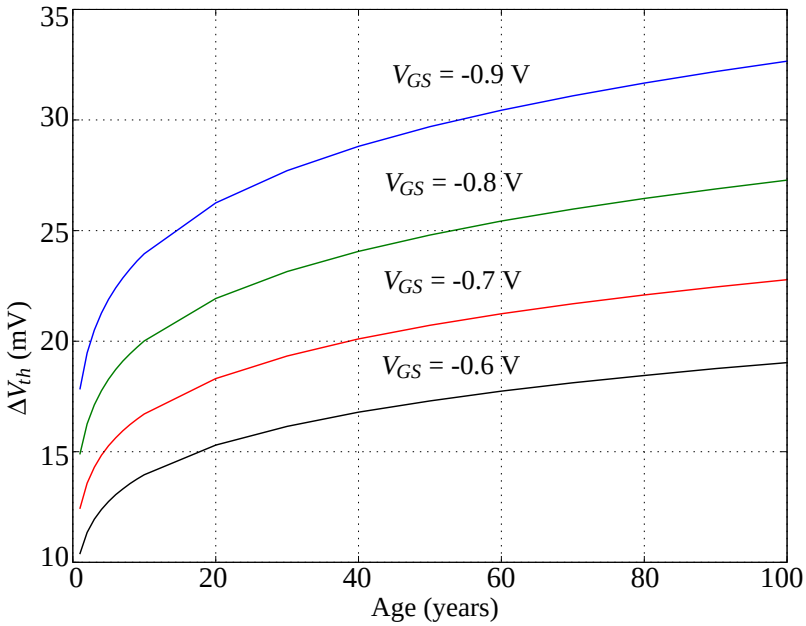


Figure 4.8: ΔV_{th} degradation for the PMOS DCO transistors stressed by NBTI at 150 ^{o}C.

Table 4.3: The dimensions of the designed DCO ($L = 0.5$ μm).

	W_X (μm)	W_Y and W_{CT} (μm)	W_{PMOS} (μm)
DCO	5	5	15.0
	4	4	12.0
	3	3	9.0
Reliable DCO	5	2.5	15.0
	4	2.0	12.0
	3	1.5	9.0

In order to reduce the phase noise, the sizes shall be increased. Table 4.4 confirms the phase noise reduction as the NMOS width increases, but it also indicates the increase of the oscillation frequency degradation for $f_{osc} = 1$ GHz and after 10 years of ageing. The reliable DCO has the oscillation frequency degradation reduced by a value between 15 % and 30 %. As a trade-off, the phase noise presented an average increase of 2.0 dBc/Hz. If such a phase noise is reasonable for a target RF standard, the reliable DCO will be the better option. Both DCOs will be designed with $W_X = 5$ μm, aiming a phase noise reduction.

4.2.4 DCO Simulation Results

Combining the trends indicated by the reliability analysis, both DCO circuits were sized ($L = 0.5$ μm and the widths are presented in the line of $W_X = 5$ μm of the Table 4.3) to cover from 600 MHz to 1.2 GHz. The classical designed DCO (at 1 GHz) consumed 1.40 mW of power without stress and 1.37 mW after 10 years of stress degradation. The reliable DCO (at 1 GHz) consumed 1.06 mW of power without stress and the same power after 10 years of stress degradation (negligible ageing).

The reliability simulation reveals in the Figure 4.9(a) and 4.9(b) the normalized f_{osc} degradation obtained from both DCOs stressed at 27 ^{o}C for different ageing times. The classical designed DCO

Table 4.4: Phase Noise (f_{osc} = 1 GHz at 1 MHz off-set) versus the f_{osc} and P ageing degradations at 10 years lifetime for each design and sizing.

	W_X (μm)	$\Delta f_{osc}/f_{osc}$ @ 1 GHz and 10 years (%)	$\Delta P/P$ @1 GHz and 10 years (%)	L(1.0 MHz) @1 GHz (dBc/Hz)
DCO	5	1.90	2.2	-94.3
	4	1.86	2.2	-93.4
	3	1.90	2.2	-92.2
Reliable DCO	5	1.6	0.4	-92.7
	4	1.7	0.4	-91.7
	3	1.3	2.1	-90.5

presented a bigger frequency range and bigger ageing degradation up to 2.8 %. Moreover, the classical designed DCO at 1 GHz results in 2.0 % of degradation limiting its lifetime at 10 years. The reliable DCO presented a lower frequency range but achieves a smaller degradation in most cases and 2.5 % at worst. Therefore, the reliable DCO can be exploited more than 50 years of lifetime in most part of the frequency range. Its lifetime is 50 years at 1 GHz for 2.0 % of degradation.

In order to present the strengths of the reliable DCO, we choose to present the reliability simulation in an accelerate environment at 150 ^{o}C. The Figure 4.10(a) and 4.10(b) reveals a bigger NBTI stress. The normalized f_{osc} degradation obtained for the classical designed DCO is more than 3 % in most cases. On the other hand, the reliable DCO has presented a f_{osc} degradation between 2 % and 3 % which represents a reduction of 30 % of the f_{osc} degradation in most cases.

The DCO and the reliable DCO phase noise simulations, presented in the Figure 4.11, are very close and only a specific RF standard choice can impose a limit to the phase noise constraint. The ageing simulations performed at each circuit lifetime (also presented in the Figure 4.11) didn't present noise performance degradation big enough to be considered a failure.

The variability simulation was performed by 1000 points of Monte Carlo simulation for both DCO programmed at f_{osc} = 1 GHz and stressed at 27 ^{o}C. First, they are simulated against only variability and the results are presented in the Figures 4.12(a) and 4.12(b). After that, they are simulated using the ageing model (presented in Figure 4.6) extracted from the nominal run for 10 years of stress (presented in the Figures 4.13(a) and 4.13(b)). All results demonstrate that the reliability has no significant impact compared to the variability, and therefore it could be neglected in our example.

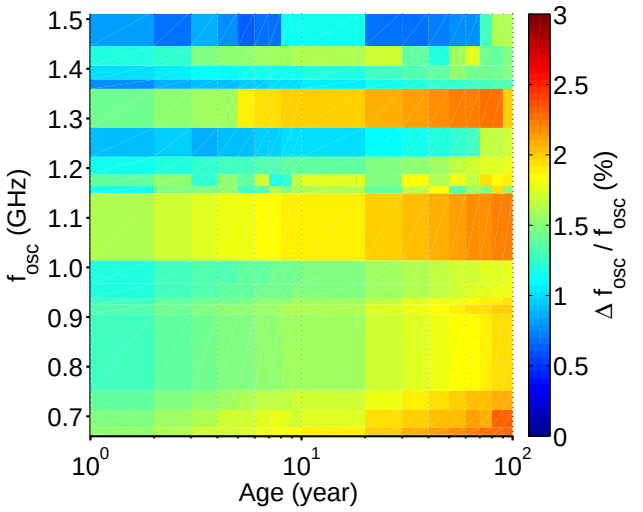
Finally, both DCOs can be characterized by the figure of merit for RF oscillators, defined as

$$\text{FoM} = L(\Delta f) - 20\log\frac{f_{osc}}{\Delta f} + 10\log P/1\text{ mW [87]}. \tag{4.36}$$

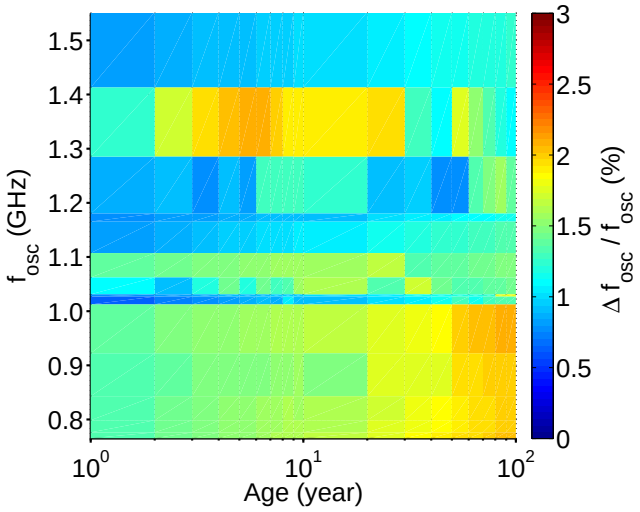
The FoM of the fresh circuits was evaluated as -152.8 dB for the classical designed DCO and -152.4 dB for the reliable DCO. If we take into account the power consumption (equation (4.33)) and the phase noise (equation (4.35)) degradation, the FoM for RF oscillators doesn't take the reliability degradation into account. Both degradations compensate each other in equation (4.36), and therefore there is no need to present the FoM for the aged circuits.

4.2.5 Reliable-DCO Design Conclusions

In this section, we validate a reliable-circuit synthesis method using a DCO design. We have considered the reliability degradation caused by circuit ageing as a design criterion, as important as the classical constraints (noise, signal range, power consumption and die area). For this purpose, we introduced the physical phenomena context and how they could be avoided. Then, we presented the reliable-DCO design and validated its reliability analysis model.



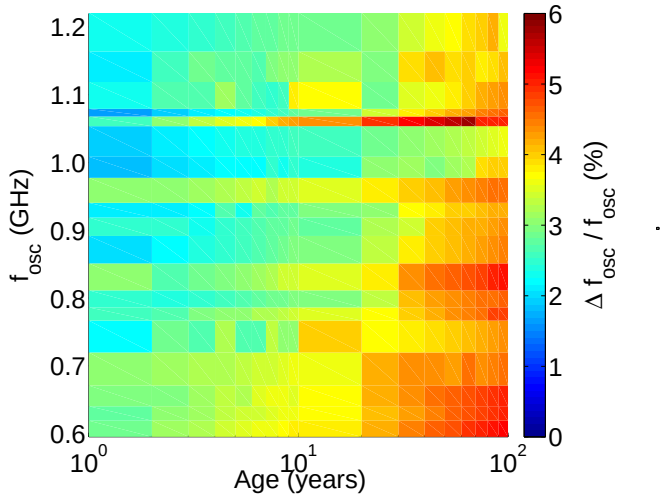
(a)



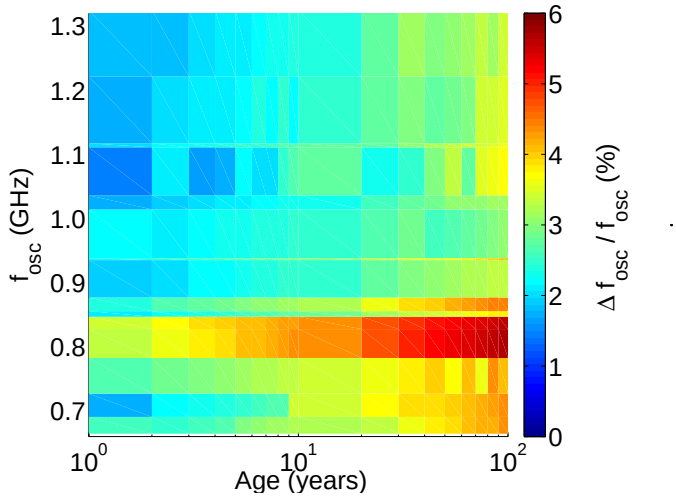
(b)

Figure 4.9: The $\Delta f_{osc}/f_{osc}$ degradation at 27 ^{o}C for (a) the classical and (b) the reliable designed DCO.

We checked that the reliability analysis gives us information to improve the method optimization, designing a more reliable circuit. Designing a classical and a reliable DCO, we achieve a frequency degradation reduction by a value between 15 % and 30 %. And also, the reliable DCO has a circuit lifetime five times longer than the classical DCO, if we fix the maximum frequency range degradation at 2.0 %. Disadvantages of the method are the increase in the phase noise and the reduction of the frequency range that could be neglected in a specific RF standard.



(a)



(b)

Figure 4.10: The $\Delta f_{osc}/f_{osc}$ degradation at 150 ^{o}C for (a) the classical and (b) the reliable designed DCO.

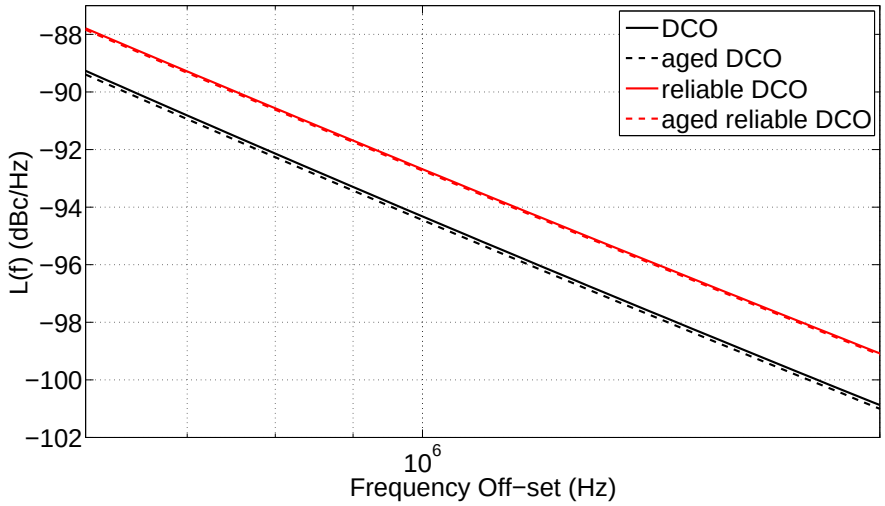
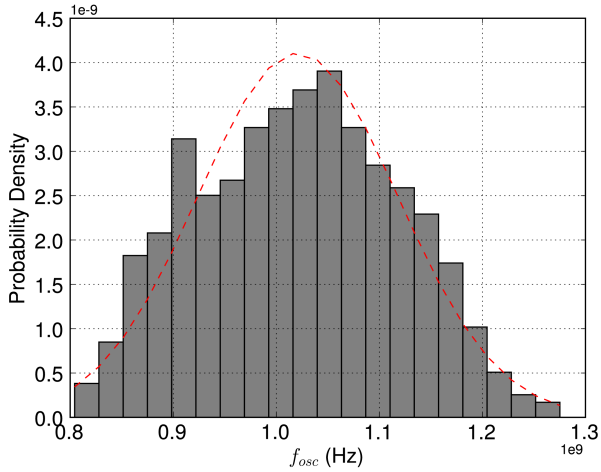
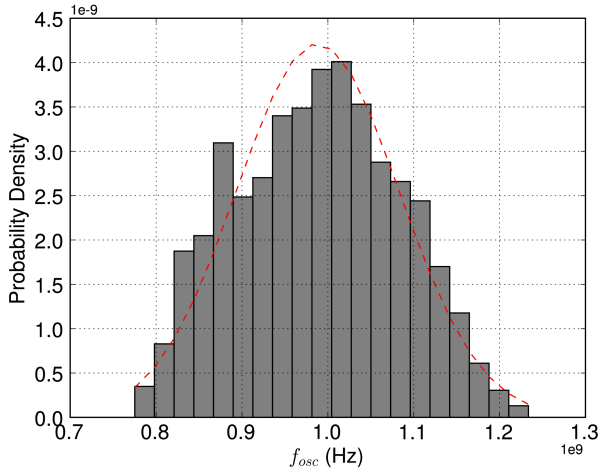


Figure 4.11: The phase noise performance of the DCO and the reliable DCO, simulated for $f_{osc} = 1$ GHz.

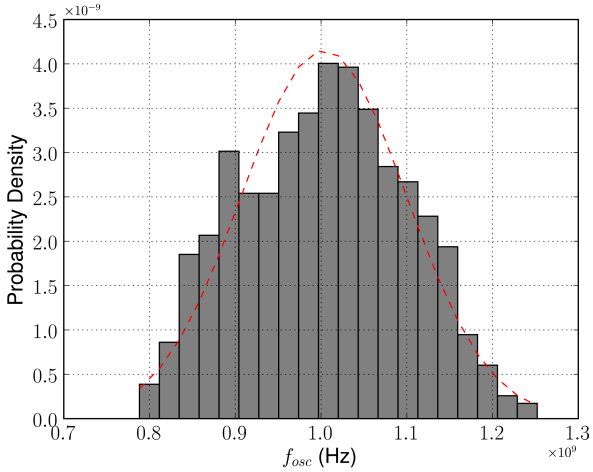


(a)

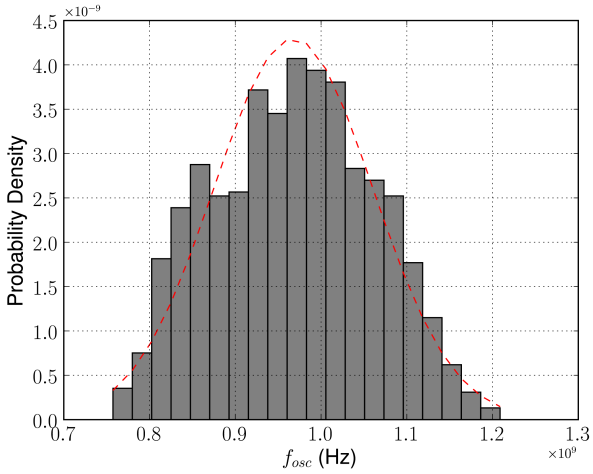


(b)

Figure 4.12: The DCO f_{osc} variability, simulated at $f_{osc} = 1$ GHz by 1000 Monte Carlo runs. (a) The classical designed circuit presented μ =1.020 GHz and σ =97.1 MHz, and (b) the reliable designed presented μ =0.988 GHz and σ =94.7 MHz.



(a)



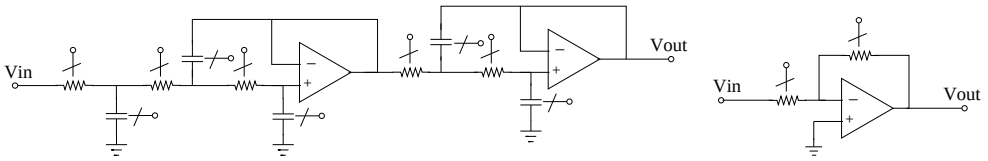
(b)

Figure 4.13: The DCO f_{osc} variability, simulated with the ageing model for 10 years of stress at $f_{osc} = 1$ GHz by 1000 Monte Carlo runs. (a) The classical designed circuit presented μ =1.003 GHz and σ =95.9 MHz, and (b) the reliable designed presented μ =0.967 GHz and σ =92.7 MHz.

4.3 RELIABLE-PGA DESIGN

4.3.1 PGA SCHEMATIC AND SPECIFICATION

In this section, we will present why the PGA is naturally a reliable circuit using a similar analysis and synthesis methodology applied at the reliable-BLIXER design. The most common topology is a multi-mode, multi-band active-RC filter and amplifier [4]. Figure 4.14 illustrates the design example of an active-RC 5th-order Sallen-Key filter presented in [4].



(a) (b)

Figure 4.14: The multi-mode, multi-band PGA schematics is composed of (a) an active-RC Sallen-Key filter and (b) a negative-feedback amplifier.

Furthermore, it is clear that all active-RC filter and amplifier will only meet multi-mode, multi-band specifications if it has reconfigurable passive components. Another possible solution is a multi-path filter with a programmable path selection (not discussed here). Using a full differential and balanced structure, the reliability degradation will hit mostly the amplifiers and the control circuits. If the control circuits are digital gates aiming to reduce the area consumption, then the circuit reliability should be increased with digital techniques. These techniques are out of the scope of this work. Hence, the amplifier is the only analog block which we can propose some reliability increase.

The PGA specifications are summarized in Table 4.5 and according to the state-of-the-art these performances are achievable. For the circuit design, we used the transistor-level characterization which will be presented in Section 6.2. This characterization consists in a I_{DS} and gm versus bias simulation for fresh normalized transistors (W = 1 μm and L = 60 nm; p type and n type) and the CMOS 65 nm integrated process ageing simulation.

Table 4.5: The PGA specifications for wireless applications.

Bandwidth	$\leq$ 25 MHz (programmable)
Gain	$\geq$ 20 dB
NF	$\leq$ 10.0 dB
IP3	$\geq$ 10.0 dBm

4.3.2 PGA MODEL EQUATION

The transfer function of an active-RC filter and an amplifier has the form:

$$H(s) = K\frac{N(s)}{D(s)}; \tag{4.37}$$

where K is the DC gain, $N(s)$ and $D(s)$ are the rational polynomial which implements the 5th-order Sallen-Key filter in the example. Assuming that K, $N(s)$, $D(s)$ are function of passive components and reconfigurable, the only source of performance variation is the amplifier. Furthermore, the amplifier performance variation may not represent a PGA failure, if the PGA could be reprogrammed reducing the variation consequences. However, the amplifier performance variation may impose PGA limitations

and it is the center of the failure estimation. The amplifier key performances are: finite DC gain, finite gain-bandwidth product, linearity and noise.

FINITE DC GAIN The Equation (4.37) assumes an ideal amplifier, which means an infinity DC gain. However, the transfer function in the case of a non-ideal amplifier suffers from an error as

$$H_{real}(s) = H_{ideal}(s)\varepsilon. \tag{4.38}$$

The error ε for a simple case of the schematics illustrated in Figure 4.14(b) is known to be

$$\varepsilon \approx \frac{G_{DC}}{G_{DC}+1}, \tag{4.39}$$

where G_{DC} is the amplifier finite DC gain. The basic amplifier of the Figure 4.15 could be decomposed in smaller amplifiers stages: differential pair amplifier and common source amplifier. The DC gain equation is written as

$$G_{DC} = gm_1 R_{L2} gm_4 R_{L3}, \tag{4.40}$$

where gm_1 and gm_4 are the transconductance of the amplifier transistors; R_{L2} and R_{L3} are the equivalent active charges of the amplifier transistor.

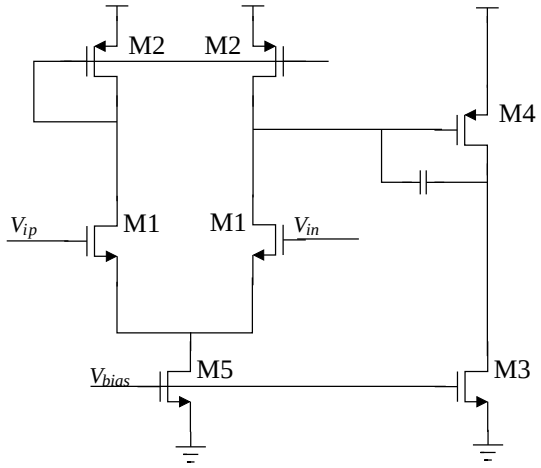


Figure 4.15: Basic operational amplifier: illustration.

