Amélioration des résolutions spatiale et temporelle des plateformes d'analyse et d'injection électromagnétiques

Julien Toulemont

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En microélectronique

École doctorale I2S

Unité de recherche LIRMM

Amélioration des résolutions spatiale et temporelle des plateformes d'analyse et d'injection électromagnétiques

Présentée par Julien TOULEMONT

Le 13 décembre 2021

Sous la direction de Philippe MAURINE et Pascal NOUET

Devant le jury composé de

Pascal NOUET	Professeur	LIRMM – Université de Montpellier	Directeur de Thèse
Philippe MAURINE	MCF	LIRMM – Université de Montpellier	Co-directeur de Thèse
Frederick MAILLY	MCF	LIRMM – Université de Montpellier	Co-encadrant de Thèse
Laurent LATORRE	Professeur	LIRMM – Université de Montpellier	Président du Jury
Jean Max DUTERTRE	MCF	ENSMSE	Rapporteur
Hélène TAP	Professeur	INP – ENSEEIHT	Rapporteur
Julien MICOLOD	Ingénieur	DGA – MI	Examinateur
Laurent SAUVAGE	MCF	Télécom ParisTech	Examinateur
Benoît GERARD	Ingénieur	DGA – MI	Invité

UNIVERSITÉ DE MONTPELLIER

Remerciements

Ce manuscrit résume le travail que j'ai effectué durant quatre années au sein du département microélectronique du LIRMM, sur les sondes et plateformes d'analyse et d'injection électromagnétiques. Ce travail de longue haleine n'aurait pas été possible sans le soutien de nombreuses personnes, que je souhaite donc remercier chaleureusement ici.

Mes premiers remerciements s'adressent à mes encadrants de thèse, Philippe MAURINE, Pascal NOUET et Frederick MAILLY pour leur soutien tout au long de ma thèse. La qualité de leur encadrement, leur bonne humeur et leurs précieux conseils m'ont permis de me sentir bien et d'avancer sereinement tout au long de ces années.

Je remercie très sincèrement Caroline LEBRUN et Faiza LAACHIR pour avoir simplifié au maximum les démarches administratives. Vous m'avez fait gagner un temps incroyable, que j'ai pu consacrer pleinement à l'avancement de mes travaux. Je suis également particulièrement reconnaissant envers Laurent DE KNYFF pour toutes les commandes de composants, la gravure des PCB et ses précieux conseils concernant l'électronique en général. J'adresse aussi toute ma gratitude à Jérémie SALLES qui a fait preuve d'une patience hors norme, en m'aidant pendant des heures sur le logiciel de conception de l'amplificateur.

Durant ces quatre années passées au LIRMM, j'ai eu l'occasion de côtoyer de nombreuses personnes qui m'ont apporté leur aide et qui ont partagé leur expérience avec moi. J'aimerais ainsi remercier Victor LOMNE, ces quatre années n'auraient pas été aussi fructueuses sans tes précieux conseils et remarques. J'ai beaucoup apprécié nos conversations enrichissantes lors des pauses repas et café.

Je me dois aussi de remercier l'ensemble des doctorants du département microélectronique que j'ai fréquenté au fur et à mesure de ma thèse, parmi eux Gwenaël CHAILLOU (mon camarade de bureau actuel), Théo SORIANO

et Paul DELESTRAC (mes camarades de salle de sport), Loïc FRANCE et Jonathan MIQUEL. Je remercie également Jr REALPE, ancien doctorant de robotique, pour son assistance à l'impression 3D.

Un petit clin d'œil à un ancien doctorant du LIRMM, Maxime COZZI, nos séances de râlerie m'ont manqué lorsque tu es parti, mais j'ai trouvé quelqu'un pour prendre la relève. En parlant de lui, un grand merci à Geoffrey CHANCEL pour son aide précieuse sur les scripts Python, mais surtout pour les nombreuses pauses café passées à s'apitoyer sur notre sort.

Mes plus profonds remerciements vont évidemment à ma mère, mon père et ma sœur qui m'ont toujours encouragé tout au long de mes études, et plus particulièrement ces dernières années. On se rappellera d'une certaine prof de Français de collège qui vous a soutenu que je n'avais pas les capacités intellectuelles pour faire de longues études.

Avant de terminer, je souhaite remercier chaleureusement trois amis Bayonnais de longue date. Nicolas MEDDOUR, preuve vivante qu'il ne sert à rien d'avoir beaucoup d'amis, mais qu'il suffit juste de trouver le bon. Nos séances de sport en visio pendant les mois du couvre-feu m'ont permis de garder le moral en cette période difficile, tout comme nos nombreuses parties d'échecs. Laurent VIGNES, ta bonne humeur constante, notamment lors de notre voyage au Portugal, m'a permis de l'apprécier encore plus. J'ai ainsi pu reprendre et terminer plus sereinement la rédaction de ce manuscrit. Lionel LARTIGUE, les quelques jours passés chez toi à la fin de chaque été m'ont permis de souffler, avant de repartir plus motivé que jamais. Les discussions sur nos thèses respectives ont également été très enrichissantes.

Enfin, je remercie le jury et les rapporteurs de me faire l'honneur d'évaluer mes quatre années de thèse. Et plus généralement, à toutes les personnes qui liront ce manuscrit, je vous souhaite une bonne lecture.

Table des matières

Table des figures

Liste des tableaux

Abréviations

AES	**A**dvanced **E**ncryption **S**tandard
CLB	**C**onfigurable **L**ogic **B**loc
CMOS	**C**omplementary **M**etal **O**xide **S**emiconductor
COTS	**C**ommercial **O**ff **T**he **S**helf
CPA	**C**orrelation **P**ower **A**nalysis
CRT	**C**hinese **R**emainder **T**heorem
CW	**C**ontinuous **W**aves
DC	**D**irect **C**urrent
DES	**D**ata **E**ncryption **S**tandard
DFA	**D**ifferential **F**ault **A**nalysis
DFF	**D** **F**lip **F**lop
DIL	**D**ual **I**n **L**ine
DoM	**D**ifference **o**f **M**eans
DPA	**D**ifferential **P**ower **A**nalysis
EM	**E**lectro**M**agnetique
EMFI	**E**lectro**M**agnetic **F**ault **I**njection

FIB	**F**ocused **I**on **B**eam
FIFO	**F**irst **I**n **F**irst **O**ut
FPGA	**F**ield **P**rogrammable **G**ate **A**rray
FWHM	**F**ull **W**idth at **H**alf **M**aximum
HIFI	**HI**gh **FI**delity
IOT	**I**nternet **O**f **T**hings
LNA	**L**ow **N**oise **A**mplifier
MIA	**M**utual **I**nformation **A**nalysis
PCB	**P**rinted **C**ircuit **B**oard
PEA	**P**rogramme d'**É**tude **A**mont
PW	**P**ulse **W**idth
RAM	**R**andom **A**ccess **M**emory
RF	**R**adio **F**réquence
RSA	**R**ivest **S**hamir **A**dleman algorithm
SNR	**S**ignal to **N**oise **R**atio
SoC	**S**ystem **o**n **C**hip
SPA	**S**imple **P**ower **A**nalysis
TA	**T**iming **A**ttack
ZIF	**Z**ero **I**nsertion **F**orce

Publications

Conférences et Workshop

— J. Toulemont, N. Ouldei-Tebina, J. M. Galliere, P. Nouet, E. Bourbao and P. Maurine. "A Simple Protocol to Compare EMFI Platforms". Cryptology ePrint Archive, Report 2020/1277, Oct 2020, https://ia.cr/2020/1277

— J. Toulemont, F. Mailly, P. Maurine, P. Nouet. "Exploring flexible and 3D printing technologies for the design of high spatial resolution EM probes". 19th IEEE International New Circuits and Systems Conference (NEWCAS 2021), Jun 2021, Toulon, France. pp.1-4

— J. Toulemont, G. Chancel, J. M. Galliere, F. Mailly, P. Maurine, P. Nouet. "On the scaling of EMFI probes". 2021 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC 2021), Sept 2021, Milan, Italy - pp. 67-73

Introduction générale

Notre communauté scientifique a été actrice et témoin, lors de la dernière décennie, de l'essor des technologies numériques et plus particulièrement de l'internet des objets (IOT). On les retrouve désormais dans la plupart des secteurs d'activité, qu'ils soient industriels ou économiques, mais également dans la vie quotidienne avec par exemple la domotique. Ces objets et systèmes, qui se comptent en milliards, manipulent des données privées et sensibles, parfois à hautes valeurs ajoutées, ce qui en fait une source de convoitise. Puisque l'ensemble de ses objets est interconnecté, il convient alors de protéger même le plus insignifiant d'entre eux ; la moindre faille de sécurité pouvant compromettre l'entièreté du système.

Les fabricants ont alors recours à la cryptographie, et plus particulièrement à des protocoles et algorithmes de chiffrement pour assurer la sécurité et la confidentialité des données de l'ensemble des systèmes électroniques. Bien que ces algorithmes (comme l'AES par exemple) soient considérés comme mathématiquement robustes, leurs implémentations sur des circuits intégrés amènent inévitablement à une fuite d'information physique pouvant prendre plusieurs formes : consommation de courant, émissions électromagnétiques, temps de calculs, etc. Cette fuite d'information a mené à l'élaboration de ce que l'on appelle aujourd'hui plus communément "attaques par canaux auxiliaires". Ce type d'attaque permet à une personne malintentionnée d'extraire la clé de chiffrement secrète d'un circuit en l'espionnant lors de son fonctionnement.

Le LIRMM est considéré depuis des années comme le laboratoire phare en ce qui concerne les attaques par médium électromagnétique. Celle par injection de fautes par médium électromagnétique est une attaque de choix car elle ne nécessite que très peu de préparation du circuit ciblé, et possède de nombreux avantages, dont le fait que le circuit n'est pas altéré ou détruit à la suite de celle-ci. C'est dans ce contexte que ma thèse a été effectuée.

Le chapitre 1 propose un état de l'Art sur les attaques par canaux auxiliaires ainsi que sur les techniques d'injection de fautes. Il est suivi par une présentation des plateformes dédiées aux attaques exploitant le médium électromagnétique, tant au niveau analyse qu'injection. Ces choix initiaux qui peuvent paraître surprenants au premier abord, sont pleinement justifiés dans le contexte du FUI CSAFE+ (projet avec lequel ma thèse a débuté) et du PEA SERGE (projet sur lequel j'ai travaillé ces 2 dernières années) finançant mes travaux. Après avoir précisé les objectifs concrets de la thèse, définis à partir des lacunes et limitations identifiées dans l'état de l'Art, mais également en prenant en compte les objectifs des projets sous-jacents, l'ensemble de mes travaux est introduit.

Le chapitre 2 présente les contributions relatives à l'amélioration de la résolution spatiale des sondes d'analyse électromagnétique. Ces contributions sont basées sur une technologie d'électronique flexible permettant la réduction de la taille des sondes. Il présente également la conception et la caractérisation d'un amplificateur faible bruit en technologie AMS $0.35\mu m$ destiné à fabriquer des sondes actives, afin de maintenir des valeurs de SNR élevées. Un prototype fonctionnel de sonde active y est également présenté. Ces travaux ont été conduits pendant les deux premières années de ma thèse.

Le chapitre 3 porte sur l'amélioration des plateformes d'injection de fautes par médium électromagnétique. Ces travaux sont basés en premier lieu sur l'amélioration de la résolution spatiale des sondes, notamment grâce à l'utilisation de la technologie d'électronique flexible précédente. Au cours de ces travaux, il est apparu que des sondes plus petites pouvaient être très intéressantes pour effectuer des attaques plus localisées et plus sélectives. Une amélioration de la résolution temporelle grâce à l'élaboration d'un système anti-rebonds simple est également proposée. L'ensemble de ces travaux a été mené de manière à conserver le caractère pratique et modulaire de ce type de plateforme. Un protocole simple de caractérisation des plateformes et des sondes y est également décrit. Ces travaux ont occupé les deux dernières années de ma thèse.

Enfin, une dernière partie conclut le travail effectué durant ces quatre années, en rappelant les principales contributions et les principaux résultats obtenus. Les perspectives y sont également développées.

Chapitre 1

État de l'Art

1.1 Résumé

Ce chapitre dresse un état de l'Art sur les attaques par observation d'un canal caché et les attaques par injection de fautes. Une attention particulière sera portée aux plateformes et moyens techniques nécessaires à la mise en œuvre d'attaques par observation du champ magnétique et par injection / induction électromagnétique. On décrira également un modèle de fautes spécifique aux injections de fautes par médium électromagnétique qui sera utile dans les chapitres ultérieurs.

1.2 Introduction

De nos jours, la plupart des secteurs d'activité reposent sur l'utilisation de circuits et systèmes électroniques. Parmi ces systèmes et circuits intégrés, on trouve des circuits dédiés à la sécurité comme les cartes à puce dont l'usage s'est démocratisé. Elles sont en effet maintenant utilisées dans de nombreux domaines comme la finance, la santé publique ou les transports pour n'en nommer que quelques-uns. Selon les secteurs, les données qu'elles traitent et stockent ont une valeur plus ou moins importante et peuvent faire l'objet de convoitise. Par conséquent, il est important d'assurer la fiabilité et la sécurité de ces systèmes, ainsi que celles des données qu'ils manipulent, contre des personnes malintentionnées. Ainsi, comme tous les systèmes sécurisés, les cartes à puce doivent garantir :

— l'authenticité des données : Les données reçues par le destinataire sont bien celles envoyées par l'émetteur,

— l'intégrité des données : Les données ne doivent pas pouvoir être remplacées ou altérées par un tiers extérieur,

— la confidentialité des données : Les données ne sont ni lisibles, ni exploitables par un tiers extérieur.

Pour garantir ces propriétés, les systèmes sécurisés intègrent de manière matérielle ou bien logicielle des protocoles et algorithmes cryptographiques. Il en existe de nombreux. Parmi ceux-ci on distingue les algorithmes de chiffrement asymétriques et les algorithmes de chiffrement symétriques. Les premiers reposent sur l'utilisation de deux clés distinctes (une clé publique et une clé privée) pour le chiffrement et le déchiffrement, alors que les derniers n'utilisent qu'une clé pour chiffrer et déchiffrer les messages.

Si ces algorithmes de chiffrement sont considérés comme mathématiquement robustes, de nombreuses failles apparaissent lors de leurs implémentations au sein de circuits intégrés. La première des ces failles est liée au fonctionnement même des circuits intégrés. En effet, ceux-ci provoquent, en fonction des données traitées, une modification de grandeurs physiques mesurables, telles que la tension d'alimentation, la température ou le champ magnétique. Si aucune précaution n'est prise, ces grandeurs peuvent alors être mesurées et exploitées par un tiers, pour extraire des données sensibles. On parle alors d'attaques par observation d'un canal caché.

La seconde faille est liée au fait que les circuits intégrés sont conçus pour fonctionner dans un domaine borné de tension d'alimentation, de température et fréquence, etc. Un tiers malveillant peut donc forcer un circuit à fonctionner, pendant un bref laps de temps, en dehors de ce domaine afin d'induire des comportements erronés qu'il exploite alors pour extraire des données secrètes ou bien leurrer un système de sécurité. On parle dans ce cas d'attaques par injection de fautes.

L'ensemble des attaques, qu'elles soient par observation ou par injection de fautes, est généralement scindé en trois grandes catégories selon le degré de pénétration physique du système nécessaire à leur application. On parle selon les cas d'attaques invasives, semi-invasives ou non invasives.

— L'application des attaques dites invasives nécessite un contact direct

avec des éléments internes constitutifs des circuits afin de les modifier, de modifier leur comportement, de les observer ou bien les neutraliser. Pour ce faire des méthodes chimiques ou mécaniques sont employées pour retirer le boîtier plastique, initialement présent pour protéger le silicium et, retirer, modifier ou altérer l'intérieur du circuit. Ces attaques sont très coûteuses et s'accompagnent souvent de la destruction des circuits cibles. Parmi ces attaques, les plus connues sont la rétro-conception (reverse engineering) [Bly+93] et le micro-sondage [Sam+02].

— Les attaques semi-invasives parmi lesquelles on trouve la plupart des attaques par injection de fautes nécessitent un minimum de préparation du circuit cible, comme par exemple une décapsulation du circuit pour avoir un accès direct à la puce, ou un amincissement de son substrat, mais ne requièrent aucune modification de la puce en elle-même. Elles utilisent le plus souvent des impulsions laser [WWM11], [Vas+17] ou électromagnétiques [Deh+12b] de manière à venir modifier temporairement la valeur d'un ou plusieurs bits. Contrairement aux attaques précédentes, le circuit n'est pas altéré ou détruit à la suite de ces attaques.

— Les attaques non invasives, parmi lesquelles on trouve les attaques par canaux cachés, peuvent être pratiquées sans aucune modification préalable des puces. Elles nécessitent donc très peu de moyens. De ce fait elles sont considérées comme les plus dangereuses.

Dans ce contexte, et compte tenu des objectifs de cette thèse relatifs au canal auxiliaire électromagnétique comme moyen d'attaque par observation ou comme médium d'injection de fautes, ce chapitre a pour objectif de dresser un état de l'Art sur les attaques tant en observation que par injection de fautes en portant une attention particulière aux plateformes matérielles permettant de réaliser ces attaques. Il est organisé comme suit. En section 1.3, un bref état de l'Art sur les attaques par canaux cachés est proposé. S'en suit section 1.4 une présentation des plateformes dédiées aux attaques exploitant le canal caché électromagnétique. La section 1.5 dresse ensuite un rapide état de l'Art sur les techniques d'injection de fautes avant de dresser section 1.6 un historique relatif aux injections de fautes par médium EM. Enfin, la section 1.7 s'intéresse aux plateformes nécessaires pour mener ce type d'injection.

Ce chapitre se termine en précisant les objectifs concrets de cette thèse, ces derniers ayant été définis à partir des lacunes et limitations identifiées dans l'état de l'Art mais également en tenant compte des objectifs du projet FUI CSAFE+ et du PEA SERGE.

1.3 Les attaques par canaux cachés

Ce premier paragraphe présente brièvement les différentes attaques par canaux auxiliaires sachant que la littérature est extrêmement riche sur ce sujet et ce notamment par rapport aux aspects "marqueurs statistiques" utilisés pour mener ce type d'attaque. L'intérêt porté lors de cet état de l'Art sur les attaques exploitant le canal électromagnétique peut paraître surprenant mais est pleinement justifié dans le contexte du PEA Serge.

1.3.1 Attaques temporelles

Les premières attaques par canaux auxiliaires ont été identifiées comme menace sérieuse depuis bien longtemps et ont fait leur apparition à la fin des années 1990, avec notamment les attaques temporelles (TA) [Koc96]. Celles-ci consistent à mesurer les temps mis par un circuit intégré pour effectuer certains calculs cryptographiques puis à analyser ces derniers pour en déduire des clés de chiffrement. Par exemple, dans [Koc96] la cible de l'attaque est un RSA s'appuyant sur une méthode d'exponentiation naïve dont le temps de calcul dépend de la valeur du bit d'exposant en cours de traitement.

1.3.2 Attaques par analyse de la consommation

Il existe deux catégories d'attaque par analyse de la consommation. Les attaques par analyse simple (SPA) de celle-ci et les attaques par analyse statistique comme l'analyse différentielle de la consommation (DPA), l'analyse par corrélation (CPA) et l'analyse par information mutuelle (MIA).

L'analyse simple de la consommation est très similaire à l'analyse des timings. En effet, tout comme les TA, elle exploite des différences entre deux macro-opérations (par exemple une mise au carré d'un mot binaire ou la multiplication modulaire de deux mots binaires). La principale différence réside dans la manière de mesurer et quantifier cette différence. Dans le cas des TA, il s'agit de la durée de l'opération alors que dans le cas de la SPA il s'agit du

pattern (forme d'onde) de la consommation de cette opération. En général, cette attaque peut être effectuée à l'œil nu (avec un oscilloscope) sur des implémentations matérielles non protégées du RSA.

Les attaques par analyse statistique de la consommation exploitent des caractéristiques beaucoup plus fines (tant en temps qu'en durée) dans les traces de consommation. Bien que très efficaces leur principe reste simple. Il consiste, en s'appuyant sur un modèle général de la consommation des circuits (en général le poids de Hamming, Hamming Weight), à construire des modèles spécifiques de la consommation d'une opération cible mélangeant un morceau de la clé cryptographique et un morceau du texte chiffré connu. Autant de modèles de consommation qu'il y a d'hypothèses sur la valeur du morceau de clé ciblée sont établis. La bonne hypothèse est finalement identifiée comme étant celle permettant d'avoir la meilleure correspondance entre le modèle et la réalité, à savoir des mesures de la consommation du circuit cible lors du traitement de l'opération cible.

Ce principe général établi dans [KJJ99] a fait l'objet de nombreuses variantes associées à différents marqueurs statistiques appelés distingueurs dans la communauté. Ainsi, dans [KJJ99], Paul Kocher utilise comme distingueur la Différence des Moyennes (DoM) dans sa fameuse Differential Power Analysis. Dans [BCO04], qui introduit la Correlation Power Analysis, le distingueur utilisé est la corrélation de Bravais-Pearson. Enfin, sans être exhaustif, la notion d'information mutuelle est utilisée dans [Gie+08].

1.3.3 Attaques par analyse thermique ou émission de photons

Le rayonnement thermique d'un circuit intégré assimilé à un corps noir peut également être exploité pour mener des attaques. En effet, un circuit émet de la chaleur lors de son fonctionnement par effet joule et ce de manière corrélée avec les données qu'il manipule. Ce flux est également lié à la nature et à la distribution spatiale des matériaux présents dans le circuit. Ainsi, le flux thermique d'un circuit peut être exploité pour extraire des informations relatives au floorplan et placement du circuit à l'aide de caméras infrarouge [Tes+07]. Cette propriété est très largement exploitée pour guider les attaques par injection de fautes effectuées avec un laser. Il peut également être utilisé pour mener des attaques de type DPA, CPA et MIA. Toutefois, ce

canal auxiliaire est peu usité pour mener des attaques car limité par ses caractéristiques temporelles. En effet, ce rayonnement s'exprime en très basses fréquences, ce qui conduit à des campagnes de mesures extrêmement longues [Hut].

Il est à noter qu'une alternative à la mesure du flux thermique existe. Elle consiste à compter les photons émis par les transistors lors de la recombinaison des porteurs. On parle alors d'analyse en photo-émission qui requière des équipements onéreux et reste caractérisée par un temps d'acquisition long [Sch+12].

1.3.4 Attaques par analyse électromagnétique

Si les procédures statistiques SPA, DPA, CPA et MIA ont été initialement introduites pour exploiter le canal auxiliaire consommation, elles peuvent également être appliquées sur le canal électromagnétique [Car+04]. En effet, la circulation de courants dans les circuits crée inévitablement un champ magnétique dont les variations sont liées au traitement des données.

Ce canal auxiliaire est probablement de nos jours le plus exploité car il offre de nombreux avantages. En premier lieu, il limite la préparation des échantillons. En effet, sa mesure est relativement aisée. Elle ne nécessite que l'utilisation d'une simple sonde (un solénoïde) et d'un amplificateur de tension et ne requière aucun contact avec la cible. En outre, les ondes se propagent aisément au travers des boîtiers des circuits. Il n'est donc pas nécessaire de les retirer.

Un second avantage du canal auxiliaire EM réside dans sa résolution spatiale. En effet, l'emploi de sondes de petites dimensions (jusqu'à $50\mu m$ dans [Ord+09] revient à utiliser un filtre spatial permettant de ne collecter qu'une partie du rayonnement EM du circuit et donc de limiter le bruit de mesure et le bruit algorithmique lié aux calculs, non ciblés par l'attaque, qui sont effectués par la partie non cryptographique du circuit au même moment.

Enfin, un troisième avantage est la très large bande passante des sondes EM qui s'étend jusqu'à $6GHz$. Elle est donc parfaitement adaptée à l'analyse de circuits dont la fréquence d'horloge est très élevée comme les puces de téléphones portables [VMC19]. Pour toutes les raisons évoquées ci-dessus,

ce type d'analyse est considéré comme une menace majeure pour la sécurité des circuits.

1.4 Plateformes et sondes d'analyse électromagnétique

Le première publication scientifique rapportant la réalisation d'une attaque CPA menée sur le canal électromagnétique date de 2001 [GMO01] et a été réalisée par la société GEMPLUS. Il est démontré que mener une telle attaque ne requière que peu de matériel. Seuls suffisent en effet une sonde électromagnétique, un amplificateur de tension large bande et faible bruit et un oscilloscope numérique. La sonde utilisée par les auteurs est en outre particulièrement rudimentaire comme le montre la Figure 1.1. Comme on peut l'observer, il s'agit d'un solénoïde de $3mm$ de long et d'un diamètre de $500\mu m$ fabriqué à la main.

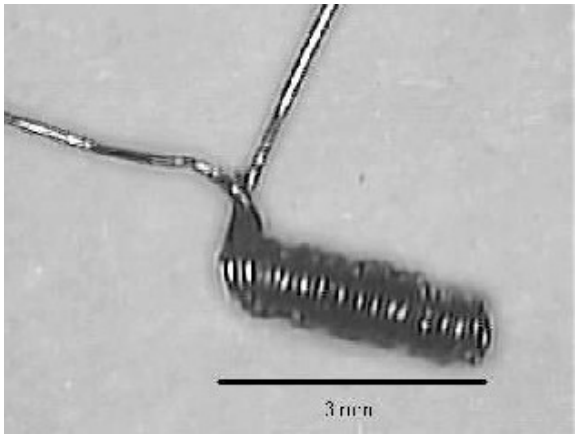


FIGURE 1.1 – Sonde artisanale utilisée dans [GMO01]

Toutefois, utiliser des sondes de telles dimensions ne permet pas d'exploiter au mieux le potentiel des analyses électromagnétiques tant à des fins d'attaque qu'à des fins de retro-ingénierie. En effet, comme indiqué dans les paragraphes ci-avant, utiliser des sondes de dimensions réduites permet de réduire le bruit ce qui augmente l'efficacité des attaques. Cela permet en outre d'obtenir des observations en champ proche magnétique dévoilant plus d'informations sur le "floorplan" des circuits cibles grâce à l'amélioration de la résolution spatiale.

C'est donc naturellement que la communauté a conçu et utilisé des sondes de dimensions de plus en plus petites. Toutefois, cette approche artisanale n'est pas adaptée à une approche industrielle reposant sur la certification des produits selon les critères communs par exemple. En effet, la reproductibilité des attaques, la fabrication des sondes et donc les caractéristiques de celles-ci ne sont absolument pas maîtrisées. En outre, depuis 2001 les technologies CMOS utilisées pour concevoir les cartes à puce et les systèmes sur puce de nos téléphones ont grandement évoluées et ce conformément à la loi de Moore. Ce qui était réalisable avec de telles sondes sur des microcontrôleurs fabriqués en technologie $0.5\mu m$ opérant sous une tension de $5V$ ne l'est plus sur des microcontrôleurs conçus en technologie $40nm$ opérant sous $1.2V$. C'est donc naturellement que les plateformes d'analyse EM utilisées ont rapidement évolué. Depuis le début des années 2010, on peut considérer qu'une plateforme standard consiste en un assemblage de dispositifs industriels tels que :

— un oscilloscope numérique d'une bande passante d'au moins $1GHz$ disposant d'un taux d'échantillonnage élevé (au moins $10GS/s$) et d'une grande profondeur mémoire (de l'ordre de 256 millions de points),

— un amplificateur faible bruit caractérisé par une bande passante d'au moins 1GHz et d'un gain d'au moins 30dB,

— un système de positionnement trois axes des sondes d'analyse de grande précision ($1\mu m$),

— des sondes d'analyse électromagnétique (typiquement fournies par la société Langer) dont le diamètre est compris entre $100\mu m$ et $500\mu m$ illustrées sur la Figure 1.2,

— et éventuellement d'une table anti-vibrations et/ou d'une cage de Faraday, afin de limiter au mieux le bruit de mesure, mais aussi de systèmes optiques pour contrôler le positionnement des sondes à la surface des circuits.

Si ce standard perdure, et s'accompagne d'une amélioration lente de la résolution spatiale (de $150\mu m$ à $100\mu m$ en dix ans), des travaux ont été conduits,

FIGURE 1.2 – Panoplie de sondes d'analyse du champ électromagnétique commercialisées par la société Langer

notamment au Japon, dans le but de développer des sondes ayant des résolutions spatiales accrues et fabriquées selon des processus industriels parfaitement maîtrisés. Parmi ces travaux, on trouve notamment ceux conduits entre 2012 et 2018 par Mai Khanh [MK+12 ; MK+15 ; MK+18].

Ces travaux décrivent l'intégration, sur une puce en technologie $180nm$, d'un système complet de mesure du champ magnétique vertical de grande résolution et de grande bande passante. Ce système sur puce intègre donc sur un même circuit le solénoïde et l'amplificateur. L'amplificateur a une bande passante s'étendant de $17MHz$ à $1.7GHz$ pour un gain de $63dB$. Le solénoïde de $100\mu m \times 100\mu m$ intégré sur la puce est constitué de 5 tours rectangulaires réalisés avec des métallisations Metal 1 de cette technologie. Il peut être surprenant d'avoir réalisé la bobine en Métal 1 compte tenu qu'il s'agit du niveau de métallisation le plus enterré. Toutefois, celle-ci est en fin de fabrication dégagée à l'aide d'un Focused Ion Beam (FIB). Dans ce travail d'une très grande qualité, même le facteur de forme du circuit réalisé a été pensé et défini de manière à favoriser les expérimentations. En effet, celui-ci est en forme de stylo comme l'illustre la Figure 1.3.

Malgré les efforts accomplis, les résultats obtenus sont décevants pour plusieurs raisons. La première est que la sonde résultante collecte le champ

magnétique parallèle à la surface des circuits alors que le champ perpendiculaire à celle-ci est en général exploité lors des attaques. La seconde raison est liée au facteur de forme de la sonde. Comme on peut le constater Figure 1.3 en observant les vis, sa taille ne permet pas d'approcher la sonde de la surface des circuits habituellement inclus dans des cavités rectangulaires de quelques mm^2. Cet inconvénient est majeur car effectuer des mesures à grande distance de la surface du circuit limite de facto la résolution spatiale et le rapport signal à bruit (SNR) des relevés expérimentaux. Enfin, on peut déplorer le coût de fabrication d'une telle sonde, coût qui même en production de masse resterait élevé car faisant appel à un FIB.

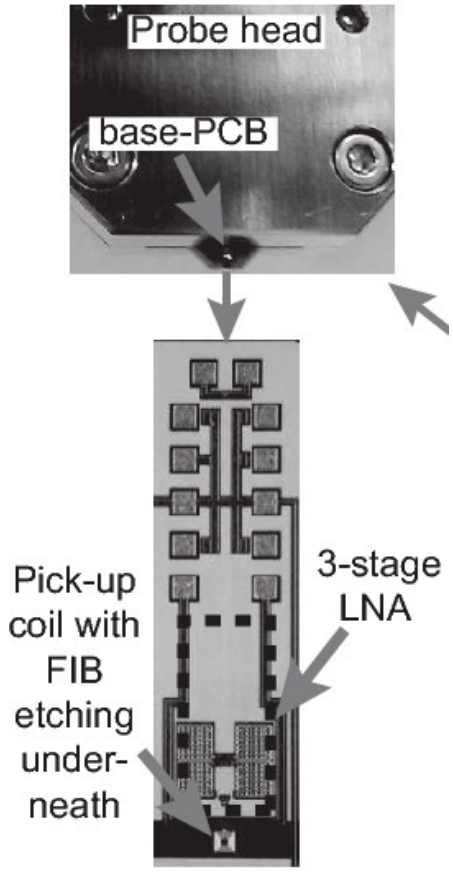


FIGURE 1.3 – Images relatives à la sonde réalisée dans le cadre des travaux décrits dans [MK+12]

1.5 État de l'Art sur les attaques par injection de fautes et techniques d'injection de fautes

Les techniques d'injection de fautes sont largement exploitées dans la communauté de la sécurité matérielle. La résistance des circuits aux attaques par injection de fautes, lors d'évaluations selon des critères communs, est d'ailleurs analysée. Injecter des fautes au sein de circuits intégrés peut être effectué à différentes fins.

1.5.1 Les attaques par injection de fautes

En effet, l'injection de fautes peut être utilisée pour dérouter des codes. Ceci permet par exemple de forcer un code à prendre un branchement plutôt qu'un autre et peut donc être utilisé pour contourner une authentification par code pin ou bien s'octroyer des droits d'administrateur ou autres privilèges sur un produit sécurisé [Vas+17].

Les attaques par injection de fautes peuvent également être utilisées pour dévoiler des secrets manipulés par les circuits tout comme les attaques par observation. On parle alors d'analyse différentielle en fautes dont les principales cibles sont les clés cryptographiques. L'exploitation différentielle et conjointe de résultats de chiffrement corrects et erronés justifie le nom donné à ce type d'attaques. Il existe différentes manières d'effectuer ces attaques différentielles. Sans prétendre à l'exhaustivité, on peut en effet effectuer :

— des analyses différentielles en fautes, au sens strict du terme (DFA) [BS97], [Gir05], qui exploitent les erreurs de calculs provoquées par une perturbation extérieure. Par exemple, dans [BS97], les auteurs récupèrent la clé de chiffrement DES en analysant entre 50 et 200 textes chiffrés erronés, générés à partir de textes non chiffrés inconnus mais connexes,

— des analyses comportementales de fautes exploitant le fait, par exemple, que des injections de fautes altèrent ou n'altèrent pas le résultat final d'un calcul [YJ00]. Il s'agit par exemple de détecter les multiplications inutiles lors de l'exécution d'un RSA de type Square and Myltiply Always [Cor99],

— des analyses de sensibilité des fautes qui, contrairement à la DFA, ne nécessitent pas de connaître la valeur du texte chiffré erroné, ni la nature de la faute injectée dans le circuit. En partant d'une condition donnant un texte chiffré correct, l'attaquant augmente graduellement l'intensité de l'attaque jusqu'à obtenir une faute (qui correspond à ce que les auteurs dans [Li+10] appellent "condition critique"). Cette condition critique, qui correspond à la sensibilité de la faute, met en lumière des dépendances entre les données sensibles manipulées et permet de récupérer la clé de chiffrement secrète,

— des attaques par collisions de textes chiffrés provoquées par injection de fautes [BK06]. Il peut s'agir par exemple de forcer le résultat d'un byte du "Addrounkey" de la première ronde d'un AES à être à égal à '00' et de rechercher ensuite, parmi les 256 possibilités, la valeur de l'octet conduisant au même résultat chiffré pour obtenir un byte de la clé,

— des attaques par modification du nombre de cycles (rondes) d'un algorithme cryptographique [CT05], [Dut+12]. Ceci permet d'avoir accès simultanément au résultat de l'avant dernière ronde et de la dernière ronde d'un AES, et de retrouver aisément la clé utilisée lors de la dernière ronde pour finalement remonter à la clé de chiffrement.

