Characterization of the systematic effects on the Planck/HFI instrument, propagation and impact on science data

Alexandre Sauvé

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *5 Décembre 2016* **par :**

ALEXANDRE SAUVÉ

Caractérisation des effets systématiques de l'instrument *Planck/HFI*, propagation et impact sur les données scientifiques

---

JURY

PETER VON BALLMOOS	Professeur d'Université	Président du Jury
ANTHONY BANDAY	Directeur de Recherche	Co-directeur de thèse
FRANÇOIS COUCHOT	Directeur de Recherche	Membre du Jury
KEITH GRAINGE	Professeur d'Université	Rapporteur
LUDOVIC MONTIER	Ingénieur de Recherche	Directeur de thèse
MICHEL PIAT	Maître de Conférences	Membre du Jury
ISABELLE RISTORCELLI	Chargée de Recherche	Membre du Jury
LOUIS RODRIGUEZ	Ingénieur	Rapporteur

École doctorale et spécialité :
SDU2E : Astrophysique, Sciences de l'Espace, Planétologie
Unité de Recherche :
IRAP / CNRS (UMR 5277)
Directeur(s) de Thèse :
Ludovic MONTIER et *Anthony BANDAY*
Rapporteurs :
Louis RODRIGUEZ et *Keith GRAINGE*

Declaration of Authorship

I, Alexandre Sauvé, declare that this thesis titled, "Caractérisation des effets systématiques de l'instrument *Planck/HFI*, propagation et impact sur les données scientifiques" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Résumé

Université Paul Sabatier
Sciences de l'Univers de l'Environnement et de l'Espace

Docteur en philosophie

Caractérisation des effets systématiques de l'instrument *Planck/HFI*, propagation et impact sur les données scientifiques

par Alexandre Sauvé

Planck est un satellite de l'ESA lancé en 2009, il représente la troisième génération d'observatoires spatiaux dans l'ère de la cosmologie de précision. Sa mission était de cartographier le fond diffus cosmologique (CMB) avec une extrême précision ($\Delta T/T < 10^{-5}$). Le niveau de précision requis nécessite un niveau de contrôle extrêmement élevé des effets systématiques introduits par les instruments embarqués. Il se trouve qu'une combinaison inattendue d'éléments a conduit à une amplification de l'impact des effets de non linéarité introduits par le composant de numérisation des données scientifiques de l'instrument *HFI*, conduisant à l'introduction d'un des effets systématiques les plus difficiles à maîtriser. Ce manuscrit présente le travail qui a conduit à la caractérisation et à la correction de cette non-linéarité.

Tout d'abord la modélisation de la réponse thermique complexe des détecteurs bolométriques sous intensité modulée est présentée. Puis la caractérisation de la réponse des détecteurs et de l'électronique de lecture est réalisée via l'utilisation du signal produit par l'impact des particules cosmiques sur les détecteurs. Dans un deuxième temps, les étapes menant à la caractérisation de précision des composants de numérisation sont détaillée. Pour pouvoir corriger l'effet de non-linéarité sur les données scientifiques, la chaîne complète de l'électronique de lecture est modélisée en prenant en compte la réponse des détecteurs sous intensité modulée et les effets non-linéaires du composant de numérisation avec le bruit. En plus de cela, il a fallu tenir compte du signal parasite complexe généré par le compresseur de l'étage de refroidissement cryogénique à 4 K, son inclusion dans la correction est détaillée. Ceci a mené à une correction qui a réduit l'impact de cet effet d'un ordre de grandeur sur le catalogue des données Planck de 2015. Finalement une étude est menée pour mesurer l'impact de cet effet de non-linéarité sur l'analyse scientifique de la Galaxie en polarisation et sur la cosmologie à travers le spectre de puissance angulaire du CMB. Une étude préliminaire est menée pour la détection de l'annihilation de nuages d'antimatière survivant à l'époque de la recombinaison en utilisant la carte d'effet Sunyaev-Zeldovich de *Planck*, et comment cette détection est affectée par la stratégie de scan de *Planck*.

Abstract

Université Paul Sabatier
Sciences de l'Univers de l'Environnement et de l'Espace

Doctor of Philosophy

Caractérisation des effets systématiques de l'instrument *Planck/HFI*, propagation et impact sur les données scientifiques

by Alexandre Sauvé

Planck is an ESA spacecraft launched in 2009, it is the third generation of spatial observatory in the era of precision cosmology. Its mission goal was to map with an exquisite precision ($\Delta T/T < 10^{-5}$) the Cosmic Microwave Background (CMB). The required level of accuracy needs an unusually high level control of the systematic effects introduced by the on-board instruments. However, an unexpected conjunction of elements has enhanced the nonlinearity introduced by the chip performing the digitization of science data by the *HFI* instrument. It resulted in the most challenging systematic effect to deal with. It is presented here the work performed to characterize and correct for this nonlinear effect.

First a detailed modeling of the complex thermal response of the bolometer detectors under AC biasing is presented. The detector response is then further characterized by using the signal produced by cosmic particles. Second, the steps leading to the in-flight accurate characterization of the digitization chip are detailed. To correct for the nonlinear effect on science data, the full electronics readout chain response is modeled, taking into account the detector response under AC biasing and the nonlinear effect of the digitization chip with the noise. Furthermore, the complex parasitic signal originating from the 4 K cryogenic stage mechanical cooler has also to be taken into account. The provided correction has been applied with success to the *HFI* data 2015 release reducing the effect by an order of magnitude. Finally it is studied how the nonlinearity effect of the digitization component affects the galactic and cosmological scientific analysis through the angular power spectrum of the CMB. A preliminary study is performed for the detection of surviving antimatter clouds annihilating with matter at the epoch of recombination with the *Planck* map of Sunyaev-Zeldovich effect, and how it is affected by the scanning strategy of the spacecraft.

Acknowledgements

Ce travaille de thèse s'est déroulé sur plus de trois ans, et avant toute autre chose je dois remercier mon fils qui du haut de ses sept ans a été extrêmement bienveillant à mon égard malgrès les très nombreuses heures que j'ai passé devant un écran au lieu de jouer avec lui.

Cette thèse est co-financée par le *CNES* et la société *Noveltis*, je tiens à les remercier et plus particulièrement Olivier Lamarle ainsi que Philippe Bru d'avoir accepté de me prendre sous leur aile.

Je remercie mon directeur de thèse Ludovic Montier de m'avoir offert l'opportunité de réaliser cette thèse ainsi que son énergie indéfectible pour débloquer tout les problèmes qui ont pu se présentér sur le chemin. Il a souffert avec beaucoup d'abnégation pour la relecture des deux articles que j'ai rédigé durant ces trois annés ainsi que ce manuscrit avec mon anglais approximatif qui n'aime pas les "s". Mon co-directeur de thèse Anthony Banday a également été d'un grand soutien et a toujours su fournir des réponses précises et des conseils pertinents chaque fois que j'en avais besoin.

J'ai une pensée particulière pour François Couchot qui a porté à bout de bras l'équipe ADC pendant plusieurs années avec sa nonchalance et sa bonne humeur qui lui vont si bien. Il m'a donné de nombreux conseils et appris énormément, il a notamment porté à ma connaissance la méthode des Lagrange Multipliers qui a débouché sur la correction d'ADC utilisé pour les données *Planck* 2015. Mais surtout il m'a fait découvrir le P'tit Cahoua du Bd St-Marcel, ce qui n'a pas de prix!

Dans le travail sur les ADC j'ai également eu beaucoup de plaisir à échanger avec Guillaume Patanchon, les discussions qu'on a eu étaient très animées et constructives.

Je tiens également à remercier toute l'équipe du café matinal (et bénévole) de l'IRAP sans qui rien de tout ceci n'aurait été possible!

The development of *Planck* has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN and JA (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). A description of the *Planck* Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at `http://www.rssd.esa.int/index.php?project=PLANCK&page=PlanckCollaboration`.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC	**A**nalog to **D**igital **C**onverter
ADU	**A**nalog **D**igital **U**nit
CMB	**C**osmic **M**icrowave **B**ackground
CPV	**C**alibration **P**erformance **V**alidation
CTP	**C**ode **T**ransition **P**oint
DAC	**D**igital to **A**nalog **C**onverter
DFT	**D**iscrete **F**ourier **T**ransform
DNL	**D**ifferential **N**on **L**inearity
DPC	**D**ata **P**rocessing **C**enter
FPU	**F**ocal **P**lane **U**nit
HFI	**H**igh **F**requency **I**nstrument
INL	**I**ntegral **N**on **L**inearity
LFI	**L**ow **F**requency **I**nstrument
LSB	**L**east **S**ignificant **B**it
PAU	**P**re **A**mplifier **U**nit
PLA	**P**lanck **L**egacy **A**rchive
PSB	**P**olarization **S**ensitive **B**olometer
PDF	**P**robability **D**ensity **F**unction
REU	**R**eadout **E**lectronic **U**nit
RMS	**R**oot **M**ean **S**quare
SWB	**S**pider **W**eb **B**olometer
SZ	**S**unyaev **Z**eldovich (effect)
TF	**T**ransfer **F**unction

List of Symbols

P	power	W ($\mathrm{J\,s^{-1}}$)
ω	angular frequency	rad
$\equiv$	equivalent to	binary operator
$\simeq$	asymptotically equivalent	binary operator
$\approx$	is approximately	binary operator
$\propto$	is proportional to	binary operator
$\langle x \rangle$	average value of x	unary operator

Introduction générale

Pour bien définir le fond du sujet qui nous intéresse commençons par un tour d'horizon de la théorie sur la formation de notre Univers qui décrit le mieux les observations. Selon la théorie de *l'inflation*, l'Univers a connu une période d'expansion très rapide après l'instant initial du Big Bang pendant ses premières 10^{-32} s. Dans les instants qui ont suivi, la matière et l'anti-matière se sont formées, leur annihilation réciproque a donné naissance à la la majeure partie des photons du fond diffus cosmologique (CMB), mais une légère asymétrie dans le rapport des deux en faveur de la matière a fait que cette dernière domine dans l'Univers moderne. Environ 380 000 ans après l'instant initial, l'expansion qui a suivi l'inflation a suffisamment refroidis le plasma dense et chaud de matière baryonique et de photons pour permettre la recombinaison des protons avec les électrons libres et former des atomes d'hydrogène majoritairement. C'est à cette époque qu'on appelle *la recombinaison* que le CMB est émis, ses photons n'étant plus soumis à la diffusion Compton sur les électrons libres ont pu se propager jusqu'à nous en ayant un minimum d'interactions avec la matière. La période qui a suivi s'appelle l'age de l'obscurité. C'est pendant ce temps que la matière a commencé à s'effondrer sur elle même à cause de l'instabilité gravitationnelle et que les premières grandes structures de l'Univers se sont formées. Par la suite la première lumière émise par les galaxies naissantes et la quasars a de nouveau ionisé l'hydrogène, c'est l'époque de *la réionisation*, ce qui a eu pour effet de diffuser à nouveau une fraction des photons du CMB à grande échelle angulaire sur les électrons de nouveau libres.

Le CMB se présente de nos jours comme une émission uniforme sur le ciel avec un spectre de corps noir à la température de 2.7255 K. Si on fait abstraction du *dipôle solaire* dû à l'effet Doppler de notre vitesse relative dans le référentiel du CMB, apparaissent de petites variations de températures de l'ordre de $\Delta T = 1 \times 10^{-5}$ K, qu'on appelle les *anisotropies*. Ces anisotropies qu'on qualifie de *primaires* sont principalement dues aux oscillations acoustiques adiabatiques produites par la compétition entre la gravité et la pression des photons à l'époque où le CMB a été émis. Elles présentent une distribution caractéristique dans leurs amplitudes et leurs échelles angulaires qu'on mesure avec le *spectre de puissance angulaire*. Ce spectre de puissance angulaire est très bien décrit avec les six paramètres du modèle cosmologique ΛCDM (Cold Dark Matter) qui fournit une estimation précise de la densité de matière et d'énergie dans l'Univers mais nécessite l'introduction d'énergie sombre et de matière sombre n'interagissant pas avec la matière ordinaire. Une petite fraction des photons du CMB sont polarisés par deux types de perturbations, les premières sont scalaires et proviennent des variations de densité de matière dues aux oscillations acoustiques, ce qui produit de la polarisation avec une orientation radiale ou circulaire qu'on appelle les *modes-E* en polarisation. Les secondes sont dues aux ondes gravitationnelles générées pendant l'inflation. Elles se manifestent également par des modes-E ainsi que par de la polarisation avec une orientation tourbillonnante qu'on appelle les *modes-B* en polarisation. Il y a une autre catégorie d'anisotropies qu'on qualifie de *secondaires* et qui sont dues aux interactions subséquentes à l'émission du CMB et qui sont principalement : l'effet de lentille gravitationnelle généré par les amas de galaxies, l'effet Sunyaev-Zeldovich (SZ) lié à la diffusion Compton inverse des photons du CMB sur le gaz d'électrons chauds des amas de galaxies, et aussi la diffusion Thompson de ces mêmes photons par les électrons libres dus à la réionisation.

La première détection accidentelle du CMB a été effectuée en 1964 par les radio astronomes Arno Penzias et Robert Woodrow Wilson avec une antenne à cornet de 6 m de large.

La radiation observée prise d'abord pour un biais était isotrope sur tout le ciel et ne présentait pas de fluctuations diurnes ou saisonnières. Puis le satellite COBE lancé en 1989 a mesuré avec l'instrument FIRAS que le CMB avait un spectre qui correspondait parfaitement à celui d'un corps noir à une température de 2.726 K. L'instrument DMR à son bord a également effectué la première observation des anisotropies et a trouvé qu'elles avaient une distribution Gaussienne, mais il n'avait cependant pas la résolution angulaire suffisante pour résoudre le premier pic acoustique du spectre de puissance angulaire. Son successeur, le satellite WMAP a véritablement marqué le début de l'ère de la cosmologie de précision en réduisant de façon remarquable le nombre de paramètres cosmologiques à six par l'intermédiaire au modèle ΛCDM qui reproduisait avec une très bonne précision le spectre de puissance angulaire du CMB. Afin de séparer l'émission des avant-plans galactiques du CMB, opération qu'on appelle la *séparation de composantes*, il a observé le ciel dans cinq bandes de fréquences (23, 33, 41, 61 et 94 GHz).

Passons maintenant au dernier satellite en date axé sur la cosmologie. *Planck* est un satellite de l'ESA bénéficiant d'une large collaboration internationale. Lancé en 2009, il a cartographié le ciel avec ses deux instruments complémentaires *HFI* et *LFI* pendant 30 mois, avec pour mission de déterminer la géométrie et le contenu de l'Univers, et de comparer les observations avec les théories sur la naissance et l'évolution de l'Univers. Le minimum requis pour la mission était de deux couvertures complètes du ciel en un an, il a fonctionné à la perfection en étendant sa mission nominale à 30 mois et en produisant cinq couvertures complètes du ciel avec ses deux instruments. Après épuisement de ses réserves d'hélium (^4He) pour le refroidissement du plan focal, l'instrument *LFI* qui pouvait fonctionner à plus haute température à continué sa couverture du ciel sur une grande partie de l'année 2013. Le décommissionnemment de *Planck* a eu lieu le 23 Octobre 2013, date à laquelle il a été déplacé sur une orbite de rebut.

Le travail présenté dans ce manuscrit est axé sur l'instrument *HFI*, et ce sera donc les caractéristiques de ce dernier qui seront détaillées ci-après. *HFI* a observé le ciel dans six bandes de fréquences (100, 143, 217, 353, 545 et 857 GHz) avec une résolution angulaire allant de $10'$ à $5'$. La couverture spectrale de *HFI* a été optimisée pour la séparation de composantes, et est particulièrement bien adaptée à l'étude de l'effet SZ qui se manifeste sur le CMB dans une gamme de fréquences autour de 217 GHz. Les 54 détecteurs bolométriques de *HFI* ont été manufacturés par JPL/Caltech et sont refroidis à 0.1 K.

Les observations ont été faites au second point de Lagrange du système Soleil-Terre et le chemin suivi par l'axe de rotation du satellite est défini relativement au plan de l'écliptique, il correspond à un mouvement en longitude qui permet de conserver une orientation anti-solaire. *Planck* a observé le ciel avec des balayages circulaires à la vitesse de rotation de un tour par minute. Les balayages circulaires sont particulièrement bien adaptés à l'étude du CMB car ils permettent une analyse naturelle des signaux d'origine cosmologique et d'une partie de ceux induits par l'instrument.

Le système de refroidissement du satellite a assuré les performances des détecteurs en les refroidissant à 0.1 K avec une grande stabilité. Le refroidissement passif a été particulièrement bien exploité avec un système d'ailettes ayant une forte émissivité à l'extérieur et une faible émissivité sur la face interne. La descente en température à 0.1 K est assurée par trois étages de refroidissement actifs. Le refroidissement de l'étage à 20 K a été assuré par un circuit fermé contenant de l'hydrogène, absorbé et désorbé par un matériaux spécifique, et produisant des températures en dessous de 20 K par expansion de Joule-Thompson (JT). L'étage suivant à 4 K a été basé lui aussi sur un circuit fermé fonctionnant à l'hélium sur le principe de l'expansion de JT. Celui-ci utilise un compresseur mécanique tournant à 20 Hz dont les vibrations sont amorties par un système électronique, cependant le blindage contre les perturbations électro-magnétiques étant insuffisant il a laissé fuir un signal parasite caractéristique surnommé *lignes 4 K* et qui sera traité plus loin en détails. Finalement

le refroidissement à 0.1 K se fait par dilution d'^{3}He dans de l'^{4}He, ce système a produit une température très stable avec des variations qui ont été seulement de l'ordre de 1×10^{-4} K, excepté pendant les éruptions solaires les plus intenses dont le flux de particules réchauffait le plan focal.

Passons maintenant aux effets systématique principaux de l'instrument *HFI* qui sont au cœur de ce travail. Par effet systématique on entend toute déviation du signal par rapport à un instrument avec une fonction de réponse angulaire de l'optique parfaitement Gaussienne et un bruit parfaitement blanc et Gaussien. Ces effets systématiques se rencontrent a priori sur la fonction de réponse angulaire de l'optique qui n'est pas parfaitement Gaussienne, et la réponse temporelle de la chaîne de détection dont l'estimation sur les données de vol dégénèrent l'une avec l'autre. La réponse temporelle a été particulièrement difficile à étalonner avec les données de vol à cause d'un excès de réponse à basse fréquence des bolomètres. Ceci a été problématique pour l'étalonnage photométrique qui se fait sur la modulation du CMB par la rotation de la terre autour du soleil pendant la mission, qu'on appelle le dipôle orbital, et qui apparaît sur le signal à une fréquence de 16 mHz. Mais il y a de nombreuses autres sources d'effets systématiques comme les *glitches* causés par les dépôts d'énergie des particules traversant les détecteurs sont aussi une source de bruit non-Gaussien. Ceux-ci doivent être enlevés du signal, mais on verra qu'ils peuvent également fournir de précieuses contraintes sur la réponse temporelle des détecteurs et sur le filtrage de l'électronique de lecture. L'effet le plus difficile à traiter qui a été rencontré vient du composant de numérisation du signal (ADC), et a été amplifié par une combinaison de plusieurs facteurs : une faible dynamique du signal numérisé, un équilibrage du signal analogique qui a fait que la numérisation se faisait exactement sur le défaut le plus important du composant, suivi par la sommation dans l'unité de traitement (DPU) des échantillons détruisant l'information nécessaire à la correction de la non-linéarité de l'ADC. Un autre effet à prendre en compte viens du compresseur de l'étage à 4 K dont les parasites se combinent au signal numérisé par l'ADC et qui rendent la correction de la non-linéarité de l'ADC beaucoup plus compliquée.

La chaîne électronique de lecture des détecteurs est un élément important pour poser le contexte de la correction de la non-linéarité de l'ADC. Il y a en fait 54 chaînes de lecture, une pour chaque détecteur, leur rôle est de produire les données scientifiques à un débit de 180 échantillons par seconde. Les détecteurs de *HFI* sont alimentés par un courant carré qui est produit grâce à une tension triangulaire dérivée par deux condensateurs. Une capacité parasite apparaît à cause de la longueur des câbles entre le bolomètre et le boîtier JFET qui fait l'adaptation d'impédance. Cette capacité parasite complexifie énormément la modélisation de la réponse du bolomètre, et nécessite l'ajout sur l'alimentation du détecteur d'une tension carée de *compensation* du courant de fuite qu'elle génère. Pour diminuer la dynamique du signal un signal carré en opposition de phase est ajouté au signal à la sortie du détecteur dans le boîtier de l'électronique de lecture (REU). Suite à cela, le signal devient très plat, ce qui est un élément critique pour la capture du signal brillant de la galaxie et des planètes sans saturation de l'électronique. Cependant la numérisation de ce signal plat se fait ensuite au centre de l'échelle de l'ADC, la partie qui contient le défaut le plus important du composant, ce qui **maximise** l'impact de la non-linéarité de l'ADC. Le signal modulé est en fait numérisé à la fréquence de 7200 Hz, les 40 échantillons par demi-période de la modulation sont ensuite sommés pour produire le signal scientifique à 180 Hz. Cette étape de sommation est importante pour améliorer le rapport signal sur bruit des données et diminuer le volume des données car la bande passantes allouée à la télémétrie est limitée. Cependant en sommant les échantillons produit par l'ADC, qu'on appelle des *codes*, leur valeur exacte est perdue en même temps que les harmoniques du parasite de l'étage à 4 K ayant une fréquence supérieure à celle de la modulation sont repliées. Une partie importante de ce manuscrit sera consacré à la tâche difficile d'estimer les valeurs perdues pour corriger de l'effet de non-linéarité de l'ADC.

Le traitement des données scientifiques est organisé en trois niveaux qui sont : L1 où la base de données exploitable par les outils informatiques est constituée, L2 où est réalisé le traitement proprement dit des données scientifiques dans le domaine temporel pour la production des catalogues des cartes du ciel, et L3 où est effectuée la séparation de composantes et la production des catalogues de données. Un niveau supplémentaire LS correspond aux simulations. Le reste du travail présenté sera localisé sur le L2, et plus particulièrement la correction de l'effet de non-linéarité qui est la toute première opération effectuée pour le traitement des données scientifiques. Les données scientifiques proprement dites sont organisées en fonction de l'horodatage au moment de leur capture sous forme d'objets qu'on appelle TOI (Time Ordered Input). Les TOI sont découpées par zones de pointage stable d'une durée approximative de 45 minutes qu'on appelle des *ring*. Les données d'un ring peuvent être projetées en fonction de leur phase à l'intérieur du cercle balayé pour produire un PBR (Phase Binned Ring) qui est un objet intermédiaire de la production des cartes et dont le niveau de bruit est réduit d'un facteur racine du nombre de tours effectués sur le ring considéré.

La production des cartes nécessite la suppression des glitches et la correction des effets systématiques comme celui de l'ADC et du compresseur de l'étage à 4 K. Les données sont ensuite déconvoluées de la réponse temporelle de la chaîne de détection, puis le pointage associé est reconstitué avec la fonction de réponse angulaire de l'optique. S'ensuit la calibration photométrique des détecteurs puis la projections des cartes au format `HEALPix` qui est une pixellisation de la sphère adapté au calcul de leur spectre de puissance angulaire. Une carte `HEALPix` à la résolution nominale contient environ 50×10^6 pixels.

Une partie des détecteurs de *HFI*, les PSB, sont sensibles à la polarisation linéaire de la lumière. Pour cela ils sont constitués de deux grilles superposées associées chacune à un canal différent et sensibles à des directions orthogonales de la polarisation. Il faut en fait trois directions différentes pour reconstituer les paramètres de Stokes (I, Q, U) de la polarisation. A cet effet, il y a toujours deux PSB associés à la même fréquence mais avec des orientations relatives de 45° sur la même ligne de balayage du plan focal. Une des problématiques majeure de ce type d'architecture de détecteur est que l'étalonnage relatif des canaux utilisés conjointement pour produire les paramètres de Stokes doit être très précis. Dans ce cas, la non-linéarité introduite par l'ADC est un facteur limitant important, on voit donc qu'il faudra une très bonne maîtrise de cet effet systématique dans le cadre de la polarisation.

Voyons maintenant comment *HFI* participe à l'étude de notre Galaxie. Le milieu interstellaire contient des grains de poussière qui présentent une grande diversité de formes et de tailles. Plusieurs mécanismes sont impliqués dans l'émission de lumière polarisée par ces grains, la connaissance de ces mécanismes est encore partielle et ils sont toujours activement étudiés. Cependant deux éléments dominants ressortent : la rotation des grains et leur alignement le long des lignes de champ magnétique. Cette rotation est aujourd'hui plutôt bien expliquée par un mécanisme qu'on appelle alignement du moment par pression radiative, c'est à dire que le grain dont la surface est irrégulière se comporte comme une éolienne que la pression de radiation non-isotrope des photons fait tourner. Ces grains une fois alignés par le champ magnétique ambiant produisent une émission thermique polarisée visible par *HFI*. Un autre mécanisme produit de la polarisation par absorption sélective de la lumière des étoiles dans les longueurs d'ondes du visible et de l'UV proche mais n'est pas visible par *Planck*. Notre Galaxie a un champ magnétique cohérent à grande échelle et donc l'émission thermique polarisée des grains est un outil unique d'étude de ce champ. De plus des modes-B en polarisation sont produits par la poussière, il est donc impératif de bien caractériser ce rayonnement pour observer les mode-B produits par les ondes gravitationnelles primordiales sur le CMB.

L'effet SZ qui a été évoqué plus haut est un outil performant pour l'étude des amas de galaxies dont la répartition de masse présente un grand intérêt pour la compréhension de

la formation des premières grandes structures de l'Univers jeune. Il se présente sous deux formes : un effet dit thermique produit par diffusion Compton inverse (tSZ) sur le gas d'électrons chauds rencontré dans les amas de galaxies, et un effet dit cinétique (kSZ) lié à la vitesse de ces mêmes amas mais qui est de second ordre. Avec les données de *Planck* on peut produire la carte de l'effet tSZ sur le ciel complet qu'on appelle la carte de paramètre Compton. Cette carte est un outil puissant d'étude et de détection des amas. En effet, l'effet SZ ne dépends pas du redshift et fournit des contraintes qu'on peut combiner avec les observations dans les longueurs d'ondes des rayons-X pour déterminer la masse des amas. Cette analyse peut être complémentée avec le spectre de puissance angulaire de la carte de paramètre Compton car il a été montré que la fonction de répartition de masse des amas peut également être contrainte avec ce spectre de puissance angulaire, ce qui est particulièrement bien adapté au cas de *Planck* où on a une résolution angulaire faible mais une grande statistique grâce à la couverture complète du ciel.

Pour finir sur les résultats, la mission *Planck* a été globalement couronnée de succès après 4 années d'opérations. *Planck* a renforcé le modèle cosmologique ΛCDM en améliorant la précision sur ses paramètres qui favorise un Univers plat dont la courbure est compatible avec zéro. L'analyse des données de la mission *Planck* n'a pas encore permit de détecter la signature des ondes gravitationelles sur le CMB, mais l'analyse conjointe avec l'équipe BICEP2/KECK a été fructueuse en permettant de poser une limite supérieure au rapport tenseur sur scalaire à $r < 0.12$. Ceci a été permit par la couverture complète du ciel à la fréquence de 353 GHz où la poussière thermique a son pic d'émission polarisée et qui permet de caractériser les modes-B qu'elle émet et qui entrent en compétition avec ceux du CMB. Le travail sur les effets systématiques présenté dans ce manuscrit a participé à améliorer la précision de ces résultats.

Chapter 1

Introduction

1.1 Context

1.1.1 History of the Universe

According to inflationary theory (**?**), the Universe experienced a period of rapid expansion during the 10^{-32} seconds after the Big Bang. This inflationary epoch is critical to understand the observed homogeneity and isotropy of the Universe, its flatness, and the absence of magnetic monopoles. Furthermore, quantum fluctuations on the Planck scale were magnified to cosmological sizes to form the seeds of galaxies and large-scale structures in the current Universe.

After inflation, while the Universe remained hot and dense although expanding at a much slower rate, particles forms with baryonic matter and antimatter were formed. During the subsequent annihilation of these particles and anti-particles, an asymmetry in their ratio led to the Universe being matter-dominated, with less than one anti-baryon for each 10^{10} baryons observed today (**?**)). The photons generated during the matter/anti-matter annihilation era correspond to the Cosmic Microwave Background (CMB) radiation observed today.

About 380 000 year after the Big Bang at redshift $z \approx 1100$, the Universal expansion resulted in the primordial plasma cooling to a temperature of about 3000 K. At this moment so-called the *recombination* epoch, protons and electrons could combine to form neutral hydrogen. The Thompson scattering process stopped allowing photons and baryonic matter to decouple. Hereafter the CMB photons have been able to propagate with minimal interaction with matter.

The period before the neutral hydrogen collapsed to form the first large scale structures and stars is called the *dark ages*. During this interval of time the CMB radiation interacts only weakly with the matter.

A second major change in the ionization state of hydrogen in the Universe occurred when the light emitted by the first galaxies and quasars again ionized neutral hydrogen. This is called the *reionization epoch* and took place about 500 Myr after the Big Bang. The mean reionization redshift lies between $z = 8.8 \pm 0.9$ and $z = 7.8 \pm 0.9$ depending on the model adopted (**?**). The increased column density of free electrons with an apparent optical depth $\tau \approx 0.06$, again caused scattering of the CMB photons. This resulted in a damping of primordial anisotropies at small angular scales, and in the production of polarization at large angular scales.

1.1.2 The Cosmic Microwave Background

The CMB is the oldest light in the Universe. Its photons remained tightly coupled to the electrons through Thompson scattering until the end of recombination, after which they free-stream to detectors today, with some modifications due to processes described below. The CMB photons are visible today at microwave wavelength with an apparent blackbody spectrum at a temperature of 2.7255 K.

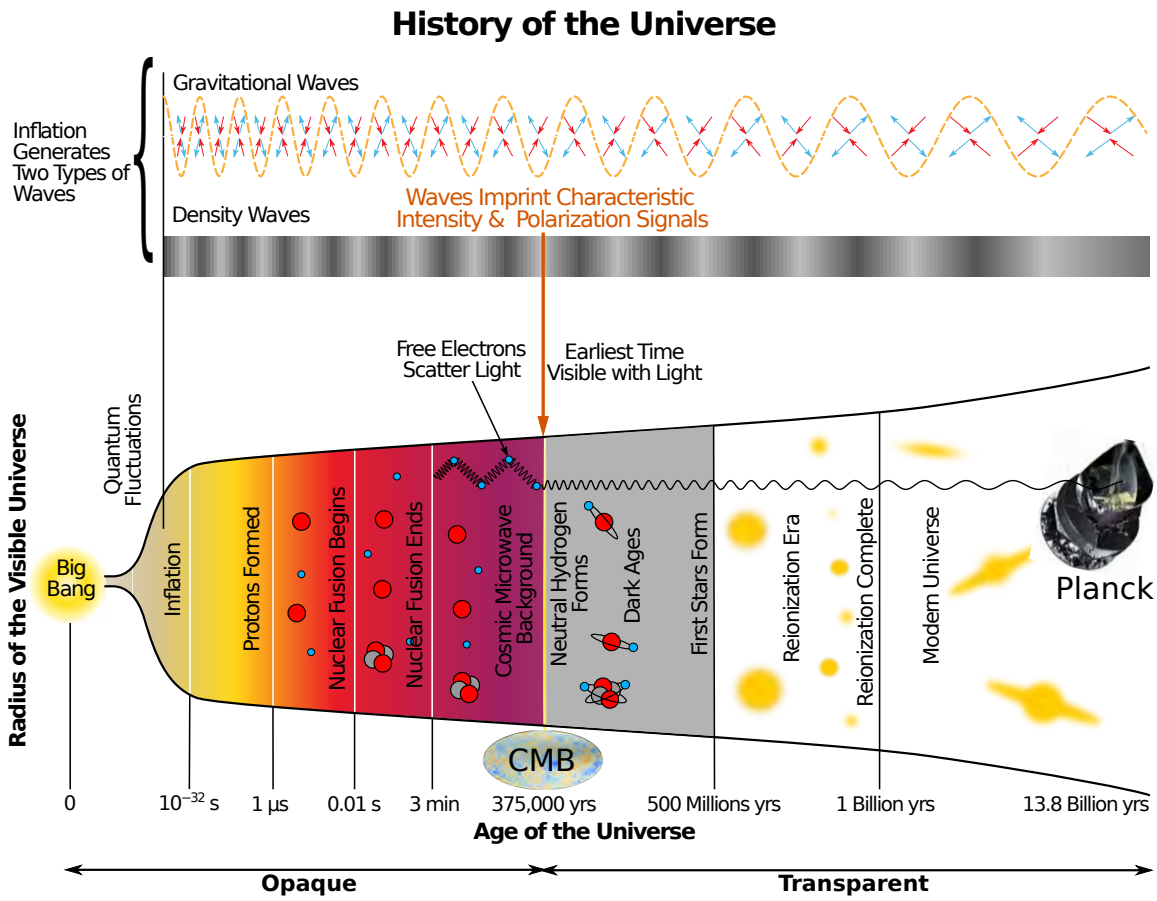


Figure 1.1: History of the Universe. Image credit: NASA `http://bicepkeck.org` (modified)

Anisotropies

The CMB is essentially uniform in temperature to one part in 10^5 (apart from the solar dipole, see below), and the fact that the small fluctuations in temperature are statistically isotropic, i.e. have the same statistical properties (variance, power, spectrum, drawn from a gaussian random field) in all directions. The analysis of these *anisotropies* provide invaluable informations to describe the history of the young Universe, **if they can be accurately measured**.

The solar dipole is due to our local motion relative to the rest frame of the CMB and has an apparent amplitude $\Delta T = 3.3 \times 10^{-3}\,\mathrm{K}$ (Planck Collaboration VIII 2016). Seasonal variations due to the motion of Earth around the sun produce a modulation of about 10% to its amplitude. This so-called orbital dipole provides an excellent calibration source for astronomical instruments using large circles to scan the sky.

Below the level of the solar dipole the *primary anisotropies* are observed with an amplitude $\Delta T < 10 \times 10^{-5}\,\mathrm{K}$. These anisotropies have a characteristic scale dependence characterized by an angular power spectrum (see Sec. 1.1.2) with peaks and troughs on a number of scales. These largely reflect the adiabatic acoustic oscillations of the baryon-photon fluid at decoupling. The statistical distribution of the primary anisotropies in intensity is very well described with the six parameters of the ΛCDM model (aka double dark theory because of the presence of non interacting dark matter and dark energy), thus providing accurate estimations of the density of matter and energy in the Universe.

A small fraction of the primitive CMB radiation is polarized by two effects. Firstly, *quadrupolar anisotropies* create linear polarization (scalar perturbation), this is referred to

as polarization E-modes. Secondly, the inflation is theorized to have generated *gravitational waves* which, following the general relativity theory are expected to leave an imprint on the CMB by creating polarization with whirling orientation (tensor perturbations), this is referred to as polarization B-modes. Unfortunately these most-wanted gravitational waves remnants are very faint and still undetected. The B-modes polarization signal be separated from the dominant foreground emission of our galaxy (**??**). The gravitational wave activity has been intense recently, with the first direct detection and proof of existence by LIGO (**?**). Polarization with *HFI* is covered in Sec. 1.4.

Secondary anisotropies, or *late time anisotropies* are caused by post-recombination interactions of the CMB photons. They are created mainly by:

- gravitational lensing by the massive galaxy clusters and dark matter clumps. This effect makes the CMB a tool of choice for the study of dark matter (**?**).
- the inverse Compton scattering of CMB photons by hot gas in galaxy clusters. This is the so-called Sunyaev-Zeldovich (SZ) effect (**?**). See Sec. 1.5.2;
- Thompson scattering caused by the column density of free electrons during the reionization, it is mainly characterized by its optical depth estimated at about $\tau =\simeq 0.06$ (**?**). This has the effect of smoothing small scales anisotropies, and causes polarization at large angular scales. Accurate measurement of large angular scale polarization provides constraints on the reionization history.

The angular power spectrum

The angular power spectrum is used for the analysis of the angular distribution of CMB anisotropies, and provides a "clean" observable for cosmological model fitting. It is similar to a Fourier spectrum, but it is applied to a sphere.

A sky map $\Delta T(\theta)$ defined over the full sky can be decomposed in spherical harmonics

$$\Delta T(\theta) = \sum_{l\geq 0} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta)\,, \tag{1.1}$$

where θ is the direction on the sphere, ℓ is called the multipole moment and $Y_{\ell m}(\theta)$ is the spherical harmonic function evaluated at direction θ. The coefficient $a_{\ell m}$ is the scalar products of the sky signal $\Delta T(\theta)$ and the spherical harmonic $Y_{\ell m}(\theta)$, it is calculated as an integral over the sphere with

$$a_{\ell m} = \int \Delta T(\theta) Y_{\ell m}(\theta) d\Omega_\theta\,. \tag{1.2}$$

With this definition of the scalar product, the functions $Y_{\ell m}(\theta)$ form an orthogonal basis on the sphere, the first harmonics are shown in Fig. 1.2. The angular power spectrum coefficients C_ℓ are given by

$$C_\ell = \frac{1}{\ell(2\ell+1)} \sum_{-m}^{m} |a_{\ell m}|^2\,. \tag{1.3}$$

The multipole moment ℓ is inversely proportional to the angular scale on the sky, it can be compared to a Fourier mode l defining scale P/l over a finite support P. The angular power spectrum of the CMB measured by *Planck* is shown in Fig. 1.3. The first peak of the spectrum is close to multipole $\ell = 250$ which indicates that the dominant angular size

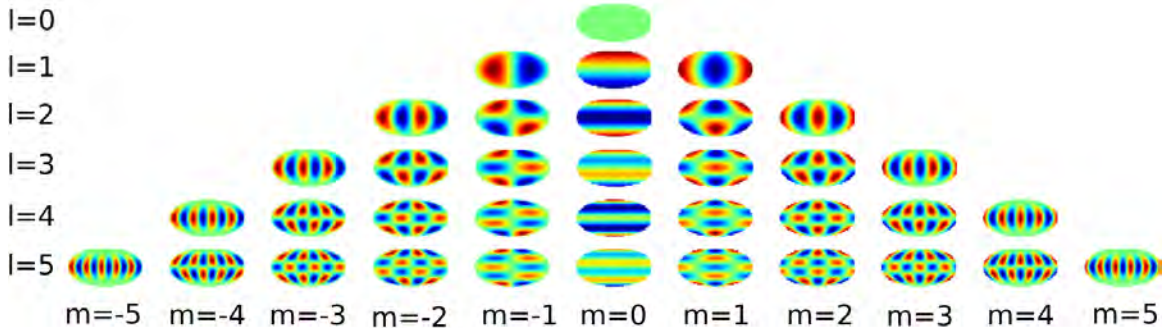


Figure 1.2: Spherical harmonics functions which form the basis for the decomposition used to produce the angular power spectrum. The harmonics functions up to $\ell = 5$ are presented in the 2D Mollweide projection which is the most commonly used in *Planck* Collaboration.

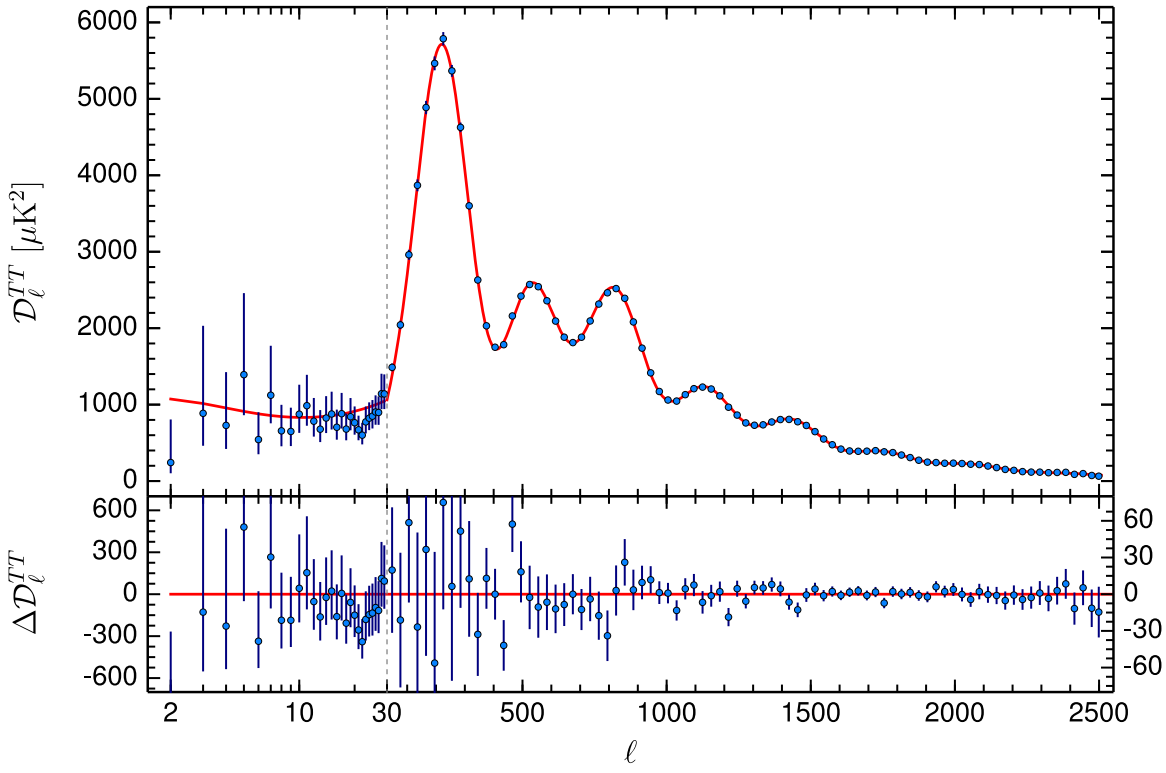


Figure 1.3: *Top*: angular power spectrum of the CMB measured by *Planck*. The values shown are $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/2\pi$. *Bottom*: residuals of the ΛCDM model fit. Image Credit: **?**).

of anisotropies is about 1° or a size of 74 Mpc in comoving[1] distance at the epoch when the CMB was released.

1.1.3 Observational history of the CMB in brief

The CMB radiation was discovered accidentally in 1964 by the American radio astronomers Arno Penzias and Robert Woodrow Wilson with the six meter horn antenna shown in Fig. 1.4. They reported an excess in the antenna noise temperature of about 3.5 K (**?**), which was "isotropic, unpolarized and free from seasonal variations".

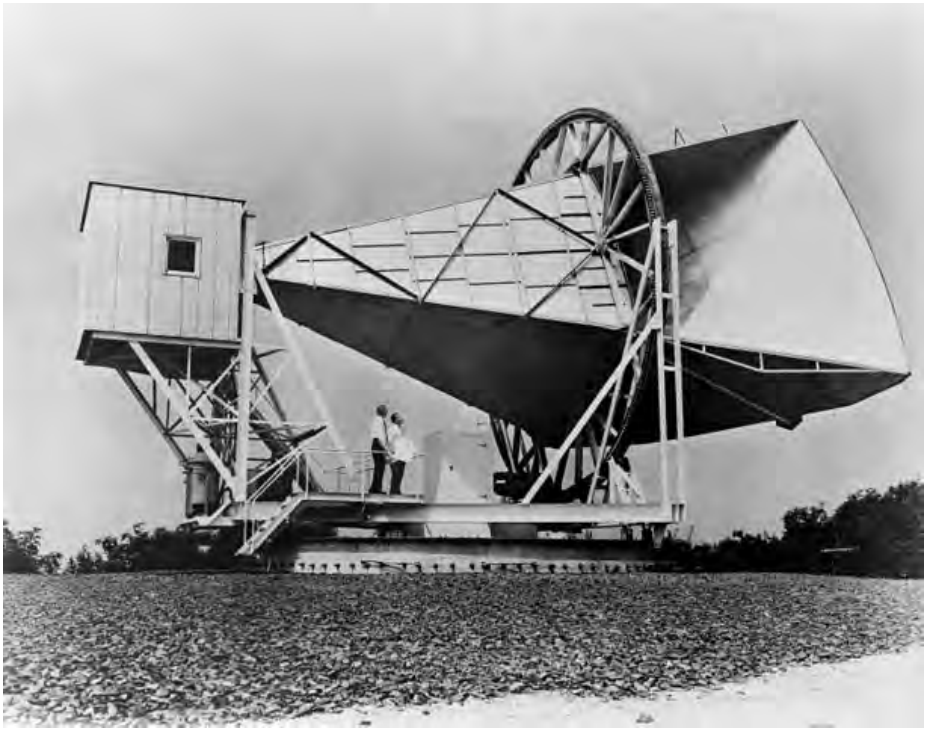


Figure 1.4: The 6m horn antenna from Bell Labs with which the CMB was first detected. Image credit: NASA

The first[2] full sky map of the CMB was made by the COBE satellite launched in 1989. The FIRAS instrument (**?**) found a perfect agreement of the CMB with a blackbody radiation at 2.726 K which pre-dated the first maps of the anisotropies. The DMR instrument mapped the CMB at three frequencies (31.5, 53 and 90 GHz) and provided the first observation of the primary anisotropies at a level $\Delta T/T = 11 \times 10^{-6}$ (**?**). These anisotropies were found to have a Gaussian distribution, but the DMR instrument did not have the required angular resolution to resolve the first acoustic peak of the CMB angular power spectrum. This was the focus in the following decade of many ground and balloon-borne experiments such as BOOMERanG (**?**) and MAXIMA (**?**).

If COBE marked the beginning of the use of CMB measurements for reliable cosmological study, it was the WMAP satellite that began the era of precision cosmology. After only three years of mission (for a total of nine years), the number of cosmological parameters was dramatically reduced to six providing an excellent fit to the ΛCDM model (**?**). To reduce the Galactic foregrounds the telescope made full sky maps at five frequencies (23, 33, 41, 61 and 94 GHz).

[1] Calculating the comoving distance consist to factoring out the expansion of the Universe

[2] «Actually, the RELIKT1 experiment made measurements pre-COBE and later claimed the detection of anisotropy. However, there were clear signs of systematic artifacts in the data.» – Anthony BANDAY

The *Planck* space observatory launched in 2009 has further refined the WMAP results with a much higher resolution (see the visual comparison in Fig. 1.5) and is discussed below. But this is not the end of the story: polarization is expected to provide a wealth of cosmological information. New upcoming CMB missions are on track, including ground-based (QUBIC, BICEP3), balloon-borne (PlanB) and space missions (LiteBird)...

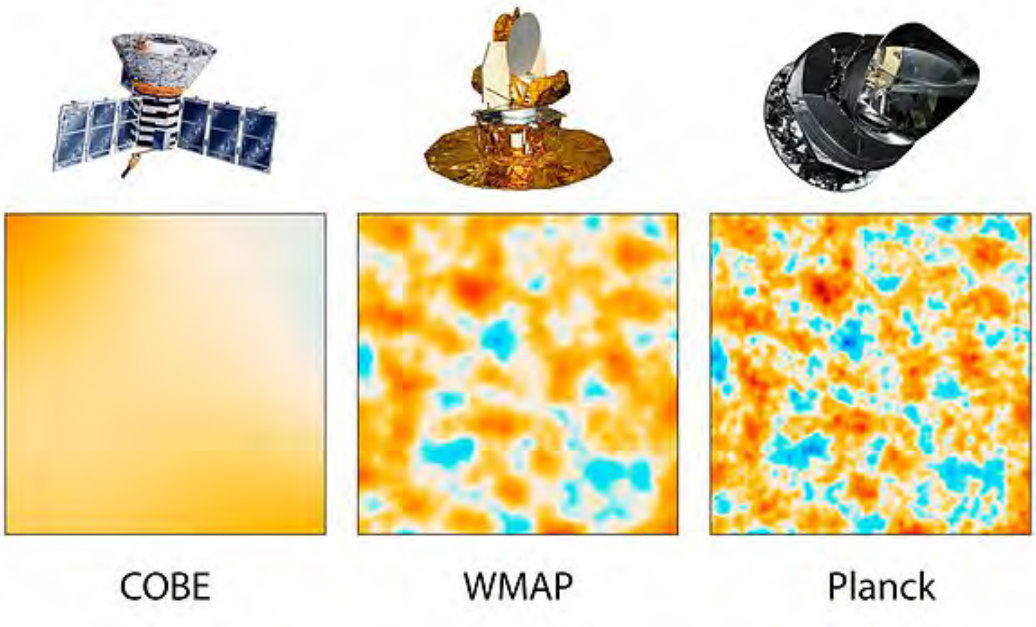


Figure 1.5: Comparison of CMB maps on a 10-square-degree patch from COBE, WMAP and *Planck*. Image Credit: NASA `http://photojournal.jpl.nasa.gov/catalog/PIA16874`

1.2 The *Planck* spacecraft

Planck is a space telescope embedding two complementary instruments *HFI* and *LFI* sharing the same focal plane and at some level the same cooling system. The work presented in this manuscript is based on the *HFI* instrument, this is why it the main focus of this section.

1.2.1 overview

Mission goals

Planck is a space mission of the European Space Agency (ESA) was designed to answer key cosmological questions using observations of the CMB at an unprecedented level of accuracy. Its main goal was to determine the geometry and content of the Universe, and and to compare observations to theories describing the birth and evolution of the Universe.

History

Planck was selected in 1995 as the third Medium-Sized Mission (M3) of ESA's Horizon 2000 Scientific Program, and later became part of its Cosmic Vision Program. It was designed to image the temperature and polarization anisotropies of the Cosmic Background Radiation Field over the whole sky, with unprecedented sensitivity and angular resolution. *Planck* observations provide a means to test theories of the early Universe and the origin of cosmic

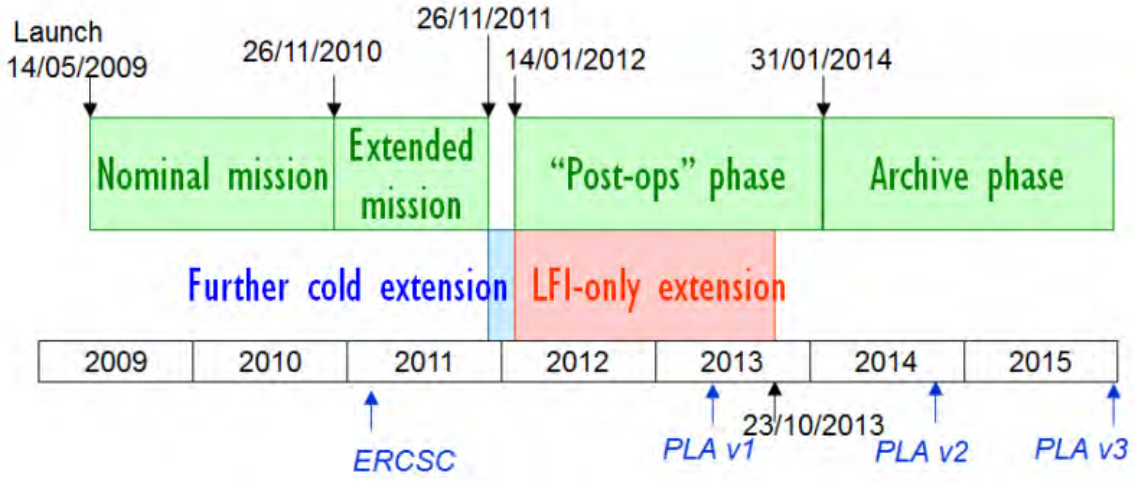


Figure 1.6: *Planck* mission time-line. Image Credit: ESA (`https://wiki.cosmos.esa.int/planckpla/index.php/The_Planck_mission`).

structure and a major source of information relevant to many cosmological and astrophysical issues. The scientific development of the mission is directed by the *Planck* Science Team[3].

Planck was launched on 14 May 2009, and the minimum requirement for success was for the spacecraft to complete two whole surveys of the sky in about one year. In the end, *HFI* and *LFI* worked perfectly for 30 months, and completed five full-sky surveys. Able to work at slightly higher temperatures than HFI, the Low Frequency Instrument (LFI) continued to survey the sky for a large part of 2013, providing even more data to improve the *Planck* final results.

Planck was turned off on 23 October 2013 (see the mission time-line in Fig. 1.6, after being moved to a graveyard orbit. The high-quality data the mission has produced will continue to be scientifically explored in the years to come.