A constraints imposed to G_{DC} is to be greater than 1000 V/V, because ε will be lower than 0.1 % in this case. However, the error Equation (4.39) is for only one amplifier, and a handy rule of thumb leads us to an error of ε^3 in the case of 3 amplifiers (e.g. two of them do the active-RC filter as Figure 4.14(a) and the last one the programmable gain as Figure 4.14(b)). Thus, the worst case of error will be increased, becoming approximately 0.3 %.

FINITE GAIN-BANDWIDTH PRODUCT The finite gain-bandwidth product (GBW) introduces an undesired pole on the transfer function. Actually, amplifier gain is frequency dependent and the Equation (4.39) shall be rewritten taking the GBW into account. The error ε for a simple case of the schematics illustrated in Figure 4.14(b) will become

$$\varepsilon(s) \approx \frac{G_{DC}}{\frac{G_{DC}}{GBW}s + G_{DC} + 1}, \tag{4.41}$$

where G_{DC} is the finite DC gain defined in Equation (4.40). The GBW is

$$GBW \approx \frac{gm_1}{C_{comp}}, \tag{4.42}$$

where C_{comp} is the compensation capacitance in Figure 4.15. Thus, the undesired pole will be

$$s = -\frac{G_{DC}+1}{G_{DC}} GBW. \tag{4.43}$$

Normally, this pole is far away of the poles of $H(s)$ becoming non dominant. As a handy rule of thumb, if such a pole is greater than four times the frequency of the dominant one, the undesired pole will be negligible. However, the worst case of error for the PGA on the example will increase the strength of the undesired pole, because such a pole will turns in three poles placed at same frequency.

LINEARITY The PGA linearity depends on the active-RC filter passive components, and the amplifier DC gain. We suppose that the passive components do not suffer from ageing and despite the variability the control circuits may correct the mismatch variations. The amplifier DC gain is responsible for the virtual ground quality, and the finite DC gain problem was already discussed. Furthermore, the state-of-the-art of PGAs [4] indicates a linearity bigger than 20 dBm, and so it will be always bigger than the specified 10 dBm of linearity. Therefore, the linearity should not be a design constraint.

NOISE The PGA noise depends on the passive components (the resistance loads), and the amplifier noise. We suppose that the passive component noise can be reduced controlling the size of the resistances and capacitances. The input referred noise of the amplifier is

$$v_{in}^2 = 2\frac{i_{n1}^2}{gm_1^2} + 2\frac{i_{n2}^2}{gm_1^2}, \tag{4.44}$$

calculated using the small signal analysis explained at Razavi's book [88] at the basic amplifier of the Figure 4.15. And thus, the PGA noise will mostly be influenced by the first-amplifier noise.

4.3.3 RELIABLE-PGA DESIGN

The reliable-PGA design can be devised into two parts: the Sallen-Key transfer function, resulting in the passive-components design values; and the amplifier design in transistor level. The Sallen-Key transfer function is a solved optimization problem and it will not be discussed in this work as mentioned before. The amplifier design in transistor level has the performance specifications summarized in Table 4.6

Table 4.6: The amplifier specifications for PGA application.

Gain-Bandwidth Product (GBW)	$\geq$ 1 GHz
Gain	$\geq$ 60 dB
v_{in}^2	$\leq 4.5\,10^{-17}$ V^2/Hz
IP3	$\geq$ 10.0 dBm

The design of an amplifier with these specifications is not a new challenge and we decided to research the state-of-the-art of miller-amplifiers. In order to achieve the specifications of the Table 4.6, we found that Miller-amplifiers shall be designed in weak or moderate inversion aiming low-power and high gain. Regarding the Figures 6.2-6.8 (see Section 6.2) obtained in the transistor level characterization, we observe that the required bias will impose severe constraints to transistor variability and almost negligible ageing degradation. Therefore, we decided to analyze the failure conditions of the amplifier only in terms of transistor variability.

4.3.4 PGA Failure Analysis

Finite DC Gain Using the minimum finite DC gain specified in Table 4.6, we found the sensitivity of the worst case of error ε^3 to the DC gain as

$$S_{\varepsilon^3}^{G_{DC}} = 0.003. \tag{4.45}$$

This too low sensitivity means that even a big amount of G_{DC} variation; the worst case error will suffer from a negligible variation. According to Equation (4.40), G_{DC} is a function of transistor *gm* and so the worst case is if the $\Delta G_{DC}/G_{DC}$ =50 % (see Figure 6.8). Using such conditions of G_{DC} and ΔG_{DC}, we estimate the $\Delta\varepsilon^3/\varepsilon^3 = 15$ %. Therefore, the finite DC gain variation is not capable to induce a PGA failure by variability effect.

Finite Gain-Bandwidth Product The finite gain-bandwidth product variation will change the position of the undesired pole combined with the finite DC gain variation. Analyzing the undesired pole variation, we found the sensitivity

$$S_s^{G_{DC}} = -9.9\,10^{-4}, \text{ and } S_s^{GBW} = 1; \tag{4.46}$$

using the performances specified in Table 4.6. The finite DC gain sensitivity is small and so negligible facing the gain-bandwidth product sensitivity. According to Equation (4.42), *GBW* is function of gm_1 and so the worst case is if the $\Delta GBW/GBW$ =50 % (see Figure 6.8). Using such conditions of *GBW* and ΔGBW, we estimate the $\Delta s/s = 50$ %. Therefore, the ΔGBW failure specification is directly connected to the *gm* variability and yield requirements.

Linearity The linearity of the PGA is dominated by the passive components, which are reconfigurable, and thus the linearity failure may be mitigated with a new gain programming. The amplifier is dependent of the differential pair *gm* linearity. According to the discussion developed in the reliable-BLIXER design, we know that the *gm* linearity will increase with the ageing stress. Regarding the *gm* variability, such variation may impose a linearity reduction. Therefore, the amplifier linearity failure is only dependent of the *gm* variability and, moreover its impacts on the PGA's linearity shall be negligible. That's why we did not do a complete analysis of the PGA's linearity.

Noise The noise of the PGA is a function of the passive components and the amplifier input referred noise (presented in Equation (4.44)). In order to estimate the noise variation, it is required a complete design of the PGA. Using previous results and the failure analysis based on the transistor level characterization, we decided to do not design the PGA and its amplifiers. In fact, the PGA's noise variation caused by the transistor-parameter variation will follow the same trends presented in the reliable-BLIXER design.

4.3.5 Reliable-PGA Design Conclusions

In this section, we analyzed the PGA and the required elements to design a reliable-PGA. Regarding the design constraints, the reliability analysis and synthesis methodology are able to point that the PGA is naturally immune to the ageing as the BLIXER was presented. Thus, the circuit variability will be the most important agent of performance variation comparing the ageing and the variability characterization proposed in Section 6.2. Hence, the PGA reliability is directly controlled by the reliability of the control circuits. Such circuits are digital circuits implemented for the gain programming reconfiguration and the reliability of digital circuits is out of the scope of this work.

CHAPTER 5

RELIABLE RF FRONT-END: TOP-DOWN DESIGN

In this chapter, we will present the design of a reliable RF front-end using a top-down approach. During Section 1.2, we discussed the architecture topology and its specifications. Now, the architecture will be modeled and analyzed using the top-down design equations and sensitivity analysis.

We propose a lower-level characterization obtained during design characterization at the bottom-up approach. The building block failure hypothetically defined will indicate the architecture failure when combined faults occur. This example will reveal possible fault masking and verify the sensitivity analysis.

Later, the RF front-end architecture will be designed with the reliability analysis and synthesis improvements. Then, the building-blocks specifications will be presented using two different reliability improvement strategies. Using the building-blocks specifications, the architecture will be simulated at behavioral level. The simulation results will be discussed and the architecture performance variation analyzed.

5.1 RF FRONT-END MODELING

Circuit modeling is the process of representing real-world circuit behavior using sets of mathematical equations. Modeling can be described in many levels; the most important among them are electrical modeling and behavioral or mathematical modeling.

Electrical model requires physical knowledge of the elements that comprise the real system, their constitutive relations, and the theoretical rules describing their interactions. Moreover, the electrical model shall represent fitted measurements, variability and ageing. These types of modeling are appropriate for transistor level simulation, and they have a trade-off between accuracy and simulation time.

Behavioral model, also called black box model, does not require prior knowledge of the physics. Its internal structure connects only input and output ideal values. The parameters of behavioral models identify the specified or estimated block characteristics. If the block specification is simulated it is possible to validate a higher-level characteristic. If it comes from estimated block characteristics, the measurement techniques and the quality of data affect the accuracy of the models. The behavioral model objective is often to speed-up the simulation time, moreover, if it is used with electrical model in a mixed simulation.

At this point, we are concentrating efforts in behavioral modeling to speed up the reliability analysis and synthesis using a top-down approach. Thus, we will consider variability and ageing as causes of RF front-end failure.

5.1.1 BEHAVIORAL MODELING

Most of building-block behavioral models [26] are available for the implementation of an RF front-end architecture. Each block has a typical set of characteristics which should be enough to represent the circuit-level behavior. However, the available models do not dispose the set of characteristics needed by the implementation of the reliable RF front-end architecture. For this purpose, we present some details

to implement each building-block behavioral model. The full VerilogA codes are available in Appendix A.

BLIXER Aggregating balun, wide-band low-noise amplifier, and I-Q mixer; the output current assignment of the block is represented as

$$I(Ip) = \frac{1}{\omega_c R_{out}} \frac{dV(Ip)}{dt} + \frac{V(Ip) - \frac{V_{RF}}{V_{LO}} V(cos)}{R_{out}}, \tag{5.1}$$

where $I(Ip)$ is the I-plus output current. The other output equations are similar. The gain and the linearity characteristics are defined by the circuit gain and IP3, and they are implemented in a conditional statement:

if (abs($V(RF)$)$< \sqrt{\frac{c_1}{3c_3}}$) **then**
  $V_{RF} = c_1 V(RF) - c_3 V(RF)^2$
else if ($V(RF) < 0$) **then**
  $V_{RF} = -\frac{2c_1}{3}\sqrt{\frac{c_1}{3c_3}}$
else
  $V_{RF} = \frac{2c_1}{3}\sqrt{\frac{c_1}{3c_3}}$
end if

The $V(RF)$ is the voltage of the node input RF node and c_i: gain of the output ith harmonic. The noise performance is defined by the circuit noise factor (F) and the input resistance (R_{in}). It is implemented by a white noise source with the power density equation

$$noiseDensity = \frac{4kT\,(F-1)}{R_{in}}, \tag{5.2}$$

where the k is the Boltzmann constant, and the T is the temperature in Kelvin.

DCO The in-phase sinusoidal output signal assignment of the DCO is

$$I(sin) = \frac{V(sin) - V_{off-set} - V_{amp}\sin(2\pi f_{LO} t + phaseNoise)}{R_{out}}, \tag{5.3}$$

and the circuit has four outputs delivering the in- and oppositional-phase for sine and cosine waves which are clocks signals delayed by 90 ^{o} spacing. The noise performance is defined by the amplitude thermal noise and the circuit phase noise ($L(f)$) by the equations

$$k_b = 10^{noiseFloor/10},\ noiseFloor \text{ is in dB} \tag{5.4}$$

$$amplitudeThermalNoiseDensity = \frac{V_{amp}^2 k_b}{2} \tag{5.5}$$

$$\Delta = 10^{L(f)/10} - k_b \tag{5.6}$$

$$k_w = \sqrt{\frac{\Delta}{\frac{1}{f^2} + \frac{f_c}{f^3}}},\ \text{the circuit phase white noise modelling} \tag{5.7}$$

$$k_f = \sqrt{f_c} k_w,\ \text{the circuit phase flicker noise modeling} \tag{5.8}$$

$$freqNoise = \textbf{whiteNoise}(V_{amp}^2 k_w^2) + \textbf{flickerNoise}(V_{amp}^2 k_f^2, 1) \tag{5.9}$$

$$phaseNoise = \frac{\sqrt{2}}{V_{amp}} 2\pi \int freqNoise\, d\theta \tag{5.10}$$

PGA For the PGA together with the low-pass filter, the I-plus output signal assignment of the block is

$$I(Ip) <+ \frac{1}{\omega_c R_{out}} \frac{dV(Ip)}{dt} + \frac{V(Ip) - V_{I_{out}}}{R_{out}}, \tag{5.11}$$

where ω_c is the low-pass filter cut-off frequency in rad/s. The gain and the linearity characteristics are defined by the circuit gain and IP3, and they are similarly modeled in the BLIXER output assignment (Equation (5.1)) and the conditional statement presented before. The noise performance is defined by the circuit NF, and it is similarly modeled in the BLIXER noise Equation (5.2).

5.1.2 Model Equations

Among frequency, bandwidth, power consumption, noise, linearity and gain characteristics; we will focus on gain, noise, and linearity. Since frequency and bandwidth are often controlled by a reliable reference, and the circuit degradation does not imply a worse power consumption as its reduction is desirable. Moreover, the power consumption decrease can be used as a measure of the degradation in the architectural level, as we presented in [28]. Thus, the equation of the power consumption is also important for the circuit analysis and will also be presented.

Power Consumption The total power consumption is a sum of all building-block power consumptions, as presented in

$$P = P_{BLIXER} + P_{DCO} + P_{PGA}, \tag{5.12}$$

and its total derivative is

$$\Delta P = \Delta P_{BLIXER} + \Delta P_{DCO} + \Delta P_{PGA}. \tag{5.13}$$

We have presented in [28] that a relationship exists between ΔP_i of the i building block and all others i building block characteristics. Thus, the measure of the power consumption degradation can be used as an estimation for all other characteristics degradations.

Gain The architecture gain is

$$G = G_{BLIXER} A_{LO} G_{PGA}, \tag{5.14}$$

where G_{BLIXER} and G_{PGA} are the gain of the BLIXER and of the PGA, and

$$A_{LO} = \frac{V_{LO}}{V_{LO}^{'}}. \tag{5.15}$$

In Equation (5.15), V_{LO} is the simulated DCO amplitude (degraded or not) and $V_{LO}^{'}$ is the expected DCO amplitude (not degraded). The total derivative of the gain performance is

$$\Delta G = A_{LO} G_{PGA} \Delta G_{BLIXER} + G_{BLIXER} G_{PGA} \Delta A_{LO} + G_{BLIXER} A_{LO} \Delta G_{PGA}. \tag{5.16}$$

Noise The architecture noise is measured by the $NF = 10 \log F$, and the total noise factor (F) is

$$F_{tot}(f) = F_{BLIXER} + \frac{F_{PGA} - 1}{G_{BLIXER}^2 A_{LO}^2} + \frac{N_{LO}(f)}{N_s} [89], \tag{5.17}$$

where $N_s = 4kTR_s$ is the noise of the input source and

$$N_{LO}(f) = \frac{R_s 10^{(P_s/10)-3}}{\left(\frac{S_{11_{BLIXER}}+1}{2}\right)^2} 10^{L(f)/10} A_{LO}. \tag{5.18}$$

In Equation (5.18), $S_{11_{BLIXER}}$ is the BLIXER S-parameter. The total derivative of F_{tot} is

$$\begin{aligned}\Delta F_{tot}(f) =& \Delta F_{BLIXER} + \frac{\Delta F_{PGA}}{G_{BLIXER}^2 A_{LO}^2} \\ &- \frac{2(F_{PGA}-1)}{G_{BLIXER}^3 A_{LO}^2}\Delta G_{BLIXER} \\ &+ \left(\frac{10^{((P_s/10)+(L(f)/10)-3)}}{4kT\left(\frac{S_{11_{BLIXER}}+1}{2}\right)^2} - \frac{2(F_{PGA}-1)}{G_{BLIXER}^2 A_{LO}^3}\right)\Delta A_{LO} \\ &- \frac{10^{((P_s/10)+(L(f)/10)-3)} A_{LO}}{4kT\left(\frac{S_{11_{BLIXER}}+1}{2}\right)^3}\Delta S_{11_{BLIXER}} \\ &+ \frac{10^{((P_s/10)+(L(f)/10)-3)} A_{LO} ln(10)}{40kT\left(\frac{S_{11_{BLIXER}}+1}{2}\right)^2}\Delta L(f).\end{aligned} \tag{5.19}$$

LINEARITY The architecture linearity is measured by the input referred interception point between the first and the third harmonic (IP3, in mili-Watts in the following equations), so it is

$$IP3 = \frac{1}{\frac{1}{IP3_{BLIXER}} + \frac{\left(1-S_{11_{BLIXER}}^2\right) G_{BLIXER}^2 A_{LO}^2 R_s}{4IP3_{PGA} R_{in_{PGA}}}} [89]. \tag{5.20}$$

The total derivative of $IP3$ is

$$\begin{aligned}\Delta IP3 =& \\ &\frac{1}{IP3_{BLIXER}^2\left(\frac{1}{IP3_{BLIXER}} - \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}R_{in_{PGA}}}\right)^2}\Delta IP3_{BLIXER} \\ &\frac{A_{LO}G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{2IP3_{PGA}R_{in_{PGA}}\left(\frac{1}{IP3_{BLIXER}} - \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}R_{in_{PGA}}}\right)^2}\Delta A_{LO} \\ &\frac{A_{LO}^2 G_{BLIXER} R_s\left(S_{11_{BLIXER}}^2-1\right)}{2IP3_{PGA}R_{in_{PGA}}\left(\frac{1}{IP3_{BLIXER}} - \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}R_{in_{PGA}}}\right)^2}\Delta G_{BLIXER} \\ &- \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}^2 R_{in_{PGA}}\left(\frac{1}{IP3_{BLIXER}} - \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}R_{in_{PGA}}}\right)^2}\Delta IP3_{PGA} \\ &\frac{A_{LO}^2 G_{BLIXER}^2 R_s S_{11_{BLIXER}}}{2IP3_{PGA}R_{in_{PGA}}\left(\frac{1}{IP3_{BLIXER}} - \frac{A_{LO}^2 G_{BLIXER}^2 R_s\left(S_{11_{BLIXER}}^2-1\right)}{4IP3_{PGA}R_{in_{PGA}}}\right)^2}\Delta S_{11_{BLIXER}}.\end{aligned} \tag{5.21}$$

5.2 BUILDING-BLOCK FAILURE ESTIMATION

The bottom-up design has presented the main blocks characteristics with reliability improvements. As presented in Section 1.2, the reliable-architecture shall agree with the multi-standard wireless applica-

tion specification [5]. Thus, we resume reliable-circuit characteristics for the main blocks: the BLIXER, the DCO, and the PGA. The BLIXER has:

- a limited 14 dB of maximum gain (G_{BLIXER}) ,
- an output bandwidth (BW_{BLIXER}) limited to 100 MHz (single pole filter model),
- a 3.5 dB of minimum noise figure (NF_{BLIXER}),
- a 1.1 dBm of 3rd harmonic interception point ($IP3_{BLIXER}$),
- a -15 dB of minimum and -10 dB of maximum input matching impedance ($S11_{BLIXER}$), and
- the mismatch between I and Q channels will not be modeled in this work.

The DCO has

- a 0.25 V of output amplitude (V_{DCO}),
- a -120 dBc/Hz of phase noise for 1 MHz (L(1 MHz)) with corner frequency at 40 kHz, and
- a 5 GHz of operational frequency (f_{DCO}).

We assume that an ideal PLL will control the oscillator in a fixed frequency point. The PGA has:

- a 20 dB programmable gain (G_{PGA}) by 2 bits,
- a 25 MHz of bandwidth (BW_{PGA}, single pole filter model),
- a 10 dB of NF_{PGA}, and
- a 10 dBm of $IP3_{PGA}$.

Using the architecture model equations presented in Subsection 5.1.2, we estimate the reliable RF front-end characteristics as:

- G = 34.1 dB,
- NF= 6.87 dB@5 MHz,
- $f_{RF} = 5.0$ GHz,
- $f_{LO} = 5.004$ GHz,
- $f_{IF} = 4$ MHz,
- IP3 =1.12 dBm, and
- S_{11} = -15.0 dB.

The RF front-end architecture failure was estimated using the degradation of the building block characteristics (see Chapter 4) and the architecture model equations (see Subsection 5.1.2). The calculated architecture characteristics degradation is:

- Δ Gain = -3 dB,
- Δ NF = 1.0 dB,
- Δ IP3 = -1.0 dBm, and
- Δ S11 = 5 dB.

5.2.1 Failure Definition

The building block failure margins are a function of the stress time and environment conditions, and they depend on the schematic used in the level below (transistor level). In order to simplify the analysis, we arbitrarily choose the failure margins in order to compare the building block sensitivities calculated analytically and a building-block failure simulation. The BLIXER block failure is defined by:

- ΔG_{BLIXER} = -3 dB,
- ΔNF_{BLIXER} = 1 dB,
- $\Delta IP3_{BLIXER}$ = 1 dBm, and
- $\Delta S_{11_{BLIXER}}$ = 5 dB.

The DCO block failure is defined by:

- ΔA_{LO} = - 0.3
- $\Delta L(1MHz)$ = 10 dBc/Hz, and
- Δf_{DCO} = 0 (by a ideal PLL).

The PGA block failure is defined by:

- ΔG_{PGA} = -3 dB,
- ΔNF_{PGA} = 5 dB, and
- $\Delta IP3_{PGA}$ = 1 dBm.

These defined failures are bigger than the maximum allowed by the architecture failure specification. So, we expect that such performances fall out of the specifications. These exaggerated variations are, however, a true building-block failure. For example, if an amplifier gain drops of 3 dB then it is interpreted as a cut-off frequency point and as the attenuation band. That is how these failures were defined; we looked for values which are always considered as a building-block failure.

5.2.2 Sensitivity Analysis

Using the model equations, we evaluate the sensitivity of all architecture characteristics for all building-block characteristics. The calculations were developed in symbolic mathematic software solving the equation

$$S_{\Phi_j}^{\phi_i} = \frac{\phi_i}{\Phi_j}\frac{\partial \Phi_j}{\partial \phi_i}, \tag{5.22}$$

and the model Equations (5.12)-(5.21). After that, the sensitivities were estimated with the design space presented in Section 5.2, and the hypothesis:

- T = 27 ^{o}C
- R_s = 50 Ω
- P_s = −60 dBm
- $S_{11_{BLIXER}}$ minimum
- $R_{in_{PGA}}$ = 100 kΩ

POWER CONSUMPTION The power consumption variation is not often considered a problem once the most common is a power consumption reduction after ageing degradation. Anyway, the power consumption sensitivities reveal important information about the circuit degradation. The power consumption sensitivities are

$$S_P^{P_{BLIXER}} = \frac{P_{BLIXER}}{P_{BLIXER} + P_{DCO} + P_{PGA}}, \tag{5.23}$$

$$S_P^{P_{DCO}} = \frac{P_{DCO}}{P_{BLIXER} + P_{DCO} + P_{PGA}}, \text{ and} \tag{5.24}$$

$$S_P^{P_{PGA}} = \frac{P_{PGA}}{P_{BLIXER} + P_{DCO} + P_{PGA}}. \tag{5.25}$$

Thus, the most sensitive building-block will be the block which has the bigger power consumption. In [4], the PGA and the low-pass filter are presented as the responsible of the most part of the power consumption budget. Therefore, the $S_P^{P_{PGA}}$ shall be the biggest; and a failure in the PGA shall be the most important to the architecture reliability.