1.5.2 Les techniques d'injection de fautes

Il existe de nombreuses techniques permettant d'injecter des fautes au sein des circuits intégrés. Les paragraphes suivants décrivent les principales.

1.5.3 Perturbation de l'horloge ou glitch d'horloge

Le principe de cette technique d'injection de fautes consiste à introduire un front parasite d'horloge, lors d'une période d'horloge spécifique, de sorte à provoquer un échantillonnage anticipé des données par les bascules DFF et donc de provoquer des fautes dites de timing et, plus précisément, de temps d'établissement des données en entrée de bascule. Ceci est en général réalisé à l'aide d'un générateur d'impulsion de tension. Toutefois, cette méthode d'injection de fautes est peu utilisée de nos jours. En effet, elle n'est applicable que si la source du signal d'horloge est externe au circuit ce qui n'est plus le cas pour les microcontrôleurs sécurisés modernes. En outre, la résolution

spatiale de cette attaque est particulièrement limitée [DCC11]. En effet, la faute est produite à la racine de l'arbre d'horloge et produit donc des fautes sur de nombreuses bascules connectées aux feuilles de celui-ci.

1.5.4 Perturbation de l'alimentation ou glitch d'alimentation

Cette technique consiste à appliquer une impulsion de tension négative sur l'alimentation du circuit cible ou bien positive sur la masse durant l'exécution d'un algorithme. Cette variation induit une réduction de la tension d'alimentation du processeur et, par voie de conséquence, une augmentation des délais de propagation. Un processeur soumis à ce type de perturbation peut alors mal interpréter une instruction, sauter une instruction ou encore effectuer une instruction en manipulant des données erronées [Bar+09], [SGD08], [Bar+10]. Bien qu'efficace, cette technique présente de faibles résolutions spatiale et temporelle. En effet, il est difficile de ne créer une faute que sur un seul élément dans un circuit et les temps de réponse (constante RC) des circuits sont élevés. En outre, des contremesures efficaces [YPA08] contre ce type d'attaques ont été mises en place.

1.5.5 Perturbation de la polarisation de substrat

En 2012, une nouvelle technique d'injection de fautes appelée Body Bias Injection [Mau+12], [Tob+13] a été proposée. Elle consiste à appliquer une impulsion de tension à l'aide d'une pointe de test à la surface du substrat. Cette impulsion dans le substrat se propage alors à travers celui-ci pour atteindre les parties actives du circuit et y générer des fautes de manière relativement locale. Peu d'informations sont, à l'heure actuelle, disponibles quant au mécanisme produisant les fautes. Toutefois, cette technique qui s'est démocratisée avec la commercialisation d'équipements idoines, par les sociétés Langer et Riscure par exemple, semble s'apparenter à de l'injection de "glitchs" de tension locaux. On notera que cette technique ne requière pas d'amincissement du substrat pour être efficace.

1.5.6 Perturbation par illumination

Parmi les techniques d'injection de fautes, les perturbations par illumination des composants sont très utilisées. Selon la gamme de longueurs d'onde considérée pour effectuer ces illuminations, la technique peut être appelée différemment. Parmi ces techniques on trouve en effet des perturbations par

injection laser et par injection de rayons X.

Historiquement, les premières injections de fautes par illumination ont été réalisées avec un flash d'appareil photo et donc en lumière blanche. Ce genre d'illumination génère en effet des photo-courants [SA03]. Ces derniers induisent des perturbations transitoires sur des signaux générés par les portes logiques, qui, si échantillonnés par une bascule, deviennent des fautes.

Compte tenu du fait que le substrat silicium des circuits intégrés est transparent dans le proche infrarouge, c'est naturellement que le laser a été largement adopté comme médium d'injection de faute privilégié afin de bénéficier de ses grandes résolutions spatiale et temporelle. Ces avantages s'accompagnent toutefois de quelques désagréments.

En premier lieu, avec l'évolution technologique et l'augmentation de la densité de métallisation dans les circuits, le laser ne peut plus être appliqué que via la face arrière des circuits et il est parfois délicat d'avoir accès à cette face selon les cibles. En second lieu, pour une efficacité maximale et une focalisation optimale du faisceau sur les jonctions PN des transistors, il est souvent nécessaire d'affiner le substrat des circuits. Enfin, sa popularité s'est accompagné du développement de nombreuses contre-mesures comme par exemple des détecteurs de tirs laser [Bas+13] que l'on retrouve en grand nombre dans certains microcontrôleurs sécurisés. Ceci rend l'utilisation laser de plus en plus complexe et justifie que certains utilisent plusieurs faisceaux laser pour mener des attaques, l'un pour générer la faute et les autres pour désactiver des contremesures.

Récemment, un faisceau de rayons X de forte énergie et fortement focalisé a été utilisé [Anc+17] pour reprogrammer des circuits sécurisés en ciblant des points mémoire flash. Si cette technique requière à l'heure actuelle l'utilisation d'équipement particulièrement onéreux (une source ID16B du synchrotron de Grenoble), ce travail ouvre la porte à l'utilisation de faisceaux de sources moins onéreuses de la marque Hamamatsu ou des sources utilisées dans des tomographes.

1.5.7 Perturbation de la température

Si les techniques d'injection de fautes produisent en général des perturbations extrêmement brèves afin de générer des fautes exploitables lors de l'exécution d'algorithmes cryptographiques, d'autres éléments d'un circuit caractérisés par des comportements longs peuvent être ciblés, comme par exemple des générateurs de nombres aléatoires. Dans ce cas, des phénomènes ou techniques caractérisés par des constantes de temps plus longues peuvent être utilisés. C'est ainsi que dans [Sou+11] des perturbations en température sont utilisées pour biaiser des générateurs de nombres aléatoires à l'aide de simples résistances chauffantes ou de modules Peltier.

1.5.8 Perturbation par médium électromagnétique

L'injection de faute par médium électromagnétique ou EMFI est une technique qui devient de plus en plus populaire et ce notamment pour attaquer des SoCs [Maj18]. Ce regain de popularité s'explique probablement par certaines de ses caractéristiques. En premier lieu, le matériel nécessaire pour effectuer une EMFI est peu onéreux : un millier d'euros pour une plateforme artisanale à deux-trois dizaines de milliers d'euros pour mettre en place une plateforme haut de gamme. En second lieu, cette technique d'injection peut être appliquée sans préparation complexe des circuits, comme le retrait du boîtier plastique ou l'amincissement du substrat, même si cela permet d'obtenir de meilleurs résultats. Enfin, son utilisation permet de contourner les détecteurs de tirs laser qui sont aujourd'hui très répandus.

Améliorer les résolutions spatiale et temporelle ainsi que la puissance des plateformes EMFI étant l'un des objectifs centraux de ma thèse, je propose d'approfondir l'étude de cette technique à travers un historique de celle-ci dans la prochaine section.

1.6 État de l'Art sur les attaques par injection de fautes EM et plateformes associées

1.6.1 Historique des attaques par médium électromagnétique de 2000 à 2020

L'injection de fautes par médium électromagnétique est une technique d'attaque bien connue et relativement récente qui devient de plus en plus populaire. Sa popularité grandissante s'explique notamment par le grand nombre d'avantages qu'elle offre. Le principal étant sa capacité d'injecter des fautes dans les circuits au travers du boîtier plastique. Elle a été initialement suggérée au début des années 2000 dans [Sam+02], [QS02] comme une menace potentielle pour les circuits électroniques sécurisés.

Dans ces articles, des injections EM sont effectuées avec une bobine réalisée en enroulant un fil autour d'une pointe de très faible diamètre. La bobine ainsi réalisée est connectée à un flash d'appareil photo afin de créer de très fortes variations du champ magnétique induisant selon les auteurs un courant de Foucault dans le circuit cible. Les auteurs affirment que les courants induits au dessus de mémoires RAM et Flash n'ont pas les mêmes caractéristiques selon que des '1' ou '0' sont stockés dans les points mémoires. Plus particulièrement, le temps de retour à la normale de la tension d'alimentation de la puce est suggéré comme étant un indicateur des valeurs mémorisées. Toutefois, les auteurs ne parviennent pas à effectuer une lecture efficace du contenu des mémoires par manque de matériel. Ces articles suggèrent également que ce type de dispositif pourrait être utilisé pour induire des fautes dans les circuits.

Après ce premier avertissement quant à la dangerosité des impulsions EM comme moyen d'injection de fautes, il faut attendre 2007 pour voir publier des résultats plus concrets dans [SH07]. Une approche différente pour générer une forte variation du champ magnétique y est proposée. Comme illustré sur la Figure 1.4, elle consiste à utiliser un allume gaz produisant un arc électrique et donc une impulsion EM intense et de courte durée. Cette technique conduit à une variation très rapide du courant ainsi qu'à un très fort rayonnement électromagnétique. Les attaques conduites sur des implémentations de l'algorithme RSA–CRT ont été une réussite, qu'elles aient été

conduites sur des circuits encapsulés ou non.

FIGURE 1.4 – Allume gaz utilisé comme injecteur EM dans [SH07].

Deux ans plus tard, en 2009, des résultats supplémentaires ont été publiés dans [Ala+09], où la résistance à l'injection de fautes EM de différents circuits a été testée. Il a été mis en évidence que les circuits intégrés étaient non seulement sensibles aux champs magnétiques, mais également aux champs électriques. Les auteurs ont également conclu que la susceptibilité d'un circuit intégré par rapport aux émissions EM reçues ne venait pas seulement des fils de câblage et du boîtier, mais était également due au couplage EM entre les injecteurs et la puce elle-même et qu'il est possible d'altérer les temps de propagation des signaux dans les puces. Pour ces travaux, les auteurs ont utilisé une plateforme dite Continous Waves (CW) couramment utilisée dans le monde de la compatibilité électromagnétique. Ce type de plateforme est constituée d'une source RF et d'un amplificateur HIFI forte puissance. Les sondes magnétiques et électriques utilisées ont été conçues à partir de câble coaxial.

C'est avec ce type de plateforme et un injecteur générant un fort champ électrique (du type pointe de test) que les auteurs de [Pou+11] montrent qu'il est possible d'injecter de manière locale de la puissance au sein des circuits et de produire des comportements fautifs. Plus précisément, les auteurs montrent qu'il est possible de modifier la fréquence de fonctionnement du circuit cible qui intègre un générateur embarqué d'horloge dont la fréquence est fixée par le contenu d'un registre. Ils démontrent donc in fine la capacité

de l'injection EM à induire des erreurs dans des registres au repos.

Un an plus tard, avec les mêmes équipements, les auteurs mettent en évidence dans [Bay+12] qu'il est possible de biaiser le flux binaire d'un générateur de nombres aléatoires de Wold [WT08] intégré sur un FPGA. Plus précisément, ils démontrent que les 50 oscillateurs en anneau de leur générateur censés osciller librement et indépendamment les uns des autres peuvent être synchronisés par application d'un champ électrique continu. Il s'agit d'un phénomène semblable au principe de Huygens relatif à des horloges se synchronisant lorsque celles-ci sont accrochées à un même support.

La même année, des attaques de Piret et Quisquater sont menées sur des implémentations matérielle et logicielle de l'AES [Deh+12b] et [Deh+12a]. L'implémentation matérielle est embarquée dans un microcontrôleur AVR. L'implémentation logicielle est mappée sur un FPGA Xilinx. Pour mener ces attaques, les auteurs ont utilisé une bobine magnétique et un générateur d'impulsion Avtech pour injecter une perturbation localisée sans venir se placer au contact du circuit ciblé.

Concernant l'injection de fautes sur le microcontrôleur, il est démontré qu'il est possible de "fauter" de manière indépendante et contrôlée chacun des bytes de l'AES grâce à la modification de l'instant d'injection. Enfin, il est suggéré que les fautes obtenues sont dues à des sauts d'instructions. Les expérimentations sur FPGA ont, quant à elles, démontré la possibilité d'injecter des fautes simple-bit et multi-bits lors du calcul de l'AES. Il est souligné que ceci est rendu possible en ajustant finement l'amplitude de la tension délivrée à la sonde d'injection ou encore en jouant sur la position de celle-ci. Enfin les auteurs suggèrent que les fautes induites par une EMFI sont des fautes de timing. Leur suggestion tient du fait qu'un détecteur d'attaques en glitch d'alimentation basé sur l'observation du délai de propagation a été déclenché par l'EMFI.

Dans [Deh+13], une analyse d'addition de round a été conduite sur une implémentation logicielle de l'AES. Des fautes ont été induites en injectant un glitch électromagnétique très court dans un microcontrôleur 32-bit ARM cortex-M3. En se focalisant sur le dernier round d'opération, les auteurs ont réussi à forcer l'exécution d'un round supplémentaire (en sautant l'instruction concernant l'incrémentation du compteur). Ces fautes leur ont permis de

récupérer la clé de chiffrement secrète avec seulement deux paires de textes chiffrés corrects et fautés.

Si les travaux précédents ont suggéré que les fautes induites par les injections EM étaient de type faute de timing, ce résultat a été remis en cause en 2015 dans [Ord+15]. Dans ce premier papier se focalisant uniquement sur l'effet des EMFI sur les circuits, les auteurs y relatent une expérience conduite sur un FPGA dans lequel une FIFO avait été implémentée, ainsi qu'un arbre d'horloge pouvant être désactivé. L'expérience conduite a alors consisté à charger la FIFO avec un contenu choisi puis à arrêter l'horloge. Après l'arrêt de l'horloge une EMFI a été appliquée avant de réactiver le signal d'horloge et d'examiner le contenu de la FIFO. L'examen de ce contenu par les auteurs mettant en évidence des fautes ne pouvant pas être expliquées par des erreurs de synchronisation (fautes de timing) puisque l'horloge était arrêtée pendant les tirs EM, les auteurs conclurent que les fautes induites par les injections EM suivent un modèle plus complexe que la simple faute de timing. En outre, lors de leurs expériences, les auteurs observèrent que selon la polarité de l'injection, les fautes étaient soit des forçages à 1 soit des forçages à 0. Enfin, il est à noter que les résultats reportés dans cette publication ont été menés avec des injecteurs EM, visibles sur la Figure 1.5, ayant différentes topologies, et ce, afin de vérifier les préconisations reportées dans [Oma+13]. Comme on peut le constater, toutes les sondes ont été fabriquées à la main et possèdent un noyau de ferrite.

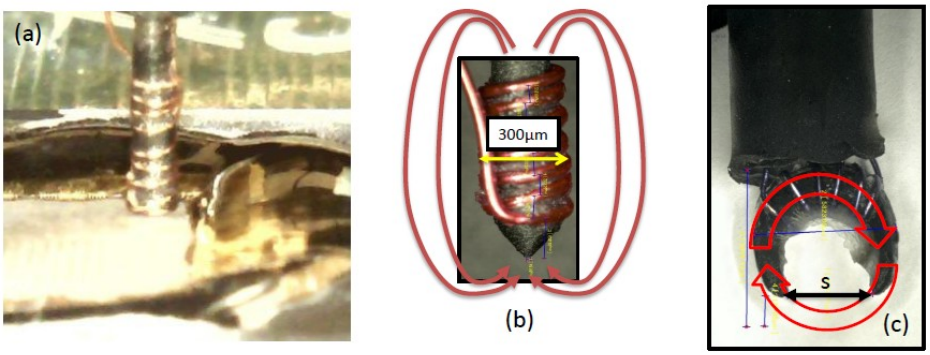


FIGURE 1.5 – Sondes d'injection EM utilisées dans [Ord+15].

Suite à ces travaux, le laboratoire de sécurité de GEMALTO disposant de la même plateforme que celle utilisée dans [Ord+15] met en évidence en

2016 [MBB16] la complémentarité de l'analyse EM et des injections EM pour mener rapidement des attaques sur des systèmes sur puce, et définit une méthodologie pour mener de manière efficace et rapide des tests de pénétration dans des SoC.

En 2016 et 2017, les auteurs de [Ord+15] poursuivent leurs travaux relatifs à l'effet des EMFI sur les circuits intégrés et publient dans [OGM15; OGM17] un modèle empirique appelé le Sampling Fault Model. Celui-ci est schématisé par la Figure 1.6 extraite de [OGM17]. Ce modèle, déduit de nombreuses expérimentations conduites tant sur FPGA que sur des microcontrôleurs STM32F439, stipule que les éléments des circuits intégrés les plus susceptibles aux EMFI sont les bascules DFF et que leur fonctionnement est directement altéré par celles-ci. Il en découle que la susceptibilité électromagnétique des circuits intégrés est périodique et de période égale à celle du signal d'horloge, et est particulièrement élevée dans de courts intervalles de temps avant les fronts montant de l'horloge. Induire des fautes par médium EM durant ces intervalles de temps nécessite des tirs moins puissants.

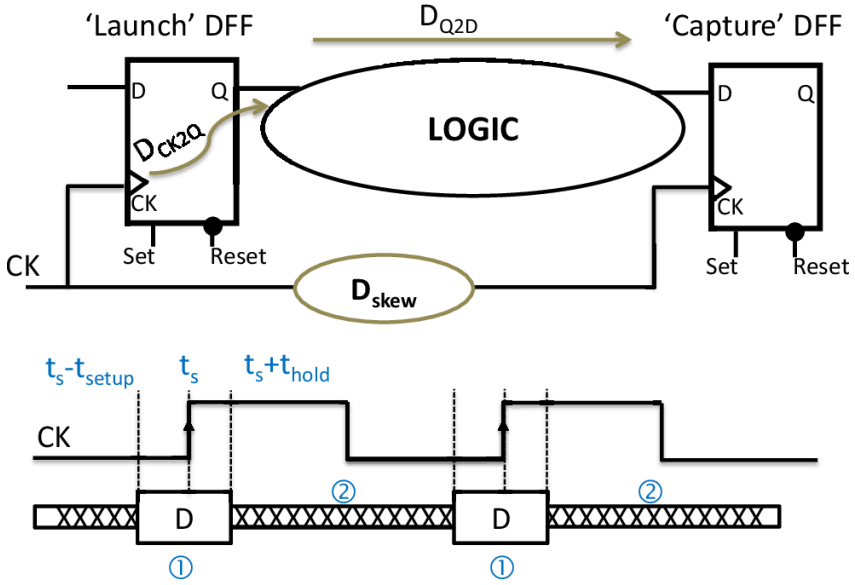


FIGURE 1.6 – Le sampling fault model décrit dans [OGM15; OGM17]

Le modèle empirique appelé Sampling Fault Model est finalement confirmé en 2019 et 2020 par les travaux de modélisation reportés dans [DLM19; DLM21]

qui expliquent comment les fautes de sampling se produisent dans les bascules mais également par [TSD19] qui reporte qu'une EMFI n'altère que très peu (9ns d'accroissement de délai pour 600 impulsions EM soit $\pm$15ps par impulsion EM environ) le délai d'un chemin purement combinatoire.

On notera que les années 2017 à 2020 sont également des années où de nouvelles plateformes EMFI émergent dans la littérature comme la plateforme faible coût développée par KUL [BAM17a], celle commercialisée par la société canadienne NewAE [O'F19] appelée ChipShouter ou encore celle développée par la société Ledger [AH20a] appelée SiliconToaster.

1.7 Plateformes et sondes d'injection de fautes électromagnétique

Durant les années qui ont suivi ces travaux précurseurs, de nombreuses plateformes d'injection de fautes électromagnétique ont été mises au point et sont présentées dans plusieurs publications, [AH20b ; BAM17b ; Bay+12 ; CH17 ; Deh+12a ; Mau12 ; Oma+13 ; O'F19 ; BAM17a ; AH20a]. A ces plateformes s'ajoutent celles commercialisées depuis 5 ans environ par les sociétés Riscure et Langer visibles sur la Figure 1.7.

Parmi l'ensemble de ces propositions, on distingue deux types de plateformes : celles dites pulsées et celles dites harmoniques ou CW (Continuous Waves). Les plateformes pulsées produisent une forte impulsion magnétique alors que les plateformes harmoniques génèrent des variations de formes arbitraires ou sinusoïdales du champ magnétique à l'aide d'un générateur arbitraire ou bien d'une source RF disposant d'options de modulation en amplitude.

Quel que soit leur type, la majorité de ces plateformes s'appuient sur les principes de l'induction magnétique pour générer des fautes et certaines sur le principe du couplage électrique ou capacitif. Celles s'appuyant sur le principe d'induction magnétique, qui sont celles nous intéressant dans cette thèse, ne sont finalement que des moyens de générer, idéalement au plus proche de la surface des circuits cibles, une variation importante et soudaine du champ magnétique. Ce flux magnétique variable traversant les milliers de boucles métalliques formées par les réseaux d'alimentation et de masse des

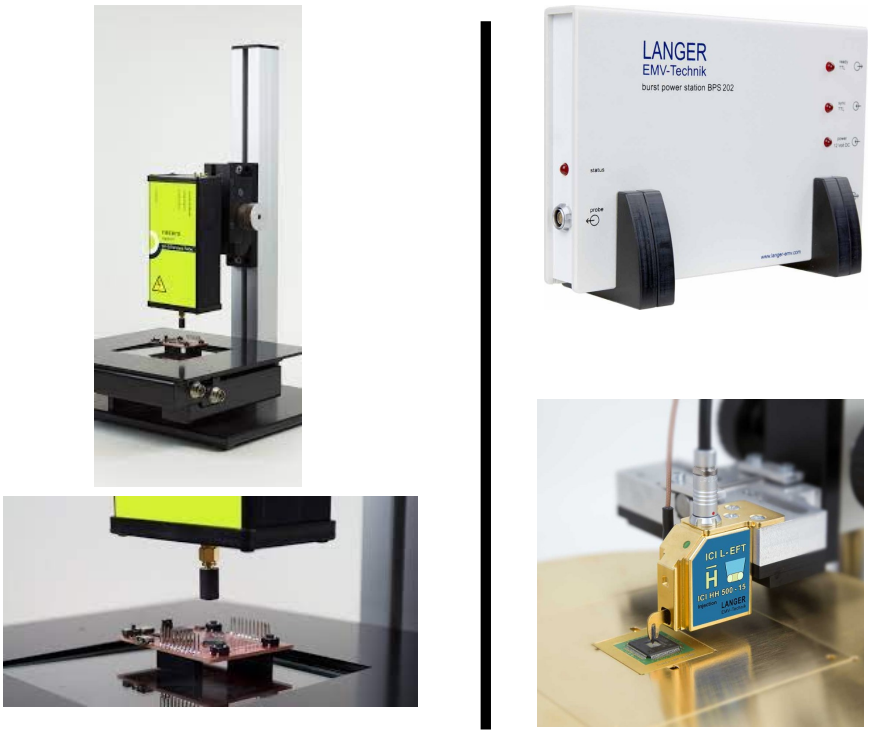


FIGURE 1.7 – Plateformes EMFI commercialisées par Riscure (à gauche) et Langer (à droite).

circuits crée aux bornes de chacune d'elles une force électromotrice et donc un courant parasite ou courant induit. L'ensemble de ces courants provoque en retour des variations internes et transitoires de la tension d'alimentation locale et des fautes.

Académique ou commerciale, une plateforme EMFI est constituée en général de trois éléments : une chaîne de contrôle, une chaîne de puissance et une chaîne d'acquisition.

La chaîne de contrôle, qui reste optionnelle mais nécessaire pour mettre en place des attaques répétables est en général constituée de deux éléments. Le premier est un banc motorisé trois axes XYZ afin de pouvoir positionner avec une grande précision et de manière répétable les sondes EMFI au dessus de la surface des circuits. Le second est un ordinateur contrôlant l'ensemble des équipements de la plateforme, et donc les chaînes de puissance et d'acquisition, mais émulant également le circuit cible.

La chaîne de puissance constitue le cœur de la plateforme. Dans le cas

d'une plateforme pulsée, elle est constituée d'un générateur d'impulsion de tension de forte puissance (plusieurs kW). Les impulsions qu'il génère sont de fortes amplitudes (plusieurs centaines de volts), de courte durée (de $5ns$ à $20ns$) et sont caractérisées par des fronts de montée et descente de courtes durées ($1ns$ à $2ns$). Cette impulsion est en général directement fournie au second élément de la chaîne d'injection à savoir une sonde EM qui transforme celle-ci en impulsion électromagnétique. Dans le cas d'une plateforme harmonique, cette chaîne d'injection est constituée d'une source RF ou bien d'un générateur de signaux arbitraires qui alimente un amplificateur HIFI large bande. Le signal amplifié est ensuite délivré à la sonde d'injection. L'intérêt de ce type de chaîne d'injection est de pouvoir parfaitement contrôler la forme et la durée du signal fourni à la sonde mais également de produire des trains rapides d'impulsions par exemple. Toutefois, ce genre de chaîne d'injection est en général limitée en puissance et relativement onéreuse. Il est en effet difficile, voire impossible, de trouver dans le commerce des amplificateurs HIFI travaillant sur une bande de fréquences d'au moins $1GHz$ et d'une puissance supérieure à $100W$.

La chaîne d'acquisition est une chaîne accessoire fort utile pour une plateforme EMFI. Elle permet en effet d'observer les comportements analogiques des circuits cibles afin par exemple de mieux contrôler les injections EM. Plus particulièrement, elle permet de positionner de manière idéale les tirs EM par rapport au déroulement d'un algorithme cryptographique ou encore l'apparition de latch-up dans les circuits cibles. Elle est en général composée d'un oscilloscope numérique, d'un amplificateur de tension faible bruit et enfin de diverses sondes de mesures permettant de mesurer le courant ou bien le champ EM émis par les circuits cibles.

Au début de ma thèse, plusieurs solutions existaient pour se doter d'une plateforme EMFI. La plus simple et rapide était d'en acheter une auprès des sociétés Riscure (depuis 2014) ou Langer (depuis 2016). Toutefois, les plateformes fournies par Riscure sont dotées de sondes dont la résolution spatiale est très faible, tandis que les plateformes Langer sont limitées en terme de précision temporelle et disposent d'injecteurs EM extrêmement fragiles et coûteux à réparer.

La seconde solution consistait à en créer une en se dotant des différents équipements nécessaires comme un générateur d'impulsion Avtech, comme

l'on fait [Deh+12a], [Oma+13] et [Ord+15] et de fabriquer nos propres injecteurs EM. Cette solution présente cependant deux inconvénients de taille. Le premier est la fabrication des sondes qui reste artisanale et limitée en résolution spatiale. Le second est lié à la désadaptation d'impédance entre le générateur et la sonde d'injection. En effet, un générateur d'impulsion tel que celui de la marque Avtech est conçu pour être relié à une charge 50Ω, alors que les sondes EM possèdent une impédance très faible (de l'ordre de l'Ohm voire inférieure). Du fait de cette désadaptation d'impédance, l'impulsion fait des allers-retours le long de la ligne de transmission, depuis le générateur jusqu'à la sonde et inversement. On observe alors des rebonds, d'amplitude de plus en plus faible jusqu'à atténuation totale de l'impulsion. Ces rebonds limitent grandement la résolution temporelle des plateformes d'injection électromagnétique qui utilisent cette solution.

La troisième approche est un peu plus compliquée à mettre en œuvre, et consiste à concevoir son propre générateur d'impulsion. De cette manière, l'adaptation d'impédance peut être réalisée de manière très précise [AH20b], [BAM17b], [CH17], [O'F19]. Cependant, cette méthode nécessite beaucoup d'investissement et de temps à mettre en œuvre. De plus, pour obtenir cette adaptation d'impédance parfaite, il faut nécessairement que le générateur d'impulsion soit situé au plus proche de la sonde. Cela pose notamment problème lorsque l'on souhaite réaliser le balayage d'un circuit électronique (pour tracer une carte de susceptibilité électromagnétique par exemple), puisqu'il faudrait alors déplacer à la fois le générateur et la sonde lors de l'acquisition des données.

1.8 Contexte et objectifs de la thèse

Ma thèse s'est inscrite dans le contexte du projet FUI CSAFE+ s'articulant autour de trois grands objectifs à savoir : l'amélioration des plateformes EMFI, la modélisation et la simulation des EMFI et enfin la définition de contre-mesures matérielles et logicielles. Ce projet qui s'est déroulé de Janvier 2017 à Juin 2021, a été le cadre de collaboration avec divers partenaires académiques et industriels : Telecom Paris, le CEA, STMicroelectronics, SecureIC, Arelis et Thales. Il est à noter également que les travaux que j'ai pu mener durant ma thèse l'ont été de sorte à répondre aux engagements pris

par le LIRMM dans le cadre d'un Plan d'Étude Amont contracté avec la Direction Générale de l'Armement et ce en partenariat avec la société SecureIC.

Dans ce contexte, les objectifs de ma thèse ont été les suivants :

— développer, avec un flot de fabrication maîtrisé et reproductible, des sondes d'analyse du champ magnétique dotées d'un préamplificateur, afin de maintenir des valeurs de SNR élevées, d'une résolution spatiale de $50\mu m$ et d'une bande passante d'au moins $1GHz$. Les sondes fabriquées devaient en outre répondre à des exigences liées à la pratique des analyses du canal caché électromagnétique comme la capacité de pénétrer les petites cavités (boîtier plastique) au fond desquelles se trouvent les circuits intégrés,

— développer, avec un flot de fabrication maîtrisé et reproductible, des sondes d'injection EM d'une résolution spatiale de $300\mu m$, les sondes fabriquées devant répondre aux mêmes exigences liées à la pratique que les sondes d'analyse,

— accroître tant la résolution temporelle que la puissance des plateformes EMFI de sorte à pouvoir injecter des fautes dans des SoCs,

— définir des protocoles simples de caractérisation des sondes d'analyse et d'injection.

Il est à noter qu'au début de ma thèse, le LIRMM disposait d'une plateforme EMFI centrée autour d'un générateur Avtech de 200V et de sondes fabriquées à la main dont la résolution spatiale était limitée à $700\mu m$ ainsi que d'une plateforme d'analyse dont la résolution spatiale était celle de la sonde Langer RF3mini, à savoir $300\mu m$.

Chapitre 2

Sondes d'analyse

2.1 Résumé

Ce chapitre présente l'intégralité des travaux réalisés pour l'amélioration de la résolution spatiale des sondes d'analyse électromagnétique. Il présente également la conception et la caractérisation d'un pré-amplificateur faible bruit en technologie AMS 0.35μm destiné à fabriquer des sondes actives, et devant permettre des mesures avec un SNR élevé. L'ensemble des travaux reportés dans ce chapitre a fait l'objet d'une publication [Tou+21a].

2.2 Introduction

La fabrication des microcontrôleurs est en constante évolution vers des technologies CMOS de plus en plus agressives. Il est donc nécessaire d'adapter les outils de caractérisation sécuritaire, comme par exemple les sondes d'analyse électromagnétique, utilisées pour la rétro-ingénierie ou la caractérisation sécuritaire des circuits. L'objectif principal est de concevoir et fabriquer des sondes d'analyse électromagnétique efficaces, ayant des résolutions spatiales de l'ordre de 50μm. Il est alors cohérent de considérer que le signal mesuré par ces sondes est très faible, d'où la nécessité d'utiliser un pré-amplificateur faible bruit pour amplifier celui-ci. Des travaux dans ce domaine [AKY06], [MK+15] ont été menés et proposent des amplificateurs différentiels faible bruit et des bobines différentielles, tous deux intégrés directement sur le silicium.

Cependant, cette approche possède un inconvénient majeur : le volume important occupé par le circuit intégré comprenant l'amplificateur et la sonde, qui ne permet pas d'approcher la bobine assez proche de la surface du circuit ciblé, limitant de facto grandement la résolution spatiale de ce genre de

prototype. Pour résoudre cette limitation importante, il a donc été décidé de séparer la partie sonde de la partie pré-amplificateur. La première partie de ce chapitre traitera de la conception et de la fabrication des sondes, la deuxième portera sur la conception et la fabrication du pré-amplificateur faible bruit, tandis que les suivantes présenteront les résultats obtenus lors des différentes caractérisations réalisées.

2.3 Sondes d'analyse flexibles

L'analyse électromagnétique du champ émis par les circuits intégrés est une pratique commune dans le domaine de la sécurité matérielle, surtout dans le cadre de la rétro-ingénierie ou bien pour retrouver des clés de chiffrement. Puisque les technologies CMOS continuent de devenir de plus en plus agressives, il émerge une vraie nécessité de réduire les dimensions des capteurs réalisant ces analyses. Dans ce contexte, la suite de cette partie explore le potentiel de l'électronique flexible et de l'impression 3D pour le design et la fabrication de sondes d'analyse électromagnétique faible coût, isolées électriquement et mécaniquement robustes (de manière à pouvoir venir au contact de la surface du circuit), ayant des résolutions spatiales de l'ordre de $50\mu m$, le standard étant actuellement autour de $100\mu m$.

2.3.1 Conception des sondes

L'évolution des technologies d'électronique flexible, tout comme les technologies CMOS, est agressive. Elles permettent désormais de dessiner des lignes de métaux sur quatre couches, avec une largeur de $50\mu m$ et un espacement minimal de $50\mu m$. Malgré l'agressivité de cette technologie, le coût de fabrication reste modéré. Il est d'à peu près 1000€ pour quatre circuits de $200cm^2$. Cela correspond à environ 800 sondes décrites dans la suite de ce document. Ainsi, plusieurs formes de sondes avec des tailles différentes ont été conçues et fabriquées.

Le premier run consistait à tester la preuve de concept, c'est la raison pour laquelle deux sortes de sondes ont été dessinées. La première, que l'on pourrait appeler "sonde classique", se compose d'une piste de métal réalisant une ou plusieurs boucles sur les différents niveaux de métaux. Cette dernière est illustrée Figure 2.1 (a).

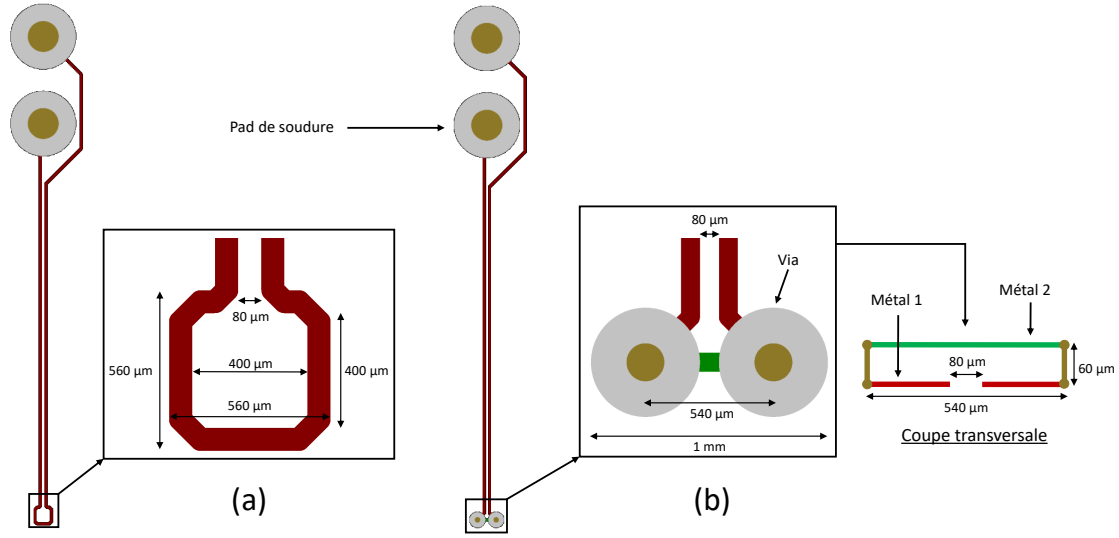


FIGURE 2.1 – (a) Sonde classique et (b) sonde dans l'épaisseur avec coupe transversale

La deuxième, que l'on a appelé "sonde dans l'épaisseur" consistait, comme son nom l'indique, à réaliser la boucle à l'intérieur du substrat du circuit, permettant ainsi de réduire de manière notable les dimensions de celle-ci. Elle est illustrée Figure 2.1 (b), avec une coupe transversale pour faciliter la visibilité. Cependant, comme on peut le voir, de telles sondes nécessitent l'utilisation de vias, qui occultent en grande partie la boucle destinée à récupérer le signal EM. Les premiers tests expérimentaux n'ont pas du tout été concluants, c'est pourquoi dans le run suivant et dans la suite de ce chapitre, seules les sondes se basant sur le design "classique"ont été considérées.