Characteristics of the *HFI* instrument

HFI observes the sky at six frequencies (100, 143, 217, 353, 545 and 857 GHz) with a $10'$ to $5'$ angular resolution. The optical frequencies observed by *LFI* are (30, 44 and 70 GHz) with a $33'$ to $14'$ angular resolution. This spectral coverage is needed for components separation in order to subtract foreground emission from the CMB (**????**). It is also powerful for the SZ effect measurement which has two components (thermal and kinetic) visible over a range of frequency around 217 GHz (see Sec. 1.5.2). *HFI* used 54 bolometric detectors cooled at 0.1 K, manufactured by JPL/Caltech.

1.2.2 Scanning strategy

Planck has been operating at the second Lagrangian point (L2) on the Sun-Earth system as shown in Fig. 1.7. The path followed by the spacecraft spin axis is defined with respect to the Ecliptic plane. It corresponds to a motion in longitude which maintains an anti-Sun pointing (about one degree per day), to which is added a cycloidal motion (precession) of the spin axis around the anti-Sun position (fiducial point).

Planck observed the sky using circular scans at a rotation speed of one revolution per minute, observing the same sky circle for typically 45 minutes – this stable pointing period is

[3] The *Planck* Science Team is a group of scientists whose general task is to monitor and advise on all aspects of *Planck* which may affect its scientific objectives. The list of its members is available at `http://www.cosmos.esa.int/web/planck/science-team`

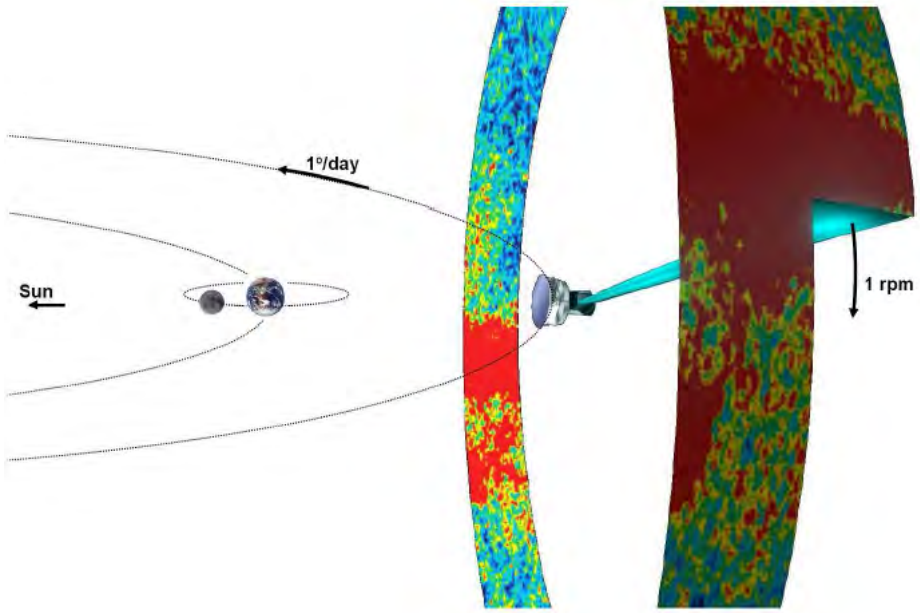


Figure 1.7: *Planck* Scanning strategy. Image Credit: ESA (`https://wiki.cosmos.esa.int/planckpla/index.php/Survey_scanning_and_performance`).

called a *ring*. Circular scans on the sky are well adapted for CMB experiments because they allow a better separation of signal from cosmological origin and some of the effects coming from the instrument (**?**).

1.2.3 The cooling system

The performance of the Planck detectors in space is enabled by their low operating temperatures, 20 K for *LFI* and 0.1 K for *HFI*, achieved through a combination of passive cooling and active coolers (**??**).

Planck makes extensive use of radiative cooling, specifically with the V-groove radiators and the telescope baffle which has a low emissivity on its inner surface and a high emissivity on the outer one. These parts are shown in a cutaway view in Fig. 1.8, indicating that the temperature decreases progressively along the spin axis of the spacecraft, down to the Focal Plane Unit (FPU) cooled to 0.1 K. This temperature is achieved with three cooling stages.

The 20 K cooler: the first subsystem of the active cooling chain is a closed cycle sorption cooler (**?**) using hydrogen as the working fluid. When hydrogen is desorbed by heating the sorption material, a Joule Thomson (JT) expansion produces temperatures below 20 K with a heat lift[4] of 1 W. This subsystem developed by the Jet Propulsion Laboratory produces a large heat lift without mechanical vibrations. It is referenced nominally as the *20 K stage*, albeit producing lower temperatures at around 18 K for the pre-cooling of the Helium fluid of the next stage. More details on the principle of the sorption cooler are are given in Sec. 6.1.3. Fig. 1.9 shows the thermal design around the focal plane of *HFI*.

The 4 K cooler: this subsystem is also based on the JT effect and works with a closed helium circuit. The 4 K cooler is described in detail in **?**, and its integration within the spacecraft in **?**. It is designed with the requirement to provide a heat lift of 9 mW and cool the structure hosting the *HFI* focal plane at a temperature of about 4.2 K. However unlike the 20 K cooler, a Stirling cycle compressor produces mechanical vibrations. Despite an

[4] The heat lift is the maximum power the system is able to extract from the lower temperature stage

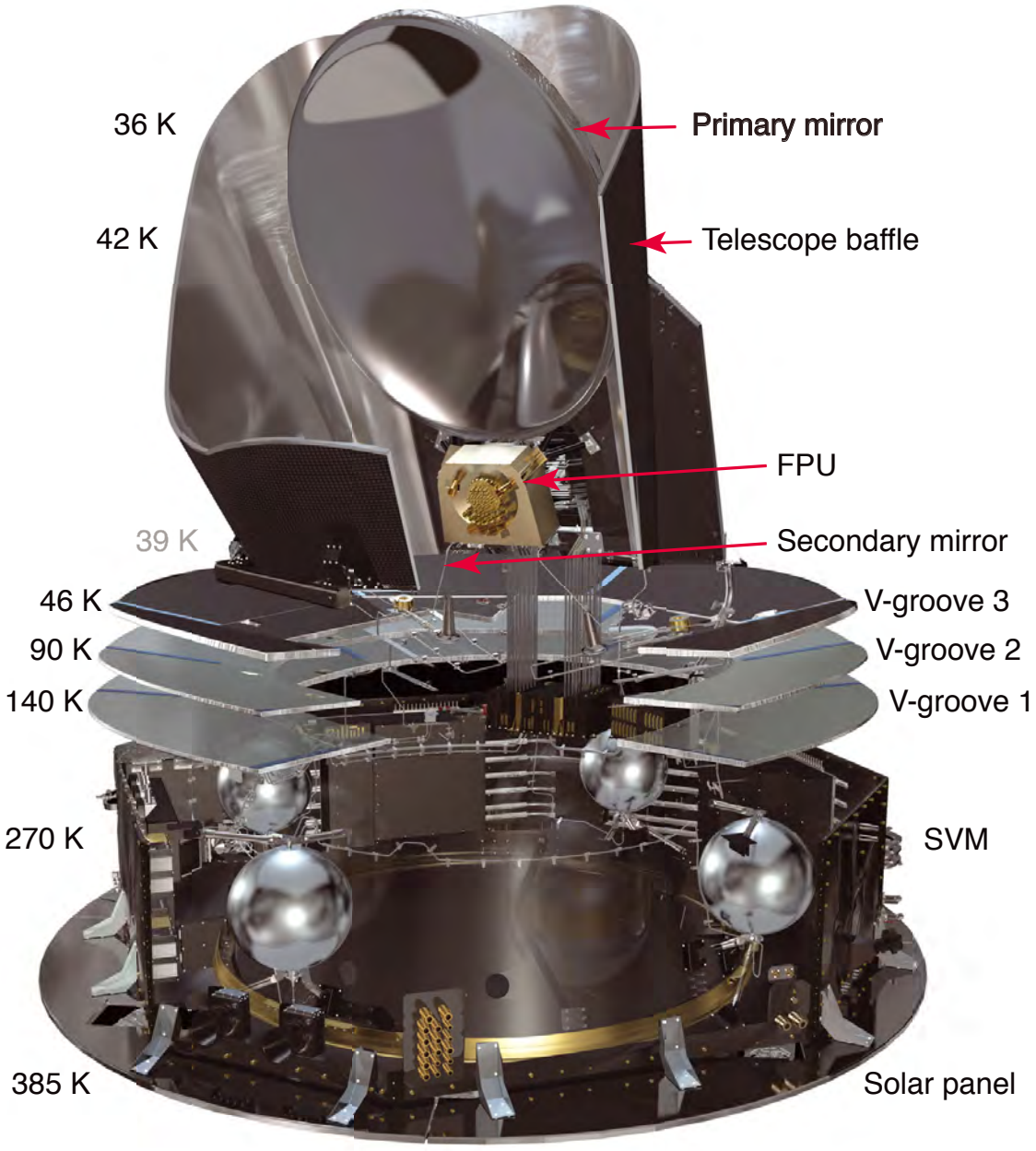


Figure 1.8: Cutaway view of *Planck*, with the temperatures of key components in flight. The solar panel at the bottom always faces the Sun and the Earth, and is the only part of the flight system illuminated by the Sun, the Earth, and the Moon. Image Credit: **?**.

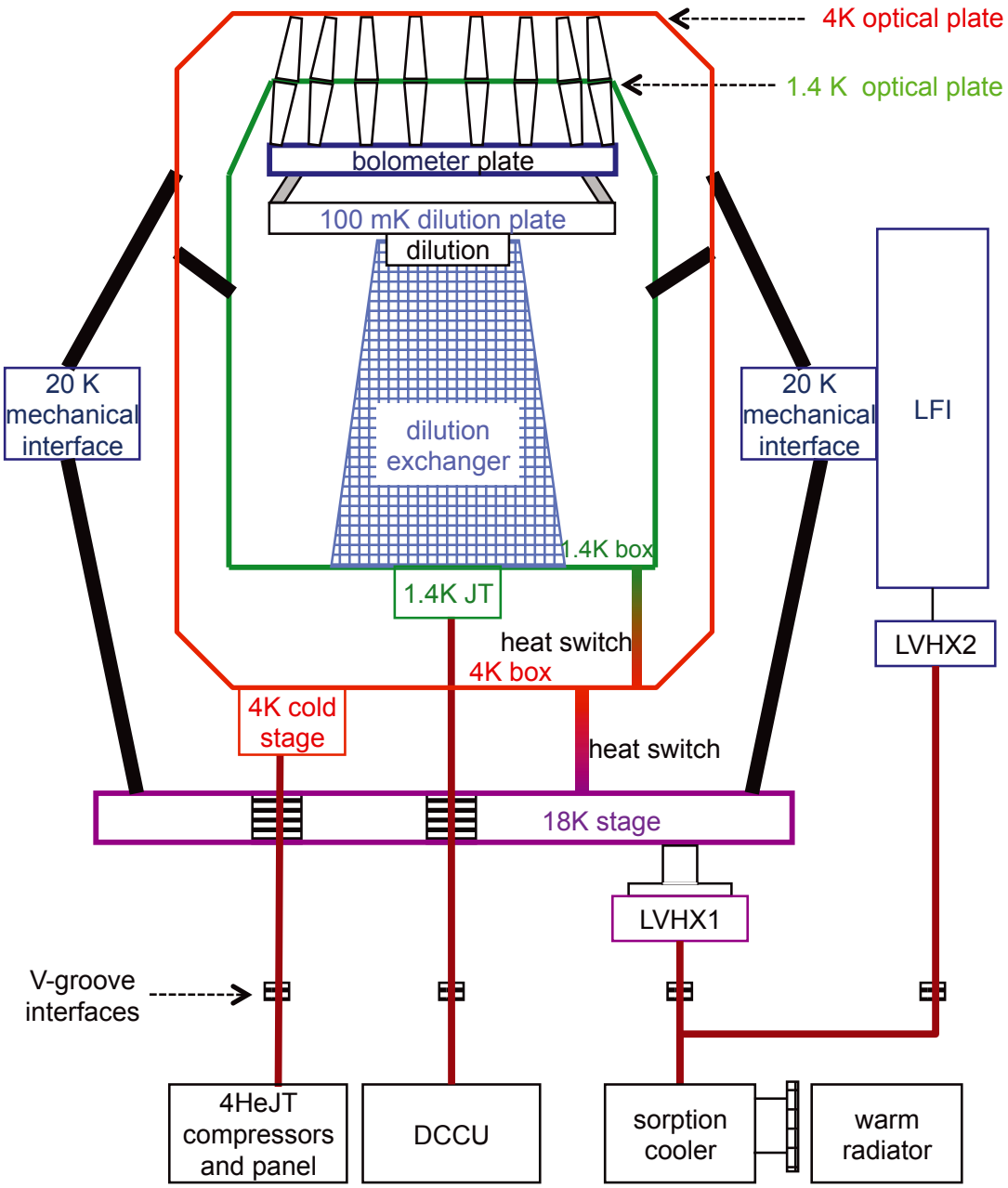


Figure 1.9: Thermal architecture of the *HFI* focal plane. Image Credit: **?**.

active damper system, which minimizes micro vibrations, it propagates some electromagnetic interferences (referred to as 4 K lines) to the electronics readout chain of the detectors.

THE 0.1 K COOLER: this is a dilution cooler which produces a temperature of 0.1 K for the bolometers through the dilution of ^{3}He into ^{4}He, with a heat lift of 1 µW. The mixture of ^{3}He-^{4}He produces a temperature of 1.4 K with a JT expansion and is released into space, this is the major limiting factor in the duration of the mission because of the finite helium supply.

This cooling architecture is very complex, and was overall a great technical success providing very high stability of the 0.1 K stage for bolometers. The variations of the 0.1 K stage over the mission are about 10^{-4} K, except during well identified solar flares events as shown in Fig. 1.10. During these events the focal plane is subject to bursts of energy deposited by particles and its temperature raises significantly. The flux of particles on the focal plane and the detectors is constant and causes a significant loss of data. Particles hits on bolometers cause important variations of signal and significant loss of data. Scientific results are not affected by this data loss, but this is one of the major systematic effect of the instrument.

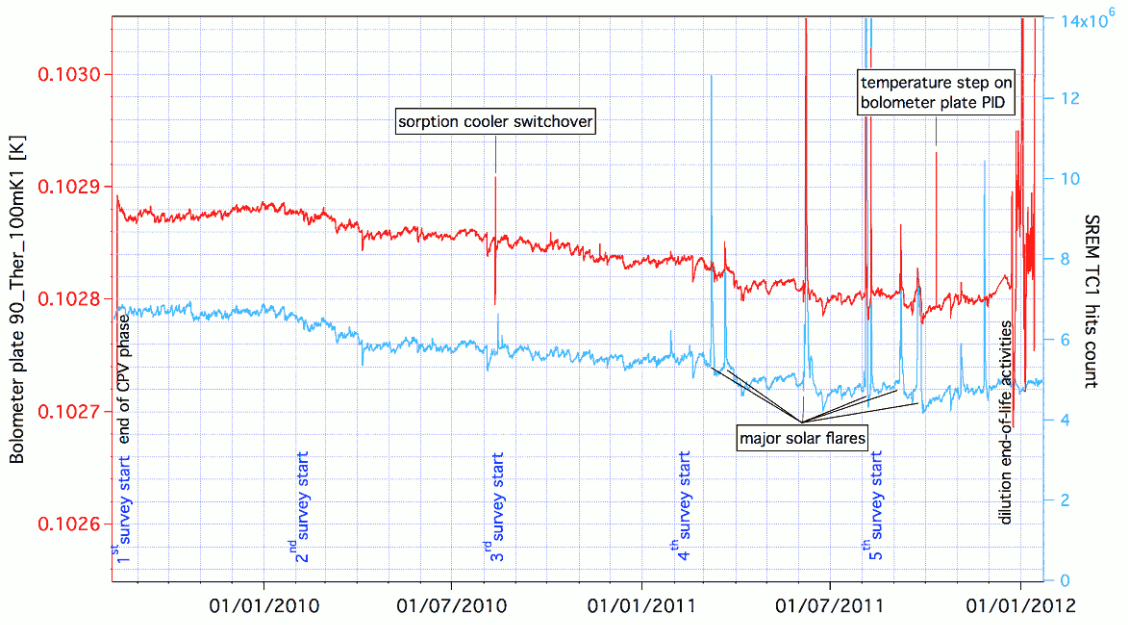


Figure 1.10: Thermal variations of the bolometer plate temperature (0.1 K stage) over the mission (red line). The blue line show the particles hit count (protons of energy > 20 MeV) monitored by the on-board Standard Radiation Event Monitor. Image credit: ESA `https://wiki.cosmos.esa.int/planckpla2015/index.php/HFI_cryogenics`

1.2.4 *HFI* Systematics

Systematic effects (in short systematics) are a major concern in the design of a precision instrument, because they have to be characterized for each channel and removed in order to get the closest estimation of the true astrophysical signal. We call systematic effect any deviation from the signal of an instrument with a perfectly Gaussian beam, a perfect pointing and a perfectly white and Gaussian noise[5]

[5] Definition given by the *Planck* working group on systematic effects. A white noise have a constant power spectral density, and the Gaussian property implies that each sample is independent and drawn from a normal distribution

Major sources of systematics

BEAMS: these describe the angular response of the optical system to a point source on the sky. In *HFI* there are several definitions:

- the *scanning beam* is defined as the beam measured from the response to a point source of the full optical and electronic system. This is measured using planet crossings (**?**). In practice it is strongly degenerate with the time transfer function and both have to be determined simultaneously. More details are given in Sec. 2.1.2. It is not possible to deconvolve *HFI* sky maps from the angular response function using this object because the computing resources involved are to high, but the scanning beam is the object relevant to map analysis at small angular scales and flux calculation;
- the *effective beam* at the map level is the overall angular response to the sky in a map pixel, which results from the combined effect of the instrumental response, the scanning strategy and the data processing. It is the kind of object which is relevant to deconvolve the average effect of the scanning beam over the CMB angular power spectrum.

TIME TRANSFER FUNCTION: this is the temporal response of the electronic readout chain to an optical power excitation of the bolometer. In *HFI* a Low Frequency Excess Response (LFER) has been detected during the ground calibration campaigns (**?**), and proved to be extremely challenging to characterize in-flight. This is a critical issue because the calibration of *HFI* detectors relies on the orbital dipole which is seen at 16 mHz resulting from *Planck* spinning frequency of 1 rpm, in a frequency range where there are no other simple calibration sources. This complex issue is caused by a mixture of the thermal behavior of the detectors, and the bolometer AC biasing. Chap. 2 is dedicated to the LFER issue and covers in details the time transfer function and how the *HFI* bolometers works.

GLITCHES: these are features in the electronic readout response caused by particles energy deposited on the highly sensitive *HFI* detectors. The optical absorber part of the detector is a metallized grid which offers in theory a small cross-section to cosmic rays. However a combination of the thermal architecture of the detectors (**?**) and a period of low solar activity cycle caused a much higher rate of glitches than expected with around one event per second (**?**). Hopefully they can be removed efficiently and have a minimal impact on *HFI* data. As a side effect, glitches result in specific excitations of the electronics readout chain of the detectors, and for this reason they provide a valuable probe of the time transfer function. This topic is covered in Chap. 3.

ANALOG TO DIGITAL CONVERTER NONLINEARITY: the ADC is a critical component which digitizes the analog signal at the end of the analog electronics readout chain, before sending samples to the Digital Processing Unit (DPU) as detailed in the next section. However a combination of factors including a low signal dynamic and acquisition at the location of the major nonlinear feature of the ADC scale caused apparent gain variations in time of science data signal. These gain variations at a level of 1-2% have been a serious concern for photometric calibration (**?**). The characterization and correction of ADC nonlinearity is central to this thesis work and is covered specifically in part II.

4 K LINES: the 4 K lines are a set of parasitic sinusoids produced by the 4 K stage cooler at well identified frequencies. In practice they appear at four distinct frequencies in the science data and can be removed easily, with the xception of *resonant rings* acquired when the spacecraft spin frequency $f_{\rm spin}$ is a multiple of the 4 K lines fundamental frequency $f_{\rm 4K}$. When working on ADC nonlinearity I found that their subtle interaction with this component was a major concern for ADC nonlinearity correction, so they had to be characterized accurately for inclusion in the correction. This topic is covered in Chap. 6.

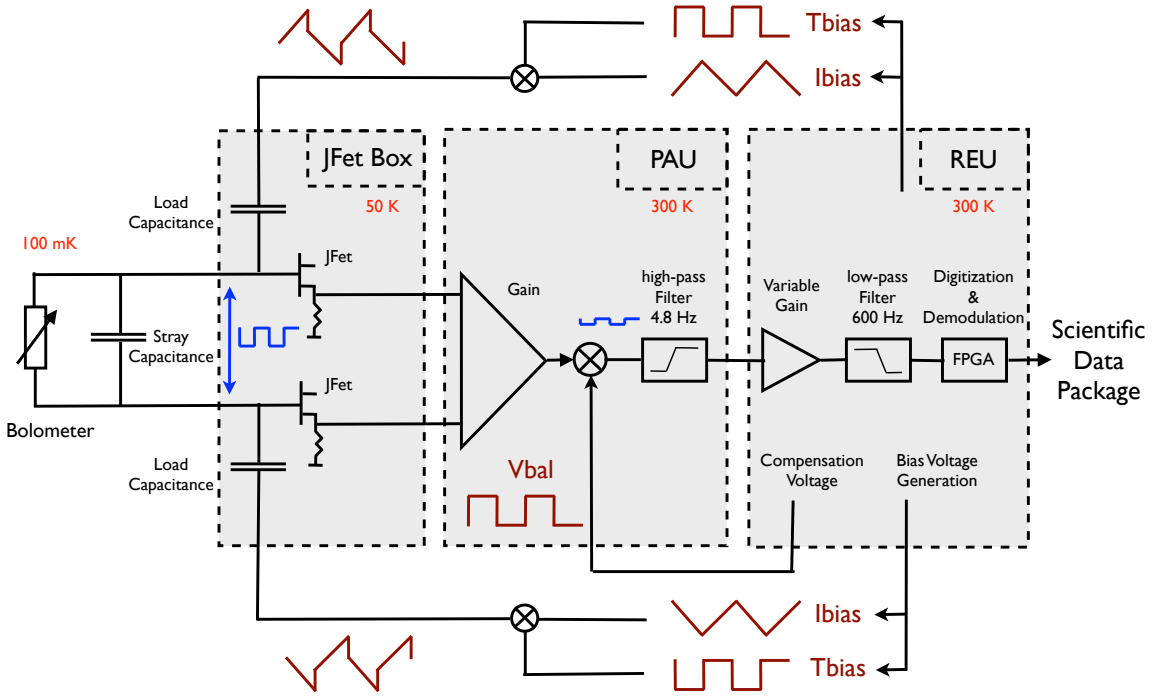


Figure 1.11: Schematic of the *HFI* readout electronic chain. Figure reproduced with the permission of Ludovic Montier, in this version we have corrected the shape of the bolometer bias signal (triangle+square) compared to the version printed in **?** where the square wave sign was inverted.

This list is not exhaustive, it attempts to covered the main systematics which are also relevant to this work. Additional sources of systematics includes bandpass mismatch (imperfect characterization of optical bandpass filters), crosstalk (cross channels parasitics), mismatch between two polarization sensitive bolometers, detectors thermal nonlinearity, etc.

1.2.5 The Calibration and Performance Validation (CPV) phase

The CPV was the last phase of the spacecraft commissioning before the beginning of the nominal mission. It consisted mainly in calibrating the *HFI* and *LFI* instrument and checking their behavior was nominal. This phase has been of particular importance to the work presented in this manuscript because it provided vital informations about the electronics readout chain.

1.2.6 The *HFI* Electronic Readout Chain

I describe here the main elements of the electronic readout chain up to the production of science data samples for the 54 science data channels of *HFI*. The shape of the analog signal, combined with the processing of digitized samples to produce science data, are critical elements that amplify the impact of the ADC nonlinearity.

Overview

The readout electronics consist of 72 channels designed to perform low noise measurements of the impedance of 52 bolometers (**?**), 16 accurate low temperature thermometers, all in the 10 MΩ range, one resistor of 10 MΩ and one capacitor of 100 pF. The readout electronic chain was split into three boxes. These are the JFET box doing source following[6], the Pre-Amplifier Unit (PAU), and the Readout Electronic Unit (REU). Each of the three boxes (JFET, PAU and REU) consists of 12 sets of six channels. These sets are electromagnetically shielded independently from each others and referred to as *belts*.

The first nine belts are dedicated to bolometers and the channels will be referenced with labels such as `02_143-1a` where the first two numbers are for the belt number and channel index between 0 and 5, followed by the three numbers of the channel frequency, and an additional channel index for this frequency with an optional letter `a` or `b` for polarized bolometers.

HFI detectors biasing

The *HFI* bolometers were developed by the Jet Propulsion Laboratory and Caltech (**?**), they were biased with an AC current at modulation frequency $f_{\mathrm{mod}} = 90.18$Hz developed to reduce JFET $1/f$ noise **?**, and tested on the *Diabolo* ground-based experiment and in the *Archeops* balloon-borne experiment (**?**).

The bolometer bias current constists of a triangular wave plus a square compensation wave, which are derived by two load capacitances as shown in Fig. 1.11. The bias current intensity can be set in-flight on a per channel basis with the `ibias`[7] parameter. An important stray capacitance arises from the length of cable between the detector and the JFET Box and effectively apply a low pass filter on the signal at the output of the detector. In practice, the compensation square wave is introduced to reduce the effect of the low pass filter, thus producing an effectively square signal at the output of the bolometer, as represented on the figure by the blue square wave.

Bolometers are thermally sensitive components with temperature varying impedance. Under AC bias the electro-thermal equilibrium of these components is extremely difficult to model analytically. However this modeling has been motivated by the LFER issue and ADC nonlinearity correction. This work is detailed in Chap.2, and led to results presented in Sauvé and Montier (2016).

Amplification and balance of detectors signal

The square signal at the output of the JFET Box presents a high dynamic range, which has to be reduced to allow the capture of the bright signal from Galactic plane and planets (the later signal is critical for beams estimation **?**). To reduce the signal dynamic range, it is balanced by adding in the PAU a square signal (the balance square) in opposition of phase. The PAU also amplifies the signal by a factor 100 and applies a second order high pass filter at 600 Hz.

The REU provides a variable gain which can be set in-flight with the electronic parameter `gamp`, to the relative amplification values 1/3, 1., 3 and 7.6. In practice the CMB channels (353 GHz and below) are set to `gamp` = 1, and the bright sub-mm channels at 545 and 857 GHz to `gamp` = 1/3, which has some impact on the relative noise and 4 K lines level.

It is worth to mentioning here that electrical crosstalk between channels is also source of parasitic signal on a given detector channel. While this electrical crosstalk is not significant

[6] Source following is performed with a circuit reproducing on its output the tension at its input without affecting the input tension

[7] channel parameters are digital values that can be set in the range $[\![0, 4095]\!]$

for science data, its amplitude is of order of the balanced signal and as a consequence renders an analytical modeling of the balanced signal shape nearly impossible.

Digitization by the ADC : the *fast samples*

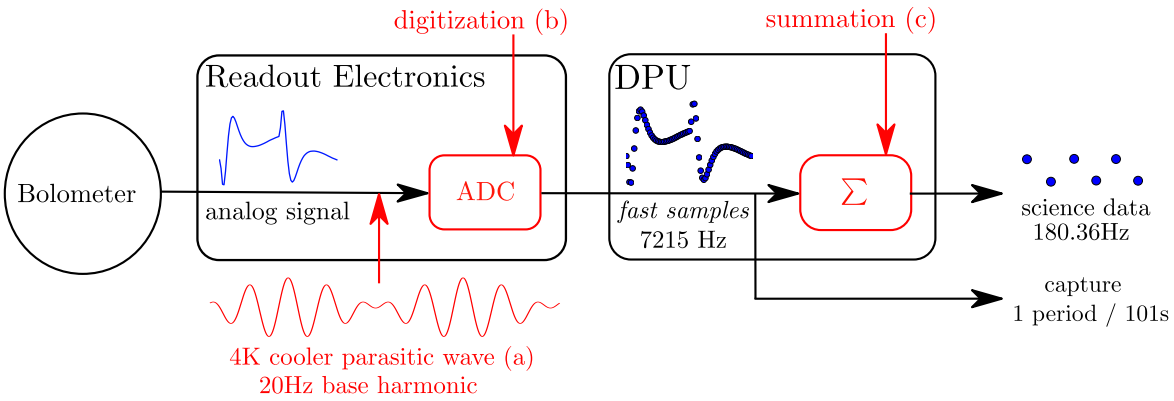


Figure 1.12: Functional view of in-flight signal processing. (a) Parasitic wave originating from the 4K stage mechanical cooler leaks into the signal through the electronic readout chain. (b) The digitization by the ADC introduce the nonlinear effect. (c) The signal summation in the DPU causes loss of information.

The analog signal, mixed with the parasitic 4 K lines is sampled at a rate of $80 f_{\text{mod}} = 7215\text{Hz}$, as shown on the functional view in Fig. 1.12. This yields the digital signal (b) counting 80 samples per modulation period. This signal is called *fast samples* and cannot be sent to Earth, because of the limited bandwidth for all the 72 high data rate channels of the HFI instrument. However every 101 seconds, and for each detector, a period of 80 *fast samples* is captured before summation by the Data Processing Unit (DPU) for further data analysis. These *fast samples* periods have been a gold mine for ADC nonlinearity characterization and correction. A clear distinction will be made hereafter between the *fast samples* and the science data samples.

The digitization is the step where the ADC nonlinearity effect is introduced, while the balancing sets the analog signal location exactly at the center of the ADC scale, at the precise location of the biggest defect[8] of the component. As a consequence, the nonlinearity of the ADC has a maximum impact, as detailed further in Chap. 5.

Signal summation

The last step to be considered in the signal processing is the half period summation (see (c) in Fig.1.12), which operates as a low pass filter for reducing electronic noise. Each science data sample is defined as the sum of the 40 *fast samples* of its corresponding half period, yielding the science data signal sampled at frequency $f_{\text{acq}} = 2 f_{\text{mod}}$.

However, the signal summation leads to a critical loss of information because of:

- Loss of ADC code value: the summation mixes the nonlinear codes[9] at the output of the ADC;
- Loss of 4K lines real signal: the spectral distribution of 4K harmonics extends to much higher frequencies than f_{mod}, and the summed modulated signal has only 4

[8] The exact nature of this defect due to the discrete nature of an ADC output is detailed in Sec. 4.2

[9] A code is the integer value in the range $[\![0, 65535]\!]$ resulting from the conversion of the analog input signal by the ADC.

visible harmonics at {20Hz, 40Hz, 60Hz, 80Hz} resulting from the folding of many higher frequency harmonics. As a consequence it is not possible to estimate the analog 4K signal directly from the science data.

Because of the DPU summation, it is clear that the knowledge of the ADC nonlinearity alone is not sufficient to correct for its effect on the science data. A model describing the bolometer response to the sky signal along with 4 K lines is necessary to reconstruct the signal as seen by the ADC. This model is described in Chap. 5.

Compression

The on-board DPU applies a tunable quantization step q on science data to reduce furthermore the volume of data. The value of q is set to about half the white noise RMS of the channel, so that this step increases the total noise of the channel by about 1% (see Sec. 5.2.3 in **?**). Every 254 science data samples the quantization step window is offset, ensuring a stochastic distribution of sample values.

This step is considered to have a negligible impact on the systematics effects discussed in this work and will not be discussed hereafter.

1.3 Data processing

1.3.1 overview

The HFI data processing pipeline is composed of four levels split as follows[10]:

- Level 1: (L1) consists of receiving the telemetry and ancillary data files and ingesting them into the Data Processing Center (DPC) data base, handled by the `DMC` proprietary software which is hosted by the `magique3` cluster at IAP.
- Level 2: (L2) is where the data are processed from timelines into maps;
- Level 3: (L3) is where the data in the form of frequency maps are converted to catalogues and full-sky astrophysical component maps;
- Level S: (LS) is the common HFI/LFI simulation software.

In practice, most of the work presented below belong to L2 where the systematics effect from the instrument are corrected for.

Time ordered data

Raw science data is stored in the DMC database (see Appendix G). More than two years of *HFI* science data for the 72 channels is available in the form of continuous Time Ordered Input (TOI) files that can be accessed by sample index, corresponding to $72 * 25 \times 10^9$ science data samples.

The TOI data can be accessed by their so-called *ring* index. A ring is the stable pointing period of about 40 minutes duration when the spacecraft continuously scans the same sky circle as described in Sec. 1.2.2. Between each ring the spacecraft changes its pointing orientation and during this period of time, typically a few minutes, the pointing is known to be *unstable* and therefore the science data is flagged as unusable. The five full sky surveys of *HFI* represent about 5000 rings per survey for a total of 27008 rings (including calibration phases at the beginning/end of the mission).

[10] See `https://wiki.cosmos.esa.int/planckpla2015/index.php/The_HFI_DPC`

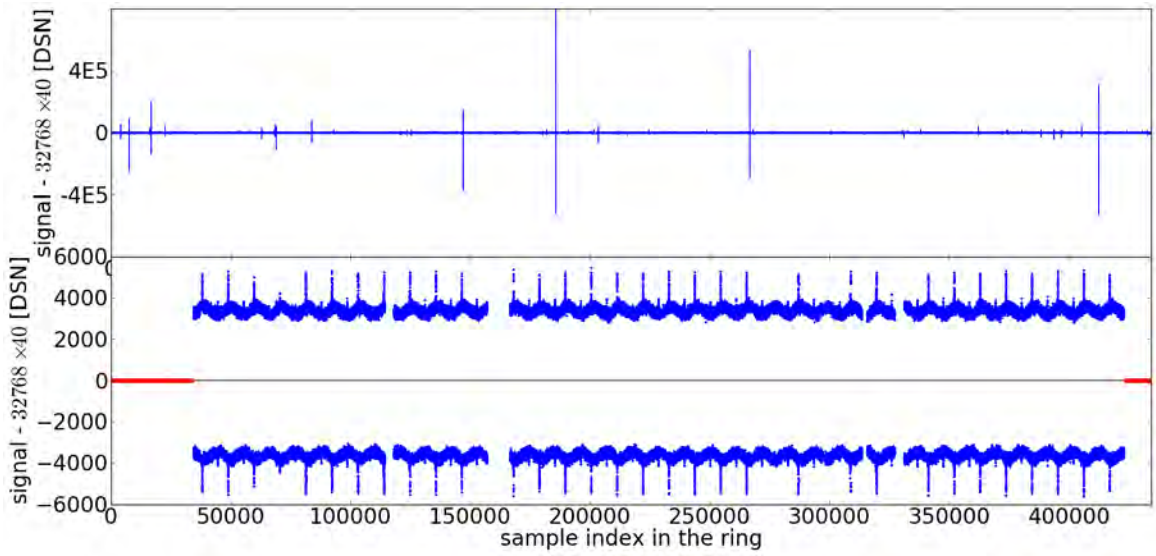


Figure 1.13: Example of raw science data TOI for channel `02_143-1a`. The median value of the digital signal scale has been subtracted for visibility. The stable pointing period is ring 1000. *Top*: raw modulated science data, the glitches visible as sharp vertical lines are the dominant dynamic feature. *Bottom*: the same data after removal of unstable pointing periods and invalid data (mainly glitch flagged data). The red line at the beginning and the end of the ring is the unstable pointing time period.

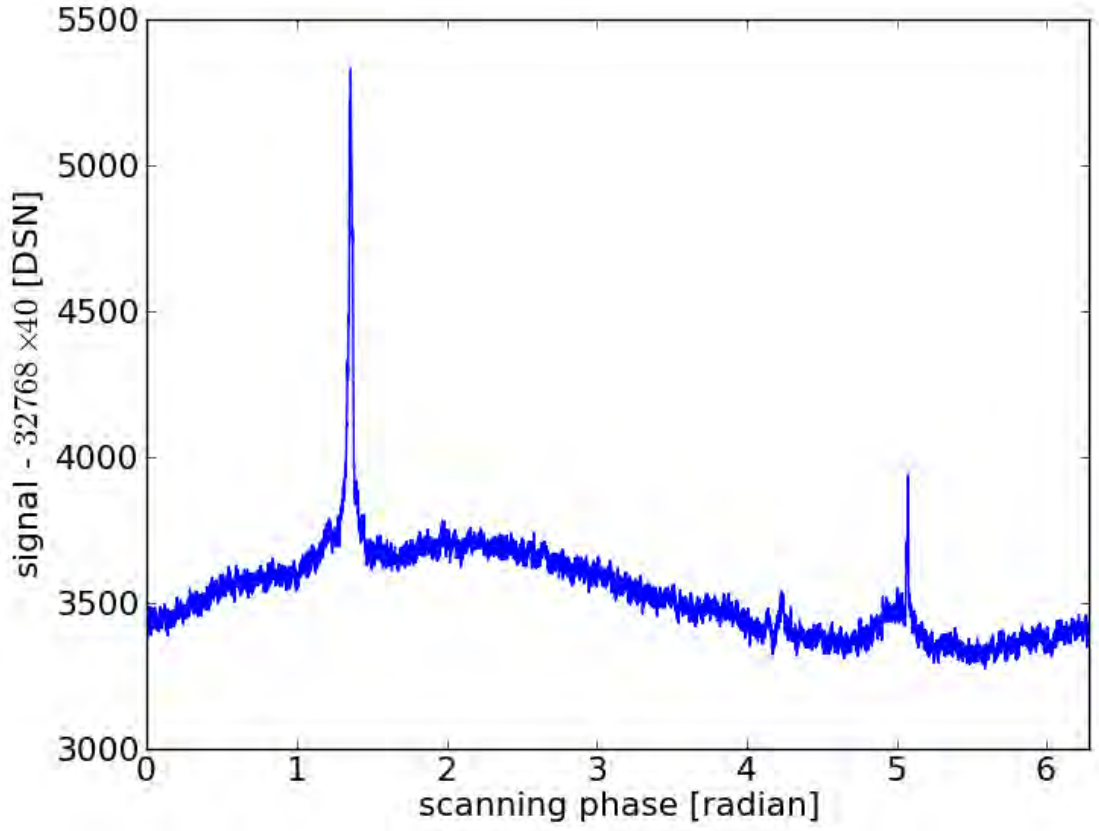


Figure 1.14: Example of Phase Binned Ring (PBR) resulting from the binning in phase of data shown in Fig. 1.13. At the nominal spacecraft spin frequency (1 rpm) the sample integration time covers about $2'$ on the sky, the corresponding number of bins for the PBR is 10800.

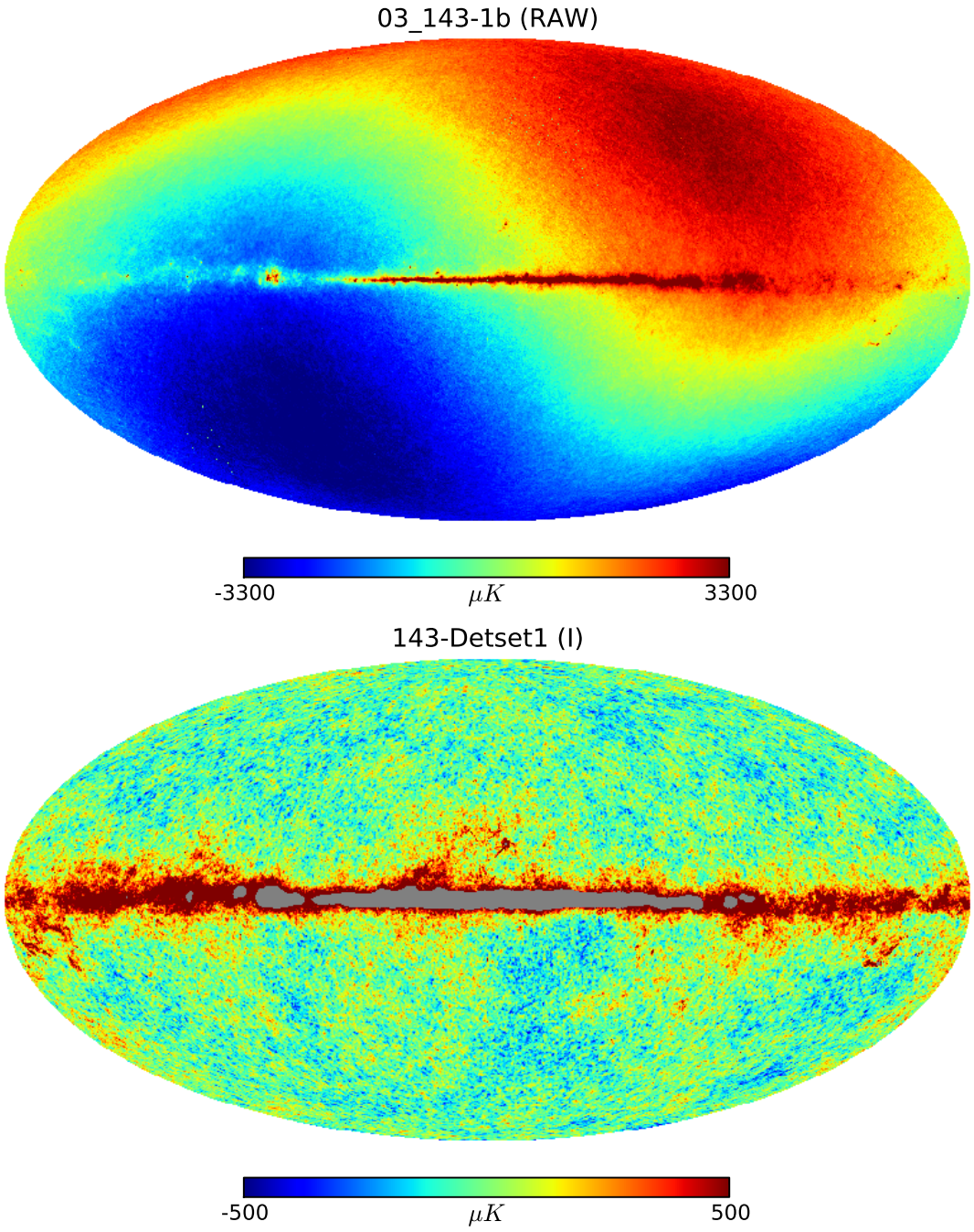


Figure 1.15: *HFI* 143 GHz intensity sky maps at different stages of processing viewed in Mollweide projection. *Left*: raw demodulated science data projected after removal of invalid data (mainly glitches) and 4 K lines. This is an unofficial product built with science data from channel `03_143-1b`. *Right*: the final version of the 143 GHz map for the 2015 data release, where the brightest features of the Galactic plane are masked in grey. This map results from the average of several channels in the 143 GHz detector set 1 of two.

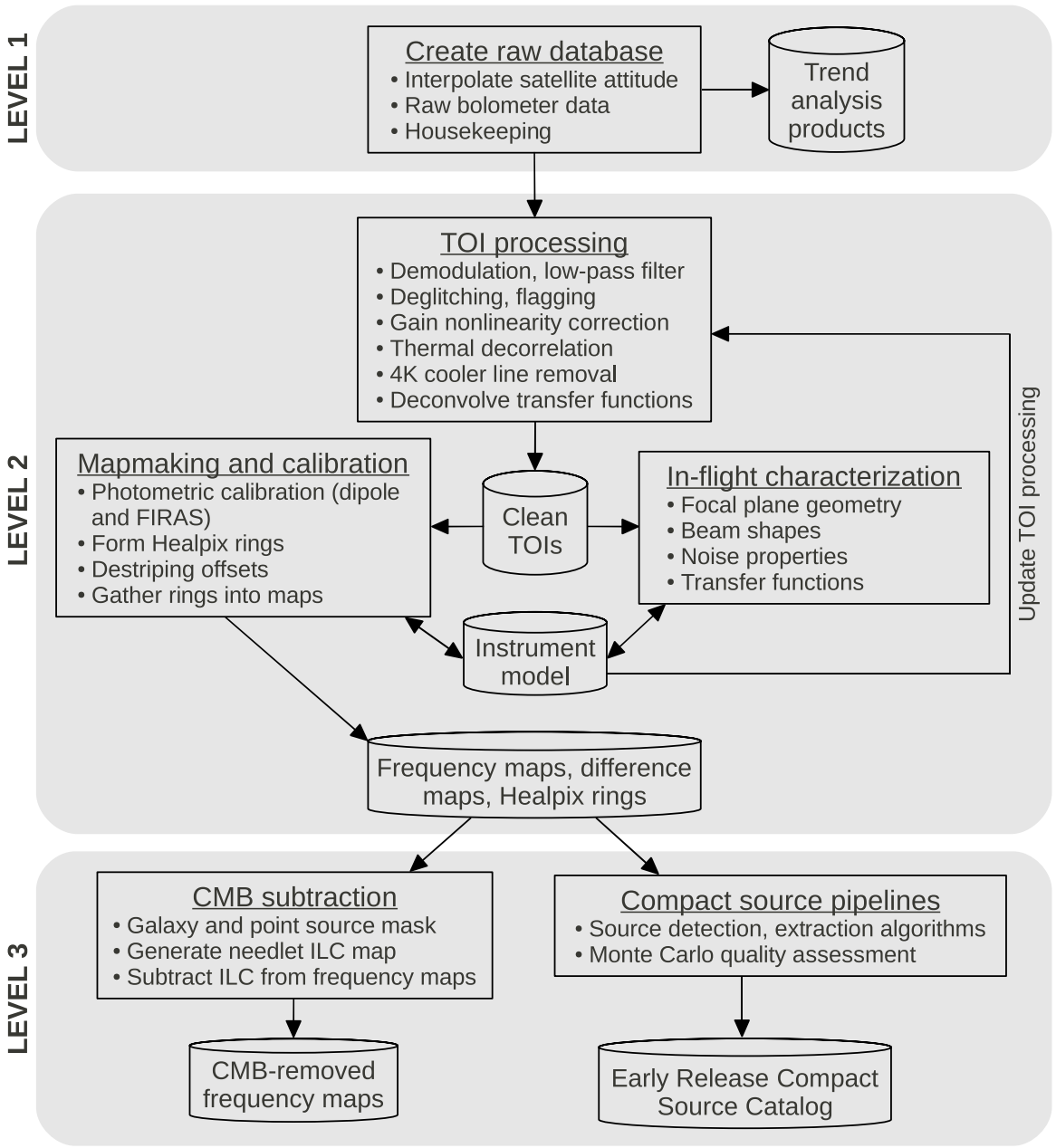


Figure 1.16: Overview of the HFI data processing pipeline. Image credit: ?.

The TOI chunk corresponding to a typical ring is shown in Fig 1.13. The raw modulated data as received via telemetry is unusable and plagued with glitches which completely dominate the signal dynamic range. After glitches and unstable pointing zone are removed there are approximately 350 000 science data samples left. The modulation caused by the solar dipole (see Sec. 1.1.2) is visible as a sinusoid wave and the spike is the brighter Galactic plane. As the detector signal is modulated (AC biasing), samples from negative parity appear mirrored.

An intermediate data representation before projection onto map is the Phased Binned Ring (PBR). An example is shown in Fig. 1.14 where the science data samples are binned with their corresponding phase in the scanned sky circle. There are about 35 sky circles averaged to produce one PBR, thus with this representation, the noise is significantly reduced. PBRs are typically used for systematics characterization and signal removal over the TOI allowing analysis of noise.

A second type of time ordered data received by telemetry from the spacecraft is *housekeeping* data. They sample at a low frequency the status of subsystems such as temperature, pressure sensors, REU setup, etc.

1.3.2 Production of *HFI* sky maps

A detailed description of data processing up to the production of maps is given in **??**. A schematic view of the *HFI* pipeline is shown in Fig. 1.16, the main steps are:

- TOI processing which includes ADC nonlinearity correction, demodulation, removal of glitches, removal of known systematics like 4 K lines or detectors nonlinearity, conversion from digital to physical units, time transfer function deconvolution and temporal noise estimation;
- pointing and beams reconstruction;
- mapmaking and photometric calibration (Planck Collaboration VIII 2016), projection to maps in the `HEALPix` (**?**) spherical pixelisation format (the `HEALPix` web homepage[11] is also a good starting point). An important step in the mapmaking process is the *destriping* stage which consist of removing the monopole value for each ring projected on the `HEALPix` sphere. The destriping is particularly sensitive to gain variations and ADC nonlinearity, because it causes spurious differences across scans over the same area of the sky.

Additional steps include the validation of the maps through angular power spectra and overall HFI validation with tests and simulations.

Examples of *Planck* sky maps are shown in Fig. 1.15. Before removal, the solar dipole is the dominant feature on the sky map. After removal of both the solar dipole and systematic effects the CMB anisotropies are visible along with the Galactic plane emission. It must be noted that any systematics preventing accurate destriping would create visible scars on the final CMB map as covered in Chap. 7.

Sky maps at each frequency are available on the *Planck* Legacy Archive[12]. Derived products such as the CMB maps in intensity and polarization are produced after removal of diffuse foregrounds emission. For this operation called *component separation* (**??**), four different algorithms are used (`Commander`, `NILC`, `SMICA` and `SEVEM`).

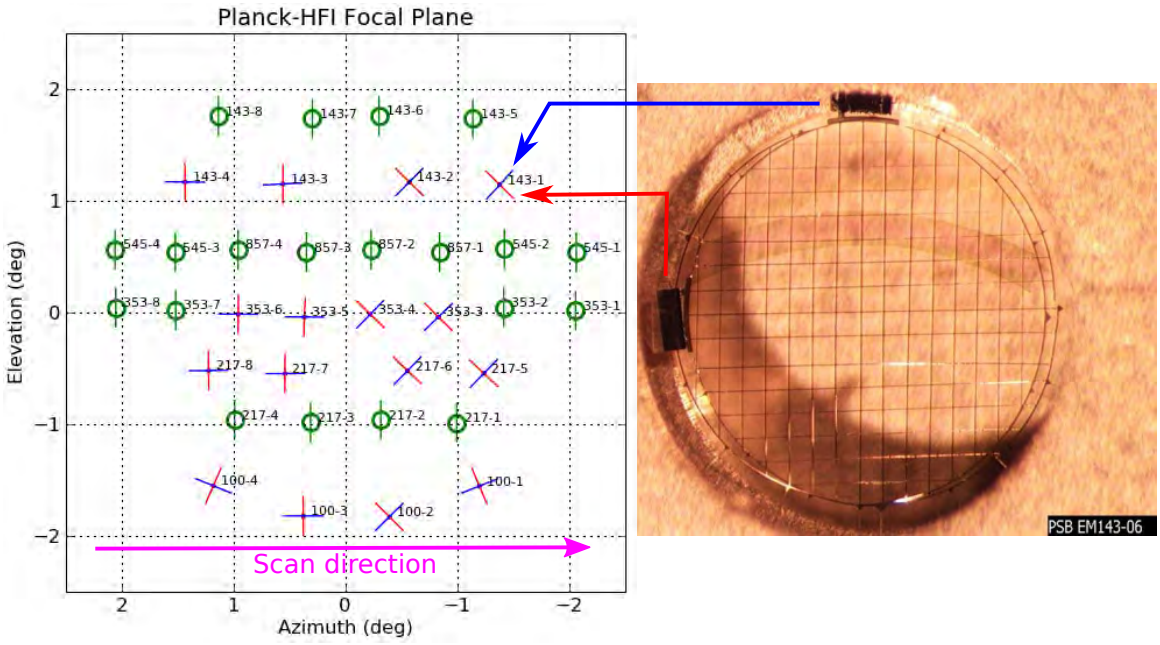


Figure 1.17: Example of an HFI PSB. The two thermometers (right) for each polarization orientation are visible at the extremities of the two superposed grids. An example of orientation of the two polarization directions is shown on the *HFI* focal plane (left).
Image credit: *Left*: ESA/PLA `https://wiki.cosmos.esa.int/planckpla2015` *Right*: ESA `http://public.planck.fr`

1.4 Polarization-Sensitive bolometers (PSB) of *HFI*

HFI is capable of measuring linear polarization, given that 16 out of the 36 optical horns of the *HFI* focal plane holds pairs of PSBs. The absorber of a PSB is a rectangular grid metallized in one direction. Electrical fields parallel to this direction develop currents and then deposit some power in the grid, while perpendicular electrical fields propagate through the grid without significant interaction (**??**). The second PSB of each pair has its orientation orthogonal to the first one. In Fig. 1.17 the rectangular grid of an PSB without its housing is visible, and the thermometers for each polarization direction are on the side, non-illuminated in the final assembly.