GAIN In the gain model equation, the equal sensitivity of the gain for all parameters is evident. Thus, the gain sensitivities are

$$S_G^{G_{BLIXER}} = 1, \tag{5.26}$$

$$S_G^{A_{LO}} = 1, \text{ and} \tag{5.27}$$

$$S_G^{G_{PGA}} = 1. \tag{5.28}$$

Hence, it is not possible to predict the most important failure onto the architecture reliability, but we can expect that the PGA can always correct the signal swing before the ADC. Thus, a failure in PGA will result in losing this degree of freedom.

NOISE The noise sensitivity calculation leads to complex symbolic equations, but estimating the sensitivity values becomes easier to evaluate the variation consequences. The noise sensitivities are

$$S_F^{F_{BLIXER}} = 0.79, \tag{5.29}$$

$$S_F^{F_{PGA}} = 0.141, \tag{5.30}$$

$$S_F^{G_{BLIXER}} = -0.25, \tag{5.31}$$

$$S_F^{A_{LO}} = -0.17, \tag{5.32}$$

$$S_F^{S_{11_{BLIXER}}} = -0.0049, \text{ and} \tag{5.33}$$

$$S_F^{L(f)} = -2.22. \tag{5.34}$$

The almost zero influence of the input matching is clear, that is the sense of the $S_F^{S_{11_{BLIXER}}}$ value. The sensitivity absolute value measures the influence magnitude , for example $S_F^{G_{BLIXER}}$ and $S_F^{A_{LO}}$ will result in a more important variation. The sensitivity signal represents the sense of the variation, for example, as G_{BLIXER} and A_{LO} decreasing, the negative signal represent a resultant noise increase, and thus a noise performance degradation. Furthermore, a G_{BLIXER} and A_{LO} decreasing is often the case in ageing stress.

The $S_F^{F_{PGA}}$ is not so small and the PGA noise degradation may result in an architecture failure. Moreover, the BLIXER failure is an F_{BLIXER} increase and a G_{BLIXER} decreasing. Such combined gain and noise variations result in the noise degradation at high frequency signals (near BW). However, the combination of A_{LO} decreasing and $L(f)$ increasing in a DCO failure result in the noise degradation at low frequency signals (near zero if direct conversion). Therefore, if one of the both failures occurs the consequence may not be big enough to result in an architecture failure. But, both failures occurring simultaneously and one of them with a failure in the PGA shall be a catastrophic failure.

LINEARITY The transistor ageing results in linearity increasing, because the transistors will work in a more linear region. This variation may be not considered prejudicial, so that the sensitivity analysis might not point important information for performance variation. The linearity sensitivities are

$$S_{IP3}^{IP3_{BLIXER}} = 1.17, \tag{5.35}$$

$$S_{IP3}^{G_{BLIXER}} = -9.5 \cdot 10^{-4}, \tag{5.36}$$

$$S_{IP3}^{A_{LO}} = -9.5 \cdot 10^{-4}, \tag{5.37}$$

$$S_{IP3}^{IP3_{PGA}} = 4.7 \cdot 10^{-4}, \text{ and} \tag{5.38}$$

$$S_{IP3}^{S_{11_{BLIXER}}} = 8.1 \cdot 10^{-7}. \tag{5.39}$$

Regarding the sensitivity analysis values, we conclude that the BLIXER is the most responsible for the IP3 variation, as $S_{IP3}^{IP3_{BLIXER}}$ is the biggest sensitivity found. Comparing to previous considerations, the BLIXER is also the less linear building-block, and thus the BLIXER variability will be the most cause of losing architecture linearity.

5.2.3 ANALYSIS RESULTS

The RF front-end architecture, modeled in VerilogA, was simulated in SpectreRF (CADENCE simulator) for fresh-circuit test and degraded test cases. We propose to investigate the architecture performance degradation for eight test cases, composed by the cases: no failure, each block failure, the combinations of two blocks failure, and the failure of all blocks together. The test cases are summarized in the Table 5.1.

Table 5.1: Failure test cases applied to the RF architecture (Figure 1.2).

Case	PGA failure	DCO failure	BLIXER failure
no failure	no	no	no
test 1	no	no	yes
test 2	no	yes	no
test 3	no	yes	yes
test 4	yes	no	no
test 5	yes	no	yes
test 6	yes	yes	no
test 7	yes	yes	yes

The architecture gain degradation is shown in the Figure 5.1 with the no failure case in solid line, the BLIXER failure case with dashed line, the DCO failure case with dotted line, and the PGA failure case with square marks. The failures obtained from the combination of blocks are shown with the combination of failure representations (eg. the BLIXER and PGA failures are represented with dashed line and square marks). The test cases were referred in the Table 5.1.

The PGA failure has caused bigger gain degradation in test cases where it takes place than the other test cases without PGA failure. Thus, the PGA block failure has the biggest impact on gain degradation. Moreover, we lost the advantages introduced by this building block, as natural anti-aliasing filter and signal swing adaptation. A failure of a single block will not represent a gain under the specification, but it would be considered as a bandwidth reduction. In these cases, the architecture bandwidth will be under 10 MHz and depending the standard application, it could be under the specification or not.

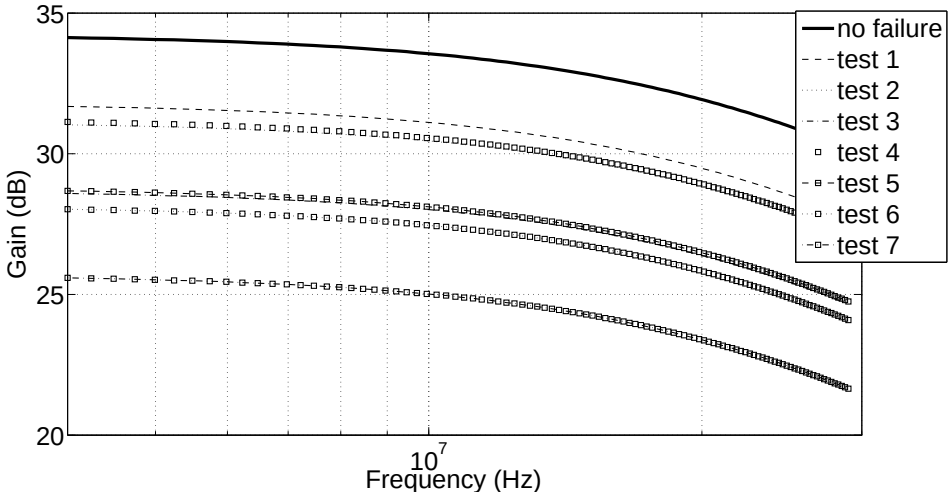


Figure 5.1: The architecture gain degradation: no failure case in solid line, the BLIXER failure case in dashed line, the DCO failure case in dotted line, and the PGA failure case with square marks.

The combination of BLIXER and DCO failures also reduces the gain under the specification, as any combination of failures with PGA failure.

The architecture NF degradation is presented in the Figure 5.2 with the no failure case in solid line, the BLIXER failure case with dashed line, the DCO failure case with dotted line, and the PGA failure case with square markers. The failures obtained from the combinations of blocks are shown with the combination of failure representations (eg. the BLIXER and PGA failures are represented with dashed line and square marks). The test cases were referred in the Table 5.1.

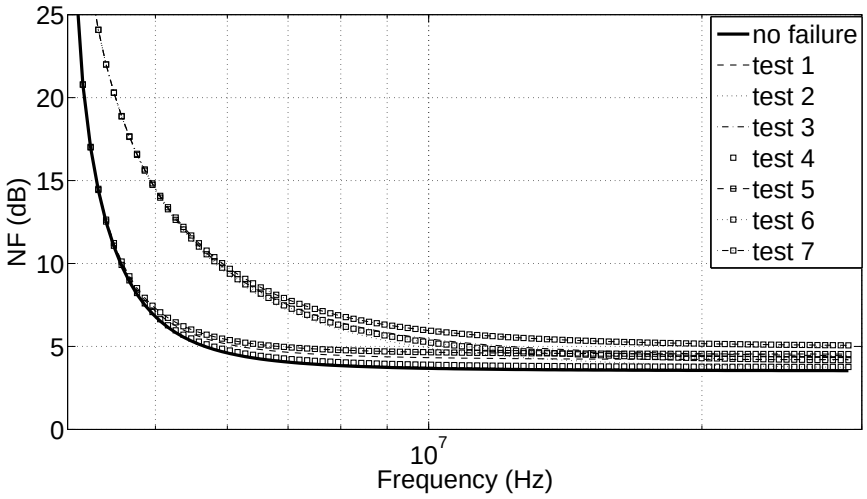


Figure 5.2: The architecture NF degradation: no failure case in solid line, the BLIXER failure case in dashed line, the DCO failure case in dotted line, and the PGA failure case with square marks.

The BLIXER failure has caused bigger noise degradation in test cases, where it takes place, than the other test cases without BLIXER failure. Moreover, if there is another failure combined, the noise degradation is increased as predicted in the Equation (5.19). The result presented in the Figure 5.2 also confirms the dominant sensitivity of the architecture's NF by the DCO in lower frequency signals.

Thus, the architecture NF reliability is also a big constraint in single failures test cases, if low-frequency performance is required. However, only the test 7 (all building block failures) is able to overcome the maximum NF barrier of 5 dB for high-frequency signals. Thus, the building-block noise characteristic limitations are not big enough to represent an architecture failure in these high-frequency signal cases.

The architecture IP3 degradation is presented in the Table 5.2 for all test cases. The results present the BLIXER failure as the most sensitive block, increasing in $\Delta IP3 = 1$ dBm. Moreover, the IP3 increase does not characterize an architecture failure as a bigger linearity is desired. One trade-off not discussed in face of IP3 increase is the signal swing decrease. Therefore, the signal-to-noise and distortion ratio (SNDR) of the architecture will not increase with the linearity increase. What's more, we analyzed a noise increasing and thus the SNDR shall decrease when the building blocks fail.

Table 5.2: IP3 results of the failure test cases applied to the RF architecture (Figure 1.2).

Case	IP3 (dBm)
no failure	1.125
test 1	2.127
test 2	1.126
test 3	2.127
test 4	1.125
test 5	2.127
test 6	1.126
test 7	2.127

5.3 Reliable-Architecture Design

The RF front-end architecture characteristics and estimated failure are shared among the building-block characteristics specifications. Using a reliable-IC synthesis method, we would like to optimize the building-block performance improving the architecture reliability. According to a reliable-architecture design strategy, we may propose a building-block characteristic variation exploring the architecture performance limits.

Table 5.3 summarizes the first-design optimization for building-block characteristics and it also presents the estimation of the RF front-end architecture characteristics. These obtained values were influenced by the previous lower-level design. Thus, the design values represent a design knowledge reuse and not a new optimization EDA run. Actually, these lower-level information is not unknown if the architecture designer decides to use a known state-of-the-art architecture topology and standard circuit schematics. Hence, these results (Table 5.3) are a start point for the reliable-architecture synthesis.

Table 5.3: Optimized building-block performance and RF front-end performance estimation for the typical lower level characterization.

BLIXER	PGA	DCO	Architecture
G = 14 dB	G = 20 dB	V_{LO} = 0.25 V	G = 34.1 dB
NF = 3.5 dB	NF = 10 dB	$L(1\,MHz)$ =	NF=
		- 120 dBc/Hz	6.87 dB@5 MHz
$f_{RF} = 5.0$ GHz		$f_{LO} = 5.004$ GHz	$f_{IF} = 4$ MHz
IP3 =1.1 dBm	IP3 =10.0 dBm		IP3 =1.12 dBm
R_{in} = 53 Ω			S_{11} = -15.0 dB

The model equations presented in Subsection 5.1.2 are able to early estimate the architecture characteristics and their variation using a gradient formulation. This formulation is suitable if the involved variables are deterministic. Moreover, the gradient formulation is accurate only if the variation is lower than 10 %. Thus, the architecture variability degradation cannot be estimated with this formulation. A better estimator shall be the variance, if we can assume that the characteristics are statistical independent variables, as it is

$$\sigma_{\Phi_j}^2 = (\Delta\Phi_j)^2 \approx \sum_{i=1}^{n} (\Delta\Phi_{ij})^2, \tag{5.40}$$

where $\Delta\Phi$ is the estimated characteristic variation. The $\Delta\Phi_{ij}$ is the estimated contribution of the lower-level characteristic of the building block i into $\Delta\Phi_j$, as the equation

$$\Delta\Phi_{ij} = \frac{\partial\Phi_j}{\partial\psi_i}\Delta\psi_i, \tag{5.41}$$

where ψ_i is the characteristic of the building block i.

Regarding the existent architecture-characteristics variation, we shall include these variations as design margins. In order to optimize the required margins for all architecture characteristics, we may better share the estimated failure variation among the building block characteristics. The reliability improvement proposed is thus sharing the estimated failure as a design margin.

First, we shall define sharing weights being the reliability improvement sharing strategy. Sharing the estimated failure, we can give the bigger variation budget for the characteristic with a lower sensitivity device. Thus, if this characteristic fails, the influence on the architecture characteristic will be as smaller as possible. The opposite strategy is to give the bigger variation budget for the characteristic with a higher sensitivity device. Thus, we can be less severe with the specification of the high-sensitive devices and more severe with the low-sensitive devices. In this strategy, we assume that if a low-sensitive device fails the influence into the architecture characteristics will be negligible.

Regarding the variability-aware criterion, if we choose to give the bigger variation budget for a lower sensitivity device, then $S_{IP3}^{G_{BLIXER}}$, $S_{IP3}^{A_{LO}}$, and $S_{IP3}^{IP3_{PGA}}$ are the smaller sensitivities (shown in Equation (5.39)). They will impose the bigger ΔG_{BLIXER}, ΔA_{LO}, and $\Delta IP3_{PGA}$ according to the allowed $\Delta IP3$ for the architecture. However, these big variations will overcome the allowed ΔNF for the architecture, regarding the noise sensitivity analysis in Equation (5.34). Thus, $\Delta IP3_{BLIXER}$ will be imposed by the $IP3$ failure and shall be tiny, but $\Delta IP3_{PGA}$ shall be bigger.

What's more, the ΔNF budget will impose the ΔG_{BLIXER}, ΔA_{LO} variations once the gain sensitivity is equal to one for all parameters (see Equation (5.28)). The $\Delta S11_{BLIXER}$ given by the noise sensitivity analysis will be too big as the variation given by the linearity analysis, therefore the estimated architecture ΔS_{11} will impose the matching variation limit. The $\Delta L(f)$, ΔF_{BLIXER}, and ΔF_{PGA} will be tiny, once they have the bigger noise sensitivity (see Equation (5.34)). Hence, the ΔG variation will be transferred to ΔG_{PGA}.

This strategy favors the PGA characteristics variation and proposes more severe constraints for BLIXER and DCO characteristics variation. The advantage of this strategy is that PGA performance variation can be negligible exploring the circuit versatility and control. Thus, if the PGA fails then it is easier to manage this failure than a BLIXER or DCO failure. Or, the larger performance variation can be used to relax the constraints on the control circuits. The disadvantage is requiring a reliable-BLIXER and a reliable-DCO, and thus this strategy may impose some power consumption or die area trade-offs.

Regarding the nominal reliability criterion, the discussion about the variation share does not change. The difference between variability-aware and nominal reliability is only how the variation is mathematically treated. The variability-aware supposes that the performances are statistical variables ($\Delta\psi = x\sigma\psi$ and x depends on the desired yield, assuming a Gaussian distribution), and the nominal reliability supposes that the performances degrade only with the time. Thus, the variability-aware will conduct to two different scenarios which are: aiming building-block characteristics as close as possible to the RF

front-end architecture standard specifications, and relaxing the building-block characteristic constraints to an overdesigned RF front-end architecture. Avoiding the over design, the architecture designer shall propose the shared variance in the first scenario for the circuit designers. In our case, $x = 1$ is proposed in variation sharing representing a 68.2 % of yield, but an analysis with more standard deviations (e.g. $x = 3$ representing 99.6 % of yield) could be desired. Table 5.4 summarizes the optimized building-block characteristics limits for the first scenario and estimates the architecture characteristics in this design scenario.

Table 5.4: Optimized building-block characteristics and RF front-end characteristics estimation for the strategy criterion to give the bigger variation budget for a lower sensitivity device.

BLIXER	PGA	DCO	Architecture
G = 13.9 dB	G = 16.8 dB	V_{LO} = 0.24 V	G = 30.6 dB
NF = 3.51 dB	NF = 10 dB	$L(1\ MHz)$ = - 120 dBc/Hz	NF= 6.65 dB@5 MHz
f_{RF} = 5.0 GHz		f_{LO} = 5.004 GHz	f_{IF} = 4 MHz
IP3 =1.09 dBm	IP3 =10.05 dBm		IP3 =1.09 dBm
R_{in} = 61 Ω			S_{11} = -10.0 dB

Regarding the variability-aware criterion, if we choose to give the bigger variation budget for a higher sensitivity device, the sensitivity analysis will lead to symmetric results for each case. The linearity sensitivity analysis (shown in Equation (5.39)) will impose a bigger $\Delta IP3_{BLIXER}$ and a smaller $\Delta IP3_{PGA}$. The noise sensitivity analysis (shown in Equation (5.34)) will impose a bigger ΔNF_{BLIXER} and $\Delta L(f)$ and a smaller ΔNF_{PGA}. Both sensitivity analysis will conduct to a negligible $\Delta S11_{BLIXER}$, ΔG_{BLIXER} and ΔA_{LO}, but the gain sensitivity analysis (shown in Equation (5.28)) will be responsible for the ΔG_{PGA} variation.

This strategy is more severe only with the PGA gain variation, but it relaxes the BLIXER and DCO characteristics variation constraints. Thus, the BLIXER could be less linear and have more noise, as the DCO could have a worse phase noise caused by variability and ageing degradation, without exceeding the architecture performance specification. Thus, the PGA high linearity is mandatory and cannot decrease. However, designing a high linear PGA imposes less constraint in power consumption and die area than a high linear BLIXER. Furthermore, a relaxed BLIXER and DCO characteristics simplify the power consumption and die area trade-off.

Regarding the nominal reliability criterion, the sensitivity analysis conducts to a similar variation share. Thus, the choice between variability-aware and nominal reliability also remains in treat or not the parameters as statistical variables. There are two scenarios to be evaluated as explained before. One of such scenarios leads to overdesigning the architecture and the other to design building-block characteristics as close as possible to the RF front-end architecture standard specification. Therefore, the architecture designer should propose the shared variance in the optimized-performance scenario for the circuit designers knowing the desired yield (in this work, $x = 1$ representing 68.2 % of yield). Table 5.5 summarizes the optimized building-block characteristics limits for the first scenario and estimates the architecture characteristics in this design scenario.

5.4 Reliable-Architecture Simulation Results

The state-of-the-art presented in Chapters 2 and 3 indicates that some ageing phenomena are still under discussion and the post-event behavior are still being studied. CADENCE simulator and ELDO simulator have only HCI and NBTI post-event models and the available PDK for CMOS 65 nm only

Table 5.5: Optimized building-block characteristics and RF front-end characteristics estimation for the strategy criterion to give the bigger variation budget for a higher sensitivity device.

BLIXER	PGA	DCO	Architecture
G = 13.9 dB	G = 16.8 dB	V_{LO} = 0.25 V	G = 30.8 dB
NF = 3.84 dB	NF = 10 dB	$L(1\ MHz)$ =	NF=
		- 119.5 dBc/Hz	7.26 dB@5 MHz
f_{RF} = 5.0 GHz		f_{LO} = 5.004 GHz	f_{IF} = 4 MHz
IP3 =0.16 dBm	IP3 =10.05 dBm		IP3 =0.16 dBm
R_{in} = 53 Ω			S_{11} = -14.8 dB

supports the ELDO models. The available PDK for CMOS 65 nm does not have the option of an ageing simulation for each Monte Carlo simulation point. Thus, we cannot simulate the variability-aware reliability analysis in commercial tools. Actually, a variability-aware reliability analysis is still a new subject in the state-of-the-art and first discussions were published in [80] and [84].

We believe that in the near future, ageing simulation tools will be part of the design process in the same way that the variability simulation tools already are. The ageing, as the variability, could be placed in early stages and could be connected with the design criteria aiming a reliable architecture. In order to fill the lacks of the commercial tools and the available PDK for CMOS 65 nm, we will present the simulation results of the RF front-end behavioral implementation. Hence, we keep pace with the required variability-aware reliability analysis, considering variability and ageing as the causes of parameter and characteristics variation.

The RF front-end architecture behavioral model was implemented in VerilogA and was presented in Subsection 5.1.1. The RF front-end architecture was simulated in SpectreRF (CADENCE simulation) for the typical lower-level characteristics and the optimized building-block characteristics obtained with reliable-architecture design (presented in Tables 5.4 and 5.5). For simulation purposes, we chose the oscillation frequency $f_{LO} = 5.004$ GHz and the low-IF frequency $f_{IF} = 4$ MHz. In this way, we made the hypothesis that all building blocks were characterized in this frequency condition and it will not be degraded, because it is often controlled by a reliable frequency reference. The simulation supposes the RF input signal around 5 GHz and the base-band output will be around the f_{IF}.

Architecture gain degradation is shown in the Figure 5.3. The classical designed RF front-end (classic design) using the typical building blocks characteristic is represented in solid line. The first reliable design strategy (strategy 1) using Table 5.4 optimized building blocks characteristics is represented in dashed line. The second reliable design strategy (strategy 2) using Table 5.5 optimized building blocks characteristics is represented in dashed-dotted line. The architecture gain was a little underestimated in comparison to Tables 5.4 and 5.5) about 0.5 dB. The simulated gain result (in Figure 5.3) does not compromise the RF front-end architecture specification and the proposed strategies reduce the design margin about 1 dB. Such a design margin reduction leads to power consumption reduction, avoiding an overdesign. However, the gain performance crosses the 30 dB limit under 25 MHz required bandwidth. Hence, if a high gain is desired, the PGA gain limitation will result in a bandwidth limitation.

Architecture noise degradation is shown in the Figure 5.4. The classical designed RF front-end (classic design) using the typical building blocks characteristic is represented in solid line. The first reliable design strategy (strategy 1) using Table 5.4 optimized building blocks characteristics is represented in dashed line. The second reliable design strategy (strategy 2) using Table 5.5 optimized building blocks characteristics is represented in dashed-dotted line. The NF for the typical building blocks characteristics is 6.9 dB at 5 MHz. The NF for design strategy of Table 5.4 is 6.8 dB at 5 MHz. The NF for design strategy of Table 5.5 is 7.3 dB at 5 MHz. This result was predicted using the analytic analysis in the design methodology. We observe that the NF at lower frequencies is dependent on RF input power and DCO phase noise characteristic. Thus, the DCO phase noise limitation will degrade the NF at lower frequencies and could be considered a noise performance limitation.