A la suite de plusieurs caractérisations, celles possédant les meilleures caractéristiques ont été sélectionnées et sont présentées Figure 2.2.

Sur cette figure, on peut s'apercevoir que pour les trois sondes, la boucle n'est plus circulaire comme c'était le cas dans les travaux précédents, mais carrée avec des "diamètres" internes égaux à $500\mu m$, $150\mu m$ et $50\mu m$. Elles ont été dessinées toutes trois avec le même objectif, qui consistait à supprimer le plus possible les boucles parasites nécessairement créées par les fils connectant la boucle de la sonde aux pads de soudure (nécessaire pour la connexion avec des équipements externes). Ceci est d'une importance capitale si l'on souhaite fabriquer des sondes possédant une très haute résolution spatiale et récolter un signal avec des valeurs de SNR très élevées.

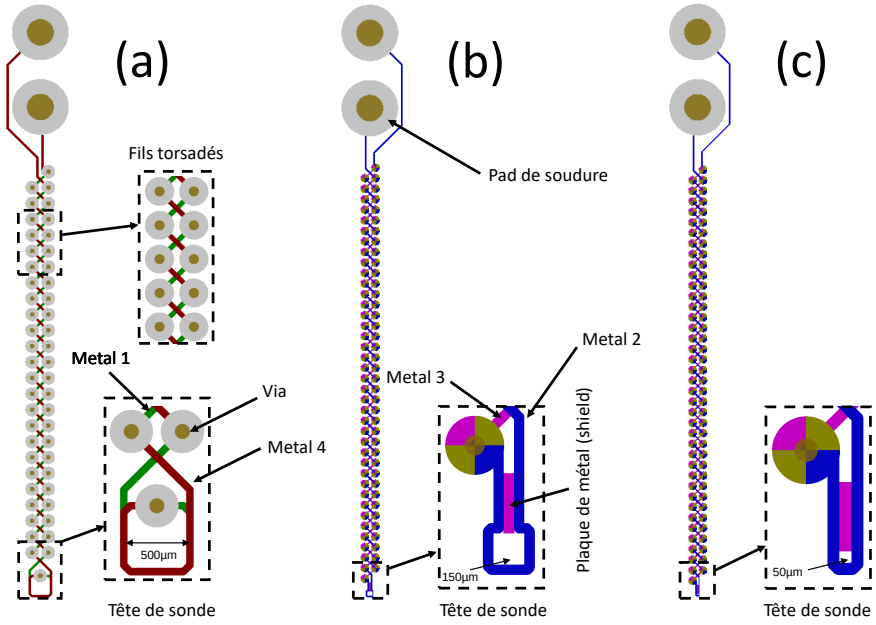


FIGURE 2.2 – Sondes flexibles possédant un diamètre interne égal à (a) 500μm et 2 boucles, (b) 150μm et 1 boucle, (c) 50μm et 1 boucle

Plusieurs solutions ont été testées en ce sens, parmi lesquelles l'utilisation de plaques de métal servant de protection (shield). Cependant, les résultats expérimentaux ont montré que la méthode la plus efficace est de torsader les fils. Ce résultat sera démontré un peu plus loin dans ce chapitre.

Il a ensuite fallu développer un support rigide pour ces sondes flexibles. Pour cela, la technologie de l'impression 3D a été explorée et utilisée pour obtenir une sonde finale avec un facteur de forme pratique. Ce support a été modélisé de telle sorte à ce que l'on puisse placer la boucle de la sonde le plus proche possible de la surface du circuit, qui se trouve généralement dans une petite cavité suite à la décapsulation de celui-ci (généralement aux moyens de procédés physico-chimiques).

A la suite du premier run (qui ne comportait que 2 couches), plusieurs angles de pliage ont été testés, et nous avons estimé que l'angle de pliage maximum pour ne pas casser les pistes de métaux était de 135°. Ainsi, pour fabriquer une sonde horizontale tout en conservant un facteur de forme pratique, le support 3D visible sur la Figure 2.3 (b) a été modélisé puis imprimé. Celui-ci possède deux trous à sa base, permettant d'y insérer un connecteur mâle 01x02 et de souder la sonde dessus. Il possède également un petit plateau à son autre extrémité, sur lequel sera collée la boucle de la sonde.

Ensuite, pour la connexion de celle-ci à des équipements externes (oscilloscope, amplificateur, etc.), un PCB rigide et très simple a été gravé. Comme on peut le voir sur la Figure 2.3 (a), il s'agit d'un connecteur mâle 01x02 relié à un connecteur SMA.

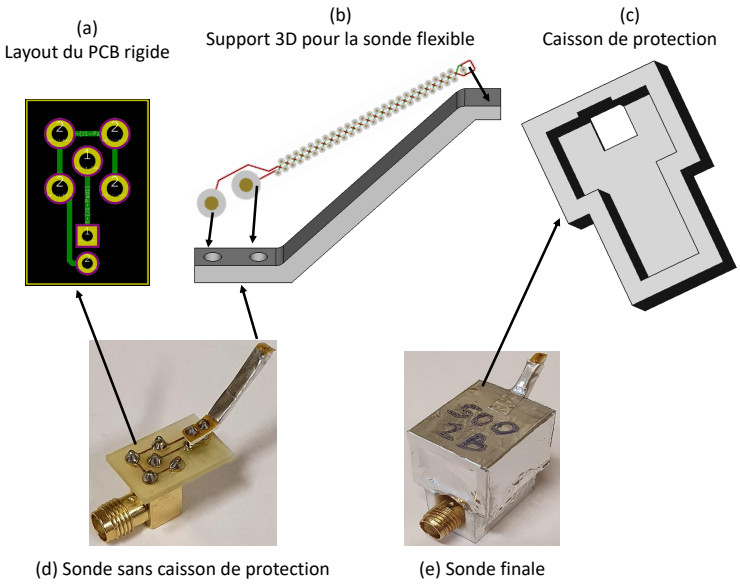
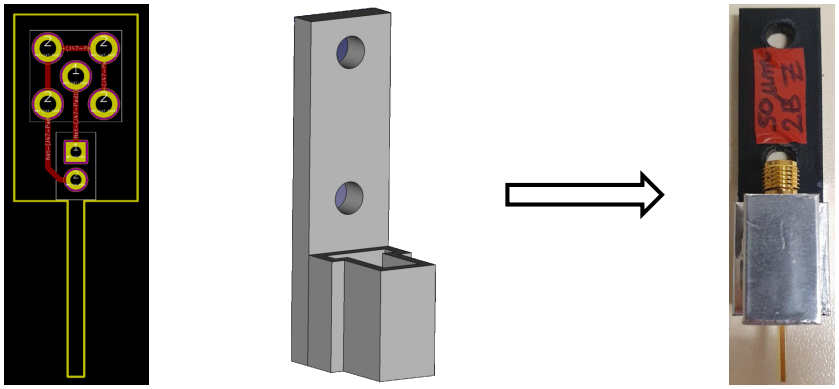


FIGURE 2.3 – Un exemple de sonde flexible avec son support et sa protection imprimés en 3D

Les premiers tests réalisés n'ont pas été concluants car trop de bruit était récolté par les sondes en question. C'est pourquoi un caisson de protection a été modélisé et imprimé, de manière à y insérer la totalité du circuit, en ne laissant apparents que la tête de la sonde et le connecteur SMA. Ce boîtier plastique a ensuite été recouvert de ruban adhésif en aluminium pour bloquer toutes les émanations parasites.

Le processus de fabrication complet est le suivant. On soude en premier lieu le connecteur SMA et le connecteur mâle 01x02 sur le PCB rigide précédemment gravé. On vient ensuite insérer sur ce connecteur le support 3D destiné à accueillir la sonde. Celui-ci est fixé solidement grâce à un point de colle. Après avoir choisi et découpé la sonde voulue, celle-ci est placée sur le connecteur et collée sur son support avant d'être soudée. Enfin, la sonde non protégée obtenue Figure 2.3 (d) est insérée à l'intérieur de son caisson de protection. Ce travail de modélisation et de fabrication a mené aux sondes présentées sur la Figure 2.3 (e). L'imprimante 3D utilisée est une Ultimaker3.

Lors du deuxième run, une nouvelle solution a été envisagée. Il s'agissait de réaliser exactement les mêmes sondes mais sur un circuit à quatre couches, et de tracer les pistes sur les couches internes (métal 2 et métal 3) au lieu des couches externes (métal 1 et métal 4). L'objectif de ce nouveau design est de pouvoir plier les sondes flexibles avec un angle plus important que le précédent, pour diminuer encore plus l'encombrement. Ceci étant rendu possible par les couches isolantes qui entourent les pistes, et qui diminuent de fait les contraintes mécaniques sur celles-ci. Après quelques tests rapides, il s'avère qu'avec ce nouveau design, il est possible de plier avec un angle de 90° sans rompre les pistes de métaux. Un nouveau PCB rigide et un nouveau caisson de protection ont donc été conçus pour l'occasion et sont illustrés Figure 2.4 (a) et (b).



(a) Layout du PCB rigide (b) Caisson de protection (c) Sonde finale

FIGURE 2.4 – Prototype de sonde flexible avec pliage à 90°

Le PCB rigide est sensiblement le même, à ceci près qu'il possède en plus une petite extension qui servira de support à la sonde et viendra donc remplacer le précédent. Le caisson de protection est également semblable au précédent. Le seul ajout est la partie supérieure percée de deux trous, permettant l'insertion de deux vis M4 pour fixer les sondes sur des équipements Newport et Thorlabs. Ici, contrairement aux prototypes précédents, la sonde est collée sur son support, mais la longueur de celui-ci a été choisie de manière à ce que la boucle dépasse. Ainsi, le pliage à 90° et le collage de la boucle sont relativement aisés. Un aperçu du prototype final est donné Figure 2.4 (c).

Enfin, de manière à pousser la protection de la sonde à son maximum, de la peinture métallique a été utilisée pour peindre le corps de la sonde et ne laisser que la boucle non protégée.

2.3.2 Caractérisation

Parmi les expérimentations menées à la suite de la fabrication des sondes, la première avait pour but de vérifier l'absence de boucles parasites dans le prototype de sonde final. C'est à ce moment-là que l'efficacité des fils torsadés a été prouvée, comme le montre la Figure 2.5.

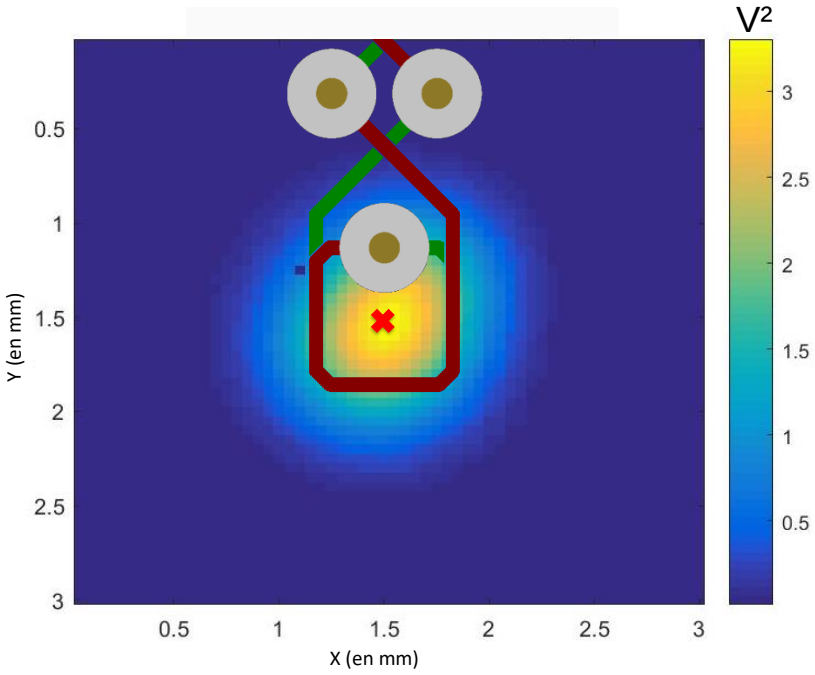


FIGURE 2.5 – Cartographie du carré de l'amplitude maximale (V^2) du signal récupéré par une sonde flexible de $500\mu m$ de diamètre à 2 boucles

Cette figure montre une cartographie du carré de l'amplitude maximale du signal récupéré par une sonde flexible de $500\mu m$, lorsqu'une impulsion EM est générée, proche de sa boucle, à une distance $h = 50\mu m$ de sa surface. Cette impulsion EM est générée en utilisant la sonde RF3 mini de Langer (de diamètre $300\mu m$) à laquelle est délivrée une impulsion EM d'amplitude et de largeur égales à $200V$ et $8ns$ respectivement. Comme attendu, l'amplitude du signal mesuré est élevée lorsque l'impulsion est générée juste au-dessus et autour des boucles de la sonde, et quasi nulle partout ailleurs, y compris lorsque l'impulsion est générée juste au-dessus des fils torsadés.

La deuxième expérimentation visait à mesurer la bande passante des sondes. Pour ce faire, un signal sinusoïdal d'amplitude $15dBm$ a été généré à l'entrée de la sonde RF3 mini de Langer. Cette dernière a été placée au contact des boucles des différentes sondes flexibles. L'amplitude du signal collecté a ensuite été mesurée entre $12.5MHz$ et $2GHz$. La Figure 2.6 donne le diagramme de Bode du système composé de la sonde Langer et de la sonde flexible. On peut remarquer une fréquence de coupure aux alentours de $75MHz$ pour les trois sondes. En sachant que la sonde RF3 mini de Langer possède une réponse en fréquence relativement plate entre $30MHz$ et $3GHz$, il semblerait que les sondes flexibles aient une bande passante qui s'étend de $75MHz$ jusqu'à plus de $2GHz$.

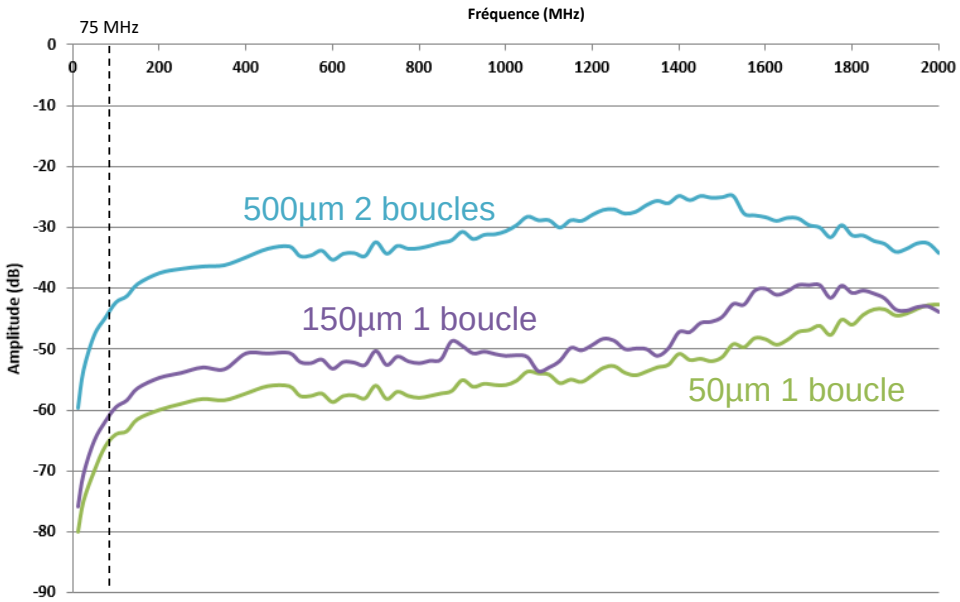


FIGURE 2.6 – Fonctions de transfert des systèmes composés de la sonde Langer RF3 mini et de chacune des trois sondes flexibles

Pour terminer, une validation fonctionnelle et pratique a été effectuée. La mesure du champ magnétique émis par un FPGA (Spartan3 1600E) dans lequel a été programmé un AES s'exécutant à $50MHz$ a été réalisé avec les trois sondes précédemment présentées. Pour ces balayages, le pas de déplacement a été fixé à $50\mu m$ et un amplificateur $40dB$ et $200MHz$ (quatre fois la fréquence d'horloge de l'AES) du fabricant Femto, visible Figure 2.7, a été utilisé.

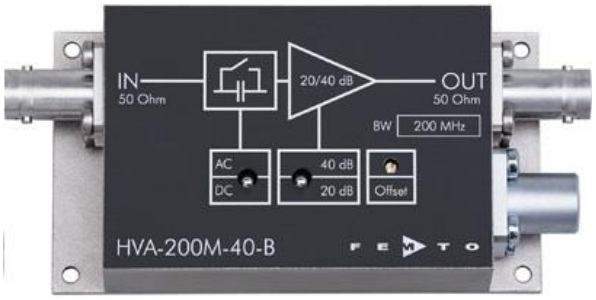


FIGURE 2.7 – Amplificateur du fabricant Femto

Les sondes ont été placées au contact du circuit, dont le boîtier plastique a été préalablement retiré. On se rend compte ici du réel avantage des sondes fabriquées qui sont électriquement isolées et mécaniquement robustes, permettant ainsi une mise en contact avec la surface du circuit. Enfin, il est important de noter que lors de ces balayages, l'AES chiffrait en continu le même texte clair, de manière à ensuite calculer le SNR en utilisant l'estimateur non biaisé introduit dans [DM18], pour finalement obtenir les cartographies SNR tracées sur la Figure 2.8. Sur celles-ci, les zones aux valeurs de SNR élevées représentent les zones du circuit actives lors de l'exécution de l'algorithme cryptographique. On y retrouve notamment les pads d'alimentation (à droite) et la position de l'AES (au centre).

L'échelle de couleurs des deux cartographies inférieures (correspondant aux sondes de diamètres égaux à $150\mu m$ et $50\mu m$) sont identiques de manière à simplifier la comparaison tandis que la dynamique de la cartographie supérieure (pour la sonde de $500\mu m$) est doublée. Les rectangles rouges donnent une vue simplifiée du placement du design, et non pas seulement la position de l'AES, qui lui occupe les deux rectangles centraux.

L'observation principale qui peut être faite sur les cartographies (b) et (c) concerne les zones bleues foncées, marquées par les chiffres (1) et (2). Comme on peut le voir, elles sont relativement similaires en forme, mais sont plus larges et plus sombres sur la cartographie correspondant à la sonde de diamètre $50\mu m$. On a ici la première preuve directe d'une plus grande résolution spatiale pour cette sonde.

En effet, elle distingue mieux ces différentes zones grâce à sa taille plus réduite. Au contraire, ces zones disparaissent partiellement sur la cartographie (a), malgré le doublement de l'échelle de couleurs. Cela s'explique aisément

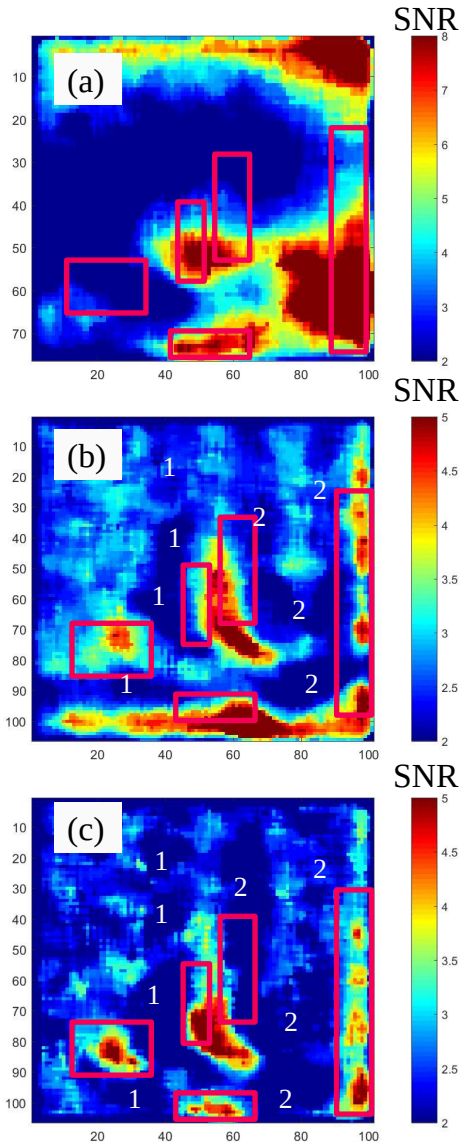


FIGURE 2.8 – Cartographies SNR obtenues avec les sondes flexibles de diamètre (a) 500μm, (b) 150μm et (c) 50μm

par les dimensions importantes de la sonde utilisée. En effet, sa surface représente environ 10% de la surface du circuit, elle récupère donc le signal d'autres sources dans le design, dont notamment celui des pads d'alimentation. C'est pourquoi il serait très difficile d'étudier précisément une partie du design qui se situerait dans le rectangle de droite avec cette sonde, à cause de la forte émission des pads d'alimentation. L'utilisation de la plus petite sonde ici serait alors pertinent, car comme on peut le voir sur la cartographie (c), les zones sont plus distinctes que sur la cartographie (a). Pour conclure, ces premiers résultats démontrent clairement que l'électronique flexible peut être utilisée pour concevoir et fabriquer des sondes d'analyse électromagnétique à faible coût, tout en améliorant la résolution spatiale.

2.4 Conception du pré-amplificateur

Cette partie présente le design d'un pré-amplificateur différentiel intégré faible bruit, développé en technologie AMS 0.35μm. Il a été conçu pour amplifier le signal récupéré par une sonde électromagnétique passive de 50μm (une bobine conçue séparément) utilisée pour la caractérisation sécuritaire des circuits électroniques. Pour cette application spécifique, il est nécessaire d'avoir un bruit équivalent en entrée très faible ainsi qu'une bande passante la plus élevée possible. Ce pré-amplificateur a été conçu à la fois en technologie CMOS et en technologie BiCMOS, pour démontrer s'il y en a, les intérêts potentiels de la technologie BiCMOS pour cette application spécifique.

2.4.1 Spécifications

L'amplificateur Femto utilisé pour les expérimentations précédentes est trop encombrant pour être placé proche de la sonde. L'idée est donc de développer un pré-amplificateur intégré qui sera ensuite utilisé directement en tant que puce nue, et connecté à la sonde à l'aide de "wire bonding" pour obtenir un prototype final de sonde active (le tout en réduisant les capacités d'interconnexion). Cette approche devrait amener à un minimum de perturbations externes, causées habituellement par les longs câbles coaxiaux servant à la connexion, et donc à une bande passante plus importante. Il sera placé au plus près de la sonde, de manière à pré-amplifier directement le signal et le bruit du capteur. On évite ainsi l'acheminement du signal brut vers un amplificateur déporté via un long câble, qui ne fait qu'accroître le bruit de mesure et détériore le SNR.

L'architecture du pré-amplificateur proposée pour cette application est sensiblement la même que celle présentée dans [AKY06] et [MK+15], c'est-à-dire un amplificateur différentiel composé de plusieurs étages, d'un buffer de sortie et d'une génération de la tension de polarisation également intégrée sur la puce. Contrairement à [AKY06], la sortie est différentielle et ne nécessite aucune capacité supplémentaire entre les étages, puisque les composantes DC et basses fréquences seront filtrées après pré-amplification. Si un gain supplémentaire est nécessaire pour une meilleure visualisation du signal, l'amplificateur à gain variable, implémenté dans [MK+15], pourra être remplacé par n'importe quel amplificateur COTS. Il sera branché à la sortie du pré-amplificateur précédent, soit au plus près dans le cas d'un amplificateur intégré soit à quelques dizaines de centimètres dans le cas d'un

amplificateur tel que le Femto. Cette politique a mené à la définition des spécifications suivantes pour le pré-amplificateur :

— Entièrement différentiel pour permettre une amplification supplémentaire avec un amplificateur COTS et une connexion de la sonde entre les deux entrées

— Gain minimum : $40dB$

— Bande passante maximale avec un minimum fixé à $200MHz$, d'où le choix d'un pré-amplificateur faible gain en boucle ouverte

— Consommation maximum : $100mA@3.3V(0.33W)$

— Impédance de sortie : 50Ω

— Offset d'entrée maximum : $10mV$

— Plancher de bruit équivalent en entrée (bruit thermique) : $1nV/\sqrt{Hz}$

Dans la suite, les architectures choisies et les flots de conception suivis pour respecter les spécifications précédentes sont présentés.

2.4.2 Premier run : Mai 2018

Architecture de l'implémentation CMOS

Pour l'application visée, un pré-amplificateur de gain $40dB$ est suffisant. La bande passante doit être la plus élevée possible, avec pour limite basse $200MHz$ (à cause de la fréquence de chiffrement des algorithmes cryptographiques). Le bruit doit être le plus faible possible tout en gardant une consommation raisonnable.

Comme expliqué précédemment, et à la lumière de ce qui existe de nos jours en terme d'amplificateur faible bruit et large bande pour caractérisation sécuritaire, une architecture entièrement différentielle semble appropriée de manière à connecter la sonde passive entre les deux entrées. La première idée a donc été de concevoir un pré-amplificateur ne possédant qu'un seul étage. Cependant, l'application visée nécessite un plancher de bruit très bas. En

négligeant le bruit basse fréquence, pour une résistance de sortie de valeur R, la densité de bruit blanc en sortie N est donnée par :

$$N = \sqrt{4kTR} \tag{2.1}$$

Avec k la constante de Boltzmann et T la température en Kelvin. Avec quelques approximations, le gain G d'un amplificateur à un étage est :

$$G \approx gm * R \tag{2.2}$$

Avec gm la transconductance de l'amplificateur. On peut alors déduire la densité de bruit équivalente en entrée à l'aide des expressions précédentes, en divisant l'équation 3.1 par l'équation 3.2, ce qui donne :

$$N_{input} = \frac{N}{G} = \frac{\sqrt{4kT}}{\sqrt{R} * gm} \tag{2.3}$$

De plus, si on travaille en régime de forte inversion, on peut écrire :

$$gm = \frac{2 * I}{V_{eff}} \tag{2.4}$$

Et en supposant un amplificateur de type paire différentielle à charge résistive :

$$R * I = V_{mc} \tag{2.5}$$

Où V_{eff} est la tension effective de grille (V_{gs}-V_t), V_{mc} la tension de mode commun de sortie et I le courant de polarisation de la paire différentielle. En remplaçant l'équation 3.4 et l'équation 3.5 dans l'équation 3.3, on arrive à :

$$N_{input} = \frac{V_{eff}}{2 * V_{mc}} * \sqrt{4kTR} \quad \alpha \sqrt{R} \tag{2.6}$$

Finalement, le bruit d'entrée équivalent est proportionnel à $\sqrt{R}$, ce qui signifie que la résistance de charge doit être réduite au maximum pour obtenir un bruit d'entrée équivalent très faible, ce qui rend impossible la réalisation d'un gain de $40dB$ avec seulement un étage. Il a donc été décidé de concevoir un pré-amplificateur différentiel à plusieurs étages. Ils sont au nombre de trois, de manière à respecter les spécifications en terme de bande passante.

Si l'on considère qu'un gain G=4 convient pour le premier étage, et puisque que l'on a deux résistances, il faut que :

$$\sqrt{2} * N_{input} = 1nV/\sqrt{Hz} \tag{2.7}$$

Ce qui nous amène à :

$$R = \frac{10^{-18} * G^2}{8 * k * T} = 483\Omega \tag{2.8}$$

En prenant une marge de conception, des résistances de 250Ω pour le premier étage ont par conséquent été choisies. L'architecture utilisée est présentée Figure 2.9.

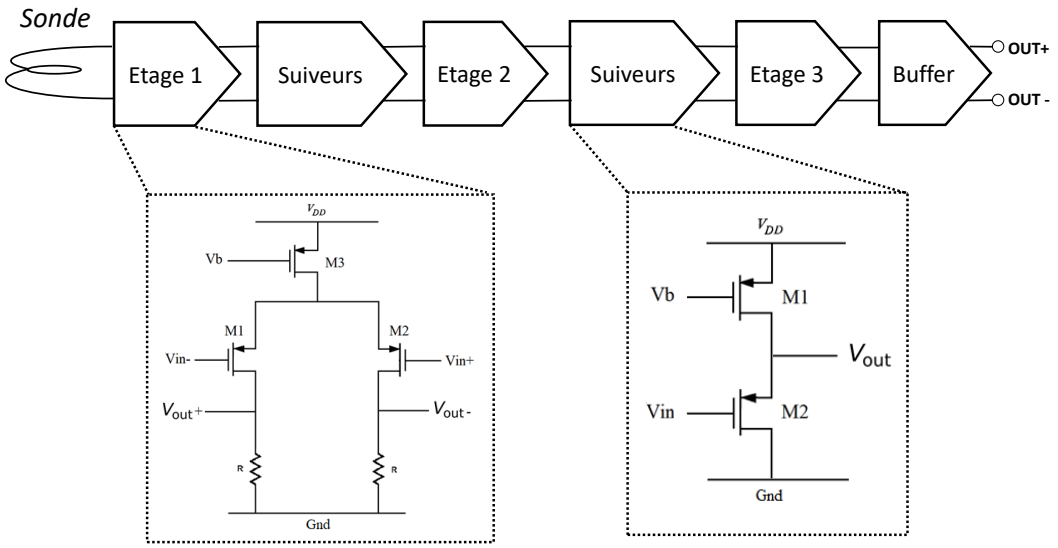


FIGURE 2.9 – Schéma bloc et schémas électroniques du pré-amplificateur CMOS conçu

Le premier étage est le plus important en terme de bruit puisqu'il est connecté directement à la sonde passive. Les résistances de cet étage ont été prises égales à 250Ω comme expliqué précédemment, c'est donc par conséquent un étage faible gain. Les étages 2 et 3 ont été dimensionnés de manière à obtenir le gain de $40dB$ désiré (avec des résistances de 500Ω et $1k\Omega$ respectivement, puisque les contraintes de bruit sont beaucoup moins importantes sur ces étages).

La consommation n'est pas une contrainte principale, mais elle doit rester raisonnable. De plus, puisqu'une grande variation de gain n'est pas problématique tant que les $40dB$ sont atteints, une architecture en boucle ouverte a

été choisie afin de maximiser la bande passante. En effet, une imprécision de $\pm 20\%$ sur le gain en tension du montage petit-signal serait acceptable. Ainsi, la tension de mode commun d'entrée sera contrôlée grâce à l'ajout de deux entrées supplémentaires, et non en utilisant le « Common Mode Feedback ». Ces deux entrées supplémentaires permettront également de réduire l'offset potentiel, à l'aide d'une simple résistance.

Ensuite, l'influence de chaque bloc sur le bloc précédent a été étudiée. Plusieurs simulations ont été effectuées pour déterminer la capacité d'entrée de chaque étage (S_i) et la capacité d'entrée de suiveurs de tension potentiels (F_i) qui pourraient être ajoutés si nécessaires. Le Tableau 2.1 donne les résultats de cette étude.

TABLE 2.1 – Capacité d'entrée des différents blocs

	S_1	F_1	S_2	F_2	S_3	F_3
C_{load} (pF)	5.38	0.12	1.07	0.12	1.08	0.12

Comme on peut le remarquer, la capacité d'entrée des suiveurs de tension est de 10 à 50 fois plus faible que celle des étages eux-mêmes. L'ajout de ceux-ci entre chacun des étages aura donc pour effet d'augmenter significativement la bande passante totale du pré-amplificateur. Mais ces suiveurs de tensions ont également une deuxième utilité. Ils agissent en tant que décaleurs de tension, de manière à assurer le bon fonctionnement de l'étage d'amplification suivant, en lui fournissant une tension de mode commun d'entrée comprise dans la gamme de fonctionnement du pré-amplificateur.

Enfin, de manière à contrecarrer les fluctuations de la tension de polarisation nécessairement induites par une alimentation externe, il a été décidé de générer cette tension directement sur la puce en utilisant un bloc déjà existant. Ce bloc (appelé BBIAS) génère un courant constant, et est couplé à un miroir de courant pour obtenir la tension de polarisation désirée.

Ce pré-amplificateur a été fabriqué en technologie CMOS $0.35\mu m$ majoritairement pour des raisons de coûts. De fait, celui-ci a été fabriqué dans le cadre d'un Wafer multi-projet de manière à réduire encore plus le coût de fabrication.

Résultats de simulations

Le Tableau 2.2 donne les résultats de simulation en fonctionnement typique pour le design réalisé. Pour obtenir ces résultats de simulations, une capacité de charge C_{load}=10pF a été considérée. Elle est censée représenter les pads de la puce ainsi que les câbles coaxiaux (SMA ou BNC) qui serviront à relier le pré-amplificateur à un oscilloscope digital. On verra dans la suite que cette valeur est bien en dessous de la réalité, menant de fait à des résultats expérimentaux décevants en termes de bande passante. Une capacité de charge cinq fois plus élevée sera par conséquent considérée lors du design du second run.

Dans les simulations qui suivent, le plancher de bruit a été mesuré à $100kHz$. Là encore, il aurait été plus judicieux d'effectuer la mesure à plus haute fréquence, pour tenir compte notamment de la fréquence de chiffrement des algorithmes cryptographiques (qui fonctionnent à des fréquences dépassant le MHz). C'est également pour cela que pour le second run, la mesure sera effectuée à $50MHz$.

TABLE 2.2 – Résultats de simulation en fonctionnement typique pour le design CMOS du run 1

Gain	Bande passante	Plancher de bruit	Consommation
44 dB	202 MHz	1.87 nV/$\sqrt{Hz}$	28.65 mA

Ces résultats sont en grande majorité en accords avec les spécifications fixées dans la partie 2.4.1. En effet, le gain est supérieur aux $40dB$ désirés, la bande passante est respectée (malgré le fait que la capacité de charge considérée soit trop faible pour être réaliste) et la consommation est bien inférieure aux $100mA$ fixés. Seul le plancher de bruit est légèrement supérieur aux spécifications, mais il reste néanmoins bien inférieur à celui de l'amplificateur du fabricant Femto que l'on utilisait jusqu'à maintenant (4.5 nV/$\sqrt{Hz}$).