In practice linear polarization has to be measured in three directions for full determination of the polarization angle. This is the reason why PSBs come associated with another pair on the same scanning line with a relative orientation of 45° as shown in Fig. 1.17. The *Stokes* parameters (I, Q, U) for two perfectly calibrated PSB pairs with 45° relative orientation are given in the local frame by the relation

$$\begin{aligned} I &= |s_{1a}|^2 + |s_{1b}|^2 = |s_{2a}|^2 + |s_{2b}|^2 \\ Q &= |s_{1a}|^2 - |s_{1b}|^2 \\ U &= |s_{2a}|^2 - |s_{2b}|^2 \,, \end{aligned} \tag{1.4}$$

where (s_{1a}, s_{1b}) is the intensity signal measured by the first pair of PSBs oriented along the axis of a reference Cartesian basis $(\vec{\boldsymbol{x}}, \vec{\boldsymbol{y}})$, and (s_{2a}, s_{2b}) is the intensity signal measured by the second pair of PSBs. The fraction p of light intensity polarized linearly and the

[11] See: `http://healpix.jpl.nasa.gov/index.shtml`
[12] See `http://pla.esac.esa.int/pla/#home`

polarization angle ψ are given by

$$p = \frac{\sqrt{Q^2 + U^2}}{I}$$
$$\psi = \frac{1}{2} \arctan\left(\frac{U}{Q}\right) . \tag{1.5}$$

With real detectors, inter calibration is not perfect, which leads to leakage of a fraction of I into Q and U thus generating spurious polarization signal. A systematic effect such as ADC nonlinearity causes gain variations as a function of signal intensity (and time to a lower extent), which results into complex leakage of the signal into Q and U.

1.5 *Planck* results

The *Planck* spacecraft operated successfully for 4 years with HFI undertaking observations for 30 months before the depletion of the 0.1 K stage Helium supply. During this period *HFI* completed five full sky surveys in intensity and polarization. *HFI* has an angular resolution up to three times better than its predecessor WMAP, and a sensitivity about one order of magnitude better with nearly photon noise limited detectors for the CMB channels. The nine frequency bands have been a major feature in the improvement of component separation and for Galactic science, with the maps from the 2015 release shown in Fig. 1.18.

1.5.1 Study of our Galaxy

ADC nonlinearity affects the analysis of *HFI* data in general, and to some extent the study of the Galaxy magnetic field. This topic is introduced here, and the ADC effect will be further described in Chap. 7.

Polarization properties of dust grains

The interstellar Galactic medium is filled with dust grains which presents variety of composition, size and shape. Several mechanisms are involved which lead to polarized emission. The exact mechanisms of this emission are not fully understood and this is a field under active study, but two main elements are involved: rotation of grains and their alignment along magnetic field.

Dust grains spin is expected to be caused dominantly by radiative torque alignment (**?**). Grains with random irregular shape exhibit some amount of helicity. Under the anisotropic flux of star light at visible and UV wavelength, the radiative pressure cause helical grains to spin. The small magnetic dipole developed by the spinning grains tends to align them along ambient magnetic field lines with their spin axis orthogonal to the larger axis of the grain.

The aligned dust grain in ambient magnetic field can produce polarized light via two identified mechanisms. First by selective absorption at near UV and visible wavelengths (**?**), in this case the polarized emission is oriented along the magnetic field lines. Secondly by thermal emission at wavelength related to the interstellar dust temperature which is about 20 K. It is this second case which is of interest with respect to the *HFI* observation frequencies.

Dust as a tracer of Galactic magnetic field

Our Galaxy has a coherent magnetic field on large scales, thus the study of polarized emission from dust grains is a unique tool for its study. Thermal dust emission over the full sky is described in Planck Collaboration Int. XIX (2015). *Planck* is the first experiment to achieve

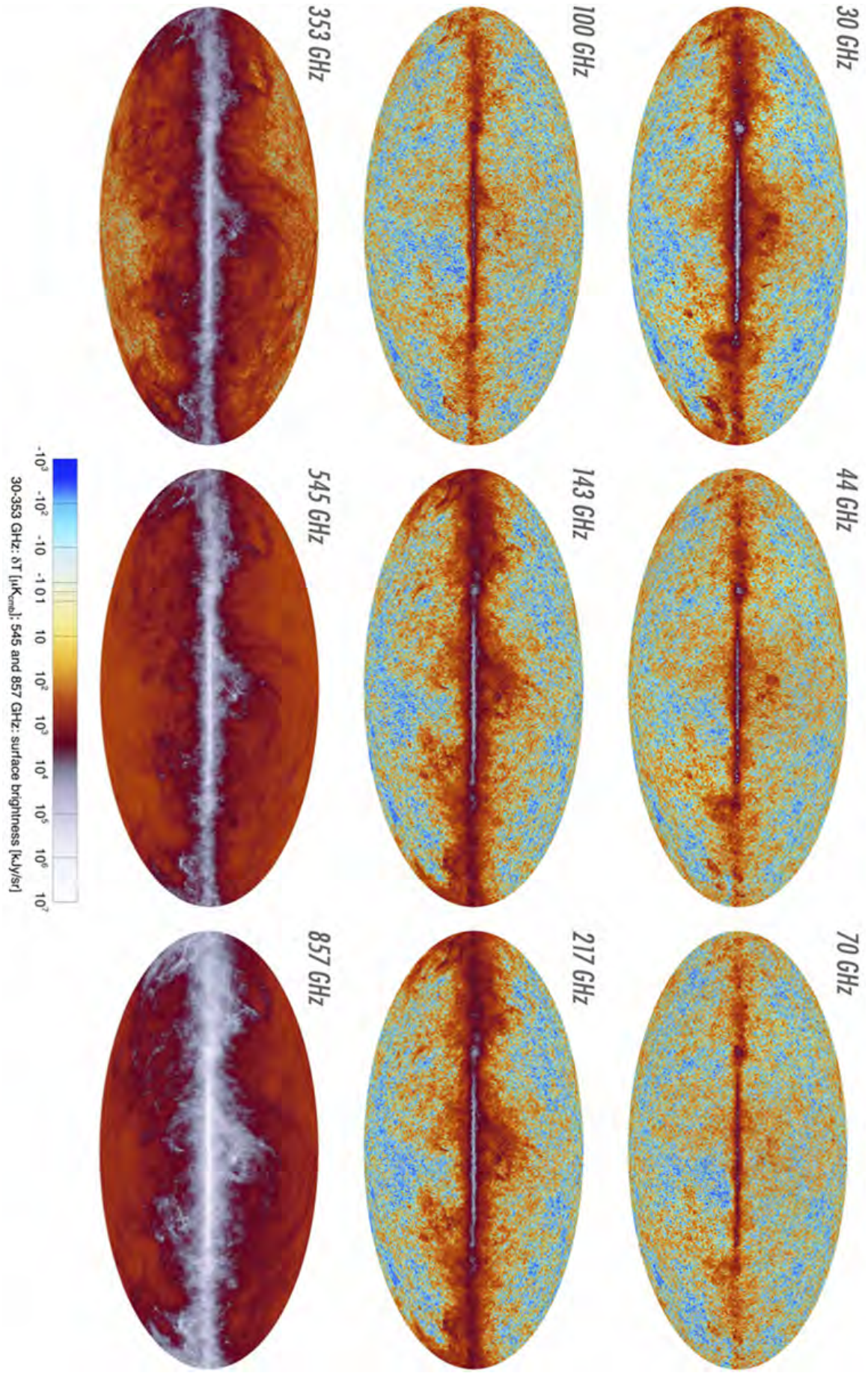


Figure 1.18: The nine Planck frequency maps in intensity show the broad frequency response of the individual channels. The color scale, based on inversion of the function $y = 10^x - 10^{-x}$, is tailored to show the full dynamic range of the maps. Image credit: **?**

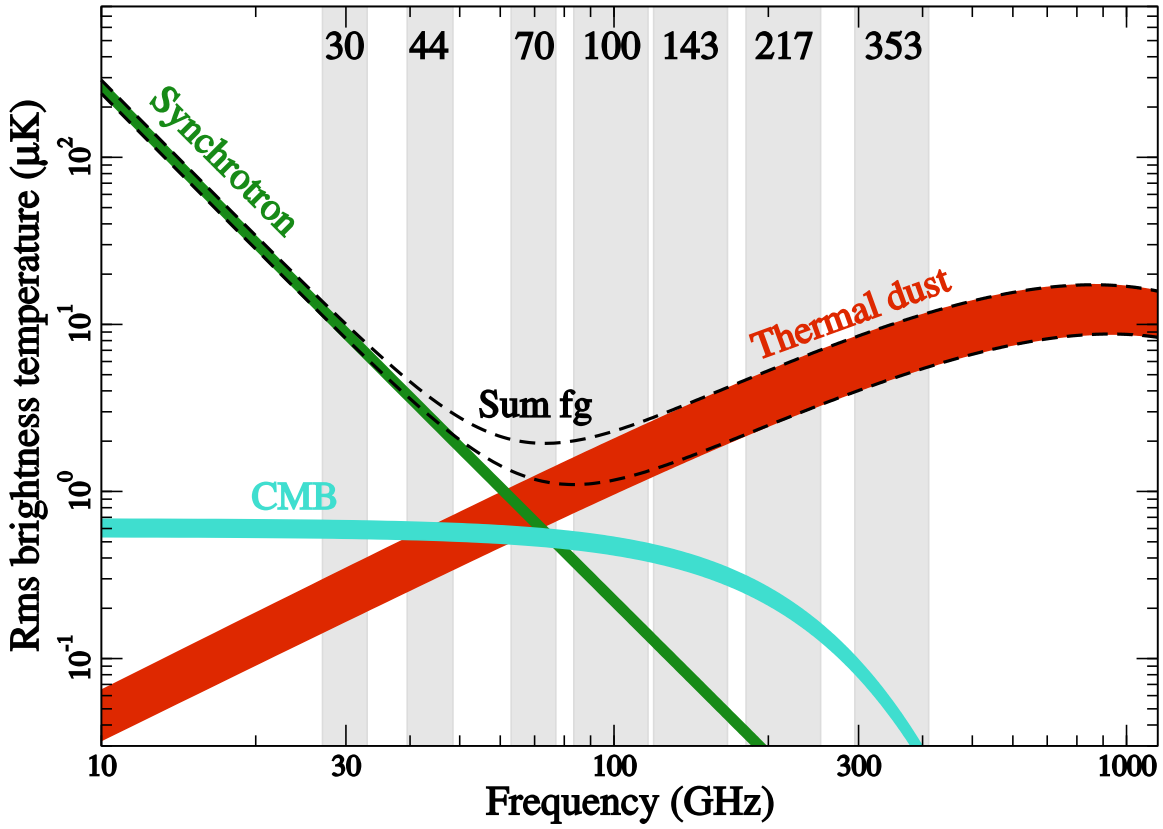


Figure 1.19: Polarized emission of foreground components in the *Planck* polarized frequency bands. Dashed lines represent the foreground brightness temperature RMS estimated using 73% and 93% of the sky respectively. Image credit: **?**

full sky coverage of dust at sub-mm wavelengths. The characterization of dust emission is also motivated by component separation required for the subtraction of foregrounds masking the CMB emission. Fig. 1.19 shows the polarized *Planck* observation frequencies and the related foreground emission, emphasizing the frequency range over which the polarized emission from thermal dust is significant. The 353 GHz frequency is the most sensitive frequency to thermal dust polarized emission, but can be significantly impacted by ADC nonlinearity, this is covered in Chap. 7.

Moreover the polarized Galactic emission generates B-modes which competes with the primordial anisotropies emissions. As of today, primordial B-modes remain undetected with the recent *Planck* and BICEP2 joint analysis indicating that an exquisite knowledge of Galactic polarized foregrounds is necessary to access the primordial B-modes (**?**).

The frequency coverage of *HFI* in the thermal dust emission bands, mainly at 353 GHz, allowed the construction of a new and unique map of the Galactic magnetic field as shown in Fig. 1.20. These measurements at 353 GHz are not perturbed by Faraday rotation thus offer a rather "clean" vision of the large scale magnetic field which is shown in Fig. 1.20.

1.5.2 The Sunyaev-Zeldovich (SZ) effect

The SZ effect (**?**), named by its discoverers, results from the transfer of energy from the intra-cluster hot gas of free electrons to CMB photons by the inverse Compton scattering effect. It is often referred to as thermal SZ or tSZ to distinguish it from the second order kinematic SZ effect or kSZ (**?**) due to bulk motions of clusters.

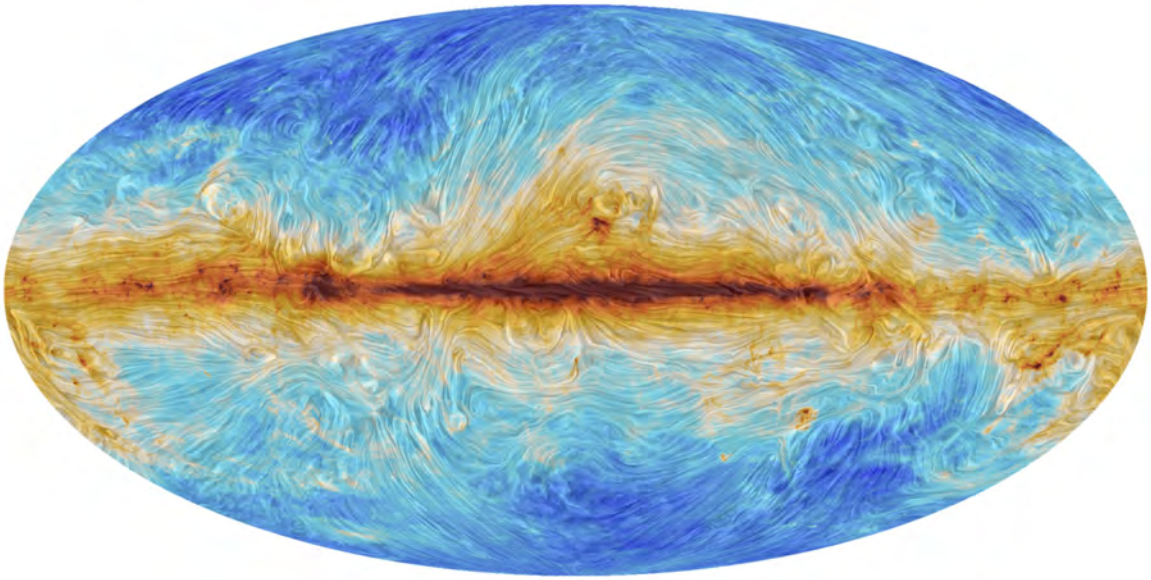


Figure 1.20: All-sky view in Mollweide projection of the angle of polarization at 353 GHz, rotated by 90° to indicate the direction of the Galactic magnetic field projected on the plane of the sky. The colors represent intensity, dominated at this frequency by thermal dust emission. The "drapery" pattern was obtained by applying the line integral convolution (LIC; **?**). Image credit: **?**

In practice a small fraction of the CMB photons "heated" by the energetic electrons are shifted to higher frequencies. This results in a characteristic frequency distribution of the effect as shown in Fig. 1.21. At frequencies below 217 GHz cold spots appear on the CMB at the location of the cluster, while at frequencies above 217 GHz it is hot spots respectively.

For quantitative analysis, the spectral distortion of the CMB expressed as a temperature change $\Delta T_{\rm SZ}$ at dimensionless frequency $x = h\nu/k_{\rm B}T_{\rm CMB}$ has the form

$$\frac{\Delta T_{\rm SZ}}{T_{\rm CMB}} = f(x)y = f(x)\int_L n_e \frac{k_{\rm B}T_e}{m_e c^2}\sigma_{\rm T} d\ell \tag{1.6}$$

where y is the Compton y-parameter, n_e is the electron number density, $\sigma_{\rm T}$ is the Thompson cross-section, T_e is the electron temperature, $k_{\rm B}$ is the Boltzmann constant, m_e is the electron mass, c is the speed of light, the integration is done along the line of sight L, and $f(x)$ is the frequency dependent function of the tSZ effect as shown in the bottom panel of Fig. 1.21.

The value of y scales with the cluster optical depth τ_e and electron temperature T_e. This is of particular interest, because electrons density is a tracer of baryon gas density which in turn is a probe of the total cluster mass. Although τ_e is degenerate with T_e, it is possible under the simplifying assumption that the gas is isothermal and using a model for the cluster gas (e.g. **?**) to fit the cluster mass from tSZ observations. Moreover combining tSZ with X-rays surface brightness observations which are also a tracer of clusters mass, provides even more constraints. Given that the tSZ effect does not depend on the redshift z of the cluster, it can provide a powerful tool for the study of large scale structures in the young Universe.

Clusters of galaxies

Clusters of galaxies are the largest gravitationally-collapsed structures in the Universe. The distribution of cluster abundance as a function of their mass, or the cluster mass function, is expected to reflect the primordial density fluctuations and provide direct access to the

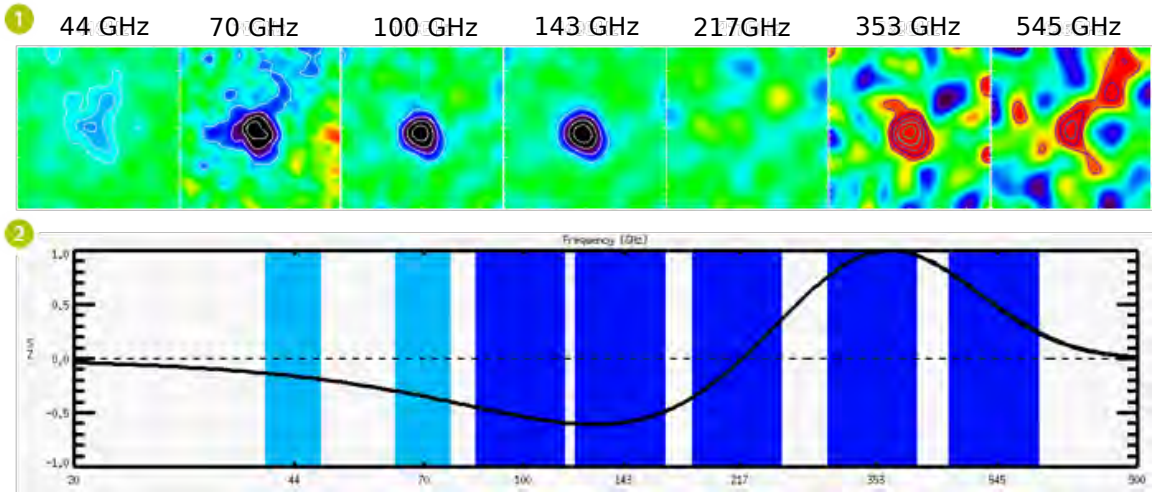


Figure 1.21: Visualization of SZ effects for the cluster Abell 2319. *Top*: The cluster as observed by *Planck*, blue is colder and red is hotter. *Bottom*: The predicted variation of flux (normalized) for each frequency band. Image credit: ESA `http://public.planck.fr/outils/astrophysique/effet-sz`

parameter σ_8[13] of the ΛCDM model (e.g. **?**). Also the evolution of σ_8 in time is driven by the the value of the matter density parameter Ω_m and to a lower extent by dark energy. This is why tSZ surveys are of great interest for cosmology, because they provide valuable constraints for σ_8, Ω_m and Ω_Λ[14] as described in **?**.

The *Planck* Consortium has published a full-sky map of the Compton y-parameter hereafter called the y-map. It is produced with a custom component separation method which is described in Planck Collaboration XXII (2016); **?**. The *Planck* y-map resolution is $5'$ to $10'$. It is low compared to SZ ground surveys from the South Pole Telescope (**?**) or the Atacama Cosmology Telescope (**?**). However the full sky coverage of *Planck* results in a larger statistical sample and allows the detection of new clusters as in the *Planck* catalogue of SZ sources (**?**) which contains 1207 confirmed clusters in the 2015 release. As an example, cluster Abel 2319 as seen in the y-map is shown in Fig. 1.22. The values of the y-parameter are typically a few 10^{-6}.

The study of the cluster mass function is complemented by the measurement of the tSZ angular power spectrum (**?**). In this case, no explicit measurement of cluster masses is required which is well suited to the *Planck* y-map with large statistics and low angular resolution.

Matter-Antimatter annihilation

Another domain of application of the y-map is the study of the antimatter (AM) annihilation with matter at the epoch of recombination. The Universe today is dominated by matter while both matter and AM are expected to have been produced in equal proportions during inflation. The disappearance of AM is still not fully understood today and in the event that some clouds composed only of AM would survive at the epoch of recombination, the thermal energy released by their annihilation with matter at the matter-AM interface is supposed to imprint characteristic scars on the y-map. This topic is specifically covered in Chap 8

[13] Variance in matter amplitude fluctuations at $8h^{-1}$ Mpc scale, where $h = H_0/100\,\mathrm{km\,s^{-1}}$ is the reduced Hubble constant and H_0 is the Hubble constant. It corresponds to $1'$ scale in the CMB angular power spectra

[14] Dark energy density parameter

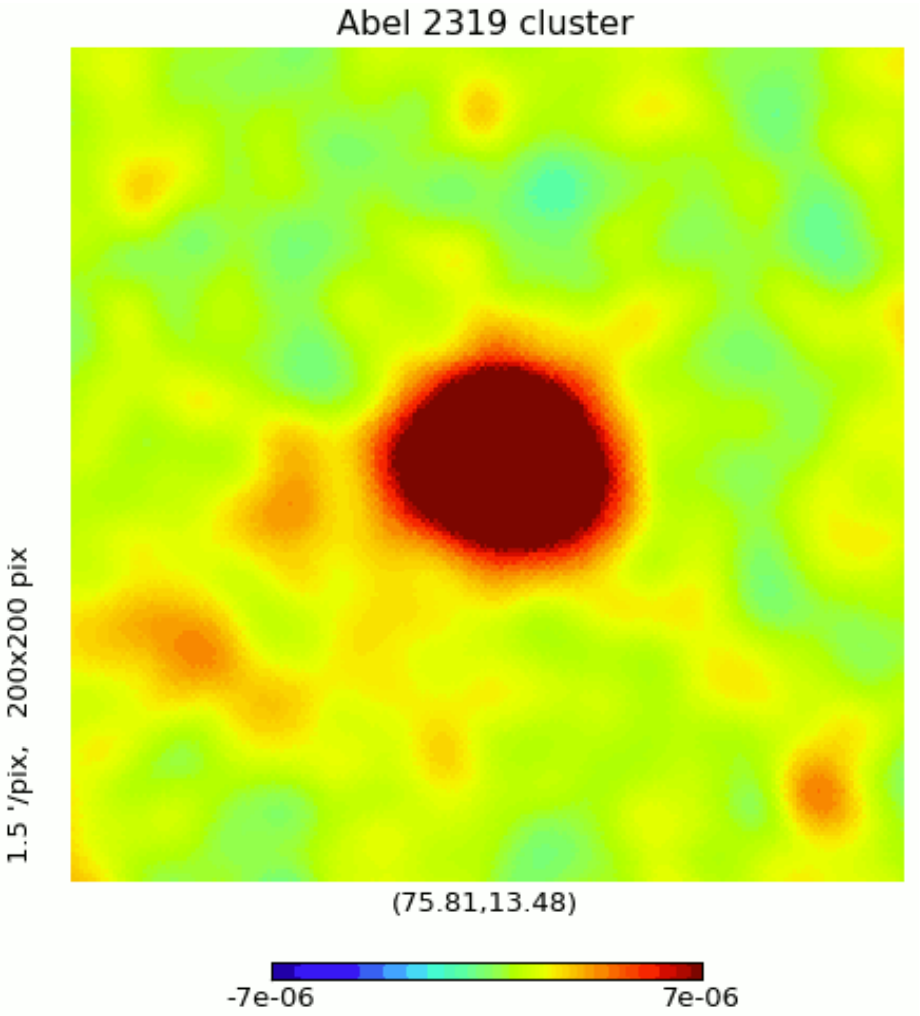


Figure 1.22: The cluster Abell 2319 as seen in the *Planck* y-map produced with `NILC`, it is the same cluster as the one shown in Fig. 1.21 in multiple frequencies. The color scale of the y-parameter is linear.

where a preliminary analysis is performed on the impact of *Planck* scanning strategy on this detection.

1.5.3 Cosmology

The *Planck* 2015 data release has strengthened the ΛCDM cosmological model with an increased accuracy on the fitted model parameters, favoring a flat Universe with a spatial curvature Ω_K compatible with zero. The CMB map obtained after components separation is shown in Fig. 1.23.

One of the expected outcome of *Planck*, the detection of the gravitational waves signature on the CMB, has been scooped by the direct proof of their existence with the direct observation of a binary black hole merger made by the Earth-based detector LIGO. The collaboration with the BICEP2/KECK team has been fruitful and set an upper limit on the tensor to scalar mode ration of $r < 0.12$. Thanks to its full sky coverage of polarized thermal dust emission at 353 GHz, *Planck* provided strong arguments in favor of the observation of B-modes produced by dust instead of primordial CMB anisotropies. These results opens the

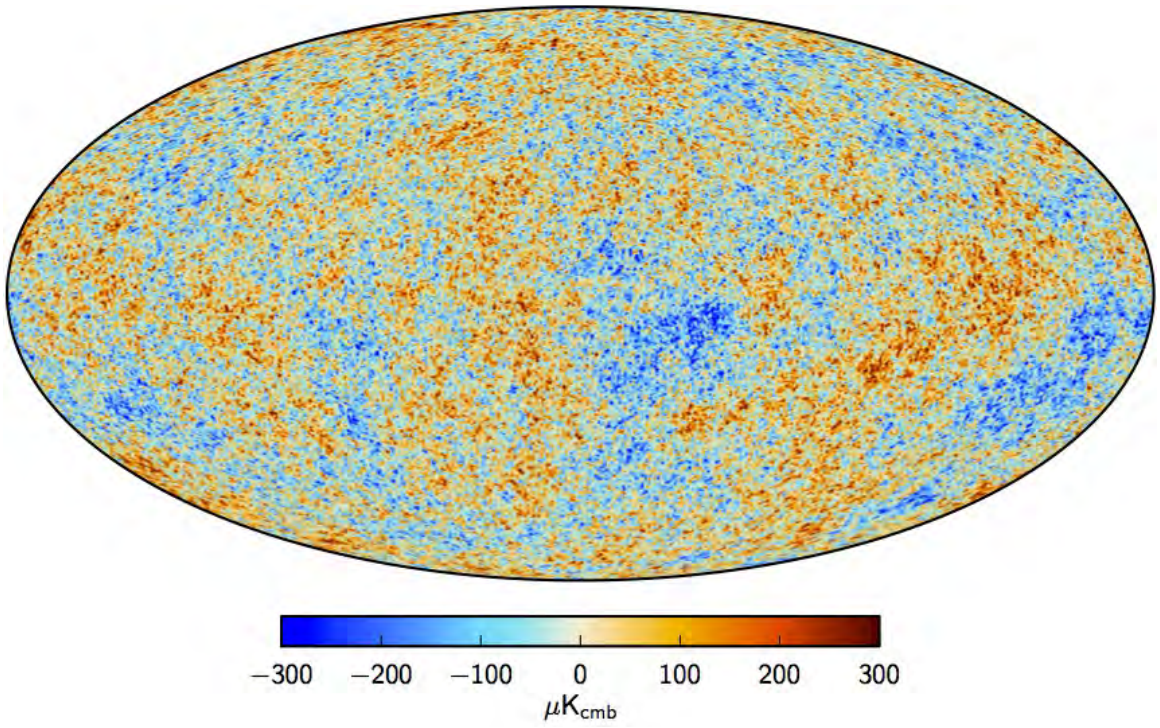


Figure 1.23: CMB intensity map in Mollweide projection at 5′ resolution derived from the joint baseline analysis of Planck, WMAP, and Haslam 408 MHz observations (**?**). A small strip of the Galactic plane, covering 1.6% of the sky, is filled with inpainting (a constrained realization that has the same statistical properties as the rest of the sky). Image credit: **?**

door for more accurate studies of the dust and foregrounds polarization properties. In particular the *HFI* full sky mapping of thermal dust provides invaluable information on column density, thus indicating which are the best locations for future CMB observations.

New constraints on the epoch of reionization have been set with a lower optical depth $\tau = 0.066 \pm 0.016$ constrained with *HFI* plus *LFI* data, and a later occurrence at $z_{\rm re} = 8.8^{+1.7}_{-1.4}$. There are still some uncertainties due to the difficulties in reducing large angular scale systematic effects (predominantly the ADC nonlinearity), but these results are in good agreement with the WMAP 9-year polarization data cleaned using the *Planck* 353 GHz polarization (**?**). The work on large angular scale systematic effects is still ongoing and updated estimations for τ and $z_{\rm re}$ are in preparation (**??**).

1.6 Lessons learned

The complex technical architecture of *Planck* and more specifically the cooling system has been a notable success despite the existence of at least one single points of failure (e.g. the 0.1 K stage). A complete technical review is available online[15] for the various subsystems of the spacecraft. Several aspect that have been dealt with in this thesis work are emphasized here:

- ADC NONLINEARITY: In the case of *HFI* it appeared critical to have a full characterization of the nonlinearity. This requires a very accurate description of the transition points of each ADC code in the nominal mission setup (for details on ADC architecture see Sec. 4.2). Such an issue could become more and more critical with successive

[15] `http://www.cosmos.esa.int/documents/387566/1178903/Planck+Lessons+Learned+Phase+1+report/bbfcb160-daf8-4d28-864b-a3e13393fe9b`

generations of instruments and increasing sensitivity. Moreover, the dynamic of the acquired signal was covering a voltage range corresponding to only ≈ 10 codes. This range should be widened significantly to lower the negative impact of unavoidable localized defects. Some relevant options are provided in Sec. 4.9.

- OVERSAMPLING AND SUMMATION: In *HFI*, the *fast samples* summation is in theory a very good option for improving the SNR of science data. However in practice, in the presence of any nonlinear systematic effect on the summed signal (such as the ADC but not limited to it), this destroys vital information for correction of the effect. Hence oversampling should be either replaced by analog integration or ADC nonlinearity should be corrected for in-flight before summation. These options are discussed further in Sec. 4.9.

- 4 K COOLER EMI: the electronic vibration damping system of the *HFI* 4 K stage mechanical compressor worked perfectly, but EMI shielding of this kind of system has to be strengthened in future mission. Some replacement solutions are promising like adiabatic demagnetization refrigeration (ADR). The synchronization of the compressor with DPU acquisition revealed to be a very good idea for removing the spurious signal propagated on science data. However the EMI leaks (designed farther as 4 K lines) in conjunction with ADC nonlinearity plus *fast samples* summation produce nonlinear effects that are extremely difficult to deal with. They remain one of the dominant source of residuals at large angular scale limiting the accuracy of scientific results for $\ell < 10$ and generating spurious signals around $\ell = 1800$. Furthermore some short timescale variations as generated by the sorption cooler bed switching also produces residuals that are difficult to deal with. This is remain an issue that can be improved for *HFI* ADC nonlinearity correction as presented in Sec. 6.5.

- COSMIC RAY SUSCEPTIBILITY: One of the surprises in the *Planck* mission was the higher than expected rate of glitches seen by *HFI* detectors. This is partly due to the coincidence between the spacecraft launch and a solar activity minima, and also to the high susceptibility of the detectors which is due to the silicon die part (**?**). The long glitches generated by hits on the silicon die have a slow thermal relaxation which, if not appropriately removed, produce non-Gaussian noise in the science data. Hence future experiments with detectors sensitive to thermal relaxation such as Transition Edge Sensor (TES) arrays should improve heat transport to the housing.

- THERMAL RESPONSE OF BOLOMETERS: Under AC biasing, bolometers have a significantly different time response than under DC biasing. Albeit behaving like an order one a low pass filter at first order they exhibit an enhanced response at low frequency which is characteristic of AC biasing. This behavior is problematic in the *HFI* case because the photometric calibration relies on the orbital dipole being visible at the very low frequency of 16 mHz. An analytical model has been developed for AC biasing of bolometers, as presented in Chap. 2 that can be used to produce valuable constraints for the low frequency part of the time transfer function.

Part I

Readout Response

Chapter 2

time transfer function analytical model

> "Oh dear! Oh dear! I shall be too late!"
>
> *White Rabbit*

The time transfer function of the *HFI* readout electronics is a critical component of the science data processing, and one of the major known source of systematic effects. Hence a very high calibration precision is required. In **?** that will be referenced hereafter as the pre-launch paper, the target was set to an accuracy level below 0.2% within the [10 mHz, 70 Hz] range. However ground calibration, albeit useful for modeling the time transfer function, is not sufficient to reach this level of accuracy in-flight where the optical load and bias parameters of the detectors are significantly different to the setup used during the ground calibration campaigns. Since the spacecraft launch, the time transfer function is approximated with an empirical model using a sum of low pass filters to describe the thermal response of the detectors. This approach presents a major issue with the calibration because the empirical model is not constraining enough. It happens that the work on ADC nonlinearity correction needs also an accurate knowledge of the time transfer function (See Chap. 5) and reveals a new and unexploited constraint for its characterization. For these reasons a work has been done to produce an analytical model of the time transfer function in frequency domain. This chapter presents this work which resulted in an article (Sauvé and Montier 2016) published in *Experimental Astronomy*, and called hereafter the AC paper (provided in Appendix A).

First the time transfer function definition and characterization within *HFI* is detailed in Sec. 2.1. Then the content of the AC paper is discussed in Sec. 2.2, and a direct application to in-flight data is presented in Sec. 2.3. Finally some perspectives are provided for an evolution of the thermal model of the simple bolometer model heat dissipation in Sec. 2.4 to meet the *HFI* accuracy goal.

2.1 Context

2.1.1 Definition of the time transfer function

In *HFI* the time transfer function, also called temporal transfer function, describes the readout electronic response for demodulated science data to an input optical excitation. It is a complex function of frequency describing the response in amplitude and phase shift of an input signal.

The readout electronic response described by the time transfer function can be decomposed in two functional subsystems which are:

- THE THERMAL RESPONSE: it is the detector time response to the signal of the incoming sky power;
- THE ELECTRONIC RESPONSE: it includes the readout electronics filtering (bandpass filter) and the DPU digital processing to produce science data samples.

This model is described in details in Sec. 2 of **?**. Following the nomenclature of this paper, the time transfer function can be written

$$TF(\omega) = \underbrace{F(\omega)}_{\text{thermal}} \underbrace{H'(\omega)}_{\text{electronic}} , \tag{2.1}$$

where $F(\omega)$ is the thermal response and $H'(\omega)$ is the electronic response[1].

The sensitive part in this decomposition is the presence of a stray capacitance due to the length of the cables between the bolometer and the Pre-Amplifier Unit (PAU, see Sec. 1.2.6). It is included in $H'(\omega)$ in **?** but in practice it participates significantly to the electro-thermal coupling of the detector and should be described within $F(\omega)$ as detailed in the AC paper.

2.1.2 Characterization

While the electronic part of the time transfer function has been very well characterized on-ground and rely on well known electronic components, the thermal part is more complex because of the AC biasing of bolometers which is not theoretically as well known as DC biasing. Hence the characterization of the time transfer function described hereafter focuses in practice on the thermal response.

Thermal response model

Under DC bias the time response of bolometers to an optical excitation can be represented by a first order low pass filter (**?**) . This theory has been extended to AC biasing with a square current that mimics the DC biasing, as described in Sec. 4 of the pre-launch paper. In practice the response in AC biasing is significantly different, and furthermore impacted by the stray capacitance. These aspects are taken into account and solved for in the AC paper.

In the case of *HFI*, the thermal part $F(\omega)$ is based on an empirical model using a sum of first order low pass filters reading

$$F(\omega) = \sum_{i=1}^{7} \frac{a_i}{1 + j\omega\tau_i} . \tag{2.2}$$

Calibration of the *HFI* thermal response

The main issue with the time transfer function is the estimation of the (a_i, τ_i) parameters of $F(\omega)$, given the strong accuracy requirements of *HFI*. Moreover a Low Frequency Excess Response (LFER) has been identified during the ground calibration campaigns, and estimated at about 1% level in amplitude in the pre-launch paper. The LFER is particularly relevant with the CMB channels channels which are calibrated with the dipole appearing at 16 mHz, but for such low frequencies there is no easy way to make calibration in-flight and the ground calibration data revealed insufficient.

In 2015 the (a_i, τ_i) parameters are estimated from heterogeneous sources (see Sec. 3.4 in **?**). The shortest time constant parameters are estimated with planets crossings, while the

[1] The values of the electronic parameters for $H'(\omega)$ are provided in the *Planck* explanatory supplement, see `https://wiki.cosmos.esa.int/planckpla2015/index.php/HFI_time_response_model`

slowest ones are estimated using the time constants fitted on short glitches and the apparent time delay between the measured dipole and its estimated location. An example of thermal response $F(\omega)$ annotated with these constraints is presented in Fig. 2.1.

This empirical approach required a tremendous amount of work of many people from the *Planck* Collaboration in order to provide enough constraints on the time transfer function to reach a level better than 0.3% accuracy for frequencies below 2 Hz (inferred by comparison of the relative amplitudes of the first peak of CMB anisotropies between frequency bands in the 2015 data release).

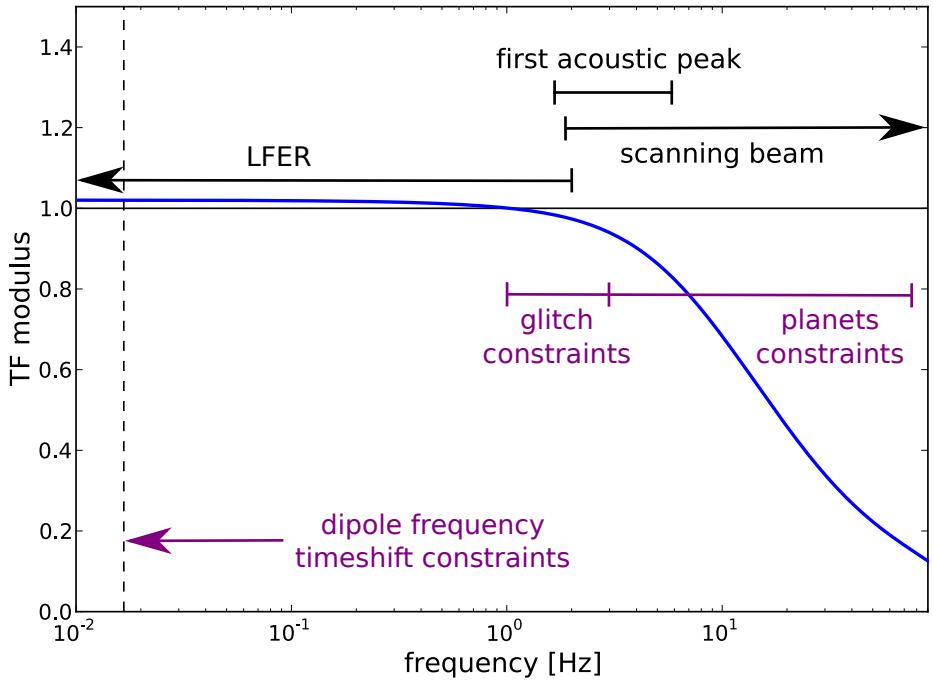


Figure 2.1: Thermal response of the time transfer function for channel `00_100-1a`. The modulus is normalized at 1 Hz to emphasize the LFER.

Degeneracy with the scanning beam

The scanning beam (see Sec. 1.2.4) is the angular response of the instrument after deconvolution by the time transfer function.

In practice, for the beam reconstruction, the time transfer function is considered as a regularization function and the scanning beam can absorb the residual error from the true time transfer function. Thus the scanning beam and the time transfer function are both degenerated. However this approach has some limitations because the beam window angular size describes only features corresponding to frequencies higher than 2 Hz on the time transfer function. It means that errors on the time transfer function for $\tau_i > 0.5$ s cannot be absorbed by the scanning beam, as it can be seen in Fig. 2.2 where the time transfer function errors leave a long asymmetric residual tail on the beam map. Hence the reduction of low frequency errors on the time transfer function is critical for CMB analysis.

2.2 Discussion of the AC paper

2.2.1 Motivations

The work presented in the AC paper has the following motivations:

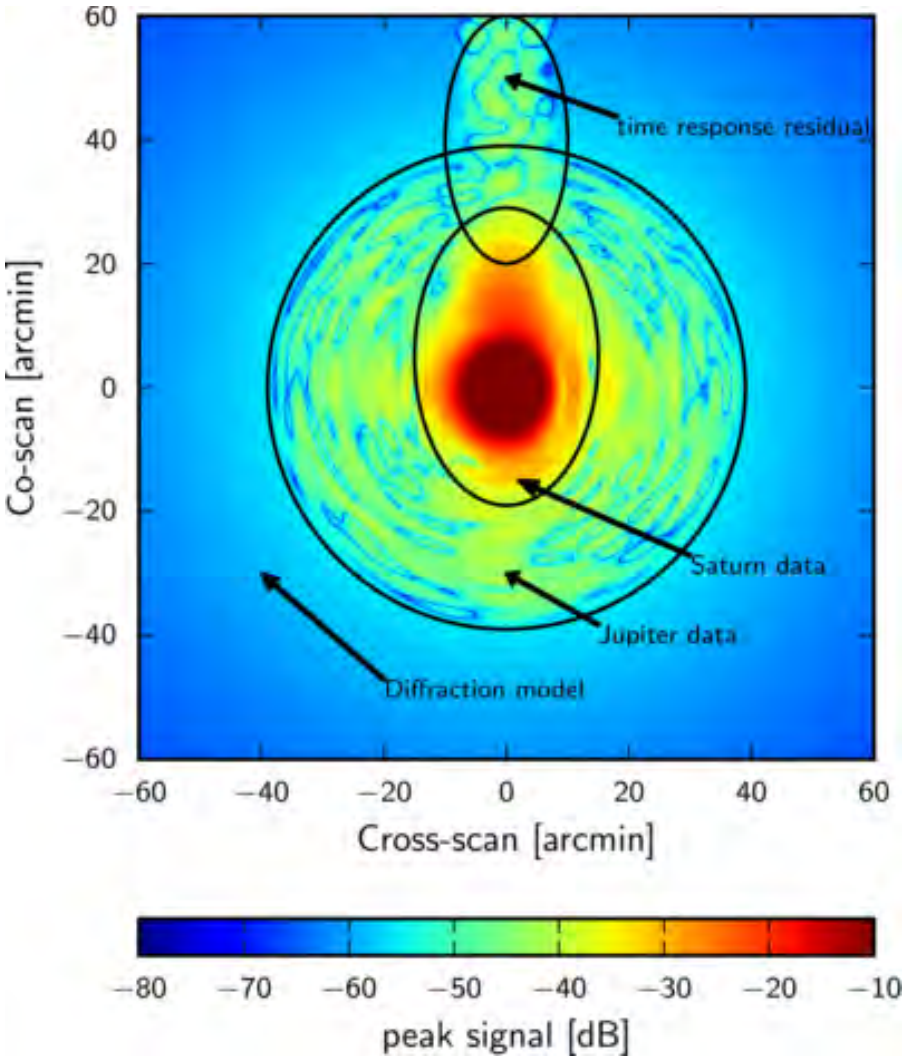


Figure 2.2: Scanning beam map for detector `42_143-6`. The upper tail is mostly due to LFER residuals. Image credit: **?**.

- the empirical model has some issues with calibration at low frequency in the LFER range and needs a high number of free thermal parameters to approximate the frequency response of the detectors in-flight. In particular there is a strong degeneracy between the (a_i, τ_i) parameters. Hence a more accurate model of the thermal response is expected to bring constraints for calibration and a reduced number of free parameters (12 to 16 per channel for the 2015 release).

- an analytical expression is suitable for data deconvolution, which is not the case with the integration tools that are mainly focused on simulations.

- the analysis of the ADC nonlinearity raised issues about the knowledge of the detector response at high frequency, in order to describe accurately the analog signal before digitization by the ADC. It also provides some insights on the detector response calibration.

Let detail a bit more the point about the detector response at high frequency which has unveiled a completely new and unused constraint for the calibration of the detector response. For ADC nonlinearity correction a first order model of the analog detector response before

digitization is needed. The detector response to the sky signal can be modeled at first order, at modulation level, using what is called the ***raw gain*** within *Planck* Collaboration, it is detailed in the AC paper and in Sec. 5.2. It was called the *steady state gain* in the AC paper for the general audience because it reflects more its nature. The interesting point about the *raw gain* is that it results from the electro-thermal equilibrium of the detector. Moreover this is an observable easy to measure from in-flight data. Expected outputs from the *raw gain* analysis are:

- a analytical model which can improve this parameter for the ADC nonlinearity correction;
- strong constraints for the calibration of the thermal response $F(\omega)$.

2.2.2 AC paper content

The main issue when building an analytical expression for AC biased bolometers comes from the product between differentials terms in coupled differential equations defining the electrical equilibrium in Eq 2.1 of AC paper and the thermal equilibrium in Eq. 2.2 of the AC paper. Steady state and optical excitation signals are considered separately, and the coupled differential equations are linearised for each case. Then the products of differential terms can be managed by using a custom splitting of the discrete frequency support for the considered signals. Finally the analytical expression of the time transfer function can be constructed with classic matrix based linear algebra.

The validation of the time transfer function analytical expression is compared to the output of the simulation tool `SEB`. The former is calculated in frequency domain and the later integrates the differential equations using finite differences.

A specific analysis is made for the detector response to slow optical signal variation. The analytical expression allows us to infer quantitative error estimation for the first order model of the detector response (using the *raw gain*). Hence the empirical approach used for ADC nonlinearity correction can be compared to the analytical model.

The analytical expression describes the signal at the output of the detector bias circuit, it is then reformulated to include the DPU integration which produce the *HFI* science data. This integrated version can be compared to the single low pass filter approximation used in the case of DC biasing.

In the Appendix B of the AC paper, the analytical expression is extended with a second heat capacity on the thermal path to the heat sink (and potentially to more components). This case is more realistic as confirmed by the glitch analysis (see Chap. 3).

2.2.3 Constraints for the LFER model

Instead of reformulating the conclusions of the AC paper here, their interpretations are reformulated below in the context of the LFER and ADC nonlinearity issues.

The integrated version of the time response of the simple bolometer model (Fig. 9 of the AC paper) differs significantly from the empirical model used to describe the LFER. While the later uses a sum $F(\omega)$ of first order low pass filters to fit the thermal response of the detector, it is clear that fitting the former with low pass filters will result in spurious long time constants. This result by itself demonstrates that LFER in the sense of the pre-launch paper is a natural behavior of AC biased bolometers within *HFI* readout. Hence a model with a correct description of the thermal dissipation within the detector is expected to be constraining for the time transfer function.

Moreover, the expression for the *raw gain* (expressed as $\overline{G}$ in Sec. 5.4 of the AC paper), shows that it results from an interpolation of the thermal response at optical frequency $\omega = 0$.

As a consequence it is a constraint for the low frequency part of the time transfer function (albeit having a high frequency description). However it is **model dependent**, meaning that it is also a tool for model tuning. This constraint at low frequency is important because it is exactly the part of the time transfer function which is the most difficult to calibrate in-flight.

2.3 Test of the analytical model against *HFI* data

This section describes a simple check of the performances of the analytical model presented in the AC paper. It is used to fit the *raw gain*, and the empirical time transfer function model used for the *HFI* 2015 data release.

2.3.1 Fit setups

Three setups are considered with the data of channel `00_100-1a` which are:

- SETUP 1: the one thermal component model is used to fit the *raw gain* only. The *raw gain* has been estimated with the best available ADC nonlinearity correction (INL version `GP41`);
- SETUP 2: the two thermal components model (described in the appendix of the AC paper) is used to fit the *raw gain* and the time transfer function of the 2015 data release. In order to have equivalent weights for both data sources, 80 data samples are used for the time transfer function in a logarithmic frequency range between 0.1 Hz and $f_{\mathrm{mod}} \times 0.95$. The *raw gain* error per sample is inferred from the symmetry error between both parities (see Sec. 5.2.3) with a value of 1.7×10^{-3} relative to the peak. For the time transfer function error, it was used the 0.3% error estimation inferred from the cross channels consistency as described in **?**, and the analytical model residuals are calculated relative to the 2015 time transfer function modulus of each data sample frequency.
- SETUP 3: the two thermal component model is used to fit the time transfer function only, with the same frequency range and number of points as for the setup 2.

There are 3 free parameters with one thermal component and five with two thermal components. For thermal parameters, starting from the nominal values C_0 and G_{s0} used in `SEB` (from in-flight and ground calibration), two free parameters are scaling factors f_{1C} applied on the bolometer heat capacity C_0, and f_{1G} on the thermal conductance parameter G_{s0} of the thermal link from which the heat flows out of the bolometer. For the two thermal components cases two more scaling parameters (f_{2C}, f_{2G}) are used to set the thermal parameters of the second thermal component with $C_{02} = f_{2C} \times C_0$ and $G_{s02} = f_{2G} \times G_{s0}$ respectively. Additionally a static gain parameter absorbs inaccuracy in the gain of the electronic readout chain, this parameter converges to 0.954 and is consistent with in-flight measurements errors. The fit performed with`mpfit`[2] converges in 5 to 25 iterations with a stopping criteria of 10^{-3} relative χ^2 variation, the quantitative fit results for the three setup are presented in Tab. 2.1. All other parameters as reported in Tab. 1 and Tab. 2 of the AC paper coming from on-ground and in-flight measurement and are left untouched.

	setup 1	**setup 2**	**setup 3**
***raw gain* relative err.**	0.0039	0.035	0.16
TF relative err.	N/A	0.0080	0.0035
$f_{0,C}$	0.443	1.856	2.248
$f_{0,G}$	0.863	1.037	1.090
$f_{1,C}$	N/A	11.47	50.12
$f_{1,G}$	N/A	3.701	9.063
$\tau_{\mathbf{eff}}$	0.00217 s	0.00807 s	0.00880 s
$\tau_{1\varphi}$	0.00337 s	0.01135 s	0.01343 s
$\tau_{2\varphi}$	N/A	54.8 s	29.4 s

Table 2.1: Results of the *raw gain* and 2015 time transfer function fit using the analytical model on channel `00_100-1a`. The time constant of the second thermal component is calculated with $\tau_{2\varphi} = G_{s02}/C_{02}$ by extrapolation of the single component formula.

2.3.2 Results

Setup 1

The results for setup 1 are shown in Fig. 2.3 and show a very good agreement between the *raw gain* data and the model. The RMS of the fit residuals for the *raw gain* is 2.3σ which is in line with a residual bias from the 4 K lines signal. This result is extremely encouraging because it means that this model can be used to improve the *raw gain* for ADC nonlinearity correction. However the time transfer function the result is not good and drifts at high frequency, and it is already known that the simple bolometer model with one time constant is not able to fit the LFER. However the new information here is that the steady state response to optical power variations or **the *raw gain* is well described by the analytical model**. The expectation here is that this is a strong model constraint, because a working model has to describe both the steady state response and the optical response for $\nu > 0$.

Setup 2

The results for setup 2 are shown in Fig. 2.4. With two thermal components the time transfer function fit is really good, with an agreement between the two transfer functions modulus within 2.7σ of the error estimated for the empirical model. However in this case the *raw gain* fit is not as good with a 3.5% relative error and the phase drift significantly between both models near the modulation frequency. It is possible that the phase drift and the *raw gain* mismatch are the expression of the same missing parts in the two thermal components version of the analytical model, or that the simplifications of the second thermal component thermal capacity are inaccurate. Anyway **the analytical model with two thermal component is very close to the in-flight data**, but tests with the deconvolution of *HFI* data are recommended to be performed only when the *raw gain* accuracy is within the same error range as for the setup 1 fit.

Setup 3

The setup 3 (see Fig. 2.4) is here to check the power of the analytical model on the in-flight time transfer function only. There is a minor improvement with the time transfer function

[2] `mpfit` perform Levenberg-Marquardt least-squares minimization, based on MINPACK-1. The version used is the python implementation available from `https://code.google.com/archive/p/astrolibpy/downloads`.

error which is about 1σ of the error estimated for the empirical model, but the phase drift is still significant and the *raw gain* is really distorted. Hence the model with two component cannot reproduce exactly the empirical model, it is expected that three thermal components, including the bolometer, could be enough to meet the data processing requirements.

It must be noticed that the time constant of the second thermal component is about 30 s (see Tab. 2.1), which is very close to a full rotation of the satellite (60 s). If we trust that the dipole time shift between detectors reported in **?** is really due to the thermal response of the detectors, then **the two thermal components version of the analytical model is consistent with a very slow time constant**.

2.4 Perspectives

The work presented in the AC paper is the result of a two years lower priority analysis. It has been concluded recently so that it was exploited only at minima. Hence it is presented below the expected roadmap to develop potential extensions to this analytical model.