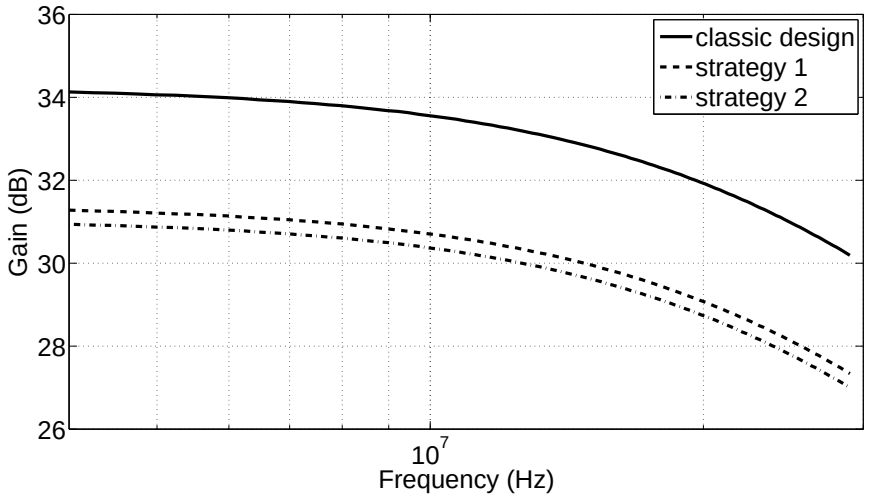


Figure 5.3: The architecture gain degradation results: the typical building blocks characteristics in solid line, the design strategy of Table 5.4 in dashed line, and the design strategy of Table 5.5 in dashed-dotted line.

Architecture linearity degradation is evaluated using the IP3 simulation result. The IP3 for the typical building blocks characteristics is 1.12 dBm. The IP3 for the design strategy of Table 5.4 is 1.64 dBm. The IP3 for the design strategy of Table 5.5 is 0.187 dBm. This result was predicted using the analytic analysis in the design methodology. We observe that the linearity does not impose performance limitation. This result shows no linearity performance failure. However, we clarify the influence of the $IP3_{BLIXER}$, and thus the BLIXER could impose linearity limitations.

Architecture matching impedance was not simulated, because this result will not be different from the estimated during the design. Moreover, the $S_{11} = S_{11_{BLIXER}}$ assumed has been characterized for the simulation frequency conditions. Using a behavioral model, the analysis concludes that the architecture matching impedance failure is conditioned to the BLIXER architecture matching impedance limitation. Such a limitation can only be estimated with a transistor-level model.

5.5 Reliable-Architecture Design Conclusion

In this chapter, we discussed the design of a reliable RF front-end, and the variation sharing strategies to avoid the overdesign. There are two important strategies highlighted using the sensitivity analysis results. The choice between both criteria is driven only by the statistical or deterministic characteristics of the parameter variation. Thus, the architecture design shall propose a feasible building-block characteristics specification, according the building blocks design experience and also avoiding the architecture overdesign.

Tables 5.4 and 5.5 present the building-block performance limits for a reliable architecture. The strategy difference between both designs is to give the bigger variation budget for the smaller or the bigger sensitive characteristic. Both designs are in the valid design space given by the lower-level characterization, but each one imposes different building-block design constraints. The results of this chapter were obtained assuming a yield coverage of 68.2 %. In order to increase the yield coverage to 99.6 % the architecture shall assume a 3 σ analysis. It shall be represented $\Delta\Phi = 3\sigma_\Phi$, and therefore the required characteristics variation shall be three times smaller than before.

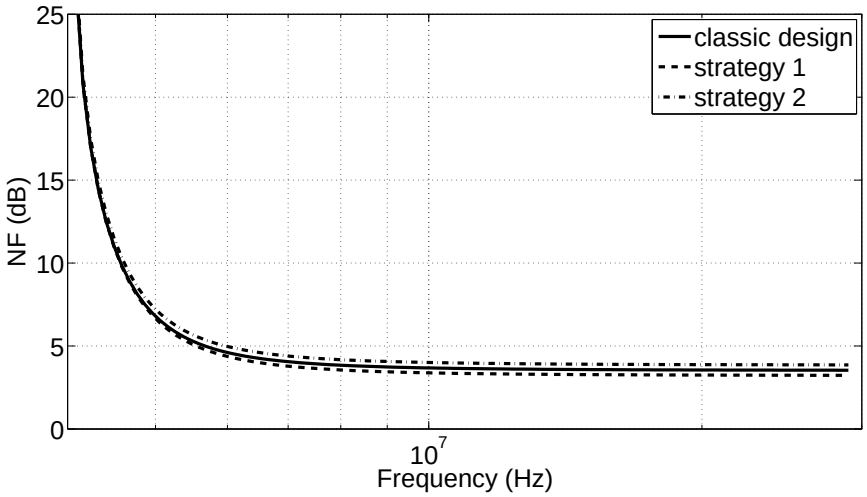


Figure 5.4: The architecture noise degradation results: the typical building blocks characteristics in solid line, the design strategy of Table 5.4 in dashed line, and the design strategy of Table 5.5 in dashed-dotted line.

CHAPTER 6

AMS/RF DESIGN FLOW FOR CIRCUIT RELIABILITY

6.1 INTRODUCTION

During the reliable RF front-end design, we notice some common steps in the AMS/RF design flow. At this point of the work, we can organize these steps in a general AMS/RF design flow for circuit reliability improvement linking bottom-up and top-down designs. Thus, we propose the design flow illustrated in Figure 6.1. After a general description of the design flow, next sections will present in details the modifications proposed in order to take the reliability into account during classical design methodologies. In this work, we do not implement a tool demonstrating the proposed design flow in an Electronic Design Automation (EDA) methodology. An EDA tool for circuit reliability is not our objective in this work, but we provide the proposed modifications details for an implementation as a future work.

The first element of the design is often defining the maximum or minimum desired performance and proposing them as the circuit specifications. The second element of the design is obtaining a description of the design space. This means associate the design options (like bias and sizing for transistor level or some state-of-the-art for building block characteristics) with the device characteristics. The design options will be referred as ψ and the device characteristics as Φ. Which is the best set of ψ for an optimum Φ? This question is answered in Figure 3.1, when the classical design methodology estimate Φ for a set of ψ described in the design space using an estimation model. After that, the optimizer tool evaluates only if the estimated Φ is the optimum.

Which is the best set of ψ leading to reliable design? To answer this question, we propose a failure evaluation which can define if the design is reliable or not using a new element: the design space variation. Actually, the set of ψ may change due to variability and ageing. Thus, the optimum Φ may not give a desired yield or reliability, not described in the classical design methodology. How the circuit reliability could be improved, if the design does not pass the failure evaluation? The answer of this question is modeling the Φ variation ($\Delta\Phi$) as a function of the ψ variation ($\Delta\psi$) and optimizing the design in terms of such variations.

The sensitivity analysis was used in bottom-up and top-down design approaches to identify the less reliable component of the device. The sensitivity analysis gives us for each component how much its characteristics variation influences the system characteristics variation and in which direction. In variability formulation, the direction information is suppressed as the characteristics are a probabilistic distribution often described by its mean and its standard deviation. In ageing formulation, the direction information is very important to show if a reduction in $\Delta\psi$ leads to a reduction in $\Delta\Phi$ in the case of a minimum desired characteristic for example. Thus, the direction information shall be compared with the ageing trend for each ψ case finding the worst case of $\Delta\Phi$, which leads to a fail in terms of maximum or minimum desired characteristic.

Knowing the variation of the characteristics ($\Delta\psi$ and $\Delta\Phi$), it is necessary to propose a new specification with some security margin and avoid the design space where the $\Delta\Phi$ is bigger than the margin. This opens the discussion of variation sharing strategies. Taking the estimated $\Delta\Phi$ as the initial security margin, the objective is proposing a set of $\Delta\psi$ for a new $\Delta\Phi$ always smaller than the initial margin. This

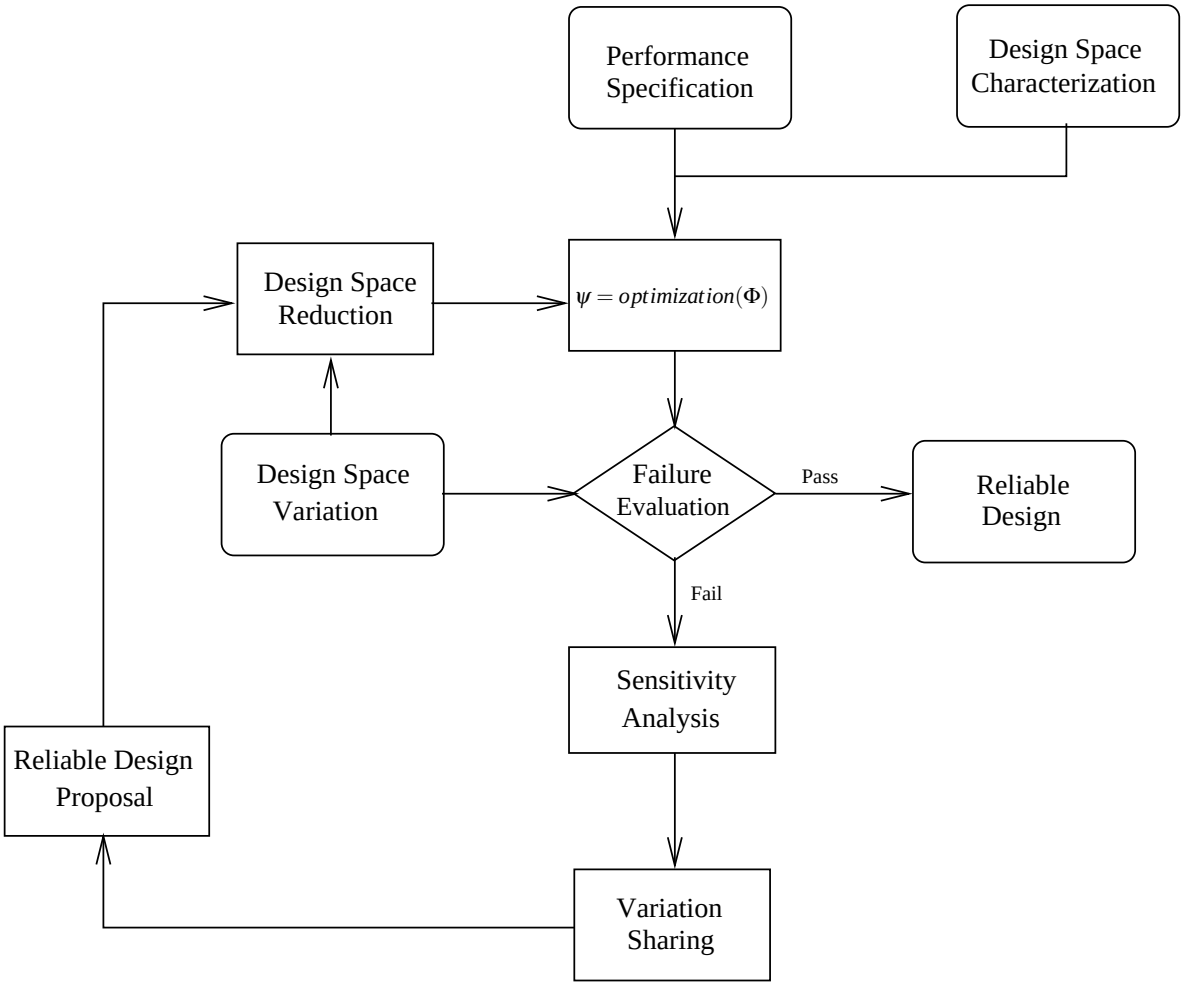


Figure 6.1: AMS/RF design flow for circuit reliability: design steps illustration.

problem will be formulated as a sharing problem and solved by the Games' Theory. The sharing problem solution can be an optimum or points of equilibrium. Assuming a solution, the designer takes some trade-offs among variability, ageing and the classical performance (e.g. die area, power consumption and speed).

The found solution in $\Delta\Phi$ will be traduced in a new reliable specification where the circuit characteristic is also optimum after some variability and ageing. The found solution in $\Delta\psi$ will be traduced in a design space reduction where the expected $\Delta\psi$ is always smaller than such solution. The reliable design specifications will feedback the optimizer with a coherent specification in terms of variability and ageing. Moreover, the reduced design space will lead the optimizer to choose a solution avoiding the variability and the ageing over the specified margin. Variability and ageing informations conduct the optimizer to converge to a new optimum design which has a higher reliability.

The optimization convergence will be now dependent of two decisions: the model of the characteristics variation and the variation sharing strategy. During the RF front-end design stages, we adopted a linear model based in the first derivative of the performance estimation model. In this case, we assumed that the statistical variables are independent for the case of the variability. We also assume that the ageing variations are smaller than 10 % as the region of convergence of the linear model. Using the design experience of reliable design, we can finally address performance, variability and ageing trade-offs. We will limit this discussion to the technology CMOS 65 nm, demonstrating these trade-offs with a ring oscillator characterization at Section 6.7.

6.2 Circuit Design Space

The first step to characterize the IC technology is extracting information about the design space. The design space in a bottom-up approach (transistor level) means how the sizing and the bias change the drain current (I_{DS}) and the transconductance (gm). Moreover, the design space shall identify ageing-stress conditions and variability influences on the two basic transistor characteristics (I_{DS} and gm). Thus, we are using the gm/ID methodology [90] as the base of the bottom-up design approach. The gm/ID methodology is a powerful sizing tool for low-voltage analog CMOS circuits. Moreover, most of methods using the design space as a criterion, characterize the design space with the basis proposed in the gm/ID methodology.

In our work, we characterized the CMOS 65 nm technology before start designing the circuit. This characterization is simulation expensive, but once did it, it would not be necessary again. Hence, this high cost is not representative after some designs using such IC technology. The first information extracted from the PDK is the I_{DS} and gm design space for a normalized transistor. We simulate a NMOS and a PMOS transistor for V_{DS} and V_{GS} bias varying from 0 to V_{DD} and for a size of W = 1 μm and L = 60 nm. The Figures 6.2 and 6.3 show the technology simulated-characterization for NMOS and PMOS respectively.

In the top-down approach, the circuit design space characterization can be obtained by a study of the state-of-the-art. The results will be very similar but in this case the target characteristics will be: die area, power consumption, gain, frequency and bandwidth, noise, and linearity. Thus, the design space in the top-down approach is strongly dependent on the application and the target architecture. The bottom-up design experience acquired in this work was used as the top-down design space, which is described in Chapter 4.

In this way, the design reuse will be the start point of the reliable architecture design. If that experience is not available, the design space should be satisfied with the published-research information of the building blocks often used in the target application. Thus, in both cases the designer could reproduce similar graphs like we present in transistor level. The design space characterization is additional information to the performance estimation model guiding the optimization.

In the bottom-up approach, the proposed reliable design flow also regards the ageing trends of I_{DS} and gm composing a design space variation for a normalized transistor. Thus, the ageing constraints impose a more complete characterization including the I_{DS} and gm degradations at the required circuit lifetime. The aged transistor was stressed by 30 years for all V_{DS} and V_{GS} bias varying from 0 to $1.1V_{DD}$ for a size of W = 1 μm and L = 60 nm. Moreover, the PMOS transistor was stressed at 150 ^{o}C putting in evidence the NBTI degradation. The Figures 6.4 and 6.5 show the aged NMOS and PMOS (respectively) degradation under HCI and NBTI stress. The Figure 6.6 completes the PMOS characterization showing the threshold voltage degradation due to NBTI only.

The information inside Figures 6.4, 6.5 and 6.6 means that bias and sizing cannot be taken only optimizing the transistor I_{DS} and gm. The required bias and sizing now shall include ΔI_{DS} and Δgm failure specification. Hence, the transistor ageing imposes new constraints reducing the design space if a failure specification is given. In a reliable-circuit synthesis point of view, the EDA optimization tool shall point the optimal transistor I_{DS} and gm including the reduced design space imposed by the reliability criterion.

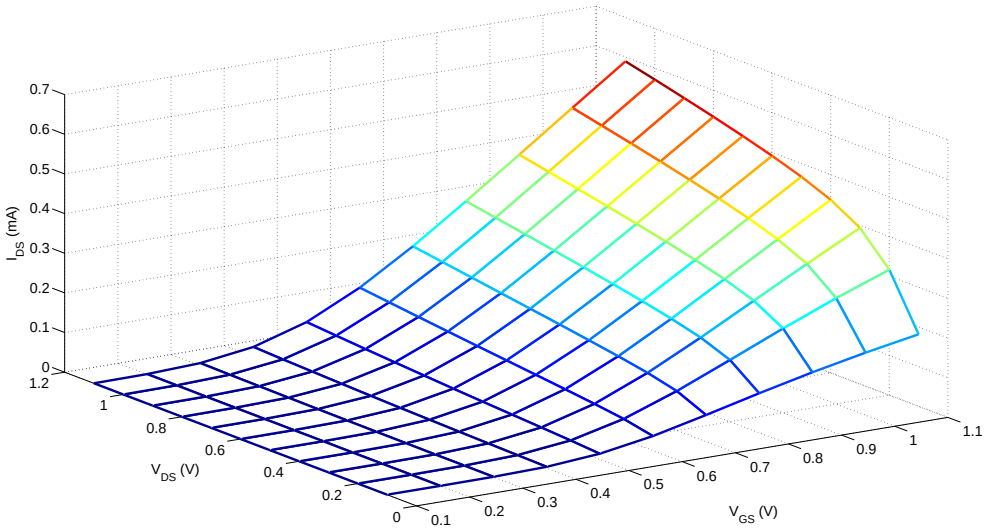
For a top-down approach, it is expected in near future more information in the state-of-the-art about the building blocks reliability. However, in this work we assumed some failure conditions for each building block in order to fill this lack. The failure hypothesis shall be coherent variations which are always considered as a block failure, like -3 dB in gain characteristics. In order to have a complete design space characterization, the building blocks shall be evaluated comparing different topologies. Hence, the top-down design space will divide the suitable building blocks topologies from the state-of-the-art for a desired reliability.

Moreover, the technology variability is increasing, and it should be taken into account. In bottom-up approach, the variability-aware design requires a characterization including the I_{DS} and gm variability.

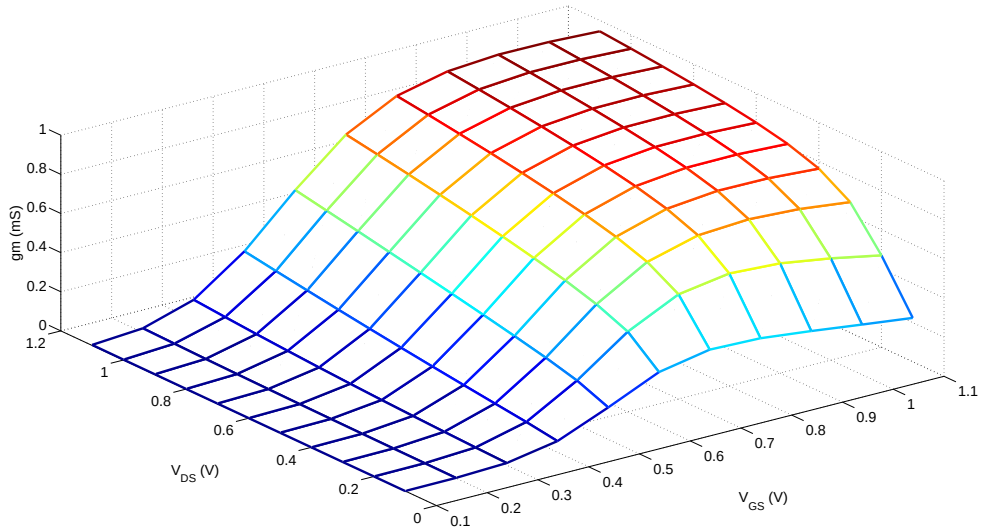
Thus, the information extracted from the PDK shall be the I_{DS} and gm mean and the design space should include the σI_{DS} and σgm informations. The Figures 6.7 and 6.8 present the percentage of the σI_{DS} and σgm over I_{DS} and gm mean, respectively. The NMOS results are presented in solid line and the PMOS results in dashed line.

In the top-down approach, the variability will appear in the standard deviation of the building blocks often characterized using Monte Carlo simulations. There are a set of known solutions to improve building blocks variability. The most applied is the double balanced and full differential topologies combined with strong layout techniques. Such solutions are responsible for reducing the IC process variation consequences, but they do not solve the mismatch variation. In this case, a higher circuit yield will be traded by increasing die area and power consumption.

The research of design space in the top-down approach suggests a characterization of different building blocks in terms of performance, variability and ageing. The reliable building blocks will have a smaller characteristics variation and will present some trade-offs often in die area and power consumption. In the state-of-the-art, the smaller variation is obtained with reconfigurable circuits, redundant paths, and error correction techniques mostly in digital domain. We will not present a complete research of design space for a top-down approach as it is out of the scope of this work.