Il est ensuite obligatoire d'effectuer des simulations de Monte-Carlo, de manière à considérer le mismatch des composants et les variations de process. En effet, l'implémentation boucle ouverte implique une grande variabilité, même si des dimensions de transistors importantes atténuent cette variation. C'est pourquoi des simulations de Monte-Carlo ont été effectuées sur

200 circuits, pour donner les résultats résumés dans le Tableau 2.3.

TABLE 2.3 – Simulation de Monte-Carlo sur 200 circuits pour le design CMOS du run 1

	Min	Max	Moyenne
Bande passante (-3dB)	126.7 MHz	330.8 MHz	205.7 MHz
Gain	76.34 (37.7dB)	259.6 (48.3dB)	159.7 (44.1dB)
Bruit (@100kHz)	1.774 nV/$\sqrt{Hz}$	3.379 nV/$\sqrt{Hz}$	2.35 nV/$\sqrt{Hz}$

On retrouve bien la grande variabilité du gain évoquée précédemment et qui est due à l'implémentation en boucle ouverte, ce n'est donc pas surprenant. Il est à noter ici que la totalité des circuits ont passé avec succès la simulation de Monte-Carlo. On a dès lors pu passer à la réalisation du layout.

Layout

Le pré-amplificateur étant organisé en plusieurs étages consécutifs connectés entre eux, c'est naturellement que la réalisation du layout a été effectuée étage par étage. Une attention particulière a été portée à la disposition des entrées et sorties de chaque bloc, pour permettre une connexion entre les différents étages relativement aisée. On retrouve ainsi, pour chaque bloc, les quatre entrées (2 $IN+$ et 2 $IN-$) sur la partie inférieure et les deux sorties sur la partie supérieure. Les rails d'alimentation (VDD et GND) sont quant à eux situés de part et d'autre du layout, pour faciliter la connexion avec les pads de la puce. Le layout final est le résultat de l'empilement des différents étages, avec le bloc générant la tension de polarisation V_{bias} situé à l'autre extrémité du layout, dans un but d'optimisation de l'espace occupé. Un aperçu du layout des différents étages et du layout final est donné Figure 2.10.

Le design complet a été envoyé en Mai 2018, mais le circuit a été reçu en Janvier 2019 avec plusieurs mois de retard suite à une erreur lors de la fabrication, indépendante de notre volonté.

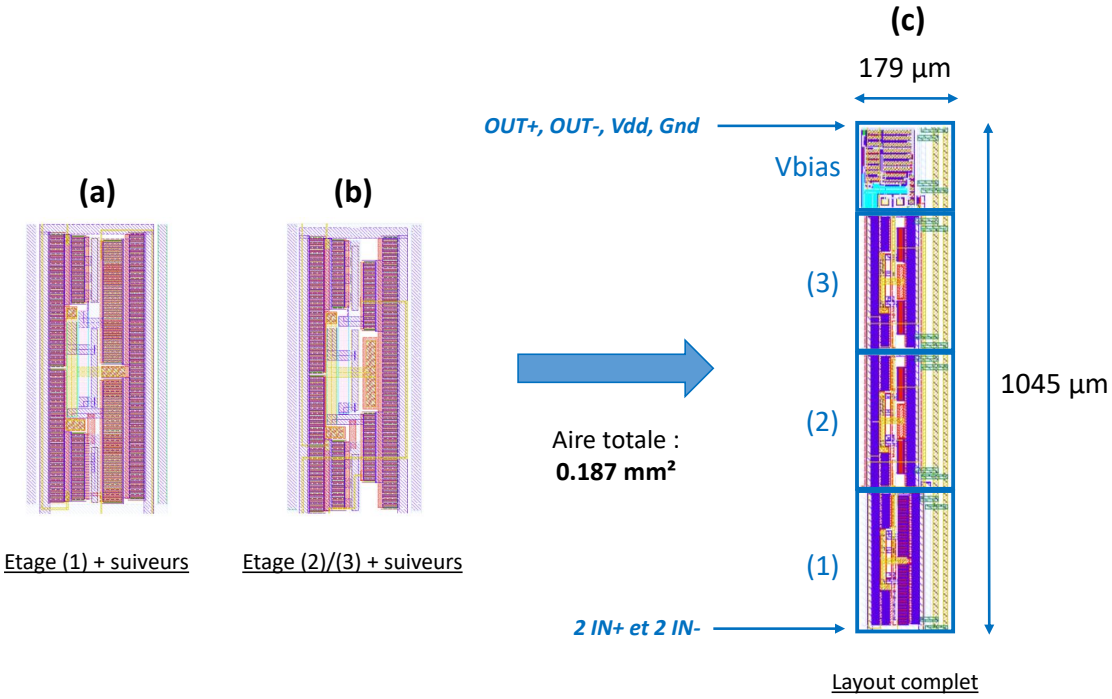


FIGURE 2.10 – Layout du pré-amplificateur CMOS conçu

Test expérimental et caractérisation en boîtier

La livraison du circuit comprenait 10 puces en boîtier (DIL48) et 50 puces nues. La première étape a été le test et la caractérisation du circuit en boîtier. Pour cela, un PCB spécifique a été conçu et gravé avant réception du circuit. Il est illustré Figure 2.11.

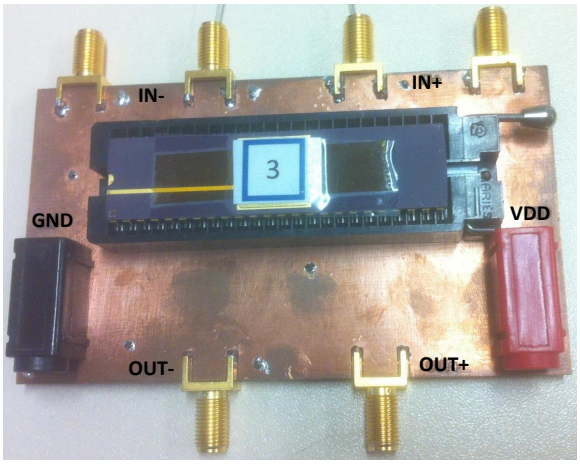


FIGURE 2.11 – PCB réalisé pour le test du pré-amplificateur en boîtier

Il se compose d'un support ZIF pour faciliter le changement des 10 circuits et pour éviter toute casse, de deux connecteurs banane pour l'alimentation (VDD et GND) et de quatre connecteurs SMA bord de carte. Deux d'entre eux correspondent aux sorties $OUT+$ et $OUT-$ et seront connectés à l'oscilloscope digital. Les quatre autres correspondent aux entrées (2 pour chacune d'elle). De cette manière, des signaux pourront être générés sur chaque entrée et une résistance pourra être ajoutée entre les deux pour atténuer le signal, et ainsi éviter une potentielle saturation du pré-amplificateur.

Pour mener la caractérisation, deux signaux sinusoïdaux d'amplitude 10 mV ont été générés sur chaque entrée. Ils ont été réglés en opposition de phase, et la tension de mode commun d'entrée a été fixée pour assurer le bon fonctionnement du pré-amplificateur. Pour chacun des 10 circuits, la consommation est comprise entre 27 et 30mA et le gain entre 46 et 47dB, ce qui respecte les spécifications de départ et est en accord avec les résultats de simulation. En revanche, en traçant les diagrammes de Bode pour chaque circuit, on s'est rendu compte que la bande passante à $-3dB$ n'était que de $5MHz$, soit bien en dessous de nos attentes et totalement en désaccord avec les simulations.

Cette bande passante très faible s'explique comme indiqué précédemment par la capacité de charge qui est bien supérieure à 10pF (celle considérée dans les simulations). En effet, celle-ci est importante, d'une part à cause des plans de masse, mais aussi à cause de la longueur des pistes, du support ZIF, du boîtier DIL mais surtout à cause des câble coaxiaux servant à la connexion avec l'oscilloscope digital (100pF/m de câble). En effet, si l'on considère les câbles utilisés pour la caractérisation, qui faisaient 50cm (donc 50pF), on ne peut espérer qu'une bande passante maximale égale à :

$$BW_{out} = \frac{1}{2 * \pi * 50\Omega * 50pF} = 64MHz \tag{2.9}$$

Ceci a conduit à la réalisation d'un deuxième prototype.

Deuxième prototype de caractérisation : utilisation de la puce nue

De façon à réduire significativement cette capacité de charge, mais également pour diminuer l'encombrement du futur prototype de sonde active, il a

été décidé d'utiliser une puce nue et non plus le circuit en boîtier. L'utilisation de la puce nue permet de se débarrasser du support ZIF et du boîtier DIL. Pour atténuer/supprimer la charge amenée par les câbles coaxiaux servant à la connexion à l'oscilloscope, nous avons réalisé un amplificateur d'instrumentation à l'aide de deux amplificateurs opérationnels (LMH6702) possédant une très faible capacité d'entrée (1.5pF). Celui-ci a été placé au plus près de la puce nue, de façon à diminuer de manière significative la longueur des pistes. Enfin, les prototypes réalisés ne possèdent plus de plans de masse, ceux-ci étant très déconseillés lors de la réalisation de PCB devant fonctionner à haute fréquence.

Concernant la puce, elle est collée sur le PCB à l'aide de laque d'argent, puis du "wire bonding" est effectué en salle blanche pour connecter les pads de la puce aux pistes du PCB. L'ensemble étant extrêmement fragile, une résine UV vient recouvrir entièrement la puce et les fils de bonding une fois le processus précédent réalisé. Un aperçu de l'un des prototypes réalisés est donné sur la Figure 2.12, avec un zoom sur la partie de la puce nue et des fils de bonding. Les parties plastiques servent à la fixation sur le banc 3 axes de la plateforme d'analyse électromagnétique.

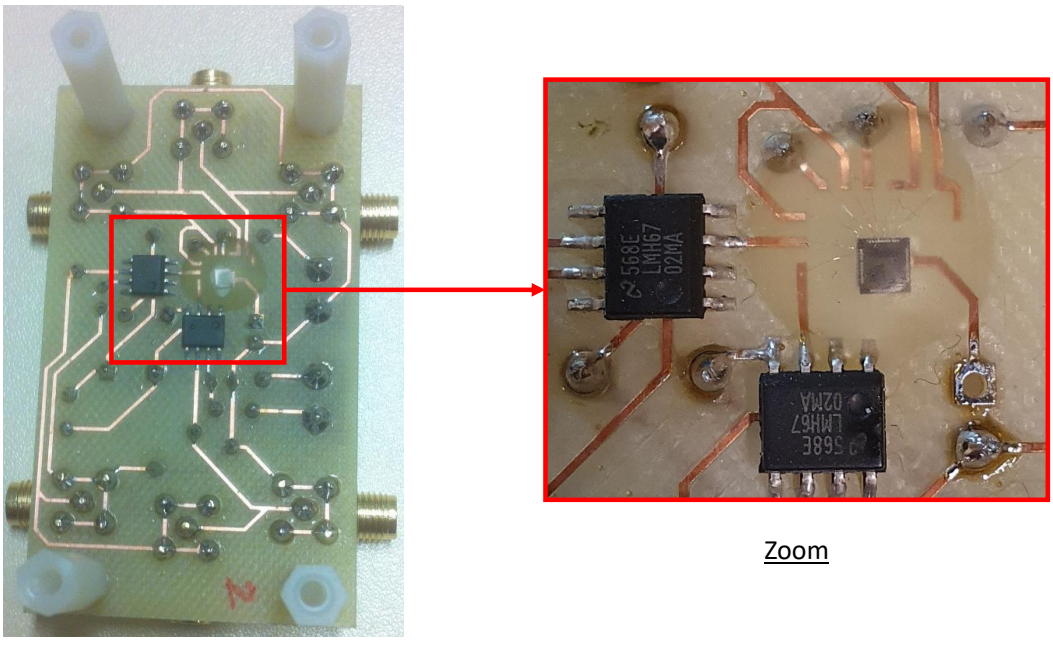


FIGURE 2.12 – Prototype utilisant directement la puce nue

A la suite de toutes ces modifications, une légère amélioration a été constatée, mais sans dépasser une bande passante de 20MHz. Cela peut être expliqué partiellement par une erreur survenue en toute fin de conception. En effet, une fois le design terminé, nous avons envoyé le layout à la personne en charge de réaliser la couronne de plots sans préciser ceux que nous voulions. Celle-ci a utilisé des plots 50Ω à la fois en entrée et en sortie. Cela a provoqué une double limitation en fréquence, avec l'ajout d'un pôle en entrée et d'un pôle en sortie. Notre pré-amplificateur possède une impédance de sortie de 50Ω, donc l'ajout de ces plots a pour conséquence de la doubler, et d'ainsi diviser par 2 la bande passante due au pôle de sortie, en plus de la limitation en entrée. Une attention toute particulière sera donc portée au choix des plots lors du deuxième run.

2.4.3 Deuxième run : Octobre 2019

Amélioration de l'implémentation CMOS

Afin d'améliorer l'implémentation CMOS précédente, une étude sur la chaîne d'amplification complète a été réalisée. Elle a permis d'identifier le dernier suiveur de tension comme étant le bloc limitant dans l'architecture. En effet, dès lors qu'une valeur réaliste de capacité de charge C_{out} est considérée ($50pF$ pour les pads et les câbles coaxiaux), la bande passante chute, comme on peut le voir dans le Tableau 2.4, à cause notamment de l'étage de sortie. Cela signifie que la sortie du pré-amplificateur n'est pas faite pour gérer à la fois une charge importante et assurer une bande passante élevée ($\geq 200MHz$).

TABLE 2.4 – Résistance de sortie, capacité de sortie et bande passante des différents blocs

	S_1	F_1	S_2	F_2	S_3	F_3 + 50pF
R_{out} (Ω)	490	48.9	403	48.5	401	48.7
C_{out} (F)	640f	2.48p	409f	2.48p	413f	51.4p
BW_{out} (Hz)	508M	1.31G	966M	1.32G	961M	63.6M

La modification seule du dernier suiveur de tension $F3$ de manière à obtenir une bande passante de $200MHz$ entraînerait des courants et des dimensions de transistors beaucoup trop élevés pour être considérés comme réalistes. C'est pourquoi un suiveur supplémentaire, utilisé comme buffer de

sortie NMOS a été ajouté en fin de chaîne, à la suite du dernier suiveur de tension.

Celui-ci a été conçu pour fixer la tension de mode commun à $\frac{Vdd}{2}$ mais également pour avoir une résistance de sortie de l'ordre de 10Ω, de façon à augmenter la bande passante. Ce buffer de sortie double la consommation du pré-amplificateur mais il permet d'accroître de manière significative la bande passante, et ce, même avec une charge de 50pF en sortie, puisque l'on peut espérer :

$$BW_{out} = \frac{1}{2 * \pi * R_{out} * C_{load}} = 318MHz \tag{2.10}$$

La consommation n'étant pas une contrainte principale, contrairement à la bande passante, c'est un compromis que nous faisons volontiers. On prend également soin de gérer le pôle d'entrée qui peut devenir un facteur limitant. En effet, le Tableau 2.1 et le Tableau 2.4 soulignent que le deuxième point faible dans l'architecture du pré-amplificateur est le premier étage et sa capacité d'entrée de $5.38pF$. Les dimensions des transistors le constituant ont donc été légèrement réduites pour diminuer cette capacité d'entrée. On déplace ainsi le pôle d'entrée à plus haute fréquence, ce qui a pour conséquence d'augmenter encore un peu plus la bande passante.

Réalisation d'une implémentation BiCMOS

L'implémentation BiCMOS de ce pré-amplificateur suit exactement la même architecture que l'implémentation CMOS, c'est-à-dire plusieurs étages avec des suiveurs de tension entre chaque. Cependant ici, seulement deux étages sont nécessaires pour obtenir un gain suffisant. Dans cette version, les transistors MOS des paires différentielles et des suiveurs de tension sont remplacés par des transistors NPN, comme indiqué sur la Figure 2.13. La génération de la tension de polarisation a également été légèrement modifiée, puisqu'elle ne nécessite désormais plus qu'un transistor NMOS monté en diode.

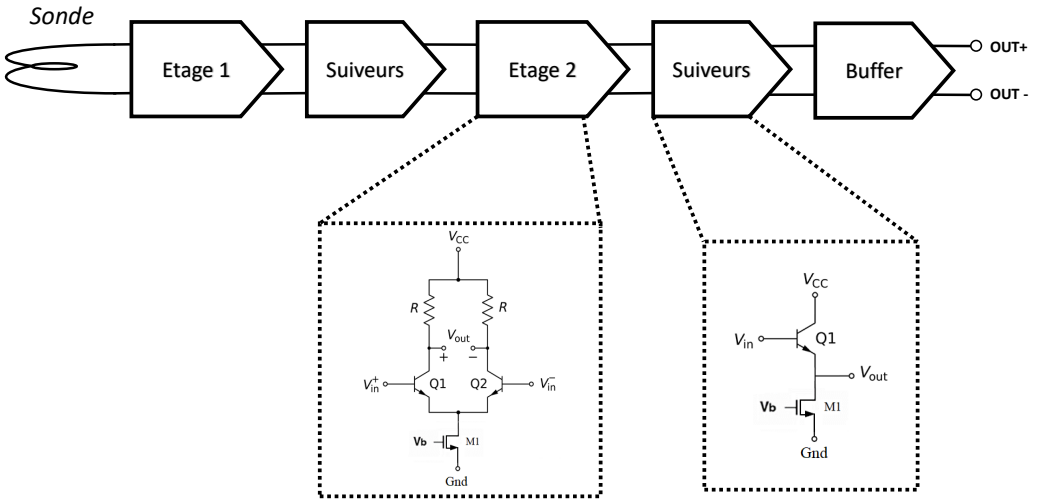


FIGURE 2.13 – Schéma bloc et schémas électroniques du pré-amplificateur BiCMOS conçu

Résultats de simulations

Pour obtenir des résultats de simulation représentatifs des performances finales des pré-amplificateurs conçus, une charge de $50pF$ a été considérée pendant toutes les simulations. Celle-ci prend en compte les pads de la puce, ainsi que les câbles coaxiaux (SMA ou BNC) qui servent à connecter le pré-amplificateur à un oscilloscope digital. Pendant ces simulations, le plancher de bruit a été mesuré à $50MHz$ car la plupart des implémentations matérielles des algorithmes de chiffrement opèrent à une fréquence d'horloge de cet ordre de grandeur. Le prochain paragraphe donne successivement les performances en fonctionnement typique pour le design CMOS et le design BiCMOS.

Le Tableau 2.5 donne les performances du pré-amplificateur CMOS en conditions typiques. Comme indiqué, la bande passante est égale à $300MHz$ (avec une charge de 50pF), ce qui respecte les spécifications. Le plancher de bruit est 10% supérieur à la valeur attendue tandis que la consommation, quant à elle, est 30% inférieure.

TABLE 2.5 – Résultats de simulation en fonctionnement typique pour le design CMOS du run 2

Gain	Bande passante	Plancher de bruit	Consommation
45.88 dB	300.5 MHz	1.11 nV/$\sqrt{Hz}$	70.32 mA

Le Tableau 2.6 donne les mêmes résultats que pour le design CMOS mais pour le design BiCMOS. On peut observer que l'on obtient un gain similaire mais pour une bande passante 1.6 fois plus large. En plus, cette bande passante plus large est obtenue avec une consommation 1.4 fois plus faible, pour un plancher de bruit similaire. Ainsi, on pourrait s'attendre à obtenir de meilleurs résultats avec le design BiCMOS lors des futures expérimentations.

TABLE 2.6 – Résultats de simulation en fonctionnement typique pour le design BiCMOS du run 2

Gain	Bande passante	Plancher de bruit	Consommation
47.42 dB	532.9 MHz	1.27 nV/$\sqrt{Hz}$	54.16 mA

Comme pour le run précédent, des simulations de Monte-Carlo ont ensuite été menées, cette fois-ci sur 1000 circuits. Les résultats pour le design CMOS amélioré sont donnés dans le Tableau 2.7. On retrouve comme pour le design précédent la grande variabilité du gain (caractéristique des circuits en boucle ouverte). On observe également que la bande passante ne descend jamais en dessous de 220MHz, et ce malgré la charge de 50pF en sortie, ce qui est une très bonne nouvelle et confirme une nouvelle fois l'efficacité du buffer de sortie ajouté dans le design. De plus, la totalité des circuits ont passé la simulation avec succès.

TABLE 2.7 – Simulation de Monte-Carlo sur 1000 circuits pour le design CMOS du run 2

	Min	Max	Moyenne
Bande passante (-3dB)	221.4 MHz	394.9 MHz	301.2 MHz
Gain	96.28 (39.67dB)	335 (50.5dB)	196.4 (45.86dB)
Bruit (@100kHz)	0.989 nV/$\sqrt{Hz}$	1.244 nV/$\sqrt{Hz}$	1.108 nV/$\sqrt{Hz}$
Offset d'entrée	-3 mV	2.2 mV	-0.004 mV
Consommation	52.54 mA	97.59 mA	71.44 mA

Pour respecter les contraintes de délai pour l'envoi du design, nous avons dû passer à la réalisation des layouts des deux circuits sans réaliser les simulations de Monte-Carlo pour le design BiCMOS. Celles-ci, réalisées à la suite de l'envoi, ont révélé une erreur de conception concernant les sources de courant des paires différentielles. En effet, celles-ci ne sont pas correctement

polarisées, ce qui entraîne 40% d'échec sur les 1000 cicruits simulés. Cette erreur a bien entendu été corrigée à la suite de sa découverte.

Layouts des deux implémentations

La stratégie adoptée pour réaliser les layouts a été la même que précédemment, à savoir d'effectuer le layout de chaque bloc indépendamment, avant de venir les connecter en cascade. Les deux layouts sont donnés sur la Figure 2.14.

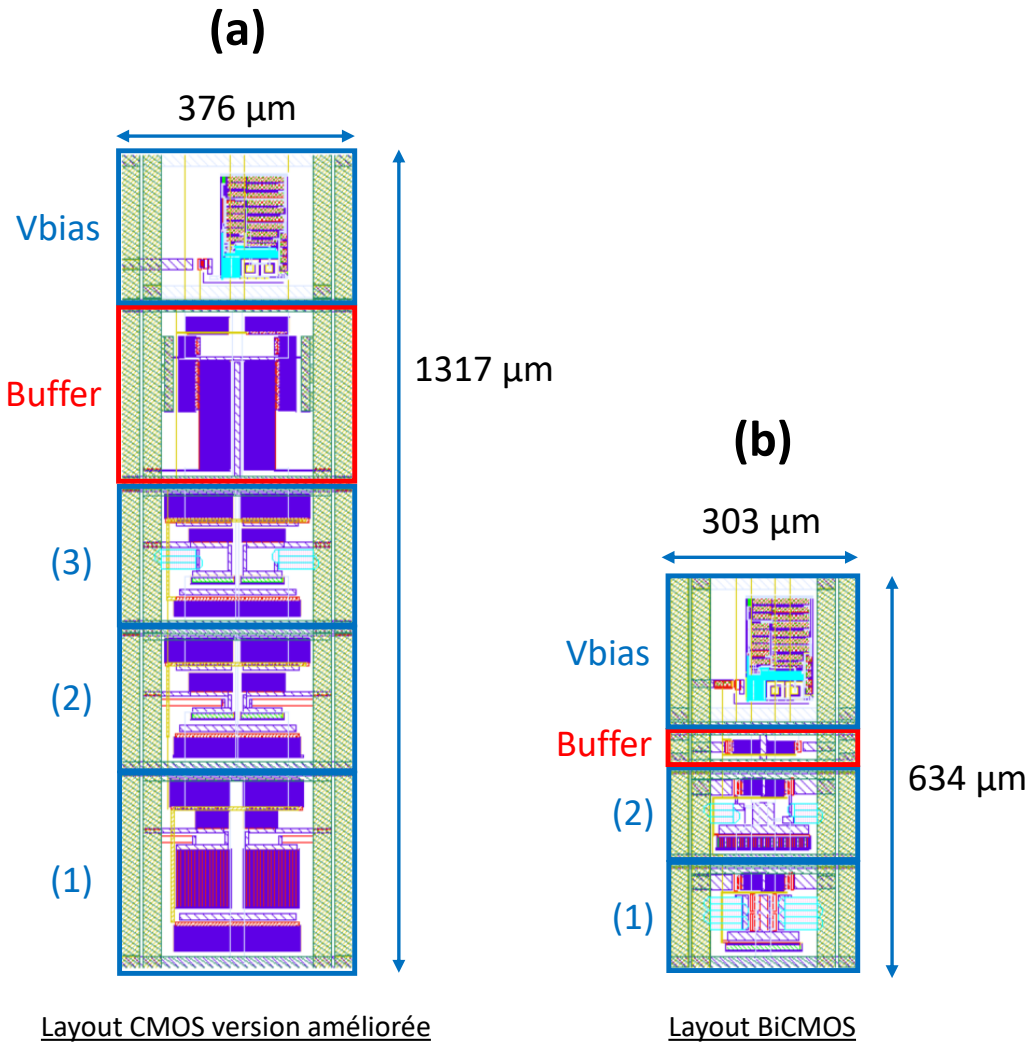


FIGURE 2.14 – Layouts des pré-amplificateurs CMOS et BiCMOS du run 2

On peut observer deux changements. Le premier est l'ajout du buffer de sortie qui résulte en un bloc supplémentaire (encadré en rouge sur la Figure). Le deuxième sont les rails d'alimentation qui sont un peu plus larges que précédemment, puisque nous avons remarqué quelques instabilités sur la tension d'alimentation lors des expérimentations sur le premier circuit. Il n'y a rien d'autre à noter si ce n'est que le design BiCMOS occupe environ deux fois moins de place que le design CMOS. Il peut alors être intéressant de choisir cette technologie si jamais la surface allouée sur le wafer est une contrainte importante (ce qui n'est pas notre cas ici).

Les designs ont été envoyés en fabrication en Octobre 2019, juste avant le début de la crise sanitaire. Comme pour le premier run, cela a résulté en plusieurs mois de retard avec une réception en Juin 2020.

Test en boîtier et validation fonctionnelle du circuit CMOS

Dans un premier temps, un test en boîtier a été réalisé dans les mêmes conditions expérimentales que précédemment. Celui-ci a permis de vérifier que le gain (47dB) et la consommation (80mA) étaient corrects, mais surtout que la bande passante était légèrement supérieure à 200MHz. On obtient une bande passante multipliée par un facteur 10 par rapport au premier run, ceci grâce à l'utilisation des bons plots, à l'ajout du buffer de sortie (permettant de réduire la résistance de sortie à 10Ω) et à la diminution de la capacité d'entrée du pré-amplificateur. Nous n'atteignons toujours pas les 300MHz attendus car il semblerait que la capacité de charge considérée (50pF) soit une nouvelle fois en dessous de la réalité.

A la suite de ce premier test, une validation fonctionnelle a été effectuée avec les sondes flexibles présentées dans la partie précédente. Elle a consisté à mesurer le champ magnétique émis par une longue boucle étroite (plusieurs centimètres de long et $1mm$ de large), gravée sur un PCB avec une piste de $500\mu m$, à l'aide d'un prototype de sonde active comprenant la puce en boîtier. Le prototype en question est visible Figure 2.15.

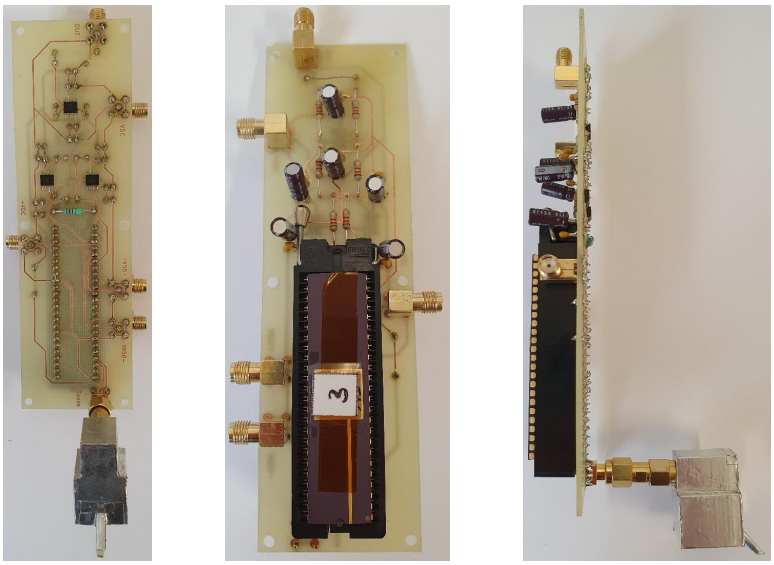


FIGURE 2.15 – Prototype de sonde active

La résistance totale de la boucle a été fixée à 50Ω à l'aide d'une résistance externe et un signal sinusoïdal de $10dBm$ à $60MHz$ lui a été délivré. La Figure 2.16 montre l'amplitude des signaux mesurée à la sortie du pré-amplificateur, lorsque les sondes de $500\mu m$ et $150\mu m$ ont été déplacées perpendiculairement au plus long côté de la boucle. Comme attendu, on observe des distributions normales. Par soucis de temps, la fréquence du signal mesuré n'a pas été modifiée. A ce stade, nous pouvons juste affirmer que la bande passante est d'au moins 60MHz, en attendant de futures expérimentations.

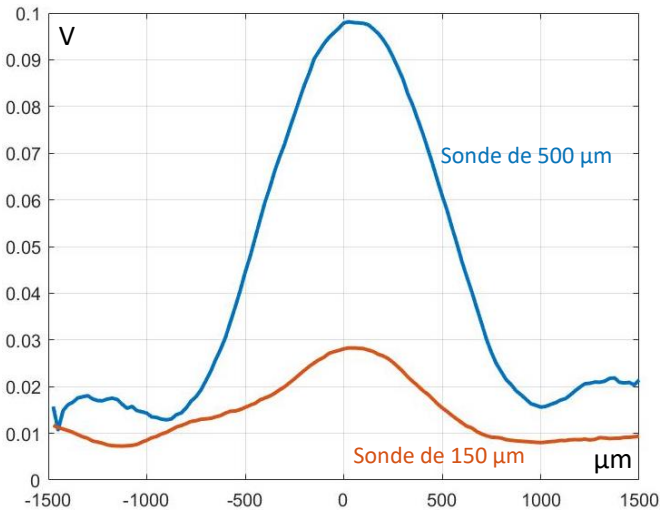


FIGURE 2.16 – Amplitude du signal de $60MHz$ mesurée à la sortie du pré-amplificateur avec les sondes de $500\mu m$ et $150\mu m$

2.4.4 Conclusion

De nombreux prototypes ont été fabriqués de manière à tester les performances du pré-amplificateur faible bruit conçu. L'évolution des prototypes est illustrée sur la Figure 2.17. A ce stade de l'étude, on peut conclure un certain nombre d'éléments.

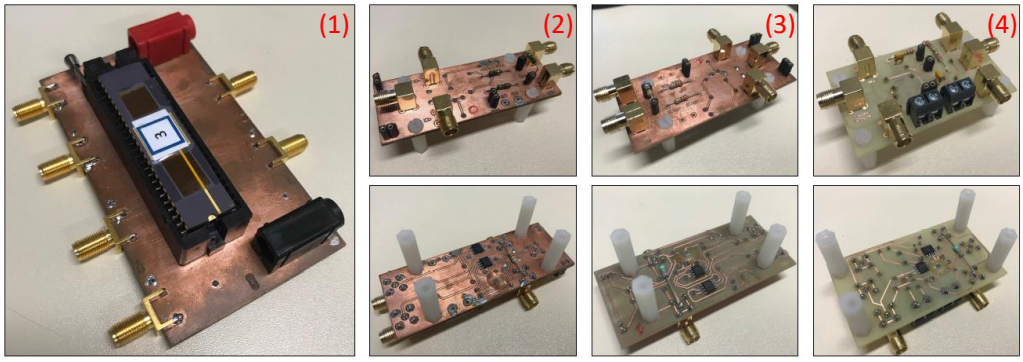


FIGURE 2.17 – Évolution des prototypes de caractérisation des pré-amplificateurs

Le circuit en technologie CMOS du premier run, bien que fonctionnel, ne respecte pas les spécifications en termes de bande passante. Plusieurs facteurs limitants ont été mis en évidence, parmi eux :

— L'utilisation des plots 50Ω qui amène une limitation en fréquence à la fois en entrée et en sortie du pré-amplificateur,

— L'impédance de sortie du pré-amplificateur (50Ω) trop importante,

— La trop faible capacité de charge considérée dans les simulations (10pF alors que la réalité est plus proche de 50-75pF),

— La capacité d'entrée du premier étage un peu trop élevée, qui accroît la limitation en entrée.

Nous avons alors décidé d'effectuer un deuxième run comprenant deux circuits, le premier utilisant la même technologie et le second en technologie BiCMOS, tout en prenant en compte les limitations du premier run, citées ci-dessus.

Le circuit BiCMOS n'a pas été caractérisé pour deux raisons. Tout d'abord, l'erreur de conception des sources de courant a amené au dysfonctionnement de plus de la moitié des circuits reçus. De plus, pour les circuits polarisés correctement, des oscillations ont été constatées. Celles-ci pourraient partiellement s'expliquer par de potentiels rebouclages internes (non voulus) lors de la réalisation du Layout.

Pour finir sur des résultats très encourageants, la deuxième version du circuit CMOS répond en tout point aux spécifications fixées en début de chapitre. De plus, il a été démontré qu'il est possible de réaliser un prototype de sonde active avec les éléments dont nous disposons (pré-amplificateur CMOS2, sonde flexible, amplificateur d'instrumentation, etc.), pour mesurer dans un premier temps le signal EM émis par une boucle réalisée sur un PCB. Ceci est très encourageant et ouvre de nouvelles perspectives.

La prochaine étape consistera donc à réaliser le même prototype, cette fois-ci avec la puce nue, tout en optimisant la taille du PCB, pour obtenir un prototype final fonctionnel et compact pour la réalisation des expérimentations futures. Il restera également à prévoir un troisième run, de manière à pouvoir tester le design BiCMOS corrigé, celui-ci étant très prometteur au vu des simulations effectuées jusqu'à présent. Enfin, une dernière étape consistera à comparer notre prototype de sonde active avec par exemple une sonde Langer passive connectée à l'amplificateur Femto, à travers la réalisation de plusieurs cartographies du champ proche électromagnétique d'un circuit.

Chapitre 3

Sondes d'injection

3.1 Résumé

Ce chapitre présente l'intégralité des travaux réalisés quant à l'amélioration des plateformes d'injection de fautes par médium électromagnétique. Ces travaux ont porté essentiellement sur l'amélioration des résolutions spatiale et temporelle avec comme contrainte de conserver le caractère pratique et modulaire de ce type de plateforme. Ce chapitre décrit également un protocole de caractérisation des sondes d'injection. Il est à noter que l'essentiel des travaux reportés dans ce chapitre a fait l'objet de deux publications [Tou+20], [Tou+21b].