2.4.1 ADC nonlinearity correction

A direct application is to use the single thermal component version as a model for *raw gain* parameter of the ADC nonlinearity correction. Many channels have significant bias from 4 K lines and it could help to get better constraint to get closer to the true *raw gain* and to estimate the bias quantitatively.

2.4.2 Three thermal component model

Another direct application is the addition of a third thermal component. This is a direct result from the analysis of the analytical model performances in Sec. 2.3, and it is quick to implement. Usually the optical power is considered to be deposit on the bolometer itself, but a possible variation would be to do it on the intermediate component as shown in Fig.2.6. In this case the first component is considered to be the thermometer and the intermediate component is considered to be the absorber grid. Additionally the second thermal component must have a heat capacity of the same order as for the thermometer to be meaningful. This architecture is motivated by the presence of thermometer specific glitches, which are detailed in Chap. 3.

If it happens that a third component is not consistent enough with the in-flight data, then another option would be to investigate a new thermal design, which is discussed in the next section.

2.4.3 The disc form factor thermal model

This section presents a discussion about the thermal model itself, the empirical time transfer function is discussed first to give the motivation for refining the thermal model, then it is proposed a new thermal model designed to comply with the observations.

Let start by having a look to the thermal parameters (a_i, τ_i) of the sum of low pass filters in Eq. 2.2. The relative scaling parameters a_i have no unit, and the parameters τ_i are time constants. The thermal filtering profile fitted with these parameters is shown in Fig. 2.7 for two 100 GHz channels. There are two striking points:

1. the first time constant at 10 ms is not only consistent with the cutoff frequency of the 100 GHz detectors estimated in the pre-launch paper, but it is also very close to the values $\tau_e \approx 9$ms fitted using the two thermal component model in setup 2 and setup 3.

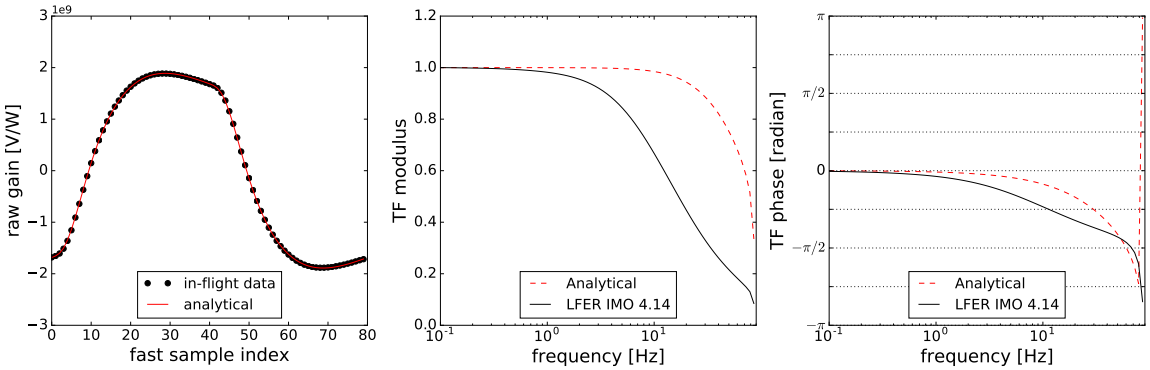


Figure 2.3: Analytical model fit setup 1: data = *raw gain* only. The 2015 time transfer function is fitted with the integrated analytical model version calculated after applying the electronic readout filtering. The transfer functions are normalized at frequency $\nu = 0$

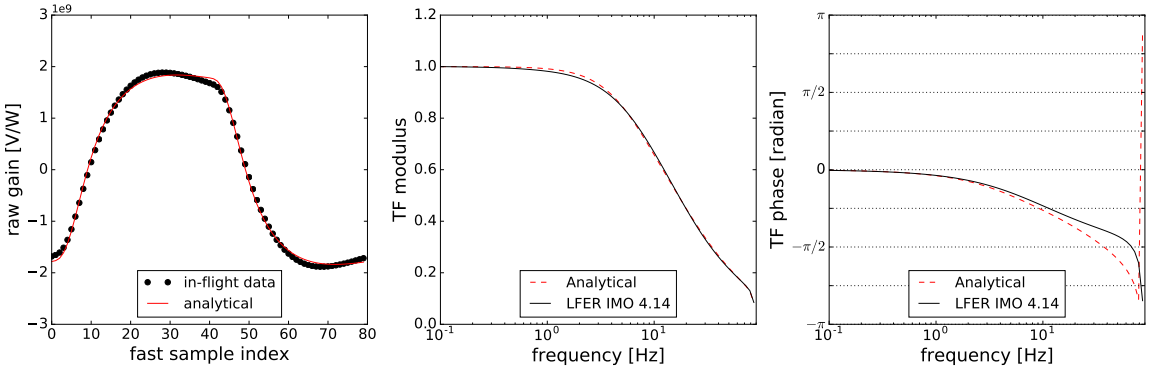


Figure 2.4: Analytical model fit setup 2: data = *raw gain* + TF. Same presentation as for Fig. 2.3.

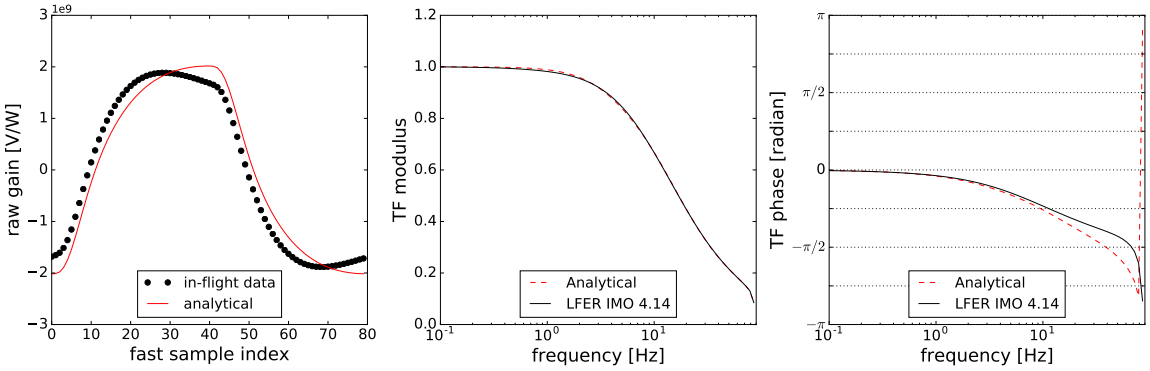


Figure 2.5: Analytical model fit setup 3: data = TF only. Same presentation as for Fig. 2.3.

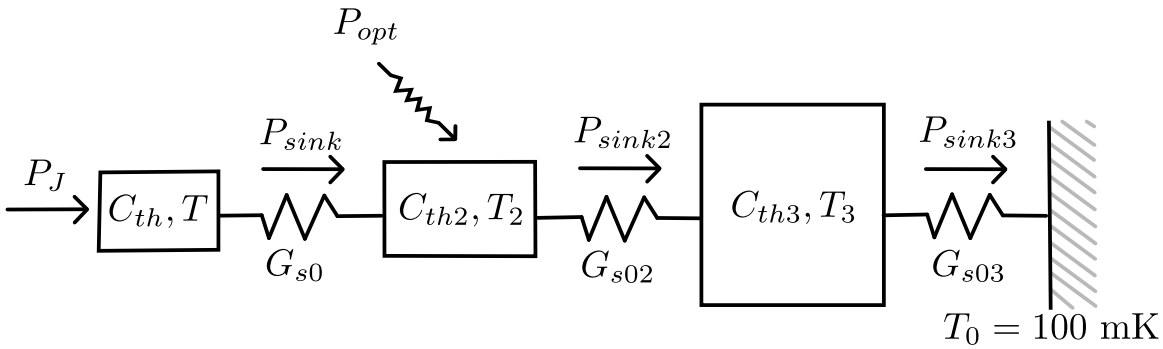


Figure 2.6: Thermal model with three components. In this version the optical power can be optionally deposited on the second component in the path.

On the contrary, the $\tau_e \approx 2$ ms obtained using the simple bolometer model alone is clearly wrong;

2. the amplitude of filters with the longest time constants clearly **follow a power law**.

The first point is consistent with a single thermal component for the bolometer (absorber grid plus thermometer), and the second point indicates that the thermal response have a continuous profile in frequency that follows a power law. This thermal profile in frequency is critical, because it obviously extends up to the dipole periodicity which is the calibration source for the CMB channels. Moreover, depending on the channel the slope is different, for example for channel `01_100-1b` the a_i parameter at dipole periodicity would be close to 10^{-3} which is significant for calibration.

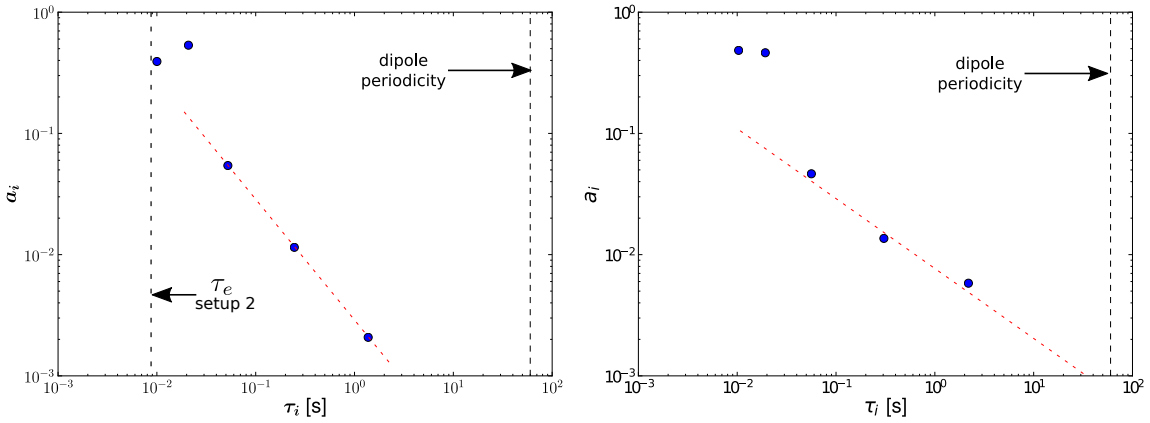


Figure 2.7: Thermal filtering profile for two 100 GHz channels. Left: channel `00_100-1a`. Right: channel `01_100-1b`. The location of the effective time constant τ_e, obtained from the setup 2 fit, has been added for the corresponding channel. The dashed red line is drawn by hand to provide an eye guide.

The previous remarks suggest either an update of the empirical model or a potential new design for the thermal dissipation. It is shown in Fig. 2.8 a thermal dissipation model using a disc-like shape instead of a thermal link with a tube-like shape. This model considers the heat conduction and the heat capacity at the same time as variables of the spacial radius parameter r and temperature T. In steady state with $P_{sink} = \text{cst}$, the temperature of the link can be computed in the same way as for the simple bolometer model by noting that $P_{sink}(r) = \text{cst}$. For small variations of the heat flow $\widetilde{P_{sink}} << P_{sink}$ the problem has to be linearized using small signals approximations. For each small element δC_{th} of the link, the

heat capacity scales with r. Using the formalism of the AC paper, the heat capacity reads

$$\delta C_{th} \sim r^{\varphi} \delta r T(r)^{\gamma} \,, \tag{2.3}$$

where $\varphi = 1$ for a disc but can be tuned between 1 (disc) and 2 (sphere) to account for different spatial geometry. The thermal equilibrium at location r can be written

$$\delta C_{th} \frac{dT}{dt} = \frac{d\widetilde{P_{sink}}}{dr} \delta r \,. \tag{2.4}$$

This model has the interesting property that the termination by a heat sink is not necessary because the thermal variations do vanish when $r \to \infty$ as $\delta C_{th} \to \infty$.

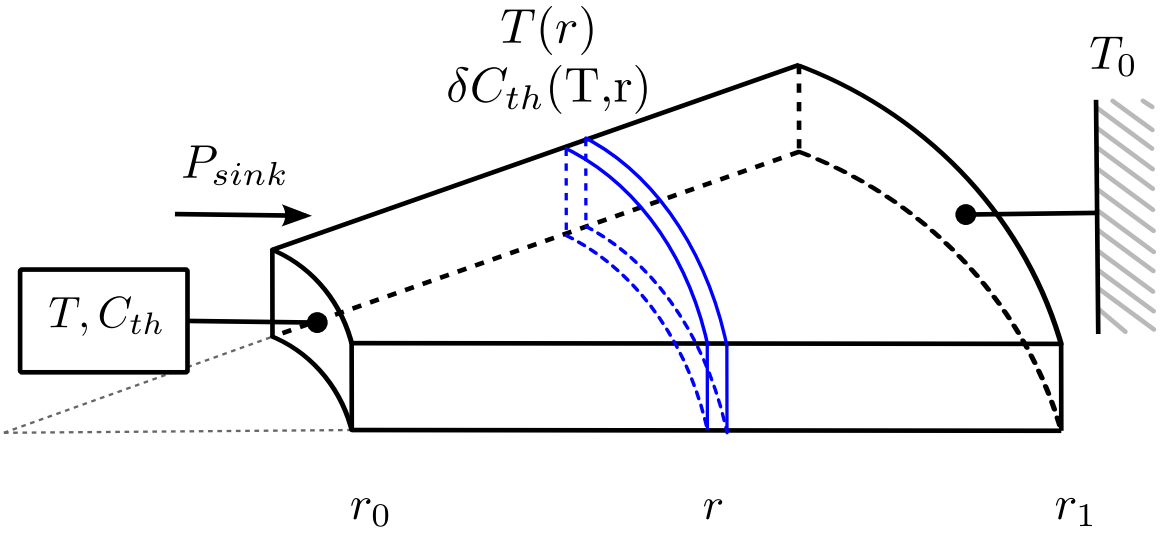


Figure 2.8: Thermal dissipation with a disc model for the link to the heat sink

2.5 Conclusions

This chapter detailed the motivations and the content of the AC paper.

The analytical model allow to build the time transfer function in frequency domain for the simple bolometer model under AC bias, and this model has been extended with a second thermal component. The single thermal component form have a direct application for ADC nonlinearity correction. While the two thermal components form is very close to the goal with a consistency of the corresponding time transfer function modulus at 2.7σ of the empirical model used for the *HFI* 2015 data release. This version needs only 5 free thermal parameters instead of 14 for the empirical model. However the analytical time transfer function phase and *raw gain* does not meet yet an acceptable consistency with the versions used for the *HFI* 2015 data release. It strongly suggests that a third thermal component would reach the objective of providing a constraining model for the issue of the LFER calibration.

Chapter 3

Glitches exploitation

Susceptibility to cosmic rays energy deposit is inherent to high sensitivity bolometers and the energy they deposit on the various part of the detector produce glitches on the measured signal. In practice this is a nuisance leading to non-Gaussian noise and loss of data. However these pretty unwanted residuals for some can become gold for others. In the case of glitches their high statistics and the stationarity in their characteristics is a great tool providing important constraints for time transfer function characterization and calibration. Usually glitches are phase-shifted to be compared to the time transfer function (e.g. see Fig. 10 in **?**). Below the contrary will be discussed, i.e. how the glitches can be used to produce the time transfer function.

Sec. 3.1 introduces the topic of glitches, from their origin to their characterization, and how they are related to the time transfer function. Then in Sec. 3.2, I present a method to build glitch templates homogeneous with the time transfer function. Finally in Sec. 3.3 direct applications of the glitch templates are shown.

The work presented below has been used to validate the empirical time transfer function LFER model (see Sec. 2 in **?**) and provided contents for two other papers which are: **?**, and its companion paper **?**.

3.1 Introduction

Glitches origin

Cosmic rays deposit energy on the 0.1 K stage detectors which is visible on the acquired data as glitches. The spacecraft operates at the second Lagrangian point where the dominant population of particles reaching the detectors is protons of energy > 39 MeV (**?**). The typical rate of glitches on a detector is about one event per second, with various characteristics depending on the physical location of the energy deposit.

Glitches follow closely the rate of particles measured by the Standard Radiation Event Monitor (SREM, see **?**). In the *HFI* instrument, the glitch rate is unexpectedly high because of the high sensitivity design of the detectors. Furthermore, it happened that the beginning of the *Planck* mission corresponded to the minimum of the 11-years cycle of solar activity, as shown in Fig. 3.1. As a consequence, the proton rate from Galactic sources which is modulated by the solar activity (**?**) was at its maximum.

Characterization of the glitch families

The location of energy deposit produces glitches with different shapes, which can be classified with the time constants of the relaxation visible after the event. Two examples are shown in Fig. 3.2: one *short* glitch and one *long* glitch in the sense of the amount of time necessary for the relaxation after the transient event.

The classification of glitches into **families** is critical, because some events have a long relaxation time leaving low frequency noise on science data. Hence individual events needs

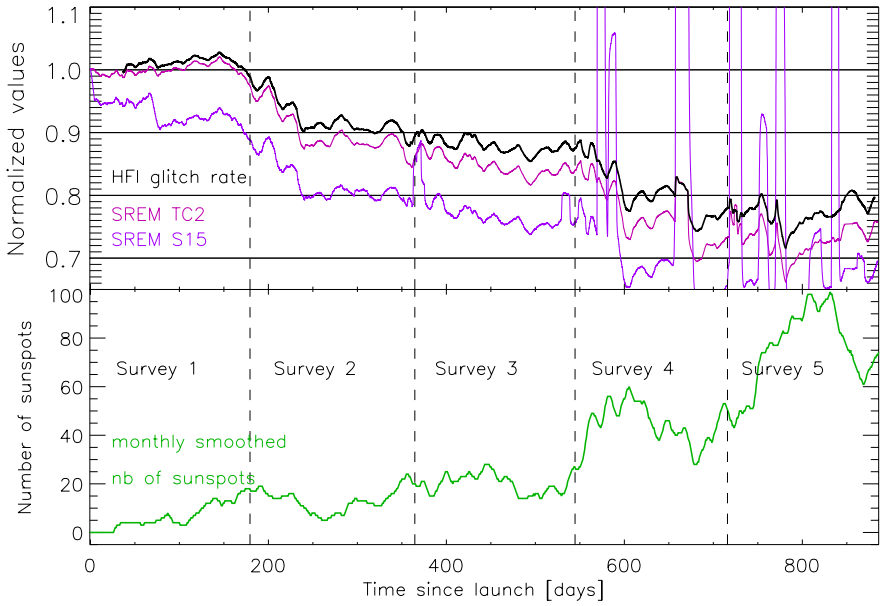


Figure 3.1: Top: Normalized glitch rate (black line) versus cosmic ray flux measured by the SREM for energy deposited > 3MeV (purple) and > 0.085 MeV (pink) as a function of time. Bottom: monthly smoothed number of sunspots. Image credit: **?**.

to be classified accurately for an optimal removal of their residual. This topic is covered in **?**. Furthermore the study of glitches provides information on the thermal path used by the heat to reach the detector, which is useful for the design of detectors with a lower sensitivity to particles hits. This topic is detailed in the companion paper (**?**).

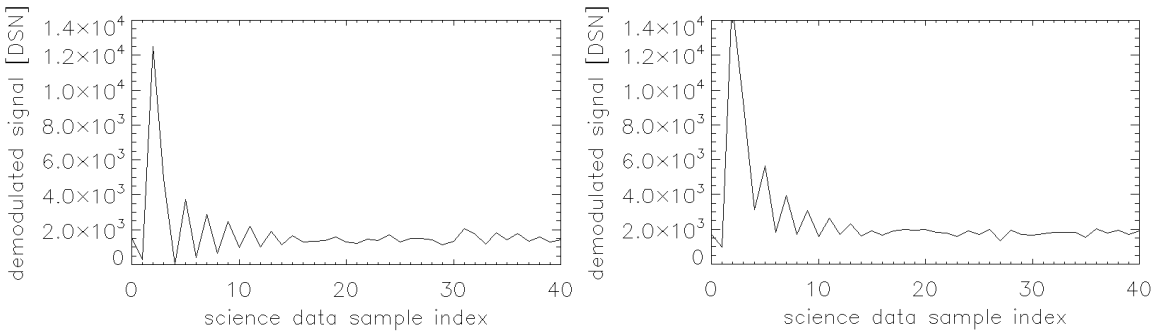


Figure 3.2: Examples of glitches on demodulated science data in DSN units for channel `00_100-1a`, there is no other processing than the demodulation. Left: a typical *short* glitch with a tail ringing around the average value of the signal. Right: a typical *long* glitch with a tail ringing around a visible relaxation.

There are three dominant glitch families (see Fig. 3.3 and the companion paper for the corresponding physical locations) which are:

- LONG GLITCHES: resulting from particle hit on the *silicon die*. this is the dominant family in number of events;
- SHORT GLITCHES: due to particles hits on the *absorber* grid. It is the family of interest for time transfer function analysis because the energy deposited uses the same thermal path as the optical power;
- HIGH ENERGY SHORT GLITCHES: due to impact of particles on the thermometer.

The 2D histogram in Fig. 3.4 allows us to distinguish these families from the repartition of the particle energy between the head of the glitch and the tail. As the real glitch energy is unknown, the *pseudo energy* quantity ΔE_g is defined as

$$\Delta E_g = \sum_{j=-2}^{19} , \tag{3.1}$$

where $j = 0$ correspond to the glitch peak value. Additionally the tail *pseudo energy* is

$$\Delta E_t = \sum_{j=8}^{19} . \tag{3.2}$$

There is a visible nonlinearity for the *long* family (the distribution is not vertical) which is completely unrelated to the bolometer nonlinearity much fainter. For the high energy *short* family there is strong numerical saturation above 10^6 DSN due to the analog signal reaching the upper bound of the ADC chip.

3.1.1 Relation with the time transfer function

When a particle hits the absorber grid at time t_0, the energy deposit can be considered as a Dirac function $\delta(t - t_0)$. Hence the output of the electronics readout after the event and *before digitization* is

$$s(t) = \delta(t - t_0) * F'(t) , \tag{3.3}$$

where $F'(t)$ is the transfer function of the electronic readout before digitization and $*$ is the convolution product. This is equivalent in Fourier domain, taking t_0 as the time reference, to

$$s(\omega) = F'(\omega) , \tag{3.4}$$

thus the Fourier transform of the signal generated by the particle provides the transfer function of the electronics readout before integration by the DPU and sampling.

Things are not so simple because of the discrete nature of science data. which is produced after integration by the DPU of the digitized signal. After integration and demodulation, the effective transfer function for the science data is $TF(\omega)$. Moreover the science data is sampled at frequency f_{acq}, which is equivalent to a convolution of $TF(\omega)$ with a Dirac comb of frequency step f_{acq}. Unfortunately the high frequency part of $TF(\omega)$ contains a significant amount of power above the modulation frequency, As a consequence the Fourier transform of an individual glitch does not produce directly the time transfer function $TF(\omega)$ but its folded version. In practice, particles hitting the grid at different phases ϕ of the modulation period will produce *short* glitches with very different shapes. This behavior is shown in Fig. 3.5 where glitches are produced with the simulation tool `SEB` (see Sec. 5.1 of the AC paper in appendix A).

The continuous Finite Impulse Response (FIR)

The folding issue can be solved by extending the discrete system response to its continuous version in time domain.

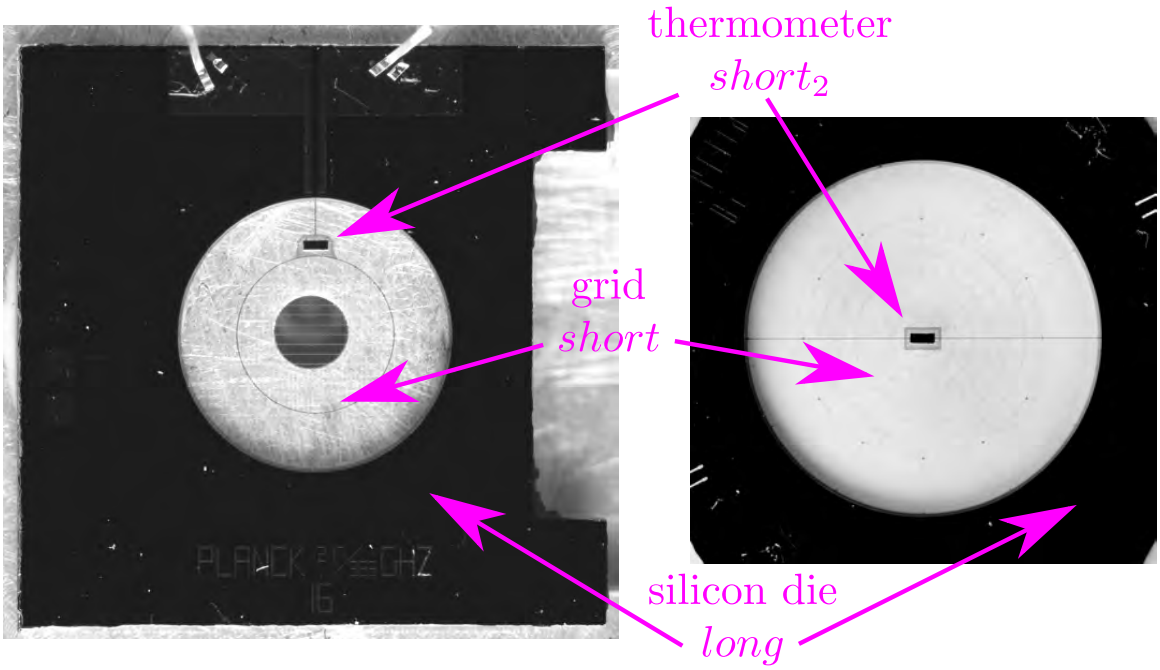


Figure 3.3: *HFI* bolometers annotated with glitch energy deposit locations and the corresponding family name. *Left*: a 143 GHz PSB (polarization sensitive) bolometer. *Right*: a 353 GHz SWB. Image credit: **?**

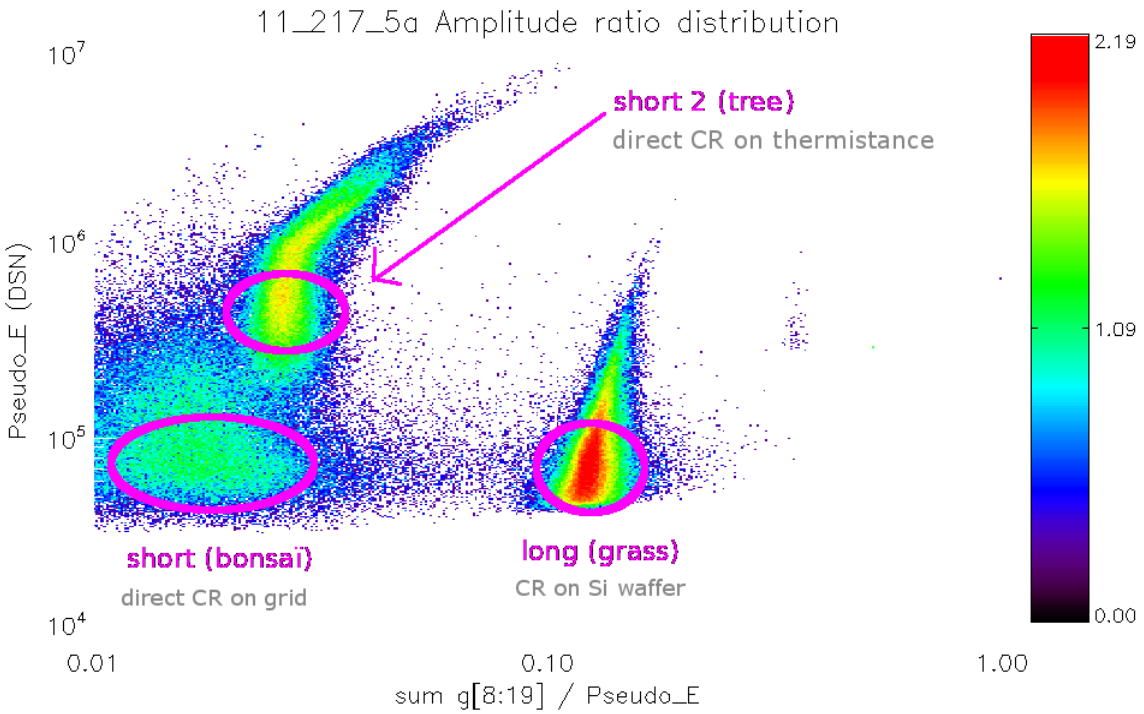


Figure 3.4: Glitch families visualization with the 2D histogram of *pseudo energy* distribution for channel `11_217-5a`. The x-axis is $\Delta E_t/\Delta E_g$ and ordinate is ΔE_g (see description in text). The z axis color code is the base 10 logarithm of the number of events per pixel. The botanical nicknames have been widely used in *Planck* Collaboration.

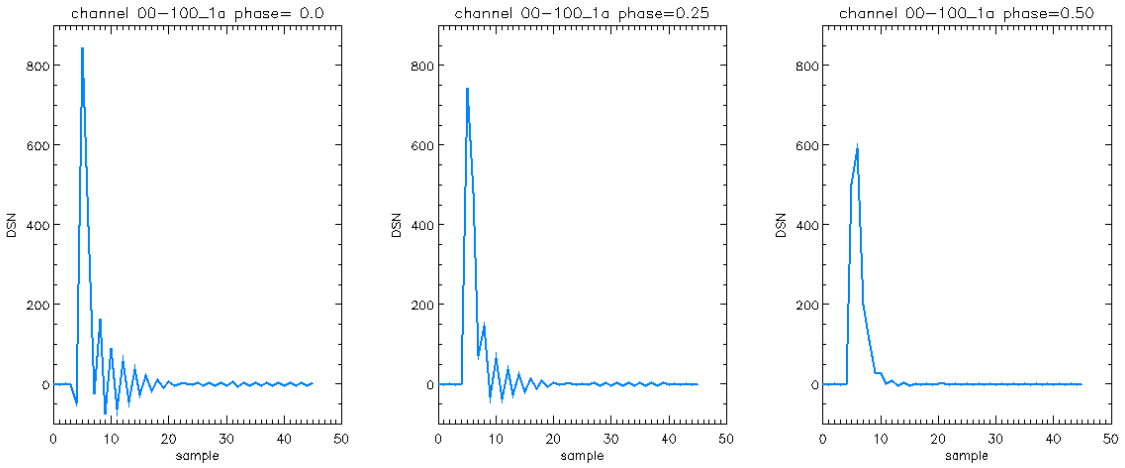


Figure 3.5: Simulation of glitch (demodulated science data) from energy deposit at different phases within the modulation period. From left to right, the energy deposit occurs at phase value 0, $\pi/2$ and π. In the rightmost case there is no ringing, this feature is also observed on real data.

A single event produces a finite number of discrete samples with a significant signal value: the FIR[1]. Let us consider that $t = 0$ is the time reference of the DPU integration, at the beginning of the first modulation half period. The location ϕ of energy deposit in the modulation period is constrained to $[0, \pi[$, because this range describes all the possible system responses on demodulated science data. The discrete response to a single event is $\mathrm{FIR}_\phi(j)$, where $j \in [\![0, n[\![$ is the index of the science data samples produced for $t > 0$ at phase ϕ, and n is the size of the time domain support in number of samples.

A continuous function of time can be produced by noting that the system response to a glitch is

$$\begin{aligned} \mathrm{FIR}_\phi(j) &= \left(\delta\left(t - \frac{\phi}{\pi f_{\mathrm{acq}}}\right) * TF(t)\right) \times \text{Ш}_{\Delta T} \\ &= TF\left(t - \frac{\phi}{\pi f_{\mathrm{acq}}}\right) \times \text{Ш}_{\Delta T}\,, \end{aligned} \tag{3.5}$$

where $\text{Ш}_{\Delta T}$ is a Dirac comb centered on zero and characterized by a time step $\Delta T = 1/f_{\mathrm{acq}}$, and $TF(t)$ is the inverse Fourier transform of $TF(\omega)$. Thus $\mathrm{FIR}_\phi(j)$ interpolates $\mathrm{TF}(t)$ at location $(1 + j)\Delta T - \frac{\phi}{\pi f_{\mathrm{acq}}}$ (the one in $(1 + j)$ comes because the sample is produced after integration by the DPU with a time shift of ΔT). This function $TF(t)$ is called hereafter the continuous FIR. An example of simulated continuous FIR is shown in Fig. 3.6, the displayed function is $\mathrm{FIR}(t) \times \text{Ш}_{\Delta T/40}$.

In practice

$$TF(t) \times \text{Ш}_{\delta t} \xrightarrow[\delta t \to 0]{} TF(t)\,, \tag{3.6}$$

which solves the folding issue we had with the nominal sampling using $\text{Ш}_{\Delta T}$. This property has been validated with `SEB` simulations by producing a version $TF_\delta(\omega)$ with sinusoids as input sky power (the classic way), and a second version $TF_{\text{Ш}}(\omega)$ with glitches as proposed

[1] In signal processing, a FIR filter has an impulse response with a finite support in time. While the *HFI* electronic readouts has an infinite impulse response in time, by abuse of language the duration for which the response is significant is called the FIR.

above. Both version are in excellent agreement, down to the nonlinearity level of the detector with a relative difference $< 10^{-4}$ for $\omega \in [0, f_{\text{acq}}[$ and 200 frequency steps.

This brings us to the next section where is presented a method for building the high frequency resolution FIR from glitches.

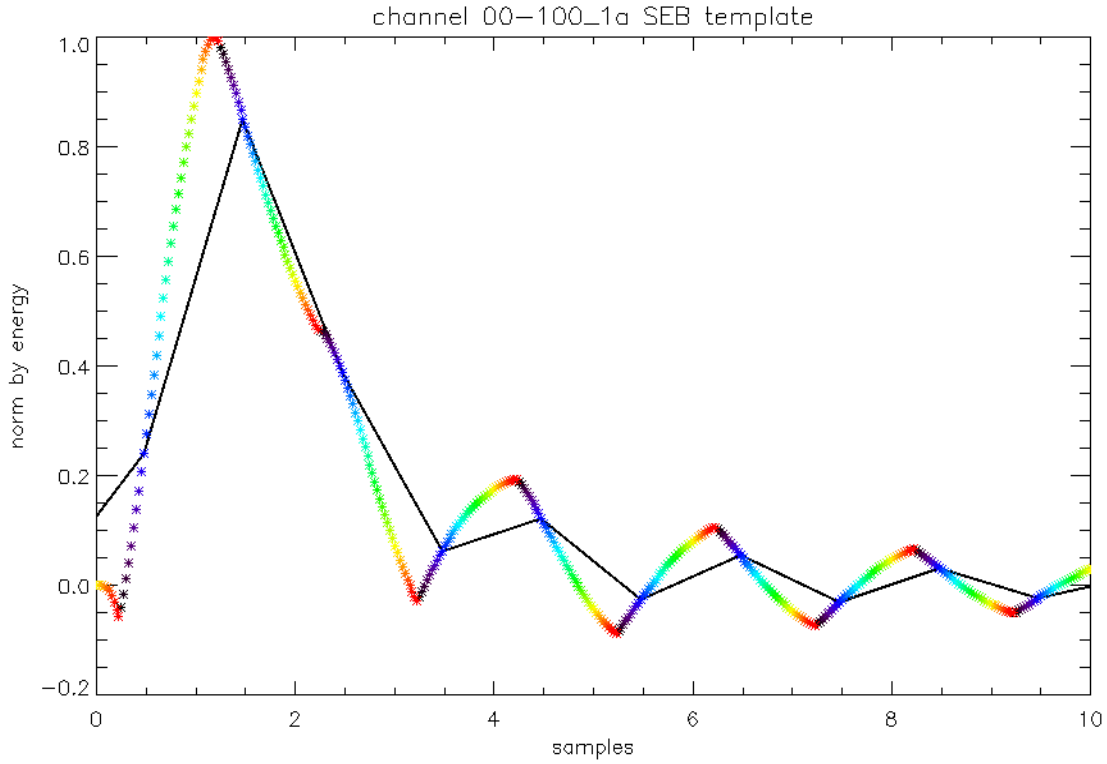


Figure 3.6: Simulation of the continuous FIR with the setup of channel `00_100-1a`. There are 40 glitches simulated with `SEB` at constant energy and sampled at frequency f_{acq}. The phase ϕ of energy deposit of each glitch in the modulation period is evenly spread over $[0, \pi[$. Each glitch is displayed with a different color code and shifted by ϕ/π, The black thick line is an eye-guide showing one of the individual glitches.

3.2 Building glitch templates

In this section is detailed the method developed to build glitch templates which are homogeneous with the continuous FIR of the selected channel. The main issue is to order the glitches of the dataset by their phase location in the modulation period.

To do so, glitches are projected in an orthogonal basis using a formalism based on Euclidean geometry. Next it is described the method which produce the template by path following of the glitch density in this basis. Finally some results are presented and template systematics are discussed.

3.2.1 Formalism

Projection of glitches in Euclidean space

Every glitch produced at phase ϕ is only an interpolation of its continuous FIR as seen previously, but another glitch produced with a phase $\phi + \delta\phi$ is very "close" to the first one. This notion of "closeness" is defined by considering glitches as points in a well chosen basis,

in which the function FIR(t) projects as a continuous curve $\mathcal{R}(\phi) : [0, \pi[\to \mathbb{R}^n$ defined by

$$\mathcal{R}(\phi) \equiv \mathrm{FIR}_\phi \tag{3.7}$$

In practice $\mathcal{R}(\phi)$ can be visualized as being all the data points of the same color in Fig. 3.6, with the phase value ϕ assimilated to the colour index. Hence, instead of reconstructing directly FIR(t) which is not simple, it is easier to reconstruct $\mathcal{R}(\phi)$ from the dataset of glitches by doing path following in $\mathbb{R}^n$, where n is the finite number of data samples for one glitch.

The distance between two points $(\boldsymbol{a}, \boldsymbol{b})$ defined as $\boldsymbol{a} = \mathcal{R}(\phi_a)$ and $\boldsymbol{b} = \mathcal{R}(\phi_b)$ is calculated with the Euclidean norm noted

$$\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) = \|\overrightarrow{\boldsymbol{ab}}\|\,, \tag{3.8}$$

where $\overrightarrow{\boldsymbol{ab}}$ is the vector defined by $(\boldsymbol{a}, \boldsymbol{b})$.

Glitch selection

Glitches present a wide energy distribution and their response is nonlinear (except for the *short* family) with respect to the particle energy. The nonlinearity causes undesirable distortion of the FIR over the energy distribution of glitches. This distortion can be minimized by selecting glitches corresponding to a short energy window. As the energy of the particle is unknown it has to be approximated with the *pseudo energy*

$$\widetilde{E}(\boldsymbol{g}) = \sum_{j=1}^{n} \boldsymbol{g}(j)\,, \tag{3.9}$$

where $\boldsymbol{g} \in \mathbb{R}^n$ is the glitch data vector extracted from science data, starting 2 samples before the peak and after removal of the sky signal baseline[2]

The issue with the *pseudo energy* is that $\widetilde{E}$ has a small variation as a function of ϕ, as shown in Fig. 3.7. It will result in a small leak of the energy distribution of glitches into the density of glitches per phase $dn/d\phi$.

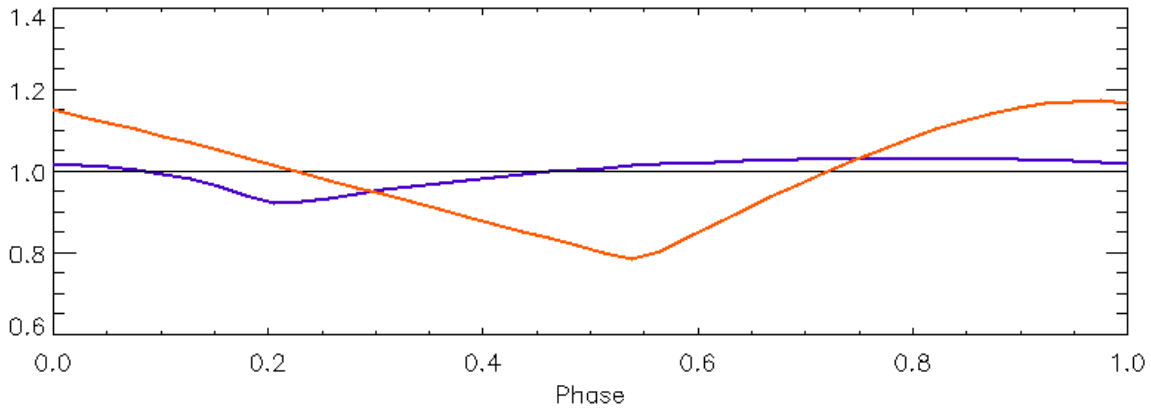


Figure 3.7: Glitch normalization with *pseudo energy* for short glitches simulated with `SEB` at constant energy E, the x-axis value is ϕ/π. Orange: peak value of the glitch relative to its average. Purple: $\widetilde{E}(\phi)/\langle\widetilde{E}(\phi)\rangle$. The flatter, the better.

[2] Sky signal baseline removal is done by subtracting the interpolated Phase Binned Ring which production is detailed in Sec. 5.5.2).

Glitch normalization

Because of the wide energy distribution of glitches, even after selection of a small energy window, the dataset of glitches in $\mathbb{R}^n$ does not form a compact curve. Instead it is rather a ribbon-like surface parametrized with (ϕ, E), where E is the real amount of energy deposited by the particle.

The *pseudo energy* normalization cannot be used here, because it will boost the noise too much. Instead, the normalization by peak can be used for the purpose of building a compact curve in $\mathbb{R}^n$. Let us consider a dataset of glitches $\{\boldsymbol{g}_i\,, i \in [\![1, m]\!]\}$. For a given glitch

$$\boldsymbol{g}_i = E \times \mathcal{R}(\phi_i)\,, \tag{3.10}$$

where ϕ_i is the unknown phase location in the modulation period where the particle deposited energy.

In practice glitch are extracted from data after peak detection and the peak is always located at index 2 of the vector $\boldsymbol{g}_i$. The glitch peak value is $g_i(2) = E \times \mathrm{FIR}((2+\phi_i/2\pi)/f_{\mathrm{acq}})$ The normalized glitch expression is

$$\begin{aligned} \widetilde{\boldsymbol{g}}_i &= \frac{\mathcal{R}(\phi_i)}{\mathrm{FIR}\left((2+\phi_i/2\pi)/f_{\mathrm{acq}}\right)} \\ &= \widetilde{T}(\phi_i)\,, \end{aligned} \tag{3.11}$$

where $\widetilde{T}(\phi)$ is the *glitch template* corresponding to the peak-normalized version of $\mathcal{R}(\phi)$.

The normalized dataset $\{\widetilde{\boldsymbol{g}}_i\}$ achieves the goal of providing a compact parametric function of ϕ, even if the template produced by path following within the normalized dataset will have to be renormalized with *pseudo energy*.

Noise radius

The noise of the data samples of one glitch $\widetilde{\boldsymbol{g}}_i$ is $\sigma_i = \sigma/max(|\boldsymbol{g}_i|)$, where σ is the TOI noise which is considered as white and stationary in time. In $\mathbb{R}^n$ the noise has a χ^2 distribution centred on the true location of the glitch $\widetilde{T}(\phi_i)$. Half of the noise distribution is contained in a n-sphere of radius r_σ. The value of r_σ is called the noise radius and is calculated with the median of a χ^2 distribution. When considering a dataset $\{\widetilde{\boldsymbol{g}}_i\}$ the normalization is done with the median of glitch amplitude, thus the noise radius of the dataset is

$$r_\sigma = \frac{\sigma\sqrt{n - 2/3}}{\mathrm{median}(\max(|\boldsymbol{g}_i|))}\,. \tag{3.12}$$

Using the median instead of the mean is necessary with a dataset subject to outliers.

In practice the noisy dataset is called the glitch *cloud* and its distribution is the convolution product of the glitch template $\widetilde{T}(\phi)$ and the χ^2 distribution of the noise. It has a tube-like shape in $\mathbb{R}^n$. Hence, it is important to consider a short energy window for glitch selection to have homogeneous noise properties over the dataset $\{\widetilde{\boldsymbol{g}}_i\}$.

Visualization of the glitch *cloud*

As the *cloud* of glitches is in $\mathbb{R}^n$ it has to be projected in 2 dimensions for visualization. For this purpose a projection axis $(\widetilde{\boldsymbol{g}}_o, \overrightarrow{\boldsymbol{e_x}})$ is defined by choosing an origin $\widetilde{\boldsymbol{g}}_o$ and a direction represented by the unit vector $\overrightarrow{\boldsymbol{e_x}}$. The x-axis value x for a glitch $\widetilde{\boldsymbol{g}}$ is defined as the distance between its orthogonal projection $\widetilde{\boldsymbol{g}}^{\perp}$ on the axis to the origin, and the y-axis value y is

defined as the distance to the axis. The expression of the projected coordinates is

$$\begin{cases} x = \overrightarrow{o\widetilde{g}^{\perp}}.\overrightarrow{e_x} \\ y = \left\| \overrightarrow{\widetilde{g}_i^{\perp}\widetilde{g}_i} \right\| \end{cases}, \tag{3.13}$$

An example of glitch cloud is presented in Fig. 3.8, while the path following (described in Sec. 3.2.2) is half way. With this representation it is clear that the two visible families are compact and well separated, but also that they have a complex shape in $\mathbb{R}^n$ which requires a specific algorithm to reconstruct accurately the glitch template. The noise radius is very small and the points around the path being followed are caused by hits on the bolometer grid presenting a large dispersion. Therefore this kind of 2D plot can be very powerful for the analysis of the thermal behaviour of the detector.

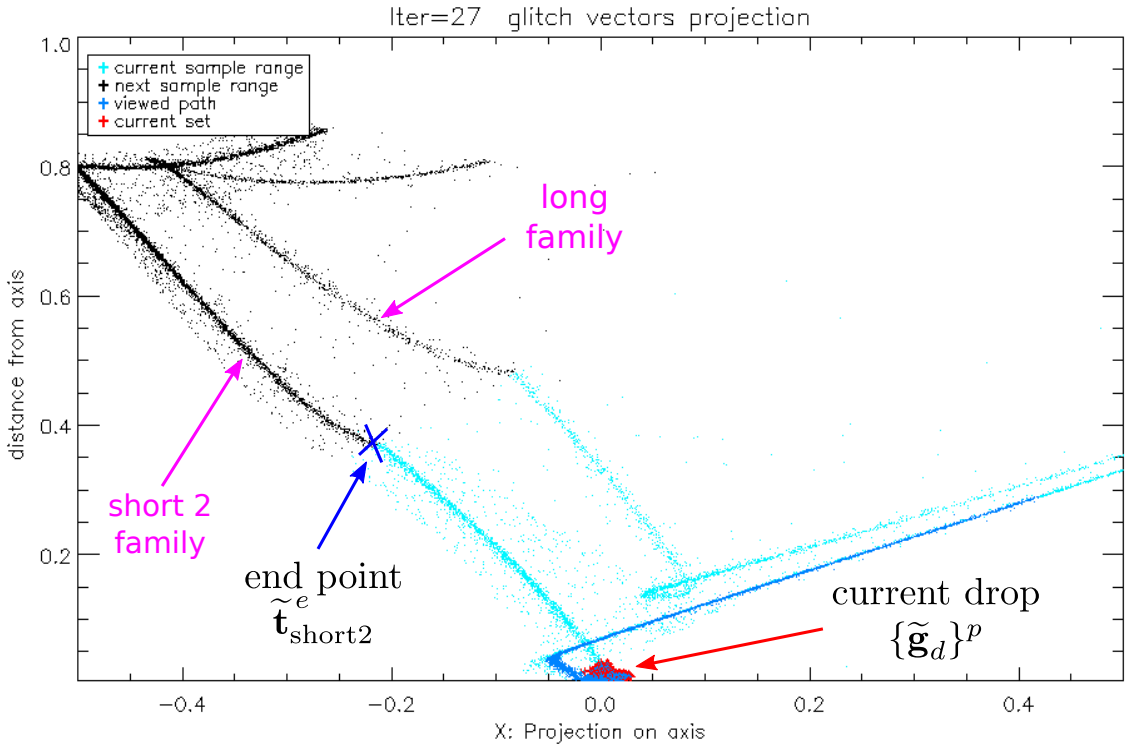


Figure 3.8: Example of 2D projection of glitch cloud for channel `02_143-2a`, while the path following algorithm (described in Sec. 3.2.2) is half way. The cyan dots form the glitch cloud $\{\widetilde{\boldsymbol{g}}_i\}$, the red dots are the small subset corresponding to the current location in the path being followed, the blue dots are the glitches forming path drops and removed from $\{\widetilde{\boldsymbol{g}}_i\}$, and the black dots are the glitch cloud time-shifted of one sample in the future. Two distinct glitch families are visible as strings with a complex shape presenting a "loop" near the origin.

3.2.2 Glitch templates reconstruction with the capillarity method

Now that the geometrical formalism is set up, this section presents how to approximate the template $\widetilde{T}_f(\phi)$ with $\{\widetilde{\boldsymbol{g}}_i\}$ for the glitch family f. First, a starting point is needed, this is done by searching a characteristic location for each family present in the *cloud*. After that, the path defined by $\widetilde{T}_f(\phi)$ is followed with a method inspired by the way water progresses by capillarity in a string.

Localisation of starting point

For a given glitch family f, the location $\widetilde{T}_f(0)$ is unknown before construction of the template, thus a *start location* characteristic of the family is needed.

A simple way of choosing it is to select empirically ϕ_0 as the *start location* where glitches g_i have two peak values, one at location $\widetilde{\boldsymbol{g}}_i(2)$ when the signal is raising, and the next one at $\widetilde{\boldsymbol{g}}_i(3)$ when the signal is falling down. Some examples of glitch in this configuration are shown in Fig. 3.9.

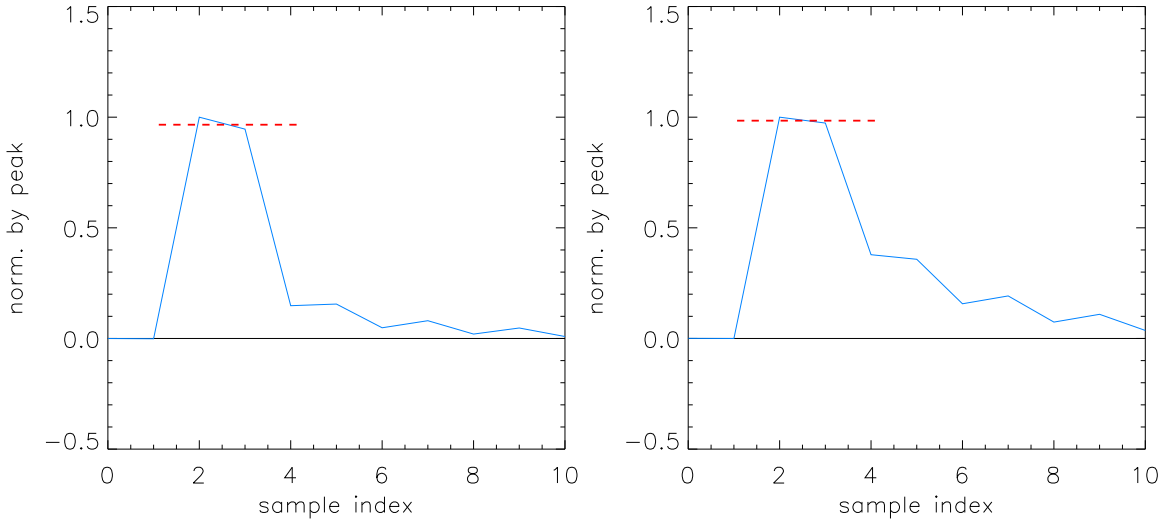


Figure 3.9: Example of real data glitch vectors $\widetilde{\boldsymbol{g}}_i$ in *start location* configuration for *short* family (left) and *long* family (right), they have been captured on channel 00_100-1a. Both examples are very close to fulfill $\widetilde{\boldsymbol{g}}_i(2) = \widetilde{\boldsymbol{g}}_i(3)$. In $\mathbb{R}^n$ these glitches define two locations very distant and unique to each family.

The property that this location is unique for each family has been validated with functions $\mathcal{R}(\phi)$ generated by `SEB`. The *long* family response can be simulated with a decreasing exponential for the input optical power instead of a Dirac. This property has been furthermore confirmed on in-flight *HFI* data.