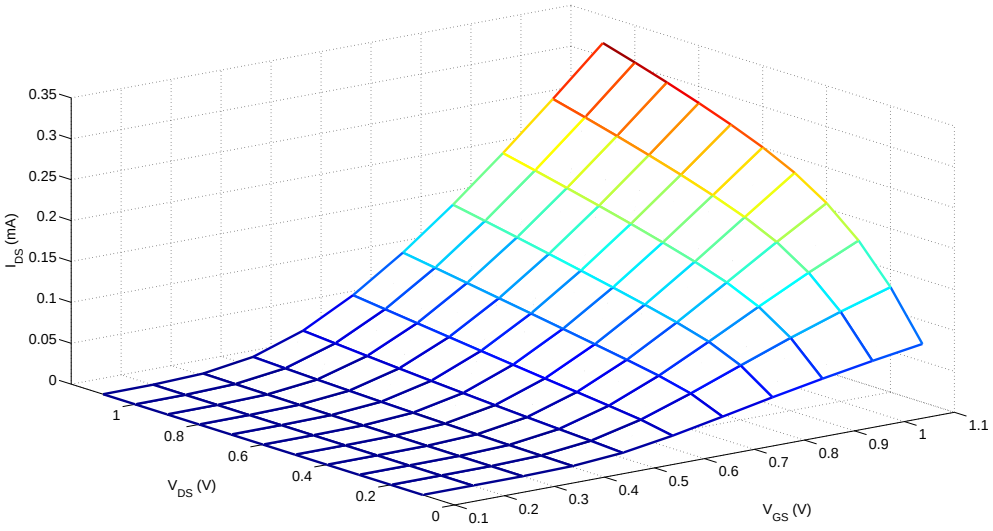


(a)

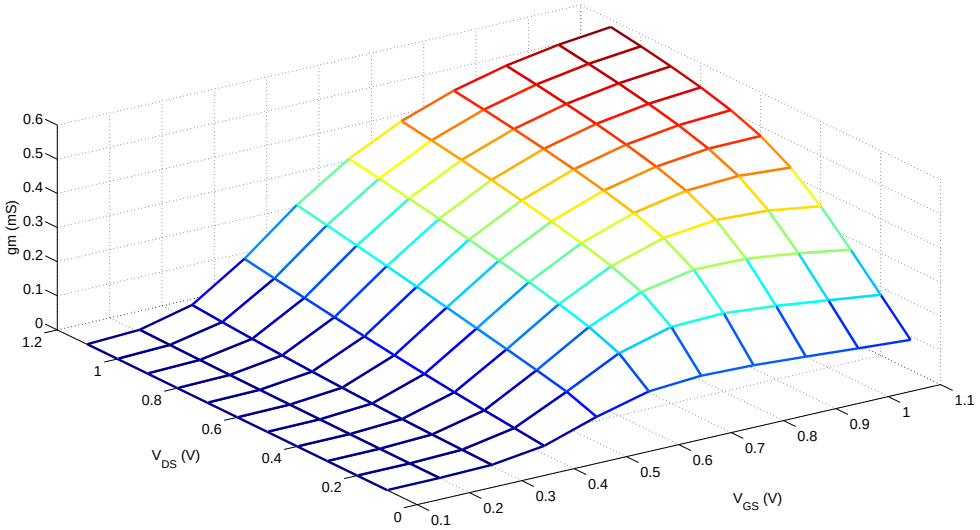


(b)

Figure 6.2: The technology simulated-characterization for normalized NMOS (W = 1 μm and L = 60 nm) biased from 0.1 V to 1.1 V for the I_{DS} (a) and the *gm* (b).

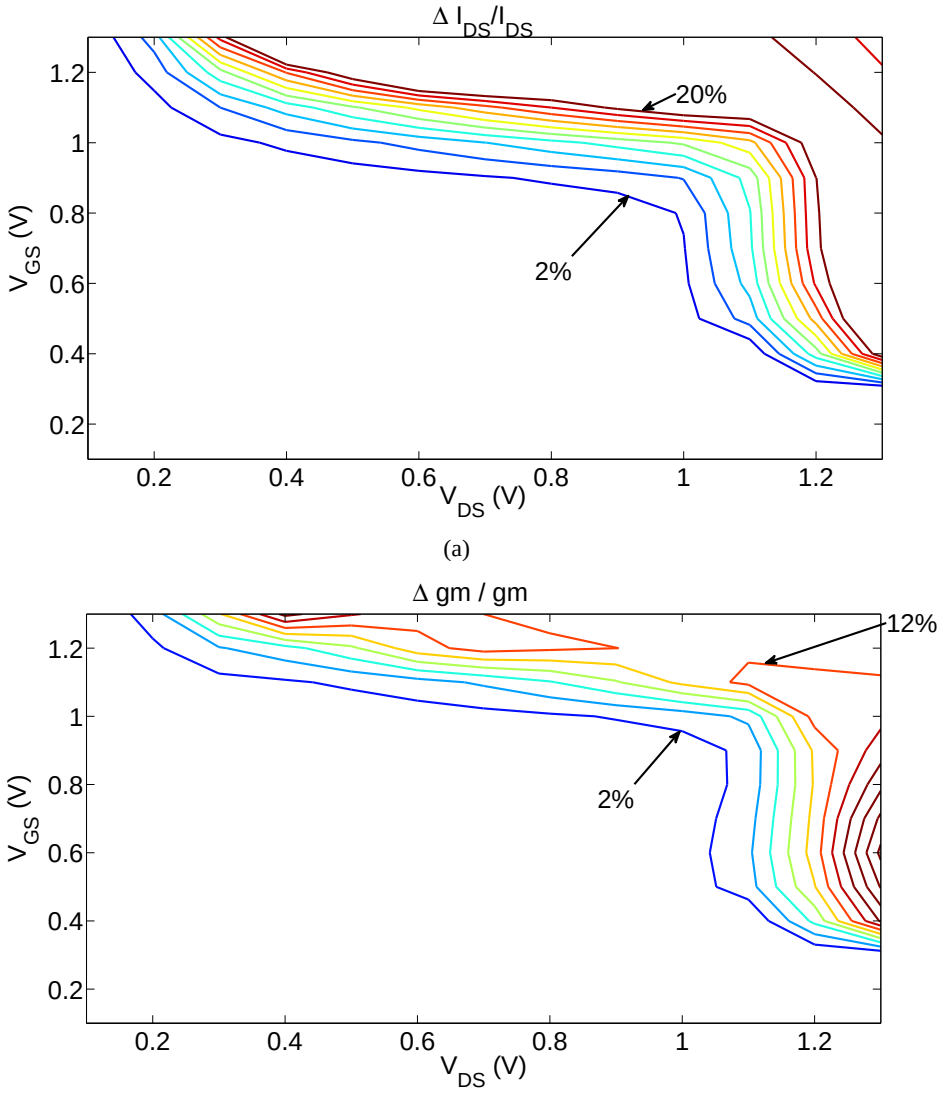


(a)



(b)

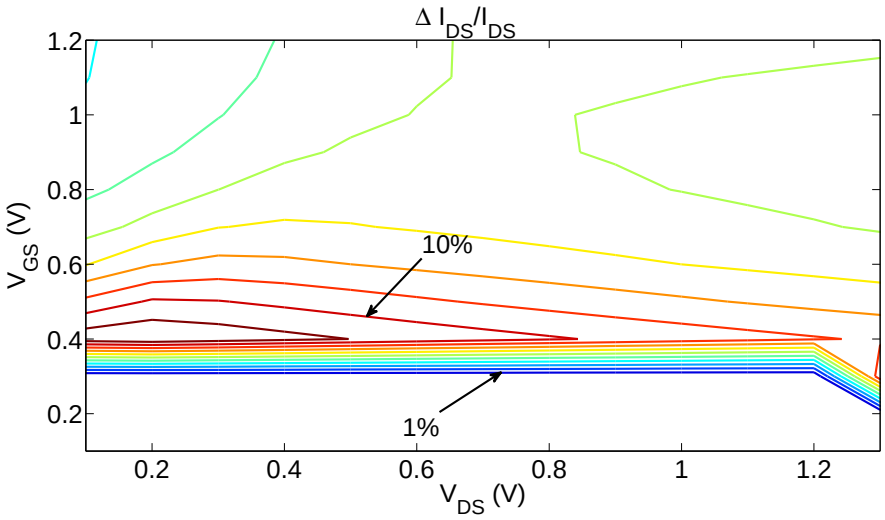
Figure 6.3: The technology simulated-characterization for normalized PMOS (W = 1 μm and L = 60 nm) biased from -0.1 V to -1.1 V for the I_{DS} (a) and the *gm* (b). The PMOS bias voltage is represented using its absolute value helping the comparison between NMOS and PMOS results.



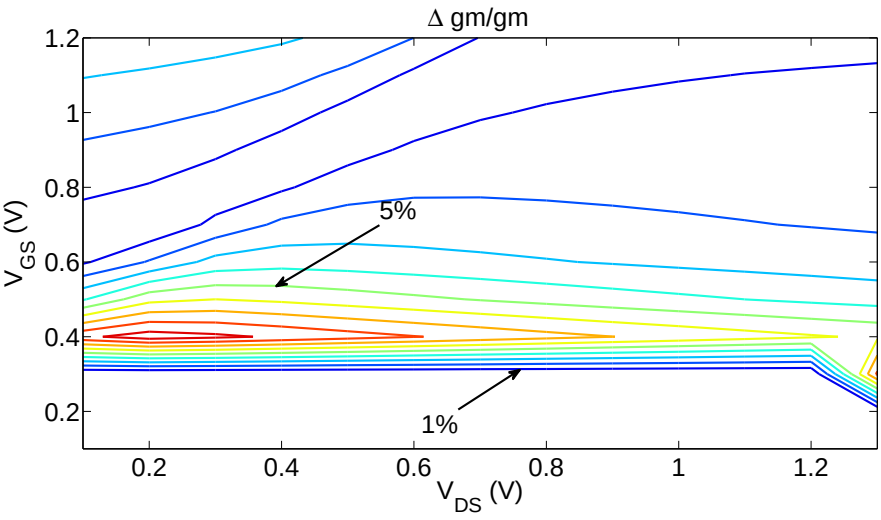
(a)

(b)

Figure 6.4: Ageing of the I_{DS} (a) and the *gm* (b) of the normalized NMOS (W = 1 μm and L = 60 nm).



(a)



(b)

Figure 6.5: Ageing of the I_{DS} (a) and the *gm* (b) of the normalized PMOS (W = 1 μm and L = 60 nm).

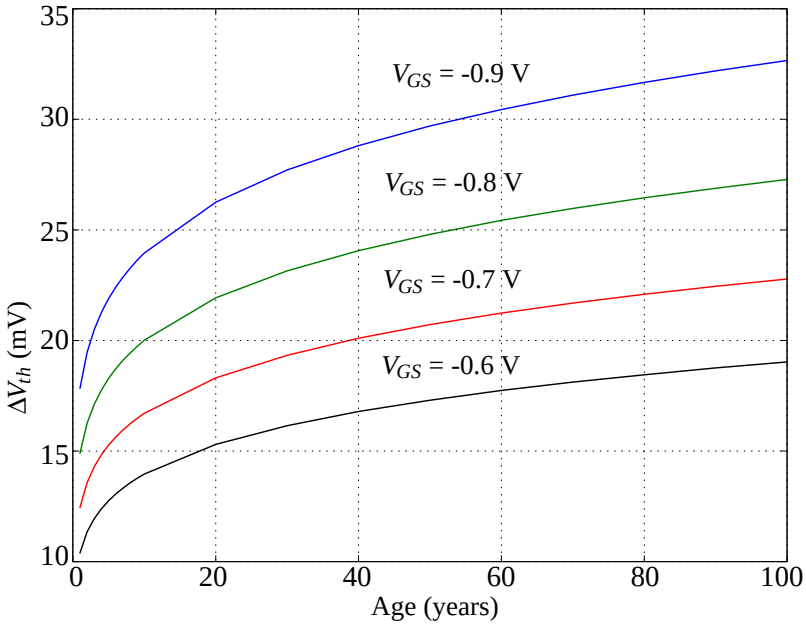


Figure 6.6: ΔV_{th} degradation for the PMOS DCO transistors stressed by NBTI at 150 ^{o}C.

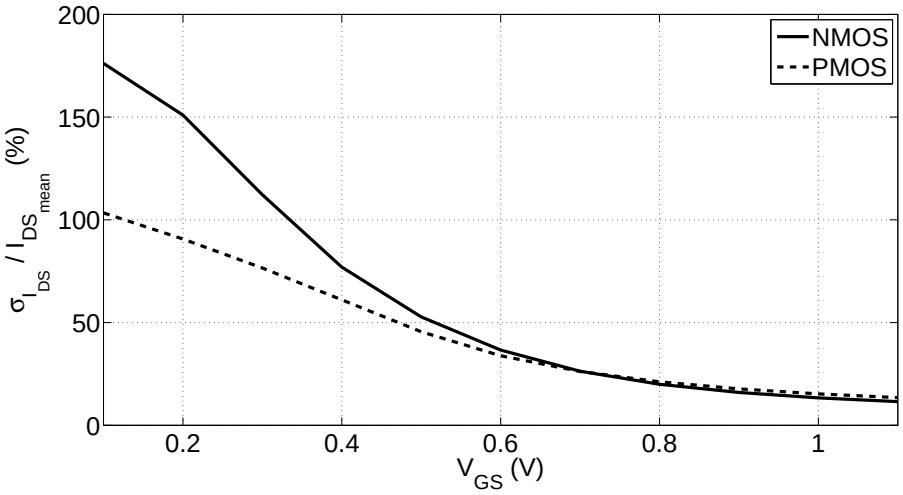


Figure 6.7: The percentage of $\sigma I_{DS}/I_{DS}$ for the NMOS (in solid line) and the PMOS (in dashed line).

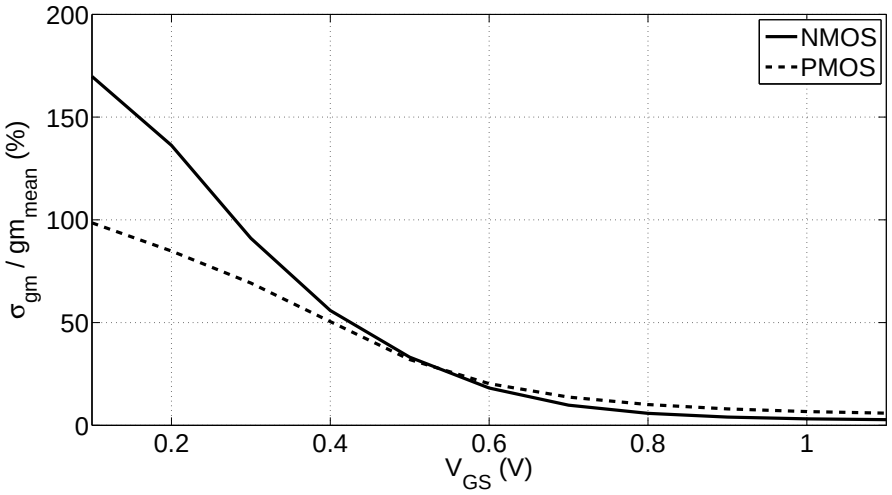


Figure 6.8: The percentage of $\sigma gm/gm$ for the NMOS (in solid line) and the PMOS (in dashed line).

6.3 FAILURE EVALUATION

During a classical design, the optimization aims the set of ψ which leads to Φ optimum. This means find the $\Phi_{typ} = \Phi_{spec}$ and consider that Φ does not change with variability and ageing. However, the sub-nanometer technologies has proven an increasing variability and ageing (see Chapter 2). The truth is the optimum Φ will fail the specification if it does not include a reasonable margin to take its variation into account. Actually, the design experience has led to an often overdesigned margin comprehensible in transistor level, but not exploitable in architecture optimization. This point of view highlights the need of an early estimation of the failure and feedback mechanisms conducting the optimizer to an also reliable design.

To estimate the failure, the state-of-the-art has proposed variability and ageing models. Most of them are electrical simulation based estimation, and then they are not suitable for an optimization (as explained in Chapter 3). One advantage of the simulation based estimation is the accuracy, often paid with long time analysis which gives us so little information about how we can achieve a better performance. Some new modeling proposals aim a behavioral model of the variation as similar as possible to the same characteristics model. Thus, the optimization effort for the characteristic variation will be similar to the own characteristics and both could be integrated in a single optimizer. However, this way out presumes that some circuits are designed and all variation data could fit the estimation model. Actually, these researches have mostly presented their behavioral model based in predictive IC technology variability and ageing models.

Both solutions are not suitable for a time-to-market reliable design, because so many efforts should be done to achieve the required information. Thus, we propose to innovate with an early estimation of the failure, reducing the accuracy and the convergence region, to design a reliable circuit at time-to-market. Moreover, we explore the top-down design proposing an increasing reliability for architectures, which was not mentioned before.

During the reliable RF front-end design, we faced small ageing degradations always under 10 %, high amount of IC process variability and lower mismatch variation. The solution is often to design full differential, and balanced circuits reducing the influence of IC process variability to always under 10 % variability due to mismatch variation. If we can consider all variables statistically independent, we can estimate the variation of a function using its derivative and the formulation of the variance of a quantity. If we consider nominal reliability degradation, which means neglect the variability; we can estimate the variation of the function using its derivative and the total derivative formulation.

Using the behavioral model estimator, the device characteristics $\Phi_j \forall j \in [1,m]$ are modeled by a function

$$\Phi_j = f(\psi_1, ..., \psi_n), \tag{6.1}$$

where $\psi_i \forall i \in [1,n]$ are the lower-level characteristics. If the lower-level characteristics changes with ageing or variability, the degraded device characteristics is estimated by

$$\Phi_{j_{degraded}} = f(\psi_1 \pm \Delta\psi_1, ..., \psi_n \pm \Delta\psi_n), \tag{6.2}$$

where $\Delta\psi_i \forall i \in [1,n]$ are the lower-level characteristics variation. Now, we assume $\Delta\Phi_{ij}$ as contribution of the $\Delta\psi_i$ at the characteristic variation (Φ_j). The variance of $\Phi_j \forall j \in [1,m]$ can be estimated by

$$\sigma^2_{\Phi_j} = (\Delta\Phi_j)^2 \approx \sum_{i=1}^{n} (\Delta\Phi_{ij})^2; \tag{6.3}$$

if we assume that the degradation $\Delta\psi_i$ represents less than 10% of the parameter ψ_i value and the set of ($\psi_i \forall i \in [1,n]$) are not correlated variables. In the case of nominal reliability, it is possible to estimate the $\Delta\Phi_j$ with a linear estimator defined by

$$\Delta\Phi_j \approx \sum_{i=1}^{n} \Delta\Phi_{ij}; \tag{6.4}$$

if we assume that the degradation $\Delta\psi_i$ represents less than 10% of the parameter ψ_i value. In both cases, $\Delta\Phi_{ij}$ is the contribution of $\Delta\psi_i$ into the Φ_j variation, defined as

$$\Delta\Phi_{ij} = \left.\frac{\partial\Phi_j}{\partial\psi_i}\right|_{\psi} \Delta\psi_i, \tag{6.5}$$

where ψ here is the solution obtained during optimization step.

Now, the failure estimation can redefine the device characteristics as

$$\Phi_{j_{degraded}} = \Phi_{j_{typ}} \pm \Delta\Phi_j. \tag{6.6}$$

If the design is reliable, then $\Phi_{j_{degraded}}$ is better than the maximum or minimum specified performance. The PGA is a design example of an optimum and reliable circuit, because we decided to use a full differential PGA reducing the variability. Moreover, the PGA ageing is negligible because the optimum design point is in a region of the space of negligible ageing degradation. What's more, the combination of both variations could be always adjusted by the digital control circuits with the reconfigurable components inside the PGA. However, some circuits can be vulnerable to characteristics variation. There are many circuits where the imposed characteristics are source of ageing stress and full differential solutions may be not suitable. For these circuits, the failure evaluation will indicate that the optimum design is not reliable. Hence, its reliability should be improved.

6.4 Reliability Improvement

In order to increase the circuit reliability, we should specify the desired characteristics taking into account its variations, and so solving the Equation (6.6) as

$$\Phi_{typ} = \Phi_{spec} \pm \Delta\Phi, \tag{6.7}$$

where $\Delta\Phi$ is the estimated variation. The reliability improvements will be achieved if the failure estimation is able to guide the optimization to a region of less variation. Moreover, the optimization will need a coherent Φ optimum agreeing with the estimated failure. The new Φ_{typ}, from the solution of the Equation (6.6) is the reliability improved design proposal.

How can we lead the optimization finally to a reliable design? Assuming we have no information about the circuit characteristics, the better option is to take

$$\Delta\psi_{ij} = \frac{\Delta\Phi_j}{n\frac{\partial\Phi}{\partial\psi_i}}, \forall i \in [1,n]. \tag{6.8}$$

This means share the Φ failure equally among the ψ failures, and estimate $\Delta\psi_i$ being the minimum in j of $\Delta\psi_{ij}$. Then, the optimization should look for a new optimum inside the space where ψ_i variations are always lesser than the estimated $\Delta\psi_i$. This simple solution will guarantee that the new design space has only reliable solutions. However, the equally $\Delta\Phi$ share is not often the best response.

The best response in $\Delta\Phi$ sharing shall be reusing the design experience by the definition of a sharing weights $W_i, \forall i \in [1,n]$. Hence, the $\Delta\Phi$ sharing can be described as a *Sharing Problem* and solved using *Games Theory* [91]. The *Sharing Problem* solution will be discussed during Section 6.5. Finally, the reliable design space will be obtained using a better estimated $\Delta\psi_i$ and a reliable solution will be evidently chosen.

6.5 Sharing Strategies

The variation sharing is an interactive decision process, which can be modeled and analyzed using a set of mathematical tools called Game Theory [91]. In variation sharing decision, we identify three primary components:

1. a set of $\psi_i \forall i \in [1,n]$, which are the lower-level circuit characteristics varying under variability and ageing influence;

2. a sharing strategy space composed by the ψ_i priority in variation sharing defined by the set of positive sharing weights W_{ij} ;

3. a set of utility functions deciding how much is $\Delta\psi_i \forall i \in [1,n]$, and defined by

$$\Delta\psi_i = \min_j \left|\Delta\psi_{ij}\right|, \tag{6.9}$$

where the variation allowed to ψ_i under the Φ_j criterion is

$$\Delta\psi_{ij} = \frac{\Delta\Phi_{ij}}{\frac{\partial\Phi}{\partial\psi_i}}, \tag{6.10}$$

and $\Delta\Phi_{ij}$ is the influence of $\Delta\psi_i$ in $\Delta\Phi_j$ depending the adopted strategy.

In a general variation sharing, the decision strategy is often choosing a characteristic $i = k$, giving it the highest priority. Then, it shares a bigger variation to this characteristic ψ_k. Next, it also shares a smaller variation to $\psi_i \forall i \in [1,n-1]$ for each Φ_{ij}. Thus, the W_{ij} shall be defined in a way that the highest priority characteristic (ψ_k) has

$$W_{kj} = \max_i W_{ij}. \tag{6.11}$$

The influence of $\Delta\psi_i$ in $\Delta\Phi_j$ is calculated by the Equation

$$\Delta\Phi_{ij} = \Delta\Phi_{ij_{max}} \frac{W_{ij}}{W_{kj}}. \tag{6.12}$$

The $\Delta\Phi_{kj}$ is the $\max_i \Delta\Phi_{ij}$ and is estimated under variability and ageing. If the design criterion is a **Nominal Reliability Analysis** (see Subsection 3.4.1), then we obtain a linear estimation solving Equation (6.4) as

$$\Delta\Phi_{kj} = \frac{\Delta\Phi_j}{\sum_{i=1}^{n} \frac{W_{ij}}{W_{kj}}}. \tag{6.13}$$

If the design criterion is a **Variability-Aware Reliability Analysis** (see Subsection 3.4.2), then we obtain a linear estimation solving Equation (6.3) as

$$\Delta\Phi_{kj} = \sqrt{\frac{(\Delta\Phi_j)^2}{\sum_{i=1}^{n} \left|\frac{W_{ij}}{W_{kj}}\right|^2}}. \tag{6.14}$$

The $\Delta\Phi_j$ is the estimated variation during the failure evaluation step of the reliable design flow (see Section 6.3).