3.2 Amélioration de la résolution spatiale

3.2.1 Contexte

L'injection de fautes par médium électromagnétique est une technique relativement récente. En effet, les premiers résultats concrets n'ont été publiés qu'en 2007 dans [SH07]. Depuis, différents types de plateformes ayant des caractéristiques et performances spécifiques, qu'elles soient académiques [BAM17b], [Deh+12a], [Ord+15] ou commerciales [AH20b], [CH17] (Langer, Riscure ou NewAE) ont démontré leur efficacité pour conduire des attaques. Bien qu'ayant des caractéristiques différentes, ces plateformes sont bâties autour de deux éléments incontournables.

Le premier est un générateur d'impulsions de tension. L'amplitude maximale, V_p^{max}, ainsi que les temps de montée et de descente des impulsions qu'il fournit sont généralement considérés et mis en avant comme les figures de mérite de la puissance d'une plateforme.

Le deuxième élément incontournable d'une plateforme EMFI est la sonde d'injection. Celle-ci est en général une simple bobine, réalisée en enroulant un fil autour d'un noyau de ferrite, dont le diamètre d varie entre $300\mu m$ et $2mm$. Ce diamètre est généralement considéré comme étant un bon indicateur de la résolution spatiale de la plateforme d'injection. Cette idée est basée sur le fait que l'amplitude du champ magnétique est maximale sous le noyau de ferrite, et suit une distribution uniforme ou normale selon la distance qui sépare la sonde du point d'observation [Oma+13].

C'est pour cette raison que la résolution spatiale des plateformes EMFI est considérée comme limitée par rapport à celle des plateformes laser, qui est de l'ordre du micromètre. Selon [Aar13], c'est le plus gros désavantage de l'injection de fautes EM comparé à l'injection de fautes laser.

3.2.2 Objectif

L'idée selon laquelle la résolution spatiale des EMFI est limitée est en contradiction avec les nombreuses attaques menées à bien sur des microcontrôleurs ou des SoC (Systems On Chip) fabriqués dans des technologies avancées, comme reporté dans [O'F19], [Tro+19], [O'F20], [Buk+18]. Elle est aussi en conflit avec la pratique de l'EMFI qui montre que la position de la sonde (même pour les plus grosses d'entre elles) doit être contrôlée très précisément pour obtenir des fautes : modifier la position de la sonde ne serait-ce que d'une dizaine de μm peut faire la différence entre obtenir une faute exploitable, un crash ou une réponse correcte.

Cette contradiction est liée à deux raisonnements certes intuitifs mais erronés, couramment adoptés faute d'une compréhension fine de l'effet des injections EM sur les circuits intégrés. Le premier de ces raisonnements consiste à considérer que la résolution spatiale d'une plateforme EMFI est déterminée par la taille de la sonde. Il n'en est rien. En effet, comme le démontre [DLM21], l'injection de fautes par médium EM s'appuie sur le principe de l'induction électromagnétique qui met en jeu non pas une antenne mais à minima deux. La résolution spatiale d'une plateforme d'injection EM n'existe donc pas (du moins au sens couramment admis pour une plateforme laser) et ne peut être définie que dans le cadre d'un couple plateforme/circuit. Les antennes réceptrices au sein des circuits intégrés étant de dimensions inférieures à celles des sondes, ce sont ces dernières qui fixent probablement la

résolution spatiale des injections.

Le second raisonnement intuitif conduisant à cette contradiction consiste à considérer qu'une EMFI produit des effets / fautes sur l'ensemble de la surface du circuit couverte par la sonde. Or il n'en est rien comme indiqué dans [DLM21] qui présente un modèle électrique des effets de l'EMFI. En effet, les auteurs montrent par la simulation que l'EMFI produit de puissantes variations locales de la tension d'alimentation interne du circuit, sous les bords de la sonde, et d'autres moins importantes sous le noyau de ferrite et aux alentours. Cela s'explique par la géométrie régulière des réseaux d'alimentation (formant des grilles) qui reçoivent l'impulsion EM. C'est pourquoi l'EMFI est censée selon [DLM21] générer des fautes sous les bords de la sonde et donc dans une zone proportionnelle aux dimensions latérales (diamètre) de la sonde et non pas à sa surface.

Compte tenu de ce constat et de notre objectif consistant à améliorer la résolution spatiale des plateformes d'injection, il nous est apparu important de s'intéresser à l'effet de la réduction des dimensions des sondes quant à l'efficacité des EMFI. Plus particulièrement, les questions auxquelles nous avons essayé d'apporter des éléments de réponse ont été les suivantes.

Étant donnée une sonde de dimension caractéristique d à laquelle on doit appliquer une impulsion d'amplitude V_P pour produire des fautes au sein d'un circuit donné, comment doit-on faire évoluer V_P de sorte à continuer à induire des fautes, et ce tout en réduisant d ? Linéairement avec d (donc linéairement avec les dimensions de la sonde) ou d^2 (donc linéairement avec sa surface), ou pire ?

Ces questions somme toute simples sont importantes. En effet, l'obtention d'une réponse indiquera s'il est envisageable de développer aisément ou pas des plateformes EMFI avec des résolutions de l'ordre de quelques dizaines de micromètres, et ce compte tenu du coût et de la difficulté à développer des générateurs d'impulsions rapides délivrant des tensions au delà de 2000V.

3.2.3 Éléments de réponses théoriques

Dans un premier temps, afin d'apporter des éléments de réponses théoriques, nous avons considéré les principes de l'induction électromagnétique

et le modèle décrit dans [DLM21]. L'objectif étant d'extraire de ce modèle un lien théorique entre V_P et d.

Selon [DLM21], l'injection de fautes par médium EM exploite le couplage entre une sonde EM et les réseaux d'alimentation et de masse des circuits intégrés de manière à induire une perturbation locale et transitoire de la tension d'alimentation, affectant de manière significative quelques centaines de portes logiques. D'après les résultats reportés dans cette publication, cette perturbation transitoire peut être considérée comme la réponse d'un circuit RC à une impulsion de tension.

Puisque les réseaux d'alimentation et de masse sont routés de manière à former des grilles régulières, ils forment un très grand nombre de boucles, et le couplage électromagnétique s'effectue entre la sonde et chacune de ces boucles. Ce dernier, comme illustré sur la Figure 3.1 dans le cas d'une seule boucle du réseau d'alimentation, peut être modélisé et simulé à l'aide d'une inductance mutuelle dotée d'un primaire et d'autant de secondaires que de boucles formées par les réseaux d'alimentation et de masse. L'inductance mutuelle entre la sonde et une des boucles a pour expression :

$$M = k \cdot \sqrt{L_P \cdot L_G} \tag{3.1}$$

Avec L_P l'inductance propre de la sonde, L_G celle d'une boucle (de hauteur h et de largeur w) du réseau d'alimentation ou de masse et k le coefficient de couplage EM.

Appliquer une impulsion d'amplitude V_P et de largeur à mi-hauteur PW à la sonde induit une différence de potentiel temporaire V_{ind}, le long de chaque boucle de section rectangulaire formée par le réseau d'alimentation. Son expression est :

$$V_{ind} = M \cdot \frac{1}{R_P} \cdot \frac{\Delta V_P}{\Delta t} = M \cdot \frac{1}{R_P} \cdot \frac{V_P}{\Delta t} \tag{3.2}$$

Avec R_P la résistance de la sonde. Cette différence de potentiel V_{ind} est répartie le long des quatre branches de chaque boucle, proportionnellement

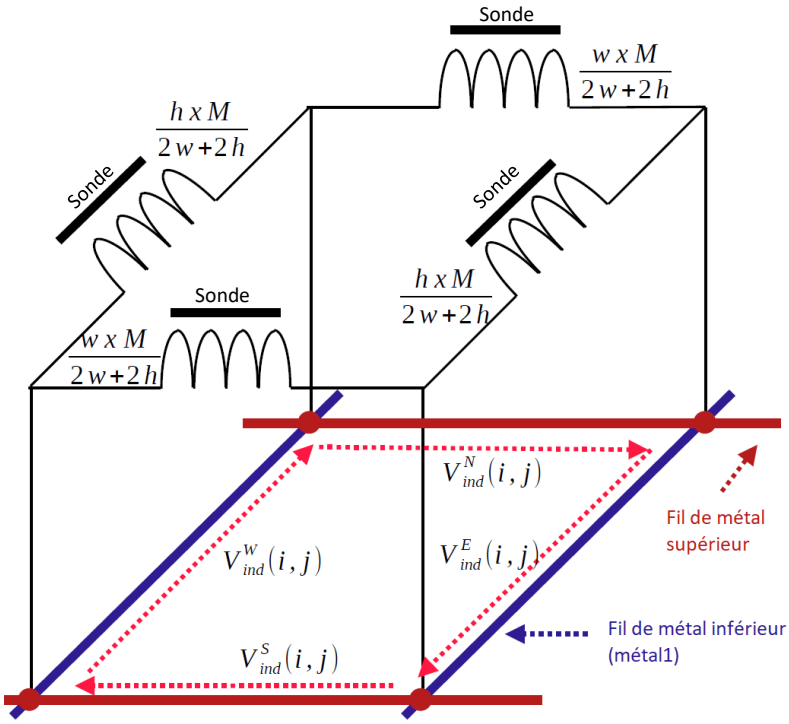


FIGURE 3.1 – Couplage EM entre la sonde et une boucle du réseau d'alimentation

à leur longueur (en supposant que la résistivité reste la même le long de la boucle) et orientée en accord avec la loi de Lentz (W, N, E, S signifiant West, North, East, South) :

$$V_{ind} = V^W_{ind} + V^N_{ind} + V^E_{ind} + V^S_{ind} \tag{3.3}$$

Avec :

$$V^W_{ind} = V^E_{ind} = \frac{h}{2h+2w} \cdot V_{ind} \qquad V^N_{ind} = V^S_{ind} = \frac{w}{2h+2w} \cdot V_{ind} \tag{3.4}$$

Lorsque les quatre branches d'une boucle sont soumises au même flux EM, alors les différences de potentiel V^W_{ind} et V^E_{ind} de cette boucle sont égales et de signe opposé. Il en est de même pour les branches V^N_{ind} et V^S_{ind}. Or, pour une boucle quelconque du réseau d'alimentation et de masse, la branche Nord correspond à la branche Sud de la boucle du dessus, et la branche Sud correspond à la branche Nord de la boucle du dessous (de même pour Est et Ouest).

De plus, les orientations des flux magnétiques sous et autour de la sonde ont des directions opposées. De fait, d'après le principe de superposition, la différence de potentiel induite s'annule partiellement ou complètement sous et autour de la sonde, mais s'ajoute à la verticale des bords de celle-ci. Par conséquent et d'après [Oma+13], si la sonde est placée au contact ou très proche de la surface du circuit, l'amplitude du flux magnétique est quasi constante sous le noyau de la sonde, et quasi nulle à l'extérieur. Dans ce cas, l'EMFI induit une variation transitoire de la tension d'alimentation seulement sous les bords de la sonde, comme illustré Figure 3.2.

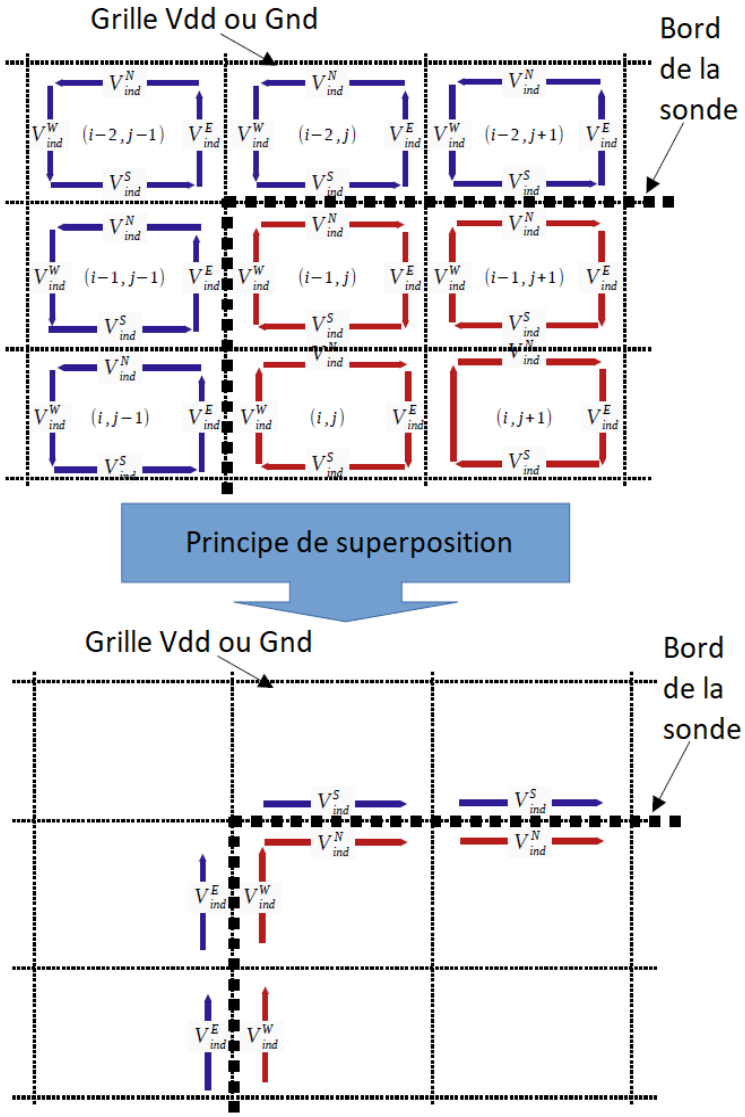


FIGURE 3.2 – Application du principe de superposition aux différences de potentiels induites, provoquées par une impulsion générant un flux d'amplitude uniforme sur la surface du circuit

Dans le cas où la sonde est placée un peu plus loin du circuit, bien que le flux magnétique suive une distribution bi-normale, cela conduit également à des effets plus importants sous les bords de la sonde, comme décrit dans [DLM21].

Si l'on tient compte de tout ce qui a été présenté jusqu'ici, l'injection de fautes EM peut être considérée comme un moyen pour injecter des courants de Foucault dans les circuits intégrés. Cependant, ces courants ne pouvant pas se déplacer librement comme dans des matériaux homogènes, ils sont forcés de suivre les grilles que forment les réseaux d'alimentation et de masse.

Compte tenu de cette modélisation des effets de l'EMFI, une solution qui apparaît comme évidente pour améliorer la résolution spatiale de l'EMFI est de réduire les dimensions des sondes. Cependant, cela signifie également réduire leurs inductances propres et donc le couplage électromagnétique avec le circuit ciblé ; réduction qu'il convient de compenser en utilisant un générateur d'impulsion plus puissant. Toutefois, à ce stade, une question demeure : Comment choisir le générateur d'impulsion le plus adapté ? Plus précisément, comment modifier, de sorte à obtenir les mêmes effets, l'amplitude de l'impulsion lorsque les dimensions des sondes varient pendant une campagne d'acquisition ?

Pour répondre à ces questions, considérons donc les équations 3.1 et 3.2, et supposons qu'un attaquant veuille induire les mêmes fautes dans le circuit ciblé, mais avec deux sondes de dimensions différentes. Il faut alors qu'il ajuste les amplitudes des impulsions délivrées aux sondes de manière à induire la même perturbation dans le circuit. En considérant l'équation 3.2, il vient :

$$\frac{M_1}{M_2} = \frac{\Delta V_{P2}}{\Delta V_{P1}} = \frac{V_{P2}}{V_{P1}} \tag{3.5}$$

En supposant, par soucis de simplification, que les sondes sont de sections carrées, il est alors possible d'estimer l'inductance propre de celles-ci d'après [Tho91], en utilisant :

$$L_p = \frac{2 \cdot N^2 \cdot \mu_0 \cdot W}{\pi} \left[log(\frac{W}{R}) - 0.524 \right] \tag{3.6}$$

Avec W la largeur de la sonde, R le rayon du fil, N le nombre de spires et μ_0 la perméabilité magnétique du vide. En injectant alors cette expression dans l'équation 3.5, on obtient :

$$\frac{V_{P2}}{V_{P1}} = \sqrt{\frac{W_1 \cdot [log(\frac{W_1}{R}) - 0.524]}{W_2 \cdot [log(\frac{W_2}{R}) - 0.524]}} \tag{3.7}$$

Qui, si l'on considère que le rayon R du fil est réduit de manière conjointe aux dimensions de la sonde, peut ensuite être approximée par :

$$\frac{V_{P2}}{V_{P1}} \simeq \sqrt{\frac{W_1}{W_2}} \tag{3.8}$$

Ce résultat est intéressant. En effet, il stipule que la puissance des générateurs d'impulsion des plateformes d'EMFI doit évoluer proportionnellement à la racine carrée des dimensions des sondes. Par exemple, réduire les dimensions d'une sonde d'un facteur quatre ne devrait nécessiter qu'une amplitude deux fois plus élevée. A ce stade, il ne reste plus qu'à confirmer ce résultat théorique et la justesse des hypothèses considérées à l'aide d'expérimentations.

Pour vérifier la validité de l'équation 3.8 ci-dessus, plusieurs expérimentations ont été menées et seront présentées dans la suite de ce document. Cependant, il convient dans un premier temps de présenter les sondes qui ont été conçues ainsi que le processus de fabrication de celles-ci.

3.3 Design et fabrication des sondes flexibles

Comme démontré dans la partie précédente, il y a un réel intérêt à réduire les dimensions des sondes d'injection pour améliorer la résolution spatiale,

sans pour autant devoir délivrer à la sonde des impulsions de plusieurs milliers de volts d'amplitude. Pour ce faire, et comme pour les sondes d'analyse décrites dans le Chapitre 2, les pistes de l'électronique flexible et de l'impression 3D ont été explorées.

La sonde se décompose en trois parties distinctes. Le PCB (Printed Circuit Board) flexible (sur lequel se trouve la boucle de la sonde), le support imprimé en 3D (qui sert de support au PCB flexible) et le PCB rigide (sur lequel se trouve les éléments nécessaires à la connexion au générateur d'impulsion). Ces trois parties ont été conçues et réalisées indépendamment et sont simplement assemblées pour réaliser le prototype de sonde final.

3.3.1 PCB flexible

La technologie utilisée possède également quatre couches de métaux et permet de dessiner des pistes de $50\mu m$ espacées de $50\mu m$. Mais contrairement aux sondes d'analyse, les sondes d'injection possèdent des vias enterrés (pour relier la couche supérieure à la couche inférieure adjacente) entre les quatre couches (et non plus seulement entre les couches 2 et 3). Par conséquent, l'épaisseur totale du circuit est légèrement supérieure et le prix s'en trouve multiplié par 2.5, du fait de sa complexité plus importante. Plusieurs designs de sondes ont donc été réalisés, toujours dans l'idée d'obtenir les meilleures performances possibles. Les deux designs présentés sur la Figure 3.3 sont deux exemples de sondes qui n'ont pas été retenus.

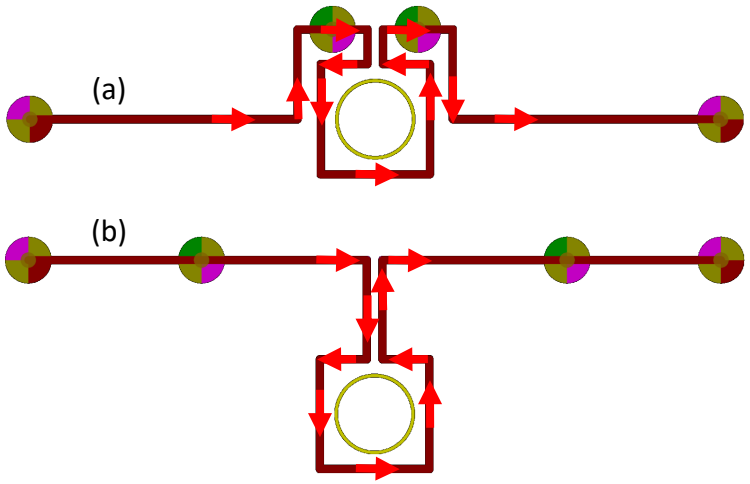


FIGURE 3.3 – Designs de sondes non retenus, avec sens de circulation du courant

En effet, on peut remarquer que le design (a) possède de nombreuses boucles parasites, et la circulation des courants dans celles-ci est opposée à celle dans la boucle principale, ce qui diminue grandement les performances finales de la sonde. Sur le design (b), ces boucles parasites sont moins importantes mais les performances finales ne sont pas non plus à la hauteur des espérances. Les autres designs proposés sont eux plus prometteurs et sont relativement simples. Pour ceux-ci, une solution permettant de se débarrasser entièrement des boucles parasites a été choisie. Un aperçu des designs, couche par couche, est donné Figure 3.4.

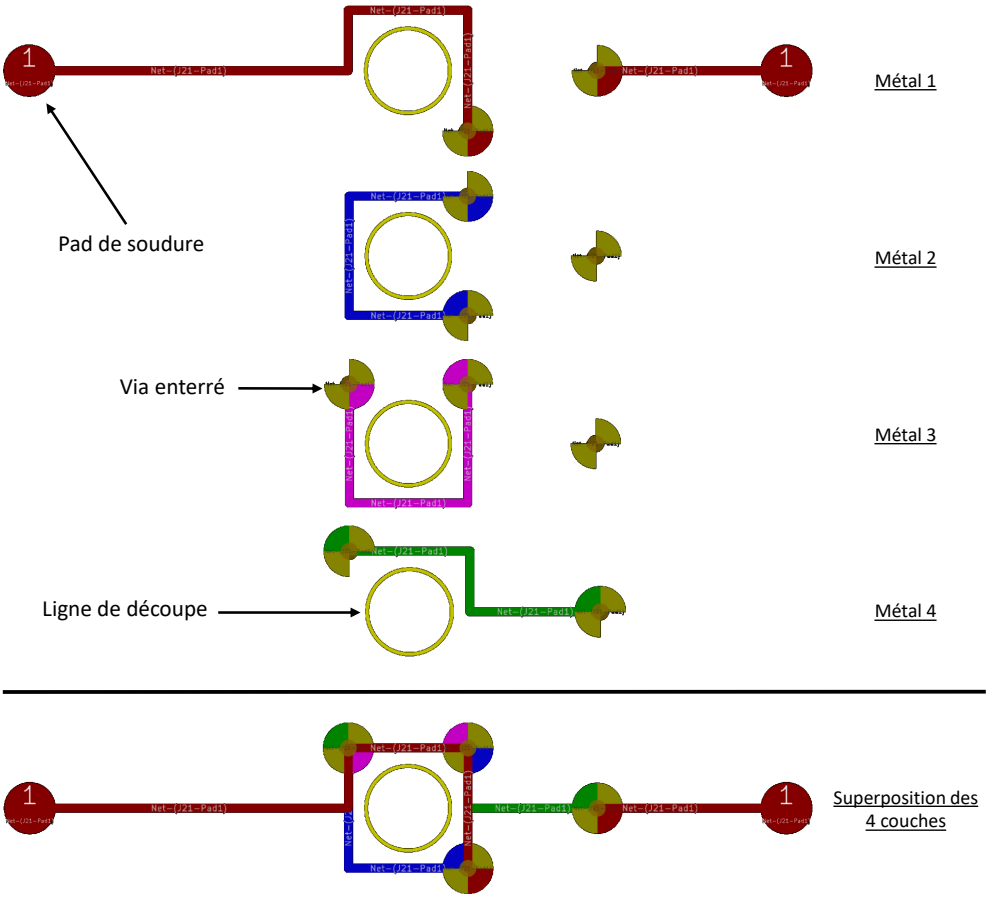


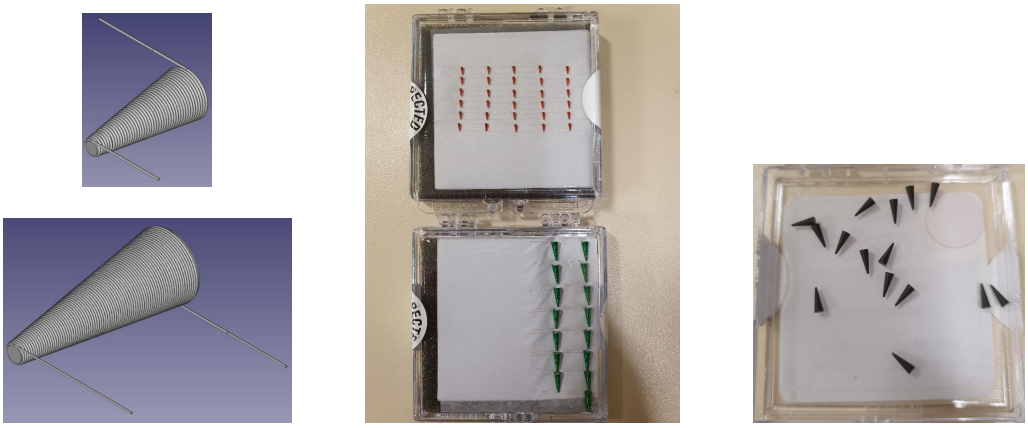
FIGURE 3.4 – Design du PCB flexible des sondes, de la couche supérieure (en haut) à la couche inférieure (en bas)

Des boucles sont réalisées sur les différents niveaux de métaux à l'aide de vias enterrés pour passer d'un niveau à l'autre. Même si le nombre optimal de boucles pour l'injection se situe entre 4 et 5, l'encombrement desdits vias (perçage de 100μm et pad de 300μm) ne permet que de réaliser 2.5 boucles par sonde au maximum si l'on considère un circuit sur quatre couches. Par

soucis de comparaison, et de manière à avoir une caractérisation complète, des sondes de taille similaire mais ne possédant que 1.5 boucles ont également été dessinées.

On se retrouve donc avec quatre tailles de sondes différentes : $650\mu m$, $520\mu m$, $400\mu m$ et $300\mu m$, et pour chacune d'entre elle soit 2.5 boucles, soit 1.5 boucles.

Pour les boucles de taille $650\mu m$ et $400\mu m$, il a été proposé des designs percés. Ce trou sert à ajouter (ou non) un cône de ferrite au centre de la boucle, de manière à améliorer significativement l'efficacité des injections (des chiffres concrets seront donnés dans la partie caractérisation des sondes). Pour obtenir ces cônes de ferrite de taille relativement petite, des inductances possédant un noyau de ferrite ont été commandées chez le fournisseur Piconics. Le fil de métal s'enroulant autour de ces noyaux de ferrite a ensuite été retiré pour ne garder que le cône. La Figure 3.5 présente le modèle 3D de ces bobines ainsi que deux photos, la première après réception des bobines et la deuxième après récupération des cônes.



Modèles 3D des inductances Piconics | Inductances après réception | Récupération des cônes de ferrite

FIGURE 3.5 – Cônes de ferrite

Pour les boucles de taille $520\mu m$ et $300\mu m$, les designs proposés sont non percés. La raison est simple. La distance minimale à respecter entre une piste de métal et une ligne de découpe est de $130\mu m$. Ainsi, si le fabricant perce lui-même les boucles, la taille des sondes s'en trouve dès lors augmentée. De façon à pouvoir atteindre des tailles de sondes encore inférieures, il a donc

été décidé de proposer des designs non percés, qui seront percés après réception des circuits, à l'aide d'une pointe de test chauffée par exemple.

Une fois la boucle des sondes dessinée et les différents paramètres choisis (taille, perçage, etc.), il a fallu réfléchir à la connexion au reste du circuit. Deux pads de 300μm de diamètre ont donc été ajoutés de part et d'autre de la boucle, de manière à pouvoir y souder deux fils de 150μm de diamètre. Tous les designs finaux sont présentés sur la Figure 3.6.

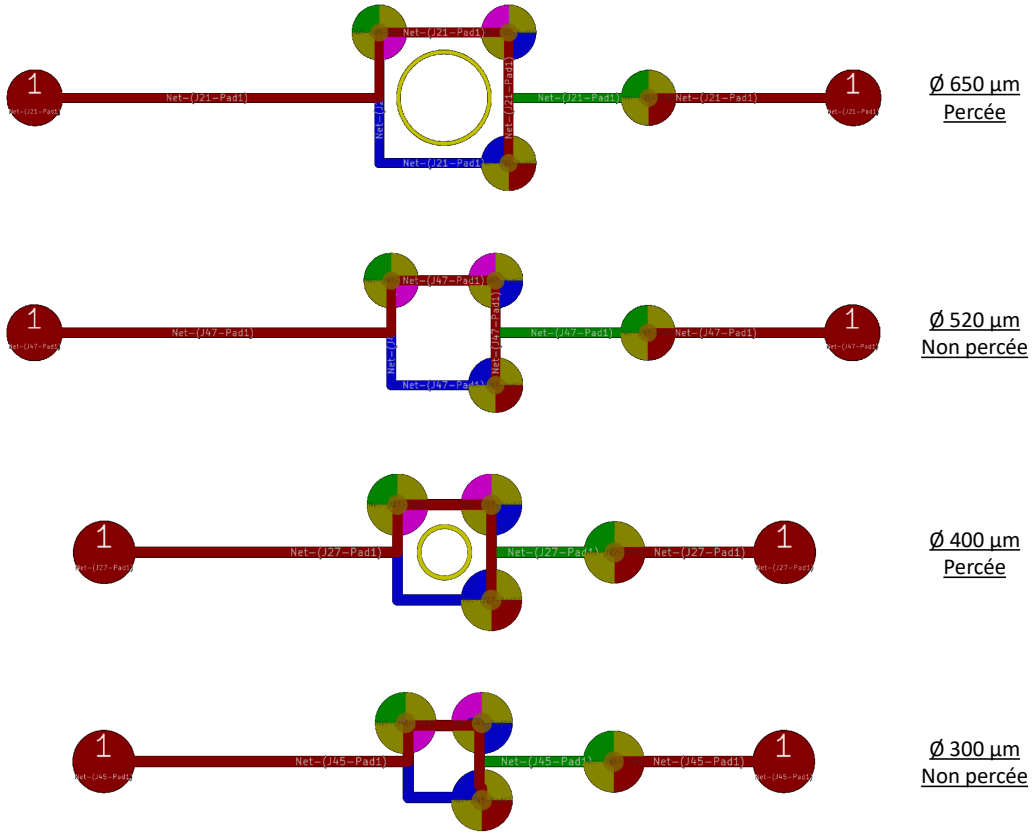


FIGURE 3.6 – Designs finaux des sondes d'injection

De manière arbitraire et par soucis de simplicité pour la future soudure des fils, l'envergure de la sonde a été fixée à 4.5mm pour les sondes de 650μm et 520μm, et à 3.5mm pour les sondes de 400μm et 300μm. Cette distance a été réduite à 2.5mm pour les quatre sondes dans le dernier run, de manière à diminuer de manière significative l'encombrement et donc de permettre une meilleure pénétration dans les cavités des circuits (causées par la décapsulation).

3.3.2 Supports 3D

Pour pouvoir fixer à la fois le circuit flexible et les fils, un support 3D en forme de cylindre a été modélisé à l'aide du logiciel de modélisation 3D FreeCAD. Ce dernier est creux afin de pouvoir y placer le cône de ferrite précédemment présenté si nécessaire. Il possède également deux rainures opposées, qui servent à guider la remontée des fils le long du support. Puisque l'envergure des sondes a été fixée à 4.5mm et 3.5mm, on distingue deux tailles de cylindre différentes. Le plus gros servira de support pour les sondes de 650μm et 520μm tandis que le deuxième servira pour les sondes de 400μm et 300μm.

Enfin, il possède une base percée de deux trous, qui serviront à le fixer sur la dernière partie du circuit à l'aide de deux vis : Le PCB rigide. Une vue des différents supports réalisés est donnée sur la Figure 3.7.

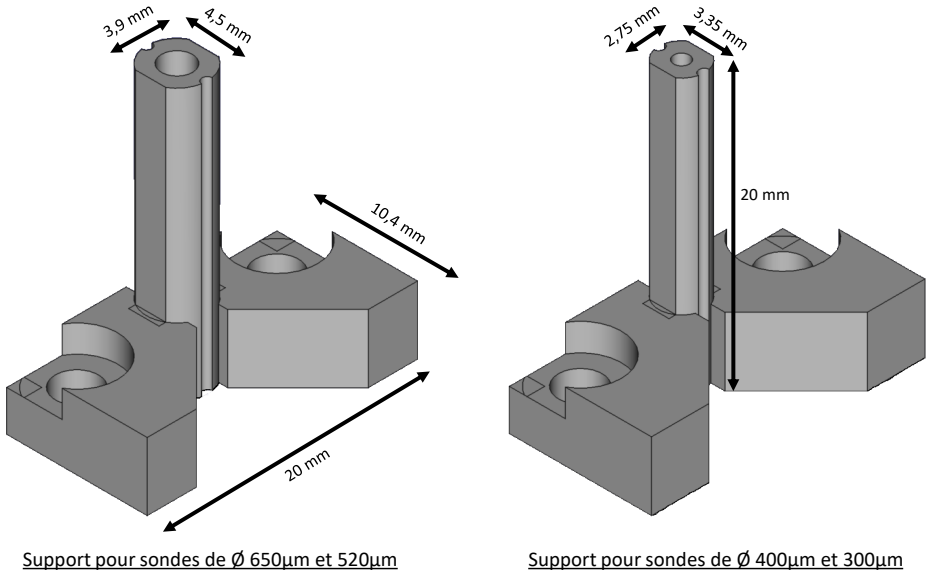


FIGURE 3.7 – Modélisation 3D des supports cylindriques

Ensuite, la question s'est posée de l'adaptation sur un banc 3 axes XYZ, de manière à pouvoir effectuer des cartographies de fautes, ou toute autre expérimentation nécessitant de devoir déplacer la sonde très précisément au-dessus d'un circuit. Un support de fixation compatible Newport et Thorlabs a donc été modélisé puis imprimé, et est présenté Figure 3.8.

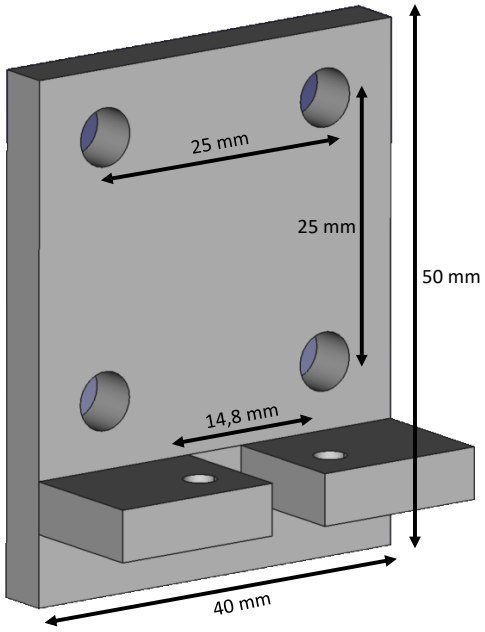


FIGURE 3.8 – Support 3D de fixation compatible Newport/Thorlabs

Il possède en tout 6 trous. Quatre d'entre eux sont de taille M4 et servent pour la fixation sur le banc XYZ, les deux autres servent à fixer solidement la sonde sur ce support. Tous les supports 3D présentés ont été imprimés à l'aide de l'imprimante 3D Ultimaker 3.