Let note $\widetilde{\boldsymbol{t}}_f^{\,\mathrm{s}}$ the *start location* in $\mathbb{R}^n$ near $\widetilde{T}_f(\phi_f)$. There is a large number of glitches within a radius r_σ of $\widetilde{T}_f(\phi_f)$ because it is at the center of the distribution of glitches produced by energy deposited at ϕ_f. It can be reached by density search in a subset of $\{\widetilde{\boldsymbol{g}}_i\}$. This subset $\{\widetilde{\boldsymbol{g}}_i\}_0$ is defined by calculating $|\widetilde{\boldsymbol{g}}_i(2) - \widetilde{\boldsymbol{g}}_i(3)|$ and keeping only glitches having this value lower than $2 \times r_\sigma/\sqrt{n}$. After that, the following iterative procedure is applied for each glitch $\widetilde{\boldsymbol{g}}_k \in \{\widetilde{\boldsymbol{g}}_i\}_0$ in order to maximize the local density and to get closer to $\widetilde{\boldsymbol{t}}_f^{\,\mathrm{s}}$ at the same time:

- the neighbors of $\widetilde{\boldsymbol{g}}_k$ at a distance lower than the noise radius are selected. Which reads $\|\overrightarrow{\widetilde{\boldsymbol{g}}_k\widetilde{\boldsymbol{g}}_i}\| < r_\sigma$;
- half of the neighbors is selected which have the lowest value for $|\widetilde{\boldsymbol{g}}_i(2) - \widetilde{\boldsymbol{g}}_i(3)|$;
- the barycenter of this half is computed and used as the new starting point candidate for step one.
- when the distance between the candidates of two iterations is lower than $0.05 r_\sigma$ the location is stored as a $\widetilde{\boldsymbol{t}}_f^{\,0}$ candidate.

The families starting points are detected at locations concentrating a large number of candidates within the noise radius. In practice the dispersion of the candidates for a given family is $\ll r_\sigma$, thus the barycenter of the family candidates is chosen as $\widetilde{\boldsymbol{t}}_f^{\,\mathrm{s}}$.

Path following

The glitch template $\widetilde{T}_f(\phi)$ has a complex shape in $\mathbb{R}^n$ in *HFI* case with loops[3] which prevent using a simple stepping algorithm based, for example, on the nearest neighbours of the current path location. Instead glitches must be selected carefully by following the centre of the path density as described below.

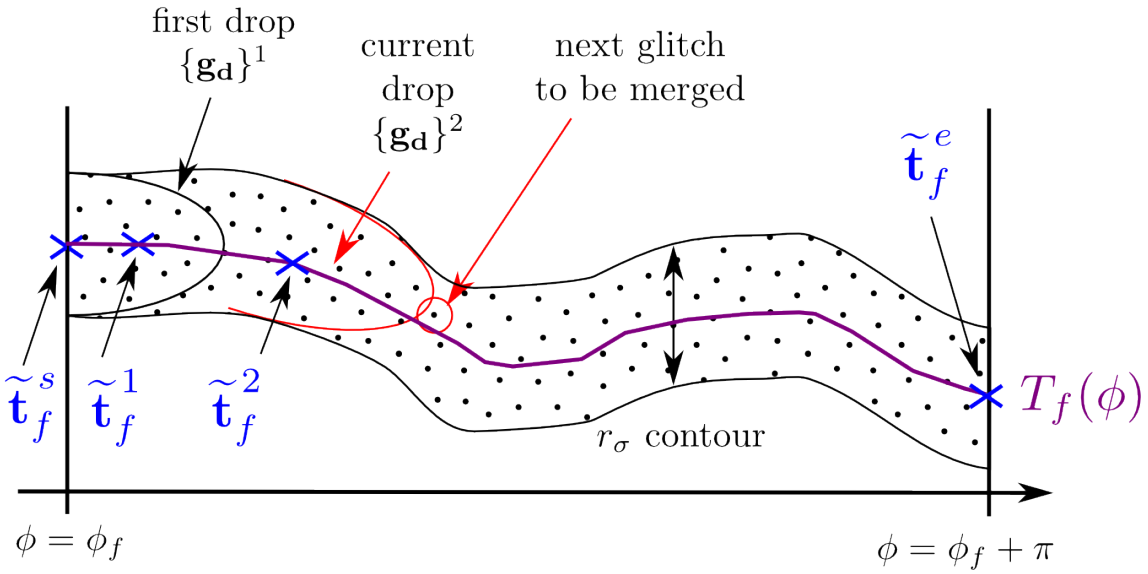


Figure 3.10: 2D representation of path following. In this configuration the current drop is $\{\widetilde{\boldsymbol{g}}_d\}^1$ and the drop expansion in formation $\{\widetilde{\boldsymbol{g}}_d\}^2$ is delimited by the red line. The path extension over the interval $[\phi_f, \phi_f + \pi[$ correspond to the cyan dots in Fig. 3.8.

The noise radius r_σ is the minimum interval between two iterations over the path, because the χ^2 distribution of the noise peaks close to this value. The algorithm boils down to building drops of glitches of radius r_σ along the path.

The initial drop: $\widetilde{\boldsymbol{t}}_f^{\,s}$ is used as a seed and all the glitches in a sphere of radius r_σ are selected to form the first drop noted $\{\widetilde{\boldsymbol{g}}_d\}^1$. The barycenter of $\{\widetilde{\boldsymbol{g}}_d\}^1$ defines the first path element $\widetilde{\boldsymbol{t}}_f^{\,1}$. Each time a drop is formed, the included glitches are removed from the global glitch set $\{g_i\}$.

Drop expansion by capillarity: the current drop $\{\widetilde{\boldsymbol{g}}_d\}^p$ $(p > 0)$ is initially empty. New glitches are added one by one to the current drop and form a merged drop $\{\widetilde{\boldsymbol{g}}_d\}^{\mathrm{M}} = \{\widetilde{\boldsymbol{g}}_d\}^{p-1} \cup \{\widetilde{\boldsymbol{g}}_d\}^p$ with the previous one. To ensure the merged drop expands along the path, a candidate glitches $\widetilde{\boldsymbol{g}}_k \in \{\widetilde{\boldsymbol{g}}_i\}$ is added only if it is the *closest* to $\{\widetilde{\boldsymbol{g}}_d\}^{\mathrm{M}}$. The *closest* property is defined for the candidate which minimizes

$$D_k = \sum_{\widetilde{\boldsymbol{g}}_q \in \{\widetilde{\boldsymbol{g}}_d\}^{\mathrm{M}}} \|\overrightarrow{\widetilde{\boldsymbol{g}}_k \widetilde{\boldsymbol{g}}_q}\|^2 . \tag{3.14}$$

A schematic view of this process is shown in Fig. 3.10. This method of expansion is very robust to the characteristic stiffness of the *HFI* time transfer function ringing.

End of expansion: when the radius of the current expanding drop $\{\widetilde{\boldsymbol{g}}_d\}^p$ reaches the noise radius r_σ, the expansion stops. The drop radius is recalculated after each glitch addition as the RMS of the distances of each glitch in $\{\widetilde{\boldsymbol{g}}_d\}^p$ to the barycenter $\widetilde{\boldsymbol{t}}_f^{\,p}$ of $\{\widetilde{\boldsymbol{g}}_d\}^p$. A new empty drop is created, it becomes the current drop and the expansion step is started again.

[3] The presence of loop and stiffness of $\widetilde{T}_f(\phi)$ is a consequence of the square modulation of *HFI* bolometers plus the summation of the signal over a half period

TERMINATION: the path is complete when the (virtual) end point $\widetilde{\boldsymbol{t}}_f^e$ is reached by the *connection drop* $\{\widetilde{\boldsymbol{g}}_d\}^c$, it is detected for $\|\overrightarrow{\widetilde{\boldsymbol{t}}_f^c\widetilde{\boldsymbol{t}}_f^e}\| < 1.5r_\sigma$. The end point is calculated from $\widetilde{\boldsymbol{t}}_f^{\mathrm{s}}$ by applying a time shift of one sample, which correspond to a shift of the vector $\widetilde{\boldsymbol{t}}_f^s$ on the left (in practice vectors are calculated with a size of $n+1$ to have content for the last value after the shift). The last drop $\{\widetilde{\boldsymbol{g}}_d\}^{c+1}$ is formed as for $\{\widetilde{\boldsymbol{g}}_d\}^1$, but in this case the remaining glitches are shared among the *connection drop* $\{\widetilde{\boldsymbol{g}}_d\}^c$ and the final drop $\{\widetilde{\boldsymbol{g}}_d\}^{c+1}$. This sharing is realized by comparing the glitches distance to $\widetilde{\boldsymbol{t}}_f^c$ and $\widetilde{\boldsymbol{t}}_f^{c+1}$.

The output of the algorithm is the path $\{\widetilde{\boldsymbol{t}}_f^{\mathrm{s}}, \widetilde{\boldsymbol{t}}_f^1, ..., \widetilde{\boldsymbol{t}}_f^{c+1}, \widetilde{\boldsymbol{t}}_f^e\}$, containing approximations of the glitch template $\widetilde{T}(\phi)$ at discrete locations for ϕ monotonously increasing.

3.2.3 Estimation of the continuous FIR

Reconstruction of the continuous FIR

Two operations are needed to produce the time domain glitch template: a phasing of the points forming the path, and the renormalization with the *pseudo energy*.

The phase of each point of the path is unknown but it can be estimated with the glitch statistic because the distribution of particles is supposed to be random uniform over the values taken by ϕ. Each drop $\{\widetilde{\boldsymbol{g}}_d\}^p$ contains n_p glitches and covers a phase range $\Delta\phi_p = \pi n_p/n_f$, where $n_f = \sum_j n_p$ is the number of glitches of the family f. The first (virtual) point defines the location $\phi = 0$. The barycenters of all drops are placed at the center of their corresponding phase interval, thus the phase location ϕ_p of a given point of the path $\widetilde{\boldsymbol{t}}_f^j$ is

$$\phi_p = \sum_{k<p} \Delta\phi_k + \Delta\frac{\phi_p}{2}\,. \tag{3.15}$$

The time domain FIR of the family f is produced after renormalization of $\widetilde{T}_f$ with the *pseudo energy* calculated for each phase. Its expression is

$$\mathrm{FIR}(j\Delta T + \phi_p) = \frac{\widetilde{\boldsymbol{t}}_f^p(j)}{\widetilde{E}(\widetilde{\boldsymbol{t}}_f^p)}\,, \tag{3.16}$$

with $j \in [\![0, n[\![$ and $\widetilde{E}$ is calculated with Eq. 3.9.

Some examples of functions $T_f(t)$ are shown in Fig. 3.11 after reconstruction of the continuous FIR for each glitch family of the channel `00_100-1a`. The zoom plot on the right column show the variety of shape for the slow component of the continuous FIR, a very slow relaxation can be seen for the *long* family (hence the name).

Bias and uncertainties

The pseudo-E normalization is a major source of error on the continuous FIR reconstructed from glitches. As it bias the selection by energy, the glitch density per phase drifts from a random uniform distribution. Let note ε_E the standard deviation of $\widetilde{E}(\phi)/\langle\widetilde{E}(\phi)\rangle$ As the glitch distribution scales with $\approx 1/E^2$ (see Sec. 2.4 of the glitch paper), for $\varepsilon_E \ll 1$ the relative error on E is $\approx 2\epsilon_E$. In practice ε_E is about 3%, this will introduce an error on the phase of about 6%. It can in theory be significantly reduced by applying a correction factor calculated from simulations.

The noise spread signal by convolving $\mathcal{R}(\phi)$ with the χ^2 distribution. This effect can be considered to be insignificant for high Signal over Noise Ratio (SNR), obtained by glitch energy selection. With a high SNR, r_σ is smaller and the number of drops becomes greater than 40, which corresponds to the *fast samples* acquisition frequency. Generally 40 drops

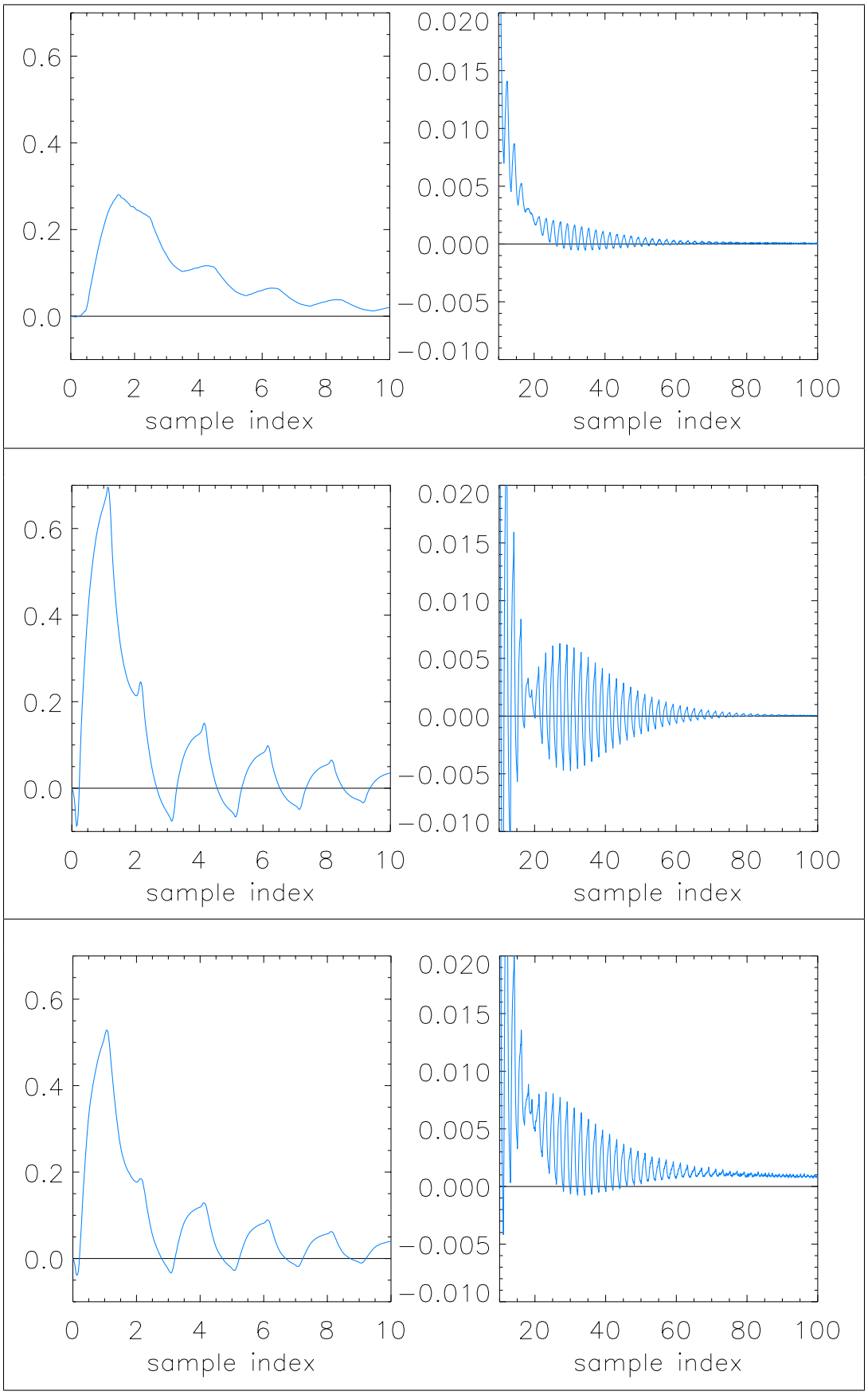


Figure 3.11: Examples of continuous FIR $T_f(t)$ reconstructed with glitch data from channel `00_100-1a` and normalized with *pseudo energy*. From top to bottom: *short* family (impact on grid), $short_2$ family (impact on thermometer) and *long* family (impact on silicon die).

are attained for *HFI* glitches amplitude over 10000 DSN (thus a SNR of ≈ 100), for a usable energy distribution ranging from 10000 DSN to 400000 DSN.

The discrete nature of the (random uniform) glitch distribution in phase introduces statistical uncertainty on $\Delta\phi_p = n_p/n_f$. The maximum uncertainty is reached for two points separated by $\pi/2$, and scales with $1/\sqrt{n_f}$. In practice datasets contains at least 10000 glitches, thus the relative phase error is smaller than 1%. This is not significant compared to the error due to *pseudo energy* normalization.

3.3 Applications

3.3.1 Validation of the the empirical time transfer function model

Since the relative error level of glitch FIR in *HFI* configuration is about a few %, they are above the target goal of 0.2% and cannot be used directly as the time transfer function. However they have been an invaluable source of information for qualitative analysis of detectors response.

The comparison of the two early empirical time transfer function model candidates with the *short* glitch FIR have shown a much better agreement with the LFER model (version *LFER8v0*) than with the *JH10* model candidate. It has been a decisive argument in favour of the LFER model, which has been used for science data time deconvolution since the 2013 data release, but updated with more thermal parameters. An example of comparison between the *short* glitch FIR and model candidates is shown in Fig. 3.12 for channel `00_100-1a`.

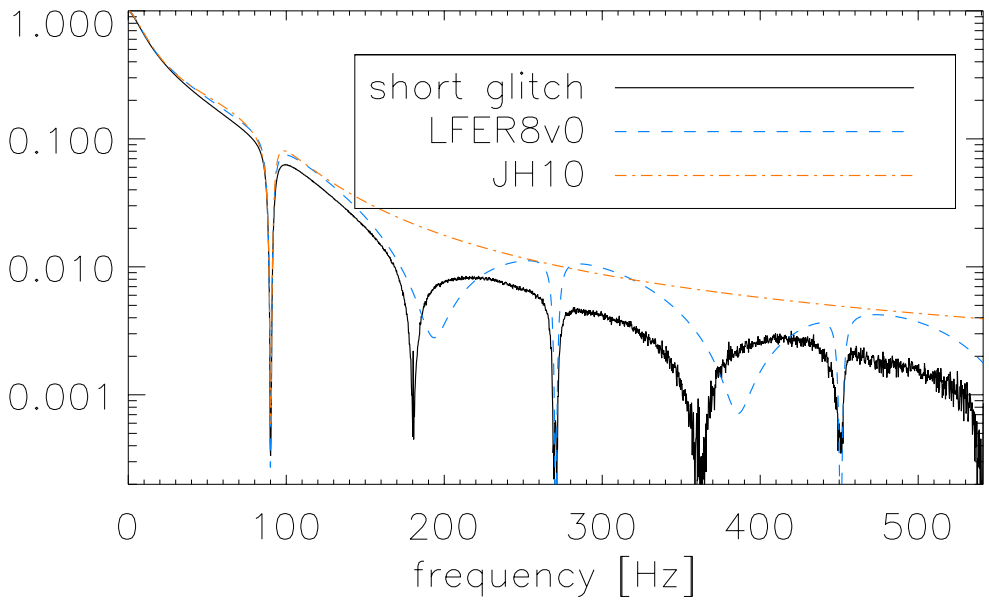


Figure 3.12: Comparison of *short* glitch FIR with two early time transfer function model candidates: the *JH10*, and the *LFER8v0* (using 3 thermal time constants at the time of this plot) normalized at 1 Hz. The *LFER8v0* is the time transfer function model which have been used for the *Planck* 2013 and 2015 data release. The difference between the *LFER8v0* model and the *short* glitch FIR is mostly due to glitch normalization with *pseudo energy*.

3.3.2 Analysis of thermal properties of detectors

The FIR construction method based on the projection in $\mathbb{R}^n$ has a higher separation power of glitch families than more classical methods based on time constant detection. The former takes advantage of both high and low frequency features, while the later are limited to the low frequency components of science data.

There has been some confusion at first between the *short* and *short*$_2$ families for channel frequencies higher than 100 GHz. This is due to the size of the detector grid which is the largest on the 100 GHz detectors, thus allowing a clear separation between *short* families. However this separation is not straightforward for higher frequencies. What is happening becomes clear with the use of the 2D projection. Let us have another look to the 2D projection shown previously in Fig. 3.8, this is the glitch cloud of a 143 GHz detector. The *short*$_2$ family is surrounded by a faint ribbon corresponding actually to the *short* family. When selecting higher energies the faint ribbon of the *short* family disappears while the *short*$_2$ distribution remains in place, which is a good indicator that we are in presence of very close thermal response that could not be distinguished otherwise.

The *short* glitch templates allows a direct estimation of the optical time transfer function, this property has been used to distinguish the location of energy deposit by the particle. The Fig. 3.13 shows the comparison of the optical time transfer function with the *short* and *long* glitch families (this figure has been published in **?**). It provides a strong demonstration that *short* glitches are energy deposit on the grid because the *short* glitch response is very consistent with the optical time transfer function model calibrated in-flight.

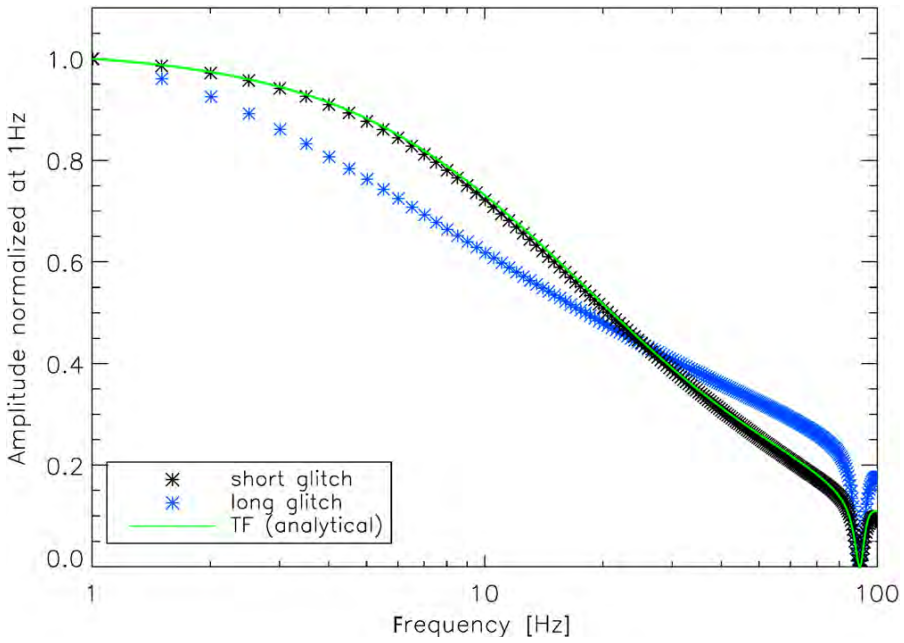


Figure 3.13: Comparison of glitch families with the optical time transfer function (version of the 2013 data release). The *short* family is completely overwritten by the optical time transfer function. The modulus of all frequency responses is normalized at 1 Hz.

3.3.3 High frequency transfer function constraints

While the thermal behavior of the detectors in response to glitches is subject to uncertainties depending on physical location of energy deposit, thermal path available and the trajectory

of the particle, the readout electronics filtering is the same for all glitches. I present in this section a preliminary work on the potential constraints that glitch FIR can provide.

Motivation

Using the *HFI* formalism for the time transfer function given in Eq. 2.1, the glitch FIR is characterized with $F(\omega)$ and the readouts electronic filtering with $H'(\omega)$ respectively. One of the components described by $H'(\omega)$ is a *Sallen-Key* High pass filter, which is setup with a unique time constant value $\tau_{\text{HP}} = 51$ ms for all channels. It has been suggested[4] to modify the value of τ_{HP} on 353 GHz channels to improve the residual ringing observed on the planet science data after time transfer function deconvolution.

With the LFER model formalism, the glitch FIR of each family can be written $F_f(\omega)H'(\omega)$. As all kind of thermal responses are filtered with $H'(\omega)$, the point is to check if glitch FIR can provide constraints on τ_{HP} by testing the consistency of the value obtained from each glitch family independently.

Model of the glitch FIR

It has been seen already that glitches are very consistent at the % level with the optical time transfer function modeled with Eq. 2.1. Hence this model can also be used to fit the glitch FIR for the estimation of τ_{HP}. There are two categories of free parameters:

- THERMAL: (a_i, τ_i) the couples (amplitude, time constant) of the low pas filters used in $F(\omega)$ (see Eq. 2.2);
- ELECTRONIC: τ_{HP} is tested along with S_{phase} which can be degenerate. The parameter τ_{stray} is always set free, because it strongly degenerates with the quickest time constant. The values of the electronic parameters and the expression for $H'(\omega)$ are provided in the *Planck* explanatory supplement[5].

After production of the FIR from the LFER model, *pseudo energy* normalization is applied to be consistent with the glitch data.

Results

For the thermal part $F(\omega)$, three low pass filters have been used (six free parameters) to adapt the thermal response of each glitch families. For the electronic part $H'(\omega)$, the influence of S_{phase} versus τ_{HP} had to be compared, so the four following cases in Tab. 3.1 have been considered.

	$\boldsymbol{\tau_{\text{HP}}}$	$\boldsymbol{S_{\text{phase}}}$
case 1	fixed	fixed
case 2	fixed	free
case 3	free	fixed
case 4	free	free

Table 3.1: Cases of free parameter considered

The χ^2 values obtained in each case is shown in Fig. 3.14 for the four test channels. By far τ_{HP} provides the most dramatic χ^2 reduction when set free, with a minor improvement when both τ_{HP} and S_{phase} are set free. The effect on the model agreement with the glitch

[4] A trace of this discussion can be found on the Tau Tiger Team portal. See `http://wiki.planck.fr/index.php/Proc/HybridTTT`

[5] See `https://wiki.cosmos.esa.int/planckpla2015/index.php/HFI_time_response_model`

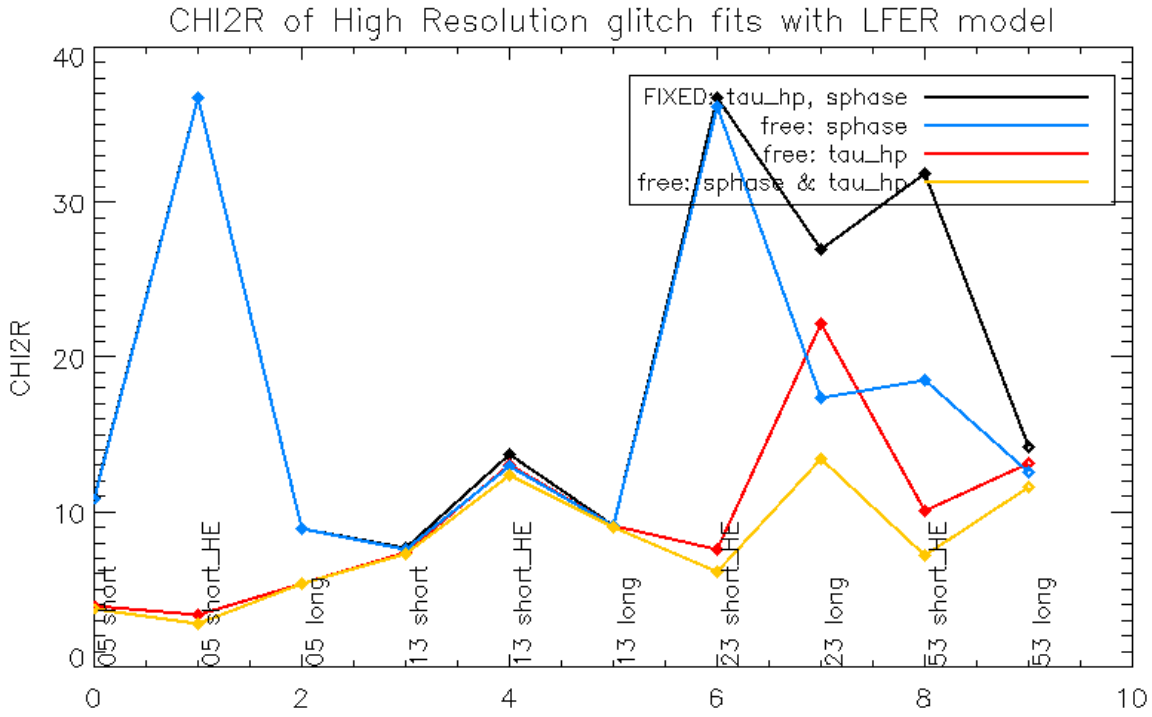


Figure 3.14: χ^2 values for different free parameters selection of the LFER model. Thermal parameters are always set free.

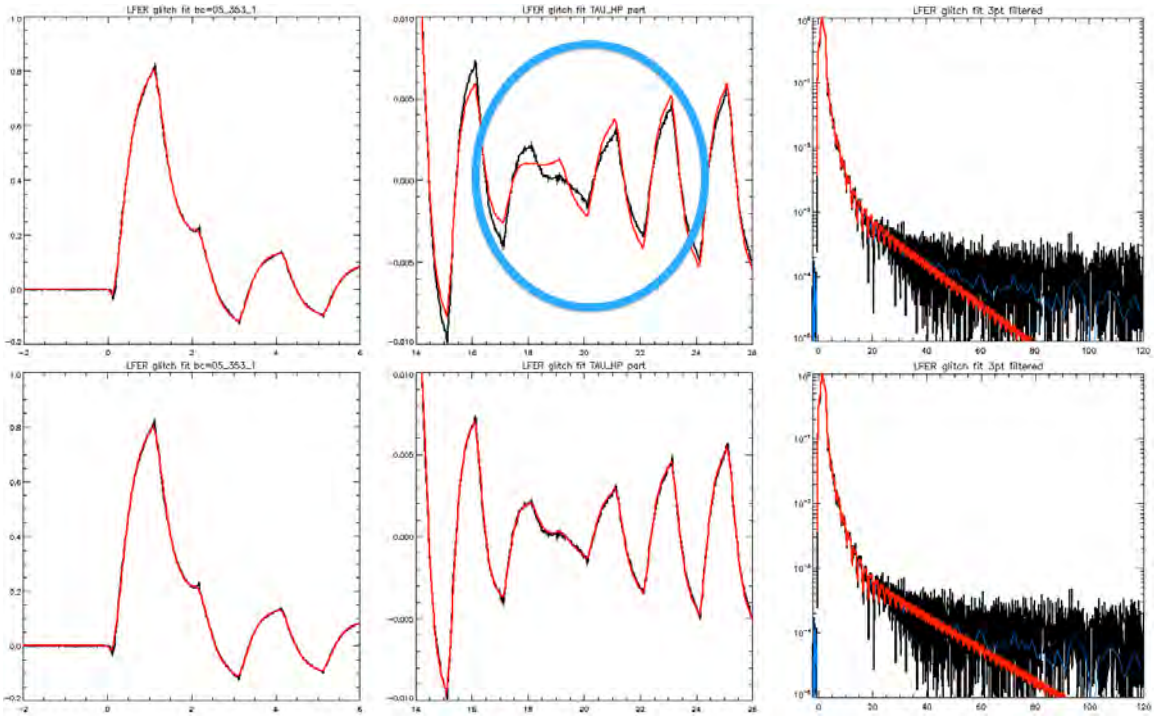


Figure 3.15: Example of fits of a *short*$_2$ glitch with the LFER model on channel `05_353-1`, in the case of τ_{HP} set to its nominal value of 51 ms (upper row) and set free (bottom row). From left to right: first six samples of the glitch FIR, glitch FIR in sample range [14, 26] corresponding to the location where τ_{HP} has the most visible impact, and glitch FIR in the sample range [0, 100] after convolution with the *Archeops* (**?**) three points digital filter [0.25, 0.5, 0.25] for visualization of the slow thermal component. The black line is the glitch FIR data, and the red line is the LFER model.

FIR is shown in Fig. 3.15. The visible effect of the $\tau_{\rm HP}$ parameter is a shift in time of the the location where there is a ringing sign inversion, it is near sample 19.

The best-fit values obtained for $\tau_{\rm HP}$ in case 4 are collected in Tab. 3.2. Error on $\tau_{\rm HP}$ is inferred from the difference in value obtained for each glitch family. In practice *short*, *short*$_2$ and *long* families could be extracted from the two first channels, for the two others channels, this was not possible due to *short*, *short*$_2$ beeing too close, thus I had to use a high energy widow to select only the *short*$_2$ family.

There is a very good agreement between the $\tau_{\rm HP}$ value estimated for each family as expected from the electronic filtering. Furthermore, I could cross-checked on channel `05_353-1` the value of $\tau_{\rm HP} = 53.03$ ms[6] obtained using the CPV `vbal` step sequence[7]. Both estimations are in one sigma agreement of the uncertainty estimated from glitches. Hence this preliminary approach shows that **glitch FIR can provide strong constraints on the electronic parameter $\tau_{\rm hp}$** and confirms the estimation method using CPV `vbal` step sequence.

channel	$\tau_{\rm HP}$
`05_353-1`	52.993 ± 0.17 ms
`13_353-2`	50.696 ± 0.09 ms
`23_353-3a`	52.545 ± 0.3 ms
`53_353-5a`	51.755 ± 0.2 ms

Table 3.2

3.4 Perspectives

At the time when this manuscript is written, there are still low frequency residuals between *HFI* sky surveys maps of the *Planck* 2015 data release (see Sec. 6.2.4 in Planck Collaboration VIII (2016)), which are an issue for science at large angular scale, mainly the study of the epoch of reionisation. The dominant source of differences between surveys is expected to be ADC nonlinearity, but non-Gaussian noise from the very long time constant of glitches residuals can also have an impact.

The results of the $\tau_{\rm HP}$ estimation shows that glitches from the three dominant families can be described to a very good extent by the LFER model on 353 GHz channels, and probably other frequencies. This may be used as an alternative for glitch removal on data, but I would go further by including the work on the detector response in Chap. 2.

The analysis of the low frequency response with the analytical model of the AC paper shows that it differs significantly from a low pass filter, and a component at the time scale of 30 to 60 seconds do appear (see Sec. 2.3). However the glitch templates used to clean the slow tail of *long* glitches is modeled with low pass filters, which estimation is limited to a few seconds. Hence a potential **non-Gaussian** residual noise power from *long* glitches with very long time constant from 30 to 60 s could be left on science data and may be removed by combining a more accurate thermal model describing the glitch FIR.

[6] This value has been obtained with a manual tweak resulting in a much better agreement between the data and the model, this is described in the Tau Tiger Team wiki. See `http://wiki.planck.fr/index.php/Proc/TTTetfCheck`

[7] the electronic response $H(\omega)$ can be measured by changing suddenly the value of the balance square used to reduce the signal dynamic of the detector modulation. This operation is roughly equivalent to the product of $H'(\omega)$ with a *Heaviside* step function in frequency domain, thus allowing estimation of $H'(\omega)$. The `vbal` step sequence has been run during the CPV and just before the EOL for this purpose.

3.5 Conclusions

Glitches have been considered in this chapter as a powerful tool for analysis and estimation of the readout electronics time transfer function and detectors thermal response.

A method has been described for reconstruction of the continuous time domain FIR corresponding to each glitch family. The glitch FIR built with this generic purpose method for the 100 GHz channels been found to be in agreement with the optical time transfer function estimated from in-flight data at percent level. Finally it has been shown that glitch FIR can provide accurate estimation for the electronics parameter τ_{HP} (but not limited to), and that the LFER model allow to describe with a good accuracy the different glitch families.

Part II

ADC characterization and correction

The ADC is a key component in the data acquisition chain of any instrument. It is the component which converts the analog signal from detectors into discrete numerical data samples. A perfect ADC would produce only *quantization* noise (see 4.2.2) on signal, but a real one also introduces nonlinearity, which is neglectable for most applications. The requirements for the CMB anisotropies study in *HFI* are very high ($\Delta T/T < 10^{-5}$). Unfortunately several elements in the design of the electronics readout chain have enhanced the effect of ADC nonlinearity. It resulted in apparent gain variations in time estimated to be greater than 1% of the solar dipole which have not been detected during on-ground validation campaigns. This is what has been called the ADC issue and-revealed an extremely challenging and time consuming problem. Furthermore the on-ground characterization of the chips has been insufficient, whereas for building a correction a very accurate characterization must be performed on chips of an already flying instrument. There are of course no dedicated on-board tools for this purpose, and the data acquisition chain is complex, thus the task is not easy.

Chap. 4 places the historical context leading to the ADC issue, and describes the characterization of ADC chip nonlinearity. After setting these basis it details the final in-flight characterization used for the 2013 and 2015 data releases. Next, the correction of science data is detailed in Chap. 5. For this purpose a simplified model of the electronics readout chain is developed. To assess the validity of the correction, specific ADC nonlinearity observables are studied, and the science data correction validated with simulations. Finally in Chap. 6 the effect of the 4 K stage mechanical cooler is taken into account. It's impact on ADC nonlinearity is characterized and the parasitic signal it generates is added to the correction model.

I have presented the content of the following chapters at the poster session 9914 of the SPIE *Astronomical + instrumentation 2016* conference. The corresponding proceeding (**?**) is a synthesis of this work and is available in Appendix B. The first characterizations of the ADC nonlinearity correction on *HFI* sky maps have been published in **?**. The 2015 version of the correction taking into account the 4 K lines is detailed in **?**.

Chapter 4

ADC chip characterization

"One hair is missing, and the world seems bald..."

Usually the digitization by the ADC within the electronic hardware is seen in a macroscopic fashion as noise, while in practice this noise is slightly non-Gaussian and introduces a bias which can be neglected. It happens that in the case of *HFI* the bias introduced by the ADC is significant compared to the astrophysical signal to be measured. As a consequence a very fine characterization of the ADC component is necessary, to correct for its inherent nonlinearity. On ground, this is a routine task which can be done with a dedicated acquisition setup. In the *Planck* case the spacecraft is operating at the second Lagrangian point, one million kilometers and half from Earth, where there is no opportunity nor budget for a service mission. As the chip cannot be brought back, the unusually fine characterization of its defects must be done in very creative ways from all the available piece of information available on Earth. This is what is related in this chapter.

The ADC nonlinearity has been a big issue in the *HFI* mission and it deserves an historical overview which is related in Sec. 4.1. After that, a synthesis of the chip algorithm and how its nonlinearity can be characterized is given in Sec. 4.2. The very first ADC nonlinearity characterization is detailed in Sec. 4.3, it is estimated from limited in-flight data acquired during the *HFI* mission. The following on-ground characterization made on *HFI* flight spare electronics in Orsay is related in Sec. 4.4. Then the knowledge of the architecture of *HFI* ADC is strengthened with the analysis of the `SPICE`[1] model from Maxwell Technologies in Sec. 4.5, and with a second finer characterization done in Cardif which is described in Sec. 4.6. Sec. 4.7 details how the nonlinearity of each detectors chips is finally characterized in-flight, and in Sec. 4.8 how the accuracy of these characterizations can be inferred from in-flight data.

It must be stressed that all this work has been done with a feeling of emergency, as the *Planck* spacecraft had completed its mission and was near its definitive end of operations. So the necessary in-flight measurements had to be scheduled quickly, while communications with the spacecraft were still possible.

4.1 Historical overview of the ADC issue

On-ground qualification : the back-door is open

The *HFI* readout electronics has been designed at IRAP (formerly known as CESR). In itself, the design and implementation of the full electronics readout chain is a great success, with high performances in terms of noise level stability and low frequency noise. The radiation

[1] `SPICE` (Simulation Program with Integrated Circuit Emphasis) is a free software for the simulation of analog electronics.

hardened ADC chips passed successfully on-ground qualification tests in 2004, but at this step, the nonlinearity measurements were set up only for probing large scale defects. An example of on-ground qualification tests output is shown in Fig. 4.1. It can be seen in the right text box that the number of sampling points is 50 while the *HFI* chip presents 65536 output values, so this acquisition probes only large scale nonlinearity above the noise level of this measurement. The ADC chips were rightly considered within the specifications at this time.

Sec. 4.2.1 will detail how ADC nonlinearity can be characterized in a more accurate way.

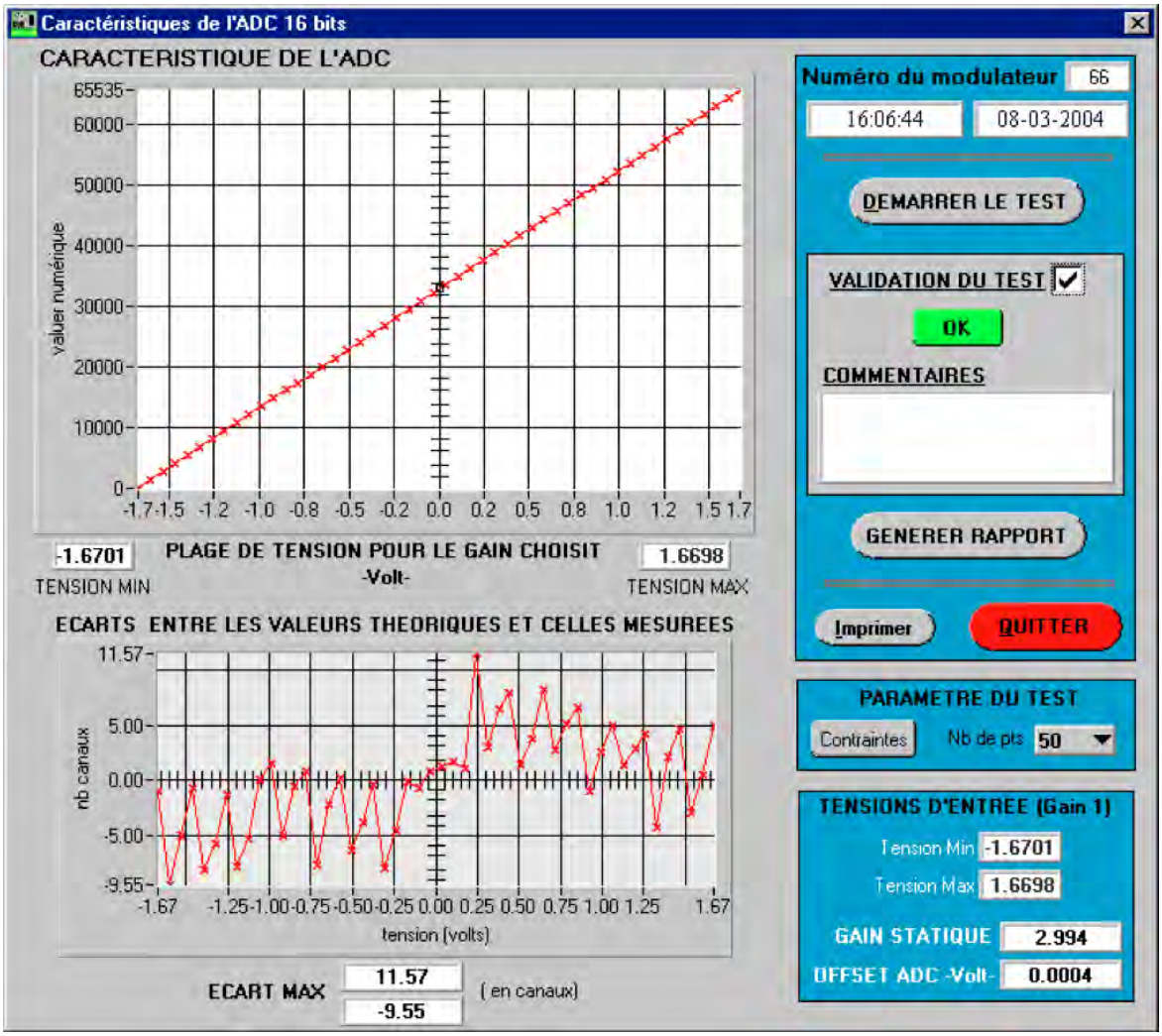


Figure 4.1: Example screenshot of the on-ground ADC characterization for one chip of *HFI*, this was run at *IRAP* (formerly known as *CESR*) in March 2008. *Top*: measurement of the ADC transfer function showing a straight line as expected, X axis is the input voltage in the range [-1.7, +1.7] , and Y axis is the returned ADC digital value for the full ADC scale [0,65535]. *Bottom*: deviation from linearity, X axis is the input voltage in the range [-1.67, +1.66], and Y axis unit is in number of digits.

The variable gain mystery

The first hint of the ADC issue appeared in In mid-2010 after one year of observation, when two complete surveys of the sky were available (`S1` and `S2`), which should be identical in theory, excepted concerning the noise. However, in the sky survey difference map built as the difference between `S1` and `S2` maps, large residuals appears following the scanning

strategy. These residuals are consistent with an apparent gain variation in time of about ±1.5% of the solar dipole (see Sec. 6.2 in **?**). These residuals were partially corrected by the application of a time-variable gain, estimated with the `BogoPix` software (Tristram et al. 2011), but this was only a first order correction, as the primary source of gain variation was in fact (as understood later) the ADC nonlinearity.

An example of a survey difference map is shown in Fig. 4.2, before and after application of the time-variable gain correction at first order. In theory the difference should be dominated by noise, but there are large scale features on the survey differences.

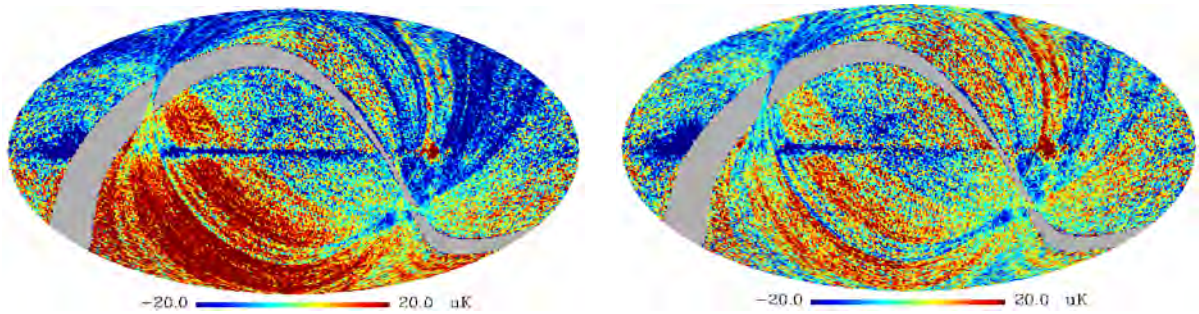


Figure 4.2: Survey sky map differences (Survey1-Survey2) for the channel `82_143-4a`. *Left*: before application of `BogoPix` time variable gain correction. *Right*: after application of `BogoPix` time variable gain correction. Color scale range from -20 $\mu\mathrm{K_{CMB}}$ (blue) to $+20\mu\mathrm{K_{CMB}}$ (red). These maps are reproduced by permission of Olivier Perdereau

ADC nonlinearity evidence from *fast samples* histograms

The origin of the effect was finally understood in December of 2011, when, after a discussion initiated by Jean-Michel Lamarre, the instrument PI of *HFI*: the ADC was noted as a potential source of variable gain because of the nonlinearity of its response.

Inspection of *fast samples* over the full *HFI* mission allowed to build histograms of ADC output values as shown in Fig. 4.3. The histogram spread is mainly caused by the thermal drift of the *HFI* bolometer plate during the mission. This thermal drift is very slow in time, so a very smooth histogram is expected, which is not the case. The high frequency jumps are caused by the ADC nonlinearity, i.e. the local difference between histogram bins reflects the inhomogeneous range in volts mapped by the corresponding integer values output by the ADC (see Sec. 4.2.1). The central bin (labeled code 32768) is nearly empty, indicating that it maps a voltage range much smaller than the average. It is the primary cause of nonlinear features visible on sky maps since it is about 10% in size compared to the amplitude of the main cosmological signal, i.e. the solar dipole the which upper and lower peak locations are shown as red lines on the ADC scale.

Dipole size compared to ADC nonlinear features

The on-ground ADC nonlinearity measurements could have been sufficient without two other factors:

1. The signal balancing was very efficient and resulted in a very flat signal acquired exactly at midscale of the ADC. Albeit this was the goal of balancing it happened that the signal was concentrated exactly where is located the biggest defect (more details in Sec. 4.2.1);

2. The CMB signal dynamic range is dominated by the solar dipole, covering a very small range compared to the central ADC defect size. The solar dipole peak-to-peak

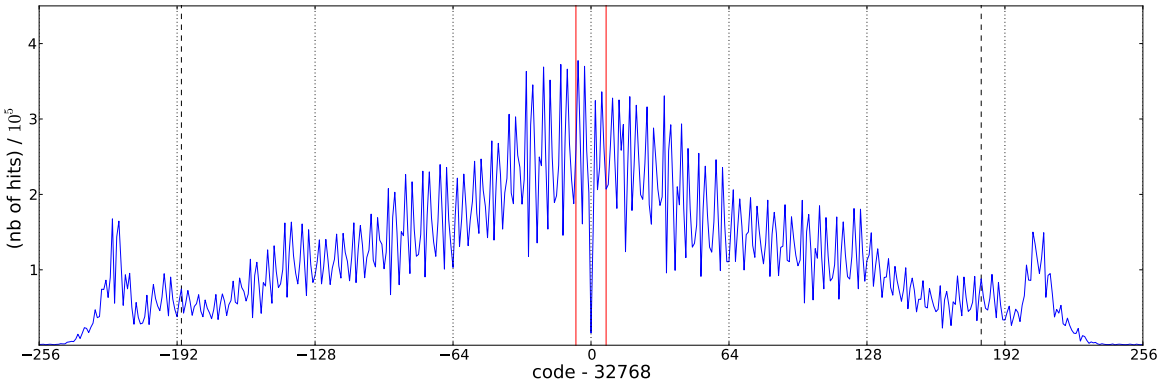


Figure 4.3: Histogram of the ADC output values over the full *HFI* mission for channel `00_100-1a`. The dashed black lines are the location of the 5% and 85% data percentiles. The red lines indicate the boundaries of the signal covered by the solar dipole.

amplitude represents on average about 10 quantization steps of the ADC, while the central defect size is about 1 quantization step.

Unfortunately, the combined effect resulted in a huge amplification impact of ADC nonlinearity on science data.

The beginning of my contribution

My contribution to the ADC issue starts at the end of *HFI* mission, when the proof of ADC implication in the time-variable gain has been shown with histograms of *fast samples*.

For the planck ADCs, several solutions had been tested for the fine characterization of the 54 on-board ADC chips used by the detectors channels. My first intuitive approach for in-flight ADC characterization was to use *fast samples* acquired during the *HFI* mission. This dataset was insufficient for the correction, but revealed that there was a generic pattern for the nonlinearity of the ADC chips used in *HFI*. This result pushed the ADC team to setup an in-flight acquisition campain dedicated to ADC nonlinearity characterization. Its output was expected to be used with a constraining model to estimate accurately ADC nonlinearity. At this moment, there was absolutely no data, at the required accuracy, available on Earth for this specific chip, but we needed a basis to work on the constraining model hypothesis. For this purpose, an acquisition campaign on the *HFI* electronics readout flight spare has been scheduled. The acquisition has been performed at IAS in Orsay were was stored the flight spare. It resulted in the first available acquisition made on a large scale of four ADC chips and confirmed observations made on in-flight data, but the full chip scale was still unobserved. It was decided at this moment to send two remaining spare chips to Rashmi V. Sudiwala at Cardiff University, who offered his help to make the acquisition on the full chip scale. We interacted several time until the right setup allowed to make an accurate characterization of the two chips. The Cardiff acquisition campain allowed to build a definitive dataset for nonlinearity constraint analysis. In the meanwhile a work has been performed on the `SPICE` electronic model of the chip (see Sec. 4.5), the results from this analysis allowed to confirm a general model for the ADC nonlinearity pattern. Using the constraints provided by the on-ground acquisition campaigns and the `SPICE` analysis, François Couchot and Guillaume Patanchon worked on the theory to implement a maximum likelihood method for ADC nonlinearity characterization. Finally, the last in-flight acquisition campaign has been setup and the output data has been used succesfully to produce the version of ADC accurate characterization.

4.2 Understanding the ADC of *HFI*

4.2.1 Basics of a Successive Approximation Register (SAR) ADC

The chip used in *Planck HFI* readout electronics is a SAR ADC from *Maxwel Technologies* referenced as `7809LPRP`. The chip specifications sheet from *Maxwel Technologies* is provided in Appendix C. To understand how an SAR ADC does work, I present here a synthesis of the online document from *Maxim Integrated* available at `https://www.maximintegrated.com/en/app-notes/index.mvp/id/1080`. The following pictures in this section come from the same source.

MSB and LSB

A wide use is made of the terms Most Significant bit (MSB) and Least Significant Bit (LSB) hereafter. They are related to the binary output values (codes) of the ADC. The LSB is the rightmost bit of these binary values, coding for one quantization step of the ADC scale. It is used hereafter as the unit of the ADC output values, but ADU are also in use instead of LSB in *Planck* Collaboration.

The MSB is the leftmost bit of the ADC output values it codes for half of the scale voltage of the component. In the ADC hardware it corresponds to the capacitor with the biggest capacitance value, thus it is the most susceptible to calibration errors, as detailed below.