The best available strategy for any characteristic $k \in [1,n]$ is the strategy that maximizes $\Delta\psi_i$ under the belief that all n characteristics do the same as well. This set of best strategies forms a Nash equilibrium [92]. Hence, the variation sharing decision at this equilibrium will always impose a smaller variation in Φ_j than the first estimation of $\Delta\Phi_j$ during failure evaluation. Moreover, this variation sharing decision converges to an optimal sharing reducing the required margin. The reliable design proposal in the proposed design flow presented in Figure 6.1 will be specifying this margin for a condition where the performance after variation is always better than the specification.

SHARING STRATEGIES EXAMPLES Identify the best strategy is proposing the best sharing priority in the variation sharing weights. The simple solution and not optimal solution is the equal share, where $W_{ij} = 1/n \forall i \in [1,n]$. This solution was presented in Equation (6.8). The best response should be reusing the design experience and proposing coherent W_{ij}. As any design decision, the W_{ij} definition will impose trade-offs among the classical performance, variability and ageing.

In order to better explain the variation sharing decision, we exemplify the reliable-circuit design in terms of a two-stage transistor-level amplifier. We evaluate only two amplifier's characteristics: gain and noise; being a function of both stages transconductance, as

$$G = g\left(gm_1, gm_2\right), \text{ and} \tag{6.15}$$
$$N = n\left(gm_1, gm_2\right). \tag{6.16}$$

The two sharing strategies for the gain variation are:

- giving the priority to gm_1 (SG_1) using

$$W_{SG_1} = [0.6, 0.4] \tag{6.17}$$

- giving the priority to gm_2 (SG_2) using

$$W_{SG_2} = [0.4, 0.6] \tag{6.18}$$

The two sharing strategies for the noise variation are:

- giving the priority to gm_1 (SN_1) using

$$W_{SN_1} = [0.9, 0.1] \tag{6.19}$$

- giving the priority to gm_2 (SN_2) using

$$W_{SN_2} = [0.1, 0.9] \tag{6.20}$$

The Equation 6.9 solution is represented by the ordered pair $(\Delta gm_1, \Delta gm_2)$. Assuming $\Delta G = 1$ and $\Delta N = 1$, we can calculate the characteristics variation using Equation (6.10), (6.12) and (6.13). Finally, we can represent the results of the variation sharing decision strategies using the Table 6.1.

Table 6.1: Variation sharing payoff defined by $(\Delta gm_1, \Delta gm_2)$, using the strategies in terms of gain and noise.

	SG_1	SG_2
SN_1	(0.6, 0.1)	(0.4, 0.1)
SN_2	(0.1, 0.4)	(0.1, 0.6)

The described *Sharing Problem* has two possible equilibria. The equilibrium assumes some trade-offs in the reliable-architecture design, which means:

- a strict design in gm_2 relaxing the gm_1 design (SG_1 and SN_1), or
- a strict design in gm_1 relaxing the gm_2 design (SG_2 and SN_2).

Thus, the best compromise shall be severe designing a building block naturally reliable and relax the design of the less reliable building block. However, point the reliability of the building blocks means design the reliable block going down to the transistor-level sizing and bias. Moreover, the number of characteristics is often bigger than the two presented; and these characteristics often present some dependency.

Define the W_{ij} is not a simple task. Such decision will involve a lot of interaction among the lower-level designers and the higher-level designers. Actually, $\Phi \in \Phi_j \forall j \in [1,m]$ are the circuit-characteristics like: die area, power consumption, gain, frequency and bandwidth, noise and linearity. Hence, this task will need a team effort finding all the W_{ij} which may be an over time-to-market solution, and so the best response cannot be used in early design steps.

In order to propose a better solution than the equal variation sharing, without a strong design experience defining W, we applied a first order sensitivity analysis in our design examples (see Subsection 5.2.2). Actually, the first order sensitivity analysis give us a reliable sharing weights if the variations are smaller than 10 % and if Φ_j defined by the model Equation (6.1) is accurate enough. Therefore, the sensitivity analysis result leads us to a better W even in a first design experience.

6.6 Sensitivity Analysis

A higher-level characteristic sensitivity to a lower-level characteristic is defined in a first order approximation as

$$S_{\Phi_j}^{\psi_i} = \frac{\psi_i}{\Phi_j}\frac{\partial \Phi_j}{\partial \psi_i}. \tag{6.21}$$

The component sensitivity contains the information of how much its variation influences the system variation. What's more, this information is normalized for the appropriate design point (Φ_j, ψ_i) under analysis.

The sensitive analysis has been proposed for AMS/RF circuit design a simple and strong tool to estimate the circuit characteristics variation trends. For example, Y Cheng and R. Fujii [93] proposed in 1992 a sensitivity analysis computation using circuit elements in symbolic format. Usually, the sensitivity analysis is applied in optimal placement of poles and zeros in the transfer function in filter design. The general idea is identifying the weak components represented by a high sensitivity and proposing improvements for these components. In this way, a smaller sensitivity is often achieved for such components.

There are two informations inside the value of $S_{\Phi_j}^{\psi_i}$. The first information is the sign of $S_{\Phi_j}^{\psi_i}$. Such sign indicates the direction of the Φ_j variation in comparison to a ψ_i variation. If a reduction in ψ_i will reduce Φ_j, so the sign of the sensitivity is positive. If an increasing in ψ_i will reduce Φ_j, so the sign of the sensitivity is negative. The second information is the magnitude of $S_{\Phi_j}^{\psi_i}$. Such magnitude measures how much the Φ_j is influenced by ψ_i.

Thus, the magnitude of the sensitive can be used as the first sharing variation weights. During the reliable RF front-end design, we proposed this W_{ij} definition to solve variation sharing in two strategies:

- giving preference to a lower $\left|S_{\Phi_j}^{\psi_i}\right|$ to have a higher $\Delta\Phi_{ij}$, so that

$$W_{ij} = \frac{\sum_{i=1}^{n}\left|S_{\Phi_j}^{\psi_i}\right|}{\left|S_{\Phi_j}^{\psi_i}\right|}, \forall i \in [1,n] \text{and} \forall j \in [1,m]; \text{ or} \tag{6.22}$$

- giving preference to a higher $\left|S_{\Phi_j}^{\psi_i}\right|$ to have a higher $\Delta\Phi_{ij}$, so that

$$W_{ij} = \frac{\left|S_{\Phi_j}^{\psi_i}\right|}{\sum_{i=1}^{n}\left|S_{\Phi_j}^{\psi_i}\right|}, \forall i \in [1,n] \text{and} \forall j \in [1,m]. \tag{6.23}$$

For each strategy, we decided for some criteria taking some constraints.

The sign of $S_{\Phi_j}^{\psi_i}$ is not an important information in a variability-aware design, because $\sigma\psi_i$ will influence $\sigma\Phi_j$ in both senses. Actually, a variability-aware design will define a worst-condition of $\Delta\Phi = x\sigma\Phi$ in the direction of the minimum or maximum performance specification. For example, if it is required a 99.73 % of yield then $x = 3$ in the case of a gaussian distribution.

However, the signal of $S_{\Phi_j}^{\psi_i}$ is very important in a reliability-aware design. With the $S_{\Phi_j}^{\psi_i}$ sign, we are able to eliminate the ψ_i ageing which do not influence in a worse Φ_j characteristic. In the reliable RF front-end for example, we used this information not to increase the circuit reliability in terms of power consumption. Actually, we always found decreasing power consumption and a positive sign for $S_{\Phi_j}^{\psi_i}$, and so such degradation will not lead to a worse power consumption expected to be bigger than the typical power consumption. We also removed the linearity from the ageing analysis, because the aged transistor is more linear and the sensitivity analysis sign is positive.

6.7 Variability and Ageing Trade-off

Designing a reliable circuit, we faced a compromise between the optimal characteristic and its variation. The convergence of the proposed design flow is dependent on the continuous reliability improvements and design space reduction. In lower-level design, the convergence is conditioned to the existence of an optimal characteristic in the design space where the transistor suffers less stress than some maximum stress. Such maximum stress causes the maximum characteristic variation allowed for the required lifetime and so the reliability specification. However, there is no guarantee that a solution exists. Thus, severe reliability requirements may lead to a deadlock during the reliability improvement iteration.

Using specific optimization tools, it is often possible to prove the existence of this solution or detect the unfeasibility during some iteration point. One idea is testing the existence of a solution before optimizing. If such a solution does not exist, the optimizer can identify why the solution is not feasible. Another way is proposing an always feasible optimization using as an example the posynomial modeling and the geometric programming. If such a solution does not exist, the model cannot be build and the unfeasibility is detected. More details in optimization tools, which are able to detect if a such solution is feasible or not, are presented in Subsection 3.2.2.

Indeed, detect the deadlock caused by a severe reliability requirement is not difficult. The problem is the deadlock diagnosis which means understand why the requirement is severe. Thus, we would be always able to propose a coherent specification for the trade-off: performance, variability and ageing. This means prefer a more reliable technology or a more reliable circuit topology when the achievable performance is suitable for the requirements. Therefore, there are two new questions to be answered:

1. How reliable are the known circuit topologies and architectures?

2. How reliable are the sub-nanometer IC technologies?

A complete reliability study of circuit topologies and architectures should be conducted using an automatic reliability analysis and design tool. Such a tool should implement the center of the proposed flow, which is the iterative failure evaluation and reliability improvement. For this future work, we contribute to presenting the physical phenomena influence into electrical circuit characteristics. The design experience demonstrated during the RF front-end design and reliability improvement may influence these future works.

The objective in future reliability works could be quantifying and increasing the reliability of circuit topologies and architectures. A state-of-the-art survey often compares performance trade-offs among the existent topologies positioning some device solution for a target application. The innovation proposed in this work will make possible such a survey comparing also the reliability among the existent topologies. Thus, some device solution will be suitable for a target performance and reliability requirements.

A complete reliability study of the sub-nanometer IC technologies should be conducted comparing the design space variation. Such a research should present the circuit-characteristics variation under

a stress environment coherent with the circuit operation conditions. For this work, we contribute to presenting the variation of the design space under the influence of variability and ageing. The usage of the characterization experience during the RF front-end design provides the required informations for the deadlock diagnosis.

The objective in future works also could be quantifying the device-characteristic variation in the operation environment conditions during the IC technology evolution. An IC technology survey often compares the achieved performance and the expected performance among different IC technologies. The innovation proposed in this work suggests the performance degradation as an element of comparison. Thus, such an IC technology will be suitable for a target application only if the performance and its degradation attend the application and reliability requirements.

During this work, we faced these two questions at same time which brings us to a different question similar to both. How reliable could be the circuits in CMOS 65 nm technology? The results achieved in this work indicate variability and ageing compromises in reliable-circuit design. We often observed a small ageing and a big IC process variability. Thus, the design solution was applying full differential and balanced circuits, bringing the circuit variability to a smaller mismatch variation. However, we cannot be sure if some reliable-circuit design is feasible in CMOS 65 nm. Therefore, the answer of this question is the key of the deadlock diagnosis avoiding the unfeasible conditions in a reliable-circuit design in CMOS 65 nm.

In order to answer this question, we decide to characterize the variability and ageing trade-off in CMOS 65 nm. For this purposes, we take a ring oscillator composed by thirteen inverter gates (the smallest inverter gate in CMOS 65 nm) which is a standard characterization circuit in CMOS IC technologies. Understanding the variability and ageing conditions in a ring oscillator, we can analyze the variability and ageing trends for a reliable-circuit design in CMOS 65 nm. What's more, such a characterization generalizes the results obtained in bottom-up and top-down design. Therefore, such a characterization may be employed in determining the variation sharing weight.

The most important characteristics of an IC technology are: die area, power consumption and speed. Using the smallest inverter gate, we assume a die area minimization which imposes the most severe trade-off between power consumption and speed. If the sizing is increased, we can often relax such a trade-off. The power consumption is reduced by a voltage source (V_{DD}) reduction. In this sense, the IC technology evolution has been proposing smaller and smaller V_{DD} as close as possible to the threshold voltage (V_{th}). The technology speed is often measured by the gate delay. The ring oscillation frequency is the inverse of the ring delay. Thus, the mean delay of an inverter gate is calculated dividing the ring delay by the number of the gates in the ring.

The Figure 6.9 exemplifies the power consumption and speed trade-off, using the mean power consumption and mean delay of an inverter gate of the ring oscillator. As expected, we should increase the power consumption to increase the speed. So that, the optimum design should be achieving $T_{delay} \leq 0.2$ ns and $P \leq 1\ \mu$W for each gate. Thus, we present the existent performance trade-off in CMOS 65 nm. This result is satisfactory to conduct to an optimal circuit design, but it is not sure if this circuit will be reliable.

Regarding Figure 6.10, we obtain the required bias characteristics $0.5 \leq V_{DD} \leq 0.7$, aiming $T_{delay} \leq 0.2$ ns and $P \leq 1\ \mu$W. Choosing a small V_{DD} close to V_{th}, the CMOS 65 nm provides the maximum speed with a minimum power consumption. However, this compromise will impose a small signal swing and noise margin which influence the other performance needs in AMS/RF circuits. A classical design will see only this trade-off and make an optimal design in these terms, which still cannot guarantee a reliable design.

The CMOS 65 nm variability will influence these trade-offs spreading the power consumption and the speed around the nominal performance. The worst case is a variation increasing the power consumption and decreasing the speed. This event is driven by lower-level characteristics, which is in this example the V_{DD}. Thus, we simulated the ring oscillator for different V_{DD} from near V_{th} to the nominal voltage source plus 10 %. For each source condition, we simulated a 1000 points Monte Carlo simulation and

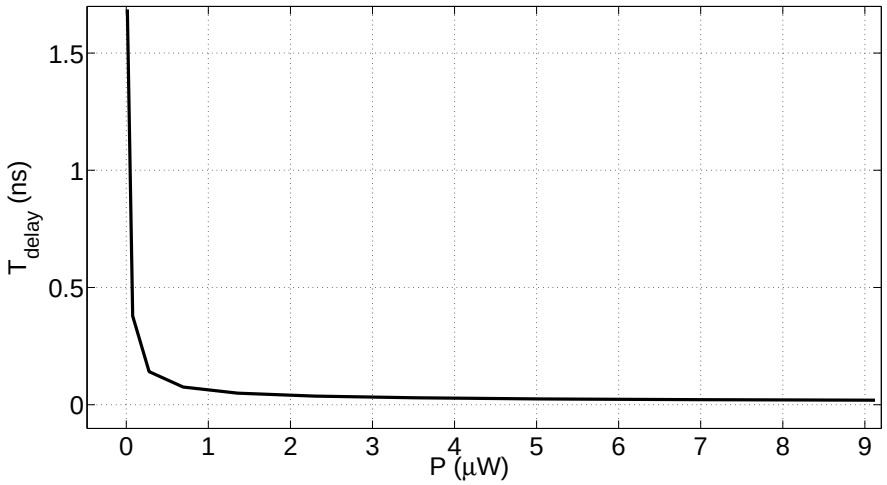


Figure 6.9: Power consumption and speed trade-off illustration using the mean power consumption and mean delay of an inverter gate of the ring oscillator.

extracted the mean and the standard deviation for each characteristic (P and T_{delay}). Through the ratio standard deviation over mean, we can analyze how much representative the variability is in the variation of the performance trade-off. The Figure 6.11 presents the expected variability influencing the power consumption and speed trade-off.

It is evident that choosing a small V_{DD} implies in a huge amount of IC process parameter variability. The expected variability influence for $0.5 \leq V_{DD} \leq 0.7$ shall be near 30 %, and the characteristic variation may represent a circuit failure. This information answers the question why full differential and balanced topologies are required, reducing the variability to values under 10 %. These topologies double the circuit area and increase the power consumption trying a variability cancellation. Actually, only the mismatch variability cannot be ideally cancelled and the result will be a smaller variability. However, this solution is not always available, and the cost in area and power consumption could be bigger than specification. Therefore, the solution is increasing the V_{DD}, increasing the power consumption and keeping the same die area. This solution does not agree with the first optimal design considerations, which were not reliable enough in terms of variability.

In order to evaluate the ring oscillator ageing, we simulated the 30 years aged ring oscillator for all V_{DD} conditions described in the variability characterization for 27^{o}C and 150^{o}C temperature stress. The aged characteristics were represented using the ratio of the absolute ageing degradation and the fresh characteristics. Understanding the ageing phenomena, we can expect that increasing V_{DD} means bigger stress conditions. Such stress leads to an increased ageing degradation as presented in Figure 6.12.

We observe that the ageing is a little more representative when V_{DD} is near V_{th}, because the bias current is small. For the CMOS 65 nm technology, we also observe an ageing influence approximately to 1 % at 27^{o}C and to 3.5 % at 150^{o}C. This result highlights the always bigger variability variation than the ageing variation. What's more, this result explains why the designed amplifiers (BLIXER and PGA amplifiers) have presented a negligible ageing. Actually, the ageing degradation can be neglected facing the more than ten times bigger variability in CMOS 65 nm.

Looking for a reliable design, we have exploited a V_{DD} increasing, reducing the variability. However, this reliable solution at the beginning of the circuit operation time naturally reduces the circuit lifetime. The expected ageing for V_{DD} between 0.9 and 1.1 times the nominal source voltage implies in an approximately 3 % at 27^{o}C and to 5 % at 150^{o}C. Under this condition, the ageing degradation is no longer negligible and it is expected that the combination of HCI and NBTI degradation leads to less reliable circuits.

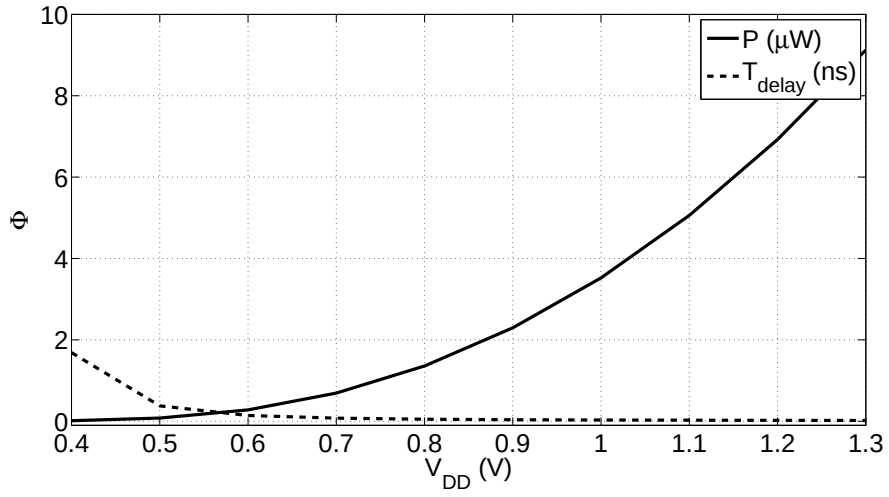


Figure 6.10: Mean power consumption and mean delay of an inverter gate of the ring oscillator as a function of the V_{DD}.

Therefore, the compromise for a reliable design should expect a performance variation at 10 % as the best case and at 50 % as the worst case according the variability trends. Moreover, the performance variation will increase at 1 % as the best case and at 5 % as the worst case according the 30 years lifetime ageing trends. The combination of the variability and the ageing trade-offs will induce an important amount of performance variation turning an optimal design in a design out of the specifications. That is why a reliable-circuit design flow is important to assure an optimal variation margin and the characteristics, according to the specification of performance and reliability.

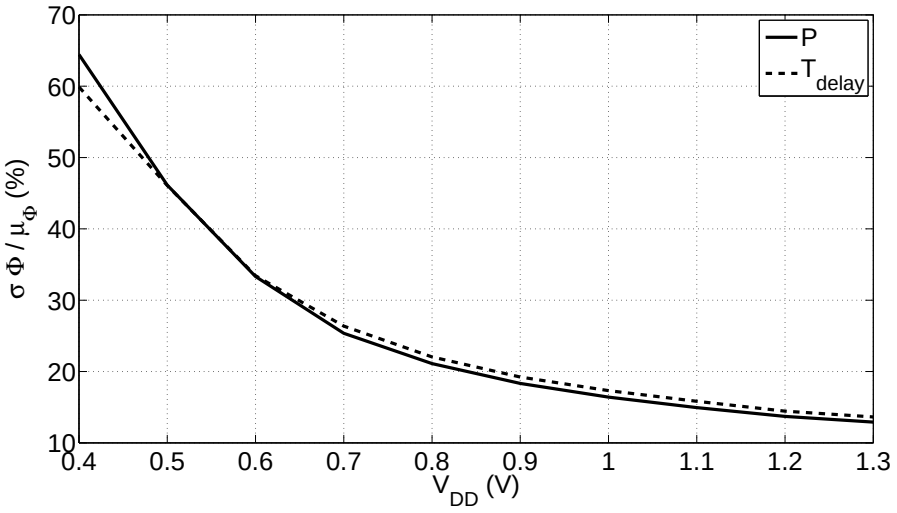
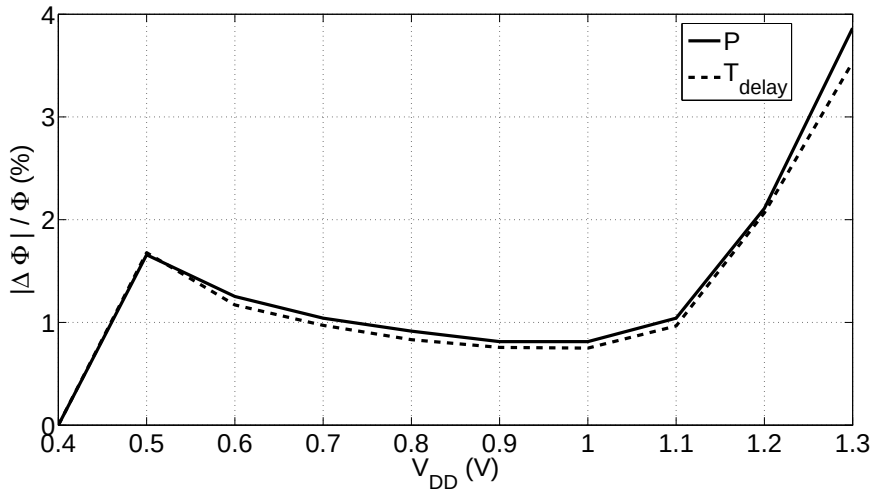
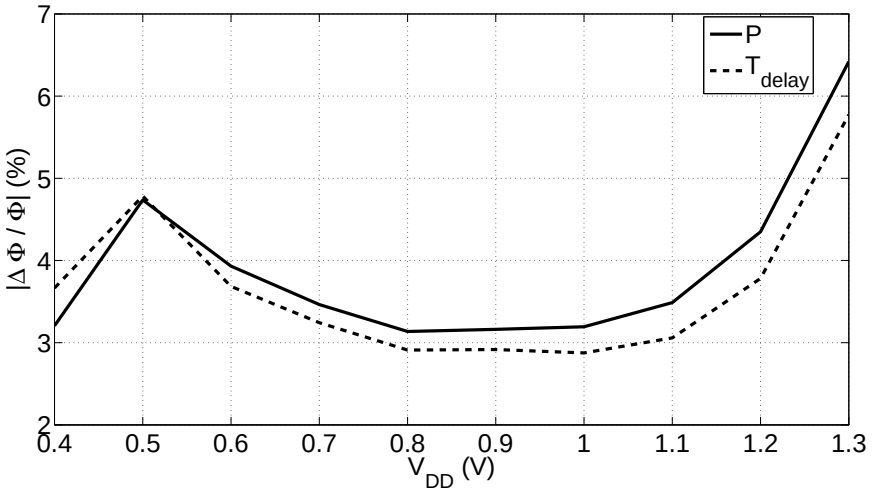


Figure 6.11: CMOS 65 nm variability influencing power consumption and speed trade-off.