3.3.3 PCB rigide

Le PCB rigide a une taille de 2cm par 2cm et possède un layout tout ce qu'il y a de plus simple. Il se compose d'un connecteur SMA bord de carte et de quatre trous. Deux d'entre eux servent à passer les deux fils, qui sont ensuite soudés directement sur le connecteur SMA. Les deux autres servent à passer deux vis, de manière à y fixer le support 3D cylindrique, et de venir fixer le tout sur le support de fixation compatible Newport/Thorlabs. Le layout du PCB rigide est donné sur la Figure 3.9.

La seule différence notable entre les deux est l'espacement des deux trous servant à laisser passer les fils. En effet, comme vu précédemment, l'espacement entre les pads de soudure diffère selon la sonde que l'on souhaite fabriquer et donc selon le support utilisé (petite ferrite ou grosse ferrite).

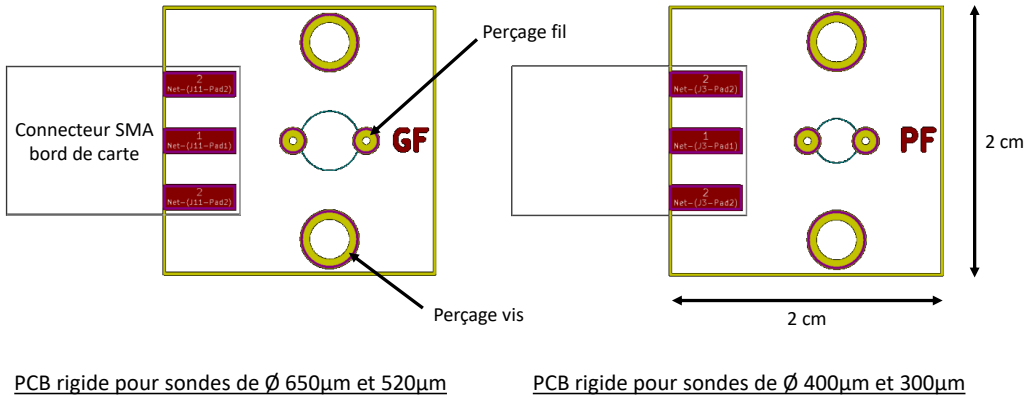


FIGURE 3.9 – Layout du PCB rigide

3.3.4 Procédé de fabrication

Le processus de fabrication est relativement simple également, mais surtout reproductible, ce qui n'était pas (ou peu) le cas dans [Ord+09]. Après avoir choisi la taille de la sonde que l'on souhaite fabriquer, la première étape consiste à découper la boucle correspondante à l'aide d'un ciseau, et de souder sur les deux pads prévus à cet effet les deux fils de 150μm. Ces soudures étant fragiles, un point de colle est rajouté par-dessus de manière à solidifier l'ensemble. Ensuite, selon que l'on souhaite réaliser une sonde avec noyau de ferrite ou non, on vient coller (ou non) le cône de ferrite dans le trou prévu à cet effet.

La deuxième étape consiste à positionner et coller le support cylindrique. Pour ce faire, son extrémité est plongée dans la colle et est ensuite positionnée au centre de la boucle en maintenant une légère pression le temps que la colle commence à sécher, le tout en prenant soin de nettoyer le surplus de colle éventuel, de manière à ne pas créer plus d'encombrement lors du séchage.

Ensuite, les fils sont guidés à l'aide des deux rainures et sont passés dans les deux trous percés à cet effet, puis l'un d'entre eux est soudé sur la masse du connecteur SMA tandis que l'autre est soudé sur le pin central. Une fois fait, la sonde est fixée au support de fixation à l'aide de deux vis.

Enfin, la dernière étape consiste à découper le substrat flexible dépassant de l'extrémité du support à l'aide d'une pince coupante, de façon à ce qu'il vienne épouser parfaitement la forme du support cylindrique. Un aperçu du

prototype final de la sonde est donné Figure 3.10.

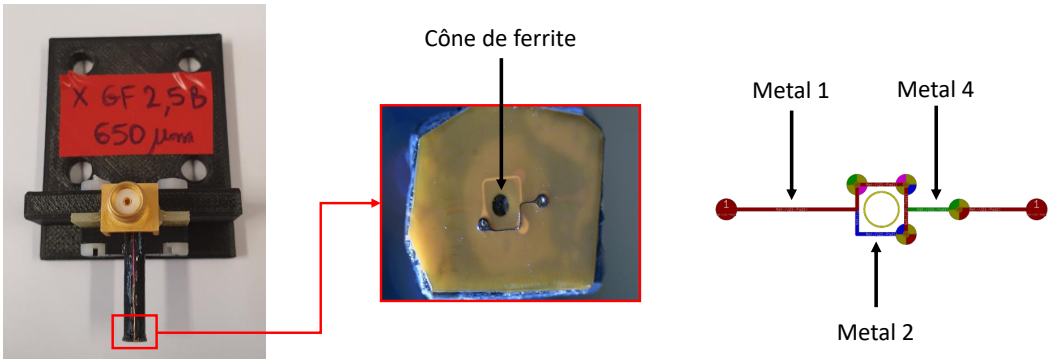


FIGURE 3.10 – Prototype final pour une sonde de 650μm de diamètre

3.4 Protocole de caractérisation de la plateforme et des sondes

3.4.1 Définition du protocole

D'après [DLM21], l'injection de fautes électromagnétique repose sur le principe de l'induction électromagnétique, et donc sur la loi de Faraday. Plus précisément, et comme illustré Figure 3.11, il y a un couplage entre la sonde électromagnétique (une simple bobine d'inductance L_P) et les boucles fermées (d'inductance L_i^{IC}) formées par les réseaux d'alimentation des circuits intégrés ciblés. Ce couplage électromagnétique crée en retour plusieurs transformateurs de tension d'inductance mutuelle $M_i \propto \sqrt{L_i^{IC} \cdot L_P}$ (une pour chaque boucle fermée dans le réseau d'alimentation) entre le générateur d'impulsion et les circuits intégrés.

Par conséquent, appliquer une impulsion à l'entrée de la sonde EM induit plusieurs impulsions (une pour chaque boucle fermée du réseau d'alimentation) de plus faibles amplitudes dans le réseau d'alimentation du circuit intégré. Toujours d'après [DLM21], à cause de la topologie régulière des grilles d'alimentation, ces impulsions possèdent une amplitude importante seulement sous les bords de la sonde EM, et une amplitude modérée sous son noyau (généralement constitué de ferrite). De fait, la résolution spatiale

d'une injection de fautes électromagnétique est déterminée par le diamètre de la sonde utilisée. Par conséquent, si l'impulsion délivrée à la sonde est d'amplitude et de largeur suffisante, des fautes transitoires apparaissent sous les bords de celle-ci.

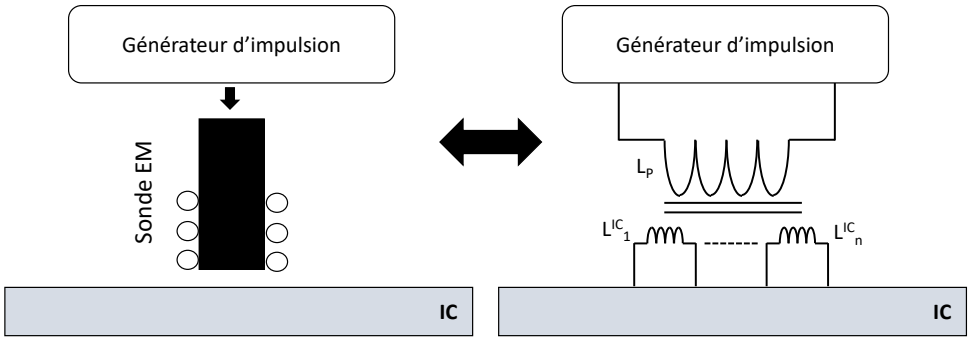


FIGURE 3.11 – Principe de l'EMFI

Selon le dernier point évoqué, pour être efficace, une plateforme d'injection électromagnétique doit être modulable (V_{pulse} et PW doivent être précisément contrôlables), et doit générer :

— Une impulsion électromagnétique puissante pour être en mesure d'injecter des fautes dans n'importe quel circuit, et ce malgré l'évolution des technologies CMOS vers des technologies de plus en plus agressives (menant de fait à des boucles de plus en plus petites à l'intérieur des réseaux d'alimentation, mais ne suivant pas la loi de Moore),

— Une impulsion électromagnétique courte pour pouvoir attaquer à la fois des circuits à haute (> $100MHz$) et basse (< $100MHz$) fréquence de fonctionnement, et ainsi ne perturber qu'un nombre restreint d'instructions (ou cycles d'horloge) à la fois.

Au vu de ces informations, un protocole simple a été adopté. Il permet de caractériser et comparer les plateformes d'EMFI en regardant l'évolution de l'impulsion délivrée en fonction de V_{pulse} et PW, et de non plus se limiter aux caractéristiques de la sonde ou du générateur d'impulsion. C'est ce protocole qui sera suivi pour caractériser les sondes d'injection présentées précédemment. Il consiste à mesurer, en fonction de V_{pulse} et PW, l'évolution de :

— L'amplitude maximale de la tension induite, notée ci-après $V_{induced}$,

— La largeur à mi-hauteur ($FWHM$: Full Width at Half Maximum),

— $t_{5\%}$ défini ici comme le temps nécessaire pour que l'impulsion induite dans la sonde de référence disparaisse (mesuré ici à 5% de $V_{induced}$). Cette sonde de référence est placée au contact de la sonde d'injection EM, et est chargée par une résistance de 50Ω (l'entrée d'un oscilloscope numérique réglé sur le mode DC 50Ω), comme illustré Figure 3.12.

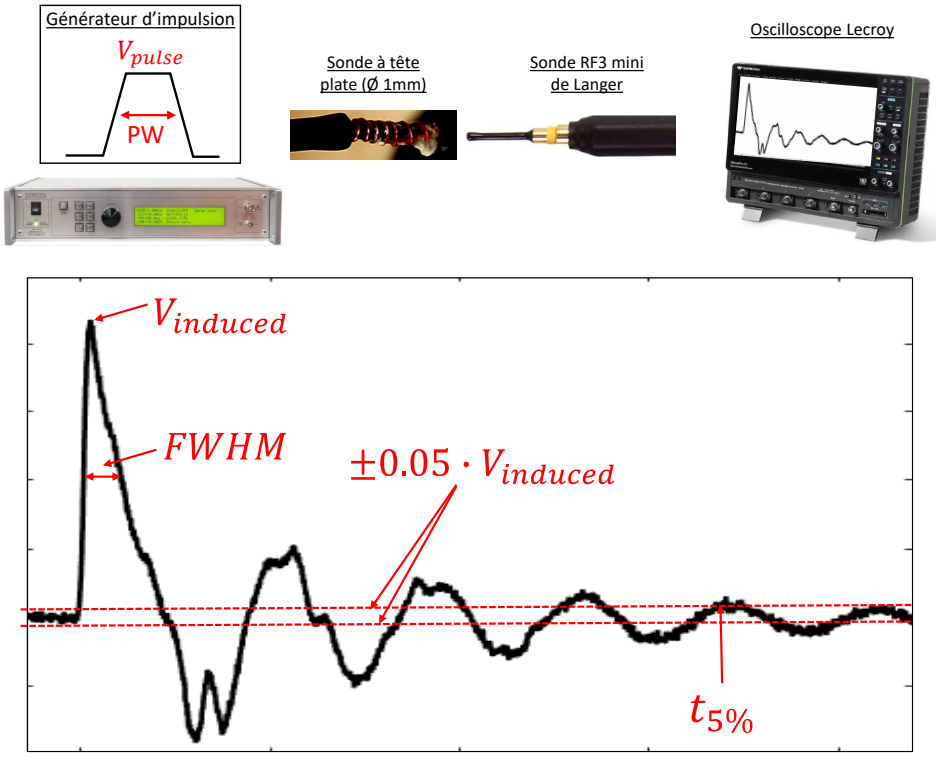


FIGURE 3.12 – Set-up expérimental et figures de mérite pour la caractérisation d'une plateforme d'EMFI

Il ne reste alors plus qu'à choisir la sonde de référence. Ici, la sonde RF3 mini de chez Langer est utilisée. Elle est adaptée 50Ω, a une bande passante très large (de $30MHz$ à $3GHz$), est résistante aux grandes variations de tension et est disponible dans la plupart des laboratoires impliqués dans la caractérisation de sécurité du fait de son prix abordable.

3.4.2 Application à deux plateformes : LIRMM et Langer

Le protocole de caractérisation précédent a été appliqué à une plateforme d'EMFI basée sur le générateur d'impulsion Avtech ainsi qu'à la plateforme développée et commercialisée par Langer. La Figure 3.13 donne, pour chacune des deux plateformes, l'évolution de $V_{induced}$, $FWHM$ et $t_{5\%}$ en fonction de V_{pulse}, pour différentes valeurs de PW.

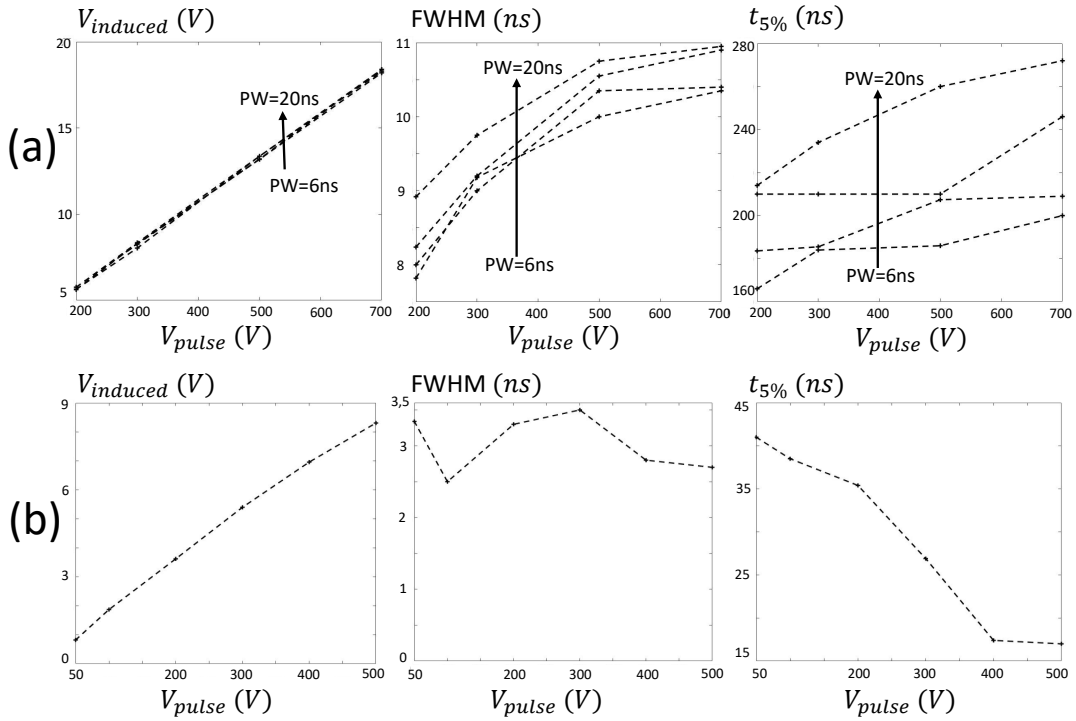


FIGURE 3.13 – Courbes de caractérisation sans système anti-rebonds

La première plateforme s'articule autour du générateur $AVRK4 - B$ de la marque Avtech, délivrant dans une résistance de 50Ω des impulsions d'amplitudes variant de $200V$ à $750V$ (avec choix de polarité positive ou négative) et de largeur comprise entre $6ns$ et $20ns$. Ce générateur possède une entrée trigger externe, permettant de générer sur la même sortie, deux impulsions indépendantes et entièrement contrôlables. Le délai D entre ces deux triggers externes et les impulsions peut être réglé entre $100ns$ et $1s$; le temps minimum séparant les deux impulsions étant de $100ns$. Le jitter entre les triggers et les impulsions est égal à $\pm 100ps + 0.03\% \cdot D$. La sonde EM utilisée sur cette plateforme a été fabriquée manuellement autour d'un noyau de ferrite et se compose de 5 boucles, pour un diamètre de $1mm$, avec une tête plate.

La plateforme commercialisée par Langer délivre des impulsions dont les amplitudes varient entre $50V$ et $500V$, à une sonde EM également fabriquée autour d'un noyau de ferrite, qui est illustrée Figure 3.14. Cette sonde possède une tête pointue pour un diamètre de $500\mu m$ (aucune information n'est donnée dans la fiche technique concernant le nombre de boucles). La largeur de l'impulsion n'est pas modifiable, et est fixée à $2ns$. Le jitter du générateur d'impulsion est égal à $\pm 1ns$.

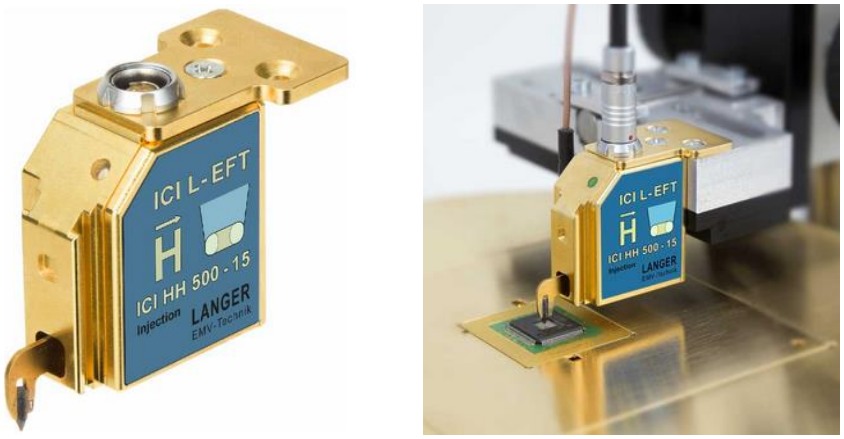


FIGURE 3.14 – Sonde EMFI de la plateforme commercialisée par Langer

Sur la Figure 3.13 (a), on remarque que l'amplitude $V_{induced}$ de la perturbation induite dans la sonde RF3 mini de langer est indépendante de PW et atteint $18V$ pour V_{pulse} égal à $750V$. Le courant induit a donc une amplitude maximale de $0.36A$ puisque l'oscilloscope est réglé sur le mode DC 50Ω.

La largeur à mi-hauteur ($FWHM$) varie entre $7.8ns$ et $11ns$ selon la valeur de PW. Cette observation est également valable pour $t_{5\%}$, dont les valeurs varient entre $160ns$ et $260ns$. Ceci est la preuve que la résolution temporelle de la plateforme est très faible. Cette faible résolution temporelle est liée à la présence de rebonds dans la perturbation EM émise, rebonds causés par la désadaptation d'impédance entre le générateur et la sonde.

La Figure 3.13 (b) montre la même évolution pour la plateforme Langer mais sur une seule courbe, la valeur de PW n'étant pas modifiable. On observe que $V_{induced}$ varie entre $0.8V$ et $8.31V$, valeurs plus faibles que pour

la plateforme basée sur le générateur d'impulsion Avtech. La largeur à mi-hauteur reste à peu près constante et proche de $3ns$ pour toutes les valeurs de V_{pulse}. Enfin concernant $t_{5\%}$, il décroît de $41ns$ à $17ns$ lorsque les valeurs de V_{pulse} augmentent.

3.4.3 Conclusion

De nombreuses plateformes d'EMFI sont désormais disponibles dans la littérature ou commercialisées en tant que COTS. Les caractéristiques listées dans les fiches techniques sont propres à chacune d'entre elles, ce qui rend compliquée leur comparaison. Dès lors, choisir la plateforme la plus adaptée à son application devient complexe. Il a donc été proposé dans cette partie un protocole de caractérisation qui peut être appliqué à l'aide d'équipement bon marché, se trouvant généralement dans les laboratoires de caractérisation sécuritaire. Le protocole défini se base sur les caractéristiques liées à l'impulsion EM, et non plus sur le matériel utilisé pour la générer, comme cela pouvait être le cas avant. Ces principaux avantages sont sa rapidité de mise en place et d'exécution, et sa simplicité. En suivant ce protocole, nous avons comparé deux plateformes. D'une part celle commercialisée par la société Langer, et d'autre part celle développée au LIRMM, qui s'articule autour d'un générateur d'impulsion de la marque Avtech.

3.5 Caractérisation de la résolution temporelle

3.5.1 Description et conception d'un système anti-rebonds

Une première caractérisation a été menée, mais un problème persiste : les rebonds du signal EM qui réduisent grandement la résolution temporelle de la plateforme d'EMFI.

Pour améliorer la résolution temporelle des plateformes utilisant un générateur commercial (de type Avtech par exemple), on pourrait penser à insérer un transformateur de tension entre la sonde EM et le générateur d'impulsion. C'est une solution classique pour résoudre le problème d'adaptation d'impédance. D'ailleurs, cette solution est celle proposée par Avtech. Elle consiste à utiliser le transformateur de tension AVX-M4-H, de manière à pouvoir adapter des charges de l'ordre de 3Ω. Cependant, même si cette solution réduit de

manière significative le nombre de rebonds indésirables, elle n'est pas parfaite puisque les sondes EM possèdent une impédance de l'ordre de 1Ω. De plus, ce transformateur divise l'amplitude de l'impulsion par un facteur 2 à l'entrée de la sonde.

Pour résoudre ce problème et ainsi grandement améliorer la résolution temporelle, il est proposé dans la suite de ce document d'utiliser une diode Transil unidirectionnelle, possédant une tension de claquage (V_{Br}) et une tension limite (V_{Clamp}) très élevées et une tension d'avalanche (V_F) inverse très faible.

Cette solution, qui ne se limite pas aux générateurs d'impulsion de la gamme Avtech, permet de supprimer les rebonds entre la sonde et le générateur d'impulsion. Son fonctionnement est décrit Figure 3.15. Comme illustré, lorsque le déclenchement s'opère, une impulsion est générée sur $V+$. Cette impulsion se propage vers l'avant, passe au travers de la sonde et atteint finalement $V-$. Arrivée à $V-$, une partie de l'impulsion est absorbée par le générateur et le reste est réfléchi. Cette fraction d'impulsion réfléchie et atténuée repasse au travers de la sonde et génère un rebond sur le signal EM émis par la sonde, ce qui limite grandement la résolution temporelle de la plateforme EMFI. En fait, il n'y a pas une impulsion EM mais plusieurs, d'amplitudes décroissantes et de polarités alternées entre positives et négatives.

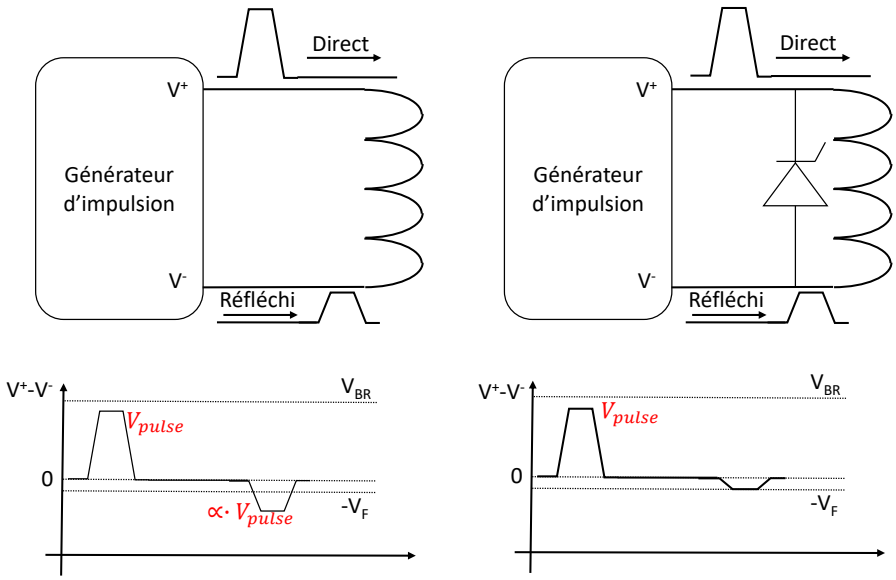


FIGURE 3.15 – Système anti-rebonds basé sur l'utilisation d'une diode Transil unidirectionnelle

Lorsqu'une diode Transil unidirectionnelle est placée comme indiqué sur la Figure 3.15, l'impulsion se propageant de $V+$ vers $V-$ passe toujours au travers de la sonde EM si :

$$V_{pulse} = (V^+ - V^-) < V_{BR}$$

Et une impulsion électromagnétique est alors toujours émise par la sonde. Cependant, l'impulsion réfléchie ne repasse pas dans la sonde et est dissipée si :

$$\alpha \cdot V_{pulse} = (V^- - V^+) > V_F$$

En effet, dans ce cas, la diode devient conductrice et court-circuite la sonde EM. Il en résulte la disparition des rebonds sur le signal émis par la sonde.

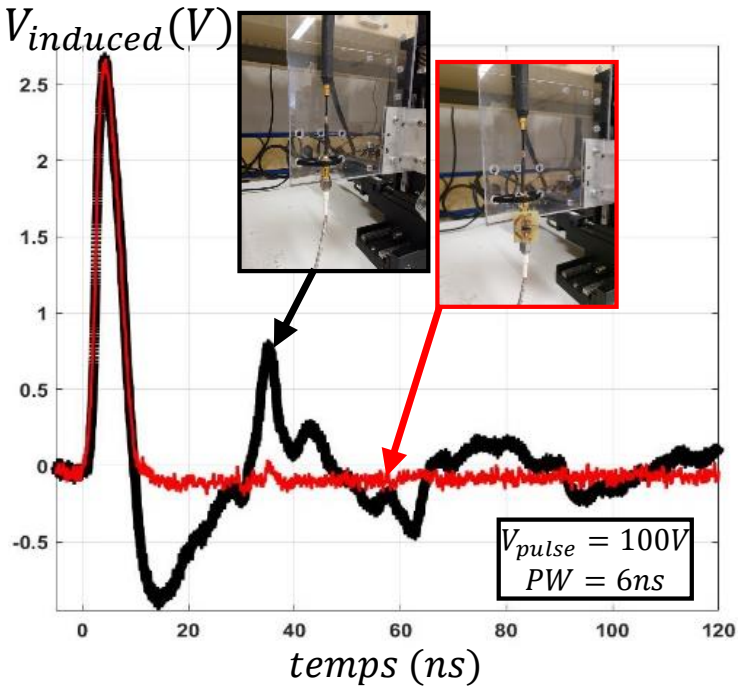


FIGURE 3.16 – Signal EM émis par la sonde dans une RF3 mini avec et sans diode Transil, pour une impulsion d'amplitude $100V$ et de largeur $6ns$

Cette solution a été implémentée sur un petit PCB ($3cm \times 2cm$) de manière à ce qu'il puisse être placé au plus proche de la sonde, i.e. directement

connecté à celle-ci. La diode Transil utilisée est une diode Transil unidirectionnelle $1.5KE600A$ fabriquée par Littelfuse. Ses caractéristiques sont les suivantes :

$$V_{Br} = 570 \text{ V}$$
$$V_{Clamp} = 860 \text{ V}$$
$$V_F = 1 \text{ V}$$

Elle peut dissiper une impulsion jusqu'à une puissance de $1500W$ et possède un temps de réponse très rapide de l'ordre de $1ps$. La Figure 3.16 illustre l'efficacité de cette solution, en montrant la tension mesurée aux bornes de la sonde RF3 mini de Langer avec et sans système anti-rebonds. Cette tension, $V_{induced}$, est représentative du courant circulant dans la sonde et donc du rayonnement EM induit par la sonde d'injection. On peut s'apercevoir que la première impulsion est entièrement transmise alors que les suivantes sont supprimées. De plus, l'amplitude de la première impulsion ne s'en trouve pas diminuée.

3.5.2 Caractérisation de la plateforme LIRMM équipée d'une sonde ancienne génération (avec système anti-rebonds)

La caractérisation de la plateforme s'articulant autour du générateur d'impulsion Avtech a été effectuée de la même manière que précédemment, mais en ajoutant le système anti-rebonds au plus près de la sonde EM.

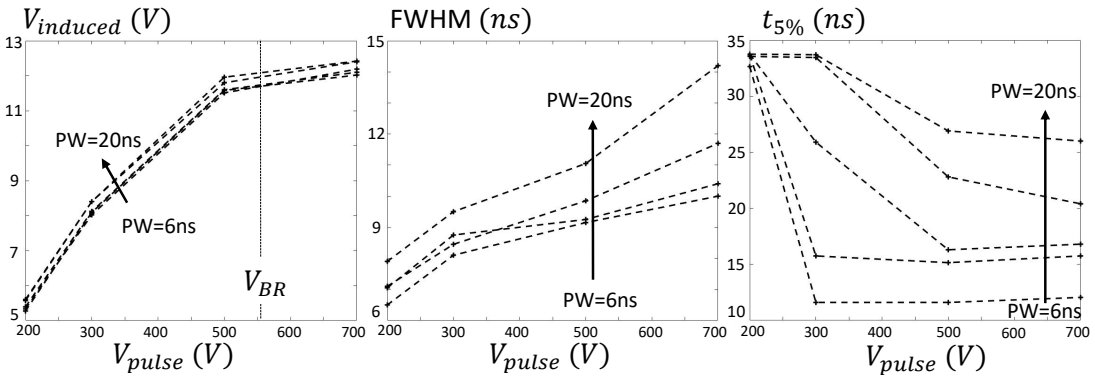


FIGURE 3.17 – Courbes de caractérisation avec système anti-rebonds

La Figure 3.17 montre l'évolution de $V_{induced}$, $FWHM$ et $t_{5\%}$ en fonction de V_{pulse}, pour différentes valeurs de PW. Si l'on compare ces résultats à ceux de la Figure 3.13 (a), on en conclu que :

— La résolution temporelle ($t_{5\%}$) est améliorée d'un facteur ≈ 5,

— La largeur à mi-hauteur (FWHM) reste inchangée,

— Il n'y a pas d'atténuation de l'impulsion EM principale pour $V_{pulse} < V_{Br}$ = 560V. Cependant, il y a une atténuation linéaire (proportionnelle à $V_{pulse} - V_{Br}$) de l'impulsion dès lors que V_{pulse} devient supérieur à V_{Br}. Cette perte de puissance peut néanmoins être diminuée, voire même supprimée, en utilisant une diode Transil avec une tension de claquage plus élevée, si cela existe.

3.5.3 Caractérisation de la plateforme LIRMM équipée d'une sonde nouvelle génération (avec système anti-rebonds)

Ce protocole de caractérisation a également été utilisé pour caractériser les sondes fabriquées et présentées un peu plus tôt dans ce chapitre. Par soucis de clarté et de simplicité, les résultats pour seulement quatre d'entre elles seront présentés dans la suite.

On considérera ainsi deux sondes de 650μm de diamètre (une avec noyau de ferrite et l'autre sans) et deux sondes de 400μm de diamètre sans ferrite (une possédant 2.5 boucles et l'autre 1.5). De cette manière, il sera possible d'évaluer l'effet de la ferrite, du diamètre et du nombre de boucles sur les différentes figures de mérite : $V_{induced}$, $FWHM$ et $t_{5\%}$. Pour chaque expérimentation, quatre valeurs de PW ont été considérées (6, 8, 10 et 20ns) pour V_{pulse} variant de 200 à 800V par pas de 100V.

La Figure 3.18 donne l'évolution de $V_{induced}$ en fonction de V_{pulse}. Si l'on commence par regarder les courbes de gauche pour lesquelles le système anti-rebonds a été utilisé, on remarque que l'évolution de $V_{induced}$ n'est pas linéaire, mais présente une légère atténuation aux alentours de $V_{pulse} = 600V$,

valeur qui correspond à la tension de claquage (V_{Br}) de la diode Transil utilisée. Malgré cette légère atténuation, les valeurs de $V_{induced}$ restent relativement similaires avec et sans système anti-rebonds.

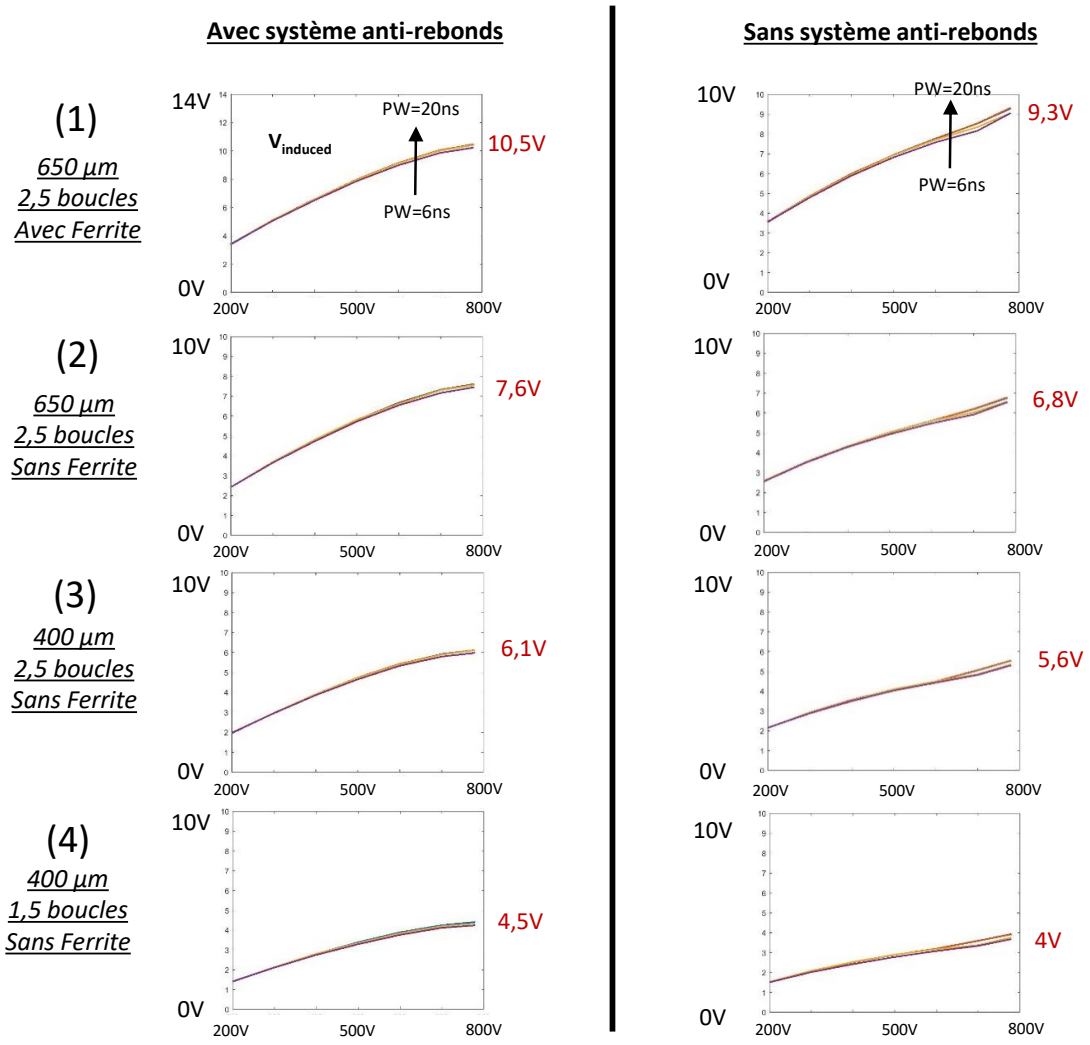


FIGURE 3.18 – Évolution de $V_{induced}$ en fonction de V_{pulse} pour différentes valeurs de PW

Ensuite, si l'on considère les sondes 1 et 2, on se rend compte que le noyau de ferrite donne des performances environ 38% plus élevées pour une sonde de même diamètre et possédant le même nombre de boucles. Le même constat est fait pour deux sondes de diamètre différents, sans ferrite et possédant 2.5 boucles (sondes 2 et 3) ; et pour deux sondes possédant un nombre de boucles différents, sans ferrite et de diamètre $400\mu m$ (sondes 3 et 4).