Algorithm

Basically a n-bit SAR ADC converts an input voltage V_{IN} into an output binary value coded on n bits. This is done by dichotomy, over the fixed voltage input range $[0;\ V_{\mathrm{REF}}]$: the Digital to Analog Converter (DAC) of the SAR generates a tension V_{DAC} which is compared at each step to V_{IN}. The dichotomy operation is shown on Fig 4.4 where the algorithm is processed with a 4-bits ADC. At the first iteration, $V_{\mathrm{DAC}} = V_{\mathrm{REF}}/2$. If $V_{\mathrm{IN}} < V_{\mathrm{REF}}/2$ then the Most Significant Bit (MSB) b_0 is set to zero, otherwise it is set to 1. The same method is used to set the second bit value by comparison of $V_{\mathrm{DAC}} = b_0 \times V_{\mathrm{REF}}/2 + V_{\mathrm{REF}}/4$ to V_{IN}, and so one until the Least Significant Bit (LSB) is reached. If we consider that the input voltage range $[0;\ V_{\mathrm{REF}}]$ is divided into 2^n quantization steps, then the output code is the index of the quantization step containing the input voltage V_{IN}.

Architecture

The global architecture of an SAR ADC is presented in Fig. 4.5. While the ADC is operating, the input voltage V_{IN} is kept constant by the *Track/Hold* function. For the considered chip, the *Track/Hold* function is provided by the capacitive Digital to Analog Converter itself (DAC) shown in Fig. 4.6. In short the capacitive DAC operates with charge redistribution as follows:

1. ACQUISITION: V_{IN} acquisition is done by connecting the common terminal to the ground and connecting the bottom plate of the capacities to V_{IN}.

2. HOLD: After that the common terminal is disconnected from the ground, and the switches of all capacitors are connected to the ground. Now $V_{\mathrm{COMMON}} = -V_{\mathrm{IN}}$ and each capacitor has trapped a charge proportional to its capacitance. Let's note here that the dummy capacitor is present only to have a total capacitance equal to 65536C.

3. COMPARISON The switch of the first bit (MSB), is connected to V_{REF}, so now $V_{\mathrm{COMMON}} = -V_{\mathrm{IN}} + \frac{1}{2}V_{\mathrm{REF}}$

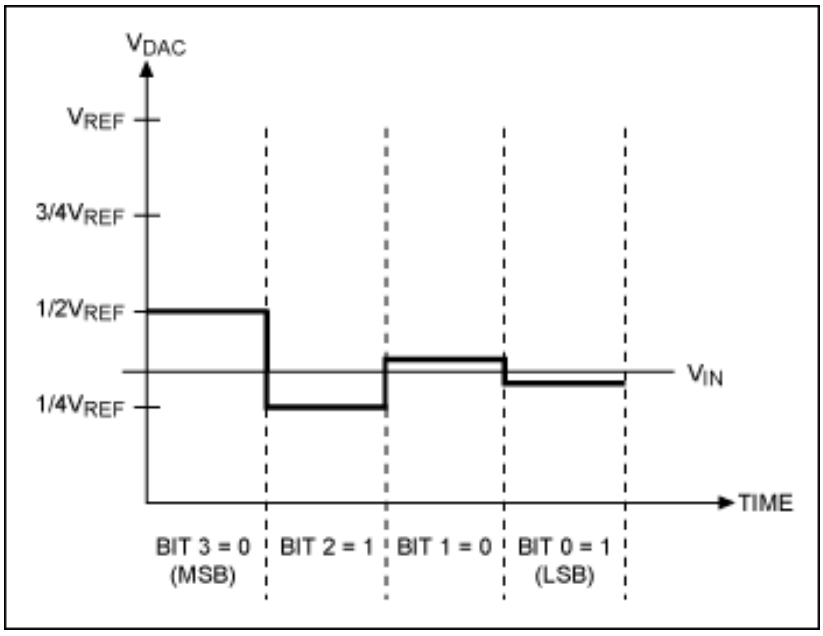


Figure 4.4: Example of SAR operation with a 4-bit ADC. V_{DAC} is the voltage generated by the DAC at each iteration to be compared with V_{IN}. In this case the output code is the binary value 0101.

4. BIT SELECTION: The comparators outputs a logic zero if $V_{\mathrm{COMMON}} < 0$, then the current capacitor switch is connected back to the ground. Otherwise it remains connected to V_{REF} and the comparator outputs a logical one.

The comparison and bit selection steps goes on until the last bit (LSB) is reached. At the end of the process we have

$$V_{\mathrm{COMMON}} = -V_{\mathrm{IN}} + \sum_{k=1}^{n} \frac{b_k}{2^k} V_{\mathrm{REF}}\,, \tag{4.1}$$

where b_k is the value of the k^{th} bit output by the *SAR Logic* function.

A non-ideal capacitive array

With non-ideal components, the relative value of the capacitances are not exactly powers of two. Thus the DAC voltage deviates from the ideal expression of Eq. 4.1, which is the cause of *high frequency nonlinearity*. Hopefully the most significant bit capacitors are production trimmed[2] to reduce errors to some extent. However this process is not perfect, leading to the kind of non-ideal behavior observed in the histogram in Fig. 4.3. So this is exactly where the nonlinear effect, that is dealt with in this manuscript, comes from.

4.2.2 ADC nonlinearity characterization in general

The formalism detailed below, will be used all along this chapter, when describing the various approaches used to characterize the ADC nonlinearity on-ground and in-flight.

[2] Trimming is a technique occurring after calibration of the component, it consists of adding or removing small fixed capacitances values to the array, with either a laser beam or by storing the appropriate states of the switches in a memory.

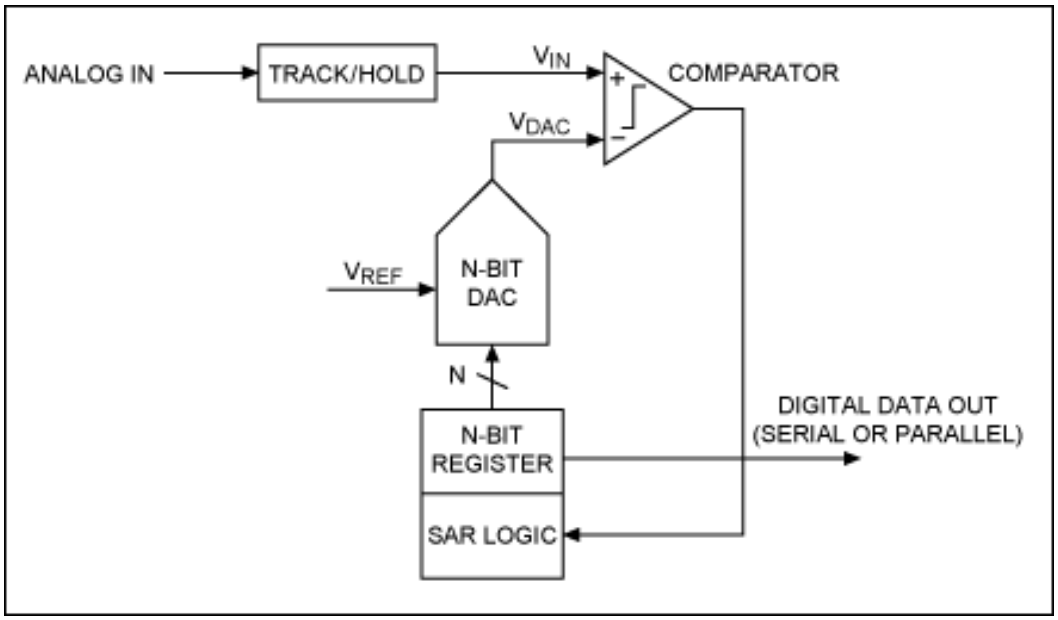


Figure 4.5: Simplified n-bit SAR ADC architecture.

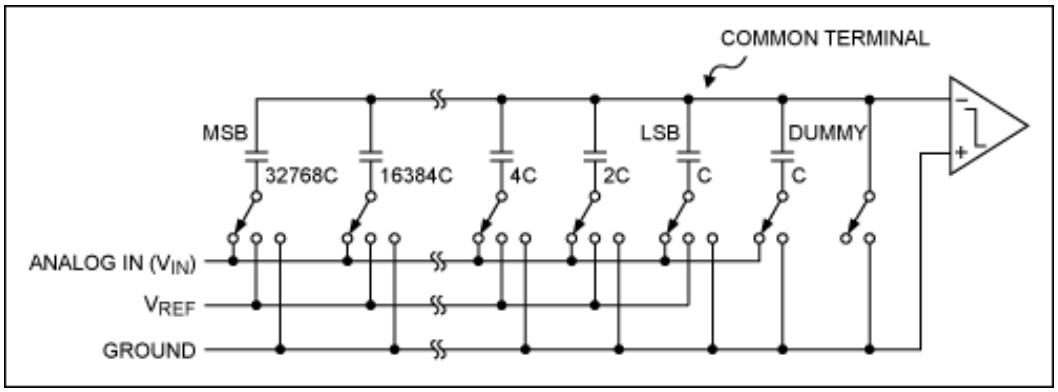


Figure 4.6: A 16-bit example of a capacitive DAC.

Quantization noise

For an ideal ADC, there are only rounding errors resulting from the conversion of the analog input into a positive integer value K (called a *code*). For a n-bit ADC the voltage input range $[0; V_{\mathrm{REF}}]$[3] is quantized into 2^n quantization steps of size $Q = V_{\mathrm{REF}}/2^n$. In practice the ADC does the operation $\mathrm{floor}(V_{\mathrm{IN}}/Q)$, and it can be shown easily that the Root Mean Square (RMS) error equals to $Q/\sqrt{12}$.

The theoretical signal-to-noise ratio (SNR) of a n-bit ADC is calculated with a full scale sine wave, thus its amplitude is $2^nQ/2$. As this value is generally expressed in decibels it

[3] In the general case the voltage range can contain negative values but a simple case is considered for the demonstration

reads

$$\begin{aligned} \mathrm{SNR} &= 20\log_{10}\frac{\text{RMS sine wave}}{\text{RMS quantization noise}} \\ &= 20\log_{10}\frac{2^n Q/2\sqrt{2}}{Q/\sqrt{12}} \\ &= 20\log_{10} 2^n \frac{\sqrt{3}}{\sqrt{2}} \\ &= 6.02n + 1.76\mathrm{dB}\,. \end{aligned} \tag{4.2}$$

The quantization noise is the minimal estimate of the noise introduced by any ADC chip. Also, for real non-ideal components, a small amount of thermal noise from internal capacitors and resistors is introduced. This is called *input-referred noise* and has a RMS value which is in general about 1 LSB.

The frequency spectrum of the quantization noise is quite complex to analyze (**?**), but in the general case the Eq. 4.2 estimate of the SNR is sufficient. In the case of *HFI* the typical frequency of the astrophysical signal is very low (0.16 mHz for the solar dipole) compared to the acquisition rate (7200 Hz), thus no frequency distortion of the astrophysical signal by the ADC is expected.

Transfer function and Code Transition Points (CTP)

The transfer function of an ADC has a stairway like shape, because the output code κ is constant inside its quantization step $Q(\kappa)$. So the transfer function can be defined only with the knowledge of the locations $v(\kappa)$ (called CTP) where the value of the output code changes. The quantization step size for code κ is related to the CTP by

$$Q(\kappa) = v(\kappa+1) - v(\kappa)\,, \quad \kappa \in [\![0, 2^n - 1]\!]\,, \tag{4.3}$$

considering $v_0 = 0$ and $v_{2^n} = V_{\mathrm{REF}}$ for an n-bit ADC.

An example of a non ideal 4-bit ADC transfer function is shown in Fig. 4.7. In this simple example, there are three defects: the quantization steps of codes 4 and 12 have a 1.4 LSB size, and the quantization step of code 8 (midscale) have 0.2 LSB size.

Differential NonLinearity (DNL)

While the transfer function is perfectly suitable when correcting for nonlinearity, for its characterization it is preferable to use a tool able to locate explicitly the biggest defects. This is generally done using the DNL which is the deviation from the ideal value Q for each step. The expression of DNL for code κ is

$$\mathrm{DLN}(\kappa) = \frac{Q(\kappa)}{Q} - 1\,, \tag{4.4}$$

where $Q(\kappa)$ is the size in volts of quantization step for code κ.

The DNL is expressed in LSB units (1 LSB = Q) and is generally not greater than 1, in absolute value, for a well designed ADC. Larger values lead to non-monotonic transfer function and to what is called *missing codes*. See Fig. 4.8 for an example of DNL calculated with the same 4-bit ADC as presented in Fig. 4.7.

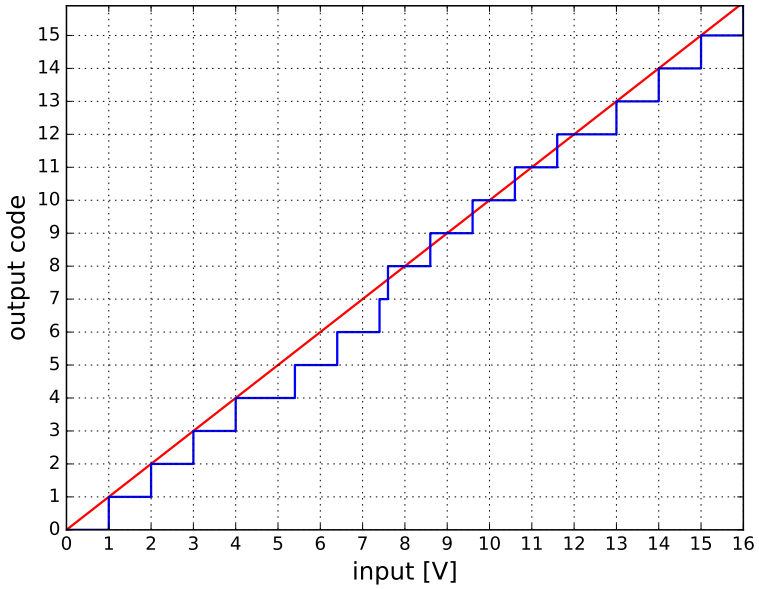


Figure 4.7: Example of a non-ideal 4-bit ADC transfer function.

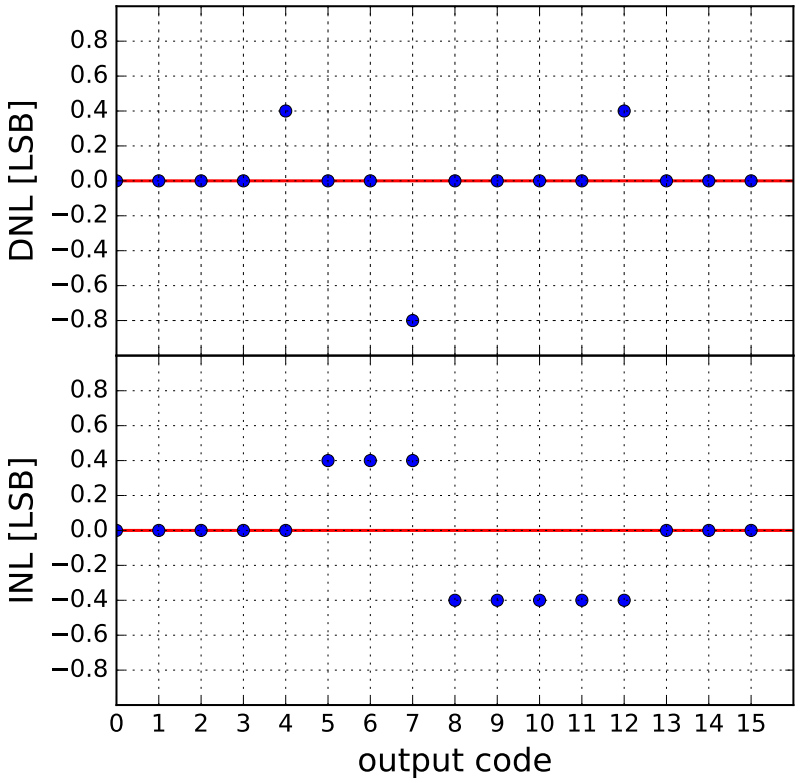


Figure 4.8: *Top panel*: DNL for the same 4-bit demonstration chip as in the previous Fig. 4.7, *Bottom panel*: corresponding INL.

Integral NonLinearity (INL)

The INL is the deviation of the ADC transfer function from the ideal linear output, expressed in LSB units. This is an important specification for an ADC, because unlike the DNL which is local, the INL accounts for large scale errors. There are generally two definitions of the INL:

1. BEST-FIT LINE INL: In this case the deviation is considered from the best-fit straight line. This is the method that gives the best results;

2. END POINTS INL: the straight line passes through the end points of the transfer function (see example Fig. 4.8).

It is the second definition that will be used for computational reasons. Hereafter INL at code K will be defined by

$$\mathrm{INL}(\mathrm{K}) = \frac{\sum_0^{\mathrm{K}} Q(\mathrm{K})}{Q} - \mathrm{K}\,, \quad \mathrm{K} \in [\![0, 65535]\!]\,, \tag{4.5}$$

where $Q(\mathrm{K})$ is the size in volt of the quantization step for the code K and Q is the mean quantization step size. However, the in-flight fit procedure already reconstructs a best-fit line, in practice, for the small visible scale of ADCs.

The INL errors have two components (**?**) that can be modeled independently:

- A Low Code Frequency component (LCF), caused by ADC analog preprocessing. This component is smooth and can be modeled with a polynomial fit.

- A High Code Frequency component (HCF), caused by the capacitances weight mismatch of the DAC within the ADC. The HCF component is multi-periodical and can be modeled using Rademacher functions (or Haar wavelet, see Eq. 4.16) with n parameters for a n-bit ADC.

The LCF component is relevant when working with a large portion of the ADC scale, which will be the case in the Cardif setup (see Sec. 4.6). For *HFI* in-flight acquisition, only a small portion of the ADC is useful for science data and the focus will be be given on HCF nonlinearity (see Sec. 4.4 and Sec. 4.7).

DNL estimation with ramp

The ideal way of characterizing an ADC nonlinearity to use an input ramp covering the full range of the ADC, ensuring $\|dV_{\mathrm{IN}}/dt\| = \mathrm{cst}$. In practice this operation is realized with a very stable triangle wave which is run over a great number of periods, in order to obtain enough statistics.

If the input noise is neglected, the DNL expression is simply

$$\mathrm{DNL}(\mathrm{K}) = \frac{h(\mathrm{K})}{\langle h \rangle}, \tag{4.6}$$

where h is the histogram of the output codes of the ADC. The relative uncertainty on $\mathrm{DNL}(\mathrm{K})$ is $\sim 1/\sqrt{h(\mathrm{K})}$. If the triangle wave is perfectly linear, there is no large scale errors as we do have with measurements using Gaussian noise (See next section).

This method can also be applied to absolute calibration if the value of $\|dV_{\mathrm{IN}}/dt\|$ is well known, but in the context of *HFI* it is relative calibration which is relevant and this will be the considered case hereafter.

DNL estimation with Gaussian noise

A white Gaussian noise is also usable to do ADC nonlinearity characterization because it's Probability Density Function (PDF) is well known. Moreover it is the only accessible source in the context of *HFI*. In this case the method consist of using the reciprocal of the source PDF to obtain the relative quantization step size. It can be shown analytically that the mean value and variance of the source do not need to be known.

Let us consider a stationary white noise source described by a PDF $\mathcal{N}_{\mu,\sigma}(V_{\mathrm{IN}})$ with a noise RMS value of σ and a mean value μ. The histogram of code hits for an acquisition session is $h(\mathrm{K})$, it can be converted to a discrete Cumulative Distribution Function (CDF)

$$H(\mathrm{K}) = \frac{\sum_{k=0}^{\mathrm{K}-1} h(k)}{\sum h} \,. \tag{4.7}$$

By construction $H(\mathrm{K}) \in [0,1]$ and is described by the normal CDF which expression is given by

$$\Phi\left(x\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(x\right)\right] \,. \tag{4.8}$$

Hence for a large number of data samples $H(\mathrm{K})$ tends to the probability of V_{IN} to fall below the CTP $v(\mathrm{K})$, which reads

$$H(\mathrm{K}) \simeq \Phi\left(\frac{v(\mathrm{K}) - \mu}{\sigma}\right) \,. \tag{4.9}$$

From this expression it comes

$$\frac{Q(\mathrm{K})}{\sigma} \simeq \Phi^{-1}\left(H(\mathrm{K}+1)\right) - \Phi^{-1}\left(H(\mathrm{K})\right) \,. \tag{4.10}$$

The expression in Eq. 4.10 is an interesting result because it shows the value of μ does not need to be known to estimate the quantization step size of the ADC normalized with σ. A schematic view of this operation is shown in Fig.4.9. Now in the case of *HFI* the absolute calibration is made with the solar dipole, thus an absolute calibration of the ADC scale is not necessary and the value of σ does not need to be known. However it will be important to check the noise stationarity hypothesis.

Unlike the ramp method, the ADC nonlinearity characterization with Gaussian noise is subject to cumulative error due to the variance of the number of hits in the range covered by two distant codes.

4.2.3 Target accuracy for INL characterization

It is necessary to know the required level of accuracy for ADC nonlinearity characterization to provide the adequate correction to keep this systematic effect at bay (the propagation of ADC nonlinearity to science is covered in 7). For an ADC nonlinearity correction it is the INL which is homogeneous with the transfer function, thus it will be mainly to the INL that will be referred to hereafter when talking about ADC nonlinearity characterization.

To infer the accuracy required for INL estimation let us consider first the maximum tolerance for the gain variations it introduces. The ADC nonlinearity propagates to science via the solar dipole residuals, thus a simple rule of thumb to estimates the relative gain tolerance $\Delta g_{\mathrm{tol}}/g$ is with

$$\frac{\Delta g_{\mathrm{tol}}}{g} = \frac{\Delta T_{\mathrm{noise}}}{\Delta T_{\mathrm{dipole}}} \,, \tag{4.11}$$

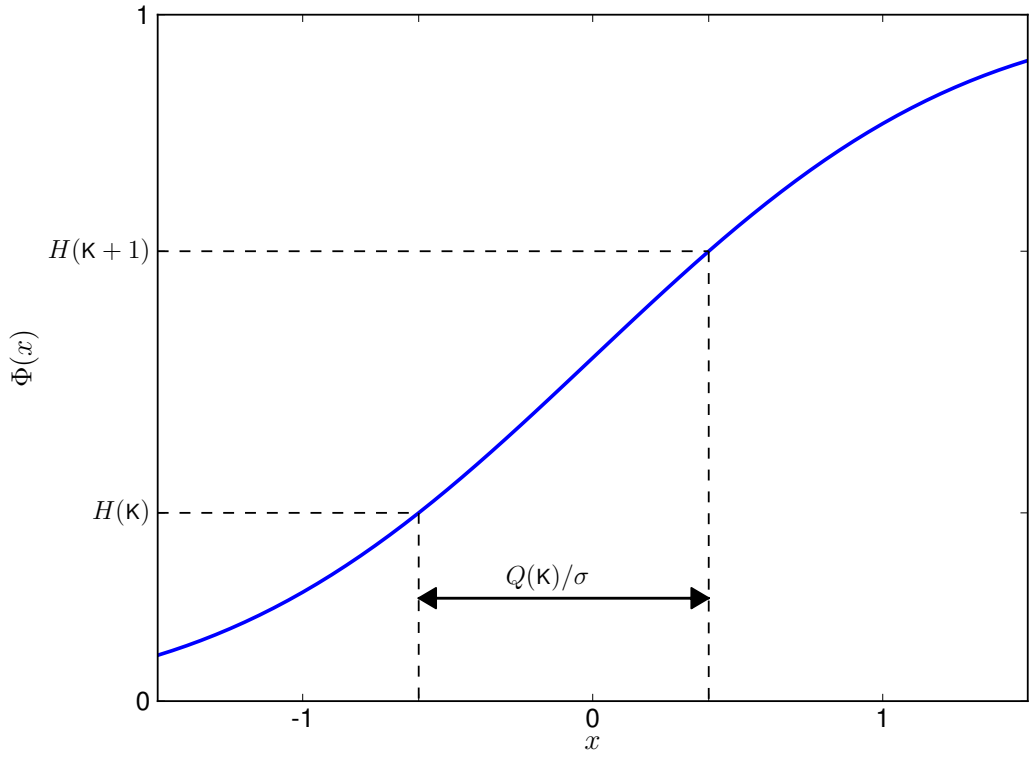


Figure 4.9: Normal CDF annotated with the local quantization step size estimate for code K. X axis is the signal domain in units of σ, and Y axis is the probability domain.

where ΔT_{noise} is the noise level on the sky maps and ΔT_{dipole} is the solar dipole amplitude. Taking as reference the estimated noise level from half ring differences[4], the noise level at large angular scales is $\Delta T_{\text{noise}} \approx 0.1\ \mu\text{K}_{\text{CMB}}$. With the solar dipole amplitude being $\Delta T_{\text{dipole}} = 3.3\ \text{mK}_{\text{CMB}}$, the target is to reach a maximum relative gain variation level of $\approx 0.3 \times 10^{-4}$.

The conversion of the tolerance on gain variations to an accuracy level for the INL estimation in LSB units is a very complex task. This is because the ADC nonlinearity is by definition a highly nonlinear effect at small voltage scales which in combination with a complex data acquisition chain, is very difficult to predict. However a simple heuristic can be inferred from the observed level of gain variations (without correction) during the mission. Assuming that the $\approx 10^{-2}$ relative gain variations are mainly caused by the central chip defect of ≈1 LSB, **the ADC nonlinearity has to be estimated with an accuracy of** $\approx 0.3 \times 10^{-2}$ **LSB** to reach the target tolerance level for gain variations.

Let us be more specific about the accuracy requirement on the INL. The input sky power variations ΔP_{sky} translates into an analog voltage Δv, and after digitization by the ADC into a variation on *fast samples* of Δs. In other words, the ADC nonlinearity error on Δs is the sum of local errors on the quantization steps size covered by Δs. Hence the target accuracy level of 0.3×10^{-2} LSB represent the maximum deviation of the estimated INL from the ADC true INL. This accuracy level must be enforced on the full range of the ADC (≈ 400 codes) covered by the cosmological signal over the two and a half years of the *HFI* mission.

[4] Half ring differences are a noise estimation method which consist to cancel the sky signal from the first half of the stable scanning period with the second half (see Sec. 6.2.4 in Planck Collaboration VIII (2016))

4.3 INL estimation from cold *fast samples*

A first quick approch to do ADC characterization is presented in this section. It has been done using the cold *fast samples* from the two and a half year of the *HFI* mission. The cold *fast samples* are named by comparison with the data acquisition after the *HFI* end of life when the detectors were warm.

4.3.1 Slow thermal variations

The slow thermal variation of the detectors plate, presents the interesting property of providing mostly a monotonic ramp over the mission. An example of the thermal drift is shown in Fig. 4.10, where the DC level is the average value of demodulated science data in digital units (DSN). The dipole dynamic is small (less than 10%) compared to the thermal variations, and does not appear on the figure, because its averaged value is zero over one ring. A few visible jumps are due to the sorption cooler switch (near ring index 10000) and to solar flares. These events have a very short timescale compared to the length of the mission.

The sky signal cannot be used as a calibration source because the deviation from linearity is too important at medium and long signal range. However the DC level is very smooth, with slow variations in time, ensuring a good coverage of the ADC scale used by the *HFI* mission. These properties allow us to make the assumption that the *fast samples* noise can be used as a calibration source, by assuming that the RMS of the *fast samples* noise is stationary throughout over the mission.

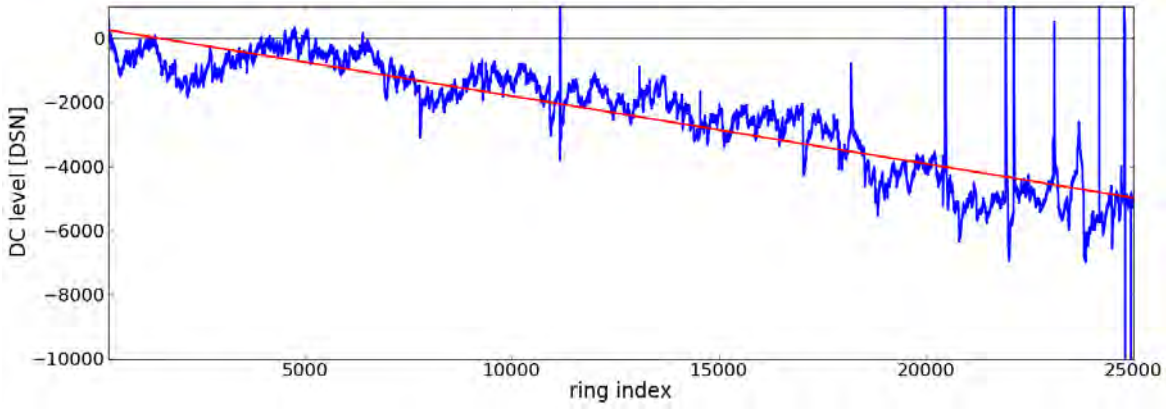


Figure 4.10: DC level variations (blue) over the mission on a per ring basis for channel `00_10-1a`. The red line is the linear fit of the data.

4.3.2 Estimation of INL with Gaussians

The estimation of the ADC nonlinearity on a given channel is a sequential procedure, stepping through each value of the demodulated science data which is expressed in DSN units from the summation of the *fast samples* integer codes values. The working hypothesis is that for a given value s_d of the demodulated science data, the PDF of the analog signal for the set of *fast sample* periods $\{\boldsymbol{p}\}|_{s_d}$ yielding the value s_d is Gaussian at each index of the period. Additionally the *fast samples* noise is considered stationary throughout the mission, and its RMS value is noted σ. From this consideration the deviation of the distribution of the *fast samples* codes from a Gaussian can be used to estimate locally the ADC nonlinearity. At the end of the procedure, all local weighted estimates are combined to produce the global INL.

Let us see an example of step on Fig. 4.11 where the *fast samples* period (a) is the average of the set of periods $\{\boldsymbol{p}\}|_{s_d}$ yielding the demodulated science data value s_d. The

demonstration histogram (b) is the distribution of the *fast samples* $\{\boldsymbol{p}(30)\}|_{s_d}$ captured only at index 30 of the period. Their distribution is affected by the non homogenous range in volt covered by each ADC code.

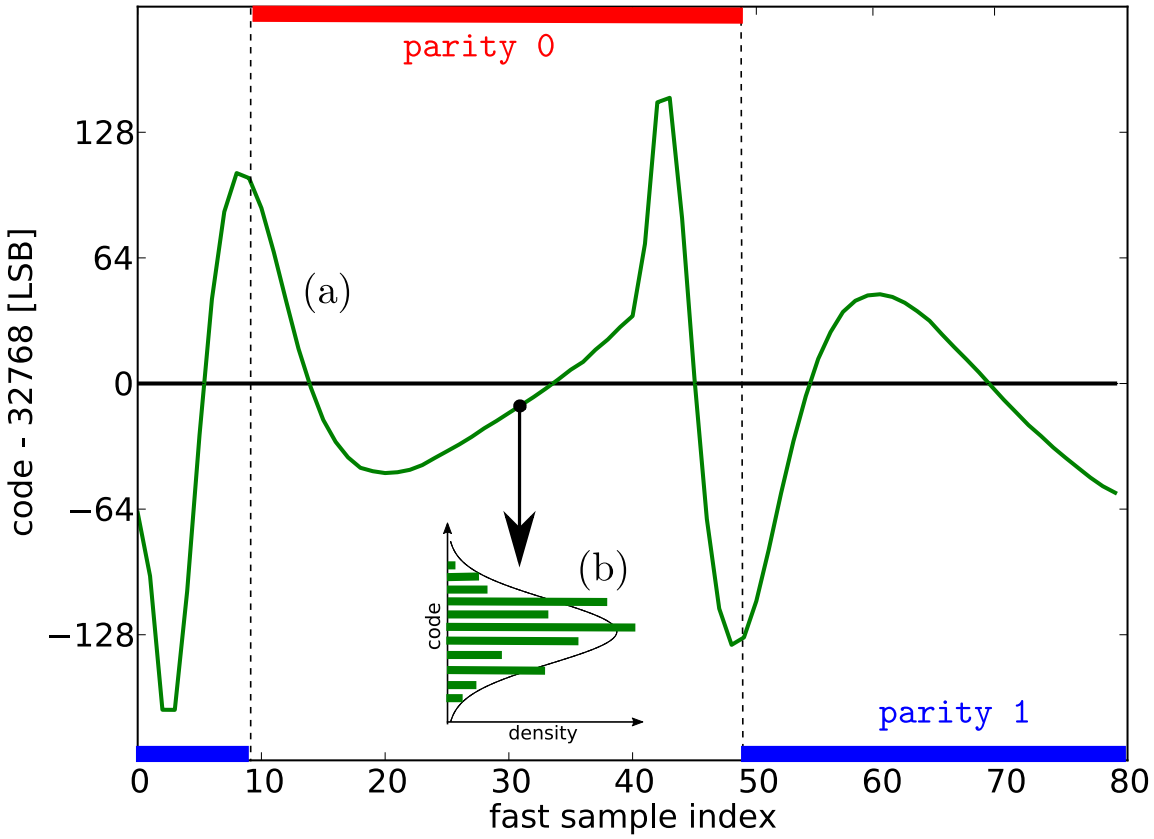


Figure 4.11: Example of *fast samples* periods selection for local INL reconstruction. The *fast samples* period (a) is the average of all selected *fast samples* periods yielding a given demodulated science data value s_d. The histogram inclusion (b) shows the non-ideal distribution of *fast samples* code values for the period index 30.

For each step the operations are the following:

1. PERIODS SELECTION: the set $\{\boldsymbol{p}\}|_{s_d}$ is built from *fast samples* periods yielding the current science data value s_d are selected. However there is only one causal half period on each *fast samples* periods (as plotted on Fig. 4.11), this is due to the timeshift ($\approx$ 10 *fast samples*) between the detector response and the modulation bias signal (which has been already detailed in Chap. 2). This is a problem because we can build a causal science data value only for the positive half period. A first order approximation is to consider the sky signal stable over the *fast samples* period $\boldsymbol{p}$. This approximation allows the calculation of one single science data value s_d from the 80 *fast samples* with

$$s_d(\boldsymbol{p}) = \frac{1}{2}\sum_{j=0}^{79}(-1)^{\mathrm{par}(j)} \times \boldsymbol{p}(j)\,, \tag{4.12}$$

   where $\mathrm{par}(j)$ is the parity value (zero or one) at *fast samples* index j;

2. LOCAL ESTIMATE OF $Q_j^{s_d}(\mathrm{K})$: The local quantization steps size $Q_j^{s_d}(\mathrm{K})$, for *fast samples* at index j, can be approximated from the *fast samples* histogram $h_j^{s_d}(\mathrm{K})$, with the

method detailed in Sec. 4.2.2. The discrete CDF $H^j_{s_d}$ is calculated from $h^j_{s_d}$ using Eq. 4.7, and the local quantization step size is obtained using Eq. 4.10 yielding

$$\frac{Q_j^{s_d}(\kappa)}{\sigma} = \Phi^{-1}(H_j^{s_d}(\kappa+1)) - \Phi^{-1}(H_j^{s_d}(\kappa))\,. \tag{4.13}$$

In practice a cut is applied on $H_j^{s_d}$ to keep only values in the (10%, 90%) percentile range , to avoid the samples distribution tails where statistics are not sufficient.

At the end of iterations, the global ADC quantization step size estimates, are calculated from the weighted local averages with

$$\frac{Q(\kappa)}{\sigma} = \frac{\sum_{j,s_d} \frac{Q_i^{s_d}(\kappa)}{\sigma} h_j^{sd}(\kappa)}{\sum_{j,s_d} h_j^{sd}(\kappa)}\,. \tag{4.14}$$

To break the degeneracy between σ and the quantization step size, a gain factor $1/\langle Q(\kappa)\rangle$ is applied to $Q(\kappa)$, so its average value is one LSB. The ADC transfer function is offset by convention so the midscale CTP is $v(32768) = 32768$.

4.3.3 Results of the INL estimation

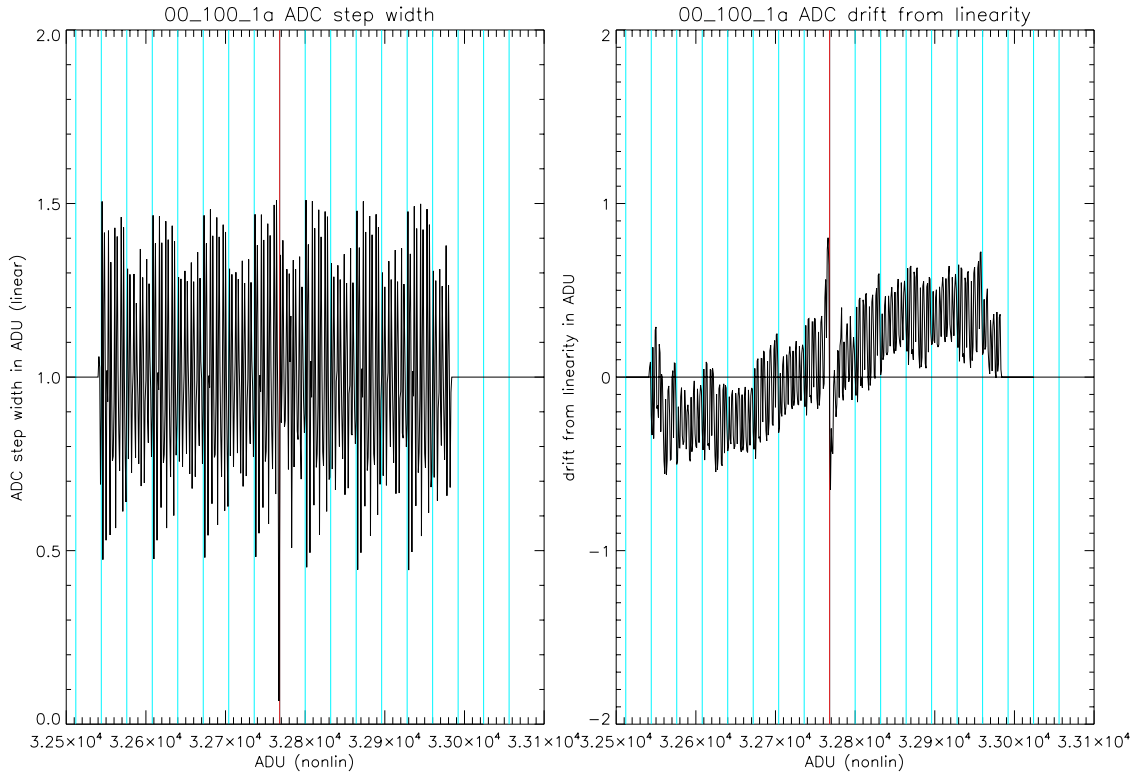


Figure 4.12: *Left*: example of DNL estimation built from cold *fast samples* for channel `00_100-1`. *Right*: corresponding INL estimation. The red line is located at midscale code 32768 and blue lines are displayed each 32 codes.

An example of the resulting DNL and INL estimated from cold *fast samples* is shown on Fig 4.12 for the first 100 GHz channel. On first sight, it confirms that the biggest defect is at midscale code 32768. An interesting pattern with a periodicity of 64 codes (**64-code**

periodicity) can also be observed on the DNL. This pattern will be discussed later, it reveals the binary nature of the underlying hardware and that the estimation is going into the right direction. The INL exhibits large scale drifts, incompatible at this scale with LCF nonlinearity (see Sec. 4.2.2), due to the roughness of the method.

The method uncertainty on each quantization step size $Q(\kappa)$ can be estimated by comparison of the values obtained using the data acquired during the first year, and with the data acquired during the second year respectively, this is shown on Fig 4.13. The yearly difference yields a value for the uncertainty on the value of $Q(\kappa)$ of about 0.01 LSB over the codes areas where statistic is high. However this uncertainty is not acceptable on the boundaries of the ADC scale covered by the *HFI* mission.

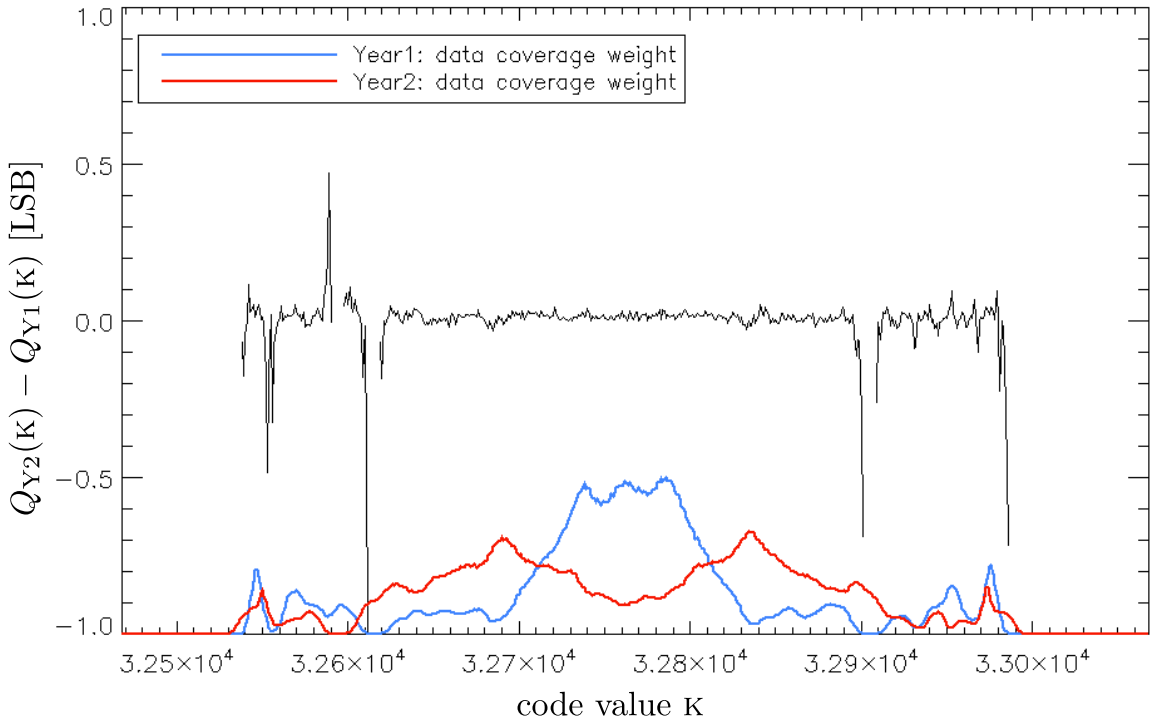


Figure 4.13: Uncertainty estimation of the quantization step in LSB units calculated from cold *fast samples*. The year one (blue line) and year two (red line) total weights are displayed with the same scale at the bottom of the figure to serve as a guide to the eye.

The comparison of the results for the 54 channels, shows that they all have different DNL characteristics albeit being in the qualification range $[-1, 1]$ LSB. As a consequence, the characterization and correction will have to be set up individually for each channels.

This first result is very encouraging, but a critical issue arises: the ADC code range at the boundaries of the mission cannot be characterized only from the in-flight cold data due to the lack of statistics. Another important issue is that the INL global drift is too large for correction. The maximum large scale INL drift should be $< 10^{-4}$, otherwise it will introduce a new bias. This large scale drift is due to:

- Inevitable leak of the unknown sky signal distribution into the Gaussian model. This could be worked out partially by using a 1-D spline model because the distribution is very smooth compared to the average ADC quantization step size. However this option has not been explored because of the critical ADC scale coverage issue and the preparation of on-ground acquisition campaigns;

- The sample selection which is biased by the ADC nonlinearity itself. This issue could be worked out with another iteration level, by applying the yielded global correction for $Q(\text{K})$ on the *fast samples* before looping over s_d calculated with this correction;
- The model for the *fast samples* signal is too simple and should include the detector response to the sky power. A first order model is detailed in Sec. 4.8.2 and further refined in Sec. 5.2;
- At this time it was not yet acknowledged that the spurious signal from 4 K lines was adding a significant bias at *fast samples* level.

Finally the preliminary in-flight estimates of the ADC chips nonlinearity suggested two ways of improvement:

1. A model of the ADC response is needed to provide better constraints for the INL. For this purpose on-ground characterization campaigns have been run to overcome issues encountered with in-flight INLs (see Sec. 4.4, Sec. 4.5 and Sec. 4.6).
2. New *fast samples* data is needed to have a proper coverage of the ADC scale, leading to the final characterization sequence described in Sec. 4.7.

4.3.4 Fit of generic SAR nonlinearity model

It is presented here the first attempt in modeling the ADC nonlinearity, using the DNLs built from cold *fast samples*. This was done using a the generic model for the DNL of an SAR ADC.

Because of the DAC architecture (see Sec. 4.2.1), the quantization steps size can only achieve n different values on an n-bit non-ideal SAR ADC (in theory).

The DNL model expression for code K can be written (**?**) as

$$\text{DNL}_\text{m}(\text{K}) = \sum_{j=1}^{n} a_j \left(\frac{r_j\left(\frac{\text{K}+1}{2^n}\right) - r_j\left(\frac{\text{K}}{2^n}\right)}{2} - \frac{1}{2} \right) , \tag{4.15}$$

where the a_j parameters are the characteristic DNL values mapping the DAC capacitances mismatch of a generic SAR ADC, and r_j are the Rademacher functions (square waves)

$$r_j(t) = \text{sign}(\sin(2^{j+1}\pi t)) , \quad t \in [0,1] ; \quad j \in \mathbb{N} . \tag{4.16}$$

The model DNL_m failed to fit the DNL over the full code range where the estimation is available, but it succeeded when considering only the 11 MSB. Even in this case the fit had to be restricted to a specific dataset defined as $\{\text{DNL}(k*32)+\text{DNL}(k*32-1), \quad k \in [\![0, 2^{11}[\![\}$ for the fit. The results are shown for two channels in Fig. 4.14. All the *HFI* bolometers channels have been fitted and the uncertainty on DNL(K) inferred with a jackknife algorithm (**?**) using 99% of the available *fast samples* data on 100 runs. The typical uncertainty inferred with the jackknife algorithm is about 0.01 LSB. This value is in very good agreement with the estimation from the yearly difference but it is too much compared to the target accuracy requirements, because the DNL error is cumulative on the INL. The main features of the best-fit modeling are:

- a quantization step size close to zero for the midscale code 32768;
- a central symmetry around each side of the midscale code 32768;

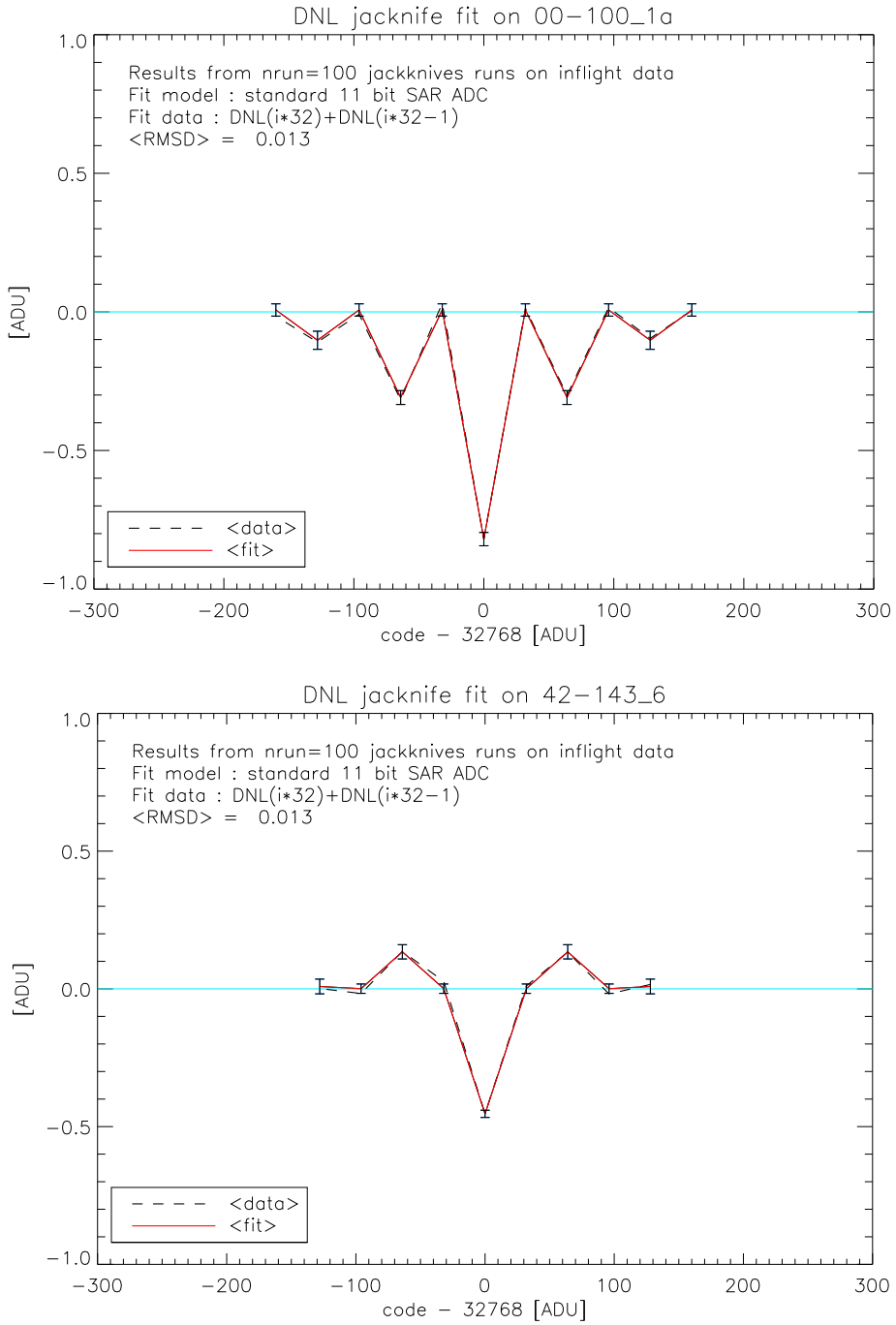


Figure 4.14: Generic SAR model fit results of the 11 first most significant bits for two channels.

- Each odd fit node apart from the central code has the same DNL value (codes 32768 ± 32, $32768 \pm 3 \times 32$, ...). The definitive explanation has been found later from the architecture of the DAC which uses a different block for the 6 LSB, thus yielding a 64-code periodic pattern (see Sec. 4.5)

4.4 INL estimations from *HFI* electronics flight spare

Thanks to the analysis performed on the cold *fast samples*, a lot of information was made available on the in-flight ADC chips, but not sufficiently to build a correction yet. Since

the chips present on the spacecraft cannot be brought back to Earth for a high precision characterization, it has been decided to start an analysis on the readout electronics of the *HFI* flight spare, assuming it embeds the same chips. To carry on the first on-ground measurements, a mission was organized to the Institut d'Astrophysique Spatiale (IAS) at Orsay where was stored the readout electronics of the the *HFI* flight spare. We were three people with Camille Parisel and Ludovic Montier to work together on the data acquisition. In this section is described the experimental setup used to produce *fast samples* for ADC nonlinearity characterization with Gaussian noise. The results are used for analysis of available constraints on the ADC nonlinearity for the `7809LPRP` chip and to validate the in-flight INL estimation method.

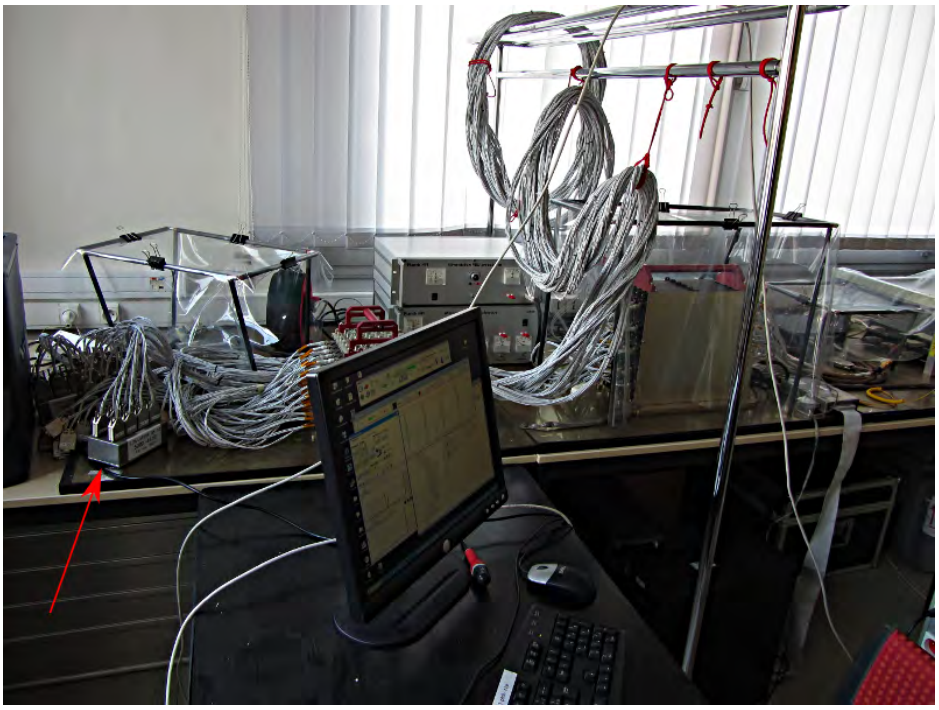


Figure 4.15: Photo of the *HFI* electronics flight spare taken at IAS in February 2012. The red arrow points to the box housing the load resistance used instead of bolometer, and the software used for data acquisition is visible on the foreground LED screen.