(a)



(b)

Figure 6.12: CMOS 65 nm ageing influencing power consumption and speed trade-off for 30 years of stress and temperatures being (a) 27^{o}C and (b) 150^{o}C.

CHAPTER 7

CONCLUSIONS

During this thesis, we proposed a new design flow for AMS/RF circuits with the aim to improve the circuit's reliability. We have exemplified this flow by designing a reliable RF front-end. Our major objective has been successfully achieved; while improving the design of AMS/RF front-end circuits based on the investigation of new trade-offs imposed by transistor variability and ageing degradation. Understanding such degradation, we proposed some sharing strategies, optimizing the design margins and taking into account this trade-off.

This research has described the physic of the ageing phenomena and the required design conditions to avoid the transistor variability and ageing. The sources of degradation (variability and ageing) and their trends in future sub-nanometer technologies have been identified. The classical design methodologies comparing the requirements imposed by the device variability and ageing have been described.

Reliability-aware design methodology is still under research and, as best as we know, there is no most accepted solution. The reliability-awareness has remained in analysis only and mostly simulation-based approach. The reason is that the ageing degradation models are still under research and the state-of-the-art efforts are concentrated in accurate modeling and simulation.

This scenario leaves an opportunity to innovate proposing a reliability-aware design methodology. The research expectations reside in increasing the optimization feedback with design equations to estimate the ageing degradation in early stages. These new equations should be integrated to a multi-objective optimization algorithm. Similar methods have been implemented in the variability-aware design methodology as proposed in [16, 60, 18]. The huge challenge is to propose an accurate design method in order to estimate the ageing degradation in early stages and to use this information for the optimum design space search. The proposed method should have in mind not only the advanced CMOS technologies, but also the future IC technologies, when circuits with even worst reliability are expected.

In this work, we have innovated by proposing AMS/RF circuit reliability improvements during the design of the multi-standard RF front-end using a bottom-up approach. First, the reliable-BLIXER design with a failure evaluation has been proposed. In the bottom-up approach, the circuit's design equations have been used to obtain an early estimation of the ageing of the circuit's characteristics. Using the CMOS 65 nm transistor ageing characterization, the sensible circuit bias condition was identified. The HCI was avoided by reducing the time in which the transistors are in strong inversion, controlling the V_{GD} and the V_{GS}. The simulation results of the typical circuit have satisfied the multi-standard RF front-end specifications.

The validation of a reliable-circuit synthesis method using a DCO design has been conducted. We have considered the reliability degradation, caused by the circuit ageing, a design criterion as important as the classical ones (noise, signal range, power consumption and die area). We have presented the reliable-DCO design and validated its reliability analysis model. We have checked that the reliability analysis gives us information to improve the optimization method, designing a more reliable circuit. Designing a classical and a reliable DCO, a reduction of the frequency degradation, by a value between 15 % and 30 %, has been observed. And also, the reliable DCO has presented a circuit lifetime five times longer than the classical DCO, if we fix the maximum frequency range degradation at 2.0 %. The disadvantages of our method are the phase noise increase and the frequency range reduction.

The PGA and the required elements to design a reliable-PGA have been analyzed. Regarding the design constraints, the reliability analysis and synthesis methodology have been able to point that the PGA is naturally immune to the ageing as the BLIXER. Thus, the circuit variability would be the most important agent of performance variation. Hence, the PGA reliability has been directly controlled by the reliability of the control circuits. Such circuits are digital circuits implemented for the gain programming reconfiguration and the reliability of digital circuits is out of the scope of this work.

Furthermore, we have innovated by proposing architecture reliability improvements during the design of the RF front-end using a top-down approach. In this case, we have discussed the design of a reliable architecture for RF front-end and the variation sharing strategies to avoid an overdesign. Two important strategies have been highlighted using the sensitivity analysis results. Thus, the architecture design should propose a feasible building-block characteristics specification, according to the building blocks design experience and also avoiding the architecture overdesign. Tables 5.4 and 5.5 have presented the building-block performance limits for a reliable architecture. For a reliable RF front-end architecture, the strategy difference between both designs was to give the bigger variation budget to the smaller or to the bigger sensitive characteristic. Both designs were in the valid design space given by the lower-level characterization, but each one has imposed different building-block design constraints. Hence, we have highlighted a new trade-off among architecture gain, noise and reliability.

Therefore, we have innovated linking top-down and bottom-up approaches in a general method which has been the proposition of a new AMS/RF design flow increasing the circuit reliability. The design of reliable circuits has highlighted a new trade-off among typical performance specification, the yield requirements and the circuit lifetime. Such a trade-off, in a reliable design, would be expecting a performance variation at 10 % as the best case and at 50 % as the worst case according the variability trends. Moreover, the performance variation would increase at 1 % as the best case and at 5 % as the worst case according to the 30 years lifetime ageing trends. The combination of the variability and the ageing will induce an important amount of performance variation turning an optimal design in a design out of the specifications. That is why a reliable-circuit design flow is important to assure optimal characteristics and an optimal variation margin, according to the specification of performance and reliability.

By demonstrating the trade-off imposed by transistor variability and ageing in CMOS 65 nm, we have been able to predict such trends in nanometer technologies. We could observe a reducing in the transistor size and the increasing of the gate electric field which have led to increase the stress conditions. Becoming the device characteristic variations more important, such variations would lead to a smaller circuit-lifetime. Hence, such a smaller lifetime has highlighted the importance to design AMS/RF circuits for reliability. Using the design flow proposed in this work, we could analyze and manage the circuit characteristic variations due to variability and ageing. Thus, reliability improvements could be proposed in early stages. Hence, such a design flow for circuit reliability should optimize the circuit performance and also improve the circuit lifetime.

7.1 CONTRIBUTIONS

In this research, we have published:

1. P. M. Ferreira, H. Petit, and J.-F. Naviner, *A New Synthesis Methodology for Reliable RF Front-End Design*, in Proc. of IEEE ISCAS, IEEE, 2011.

2. P. M. Ferreira, H. Petit, and J.-F. Naviner, *A Synthesis Methodology for AMS/RF Circuit Reliability: Application to a DCO Design*, Microelectronics Reliability, Vol. 51, No. 4, pp. 765-772 (2010), 10.1016/j.microrel.2010.11.002.

3. P. M. Ferreira, H. Petit, and J.-F. Naviner *WLAN / WiMAX RF Front-End Reliability Analysis*. in: Proc. IEEE of EAMTA. IEEE; 2010:46-49.

4. P. M. Ferreira, H. Petit, and J.-F. Naviner, *Méthodologie de Conception de Circuits Mixtes et Radiofréquences pour la Fiabilité*, in Proc. JNRDM, 2010.

5. P. M. Ferreira, H. Petit, and J.-F. Naviner, *AMS/RF Reliability Simulation*, in Proc. GDR SoC-SiP, no. 4, 2010.

6. P. M. Ferreira, H. Petit, and J.-F. Naviner, *AMS and RF Design for Reliability Methodology*, in Proc. of IEEE ISCAS, IEEE, 2010.

7. P. M. Ferreira, H. Petit, and J.-F. Naviner, *CMOS 65 nm wideband LNA reliability estimation*, in Proc. of IEEE NEWCAS-TAISA, IEEE, 2009.

8. P. M. Ferreira, H. Petit, and J.-F. Naviner, *Conception de Circuit RF pour Fiabilité*, in Proc. JNRDM, 2009.

9. P. M. Ferreira, H. Petit, and J.-F. Naviner, *AMS/RF Reliability Simulation*, in Proc. GDR SoC-SiP, no. 3, 2009.

7.2 Future Perspectives

In this work, we have found many challenges to propose variability-aware and reliability-aware design methodologies. Thus, we could point some research perspectives in: new analysis tools, new design models, and new synthesis methods; linking variability and ageing. Both are agents of one single consequence which is circuit characteristics variation and so reliability degradation.

This work could be continued with the implementation of an automatic reliability analysis and design tool in agreement with the commercial tools. Such a tool should implement the center of the proposed design flow, which is the iterative failure evaluation and reliability improvement. For this future work, we have contributed by presenting the physical phenomena influence into electrical circuit characteristics. The design experience demonstrated during the RF front-end design and reliability improvement might influence these future works. Furthermore, we have contributed by presenting the variation of the design space under the influence of variability and ageing. The usage of the characterization experience in the RF front-end design has provided the required informations to propose a feasible and more reliable circuit.

By considering the variability and ageing trade-off, we have introduced the discussion of two main questions:

1. How reliable are the known circuit topologies and architectures?

2. How reliable are the sub-nanometer IC technologies?

First, the objective in future reliability works could be quantifying and increasing the reliability of circuit topologies and architectures. A state-of-the-art survey often compares performance trade-offs among the existent topologies positioning some device solution for a target application. The innovation proposed in this work would make possible such a survey comparing also the reliability among the existent topologies. Thus, some device solution would be classified as suitable or not for a target performance and reliability requirements.

Finally, a complete reliability study of different sub-nanometer IC technologies should be conducted comparing the design space variation. Such a research should present the circuit-characteristics variation under a stress environment coherent with the circuit operation conditions. An IC technology survey often compares the achieved performance and the expected performance among different IC technologies. The innovation proposed in this work has suggested the performance degradation as an element of comparison. Thus, such IC technology would be suitable for a target application only if the performance and its degradation have attended the application specification and reliability requirements.

APPENDIX A

VERILOGA BEHAVIORAL MODEL

A.1 BLIXER

```
//Verilog-AMS HDL for "RFreliability", "BLIXER" "veriloga"

`include "constants.vams"
`include "disciplines.vams"

module BLIXER (RF, Ip, In,Qp, Qn, sinA,sinB,cosA,cosB,vdd);
  input sinA;
  output Qn ;
  output In ;
  input cosB ;
  input cosA ;
  inout RF;
  output Qp ;
  output Ip ;
  input sinB ;
  inout vdd;

  electrical RF;
  electrical Ip;
  electrical In;
  electrical Qp;
  electrical Qn;
  electrical sinA;
  electrical sinB;
  electrical cosA;
  electrical cosB;
  electrical vdd;

parameter real power = 1.4; //Power consumption in mW
parameter real gain = 30; // Gain from input to one output, in dB.
parameter real epsilon = 0; // Gain mismatch in percent.
parameter real match = 0; // Balanced gain mismatch in percent.
parameter real I_ip3 = 0; // Input referred 3rd order intercept for I-output.in dBm
parameter real Q_ip3 = 0; // Input referred 3rd order intercept for Q-output.in dBm
parameter real rin = 50 from (0:inf); // Input resistance in Ohms.
parameter real rout = 300 from (0:inf); // Output resistance in Ohms.
parameter real fp =  100e6 from (0:inf); // Pole of the 1st order filter.
```

```
parameter real nf = 4 from [0:inf]; // Noise figure in dB.
parameter real msin = 0.25 from [0:inf]; //Sin wave amplitude

real I_a, Q_a, I_b, Q_b, I_ip, Q_ip, I_vrfoutmax;
real Q_vrfoutmax, I_vrfinmax, Q_vrfinmax, I_vrfout, Q_vrfout; //Non-linearity model variables
real noise_current_squared, rnf; //Noise model variables
real vrf,vsin,vsinB,vcos,vcosB; // They are the respective node voltage
real Wp;

analog begin
// Convert the input parameters from engineering units to implementation units.
@(initial_step) begin
//Compute the I path gain of the first harmonic
I_a = pow(10,gain/20)*(1+epsilon/200);
//Compute the I path interception point (1mW normalization)
I_ip = sqrt(pow(10,I_ip3/10-3)*2*rin);
//Compute the I path gain of the third harmonic
I_b = I_a/(I_ip*I_ip)*4.0/3.0;
// Compute the I path critical point.
I_vrfinmax = sqrt(I_a/(3.0*I_b));
I_vrfoutmax = (2.0*I_a/3.0)*I_vrfinmax;
//Compute the  Q path gain of the first harmonic
Q_a = pow(10,gain/20)*(1-epsilon/200);
//Compute the Q path interception point (1mW normalization)
Q_ip = sqrt(pow(10,Q_ip3/10-3)*2*rin);
//Compute the  Q path gain of the third harmonic
Q_b = Q_a/(Q_ip*Q_ip)*4.0/3.0;
// Compute the  Q path critical point.
Q_vrfinmax = sqrt(Q_a/(3.0*Q_b));
Q_vrfoutmax = (2*Q_a/3)*Q_vrfinmax;
//Compute the noise factor
rnf = pow(10,nf/10.0);
noise_current_squared = 4.0*(rnf-1)*‘P_K*$temperature/rin;
//Compute the pole of the 1st order filter
Wp = 2*‘M_PI*fp;
end
  // Assign the input voltage to the variables: vrf,vsin,vsinB,vcos,vcosB
vrf = V(RF);
vsin= V(sinA);
vsinB= V(sinB);
vcos= V(cosA);
vcosB= V(cosB);
  // Apply the third order nonlinearity. Clamp the output for extreme inputs.
if ( abs(vrf) < I_vrfinmax ) I_vrfout = (I_a - I_b*vrf*vrf)*vrf;
else if (vrf < 0) I_vrfout = -I_vrfoutmax;
else I_vrfout = I_vrfoutmax;
if ( abs(vrf) < Q_vrfinmax ) Q_vrfout = (Q_a - Q_b*vrf*vrf)*vrf;
else if (vrf < 0) Q_vrfout = -Q_vrfoutmax;
else Q_vrfout = Q_vrfoutmax;
```

```
//Input assignment
I(RF) <+ V(RF)/rin;
I(RF) <+ white_noise(noise_current_squared, "BLIXER");
//1st order filter Balanced output assignment
//Non balanced eq
//I(I) <+ ddt(V(I)/(Wp*rout))+(-2*I_vrfout*vcos + V(I))/rout;
//Positive
I(Ip) <+ ddt(V(Ip)/(Wp*rout))+(-I_vrfout*(1+match/200)*vcos/msin + V(Ip))/rout;
//Negative
I(In) <+ ddt(V(In)/(Wp*rout))+(-I_vrfout*(1-match/200)*vcosB/msin + V(In))/rout;
//I(Q)<+ ddt(V(Q)/(Wp*rout))+(2*Q_vrfout*vsin + V(Q))/rout;
//Positive
I(Qp) <+ ddt(V(Qp)/(Wp*rout))+(Q_vrfout*(1+match/200)*vsin/msin + V(Qp))/rout;
//Negative
I(Qn) <+ ddt(V(Qn)/(Wp*rout))+(Q_vrfout*(1-match/200)*vsinB/msin + V(Qn))/rout;
//Power consumption assignment
I(vdd) <+ 0.001*power/V(vdd);

end
endmodule
```

A.2 DCO

```
//Verilog-AMS HDL for "RFreliability", "DCO" "veriloga"

`include "constants.vams"
`include "disciplines.vams"


`define db20_real(x) pow(10, (x)/20)
`define db10_real(x) pow(10, (x)/10)


module DCO(sinA,sinB,cosA,cosB,vdd);
inout vdd;
inout sinA;
inout sinB;
inout cosA;
inout cosB;
electrical vdd;
electrical sinA;
electrical sinB;
electrical cosA;
electrical cosB;

parameter real power = 1.4; //Power consumption in mW
parameter real amp = 1 ;// amp: LO amp when matched(V) default: 1 V
parameter real flo = 1e+09;// flo: LO frequency (Hz) default: 1 GHz
parameter real rout = 50 ;// rout: output impedance (Ohm) default: 50 Ohm
parameter real offset = 0;//offset: DC voltage offset at the output
```

```
parameter real noiseFloor = -180 ;// noiseFloor: the noise floor (dBc/Hz) default:-180 dBc/Hz
parameter real f1 = 1e6 ;// f1: frequency point for Lf1 (Hz) default: 1MHz
parameter real Lf1 = -100 ;// Lf1: phase noise at f1(dBc/Hz) default: -100 dBc/Hz
parameter real fc = 0 ;// if fc is not set to nonzero, then the flicker noise is ignored.
//parameter string freqName="LO" ;
// fundname: the name of the fundamental frequency default: LO

electrical gnd;
electrical int;
//electrical intSin;
ground gnd;

  real vc;
  real kw, kf;   // coefficients of white and flicker noise
  real kb;       // white noise coeff for the noise floor
  real y1;

analog begin

@(initial_step ) begin
kb = `db10_real(noiseFloor);
if(fc == 0) begin // no flicker noise
  y1 = `db10_real(Lf1)-kb;
  kw = f1*sqrt(y1);
  kf = 0;
end else begin
  y1 = `db10_real(Lf1)-kb;
  kw = sqrt(y1/(1/(f1*f1)+fc/(f1*f1*f1)));
  kf = sqrt(fc)*kw;
end
end
//V(intSin)<+I(intSin)*50;
// form the integral of noise
I(int) <+ white_noise(amp*amp*kw*kw, "DCOtermal");
I(int) <+ flicker_noise(amp*amp*kf*kf, 1, "DCOflicker");
I(int) <+ 1/(2*`M_PI)*ddt(V(int));

vc = sqrt(2.0)*V(int)/amp;

I(sinA) <+ (V(sinA)-offset)/rout;
I(sinB) <+ (V(sinB)-offset)/rout;
I(cosA) <+ (V(cosA)-offset)/rout;
I(cosB) <+ (V(cosB)-offset)/rout;
// insert the phase noise into the real signal
I(sinA) <+ -amp*sin(2*`M_PI*flo*$abstime + vc);
I(sinB) <+ -amp*sin(2*`M_PI*flo*($abstime+0.5/flo)+vc);
I(cosA) <+ -amp*cos(2*`M_PI*flo*$abstime+vc);
I(cosB) <+ -amp*cos(2*`M_PI*flo*($abstime+0.5/flo)+vc);
// insert the white noise floor
I(sinA) <+ white_noise(amp*amp*kb/2, "DCO");
```

```
I(sinB) <+ white_noise(amp*amp*kb/2, "DCO");
I(cosA) <+ white_noise(amp*amp*kb/2, "DCO");
I(cosB) <+ white_noise(amp*amp*kb/2, "DCO");
//Power consumption assignment
I(vdd) <+ 0.001*power/V(vdd);
end
endmodule
```

A.3 PGA

```
//Verilog-AMS HDL for "
//reliability", "PGA" "veriloga"

`include "constants.vams"
`include "disciplines.vams"

module PGA (inIp, inIn,inQp, inQn, Ip, In,Qp, Qn, vdd);
  input inIp;
  output Qn ;
  output In ;
  input inIn ;
  input inQp ;
  output Qp ;
  output Ip ;
  input inQn ;
  inout vdd;

  electrical Ip;
  electrical In;
  electrical Qp;
  electrical Qn;
  electrical inIp;
  electrical inIn;
  electrical inQp;
  electrical inQn;
  electrical vdd;

parameter real power = 1.4; //Power consumption in mW
parameter real gain = 30; // Gain from input to one output, in dB.
parameter real epsilon = 0; // Gain mismatch in percent.
parameter real I_ip3 = 0; // Input referred 3rd order intercept for I-output.in dBm
parameter real Q_ip3 = 0; // Input referred 3rd order intercept for Q-output.in dBm
parameter real rin = 50 from (0:inf); // Input resistance in Ohms.
parameter real rout = 50 from (0:inf); // Output resistance in Ohms.
parameter real fp =  25e6 from (0:inf); // Pole of the 1st order filter.
parameter real nf = 4 from [0:inf]; // Noise figure in dB.

real I_a, Q_a, I_b, Q_b, I_ip, Q_ip, I_voutmax;
//Non-linearity model variables
real Q_voutmax, I_vinmax, Q_vinmax, I_vpout,I_vnout, Q_vpout, Q_vnout;
```

```
real noise_current_squared, rnf; //Noise model variables
real vinIp, vinIn, vinQp, vinQn; // They are the respective node voltage
real Wp;

analog begin
// Convert the input parameters from engineering units to implementation units.
@(initial_step) begin
//Compute the I path gain of the first harmonic
I_a = pow(10,gain/20)*(1+epsilon/200);
//Compute the I path interception point (1mW normalization)
I_ip = sqrt(pow(10,I_ip3/10-3)*2*rin);
//Compute the I path gain of the third harmonic
I_b = I_a/(I_ip*I_ip)*4.0/3.0;
// Compute the I path critical point.
I_vinmax = sqrt(I_a/(3.0*I_b));
I_voutmax = (2.0*I_a/3.0)*I_vinmax;
//Compute the  Q path gain of the first harmonic
Q_a = pow(10,gain/20)*(1-epsilon/200);
//Compute the Q path interception point (1mW normalization)
Q_ip = sqrt(pow(10,Q_ip3/10-3)*2*rin);
//Compute the  Q path gain of the third harmonic
Q_b = Q_a/(Q_ip*Q_ip)*4.0/3.0;
// Compute the  Q path critical point.
Q_vinmax = sqrt(Q_a/(3.0*Q_b));
Q_voutmax = (2*Q_a/3)*Q_vinmax;
//Compute the noise factor
rnf = pow(10,nf/10.0);
noise_current_squared = 4.0*(rnf-1)*‘P_K*$temperature/rin;
//Compute the pole of the 1st order filter
Wp = 2*‘M_PI*fp;
end
  // Assign the input voltage to the variables: vrf,vsin,vsinB,vcos,vcosB

vinIp= V(inIp);
vinIn= V(inIn);
vinQp= V(inQp);
vinQn= V(inQn);
  // Apply the third order nonlinearity. Clamp the output for extreme inputs.
if ( abs(vinIp) < I_vinmax ) I_vpout = (I_a - I_b*vinIp*vinIp)*vinIp;
else if (vinIp < 0) I_vpout = -I_voutmax;
else I_vpout = I_voutmax;
if ( abs(vinIn) < I_vinmax ) I_vnout = (I_a - I_b*vinIn*vinIn)*vinIn;
else if (vinIn < 0) I_vnout = -I_voutmax;
else I_vnout = I_voutmax;
if ( abs(vinQp) < Q_vinmax ) Q_vpout = (Q_a - Q_b*vinQp*vinQp)*vinQp;
else if (vinQp < 0) Q_vpout = -Q_voutmax;
else Q_vpout = Q_voutmax;
if ( abs(vinQn) < Q_vinmax ) Q_vnout = (Q_a - Q_b*vinQn*vinQn)*vinQn;
else if (vinQn < 0) Q_vnout = -Q_voutmax;
else Q_vnout = Q_voutmax;
```

```
//Input assignment
I(inIp) <+ V(inIp)/rin;
I(inIn) <+ V(inIn)/rin;
I(inQp) <+ V(inQp)/rin;
I(inQn) <+ V(inQn)/rin;

I(inIp) <+ white_noise(noise_current_squared, "PGA");
I(inIn) <+ white_noise(noise_current_squared, "PGA");
I(inQp) <+ white_noise(noise_current_squared, "PGA");
I(inQn) <+ white_noise(noise_current_squared, "PGA");
//1st order filter Balanced output assignment

I(Ip) <+ ddt(V(Ip)/(Wp*rout))+(-I_vpout + V(Ip))/rout; //Positive
I(In) <+ ddt(V(In)/(Wp*rout))+(-I_vnout + V(In))/rout; //Negative

I(Qp) <+ ddt(V(Qp)/(Wp*rout))+(Q_vpout + V(Qp))/rout; //Positive
I(Qn) <+ ddt(V(Qn)/(Wp*rout))+(Q_vnout + V(Qn))/rout; //Negative

//Power consumption assignment
I(vdd) <+ 0.001*power/V(vdd);

end
endmodule
```

BIBLIOGRAPHY

[1] G. Gielen, P. D. Wit, E. Maricau, J. Loeckx, J. Martín-Martínez, B. Kaczer, G. Groeseneken, R. Rodríguez, and M. Nafría, "Emerging Yield and Reliability Challenges in Nanometer CMOS Technologies," in *Proc. Design, Automation, and Test in Europe*. Munich: ACM, 2008, pp. 1322–1327.