La Figure 3.19 donne l'évolution de $FWHM$ en fonction de V_{pulse}. Ici, pas de différences notables si ce n'est que la largeur à mi-hauteur reste plus "stable" pour des valeurs de $V_{pulse} > 600V$ lorsque le système anti-rebonds est utilisé.

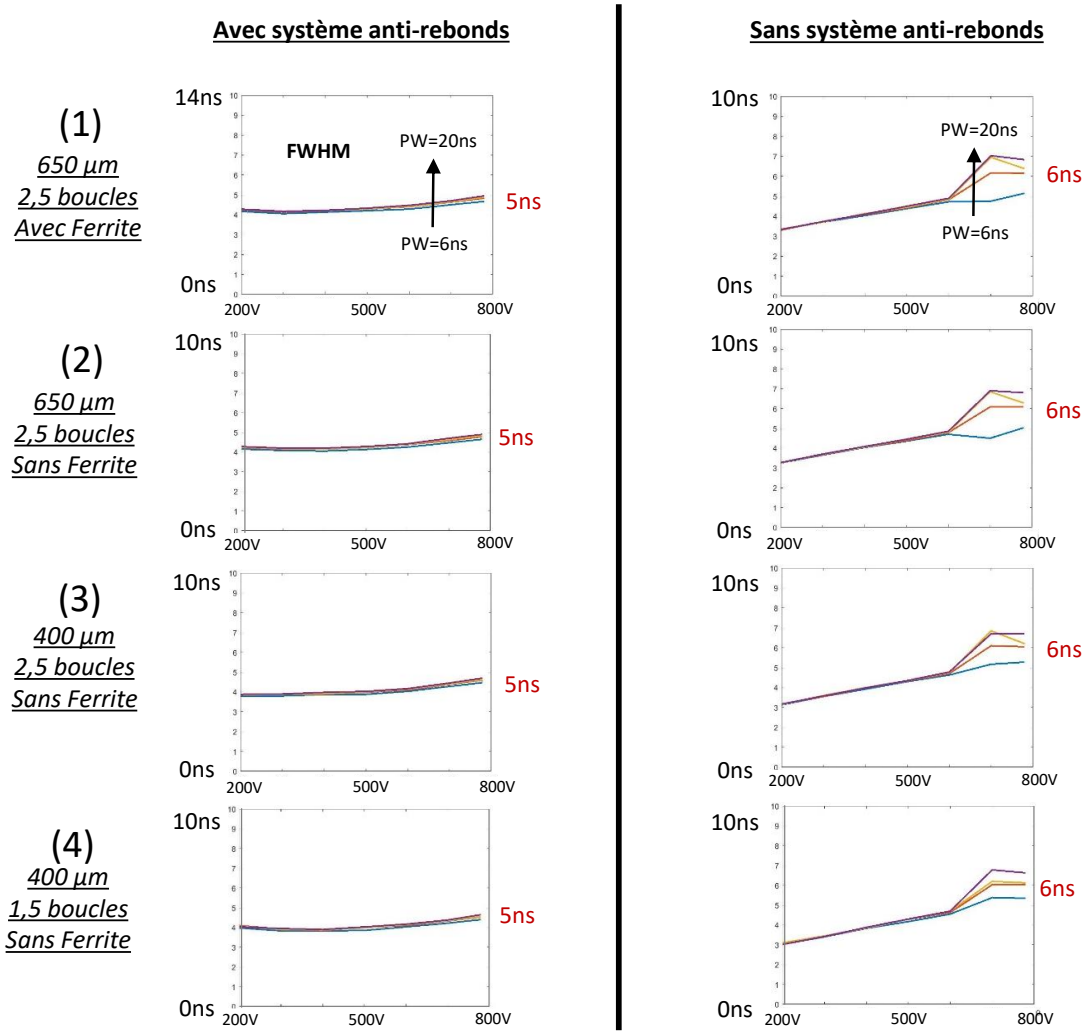


FIGURE 3.19 – Évolution de $FWHM$ en fonction de V_{pulse} pour différentes valeurs de PW

La Figure 3.20 donne l'évolution de $t_{5\%}$ en fonction de V_{pulse}. L'efficacité du système anti-rebonds est encore une fois mise en évidence ici, avec des temps de réponse à 5% divisés environ par 2.

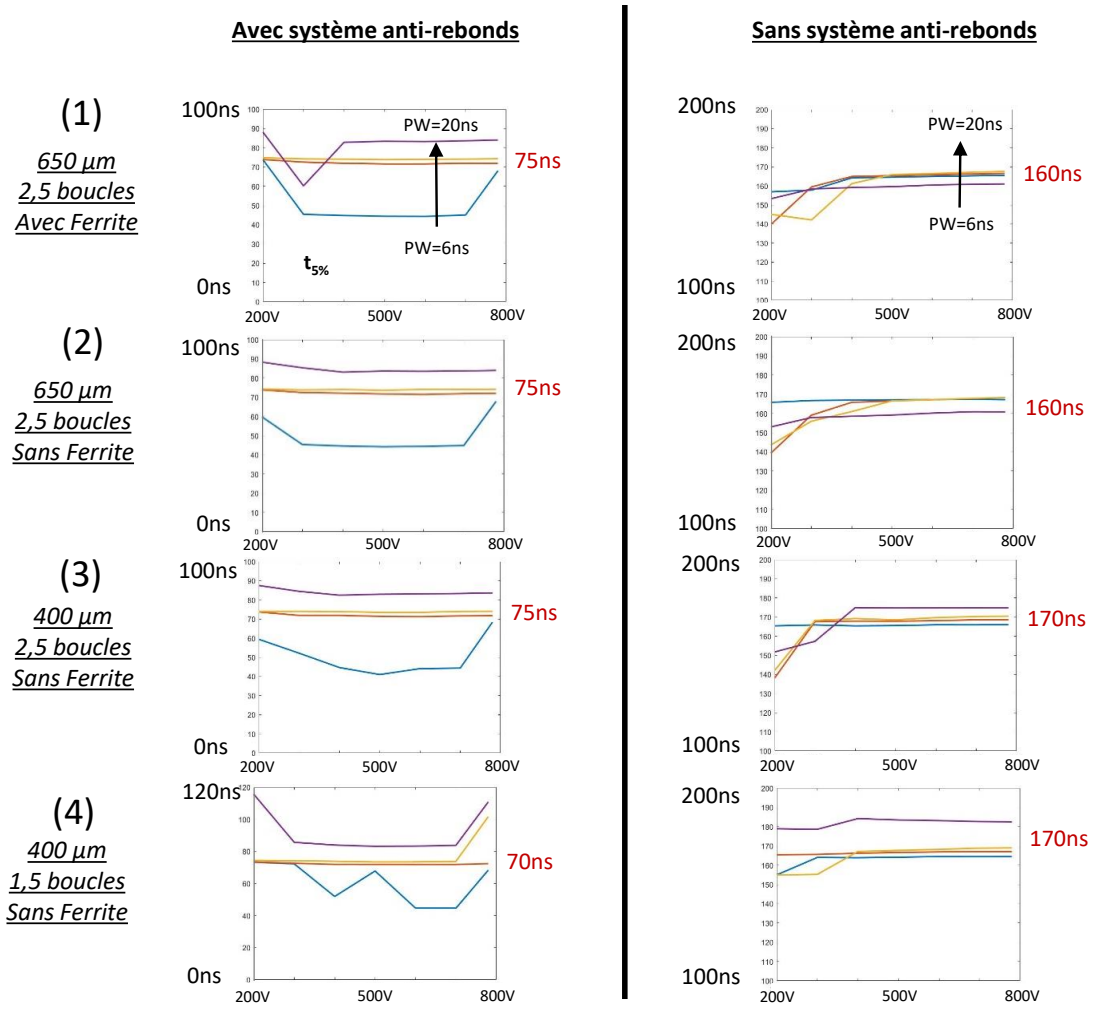


FIGURE 3.20 – Évolution de $t_{5\%}$ en fonction de V_{pulse} pour différentes valeurs de PW

3.5.4 Comparaison et discussion

Cette partie a été dédiée à la conception et à la réalisation d'un système simple et compact permettant de supprimer les rebonds sur le signal EM émis par la sonde d'injection. Ce système, à base de diode Transil, permet la suppression complète de tous les rebonds sans atténuation de l'amplitude de la première impulsion.

Suite à cela, la plateforme du LIRMM a été caractérisée à l'aide du protocole défini dans la partie 3.4.1, dans un premier temps avec une sonde ancienne génération dont le LIRMM disposait avant mon arrivée en thèse, puis avec une sonde nouvelle génération issue de mes travaux de recherche. Chacune de ces caractérisations a été menée avec et sans système anti-rebonds.

Les courbes de caractérisation montrent que le système anti-rebonds fabriqué ne modifie ni la largeur à mi-hauteur FWHM, ni la tension induite $V_{induced}$ (si l'on ne prend pas en compte la légère atténuation au delà de la valeur de V_{Br}, qui reste somme toute négligeable). En revanche, il diminue d'un facteur 2 à 5 les temps de réponse à 5%, $t_{5\%}$. Ici, il est important de noter que le gain en résolution temporelle est plus important sur les sondes ancienne génération (amélioration d'un facteur 5) que sur les sondes nouvelle génération (amélioration d'un facteur 2). En revanche, ces dernières sont 2 à 4 fois plus petites, ce qui permet une meilleure sélectivité et une résolution spatiale accrue.

3.6 Caractérisation de la résolution spatiale

La plateforme utilisée pour les expérimentations s'articule autour du générateur AVRK4 de la marque Avtech. Celui-ci délivre des impulsions d'amplitude comprise entre $100V$ et $750V$. Son slew rate dans une charge de 50Ω est constant et égal à $187V/ns$. Ainsi, les temps de montée et de descente pour une impulsion d'amplitude $V_P = 750V$ sont de $4ns$. Ce générateur possède une résistance de sortie de 50Ω. On retrouve donc le problème de désadaptation d'impédance qui est corrigé à l'aide du système anti-rebonds présenté dans la partie précédente.

Les sondes utilisées pour cette caractérisation sont celles présentées précédemment. Pour des raisons de simplicité, mais sans pour autant perdre en généralité, seules les sondes de diamètres $300\mu m$, $400\mu m$ et $650\mu m$ sans noyau de ferrite seront considérées. Des résultats similaires sont obtenus pour les sondes possédant un noyau de ferrite, avec des valeurs de V_P 30% moins élevées.

Plusieurs campagnes d'acquisition ont été menées avec les différentes sondes, sur le même circuit. Le circuit attaqué est un FPGA de Xilinx (Spartan3E–1600), dans lequel a été programmé un AES 128 bits s'exécutant à $50MHz$. Le boîtier plastique du circuit a été retiré à l'aide de procédés chimiques, et la sonde a donc été placée précisément à une distance $h = 50\mu m$ de la surface de la puce.

3.6.1 Cartographies V_P^{min}

Dans une première expérimentation, des balayages d'une même partie de la puce contenant l'AES ont été réalisés à l'aide des trois sondes considérées. Ces balayages ont été réalisés de manière à mesurer l'amplitude minimale V_P^{min}, à appliquer à la sonde à chaque position pour induire une faute (si possible) lors de l'exécution de l'AES. Le déplacement de la sonde a été réglé à $100\mu m$ et la recherche de V_P^{min} a été limitée à la plage $V_P \in [100V, 500V]$, par pas de $20V$.

Les résultats obtenus ont ensuite été traités pour obtenir la Figure 3.21, qui montre les histogrammes de V_P^{min} obtenus avec les trois sondes.

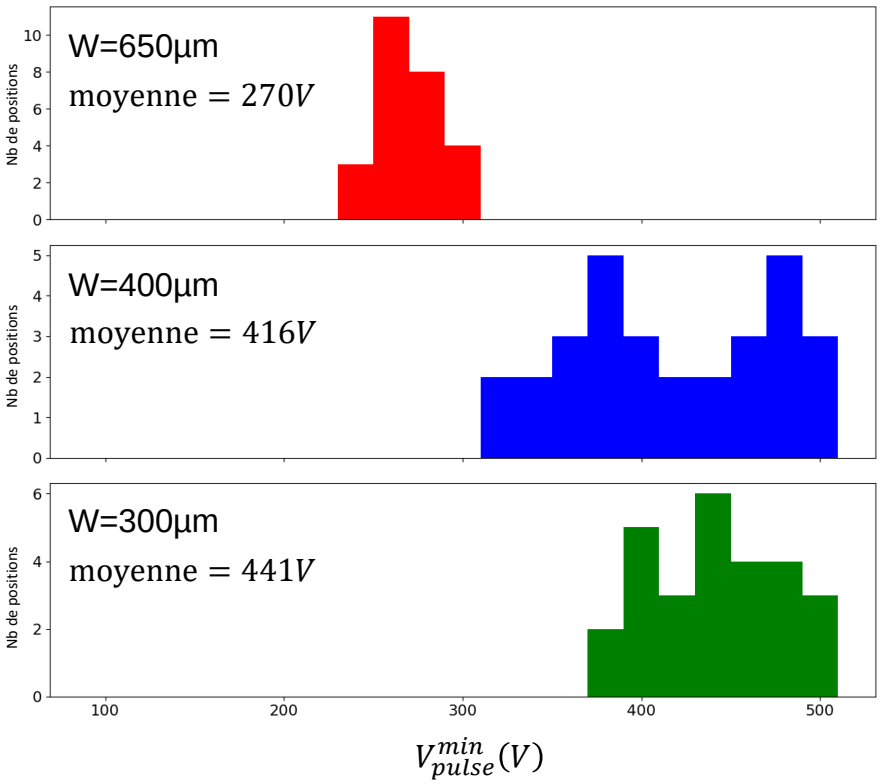


FIGURE 3.21 – Histogrammes de V_P^{min} obtenus avec les trois sondes (W = $650\mu m$, $400\mu m$ et $300\mu m$) à une hauteur $h = 50\mu m$. Pendant les balayages, PW est fixé à $10ns$

Comme attendu, on observe un décalage des histogrammes vers la droite lorsque l'on diminue les dimensions de la sonde. En considérant que les valeurs minimales de V_P^{min} pour chaque histogramme sont fiables malgré un

possible désalignement des cartographies, les valeurs expérimentales suivantes ont été calculées :

$$\frac{V_{P2}^{Exp}}{V_{P1}^{Exp}} \simeq \frac{min(V_{P2}^{min})}{min(V_{P1}^{min})} \tag{3.9}$$

Elles ont été comparées à l'équation 3.8 mais également à $\frac{W_2}{W_1}$, pour obtenir les tendances tracées sur la Figure 3.22.

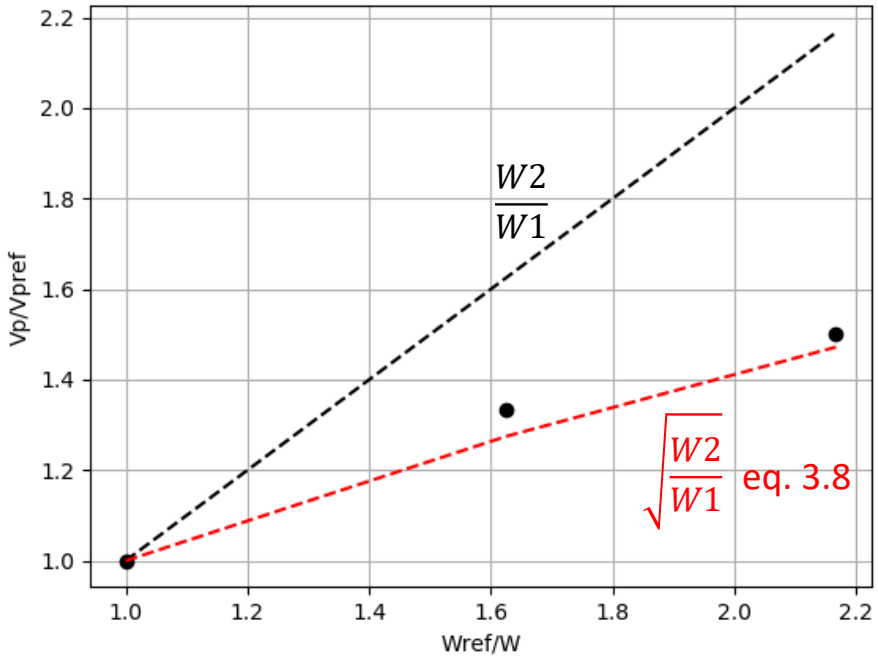


FIGURE 3.22 – Valeurs théoriques et expérimentales de $\frac{V_{P2}}{V_{P1}}$ provenant de l'équation 3.8 et de la Figure 3.21

Comme on peut le remarquer, les valeurs expérimentales sont très proches de celles prédites par l'équation 3.8, et bien inférieures à la tendance $\frac{W_2}{W_1}$. Cela confirme donc la règle d'évolution de V_P énoncée précédemment et montre clairement que la puissance du générateur d'impulsion évolue de manière sous-linéaire par rapport au ratio des dimensions de la sonde $\frac{W_2}{W_1}$.

Cependant, tracer une tendance avec seulement trois points n'est pas très rigoureux. C'est pour cette raison que nous avons été amenés à considérer trois sondes supplémentaires. La même expérimentation a donc été menée

avec les sondes d'analyse EM présentées dans le Chapitre 2. Tous les paramètres et réglages de la plateforme sont restés identiques, seul le design des sondes diffère d'une campagne d'acquisition à l'autre. Les sondes d'analyse EM n'ayant pas une forme carrée, mais plutôt une forme polygonale (une section carrée correspondant à la boucle et une section rectangulaire correspondant à la remontée des deux fils vers les pads de soudure), il convient désormais de considérer une formule plus générale pour leurs inductances propres, qui prend en compte le périmètre de la sonde. Toujours d'après [Tho91], pour un polygone de périmètre P et d'aire A, on a :

$$L_p = \frac{\mu_0 \cdot P}{2 \cdot \pi} \left[ln(\frac{2 \cdot P}{R}) + 0.25 - ln(\frac{P^2}{A}) \right] \tag{3.10}$$

Le rapport des tensions V_{P2} et V_{P1} peut alors s'approximer comme précédemment, mais cette fois-ci en fonction des périmètres des sondes :

$$\frac{V_{P2}}{V_{P1}} \simeq \sqrt{\frac{P_1}{P_2}} \tag{3.11}$$

Les tensions minimales V_P^{min} nécessaires pour induire des fautes sur la même partie du circuit ont donc été mesurées pour les sondes de $150\mu m$, $100\mu m$ et $50\mu m$. Elles sont égales à 340V, 380V et 450V respectivement. Nous avons ainsi pu compléter la Figure 3.22 avec 3 points et une tendance supplémentaire, comme illustré sur la Figure 3.23.

On remarque que nos six points expérimentaux suivent bien l'évolution de la tendance $\sqrt{\frac{P_2}{P_1}}$, ce qui nous conforte dans l'idée que ce facteur de dimensionnement est celui à suivre lors de la réduction de la taille des sondes. Ce dernier graphique nous permet également de conclure qu'il est plus intéressant, pour des sondes de petites tailles, de privilégier des formes rectangulaires ou polygonales plutôt que carrées. En effet, dans notre exemple, pour une sonde de $50\mu m$, une tension de $450V$ semble suffisante pour induire des fautes dans le circuit, alors qu'il aurait théoriquement fallu au minimum $865V$ si la sonde avait été carrée.

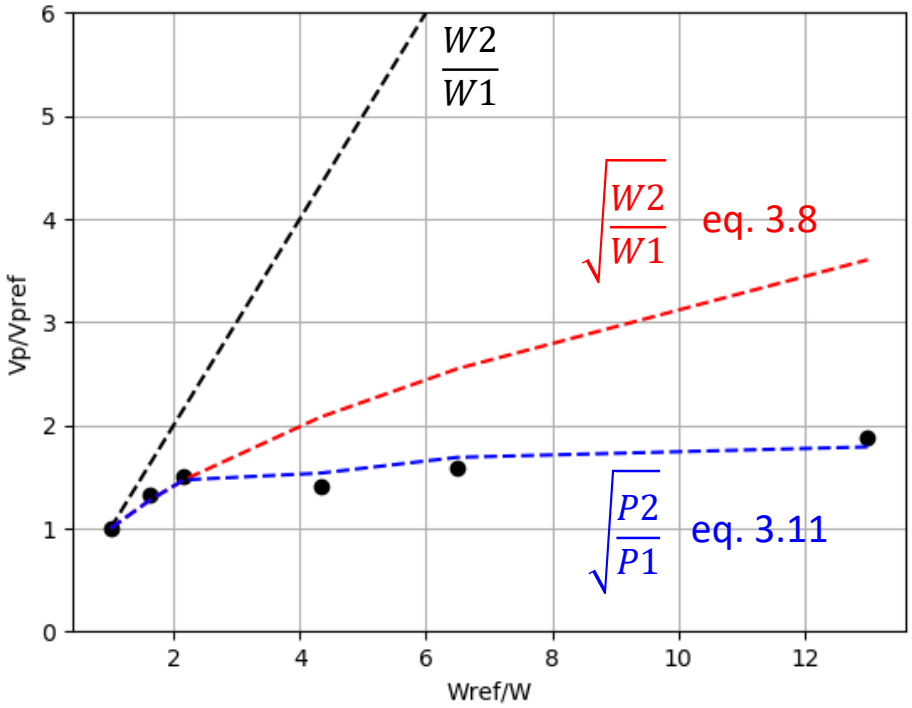


FIGURE 3.23 – Valeurs théoriques et expérimentales de $\frac{V_{P2}}{V_{P1}}$ provenant de l'équation 3.11

3.6.2 Probabilité de fautes

Pour confirmer la fiabilité de l'équation 3.8, trois balayages ont été effectués avec les trois sondes fabriquées pour l'EMFI :

— Le premier en utilisant la plus grosse sonde ($W = 650\mu m$) avec $V_P = 240V$,

— Le deuxième en utilisant la sonde intermédiaire ($W = 400\mu m$) avec $V_P = 330V$,

— Le dernier en utilisant la sonde la plus petite ($W = 300\mu m$) avec $V_P = 400V$.

Les ratios $\frac{V_{P2}}{V_{P1}}$ sont alors égaux à 1.37 (contre 1.27 théoriquement) et 1.67 (contre 1.47 théoriquement) respectivement. En effet, une marge de sécurité (restant tout de même inférieure à 15%) a été prise en compte, car les expérimentations précédentes ont montré que l'équation 3.8 sous-estime légèrement le vrai facteur d'évolution de V_P. Pour mieux estimer l'importance de

la marge considérée, il est important de noter que les rapports auraient été égaux à 1.62 et 2.16 si l'évolution avait été linéaire ($\frac{W_2}{W_1}$), avec donc des V_P égaux à $388V$ et $518V$ respectivement.

Pendant les balayages, plusieurs injections ont lieu à chaque position de la sonde au-dessus de la surface du circuit (6 pour être précis), de manière à calculer la probabilité d'induire une faute. Le FPGA est reprogrammé à chaque changement de position de la sonde pour éviter que des fautes persistantes ne viennent fausser les résultats. La Figure 3.24 montre les cartes de probabilités obtenues à la fin des trois balayages. Les pixels rouges correspondent aux positions de la sonde provoquant un crash du FPGA.

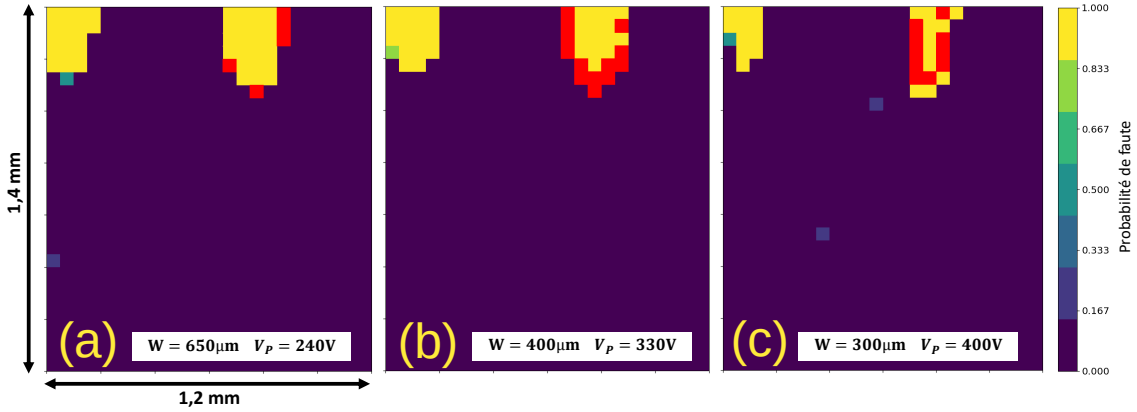


FIGURE 3.24 – Cartes de probabilités pour (a) $W = 650\mu m$ et $V_{pulse} = 240V$, (b) $W = 400\mu m$ et $V_{pulse} = 330V$ et (c) $W = 300\mu m$ et $V_{pulse} = 400V$. Les pixels rouges correspondent aux positions de la sonde provoquant un crash du FPGA.

On peut d'ores et déjà observer que l'EMFI possède une très grande reproductibilité. En effet, la probabilité de faire une faute est soit égale à 1, soit égale à 0 la plupart du temps. Cependant, l'observation la plus importante est le fait que les trois cartographies sont très similaires, alors qu'elles ont été effectuées avec trois sondes de tailles différentes. Ceci confirme une nouvelle fois la fiabilité de l'équation 3.8, ou tout du moins que l'évolution de V_P est sous-linéaire.

Cela signifie qu'il y a possibilité d'être plus sélectif en utilisant des sondes d'injection EM plus petites, et donc de perturber le signal d'alimentation sur une zone plus restreinte. Ce résultat est d'autant plus intéressant qu'il signifie également qu'il est possible d'améliorer la résolution spatiale des plateformes d'EMFI, sans nécessairement entrer dans une course impossible vers

des générateurs d'impulsion de plus en plus puissants (ce qui aurait été le cas avec des facteurs de dimensionnement égaux à $\frac{W_2}{W_1}$ ou $\frac{{W_2}^2}{{W_1}^2}$).

Pour pousser l'étude encore plus loin, nous avons également considéré les trois sondes initialement fabriquées pour faire de l'analyse EM. La même expérimentation a été menée, ce qui a donné lieu à trois balayages supplémentaires :

— Le premier en utilisant la sonde de $150\mu m$ avec $V_P = 520V$,

— Le deuxième en utilisant la sonde de $100\mu m$ avec $V_P = 630V$,

— Le dernier en utilisant la sonde de $50\mu m$ avec $V_P = 780V$.

Comme sur la Figure 3.24, on retrouve sur la Figure 3.25 les deux lobes de fautes sur la partie supérieure des cartographies.

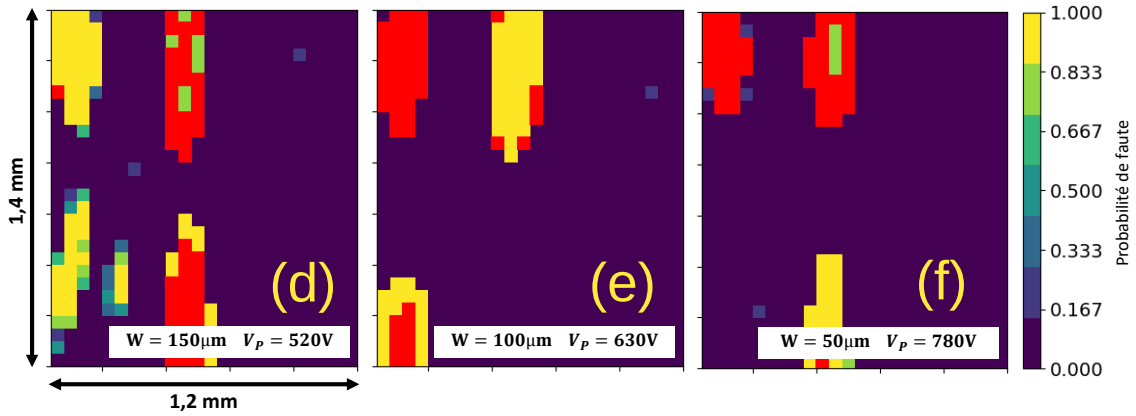


FIGURE 3.25 – Cartes de probabilités pour (d) $W = 150\mu m$ et $V_{pulse} = 520V$, (e) $W = 100\mu m$ et $V_{pulse} = 630V$ et (f) $W = 50\mu m$ et $V_{pulse} = 780V$. Les pixels rouges correspondent aux positions de la sonde provoquant un crash du FPGA.

Il y a cependant plusieurs différences notables concernant ceux-ci. La première est que le lobe de droite semble s'être légèrement décalé sur la gauche. On remarque également qu'il y a beaucoup plus de positions qui provoquent un crash du FPGA. La deuxième observation qui peut être faite est l'apparition de zones de fautes supplémentaires par rapport à l'expérimentation précédente (sur la partie inférieure des cartographies). Cela est notamment dû à une tension V_P trop importante. En effet, ces tensions ont été choisies

en suivant le dimensionnement en $\sqrt{\frac{W_1}{W_2}}$ et non pas celui en $\sqrt{\frac{P_1}{P_2}}$. Les tensions V_P adaptées auraient été $410V$ (contre $520V$ ici), $445V$ (contre $630V$ ici) et $475V$ (contre $780V$ ici) respectivement, en considérant la même marge de sécurité que les trois balayages précédents. Cela explique aussi pourquoi les lobes supérieurs sont plus étendus. A noter également que le design des sondes n'est pas le même, et que les conditions expérimentales ont également été modifiées, notamment la température, car les expérimentations ont été menées avec plusieurs semaines d'écart, et proche de la période estivale.

Dans le cadre des expérimentations précédentes, l'efficacité du système anti-rebonds a pu être observée une nouvelle fois. En effet, deux cartographies de fautes ont été effectuées avec exactement les mêmes paramètres, avec présence ou non du système anti-rebonds. La Figure 3.26 montre le résultat obtenu, avec à gauche la cartographie sans système anti-rebonds et au milieu celle avec (pour la même amplitude d'impulsion de $450V$). On observe bien qu'aucune faute n'est induite dans le circuit lorsque le système anti-rebonds n'est pas ajouté à la sonde, contre deux lobes de fautes distincts lorsque celui-ci est présent. De plus, on remarque également l'élargissement de la zone de fautes lorsque l'amplitude de l'impulsion passe de $450V$ à $500V$.

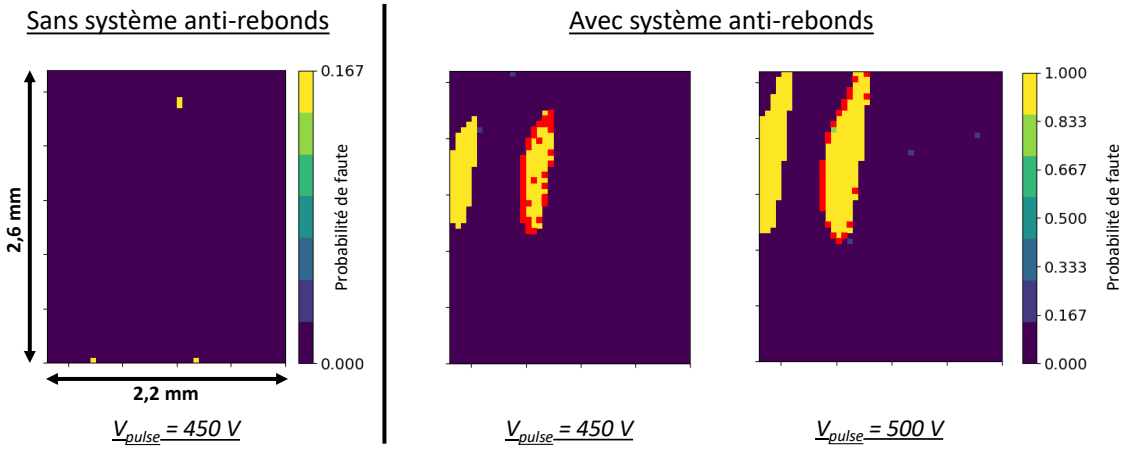


FIGURE 3.26 – Efficacité du système anti-rebonds et élargissement de la zone de faute avec augmentation de V_P

Nous n'avons pour l'instant pas d'explication scientifique concernant cette efficacité accrue, même si l'on suppose que le premier rebond viendrait contrebalancer en grande partie l'effet induit dans le circuit par la première impulsion, du fait de sa polarité opposée (voir Figure 3.16).

Les conséquences d'une diminution des dimensions des sondes d'injection ont été étudiées, théoriquement dans la partie 3.2.3 et expérimentalement dans celle-ci. On aurait pu s'attendre à devoir utiliser des générateurs d'impulsion de plus en plus puissants, car il semblait instinctif de penser que la puissance évolue proportionnellement aux dimensions de la sonde (donc quadratiquement avec sa surface). Les résultats obtenus prouvent qu'il n'en est rien, et que la puissance doit croître proportionnellement à la racine carrée des dimensions de la sonde. C'est une excellente nouvelle qui permettra donc d'améliorer de manière significative la résolution spatiale des plateformes d'EMFI, sans avoir à délivrer aux sondes d'injection des impulsions de plusieurs milliers de volts d'amplitude.