4.4.1 Experimental setup

The on-ground acquisition campaign uses a resistor as a stable Gaussian noise source (Johnson noise) for calibration. For this purpose two changes have been made to the electronic readout chain of the flight spare, compared to the in-flight setup:

1. the JFET box has been removed, it has a 1.0 gain and is only useful with real detectors;
2. the bolometer with its two bias capacitances have been replaced by a resistor with an impedance equivalent to the in-flight cold detectors at 100 mK and two equivalent capacitances.

Two different setups have been defined:

- Setup A (see Table 4.1) consistent with the end-of-life (EOL) in-flight sequence setup scheduled for the definitive ADC characterization (detailed in Sec. 4.7.1);

- Setup B (see Table 4.2) the purpose of which is to do an accurate characterization on several thousand codes around midscale code 32768.

The INL estimations made from the data acquired with Setup A and Setup B respectively will be compared in Sec. 4.4.5.

Parameter	Value	Comment
R	500 Ω	
`gamp`	3	gain=7.6
`ibias`	from 0 to 49	bias current
$C1$, $C2$	4.7 pF	bias capacitances

Table 4.1: Setup A : Mimics the EOL sequence.

Parameter	Value	Comment
R	22 MΩ	
`gamp`	3	gain=7.6
`ibias`	0	bias current
$C1$, $C2$	4.7 pF	bias capacitances

Table 4.2: Setup B : INL estimation with very large Gaussians.

4.4.2 Data acquisition

The *HFI* readout electronics are controlled with the Electrical Ground Support Equipment (EGSE) software developed at IRAP, and installed on a Win95 PC connected to the flight spare DPU. The EGSE has two functions: it sends programmable commands to the DPU and reads data as packets. These packets are in the same format as for the telemetry sent to Earth by the in-flight instrument. Hence we developed with Camille Parisel a piece of code in C language to extract the *fast samples* from the telemetry packets, while the data was acquired.

There is a complete belt (six channels) in the flight spare, from which the four remaining functional channels are acquired simultaneously at the *fast samples* data rate of 7200 Hz (no signal summation). This rate allows us to collect in one hour and half, as much *fast sample* as it would be possible in 10 months (duration of the EOL sequence) at the in-flight rate of one period each 101 seconds. The details of channel acquisitions and dataset nomenclature is provided in Appendix D.

4.4.3 Estimation of INL from Setup B

An example of a raw histogram of data is shown in Fig. 4.3.2 and has a very nice Gaussian shape, because the hits spread is dominated by the noise RMS (1500 LSB). The central defect and some other significant defects are visible at equidistant steps of 1024 codes apart from midscale code 32768, and the HCF nonlinearity appears as the hairy features.

The INL is estimated with the same method as described in Sec. 4.3.2. except that in the case of the IAS dataset, the "sky" signal is supposed to be constant during the whole acquisition. Hence there is only one "science data" signal bin. Two acquisition sessions were performed, labelled `B-60-I0` and `B-60-I1-V2`, see Appendix D. The quantization step sizes have been estimated independently for each acquisition and merged using Eq. 4.14 to produce the final estimate. These operations have been repeated for each of the four channels that were functional (labelled `FS60`, `FS62`, `FS64`, `FS65`).

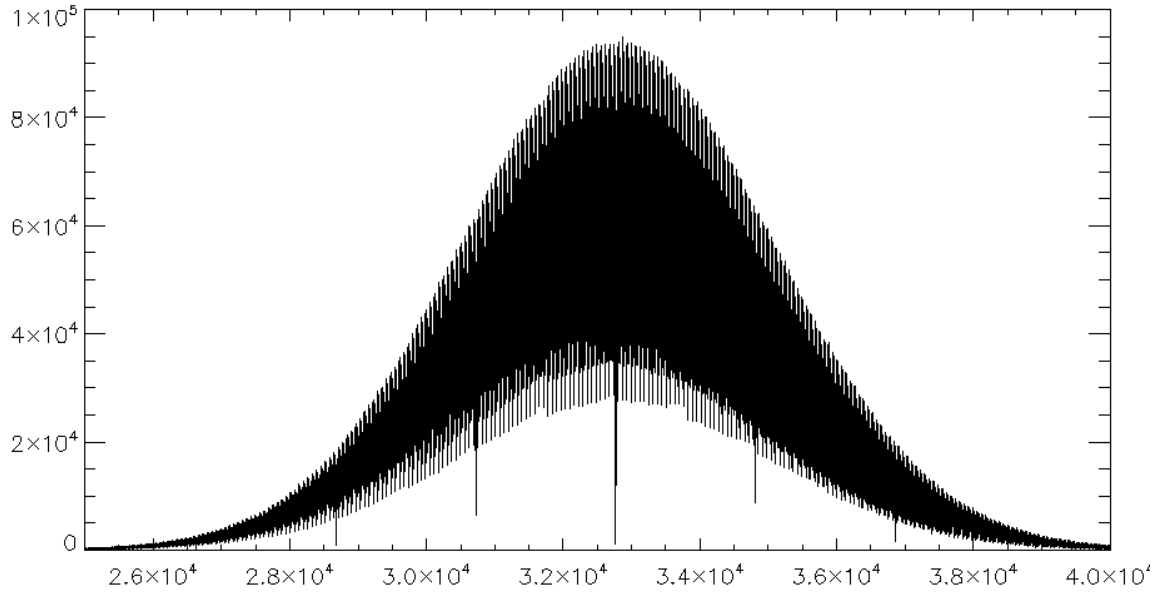


Figure 4.16: Histogram of *fast samples* for flight spare channel FS60 using dataset `B-60-I0` (see Appendix D).

The resulting INL is shown in Fig. 4.17 for channel FS60. Over the INL, there is a visible LCF drift, it can be removed efficiently (see Fig. 4.18) with a fourth order polynomial fit, considering it is low frequency compared to the INL features. This drift cannot be produced by the chip itself for the reason that its amplitude is too large and inconsistent with the chip specification, and because it is also too localized. The possible origins of the LCF drift are a variation in the noise RMS during the acquisition (that should appear as a Gaussian shape), or a drift of the mean signal value (that should appear as a Gaussian derivative shape). For more details on the impact on the INL see Sec. 4.7. However in this case, it is preferable to use a posterior correction than an estimation of the signal drift during the acquisition, because the value of the RMS is too high for an accurate determination of signal variations on the 80 *fast samples* locations. Signal drift and noise stationarity are typical difficulties encountered when using noise as a calibrator.

4.4.4 The 64-code periodic pattern

As observed on the INL estimated with the cold *fast samples* data, the 64-code periodicity is still present in the ground measurements and characteristic of the `7809LPRP` ADC chips. As this 64-code pattern cannot yet be modeled, it strengthens the basis for an empirical approach rather than a parametric one.

After removal of the LCF component (see Fig. 4.19), DNL and INL are very regular and lead to accurate analysis of the 64-code periodic pattern. Fig. 4.20 displays (on the left) the overplot of 65 DNL periods of 64-code (starting at code $k * 64$), showing on average a good agreement between the periods. The only exception to the 64-code periodicity is the first 64 codes after midscale.

The 64-code periodic pattern does not exhibit any symmetry that can be modeled with the standard SAR ADC model detailed in Sec. 4.3.4. However, the generic SAR model multi-periodicity is visible when looking at the DNL jumps located at the boundary between each 64-code pattern (see the right panel in Fig. 4.20). These jumps are referred to hereafter as the ”64-code jumps“, their value is calculated as the sum of the DLN values of nine codes centered on the jump location, minus the average value. This was a bit puzzling, because

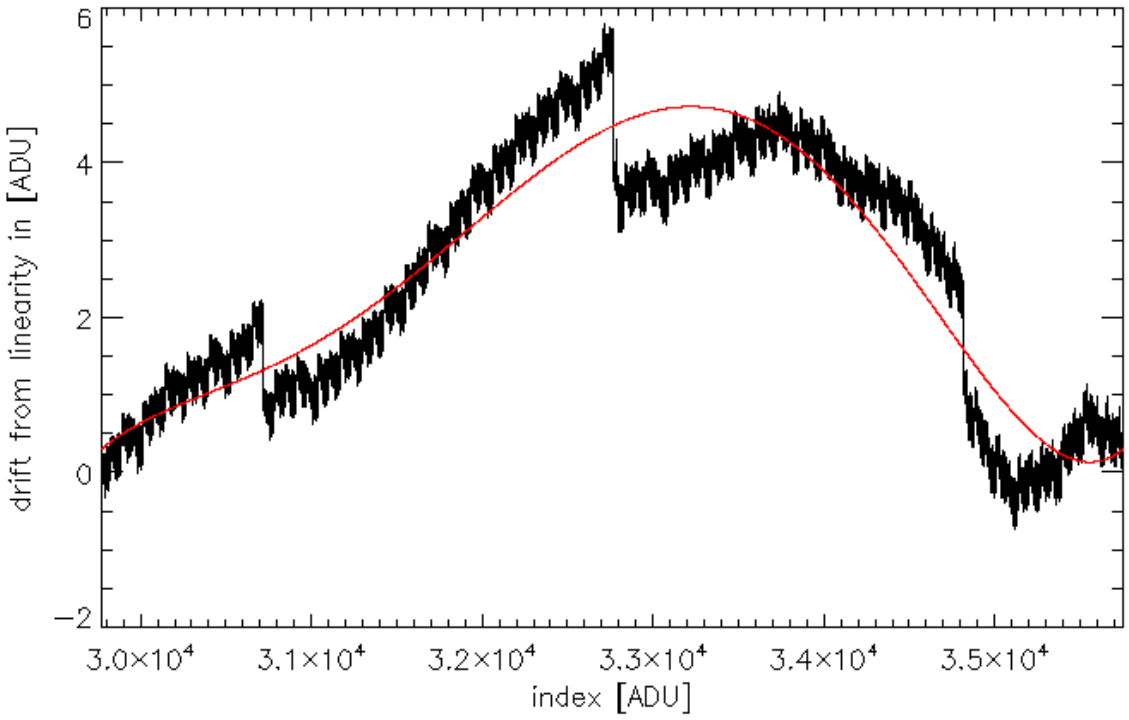


Figure 4.17: LCF drift on the INL for channel FS60. The red thick line is the fourth order polynomial fit.

the 64-code jumps does not affect only one code, but several ones around the theoretical location. The jump effect is the highest on the expected location then decrease with the distance to the jump location.

The on-ground acquisition strengthens the first in-flight hypothesis for the 64-code pattern, this pattern is visible on a very large code range around midscale (except for the first 64 codes after midscale). The generic SAR INL jumps also appears for the 10 MSB, but they are not localized as it would be in theory, the jump affects a few codes around the expected location.

4.4.5 Uncertainty estimation for the EOL sequence

The EOL (End Of Life) sequence is the final in-flight estimation of INL for each channel ADC, it is detailed in Sec. 4.7. In this section the accuracy of the scheduled in-flight measurement is estimated by comparison of the estimations yielded by the two acquisition setups A and B. The $\mathrm{DNL_B}$ obtained with the Setup B processing (B like Big Gaussian) is taken as a reference, it is because the large value of the noise is expected to introduce a smaller bias on the high frequency components of the DNL. Then $\mathrm{DNL_B}$ is compared with $\mathrm{DNL_A}$ yielded by Setup A which is consistent with the in-flight data acquisition of the EOL sequence.

The DNL relative error is, by construction, proportional to the histogram bin error which is $\sim 1/\sqrt{n_h}$, with n_h the number of hit counts for the considered ADC code. For Setup B, the average error for a DNL bin is $\varepsilon_B \approx 3.5 \times 10^{-3}$ LSB, and for Setup A respectively $\varepsilon_A \approx 3.3 \times 10^{-2}$ LSB, with the standard deviation of $\mathrm{DNL_B} - \mathrm{DNL_A}$ yielding 3.2×10^{-2} LSB. The correlation plot between $\mathrm{DNL_B}$ and $\mathrm{DNL_A}$ is shown in Fig. 4.21, it can be seen on this figure that both estimates are in very good agreement within 1σ of the noise and that there is no detectable bias correlated to the value of the DNL.

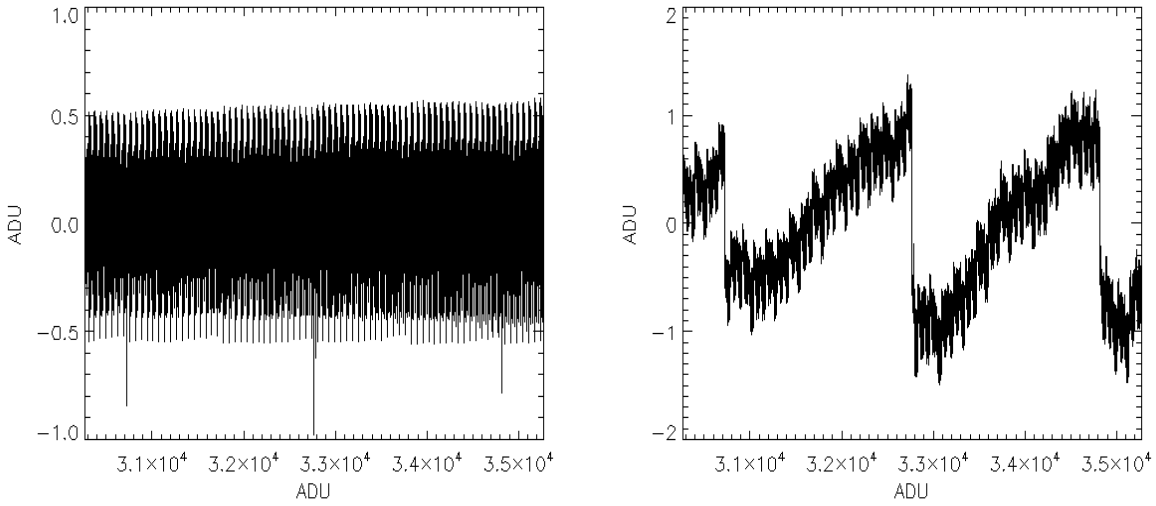


Figure 4.18: Resulting DNL (left) and INL (right) for FS60 after LCF component removal.

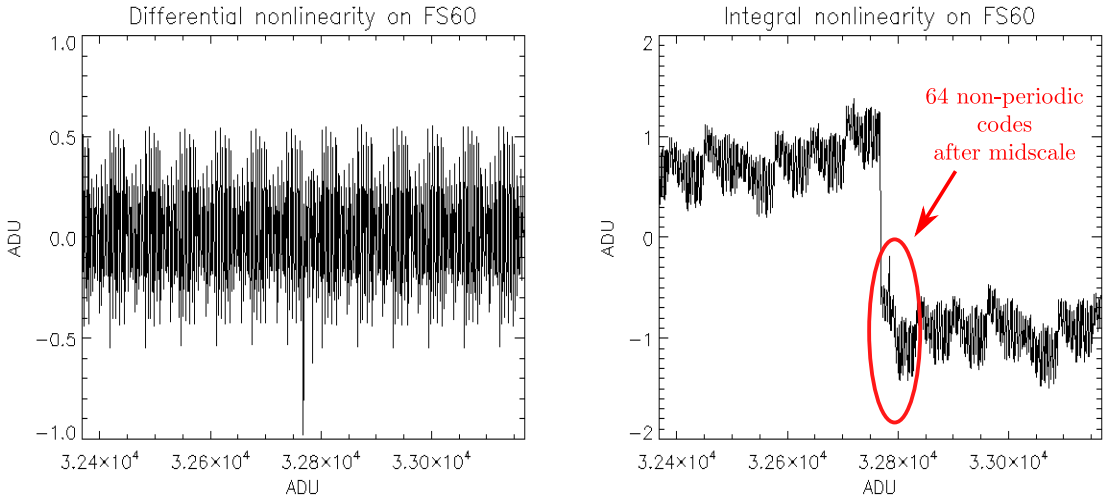


Figure 4.19: Same as Fig. 4.18, but the code range is zoomed on the typical ADC range used to digitize the cosmological signal all over the *HFI* mission.

Hence the result is in complete agreement with the expected noise level, this is an important validation of the INL estimation method for the future in-flight EOL sequence that will run with a limited amount of time before the spacecraft definitive end of operations.

4.5 SPICE model approach

My next option has been to make use of the electronic model of the 7809LPRP chip with the objective to work on a parametric model of its nonlinearity and more specifically the 64-code pattern.

4.5.1 Workplan

Ludovic Montier contacted Maxwell Technologies, and negotiated to obtain the electronic SPICE model of the 7809LPRP chip. In exchange we signed a non disclosure agreement with

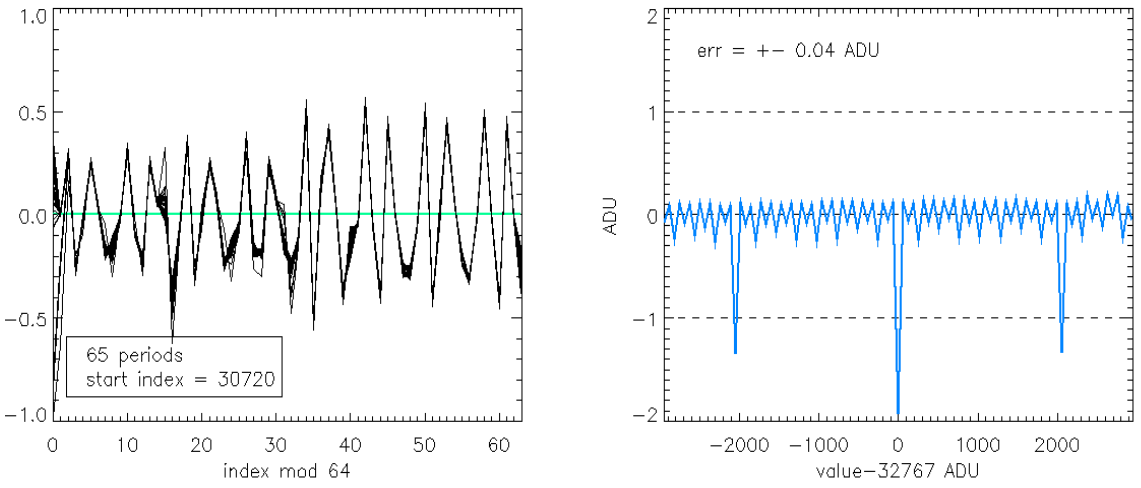


Figure 4.20: Visualization of the 64-code periodic pattern. *Left*: overplot of 65 DNL periods of 64-code. *Right*: DNL jumps each 64 codes on the $\pm 2\sigma$ code range of the ADC around midscale.

Maxwell technologies. This is the reason why this section is nearly empty, as it is not allowed to describe the electronic design of the chip.

The work plan was:

1. Analysis of the `SPICE` model to understand the underlying DAC architecture for the purpose of building an analytical model of the chip nonlinearity;

2. Use the `SPICE` model simulations with the `NGSpice`[5] software in a fit procedure, with DAC capacitance mismatch as free parameters to model in-flight chips nonlinearity. `NGSpice` has a huge advantage that it is a command line tool, so it is possible to run it in a naive parallel implementation to speed up considerably fit procedures.

However, after reception of the `SPICE` model, a missing file created a major issue for running the simulations. The original author who had retired Maxwel technologies a long time ago could not be contacted any more. The assistance of an electronics specialist was required, who was provided by the CNES in the person of Cecile Fiachetti[6]. With the help of Cecile we could work quickly on the `SPICE` model to extract the global DAC architecture. Unfortunately it was impossible to supply the subcomponent provided by the missing file. This subcomponent was used everywhere in the chip model, hence it was impossible to run the `NGSpice` simulation. It was a major drawback, because without electronic simulations, the understanding of the `7809LPRP` algorithm became extremely difficult and time consuming, so we agreed to stop at the DAC architecture level.

4.5.2 DAC architecture of the `7809LPRP`

Fig. 4.22 show the capacitive design of the `7809LPRP` chip. It confirms the 64-code periodicity of the DNL, because the DAC is segmented in two blocks. The lower bits coding for 64 values are handled by the area labeled `B'`, while the 10 MSB are organized as a complete and independent DAC. The presence of `CD10` and `CD10A` capacitances (surrounded by blue dashes) presenting the same value are the signature of the termination bit of an SAR DAC. As a consequence, it is clear that the DAC operates by determining the 6 LSB with a different

[5] `NGSpice` is a free software implementation of `SPICE` with more analysis features. `http://ngspice.sourceforge.net/`

[6] CNES Centre Spatial de Toulouse, 18 avenue Edouard Belin, 31 401 TOULOUSE CEDEX 9, FRANCE.

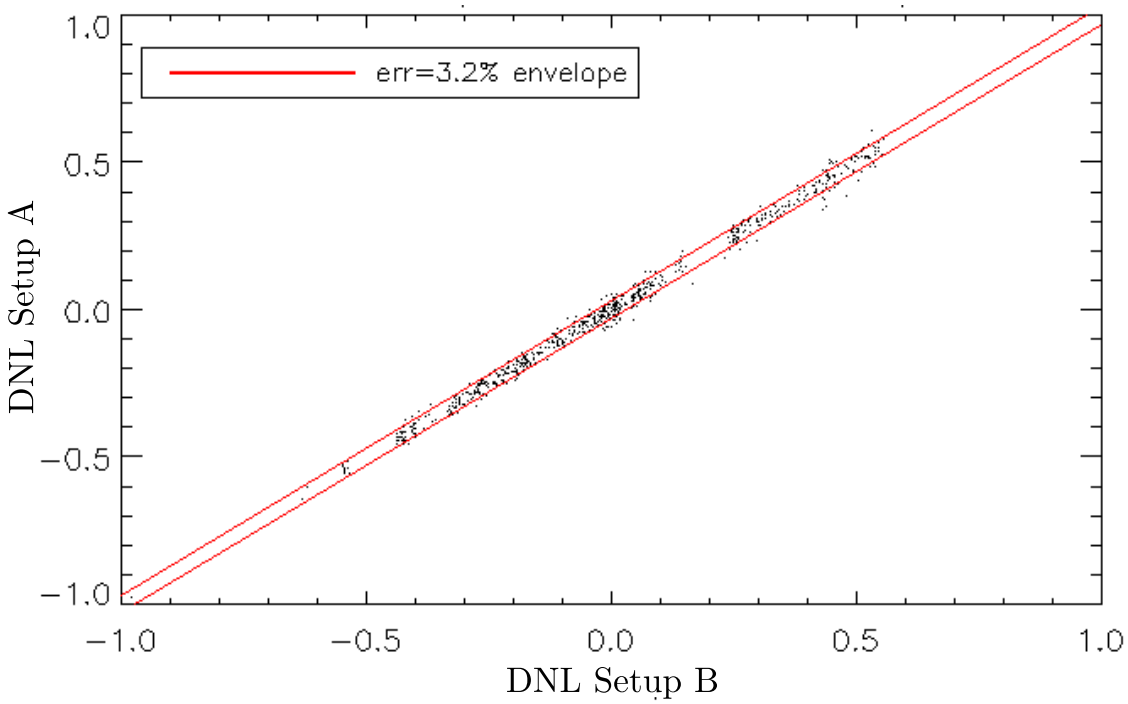


Figure 4.21: Correlation plot of DNL yielded by Setup A and B respectively. The two red lines bounds the 1σ envelope of the error, and are drawn with a slope of 1.0.

algorithm (maybe inside a pipeline). The unconnected capacitances of the diagram are linked to the SAR logic.

Rashmi Suddiwala tried to build a parametric model for this architecture, but could not obtain a fit for the available INLs.

As a conclusion, the DAC analysis shows that **the 64-code pattern of the six LSB has to be handled with a different model** than the ten MSB which obey the classical SAR nonlinearity model.

4.6 Cardiff full scale acquisition

After the IAS campaign in Orsay, two remaining spare `7809LPRP` chips were sent to Rashmi Suddiwala at Cardiff University for full scale characterization.

We interacted a few times to find an acquisition setup adequate to meet the required accuracy. The results presented in this section have been obtained by Rashmi Suddiwala with a setup using a 24-bit DAC to generate a triangle wave for *fast samples* acquisition by the `7809LPRP` ADC.

4.6.1 Acquisition setup

The input triangle wave at 0.0005 Hz was generated with a NI 24-bit DAC and a 5.17 V amplitude at ADC input. The ADC input range was set to ±5 V and its sampling rate to 26 kHz. For the DNL, the statistical error for the dataset is better than 0.01 LSB, and the great advantage of the calibration with a triangle wave is that it is also the absolute error for the INL. This is excellent for INL analysis at medium (64 to 2048 codes) scale.

Fig. 4.23 shows the resulting INL for one chip, there is a small low frequency drift between the two datasets of the figure, but the overall agreement between the two datasets is excellent. This low frequency drift is characteristic of the stability of the acquisition setup but is not

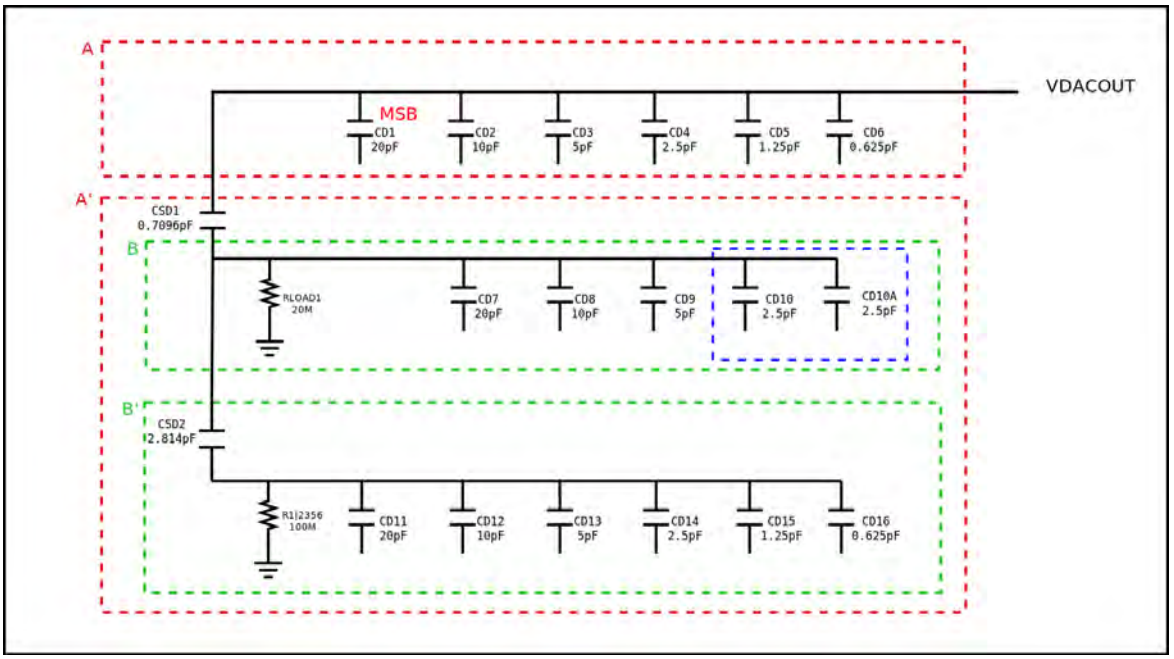


Figure 4.22: Diagram of the 7809LPRP DAC, extracted by hand from the `SPICE` model analysis. VDACOUT connects to the voltage generated by the DAC.

critical for the chip analysis in the case of *HFI* mission code range. The large scale behavior shows patterns at 2048 codes periods which are typical of the generic SAR ADC nonlinearity. There is also a drift (approximately cone shaped) in the high frequency pattern, but it is considered negligible on the typical code range used to digitize the cosmological signal during the *HFI* mission.

4.6.2 64 codes periodicity

François Couchot experimented with empirical approach by testing the 64 codes periodicity. When removing from the INL the average 64-code pattern and the jumps each 64 codes, the residuals are very flat, see Fig. 4.24. A few local spikes appear, but they are second order, compared to the jumps effect, because the spikes are smoothed by the electronic noise. Some residual jumps are visible, this is the same feature that were described when calculating them for Fig. 4.20, i.e. the jumps located each 64 codes do affect the few $\pm$ 4 codes around instead of just one. However most of the jump effect is still located at the expected location.

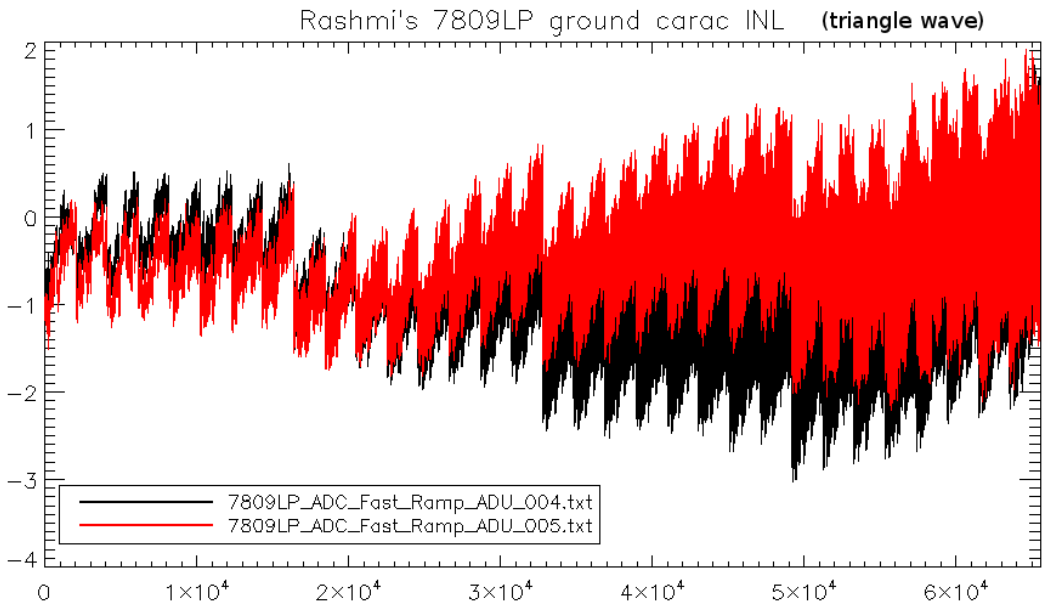


Figure 4.23: Fullscale INL estimated for one chip from Cardiff measurements. The black and red lines are from two independent data sets acquired with the same setup.

4.7 INL estimation from in-flight warm data

"Size does matter!"

After many tries to find a parametric model for `7809LPRP` nonlinearity, this option has been discarded in favor of an empirical description based on the three following features:

- The 64-code periodicity;
- INL jumps each 64 codes;
- The non periodic behavior of the 64 codes just after midscale.

François Couchot and Guillaume Patanchon worked a on a maximum likelihood method to produce the best possible characterization from the in-flight warm data. These results have been used to optimize the commands sent to the spacecraft to produce the measurements from the EOL sequence. This sequence was dedicated to the characterization of the ADC INLs for the 54 detector channels over the useful range of the ADC. I've been deeply involved in this process with the *fast samples* model, the detectors response in warm state and the analysis of potential residual pollution by the 4 K lines in warm state. A work was also performed on observables to infer the accuracy of the INL estimation candidates which is described in Sec. 4.8.

This section reports the main steps of the the EOL sequence, and the method used to produce the definitive fine ADC characterization.

4.7.1 The EOL sequence

EOL warmup

The discovery of the ADC nonlinearity as the origin of the time-variable gain happened only one month before the EOL warmup of *HFI* detectors. The EOL was a scheduled consequence

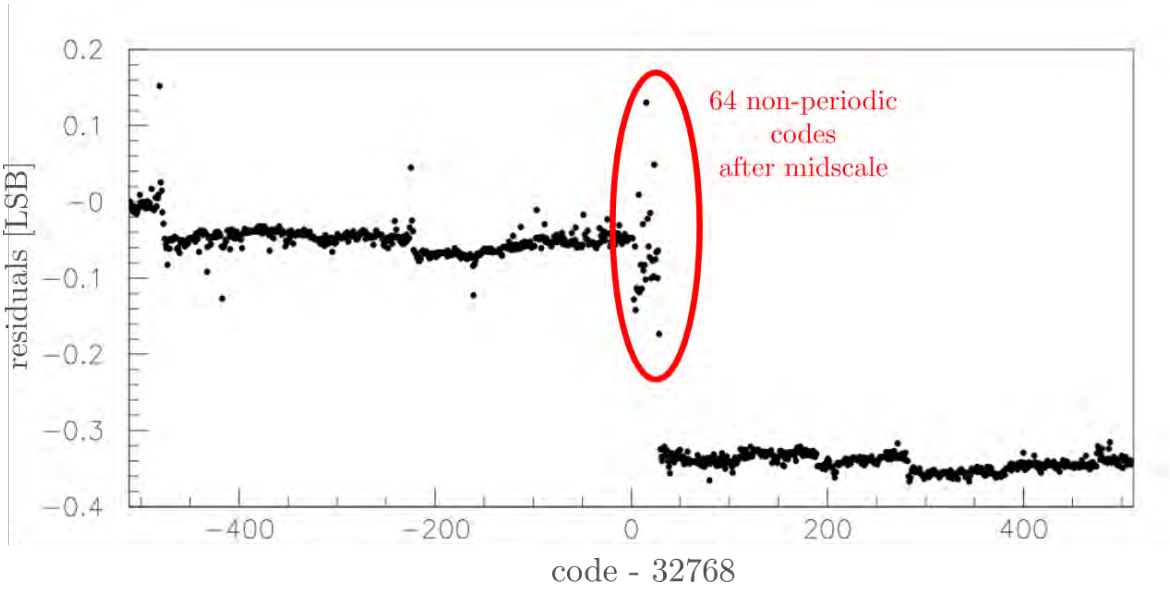


Figure 4.24: INL with the 64-code average pattern removed. The 64-code jumps are also partially removed. Figure reproduced with the permission of François Couchot.

of the finite amount of Helium (isotope 3 and 4) available for the dilution cooling system of the 0.1 K detector plate. During the warmup phase, the detector output signal is not stable. The temperature given by the two thermometers on the bolometer plate is shown in Fig. 4.25. Starting from 0.1 K on 01/10/2010, the temperature raises during two months before stabilizing above 1 K. During this period the raw unprocessed signal varies quickly and is not usable for the characterization of ADC nonlinearity.

After temperature stabilization around 1 K, the detectors are blind, in other words their gain response to sky power is negligible. This is in fact a very good thing for ADC characterization, because the only input signal of the ADCs becomes electronic white noise which is a perfectly known source of calibration (consistent with the Setup A described in Sec. 4.4.1).

HFI DPU reboot versus *LFI* requirements

The *HFI* readout electronics can be set in-flight on a *fast samples* data telemetry mode. In this mode the *fast samples* acquired at 7200 Hz rate for a single channel are sent to Earth at 40 times the nominal science data acquisition rate.

Unfortunately, after switching to the single channel *fast samples* telemetry mode, the DPU experienced reboots. Doing this repeatedly was not acceptable for the *LFI* instrument sharing the spacecraft which was still operating. As a consequence, it was decided to remain in the nominal telemetry mode to guarantee safe science data acquisition for *LFI*. Hence the *fast samples* were only available at the very low rate of 42 Hz on average, with non causal periods interleaved in time for each detector. This had a strong impact on the methodology, the drawback being that the hypothesis of Gaussian noise and others potential signal biases (4K harmonics) are very difficult to check. This was an all but optimal setup, nevertheless we could gather ten months of statistics at the slow rate.

EOL sequence

The EOL sequence consists of sending to the spacecraft a batch of `ibias` commands to change the detector bias current value, so that the analog signal can cover the ADC range

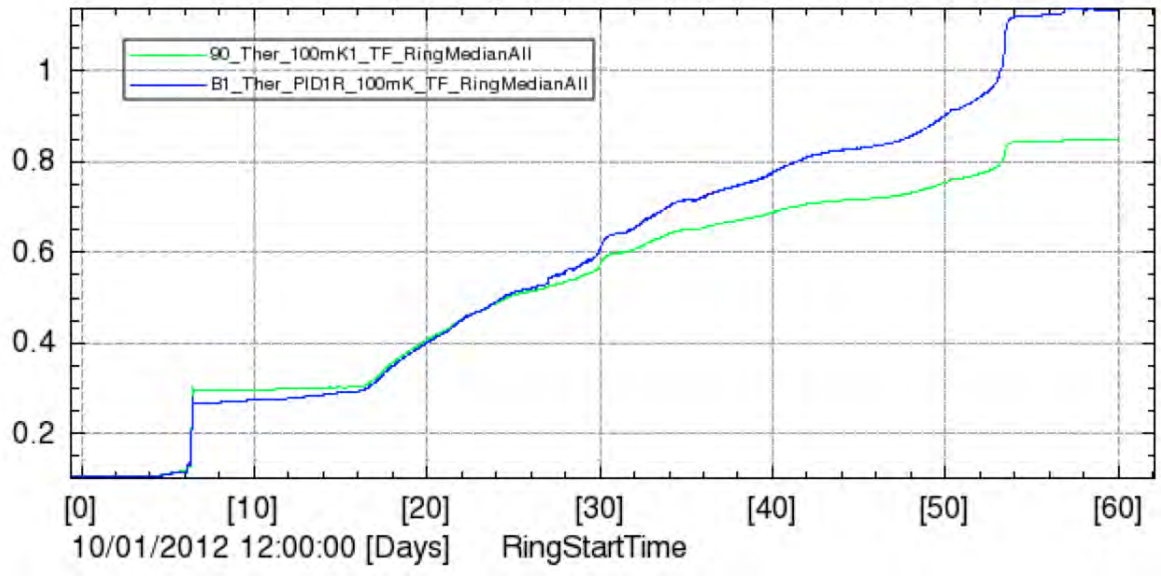


Figure 4.25: Temperature of the detectors plate during the first two month of EOL warmup, with x-axis is in days and temperature in Kelvin. The green and blue lines are the measurements from the two detectors present on the plate. Picture reproduced by permission of Laurent Vibert

spanned by the astrophysical signal during the HFI mission (±512 codes around midscale). During the EOL sequence, 20 `ibias` values are used, starting from zero, and the electronic readout gain is set to its maximum value so that the RMS value of the noise can reach the maximum possible value of 20 LSB. This sequence ran for ten months (see Fig. 4.26) and *fast samples* periods were captured at the same rate as during the HFI mission.

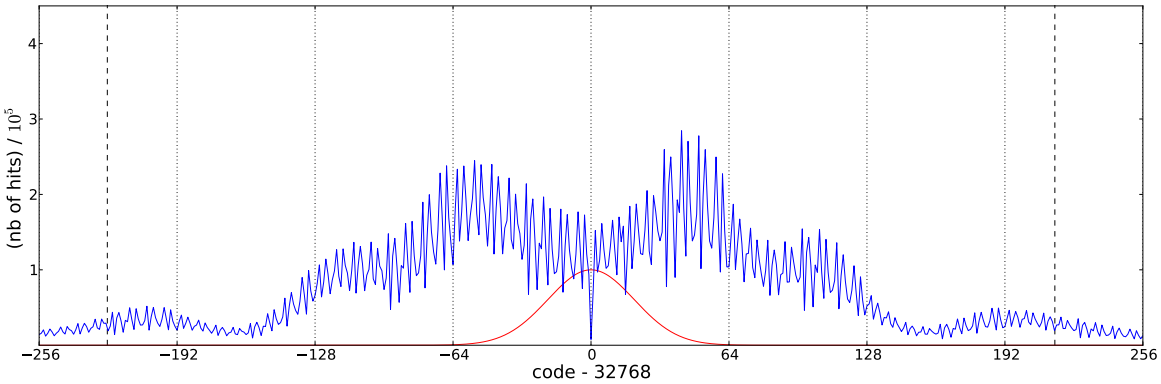


Figure 4.26: Histogram of the *fast samples* acquired over the 10 months of the EOL sequence duration. The two black vertical dashed lines are the 5% and 95% percentile boundaries. The red line shows the typical scale of the Gaussian noise distribution. The histogram scale is the same as for Fig.4.3.

Two strong hypotheses must be made for the *fast samples* acquired in the EOL sequence to be used as a source of calibration for the ADC nonlinearity:

- STABILITY OF THE DNL IN TIME: The data acquisition is performed after *HFI* EOL, but the REU hosting the chip is in the 300 K stage of the spacecraft. As the REU temperature is stable enough over the mission (measured variations are of order 1 K), we consider that the ADC nonlinear features are stable over the mission. Fig. 4.27 shows the variation of the REU temperature for the nine electronic channel belts over

the mission. The slow temperature modulation is due to the yearly motion of Earth around the sun.

- GAUSSIAN NOISE: The noise is expected to be Gaussian and stationary, in other words the 1σ value of the noise is expected to be constant at least in each stable `ibias` acquisition period. This includes the fact that additional bias from the 4 Kelvin mechanical compressor is insignificant, and that the electronic signal shape is constant to a very high precision. Furthermore the warm detectors are expected to have insignificant sky power response, since at a temperature greater than 1 K they have a very low impedance (about 1 kΩ compared to the average 12 MΩ at the nominal mission temperature of 100 mK).

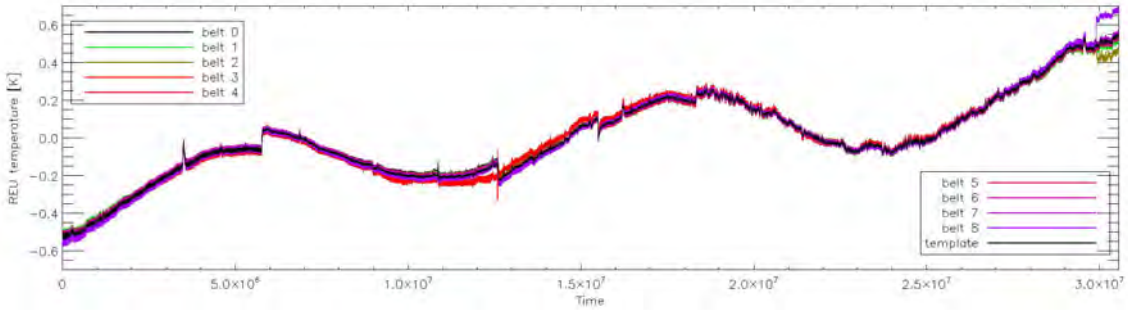


Figure 4.27: Variation of the REU temperature for the nine electronic channel belts over the two and a half year of the mission. The average value (about 300K) has been subtracted and the y-axis ranges from -0.6 K to 0.6 K. The visible modulation is caused by variation of the distance to the Sun during the yearly motion of the spacecraft. Picture reproduced by permission of Cécile Renault

ADC scale coverage with `ibias` parameter

Fig. 4.28 shows the mostly flat signal shape for a warm detector, and the square shape of the signal extension for an `ibias` increment. The spike at the beginning of the period is due to the non-zero `vtran` parameter (related to the square signal used to compensate the stray capacitance of the cables) which has not been changed since the cold mission. As the signal extension is $\approx$ 30 LSB, this property allows the desired ADC scale to be covered in `ibias` increments, under the condition that the 1σ value of the noise is of the same order as the step size.

The ADC range needed to cover the astrophysical signal dynamical range (mainly due to the solar dipole) is fully contained in the range $\pm$512 codes around midscale, so that 20 `ibias` increments are sufficient.

Gaussian noise and channel `gamp` parameter

The coverage of the ADC range must be homogeneous for the calibration with Gaussian noise. Hence it is important to tune the 1σ value of the noise to the same size as the `ibias` step increment (30 LSB) (see right plot in Fig. 4.28), so that the widths of the small Gaussians at each `ibias` step overlap with the previous one.

The `gamp` parameter of the REU allows us to change the amplification gain before digitization. For the EOL sequence it has been set to the maximum amplification value of 7.6 (see Tab.4.3). With this value, $\sigma \approx 20$ LSB, which is in good agreement with the requirements for an homogenous ADC scale coverage.

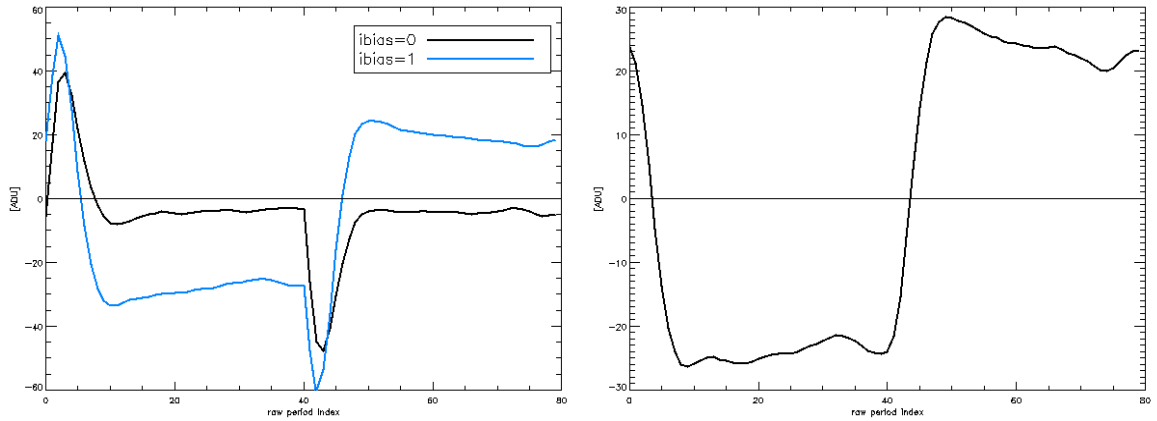


Figure 4.28: Signal shape for a warm detector. The x-axis ranges from 0 to 80 *fast samples* capture index. *Left*: one *fast samples* period for `ibias` = 0 (black) and `ibias` = 1 (blue) on channel `00_100-1a`. The y-axis ranges from −60 LSB to 60 LSB around ADC midscale value (32768). *Right*: signal for `ibias` = 0 minus signal for `ibias` = 1. The y-axis ranges from −30 LSB to 30 LSB around ADC midscale value.

`gamp` parameter	gain	comment
0	1/3	nominal value for submm channels
1	1	nominal value for CMB channems
2	3.0	
3	7.6	EOL sequence

Table 4.3: Possible values for the gain of the electronic readout chain.

4.7.2 INL estimation with maximum likelihood

In this section is summarized Guillaume Patanchon's work for his *Habilitation à Diriger des Recherches* (HDR).

Signal model for EOL sequence

The signal at the input of the ADC is supposed to be white and Gaussian, but only when considering each of the 80 samples index of the *fast samples* period separately. For the purpose of modeling this signal, let us state some formalism.

A data sample $d_i(t)$ is acquired at time t, where the index i maps to a given acquisition session in the EOL sequence corresponding to a fixed `ibias` value and *fast samples* index in the range $[\![0, 80[\![$ for a given detector channel. Hence for each acquisition session at a fixed `ibas` value will correspond 80 data sets. The samples acquired are the result of the digitization of the analog input signal by the ADC transfer function $\mathcal{A}$ such as $d_i(t) = \mathcal{A}(s_i(t))$, and are captured at times separated by 101 seconds by the DPU. The corresponding dataset is denoted $\{d_i(t)\}$.

Using the hypothesis that the input signal noise is Gaussian, $s_i(t)$ is described by the normal probability density function reading

$$p_i\left(s_i(t)|\{\mu_i, \sigma_i\}\right) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(s_i(t)-\mu_i)^2}{2\sigma_i^2}}, \tag{4.17}$$

where μ_i and σ_i are the average and standard deviation in volt of $s_i(t)$.

Using these notations, the probability of the measured dataset $\{d_i(t)\}$ is

$$q_i\left(\{d_i(t)\} = \mathrm{K} \mid \theta\right) \int_{v(\mathrm{K})}^{v(\mathrm{K}+1)} p_i(x|\{\mu_i, \sigma_i\})dx\,, \tag{4.18}$$

with $\theta = \{\mu_i, \sigma_i, v(\mathrm{K})\}$ being the unknown parameter set of size 65537.

Using the mean and standard deviation of the dataset as measured by the ADC, an initial (biased) estimate (μ_i^0, σ_i^0) can be calculated. As these parameters are degenerate with the code transition points $\{v(\mathrm{K})\}$, they must be estimated together. The standard deviation of the noise σ_i is degenerate with the local variation of the ADC transfer function gain, and the mean of the input signal μ_i is degenerated with a local distortion of the ADC transfer function similar to the gauss function derivative. Hopefully these degeneracies can be partially broken with

- The overlap of the $\{d_i(t)\}$ datasets;
- The 64-code periodicity constraint (except for the first 64 codes located just after midscale).

A global degeneracy between the values of σ_i is fixed by scaling the average value of $Q(\mathrm{K})$ to 1LSB.

Likelihood function

Let us consider the set $\{h_i(\mathrm{K})\}$, which is the histogram of dataset $\{d_i(t)\}$ containing N_i elements. We assume that data samples are independent from each others, so there is no loss of information when using the histogram instead of the dataset.

To build an estimator for the unknown parameters θ, we can derive a likelihood function from the histograms. For this purpose the histogram expected value can be put in relation with its previously defined probability as

$$\mathrm{E[h_i(K)|\theta]} = \mathrm{N_i q_i(K|\theta)}\,. \tag{4.19}$$

The likelihood function of a histogram is described by a multinomial distribution that we approximate with a Gaussian function, which is justified by the fact that the dataset used contains a large statistic. Given this approximation, the global log likelihood function for all the histograms can be written

$$\Phi(\theta) = \sum_i \sum_{\mathrm{K}} \frac{\left(h_i(\mathrm{K}) - \mathrm{E[h_i(K)|\theta]}\right)^2}{2\mathrm{E[h_i(K)|\theta]}} + \log \mathrm{E[h_i(K)|\theta]} \tag{4.20}$$

A second approximation will be needed for the expected value of the histograms hit count. To do this, we note that the value of σ_i is about 20 LSB. So the integral of the very smooth probability density function on the short range $[v(\mathrm{K}), v(\mathrm{K}+1)]$ (about 1 LSB) can be approximated with the trapezoidal rule. Then the first order expansion of Eq. 4.19 can be written

$$\mathrm{E}[h_i(\mathrm{K})|\theta] \simeq N_i Q(\mathrm{K}) p_i\left(\frac{v(\mathrm{K}) + v(\mathrm{K}+1)}{2}\middle|\{\mu_i, \sigma_i\}\right)\,, \tag{4.21}$$

with $Q(\mathrm{K}) = v(\mathrm{K}+1) - v(\mathrm{K})$ the quantization step size of code K.

Free parameters

There are two categories of parameters, the problem to solve itself described by the global code transition points $\{v(\mathrm{K})\}$, and the local small Gaussians described by $\{\mu_i, \sigma_i\}$ which are used for calibration. The INL can be characterized with either the quantization step sizes $Q(\mathrm{K}) = v(\mathrm{K}+1) - v(\mathrm{K})$, or the code transition points $v(\mathrm{K})$, both notations will be used when appropriate.

The initial problem brings a part of the solution: as the low signal dynamic of the astrophysical signal over the HFI mission is concentrated on ±512 codes around midscale, the full scale 2^{16} code transition points of the ADC does not need to be determined. This is also true for the bright signal of the Galaxy and planets where the ADC effect can be considered negligible, since it is averaged and canceled over the large dynamic range of these signals. Furthermore the hypothesis of the 64-code periodicity reduces the number of parameters, since each period can be described with 62 quantization step size, and 2 more parameters for the transitions located every 64 codes. However the 64 codes just after midscale do not follow the 64-code periodicity and have to be described separately with 62 parameters.

Concerning the sources of Gaussian noise, each `ibias` measurement, is described by 80 values of μ_i corresponding to the shape of the electronic signal period, all of them sharing the same value for σ_i. Hence acounting for the 20 different values of `ibias` used, 20 * (80) + 20 parameters are needed, leading to 1620 parameters.

Finally to describe $\{v(\mathrm{K})\}$ we need 62 parameters for the pattern periodicity, 62 more for the codes after midscale that do not follow the 64-code pattern, and 32 for the transitions at each period boundary in the range ±512 codes around midscale.