[2] S. Blaakmee, E. Klumperink, D. Leenaerts, and B. Nauta, "A Wideband Balun LNA I/Q-Mixer combination in 65nm CMOS," in *Proc. IEEE Int. Solid-State Circuits*, 2008, pp. 326–328.

[3] W. Altabban, "Conception portable d'une ADPLL pour des applications TV," Ph.D. dissertation, TELECOM ParisTech, 2009.

[4] K.-y. Lee, "Multi-Mode, Multi-Band Active-RC Filter and Tuning Circuits for SDR Applications," in *Advances in Solid State Circuits Technologies*. INTECH, 2010, no. April, ch. 6, pp. 95–106.

[5] J. M. D. L. Rosa, R. Castro-López, A. Morgado, E. C. Becerra-Alvarez, R. D. Río, F. V. Fernández, and B. Pérez-Verdú, "Adaptive CMOS analog circuits for 4G mobile terminals - Review and state-of-the-art survey," *Microelectron. J.*, vol. 40, no. 1, pp. 156–176, 2009.

[6] C. Forzan and D. Pandini, "Statistical static timing analysis: A survey," *Integration, the VLSI Journal*, vol. 42, no. 3, pp. 409–435, Jun. 2009.

[7] K. A. Bowman, A. R. Alameldeen, S. T. Srinivasan, and C. B. Wilkerson, "Impact of die-to-die and within-die parameter variations on the throughput distribution of multi-core processors," in *ACM Proc. of ISLPED*. New York, New York, USA: ACM Press, 2007, pp. 50–55.

[8] S. Sahhaf, R. Degraeve, P. J. Roussel, B. Kaczer, T. Kauerauf, and G. Groeseneken, "A New TDDB Reliability Prediction Methodology Accounting for Multiple SBD and Wear Out," *IEEE Trans. Electron Devices*, vol. 56, no. 7, pp. 1424–1432, 2009.

[9] M. Alam and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectron. Reliab.*, vol. 45, no. 1, pp. 71–81, 2005.

[10] J. H. Stathis, "Reliability limits for the gate insulator in CMOS technology," *IBM J. Res. & Dev.*, vol. 46, no. 2, pp. 265–286, 2002.

[11] B. Kaczer, R. Degraeve, M. Rasras, K. V. D. Mieroop, P. J. Roussel, and G. Groeseneken, "Impact of MOSFET Gate Oxide Breakdown on Digital Circuit Operation and Reliability," *IEEE Trans. Electron Devices*, vol. 49, no. 3, pp. 500–506, 2002.

[12] J. R. Black, "Electromigration-A Brief Survey and Some Recent Results," *IEEE Trans. Electron Devices*, vol. 16, no. 4, pp. 338–347, 1969.

[13] J. B. Bernstein, M. Gurfinke, X. Li, J. Walters, Y. Shapira, and M. Talmor, "Electronic circuit reliability modeling," *Microelectron. Reliab.*, vol. 46, no. 12, pp. 1957–1979, 2006.

[14] G. G. E. Gielen, “Design methodologies and tools for circuit design in CMOS nanometer technologies,” in *IEEE Proc. of ESSDERC*, 2006, pp. 21–32.

[15] S. R. Nassif, “Design for variability in DSM technologies,” in *Proc. of IEEE ISQED*. IEEE Comput. Soc, 2000, pp. 451–454.

[16] G. Yu and P. Li, “Yield-aware analog integrated circuit optimization using geostatistics motivated performance modeling,” in *IEEE/ACM Proc of ICCAD*, Nov. 2007, pp. 464–469.

[17] H. Onodera, “Toward Variability-Aware Design,” in *IEEE Proc. of Symp. on VLSI Tech.*, vol. 4, Jun. 2007, pp. 92–93.

[18] V. Wang, K. Agarwal, S. R. Nassif, K. J. Nowka, and D. Markovic, “A Simplified Design Model for Random Process Variability,” *IEEE Trans. Semiconductor Manufacturing*, vol. 22, no. 1, pp. 12–21, Feb. 2009.

[19] B. Mongellaz, “Contribution à l’intégration de la fiabilité dans le flot de conception des circuits intégrés fondée sur l’utilisation d’un langage de description comportementale VHDL-AMS,” Ph.D., Universite Bordeaux I, 2004.

[20] M. Ruberto, O. Degani, S. Wail, A. Tendler, A. Fridman, and G. Goltman, “A reliability - aware RF power amplifier design for CMOS radio chip integration,” in *Proc. IEEE Int. Reliability Physics Symposium*, Phoenix, 2008, pp. 536–540.

[21] J. Yuan and H. Tang, “CMOS RF Design for Reliability Using Adaptive Gate-Source Biasing,” *IEEE Trans. Electron Devices*, vol. 55, no. 9, pp. 2348–2353, Sep. 2008.

[22] B. Wunderle and B. Michel, “Lifetime modelling for microsystems integration: from nano to systems,” *Microsyst. Technol.*, vol. 15, no. 6, pp. 799–812, 2009.

[23] P. M. Ferreira, H. Petit, and J.-f. Naviner, “A New Synthesis Methodology for Reliable RF front-end Design,” in *Proc. IEEE ISCAS*, 2011, pp. 1–4.

[24] F. H. Gebara, J. D. Schaub, A. J. Drake, K. J. Nowka, and R. B. Brown, “4.0 GHz 0.18 um CMOS PLL Based on an Interpolative Oscillator,” in *Proc. Symposium on VLSI Circuits*, 2005, pp. 100–103.

[25] P. M. Ferreira, H. Petit, and J.-F. Naviner, “A synthesis methodology for AMS/RF circuit reliability: Application to a DCO design,” *Microelectronics Reliability*, vol. 51, no. 4, pp. 765–772, Dec. 2010.

[26] J. E. Chen, “Modeling RF Systems,” pp. 1–41, 2006. [Online]. Available: http://www.designers-guide.org/Modeling/modeling-rf-systems.pdf

[27] P. Maris Ferreira, H. Petit, and J.-F. Naviner, “WLAN / WiMAX RF Front-End Reliability Analysis,” in *Proc. IEEE of EAMTA*. IEEE, 2010, pp. 46–49.

[28] ——, “AMS and RF Design for Reliability Methodology,” in *Proc. IEEE ISCAS*. IEEE, 2010, pp. 3657–3660.

[29] A. Birolini, *Quality and Reliability of Technical Systems, Theory - Practice - Management*, 1st ed. Springer-Verlag, 1994.

[30] A. Papanikolaou, H. Wang, M. Miranda, F. Catthoor, and W. Dehaene, “Reliability Issues in Deep Deep Submicron Technologies: Time-Dependent Variability and its Impact on Embedded System Design,” in *VLSI-SoC: Research Trends in VLSI and Systems on Chip*, G. De Micheli, S. Mir, and R. Reis, Eds. Boston: Springer, 2007, vol. 249, pp. 119–141.

[31] International Roadmap Committee, “International Technology Roadmap of Semiconductors,” 2009. [Online]. Available: http://www.itrs.net

[32] H. Mendez, “Comparing SOI and bulk FinFETs: Performance, manufacturing variability, and cost,” 2010. [Online]. Available: http://www.electroiq.com/articles/sst/print/volume-52/issue-11/features/Cover_Article/Comparing_SOI_and_bulk_F

[33] H. Hara, Y. Okamoto, and H. Ohnuma, “A New Instability in MOS Transistor Caused by Hot Electron and Hole Injection from Drain Avalanche Plasma into Gate Oxide,” *Jpn. J. Appl. Phys.*, vol. 9, no. 9, pp. 1103–1112, 1970.

[34] X. Li, J. Qin, and J. B. Bernstein, “Compact Modeling of MOSFET Wearout Mechanisms for Circuit-Reliability Simulation,” *IEEE Trans. Device Mat. Rel.*, vol. 8, no. 1, pp. 98–121, 2008.

[35] E. Takeda, N. Suzuki, and T. Hagiwara, “Device Performance Degradation Due to Hot-Carrier Injection at Energies Below the Si-SiO_2 Energy Barrier,” in *Proc. IEEE Int. Electron Devices Meeting*, 1983, pp. 396–399.

[36] W. Wang, V. Reddy, A. T. Krishnan, R. Vattikonda, S. Krishnan, and Y. Cao, “Compact Modeling and Simulation of Circuit Reliability for 65-nm CMOS Technology,” *IEEE Trans. Device Mat. Reliab.*, vol. 7, no. 4, pp. 509–517, 2007.

[37] B. C. Paul, K. Kunh, H. Kufluoglu, M. A. Alam, and K. Roy, “Negative Bias Temperature Instability: Estimation and Design for Improved Reliability of Nanoscale Circuits,” *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 4, pp. 743–751, 2007.

[38] Y. Miura and Y. Matukura, “Investigation of Silicon-Silicon Dioxide Interface Using MOS Structure,” *Jpn. J. Appl. Phys*, vol. 5, no. 2, p. 180, 1966.

[39] J. H. Stathis and S. Zafar, “The negative bias temperature instability in MOS devices: A review,” *Microelectron. Reliab.*, vol. 46, no. 4, pp. 270–286, 2006.

[40] S. Mahapatra and M. A. Alam, “Defect Generation in p-MOSFETs Under Negative-Bias Stress: An Experimental Perspective,” *IEEE Trans. Device Mat. Rel.*, vol. 8, no. 1, pp. 35–46, 2008.

[41] M. Denais, A. Bravaix, V. Huard, C. Parthasarathy, G. Ribes, F. Perrie, Y. Rey-Tauriac, and N. Revil, “On-the-fly characterization of NBTI in ultra-thin gate oxide PMOSFET’s,” in *Proc. IEEE Int. Electron Devices Meeting*, 2004, pp. 109– 112.

[42] B. Kaczer, R. Degraeve, P. Roussel, and G. Groeseneken, “Gate oxide breakdown in FET devices and circuits: From nanoscale physics to system-level reliability,” *Microelectron. Reliab.*, vol. 47, no. 4, pp. 559–566, 2007.

[43] B. Weir, P. Silverman, D. Monroe, K. Krisch, M. Alam, G. Alers, T. Sorsch, G. Timp, F. Baumann, C. Liu, Y. Ma, and D. Hwang, “Ultra-Thin Gate Dielectrics: They Break Down, But Do They Fail?” in *Proc. IEEE Int. Electron Devices Meeting*, 1997, pp. 73–76.

[44] E. Miranda, J. Suñé, R. Rodrígues, N. Montserrat, and X. Aymerich, “A Function-Fit Model for the Soft Breakdown Failure Mode,” *IEEE Electron Device Lett.*, vol. 20, no. 6, pp. 265–267, 1999.

[45] J. R. Black, “Physics of Electromigration,” in *Proc. IEEE Reliability Physics Symposium*, 1974, pp. 142–149.

[46] S. Zafar, B. H. Lee, J. Stathis, A. Callegari, and T. Ning, “A Model for Negative Bias Temperature Instability (NBTI) in Oxide and High k pFETs,” in *Proc. IEEE Symposium on VLSI Technology*, 2004, pp. 208–209.

[47] S. Zafar, "A Model for Negative Bias Temperature Instability in Oxide and High-k pFETs," in *Proc. IEEE Integrated Circuit Design and Technology*, 2007, pp. 1–5.

[48] V. Chandra and R. Aitken, "Impact of Voltage Scaling on Nanoscale SRAM Reliability," in *IEEE Proc. of DATE*, 2009, pp. 387 – 392.

[49] E. Ogawa, J. McPherson, J. Rosal, K. Dickerson, T.-C. Chiu, L. Tsung, M. Jain, T. Bonifield, J. Ondrusek, and W. McKee, "Stress-induced voiding under vias connected to wide Cu metal leads," in *Proc. of IEEE International Reliability Physics Symposium*. IEEE, 2002, pp. 312–321.

[50] A. Heryanto, K. Pey, Y. Lim, W. Liu, J. Wei, N. Raghavan, J. Tan, and D. Sohn, "Study of stress migration and electromigration interaction in copper/low-κ interconnects," in *Proc. of IEEE IRPS*. IEEE, May 2010, pp. 586–590.

[51] International Roadmap Committee, "PROCESS INTEGRATION, DEVICES, AND STRUCTURES," International Technology Roadmap for Semiconductor, Tech. Rep., 2001.

[52] C. G. V. D. Walle and B. R. Tuttle, "Microscopic Theory of Hydrogen in Silicon Devices," *IEEE Trans. Electron Devices*, vol. 47, no. 10, pp. 1779–1786, 2000.

[53] J. Ko and S. Cm, "Negative bias stress of MOS devices at high electric elds and degradation of MNOS devices," *J. Appl. Phys.*, vol. 48, no. 5, pp. 2004–14, 1977.

[54] B. Liu, Y. Wang, Z. Yu, L. Liu, M. Li, Z. Wang, J. Lu, and F. Fernandez, "Analog circuit optimization system based on hybrid evolutionary algorithms," *Integration, the VLSI Journal*, vol. 42, no. 2, pp. 137–148, Feb. 2009.

[55] D. J. Allstot, K. Choi, and J. Park, *Parasitic-Aware Optimization of CMOS RF Circuits*. Kluwer Academic Publishers, 2003.

[56] J. Cong, G. Luo, and E. Radke, "Highly Efficient Gradient Computation for Density-Constrained Analytical Placement," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 12, pp. 2133–2144, Dec. 2008.

[57] J. Cong and G. Luo, "An analytical placer for mixed-size 3D placement," in *Proc. of ACM International Symposium on Physical Design*. New York, New York, USA: ACM Press, 2010, p. 61.

[58] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi, "A tutorial on geometric programming," *Optimization and Engineering*, vol. 8, no. 1, pp. 67–127, Apr. 2007.

[59] L. Vandenberghe, "Techniques for improving the accuracy of geometric-programming based analog circuit design optimization," in *Proc. of IEEE/ACM International Conference on Computer Aided Design*. Ieee, 2004, pp. 863–870.

[60] Y. Xu, K.-L. Hsiung, X. Li, L. T. Pileggi, and S. P. Boyd, "Regular Analog/RF Integrated Circuits Design Using Optimization With Recourse Including Ellipsoidal Uncertainty," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst*, vol. 28, no. 5, pp. 623–637, May 2009.

[61] G. Gielen, H. Walscharts, and W. Sansen, "Analog circuit design optimization based on symbolic simulation and simulated annealing," *IEEE Journal of Solid-State Circuits*, vol. 25, no. 3, pp. 707–713, Jun. 1990.

[62] R. Gupta, B. Ballweber, and D. Allstot, "Design and optimization of CMOS RF power amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 2, pp. 166–175, 2001.

[63] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., 1989.

[64] A. P. Vaze, "Analog Circuit Design using Genetic Algorithm : Modified," *World Academy of Science, Engineering and Technology*, vol. 14, pp. 62–64, 2006.

[65] A. Jafari, M. Zekri, S. Sadri, and A. Mallahzade, "Design of Analog Integrated Circuits by Using Genetic Algorithm," in *Proc. of International Conference on Computer Engineering and Applications*. IEEE, Mar. 2010, pp. 578–581.

[66] A. Mitev, M. Marefat, D. Ma, and J. Wang, "Parameter reduction for variability analysis by slice inverse regression method," *IET Circuits, Devices & Systems*, vol. 2, no. 1, p. 16, 2008.

[67] T. McConaghy, T. Eeckelaert, and G. Gielen, "CAFFEINE: Template-Free Symbolic Model Generation of Analog Circuits via Canonical Form Functions and Genetic Programming," in *IEEE Proc. of DATE*. IEEE, 2005, pp. 1082–1087.

[68] CADENCE, "Accelerating Analog Simulation with Full SPICE Accuracy," Cadence, Tech. Rep., 2008.

[69] ——, "Virtuoso Multi-Mode Simulation," Cadence, Tech. Rep., 2007.

[70] ——, "RELIABILITY SIMULATION IN INTEGRATED CIRCUIT DESIGN," Cadence, Tech. Rep., 2003.

[71] MentorGraphics, "ELDO Integrated Circuit Simulation," Mentor Graphics, Tech. Rep., 2009. [Online]. Available: www.mentor.com

[72] ——, "ELDO RF Integrated Circuit Simulation," Mentor Graphics, Tech. Rep., 2010.

[73] R. Daniels, H. V. Sosen, and H. Elhak, "Accelerating Analog Simulation with HSPICE Precision Parallel Technology," Synopsys, Tech. Rep. September, 2010.

[74] Magma, "The Titan Unified , Automated , Full-Chip Mixed-Signal Design Solution," Magma Design Automation, Tech. Rep., 2008.

[75] L. G. e Silva, J. Phillips, and L. M. Silveira, "Effective Corner-Based Techniques for Variation-Aware IC Timing Verification," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst*, vol. 29, no. 1, pp. 157–162, Jan. 2010.

[76] J. Jaffari and M. Anis, "Correlation Controlled Sampling for Efficient Variability Analysis of Analog Circuits," in *Proc. of IEEE Design, Automation and Test in Europe*, no. 1, 2010, pp. 1305 – 1308.

[77] C. M. A.-C. Raymond H. Myers, Douglas C. Montgomery, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, 3rd ed. Wiley, 2009.

[78] R. H. Tu, E. Rosenbaum, W. Y. Chan, C. C. Li, E. Minami, K. Quader, P. K. Ko, and C. Hu, "Berkeley Reliability Tools-BERT," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, no. 10, pp. 1524–1534, 1993.

[79] E. Xiao and J. S. Yuan, "RF Circuit Design in Reliability," in *Proc. IEEE Int. Microwave Symposium Digest*, 2003, pp. A61 – A64.

[80] E. Maricau and G. Gielen, "Variability-Aware Reliability Simulation of Mixed-Signal ICs with Quasi-Linear Complexity," in *IEEE Proc. of DATE*, 2010, pp. 1094 – 1099.

[81] C. R. Parthasarathy, M. Denais, V. Huard, G. Ribes, D. Roy, C. Guerin, F. Perrier, E. Vincent, and A. Bravaix, "Designing in reliability in advanced CMOS technologies," *Microelectron. Reliab.*, vol. 46, no. 10, pp. 1464–1471, 2006.

[82] G. T. Sasse, M. Acar, F. G. Kuper, and J. Schmitz, “RF CMOS reliability simulations,” *Microelectron. Reliab.*, vol. 48, pp. 1581–1585, 2008.

[83] Y.-l. R. Lu, Y.-h. Lee, W. J. Mcmahon, and T.-c. Fung, “Robust Inductor Design for RF Circuits,” in *Proc. IEEE Custom Intergrated Circuits Conference*, no. Cicc, 2006, pp. 571–574.

[84] X. Pan and H. Graeb, “Reliability analysis of analog circuits using quadratic lifetime worst-case distance prediction,” in *IEEE Proc of Custom Integrated Circuits*. IEEE, Sep. 2010, pp. 1–4.

[85] P. Maris Ferreira, H. Petit, and J.-F. Naviner, “CMOS 65 nm wideband LNA reliability estimation,” in *Proc. IEEE NEWCAS-TAISA*. IEEE, 2009, pp. 1–4.

[86] A. A. Abidi, “Phase Noise and Jitter in CMOS Ring Oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1803–1816, Aug. 2006.

[87] ITRS, “SYSTEM DRIVERS,” International Technology Roadmap for Semiconductor, Tech. Rep., 2009.

[88] B. Razavi, *Design of Analog CMOS Integrated Circuits*. McGraw Hill Higher Education, 2003.

[89] Q. Gu, *RF System Design of Transceivers for Wireless Communications*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[90] P. Jespers, *The gm/ID Methodology, a sizing tool for low-voltage analog CMOS Circuits*. Boston, MA: Springer US, 2010. [Online]. Available: http://www.springerlink.com/index/10.1007/978-0-387-47101-3

[91] C.-H. Ko and H.-Y. Wei, “Game Theoretical Resource Allocation for Inter-BS Coexistence in IEEE 802.22,” *IEEE Transactions on Vehicular Technology*, vol. 59, no. 4, pp. 1729–1744, May 2010.

[92] J. F. Nash, “EQUILIBRIUM POINTS IN N-PERSON GAMES,” *Proc. Nat. Acad. Sci.*, vol. 36, no. 1, pp. 48–49, 1950.

[93] Y. Cheng and R. Fujii, “SAUCES: A SENSITIVITY ANALYSIS PROGRAM FOR ANALOG CIRCUIT DESIGN,” in *Proc. IEEE ISCAS*, 1992, pp. 1175–1178.