3.7 Conclusion

Ce chapitre a fait l'objet de nombreuses améliorations sur les plateformes d'injection EM. Dans un premier temps, une étude a été menée pour justifier l'intérêt d'une diminution de la taille des sondes de manière à augmenter la résolution spatiale, sans pour autant devoir fournir à la sonde des impulsions de plusieurs milliers de volts d'amplitude. Ensuite, la conception des sondes ainsi que leur fabrication (utilisant les technologies de l'électronique flexible et de l'impression 3D) ont été présentées en ce sens. Suite à cela, un protocole simple de caractérisation des plateformes d'EMFI a été proposé, pour pouvoir comparer notamment la plateforme commercialisée par Langer et la plateforme du LIRMM, qui s'articule autour d'un générateur d'impulsion de la marque Avtech. Cette caractérisation a été suivie du développement et de la fabrication d'un système anti-rebonds, permettant d'accroître significativement la résolution temporelle des plateformes. Chacune des sondes fabriquées à fait l'objet de la caractérisation précédente, à la fois avec et sans système anti-rebonds. Enfin, plusieurs expérimentations ont été menées de manière à confirmer l'étude théorique sur la résolution spatiale énoncée en premier point.

Pour terminer, il est intéressant de regarder quels ont été les progrès du LIRMM sur l'injection de fautes au cours des années. En 2014, trois ans avant le début de ma thèse, une expérimentation avait été menée sur un Spartan3E-1600 décapsulé de chez Xilinx. Elle avait pour but de tracer la probabilité d'injecter une faute, le tout en balayant la totalité du circuit avec une sonde à noyau de ferrite de diamètre $750\mu m$. Le résultat de la cartographie peut être

observé à gauche de la Figure 3.27. On se rend alors compte que la probabilité d'injecter une faute est soit égale à 1, soit nulle, et ce sur la totalité de la puce. Sur la droite de la Figure 3.27 se trouve la cartographie effectuée 7 ans plus tard sur le même circuit, mais avec une sonde nouvelle génération de $100\mu m$ de diamètre, dont le design a été présenté dans le chapitre 2. On y retrouve les zones où la probabilité d'injecter une faute est égale à 1 ou 0, ainsi que certaines zones de "crash" (position où l'injection provoque une non réponse du FPGA) représentées par les pixels rouges. Cependant, l'observation principale et la plus importante qui peut être faite est l'apparition des deux bandes vertes, correspondant à des positions où la probabilité de fautes est exactement égale à 0.5.

Ce résultat intéressant est relativement récent, c'est pourquoi nous n'avons pas d'explications concrètes à présenter pour le moment. Cependant, une imagerie infrarouge de la puce du circuit a été effectuée, et il semblerait que cette probabilité de fautes de 0.5 soit due à la perturbation des blocs RAM interposés avec les CLBs (Configurable Logic Bloc) du FPGA. Des expérimentations futures permettront de confirmer ou non la véracité de cette supposition.

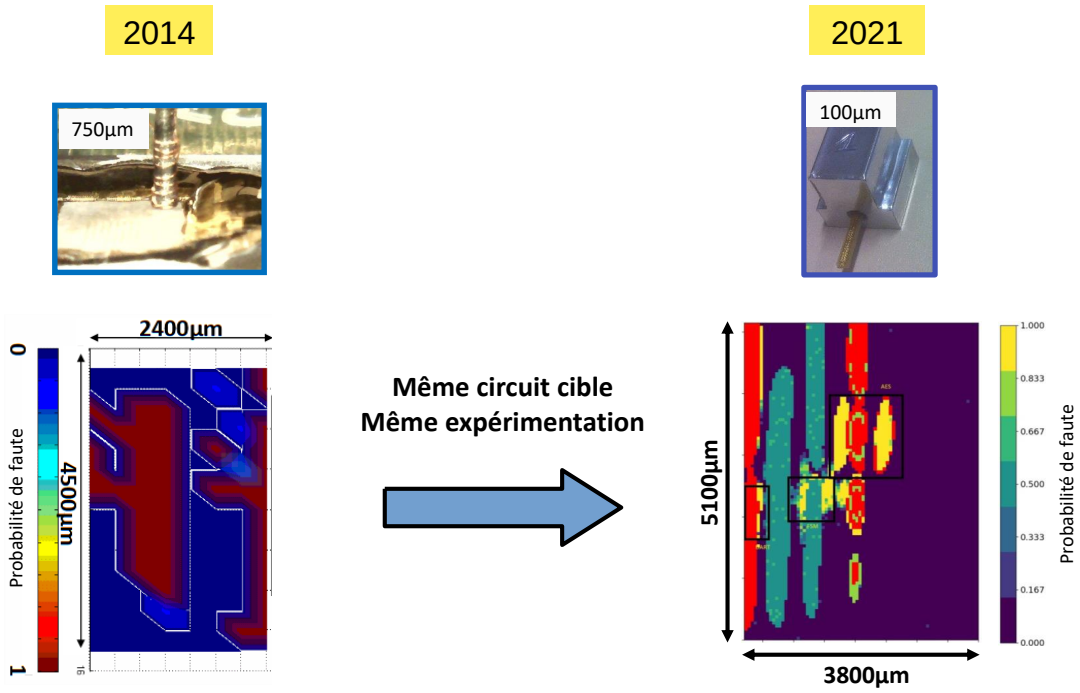


FIGURE 3.27 – Progrès du LIRMM sur l'injection de fautes EM entre 2014 et 2021

Conclusion générale

Conclusion

Ce document a dans un premier temps dressé un état de l'art succinct des différentes attaques pouvant être menées sur des circuits électroniques intégrés. Le premier chapitre s'est concentré sur les attaques par canaux auxiliaires et sur les attaques par injection de fautes. Une attention toute particulière a été portée aux plateformes et moyens techniques nécessaires à la mise en œuvre d'attaques par observation du champ magnétique et par injection électromagnétique. Un modèle de faute spécifique aux injections de fautes par médium électromagnétique a également été présenté au vu de son utilité pour les chapitres suivants.

Le deuxième chapitre a porté sur l'amélioration de la résolution spatiale des sondes d'analyse électromagnétique. Une nouvelle technologie d'électronique flexible a été explorée en ce sens, et l'intégralité du travail de conception et de fabrication a été détaillée. En combinant cette technologie avancée avec l'impression 3D, des prototypes de sondes d'analyse fonctionnels et très prometteurs ont été réalisés. Les premières mesures de SNR à l'aide d'un amplificateur COTS (du fabricant Femto) ont été illustrées et mettent en évidence une résolution spatiale plus importante lorsque les plus petites sondes ($50\mu m$) sont utilisées.

Ce chapitre a également présenté la conception et la caractérisation d'un pré-amplificateur faible bruit destiné à être utilisé au sein de sondes actives. L'objectif est ici de permettre des mesures sans dégradation du SNR de la sonde. Les premières expérimentations avec la puce en boîtier ont donné de bons résultats et un démonstrateur fonctionnel de sonde active est présenté en guise de conclusion de ce chapitre.

Le troisième chapitre a présenté l'intégralité des travaux réalisés quant à l'amélioration des plateformes d'injection de fautes par médium électromagnétique. Ils ont porté essentiellement sur l'amélioration de la résolution spatiale et temporelle, avec comme contrainte principale de conserver le caractère pratique et modulaire de ce type de plateforme. La résolution temporelle a été améliorée grâce à la conception et à la fabrication d'un système anti-rebonds compact. Ce système à base de diode Transil, a prouvé son efficacité à maintes reprises et a été préféré aux systèmes industriels déjà existants sur le marché. Il a permis de supprimer l'intégralité des rebonds sur le signal EM sans atténuer l'amplitude de l'impulsion principale.

Une étude théorique a ensuite été menée et a permis de justifier l'intérêt d'une diminution de la taille des sondes de manière à augmenter la résolution spatiale, et ce, sans avoir à fournir des impulsions d'amplitudes démesurées (plusieurs milliers de volts) à la sonde. Contrairement à ce que l'on aurait pu attendre, celles-ci doivent évoluer selon un facteur racine carrée avec la réduction de la dimension des sondes. C'est une excellente nouvelle qui permet d'éviter une course effrénée vers des générateurs d'impulsion toujours plus puissants. Suite à cette étude, les mêmes technologies d'électronique flexible et d'impression 3D ont été utilisées de manière à améliorer la résolution spatiale des sondes d'injection électromagnétique.

Enfin, afin de caractériser les plateformes et les sondes, un protocole simple de caractérisation a été mis en place. Il se base sur les caractéristiques de l'impulsion générée, et non plus sur les caractéristiques de l'équipement utilisé pour la générer. Ce protocole a permis, avec l'aide d'Eric BOURBAO de la société Thales DIS, de comparer deux plateformes d'EMFI. La première est commercialisée par la société Langer, et la deuxième est celle qui a été développée au LIRMM, qui s'articule autour d'un générateur d'impulsion de la marque Avtech. Les sondes ancienne génération (fabriquées et utilisées au LIRMM avant mon arrivée en thèse) et les sondes nouvelle génération (celles présentées dans ce document) ont également été caractérisées en suivant ce protocole. On a ici pu démonter l'idée initiale de l'étude, qui était que des sondes plus petites pouvaient être utilisées pour effectuer des attaques plus localisées et sélectives.

Pour terminer, les performances du banc d'injection de fautes par médium électromagnétique du LIRMM avant mon arrivée en thèse ont été comparées à celles de 2021 au travers de la même expérimentation. De nouveaux résultats, non expliqués jusqu'ici et donc très intéressants ont été observés avec les sondes nouvelle génération. Des expérimentations supplémentaires sont prévues en vue de les comprendre.

Perspectives à court terme

La fin de ma thèse ne signifie pas pour autant la fin de ma contribution sur l'ensemble des sujets évoqués dans ce document. En effet, le PEA SERGE ne prend fin qu'en Décembre 2023, et de nombreuses expérimentations doivent encore être menées pour compléter les résultats actuels.

La première d'entre elles concerne les sondes elles-mêmes. En effet, dans ce document, et jusqu'à aujourd'hui, la totalité des expérimentations ont été menées avec des sondes horizontales. Or, une partie de mon travail a aussi porté sur les sondes verticales. Il faudra donc fabriquer ces sondes et les soumettre aux mêmes expérimentations de manière à pouvoir en tirer des conclusions.

Ensuite, l'intégralité des résultats en analyse et en injection ont été obtenus sur un FPGA (Spartan 3E-1600). Cependant, nous avons à notre disposition au LIRMM des STM32F439 ouverts face avant et face arrière, intégrant un AES. Il est donc prévu dans un futur proche de changer de cible pour compléter et comparer les résultats sur FPGA et microcontrôleur. Des tests préliminaires ont toutefois permis de vérifier que des sondes de $50\mu m$ permettent effectivement d'injecter des fautes au sein de tels microcontrôleurs, et ce avec des tensions comprises entre 350 et 400V.

Deux autres expérimentations concernant les tensions minimales V_P^{min} à appliquer à la sonde pour obtenir des fautes sont également prévues. La première campagne d'acquisition sera menée en fonction de h, la hauteur effective de la sonde par rapport à la surface du circuit. Cela permettra de se rendre compte si la décapsulation des circuits attaqués est nécessaire à la réussite de l'attaque (avec les sondes présentées dans ce document). La deuxième sera menée sur une période de 48 à 72 heures et mettra en évidence l'influence de la température sur une même injection (mêmes paramètres du

générateur d'impulsion et même position de la sonde).

Enfin, une dernière expérimentation visera à vérifier ou non la véracité de la supposition faite dans la conclusion du troisième chapitre, à savoir que la probabilité de faute de 0.5 serait due à une perturbation des blocs RAM du FPGA Spartan 3E-1600.

Pour finir sur l'objectif initial de ma thèse, nous avons désormais à notre disposition tous les éléments nécessaires à la réalisation de sondes actives pouvant aller jusqu'à un diamètre de $50\mu m$. La fabrication de plusieurs prototypes de tailles différentes est prévue, de manière à pouvoir les comparer avec la sonde Langer équipée de l'amplificateur Femto.

Bibliographie

[Aar13] M. AARTS. « Electromagnetic fault injection using transient pulse injections :a comparison of EM-FI and optical-FI on smart cards ». In : 2013.

[AH20a] Karim M. ABDELLATIF et Olivier HÉRIVEAUX. « SiliconToaster : A Cheap and Programmable EM Injector for Extracting Secrets ». In : *2020 Workshop on Fault Detection and Tolerance in Cryptography (FDTC)*. 2020, p. 35-40. DOI : 10.1109/FDTC51366.2020.00012.

[AH20b] Karim M. ABDELLATIF et Olivier HÉRIVEAUX. « SiliconToaster : A Cheap and Programmable EM Injector for Extracting Secrets ». In : *2020 Workshop on Fault Detection and Tolerance in Cryptography (FDTC)*. 2020, p. 35-40. DOI : 10.1109/FDTC51366.2020.00012.

[AKY06] S. AOYAMA, S. KAWAHITO et M. YAMAGUCHI. « Fully integrated active magnetic probe for high-definition near-field measurement ». In : *2006 IEEE International Symposium on Electromagnetic Compatibility, 2006. EMC 2006.* T. 2. 2006, p. 426-429. DOI : 10.1109/ISEMC.2006.1706340.

[Ala+09] Ali ALAELDINE et al. « Assessment of the Immunity of Unshielded Multi-Core Integrated Circuits to Near-Field Injection ». In : *2009 20th International Zurich Symposium on Electromagnetic Compatibility*. 2009, p. 361-364. DOI : 10.1109/EMCZUR.2009.4783465.

[Anc+17] Stéphanie ANCEAU et al. « Nanofocused X-Ray Beam to Reprogram Secure Circuits ». In : *Cryptographic Hardware and Embedded Systems – CHES 2017*. T. 10529. Lecture Notes in Computer Science. Springer, 2017, p. 175-188. DOI : 10.1007/978-3-319-66787-4_9.

[BAM17a] Josep BALASCH, Daniel ARUMÍ et Salvador MANICH. « Design and validation of a platform for electromagnetic fault injection ». In : *2017 32nd Conference on Design of Circuits and Integrated Systems (DCIS)*. 2017, p. 1-6. DOI : 10.1109/DCIS.2017.8311630.

[BAM17b] Josep BALASCH, Daniel ARUMÍ et Salvador MANICH. « Design and validation of a platform for electromagnetic fault injection ». In : *2017 32nd Conference on Design of Circuits and Integrated Systems (DCIS)*. 2017, p. 1-6. DOI : 10.1109/DCIS.2017.8311630.

[Bar+09] Alessandro BARENGHI et al. « Low Voltage Fault Attacks on the RSA Cryptosystem ». In : *2009 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*. 2009, p. 23-31. DOI : 10.1109/FDTC.2009.30.

[Bar+10] Alessandro BARENGHI et al. « Low voltage fault attacks to AES ». In : *2010 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*. 2010, p. 7-12. DOI : 10.1109/HST.2010.5513121.

[Bas+13] R. Possamai BASTOS et al. « A bulk built-in sensor for detection of fault attacks ». In : *2013 IEEE International Symposium on Hardware-Oriented Security and Trust (HOST)*. 2013, p. 51-54. DOI : 10.1109/HST.2013.6581565.

[Bay+12] Pierre BAYON et al. « Contactless Electromagnetic Active Attack on Ring Oscillator Based True Random Number Generator ». In : *Constructive Side-Channel Analysis and Secure Design*. Sous la dir. de Werner SCHINDLER et Sorin A. HUSS. Berlin, Heidelberg : Springer Berlin Heidelberg, 2012, p. 151-166. ISBN : 978-3-642-29912-4.

[BCO04] Eric BRIER, Christophe CLAVIER et Francis OLIVIER. « Correlation Power Analysis with a Leakage Model ». In : *Cryptographic Hardware and Embedded Systems - CHES 2004*. Sous la dir. de Marc JOYE et Jean-Jacques QUISQUATER. Berlin, Heidelberg : Springer Berlin Heidelberg, 2004, p. 16-29. ISBN : 978-3-540-28632-5.

[BK06] Johannes BLÖMER et Volker KRUMMEL. « Fault Based Collision Attacks on AES ». In : *Fault Diagnosis and Tolerance in Cryptography*. Sous la dir. de Luca BREVEGLIERI et al. Berlin, Heidelberg : Springer Berlin Heidelberg, 2006, p. 106-120. ISBN : 978-3-540-46251-4.

[Bly+93] S. BLYTHE et al. « Layout reconstruction of complex silicon chips ». In : *IEEE Journal of Solid-State Circuits* 28.2 (1993), p. 138-145. ISSN : 1558-173X. DOI : 10.1109/4.192045.

[BS97] Eli BIHAM et Adi SHAMIR. « Differential fault analysis of secret key cryptosystems ». In : *Advances in Cryptology — CRYPTO '97*. Sous la dir. de Burton S. KALISKI. Berlin, Heidelberg : Springer Berlin Heidelberg, 1997, p. 513-525. ISBN : 978-3-540-69528-8.

[Buk+18] Sebanjila K. BUKASA et al. « Let's Shock Our IoT's Heart : ARMv7-M under (Fault) Attacks ». In : 2018. ISBN : 9781450364485.

[Car+04] Vincent CARLIER et al. « Electromagnetic Side Channels of an FPGA Implementation of AES. » In : *IACR Cryptology ePrint Archive* 2004 (jan. 2004), p. 145.

[CH17] Ang CUI et Rick HOUSLEY. « BADFET : Defeating Modern Secure Boot Using Second-Order Pulsed Electromagnetic Fault Injection ». In : *11th USENIX Workshop on Offensive Technologies (WOOT 17)*. Vancouver, BC : USENIX Association, août 2017. URL : https://www.usenix.org/conference/woot17/workshop-program/presentation/cui.

[Cor99] Jean-Sébastien CORON. « Resistance against Differential Power Analysis for Elliptic Curve Cryptosystems ». In : *Cryptographic Hardware and Embedded Systems, First International Workshop, CHES 99, Worcester, MA, USA, August 12-13, 1999, Proceedings*. T. 1717. Lecture Notes in Computer Science. Springer, 1999, p. 292-302. DOI : 10.1007/3-540-48059-5_25.

[CT05] Hamid CHOUKRI et Michael TUNSTALL. « Round reduction using faults ». In : (jan. 2005), p. 13-24.

[DCC11] Zhi-bo DU, Yun CHEN et Ai-dong CHEN. « The Impact of the Clock Frequency on the Power Analysis Attacks ». In : *2011 International Conference on Internet Technology and Applications*. 2011, p. 1-4. DOI : 10.1109/ITAP.2011.6006291.

[Deh+12a] A DEHBAOUI et al. « Injection of transient faults using electromagnetic pulses Practical results on a cryptographic system ». In : (jan. 2012).

[Deh+12b] Amine DEHBAOUI et al. « Electromagnetic Transient Faults Injection on a Hardware and a Software Implementations of AES ». In : *2012 Workshop on Fault Diagnosis and Tolerance in Cryptography*. 2012, p. 7-15. DOI : 10.1109/FDTC.2012.15.

[Deh+13] Amine DEHBAOUI et al. « Electromagnetic Glitch on the AES Round Counter ». In : *Constructive Side-Channel Analysis and Secure Design*. Sous la dir. d'Emmanuel PROUFF. Berlin, Heidelberg : Springer Berlin Heidelberg, 2013, p. 17-31. ISBN : 978-3-642-40026-1.

[DLM19] Mathieu DUMONT, Mathieu LISART et Philippe MAURINE. « Electromagnetic Fault Injection : How Faults Occur ». In : *2019 Workshop on Fault Diagnosis and Tolerance in Cryptography, FDTC 2019, Atlanta, GA, USA, August 24, 2019*. IEEE, 2019, p. 9-16.

[DLM21] Mathieu DUMONT, Mathieu LISART et Philippe MAURINE. « Modeling and Simulating Electromagnetic Fault Injection ». In : *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* 40.4 (2021), p. 680-693.

[DM18] Gilles R. DUCHARME et Philippe MAURINE. « Estimating the Signal-to-Noise Ratio Under Repeated Sampling of the Same Centered Signal : Applications to Side-Channel Attacks on a Cryptoprocessor ». In : *IEEE Transactions on Information Theory* 64.9 (2018), p. 6333-6339. DOI : 10.1109/TIT.2018.2851217.

[Dut+12] Jean-Max DUTERTRE et al. « Fault Round Modification Analysis of the advanced encryption standard ». In : *2012 IEEE International Symposium on Hardware-Oriented Security and Trust*. 2012, p. 140-145. DOI : 10.1109/HST.2012.6224334.

[Gie+08] Benedikt GIERLICHS et al. « Mutual Information Analysis ». In : *Cryptographic Hardware and Embedded Systems – CHES 2008*. Sous la dir. d'Elisabeth OSWALD et Pankaj ROHATGI. Berlin, Heidelberg : Springer Berlin Heidelberg, 2008, p. 426-442. ISBN : 978-3-540-85053-3.

[Gir05] Christophe GIRAUD. « DFA on AES ». In : *Advanced Encryption Standard – AES*. Sous la dir. d'Hans DOBBERTIN, Vincent RIJMEN et Aleksandra SOWA. Berlin, Heidelberg : Springer Berlin Heidelberg, 2005, p. 27-41. ISBN : 978-3-540-31840-8.

[GMO01] Karine GANDOLFI, Christophe MOURTEL et Francis OLIVIER. « Electromagnetic Analysis : Concrete Results ». In : *Cryptographic Hardware and Embedded Systems — CHES 2001*. Sous la dir. de Çetin K. KOÇ, David NACCACHE et Christof PAAR. Berlin, Heidelberg : Springer Berlin Heidelberg, 2001, p. 251-261. ISBN : 978-3-540-44709-2.

[Hut] Michael HUTTER. *The Temperature Side Channel and Heating Fault Attacks*.

[KJJ99] Paul KOCHER, Joshua JAFFE et Benjamin JUN. « Differential Power Analysis ». In : *Advances in Cryptology — CRYPTO' 99*. Sous la dir. de Michael WIENER. Berlin, Heidelberg : Springer Berlin Heidelberg, 1999, p. 388-397. ISBN : 978-3-540-48405-9.

[Koc96] Paul C. KOCHER. « Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and Other Systems ». In : *Advances in Cryptology — CRYPTO '96*. Sous la dir. de Neal KOBLITZ. Berlin, Heidelberg : Springer Berlin Heidelberg, 1996, p. 104-113. ISBN : 978-3-540-68697-2. DOI : 10.1007/3-540-68697-5_9.

[Li+10] Yang LI et al. « Fault Sensitivity Analysis ». In : *Cryptographic Hardware and Embedded Systems, CHES 2010*. Sous la dir. de Stefan MANGARD et François-Xavier STANDAERT. Berlin, Heidelberg : Springer Berlin Heidelberg, 2010, p. 320-334. ISBN : 978-3-642-15031-9.

[Maj18] Fabien MAJERIC. « Etude d'attaques matérielles et combinées sur les "System-on-chip" ». Thèse de doct. Nov. 2018.

[Mau12] Philippe MAURINE. « Techniques for EM Fault Injection : Equipments and Experimental Results ». In : *2012 Workshop on Fault Diagnosis and Tolerance in Cryptography*. 2012, p. 3-4. DOI : 10.1109/FDTC.2012.21.

[Mau+12] Philippe MAURINE et al. « Yet Another Fault Injection Technique : by Forward Body Biasing Injection ». In : *YACC'2012 : Yet Another Conference on Cryptography*. Porquerolles Island, France, sept. 2012. URL : https://hal-lirmm.ccsd.cnrs.fr/lirmm-00762035.

[MBB16] F. MAJÉRIC, E. BOURBAO et L. BOSSUET. « Electromagnetic security tests for SoC ». In : *2016 IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. 2016, p. 265-268. DOI : 10.1109/ICECS.2016.7841183.

[MK+12] Nguyen Ngoc MAI-KHANH et al. « An integrated high-precision probe system for near-field magnetic measurements on cryptographic LSIs ». In : *SENSORS, 2012 IEEE*. 2012, p. 1-4. DOI : 10.1109/ICSENS.2012.6411173.

[MK+15] Nguyen Ngoc MAI-KHANH et al. « A Near-Field Magnetic Sensing System With High-Spatial Resolution and Application for Security of Cryptographic LSIs ». In : *IEEE Transactions on Instrumentation and Measurement* 64.4 (2015), p. 840-848. ISSN : 1557-9662. DOI : 10.1109/TIM.2014.2373472.

[MK+18] Nguyen Ngoc MAI-KHANH et al. « Noninvasive Localization of IGBT Faults by High-Sensitivity Magnetic Probe With RF Stimulation ». In : *IEEE Transactions on Instrumentation and Measurement* 67.4 (2018), p. 745-753. ISSN : 1557-9662. DOI : 10.1109/TIM.2017.2789038.

[O'F19] Colin O'FLYNN. « MIn()Imum Failure : EMFI Attacks against USB Stacks ». In : *Proceedings of the 13th USENIX Conference on Offensive Technologies*. WOOT'19. Santa Clara, CA, USA : USENIX Association, 2019, p. 15.

[O'F20] Colin O'FLYNN. *BAM BAM!! On Reliability of EMFI for in-situ Automotive ECU Attacks*. Cryptology ePrint Archive, Report 2020/937. 2020.

[OGM15] Sébastien ORDAS, Ludovic GUILLAUME-SAGE et Philippe MAURINE. « EM Injection : Fault Model and Locality ». In : *2015 Workshop on Fault Diagnosis and Tolerance in Cryptography, FDTC 2015, Saint Malo, France, September 13, 2015*. Sous la dir. de Naofumi HOMMA et Victor LOMNÉ. IEEE Computer Society, 2015, p. 3-13.

[OGM17] Sébastien ORDAS, Ludovic GUILLAUME-SAGE et Philippe MAURINE. « Electromagnetic fault injection : the curse of flip-flops ». In : *J. Cryptogr. Eng.* 7.3 (2017), p. 183-197.

[Oma+13] R. OMAROUAYACHE et al. « Magnetic microprobe design for EM fault attack ». In : *2013 International Symposium on Electromagnetic Compatibility*. 2013, p. 949-954.

[Ord+09] Thomas ORDAS et al. « Near-Field Mapping System to Scan in Time Domain the Magnetic Emissions of Integrated Circuits ». In : *Integrated Circuit and System Design. Power and Timing Modeling, Optimization and Simulation*. Sous la dir. de Lars SVENSSON et José MONTEIRO. Berlin, Heidelberg : Springer Berlin Heidelberg, 2009, p. 229-236. ISBN : 978-3-540-95948-9. DOI : 10.1007/978-3-540-95948-9_23.

[Ord+15] S. ORDAS et al. « Evidence of a Larger EM-Induced Fault Model ». In : *Smart Card Research and Advanced Applications*. Sous la dir. de Marc JOYE et Amir MORADI. Cham : Springer International Publishing, 2015, p. 245-259. ISBN : 978-3-319-16763-3.

[Pou+11] F. POUCHERET et al. « Local and Direct EM Injection of Power Into CMOS Integrated Circuits ». In : *2011 Workshop on Fault Diagnosis and Tolerance in Cryptography*. 2011, p. 100-104. DOI : 10.1109/FDTC.2011.18.

[QS02] Jean-Jacques QUISQUATER et D. SAMYDE. « Eddy current for magnetic analysis with active sensor ». In : (jan. 2002).

[SA03] Sergei P. SKOROBOGATOV et Ross J. ANDERSON. « Optical Fault Induction Attacks ». In : *Cryptographic Hardware and Embedded Systems - CHES 2002*. Sous la dir. de Burton S. KALISKI, çetin K. KOÇ et Christof PAAR. Berlin, Heidelberg : Springer Berlin Heidelberg, 2003, p. 2-12. ISBN : 978-3-540-36400-9.

[Sam+02] D. SAMYDE et al. « On a new way to read data from memory ». In : *First International IEEE Security in Storage Workshop, 2002. Proceedings.* 2002, p. 65-69. DOI : 10.1109/SISW.2002.1183512.

[Sch+12] Alexander SCHLÖSSER et al. « Simple Photonic Emission Analysis of AES ». In : *Cryptographic Hardware and Embedded Systems – CHES 2012*. Sous la dir. d'Emmanuel PROUFF et Patrick SCHAUMONT. Berlin, Heidelberg : Springer Berlin Heidelberg, 2012, p. 41-57. ISBN : 978-3-642-33027-8.

[SGD08] Nidhal SELMANE, Sylvain GUILLEY et Jean-Luc DANGER. « Practical Setup Time Violation Attacks on AES ». In : *2008 Seventh European Dependable Computing Conference*. 2008, p. 91-96. DOI : 10.1109/EDCC-7.2008.11.

[SH07] Jörn-Marc SCHMIDT et Michael HUTTER. « Optical and EM Fault-Attacks on CRT-based RSA : Concrete Results ». English. In : *Austrochip 2007, 15th Austrian Workhop on Microelectronics, 11 October 2007, Graz, Austria, Proceedings*. Austrochip 2007 ; Conference date : 11-10-2007 Through 11-10-2007. Verlag der Technischen Universität Graz, 2007, p. 61-67. ISBN : 978-3-902465-87-0.

[Sou+11] Mathilde SOUCARROS et al. « Influence of the temperature on true random number generators ». In : *2011 IEEE International*

Symposium on Hardware-Oriented Security and Trust. 2011, p. 24-27. DOI : 10.1109/HST.2011.5954990.

[Tes+07] G. TESSIER et al. « Back side thermal imaging of integrated circuits at high spatial resolution ». In : *Applied Physics Letters* 90.17, 171112 (avr. 2007), p. 171112. DOI : 10.1063/1.2732179.

[Tho91] M. THOMPSON. « Inductance Calculation Techniques—Part II : Approximations and Handbook Methods ». In : 1991.

[Tob+13] K. TOBICH et al. « Voltage Spikes on the Substrate to Obtain Timing Faults ». In : *2013 Euromicro Conference on Digital System Design*. 2013, p. 483-486. DOI : 10.1109/DSD.2013.146.

[Tou+20] J. TOULEMONT et al. *A Simple Protocol to Compare EMFI Platforms*. Cryptology ePrint Archive, Report 2020/1277. https://ia.cr/2020/1277. 2020.

[Tou+21a] J. TOULEMONT et al. « Exploring flexible and 3D printing technologies for the design of high spatial resolution EM probes ». In : *2021 19th IEEE International New Circuits and Systems Conference (NEWCAS)*. 2021, p. 1-4. DOI : 10.1109/NEWCAS50681.2021.9462763.

[Tou+21b] J. TOULEMONT et al. « On the scaling of EMFI probes ». In : *2021 Workshop on Fault Detection and Tolerance in Cryptography (FDTC)*. 2021, p. 67-73. DOI : 10.1109/FDTC53659.2021.00019.

[Tro+19] Thomas TROUCHKINE et al. « Electromagnetic fault injection against a System-on-Chip, toward new micro-architectural fault models ». In : *CoRR* abs/1910.11566 (2019).

[TSD19] Oualid TRABELSI, Laurent SAUVAGE et Jean-Luc DANGER. « Impact of Intentional Electromagnetic Interference on Pure Combinational Logic ». In : *2019 International Symposium on Electromagnetic Compatibility - EMC EUROPE*. Barcelone, Spain, 2019, p. 398-403. URL : https://hal.telecom-paris.fr/hal-02318731.

[Vas+17] Aurélien VASSELLE et al. « Laser-Induced Fault Injection on Smartphone Bypassing the Secure Boot ». In : *2017 Workshop on Fault Diagnosis and Tolerance in Cryptography (FDTC)*. 2017, p. 41-48. DOI : 10.1109/FDTC.2017.18.

[VMC19] Aurélien VASSELLE, P. MAURINE et Maxime COZZI. « Breaking Mobile Firmware Encryption through Near-Field Side-Channel Analysis ». In : *Proceedings of the 3rd ACM Workshop on Attacks and Solutions in Hardware Security Workshop* (2019).

[WT08] Knut WOLD et Chik How TAN. « Analysis and Enhancement of Random Number Generator in FPGA Based on Oscillator Rings ». In : *2008 International Conference on Reconfigurable Computing and FPGAs*. 2008, p. 385-390. DOI : 10.1109/ReConFig.2008.17.

[WWM11] Jasper G.J. van WOUDENBERG, Marc F. WITTEMAN et Federico MENARINI. « Practical Optical Fault Injection on Secure Microcontrollers ». In : *2011 Workshop on Fault Diagnosis and Tolerance in Cryptography*. 2011, p. 91-99. DOI : 10.1109/FDTC.2011.12.

[YJ00] Sung-Ming YEN et M. JOYE. « Checking before output may not be enough against fault-based cryptanalysis ». In : *IEEE Transactions on Computers* 49.9 (2000), p. 967-970. ISSN : 1557-9956. DOI : 10.1109/12.869328.

[YPA08] Asier Goikoetxea YANCI, Stephen PICKLES et Tughrul ARSLAN. « Detecting Voltage Glitch Attacks on Secure Devices ». In : *2008 Bio-inspired, Learning and Intelligent Systems for Security*. 2008, p. 75-80. DOI : 10.1109/BLISS.2008.26.

Abstract

Nowadays, most of the sectors in the industry rely on the use of electronic circuits and systems. Among these systems and integrated circuits, there are circuits dedicated to security, such as smart cards. The data they handle are very valuable and can be desired. Therefore, it is important to ensure the reliability and the security of those systems against malicious attackers. That is why it is necessary to adapt the security characterization tools, especially considering the constant evolution of technologies and threats. In this context, my PhD focused on the development of platforms dedicated to electromagnetic attacks, through : The enhancement of the spatial resolution of electromagnetic analysis probes thanks to an advanced flexible electronic technology paired with 3D printing, but also thanks to the design of a low noise amplifier in CMOS $0.35\mu m$ technology ; The enhancement of both timing and spatial resolutions of electromagnetic fault injection platforms, by exploiting the same technologies, but also thanks to the development of simple electronic modules. All those modules, allowing the improvement of the platforms, were developed, validated and characterized. Their exploitability for hardware cryptanalysis purposes was also experimentally demonstrated.

Résumé

De nos jours, la plupart des secteurs d'activité reposent sur l'utilisation de circuits et systèmes électroniques. Parmi eux, on trouve des circuits dédiés à la sécurité comme les cartes à puce. Les données qu'elles manipulent ont une valeur importante et peuvent faire l'objet de convoitise. Il est donc important d'assurer la fiabilité et la sécurité de ces systèmes contre des personnes malintentionnées. C'est pourquoi il est nécessaire d'adapter les outils de caractérisation sécuritaire pour tenir compte de l'évolution des technologies et des menaces. Dans ce contexte, ma thèse s'est concentrée sur le développement de plateformes dédiées aux attaques exploitant le médium électromagnétique, au travers de : L'amélioration de la résolution spatiale des sondes d'analyse électromagnétique, et ce, grâce à l'exploitation de technologies d'impression 3D, d'électronique flexible mais également grâce au développement d'un amplificateur très faible bruit en technologie CMOS $0.35\mu m$; L'amélioration des résolutions spatiale et temporelle des plateformes d'injection de fautes par médium électromagnétique, là encore, en exploitant les mêmes technologies, mais également grâce au développement de modules électroniques simples. L'ensemble des dispositifs permettant d'améliorer ces plateformes a été développé, validé et caractérisé. Leur utilisation à des fins de cryptanalyse matérielle a également été démontrée expérimentalement.

LIRMM - 161 rue Ada 34095 Montpellier Cedex 5 - France