We end up with $1620 + 156 = 1776$ free parameters.

INL estimation method

The number of free parameters is too large to allow a global analytical solution. Hence the estimation is performed with a three steps iterative process, and the estimates calculated at iteration index j will be denoted by the exponent j.

At the beginning of the process, the ADC is supposedly linear with $v^0(\mathrm{K}) = \mathrm{K}$, and the initial parameters for the noise are initialized with their biased estimates, σ_i^0 and μ_i^0. At each iteration, the steps are the following:

1. ESTIMATION OF $\{\mu_i^j\}$: The position of the average electronic signal μ_i^j is estimated using the intermediate parameters $\{v_i^j(\mathrm{K})\}$ which are the local CTP derived from $\{d_i(t)\}$ independently. The local CTP are estimated by ensuring $\mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta_\mathrm{i}^\mathrm{j}] = \mathrm{h_i}(\mathrm{K})$ (using Eq. 4.19), and using the global solution $\{v^{j-1}(\mathrm{K})\}$ obtained from the previous iteration. To do so, for each dataset $\{d_i(t)\}$, the difference between the local estimate $\{v_i^j(\mathrm{K})\}$ and the previous global estimate $\{v^{j-1}(\mathrm{K})\}$, is minimized iteratively. The expression to be minimized is given by

$$\Phi_i = \sum_{\mathrm{K}} \left(Q_i^j(\mathrm{K}) - Q^{j-1}(\mathrm{K})\right)^2 p_i \left(\frac{v_i^j(\mathrm{K}) + v_i^j(\mathrm{K}+1)}{2} \middle| \{\mu_i^j, \sigma_i^{j-1}\} \right) , \tag{4.22}$$

   where Φ_i is a criterion derived from Eq. (4.17, 4.20, 4.21), $Q_i^j(\mathrm{K}) = v_i^j(\mathrm{K}+1) - v_i^j(\mathrm{K})$, $Q^{j-1}(\mathrm{K}) = v^{j-1}(\mathrm{K}+1) - v^{j-1}(\mathrm{K})$. In practice, 100 iterations of a custom Monte Carlo Markov Chain (MCMC) process are needed at this step.

2. ESTIMATION OF $\{v^j(\mathrm{K})\}$: The global estimate for code transition points is obtained by merging the local estimates $\{\{v_i^j(\mathrm{K})\}\}$ of step 1. In practice the retained effective

solution is a linear combination of the local estimates weighted by the local statistics, and the global quantization step size reads

$$\begin{cases} Q^j(\mathrm{K}) = \left(\sum_i \mathrm{W}_i(\mathrm{K})\right)^{-1} \left(\sum_i \mathrm{W}_i(\mathrm{K}) Q_i^j(\mathrm{K})\right) \\ \mathrm{W}_i(\mathrm{K}) = h_i(\mathrm{K})\,, \end{cases} \tag{4.23}$$

where the expression for the local solution weight $\mathrm{W}_i(\mathrm{K})$ can be derived using again Eq. 4.20, and the first order approximation is equivalent to use the histogram statistics. The 64 code periodicity is enforced at the end of this step by averaging the quantization steps $Q_i^j(\mathrm{K} \bmod 64)$ for the 62 codes between two transitions.

3. ESTIMATION OF $\{\sigma_i^j\}$: The value of the local standard deviations is adjusted by minimizing the discrepancy between all the local $\{\{v_i^j(\mathrm{K})\}\}$ found at step 2. As stated before, changing the value of σ_i is equivalent to a local gain change, which can be written using the quantization step size as $Q_i^j(\mathrm{K}) = \frac{\sigma_i^j}{\sigma_i^{j-1}} Q_i^{j-1}(\mathrm{K})$. Then the values of each σ_i^j independently is adjusted with a factor α_i^j to minimize the expression

$$f(\alpha_i^j) = \sum_{\mathrm{K}} w_i(\mathrm{K}) \left(Q_i^j(\mathrm{K}) - \alpha_i^j Q^j(\mathrm{K})\right)^2\,, \tag{4.24}$$

and using as weight $w_i(\mathrm{K}) = h_i(\mathrm{K})$, for $\frac{\partial f}{\partial \alpha_i^j} = 0$ it yields the solution

$$\begin{cases} \sigma_i^j &= \alpha_i^j \sigma_i^{j-1} \\ \alpha_i^j &= \dfrac{\sum_{\mathrm{K}} w_i(\mathrm{K}) Q_i^j(\mathrm{K}) Q^j(\mathrm{K})}{\sum_{\mathrm{K}} w_i(\mathrm{K}) Q(\mathrm{K})^2} \end{cases} \tag{4.25}$$

A total of 6 iterations with these three steps is sufficient for the convergence of parameters. The final output of the reconstruction method is presented in Fig 4.29. The high frequency defects (thick hairy features) tend to be averaged out by noise on the science data, while the jumps every 64 code do not, the most visible being the central one which is about 1 LSB in size. Also the 64 codes periodicity constraint helps to reduce the random walk visible as a black line on the left intermediate INL.

4.7.3 INL uncertainty

It is provided here an estimate of the uncertainty on the INL built with the likelihood method. To do so, the potential biases from the Gaussian parameters used in the likelihood are neglected, and the calculations are produced for the channel `00_100-1a`. There are three kind of code parameters to be considered:

1. the 64-code jump locations;
2. the 64 codes just after midscale;
3. the codes described by the 64-code periodic pattern.

For a specific quantization step Q_{K}, the error is described by the statistics of the EOL sequence histogram h_{EOL} filled with Gaussian noise. Hence the relative uncertainty for Q_{K} is $\mathrm{err}(Q_{\mathrm{K}}) = 1/\sqrt{n_{\mathrm{K}}}$, where n_{K} is the total number of hits for code K. The mean number of hits in the range 32768 ± 512 is $\approx 50 \times 10^3$ hits, thus the average uncertainty is $\approx 0.44 \times 10^{-2}$ LSB for the jumps every 64 codes.

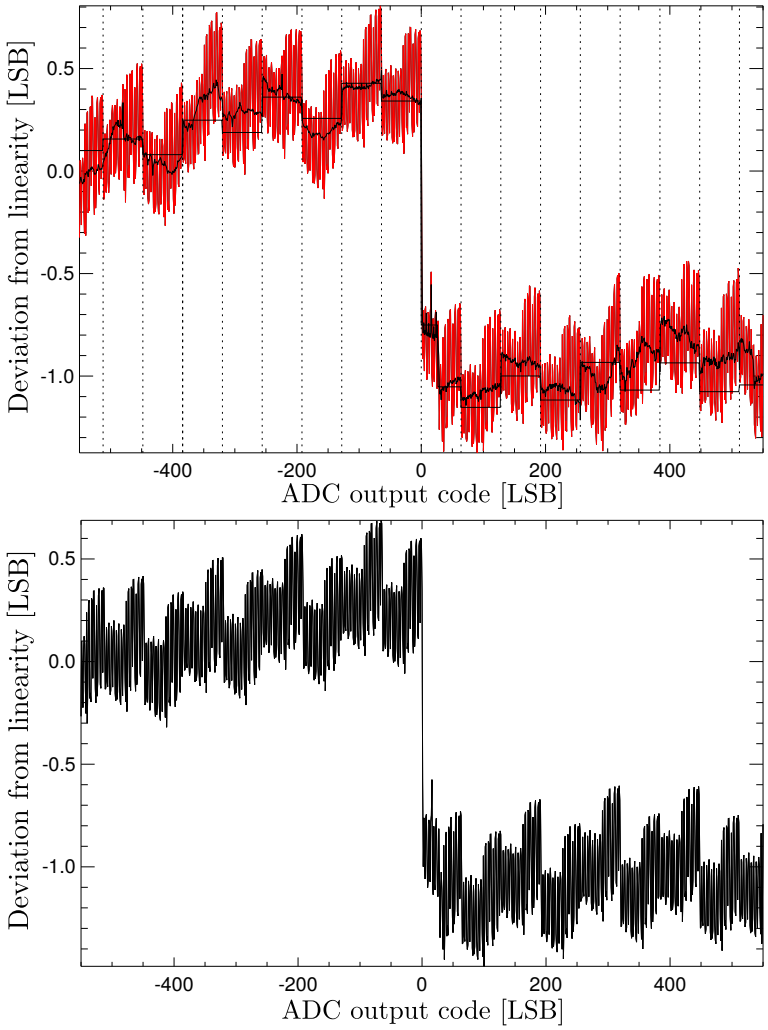


Figure 4.29: Estimated INL for the ADC of the channel `00_100-1a`, the 32768 value has been subtracted from the output code value. *Top*: the INL is taken from an intermediate step, when the 64-code periodicity is not enforced, the black line is the deviation from the averaged 64 periodic pattern. *Bottom*: the INL is the final result after convergence of the iterative procedure.

For the 64 codes just after midscale, the average bin filling of h_{EOL} is about $\approx 200 \times 10^3$ hits. Considering random walk statistics, the uncertainty is $\sqrt{64}/\sqrt{200 \times 10^3} = 1.7 \times 10^{-2}$ LSB over these 64 codes.

The 64-code periodicity allows the averaging of the statistics of all the remaining parts of the histogram. Excepted for the 64 codes just after midscale, the average number of hits for a code falling in the periodic pattern is 700×10^3. The periodicity of the pattern enforces that the maximum error is for a distance of 32 codes instead of 64, so the random walk error in this case is $\sqrt{32}/\sqrt{700 \times 10^3} = 0.68 \times 10^{-2}$ LSB.

Those uncertainties are consistent with the goal targeted in Sec. 4.2.3, except for the 64 codes just after midscale which are the weakest portion of the estimated INL due to lack of statistics.

4.8 Validation of a candidate INL

There are very few options to check the accuracy of the INL estimated with in-flight data, because there are no electronic calibration sources known with enough accuracy to do so.

The naive way is to estimate ADC residuals on the maps projection of science data after correction (as detailed in Sec. 5.5, but some quick estimators have also been developed which are presented below.

This section presents an histogram based method to estimate qualitatively the accuracy of a candidate INL. A second quantitative method is detailed, it is based on the calibration with the shape of the detector response (the *raw gain*), and result from the analysis of the *fast samples* acquired during the mission (cold *fast samples*).

4.8.1 The smooth histogram method

This method is based on the expected smoothness of the *fast samples* histogram over the full mission. If considering a perfect ADC, the bins are filled with Gaussian noise (RMS $\approx$ 4 LSB) and the average value vary slowly, dominated by the thermal fluctuations of the bolometer plate ($\approx$ 100 LSB for a given *fast samples* period index). Hence a correction of the histogram, using the estimated ADC nonlinearity, would yield a smooth histogram as it would be the case with a perfect ADC.

To check the smoothness hypothesis let us consider $h(\mathrm{K})$ the histogram of *fast samples* codes values over the *HFI* mission. If the averaged analog input signal $\langle s \rangle$ is constant over $h(\mathrm{K})$ and its adjacent bins, then the bin content is proportional to the quantization step size $h(\mathrm{K}) \sim \langle s \rangle Q(\mathrm{K})$. Hence the order zero ADC corrected histogram is

$$h_0(\mathrm{K}) = \frac{h(\mathrm{K})}{Q(\mathrm{K})}\,. \tag{4.26}$$

The resulting h_0 histogram is shown in Fig. 4.30 for the channel 00_100_1a, which is known to have a very low bias coming from other systematics such as 4 K lines. After correction, the result is impressive compared to the original version (see Fig. 4.26). The high frequency features have nearly disappeared. The few remaining spikes are due to local variations in the 64-code periodicity and are expected to be negligible after averaging by the noise of the analog signal (aka noise dithering). It can be seen in the high frequency residuals that the 64 codes after midscale look much better corrected (albeit with a noise level consistent with the statistics), this is expected because these codes have an "individual" estimate and do not follow the 64-code pattern. The remaining negative spike at midscale is a bit more problematic, it is not compensated by a positive spike on an adjacent code, hence this may be a residual bias from the likelihood method.

Finally the comparison between h_0 from the cold mission data, and from the EOL sequence respectively yields two important pieces of information:

- The ADC nonlinearity has not changed significantly over more than 3 years duration of the mission plus EOL sequence;
- The residual high frequency features are consistent between both datasets, this is an excellent crosscheck of the correction efficiency with uncorrelated data.

4.8.2 The *raw gain* method

This second method is based on the stability of the detector response also designed as the *raw gain* (this has been covered in details in Chap. 2). The hypothesis is that the *raw gain* can be used as a calibration source to estimate the efficiency of ADC nonlinearity corrections. As a consequence the consistency of the *raw gain* calculated at different moments in the mission lifetime should be a tracer for ADC nonlinearity residuals.

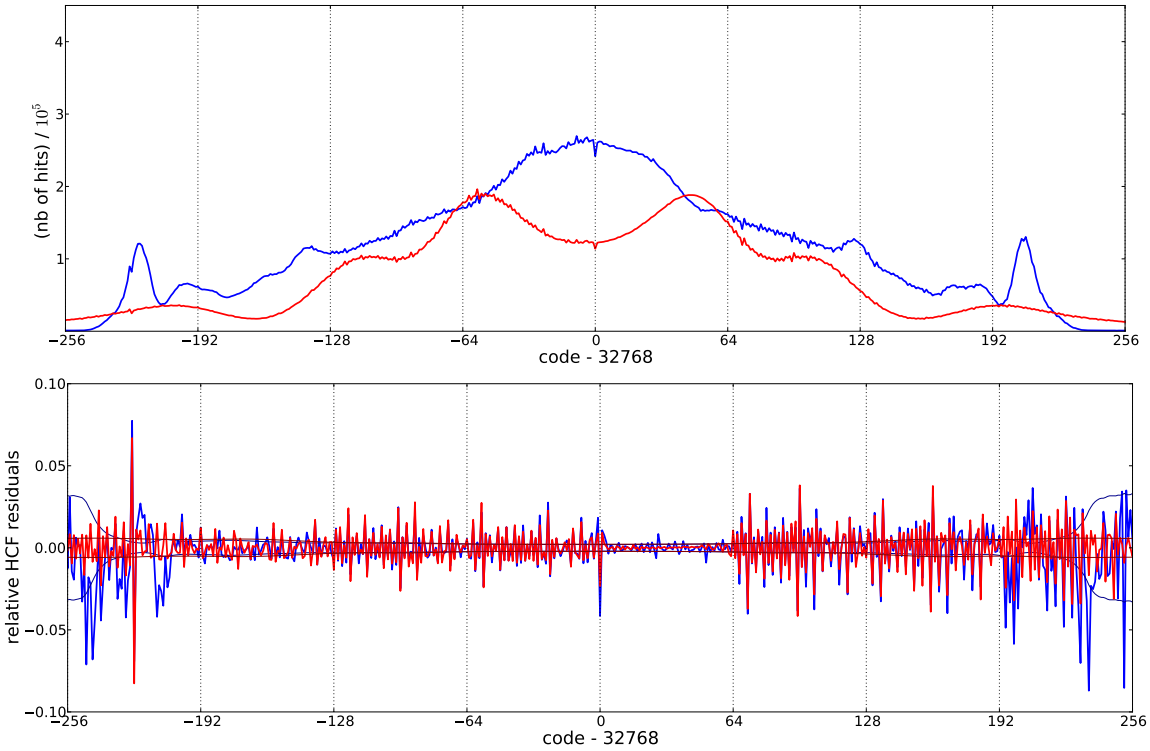


Figure 4.30: *Top*: Histogram h_0 of *fast samples* from channel `00_100-1a` corrected for ADC nonlinearity. The blue line (more statistics) is for cold data of *HFI* mission (see uncorrected version in Fig. 4.3). The red line is for EOL sequence data (see uncorrected version in Fig. 4.26).
Bottom: high frequency relative residuals calculated by removing a smoothed baseline using a gaussian kernel with a variance of one LSB. The thin darker lines are the 1σ uncertainty from the histograms filling.

ADC nonlinearity correction of *fast samples*

Before calculating the observable, the ADC nonlinearity correction must be applied on the *fast samples*.

In a typical SAR ADC, the transfer function $\mathcal{A}$ behaves like a floor function, so the average value of the input signal yielding the code κ can be considered to be at the center of the quantization step Q_κ. Hence the ADC correction function reads

$$\mathcal{A}^{-1}(\kappa) = \frac{v(\kappa) + v(\kappa+1)}{2}\,. \tag{4.27}$$

The result returned by $\mathcal{A}^{-1}$ is expressed in LSB because in the *HFI* case, the ADC corrections are calibrated to have an average quantization step size of 1.0 LSB.

Cold *fast samples* model of signal response

Now that the ADC nonlinearity correction can be applied on the *fast samples*, let us see how the sky signal variations map to the *fast samples*.

The detector response is supposed to be linear all over the mission, it can be written as a first order approximation $\Delta s_d / \Delta P_{\text{sky}} = g$, where s_d (summed *fast samples* in DSN units) is the demodulated sky signal, P_{sky} is the input sky power and g is the scalar detector gain. At the time of this analysis, the detector response model was not yet as polished as presented in Chap. 2, but analysis of *fast samples* shows already that, at modulation period level, the static gain translates to a vector $\mathbf{g}$ defining the modulation period extension as a

response to a variation of $P_{\rm sky}$. The *fast samples* version of the gain is called the *raw gain* and its expression is $\mathbf{g} = \Delta\boldsymbol{p}/\Delta P_{\rm sky}$, where $\boldsymbol{p}$ is the *fast samples* period corrected for ADC nonlinearity with $\mathcal{A}^{-1}$. It is simpler to use LSB/DSN units by calculating the gain with

$$\mathbf{g} \sim \Delta\boldsymbol{p}/\Delta s_d\,. \tag{4.28}$$

Calculating the *raw gain*

The *raw gain* has been defined in Sec. 4.8.2 as a 80 samples vector $\mathbf{g} \sim \Delta\boldsymbol{p}/\Delta s_d$, where $\Delta\boldsymbol{p}$ is the variation of the *fast samples* period shape corresponding to the variation of the demodulated science data sample value $\Delta s_d \sim \Delta P_{\rm sky}$. Using only data from *fast samples* periods, the science data value cannot be calculated exactly for both parities, because only one causal half period is visible. However if the input sky power is considered stable over the duration of the period, an estimate for the value of the demodulated science data sample is given by

$$s_d(\boldsymbol{p}) = \frac{1}{2}\sum_{j=0}^{79} \mathrm{sign}(\mathbf{g}(j)) \times \boldsymbol{p}(j)\,. \tag{4.29}$$

This equation is consistent with the previous version given in Eq. 4.12. In practice the location where the sign of $\mathbf{g}$ changes has been calculated during the corresponding CPV sequence as the `sphase` parameter. This parameter is obtained by adjusting the index in $\boldsymbol{p}$, where the science data summation starts, in order to maximizing the gain of the full readout chain. For a nominal science data sample with parity zero the calculation done by the DPU is $s_0 = \sum_{j=0}^{40} \boldsymbol{p}[j + \texttt{sphase}]$.

The final expression for the raw gain is

$$\mathbf{g}(j) = \frac{\Delta\boldsymbol{p}(j)}{s_d(\Delta\boldsymbol{p})}\,, \quad j \in [\![0, 79]\!]\,. \tag{4.30}$$

Overcoming the 4 K lines bias

An efficient way to calculate the raw gain is to average the *fast samples* periods falling within a given range of demodulated science data value $[S_k, S_k + \Delta s_d]$, where k is the index for the partitioning of the science data values taken over the mission in increments of Δs_d (with a [5%, 90%] percentile cut). In practice, an averaged period $\langle\boldsymbol{p}\rangle_k$ for a science data range index k is calculated with

$$\langle\boldsymbol{p}\rangle_k(j) = \langle\boldsymbol{p}_i(j)\rangle\,, \quad s_d(\boldsymbol{p}_i) \in [S_k, S_k + \Delta s_d]\,, \tag{4.31}$$

where i is the time ordered *fast samples* period index. With this binning the raw gain is given by

$$\mathbf{g}_k(j) = \frac{\langle\boldsymbol{p}\rangle_{k+q}(j) - \langle\boldsymbol{p}\rangle_k(j)}{s_d\big(\langle\boldsymbol{p}\rangle_{k+q} - \langle\boldsymbol{p}\rangle_k\big)}\,, \tag{4.32}$$

where q allow to select empirically the SNR by increments of Δs_d.

However the variations over the mission of the 4 K lines spurious signal (up to 5 LSB) bias significantly the estimate of the *raw gain*. An example is shown in Fig. 4.31 for a channel moderately affected by 4 K lines. The calculated $\mathbf{g}_k$ are completely dominated by the 4 K lines variations which are strongly correlated with the bolometer plate thermal variations.

One solution to avoid the 4 K lines bias is to use the CMB dipole modulation (solar plus orbital) which is not correlated to the 4 K lines signal. This can be done by using the difference of two consecutive periods $\delta\boldsymbol{p}_i = \boldsymbol{p}_{i+1} - \boldsymbol{p}_i$. The $\delta\boldsymbol{p}_i$ differences are dominated by the CMB dipole modulation amplitude, and filters out other analog signal variations

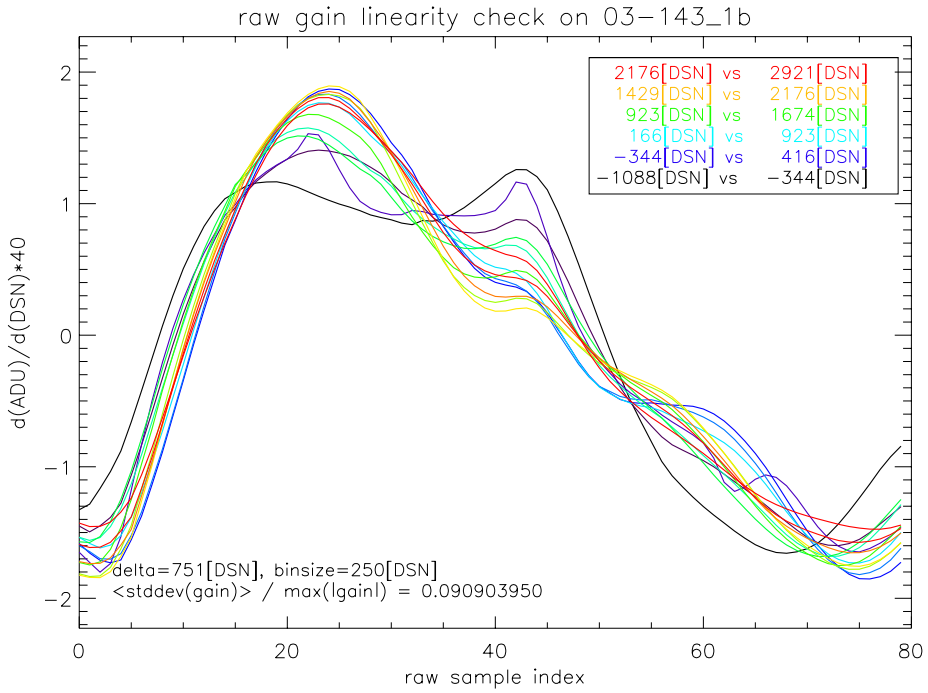


Figure 4.31: Raw gain calculated with *fast samples* periods from different science data bins for channel `03_143-1a` (biased by 4 K lines signal). The *fast samples* periods are averaged over science data bins of size $\Delta s_d = 250$, and $\Delta \boldsymbol{p}/\Delta s_d$ is calculated for consecutive averaged periods with $\Delta s_d > 500$.

including 4 K lines at larger timescales. The scheduling of *fast samples* periods capture, each 101 seconds, enforces a constant phase difference of about $\pi/3$ in the phase of the CMB dipole sinusoid between two captures (a complete rotation of the spacecraft over the sky takes about 60 seconds). This phase difference translates on average to a 7 LSB amplitude on $\delta \boldsymbol{p}_i$, which is enough for precise estimations given the total number of available periods ($\approx 750 \times 10^3$).

Now, the *raw gain* is calculated with time ordered period differences $\delta \boldsymbol{p}_i$. The slow thermal variations in temperature of the bolometers thermal plate (see Fig. 4.10) ensures that each bin in time probes a different region of the ADC. The binned *raw gain* for bin index m reads

$$\langle \mathbf{g} \rangle_m(j) = \left\langle \frac{\delta \boldsymbol{p}_i}{s_d(\delta \boldsymbol{p}_i)} \right\rangle , \quad \left\{ \begin{array}{l} i \in [\![m\frac{n_p}{n_{\rm bin}}, (m+1)\frac{n_p}{n_{\rm bin}} [\![\\ m \in [\![0, n_{\rm bin} [\![\end{array} \right. , \tag{4.33}$$

where n_{bin} is the number of bins in time, and n_p is the total number of *fast samples* periods used.

Stability of the *raw gain* over the mission

The Fig. 4.32 shows an impressive consistency for the *raw gain* calculated with Eq 4.33 and using five time ordered bins. The ADC nonlinearity residuals are visible as deviations from the average near the peak. This result is excellent if we consider that the signal variations contained in $\delta \boldsymbol{p}_i$ are very small compared to the global thermal fluctuation of the bolometer plate. The relative error can be estimated with the normed RMS of the standard deviation

of *raw gains* at each *fast samples* index

$$\mathrm{err}(\langle \mathbf{g} \rangle_k) = \frac{\mathrm{RMS}\Big(\mathrm{stddev}(\langle \mathbf{g} \rangle_k(j))_{k \in [\![1, n_{bin}]\!]}\Big)_{j \in [\![0,79]\!]}}{\max(|\mathbf{g}|)} \,. \tag{4.34}$$

The application of the nonlinearity correction (version 2015 `GP41`) with $\mathcal{A}^{-1}$, reduces the estimated error by a factor $\approx$ 5. This demonstrate that the nonlinearity correction increases the consistency of the *raw gain* calculated at different times of the *HFI* mission, and simultaneously establishes an upper limit on potential variations in the detector response over the mission.

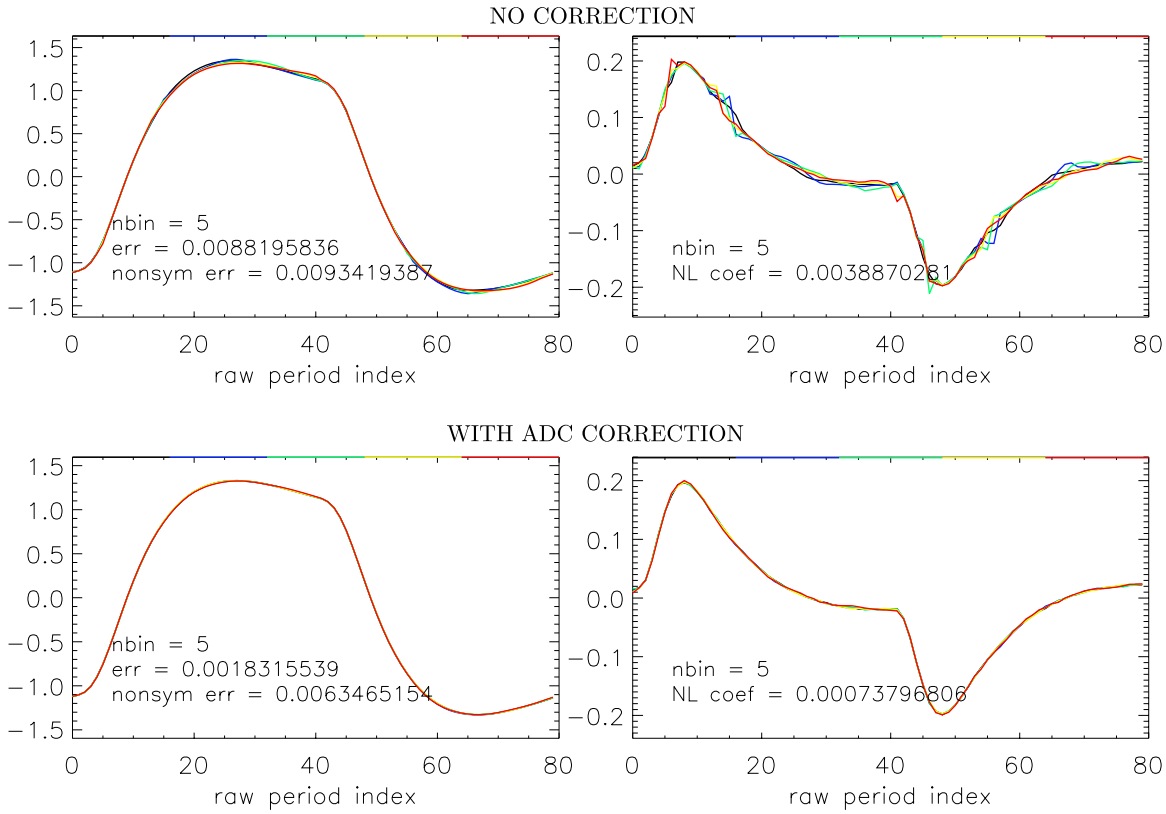


Figure 4.32: Effect of ADC nonlinearity correction on *raw gain* stability over the *HFI* mission for channel `03_143-1a` (unbiased). *Left column*: binned *raw gains* $\langle \mathbf{g} \rangle_k$ with $n_{bin} = 5$ and the color code for each bin is over-plotted on top of each plot, ranging from black (beginning of *HFI* mission) to red (end of *HFI* mission). *Right column*: time domain differences $\langle \mathbf{g}' \rangle_k$. *Top row*: uncorrected *fast samples*. *Bottom row*: *fast samples* corrected using $\mathcal{A}^{-1}$ parametrized with the INL labeled `GP41` which is the best estimate of year 2015.

The relative error (Eq. 4.34) can also be estimated with the time domain differences of consecutive *fast samples* of the *raw gain* which is homogeneous with a time domain derivative of $\langle \mathbf{g} \rangle_k$. The expression for this new estimation is

$$\langle \mathbf{g} \rangle'_k(j) = \langle \mathbf{g}_i \rangle_k((j+1) \bmod 80) - \langle \mathbf{g}_i \rangle_k(j) \,. \tag{4.35}$$

Despite an increase of the noise by a factor $\sqrt{2}$, this processing presents the advantage of filtering out low frequency biases due to residual 4 K lines and glitch tails. Hence this second indicator is very useful for qualitative analysis.

The relative errors based on the *raw gains* for all channels are compiled in Fig. 4.33. The correction of the central defect alone considerably improves the consistency of the *raw*

gain estimates as expected, but the more accurate estimation of the EOL sequence yields a significantly better improvement. However this is only an upper limit because the accuracy of the *raw gain* based estimator is limited by residual biases. The main bias comes from the 4 K lines, which are known to have periodic amplitude variations following the periodicity of the 20 K stage operations on a 15 minutes scale (this is detailed in Sec. 6.1.3).

These results are strong indicators that:

- the INL estimated from the EOL sequence brings a significant improvement to the data by nearly one order;
- the *raw gain* shape is very stable over the mission at the *fast samples* period level. An upper limit for its variations, after ADC nonlinearity correction, can be set at 10^{-3} level.

The main drawback of this method is that the estimation of *raw gain* shape via *fast samples* alone, cannot track scalar gain variations of $g = \Delta s_d / \Delta P_{\rm sky}$ over the mission.

4.9 Perspectives

4.9.1 Avoiding ADC nonlinearity

Some future missions, such as the QUBIC (**?**) experiment will make use of sample summation, or similar processing to increase the SNR. It is discussed in this section some possible design to prevent ADC nonlinearity issues and avoid loss of information necessary for nonlinearity correction.

Increasing signal dynamic

A trivial solution is to increase the signal dynamic, compared to the ADC defects in LSB units. The SNR for the quantization noise is generally calculated in decibels with the formula $SNR = 6.02n + 1.76\text{dB}$, where n is the number of bits of the ADC. But this theoretical number holds only in the case of an ideal component with the quantization noise averaging out, and considering the full input scale of the component. For a non-ideal component, a useful figure to characterize bias level is the maximum value of the INL, or in other words the maximum transfer function drift from linearity in LSB units. In this case the Signal-to-Bias Ratio, which does not average out with statistics, would read

$$SBR = 20 \log_{10} \frac{S}{\text{INL}} \text{dB}\,, \tag{4.36}$$

where S is the considered signal dynamic converted to LSB units. So in the case of HFI, with a best case signal dynamic of 14LSB, for the solar dipole, and an in-flight measured INL value about 1LSB, we have SBR = 23dB. If we would like to reach a requirement SBR= 80dB we would need at least 10 more bits with the same INL value of 1LSB. It is possible to discard insignificant bits in the on-board DPU, but as the accuracy requirement also increases with new instruments, discarding too many bits in the acquired samples may not be possible.

Increasing moderately the steady state signal amplitude is also useful. In the case of HFI, the signal balancing was a success in making the resulting analog signal very flat (70% of the *fast samples* for a half period cover only 30LSB). But a flat signal can concentrate signal acquisition on a potentially important ADC defect. Changing the design of the balancing to obtain for example a triangular shape, as presented in Fig. 4.34, for the acquired signal would be a good solution to ensure acquisition is done on different ADC areas. The result would be a reduced bias level by averaging ADC defects at different locations.

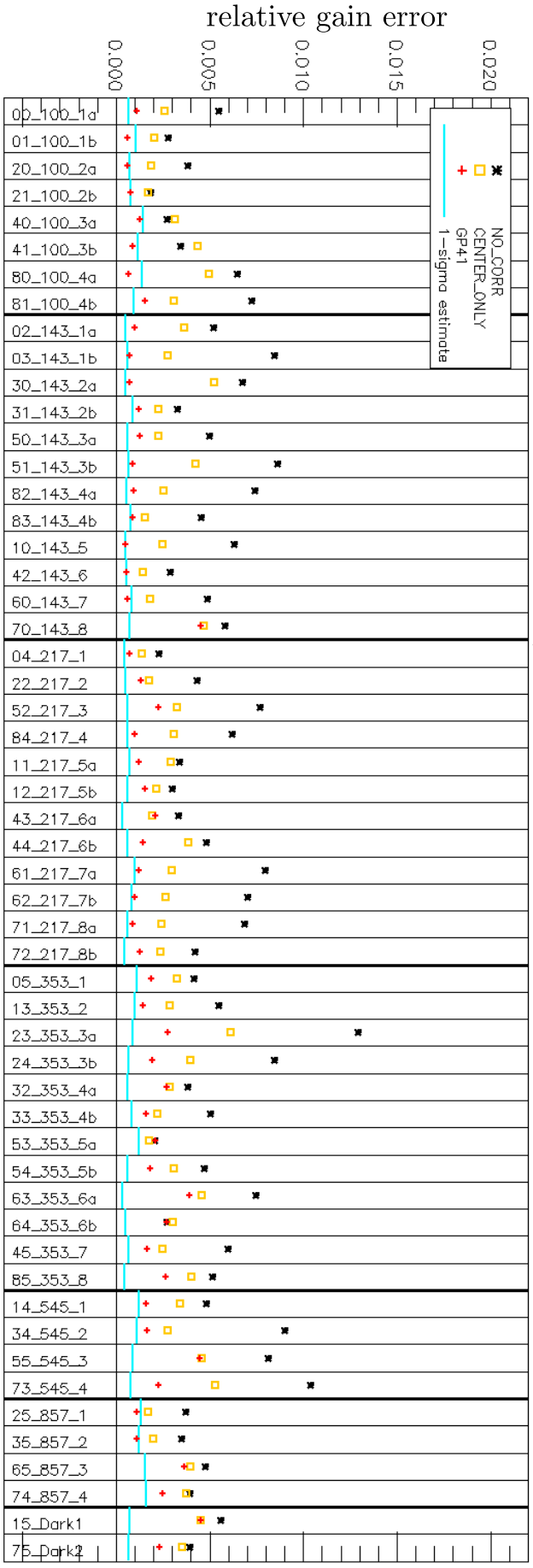


Figure 4.33: Estimation of ADC nonlinearity from *raw gain* consistency over *HFI* mission ($n_{bin} = 5$), sorted by channel frequency. A better INL correction is expected to yield a lower value. The black asterisk (NO_CORR) is estimated with no correction. The orange square (CENTER_ONLY) is for correction of midscale defect (code 32768) only. The red cross is for the correction with the INL labeled GP41 which is the best estimate of year 2015.

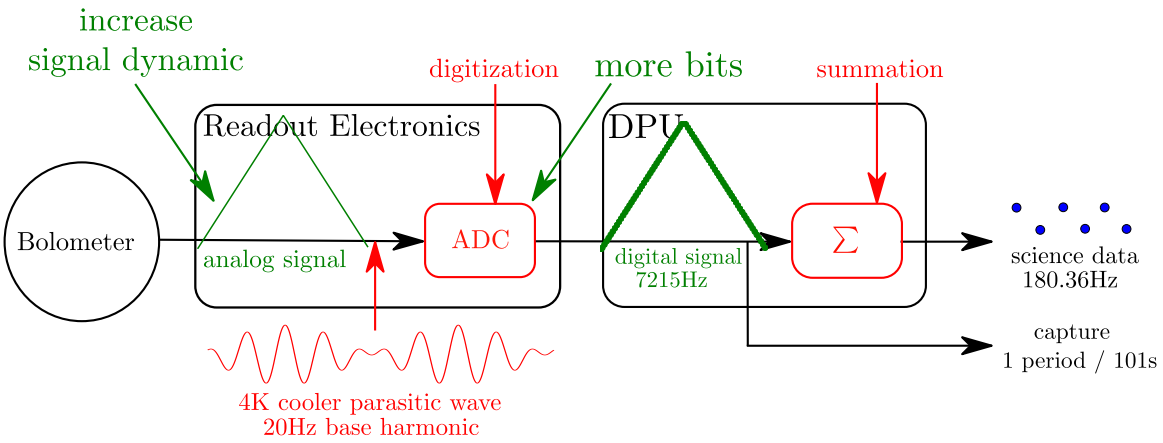


Figure 4.34: Functional diagram of readout electronics with increased signal dynamic

Using analog integration

The digital summation implemented in the DPU is designed to improve the SNR of the science data. But another function has been implemented in the DPU with the `nblank` parameter, still not used in the HFI mission, to allow the removal of `nblank` *fast samples* at transient locations. If we forget about the `nblank` option, the digital summation function can be replaced with an equivalent analog integrator in the readout chain as shown in Fig. 4.35. The square signal of detectors would integrate into a triangle wave, and the acquisition could be done near the peak location once for each half period at twice the modulation frequency. This design does not destroy information after digitization, so nonlinear correction of the ADC component can be done a posteriori.

The main problem with analog integrators is their potential signal instability, thus they have to be designed carefully. They are also an additional source of noise.

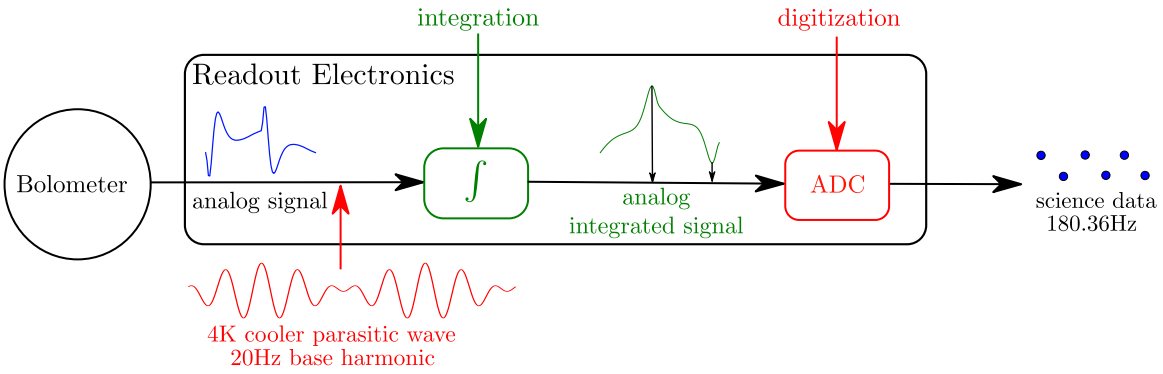


Figure 4.35: Functional diagram of readout electronics with analog integration. The analog integrated signal is calculated with the 40 *fast samples*-large boxcar average of a *raw constant*.

DPU correction

It is possible to make the the ADC nonlinearity correction in the DPU as shown in Fig. 4.36. This method requires the characterization of each ADC and the provision of this information to the DPU as data tables, one value for each ADC code being sufficient. Also, to avoid floating point operations in the FPGA, the tables must be built with significantly more bits

than the ADC output. The main advantage is that complex digital processing is still available with this design, and it has a better SNR than with an additional analog component.

For this solution, the hypothesis is that the nonlinear characteristics of the ADC are stable over the mission. Otherwise a runtime characterization must be implemented with a perfectly known calibration source such as a ramp signal.

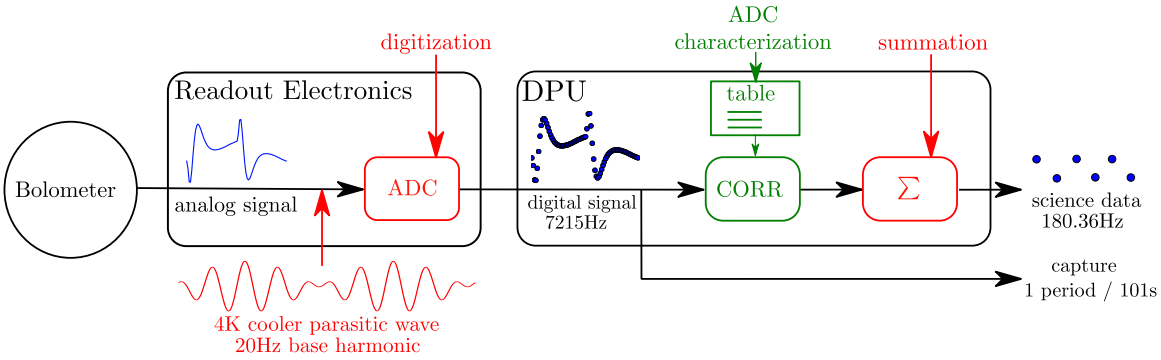


Figure 4.36: Functional diagram of readout electronics with in-flight ADC correction

4.9.2 New track for the INL estimation

It should be possible to produce the INL estimation without any free parameters, and that should significantly improve the result in the location of the 64 codes just after midscale. This is because the probability of leakage from the source noise distribution into the INL is very high with the maximum likelihood method described in Sec. 4.7.2.

In practice the INL estimation from Gaussian noise has been implemented differently by Guillaume Patanchon than with the method described in Sec. 4.2.2. It was acknowledged at the time when synthesizing his manuscript in March of 2016. The main difference is that in the likelihood case 1620 free parameters have to be fitted and degenerate with the INL estimated at each iteration. On the contrary, the method developed in Sec. 4.2.2 does not rely on the Gaussians mean value, albeit the σ_i parameters of the gaussians still need to be inter-calibrated.

The problem with the inter-calibration of the σ_i parameters can be solved by using the science data TOIs which have been produced at the usual rate during the EOL sequence. Unfortunately the PDF of the noise on these TOIs is also affected by ADC nonlinearity, but it is possible to apply an ADC nonlinearity correction **on these PDF directly**. By doing so, the 10 months of the EOL sequence can be ADC corrected iteratively very quickly and produce updated values for the σ_i parameters. Hence with this methodology the INL can be estimated by direct calculation.

A notice must be added for the compression (see Sec. 5.2.3 in **?**) of science data which is neglected most of the time and has been found to have a negligible impact on the ADC nonlinearity (unpublished result from the *Planck* team at IAP). In this case it can be considered as a convolution of the science data noise PDF with a boxcar average function, thus it must be added in the noise model of the science data TOIs for high accuracy results.

4.10 Conclusions

This chapter covers the basics of how *HFI* ADCs works and why they had such a large impact on the science data of *HFI*. The various attempts, on-ground and in-flight, for ADC

characterization and modeling have been described. It led to the final and accurate characterization of ADC INLs using a likelihood method on the data acquired during the EOL sequence. Finally it has been shown how the limited dataset of the cold *fast samples* provides insights on the accuracy of the estimated INLs.

The in-flight characterization of ADC nonlinearity, despite many difficulties, was a success. It allowed the measurement of INLs for the 54 detectors channels with an absolute accuracy close to the target ($\approx 0.3 \times 10^{-2} LSB$). The probabilities to get such a good result were low and it required the effort of many peoples within the *Planck/HFI* consortium over several years.

During the analysis of *fast samples*, the basis of a first order signal model for the detector response at modulation level has been developed with the concept of *raw gain*, this will be detailed in the next chapter. The stability of the *raw gain* shape throughout the mission has been estimated to be better than $\approx 10^{-3}$ on most channels where the 4 K lines level is moderate.

Chapter 5

Correction of science data

The ADC characterization was a big issue and globally a success, thanks to the contribution of many people from the *Planck* Collaboration, but this is only the first step to the correction of *HFI* science data. The complex readout chain of *HFI* includes a summation that destroys the information of the exact value of single ADC output codes. Therefore to correct the ADC nonlinearity effect on science data samples, the analog signal before digitization must be modeled with an **exquisite precision**.

The principle of the *HFI* electronics readout modeling is introduced in Sec. 5.1. The model for the analog signal and the production of its parameters are detailed in Sec. 5.2. When digitizing this analog signal and summing the *fast samples*, the noise needs a specific modeling, it is explained in Sec. 5.3. Next, the science data transfer function $\mathcal{F}$ is produced in Sect. 5.4, and the noise model is validated with simulations. Finally the correction is validated in Sec. 5.5 against real data and simulations.

The 4 K lines effect on the signal and the modifications needed to include its effect in the science data transfer function will be described in Chap. 6.

5.1 Principle

It is possible to apply the ADC nonlinearity effect on the analog signal model, to build the forward transfer function $\mathcal{F}(P_{\rm sky})$ of the *HFI* readout chain which associates to the input sky power a corresponding nonlinear science data sample. The science data correction is a straightforward derivation of this function.

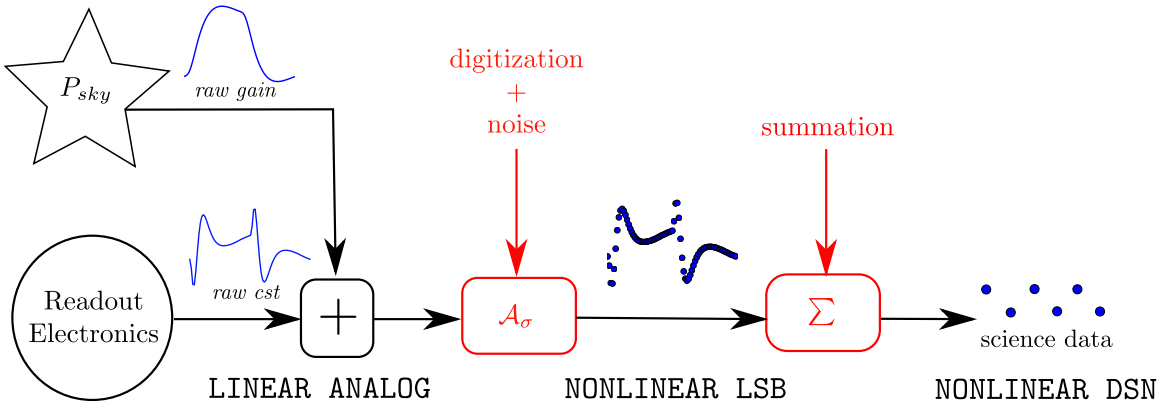


Figure 5.1: Simplified model of the *HFI* readout electronics with ADC nonlinearity

To build the analytical transfer function of the readout chain, some simplifications are necessary. A simplified version of the readout chain is presented in Fig. 5.1, where all components are considered linear, excepted the digitization by the ADC.

The analog signal is considered in steady state, and can be represented with a periodic signal, the *raw constant*. This signal is considered just before digitization, and has been filtered by the amplification chain (see Appendix A: Electronics model in **?**)). The sky power P_{sky} falling on the detector adds a periodic signal described with the *raw gain*.

The summation step averages the noise digitized by the ADC with the analog signal, thus the digitization step has to be considered along with the noise model. The digitization plus noise function will be noted $\mathcal{A}_\sigma$, and its nonlinear output is in LSB units.

Finally the nonlinear science data samples in DSN units are produced by the summation of the analog signal digitized with the function $\mathcal{A}_\sigma$.

5.2 Analog signal model

5.2.1 Slow sky signal variations

For the cosmological channels (100 GHz to 217 GHz included), the sky signal variations are dominated by the solar dipole, except for strong point sources and the bright Galactic center which are both masked for CMB analysis.

Fig. 5.2 shows the typical demodulated science data (s_d) signal for a 143 GHz channel at a ring index where the solar dipole at its maximum amplitude. A complete sky circle is shown. To estimate the dynamic of the signal at *fast samples* level, a factor 1/40 has been applied on the demodulated science data values. In the lower panel, the signal derivative is shown. As the solar dipole dynamic is described very accurately by a sinusoid, then for an amplitude of ≈ 10 LSB, the maximum variation of the *fast samples* signal Δs considering a sinusoid periodicity of 10800 science data samples is $\approx 1 \times 10^{-3}$ LSB. The signal from the primary anisotropies is neglected compared to the one of the solar dipole.

Hence the signal dynamic is dominated by the noise. The sky signal can be approximated as a constant over a modulation half period. This assumption is necessary to build the first order model of the analog signal.

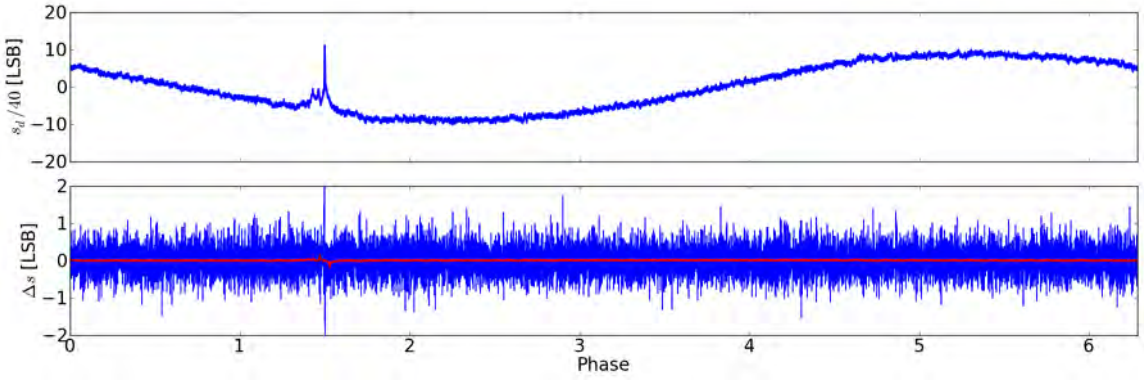


Figure 5.2: Typical signal dynamic of the solar dipole on channel 10_143-2a. The TOI data is binned in phase with the same angular resolution as for the acquisition to reduce the noise. *Upper panel*: phase binned ring of the demodulated science data signal at ring index 3100 (glitches flagged out and average value subtracted). *Lower panel*: is the signal variation between one bin and the next one (the blue line). The red thick line is for the same data after applying a boxcar average of 300 bins. The X axis is the phase of the satellite, in the current sky circle being acquired in the corresponding stable pointing period.

5.2.2 Signal Model

The results from the analytical model analysis described in Chap. 2 are used to write the first order approximation of the signal. It reads

$$s(t)|_{\langle P_{\rm sky}\rangle+\delta P_{\rm sky}} = \overline{s}(t) + \delta P_{\rm sky} \times g(t) + \mathcal{O}({\delta P_{\rm sky}}^2)\,, \tag{5.1}$$

where $\overline{s}(t)$ is the signal in steady state called the *raw constant*, $g(t)$ has been already defined as the *raw gain*, and $\delta P_{\rm sky}$ is the small input sky power variation from the average value $\langle P_{\rm sky}\rangle$. It has already been shown from the detector response analysis (see Sec. 5.4 in Sauvé and Montier 2016) that the slowly varying signal approximation holds for $\delta P_{\rm sky}$ pulsations up to 2 Hz with a precision better than 1% relative to the term $\delta P_{\rm sky} \times g(t)$.

Heareafter, $P_{\rm sky}$ will be calculated in DSN units (summed LSB) to be consistent with *HFI* TOI data, and g in LSB/DSN units. Therefore to a signal variation Δs_d on demodulated science data, the corresponding analog signal variation $\Delta s(t)$ can be written $\Delta s_d \times g(t)$.

5.2.3 The *raw gain*

Estimation with a Principal Component Analysis (PCA)

The *raw gain* is estimated with a PCA (a clear synthesis of the method is provided in **?**) that is run on the difference between consecutive *fast samples* periods δp_i (see Sect. 4.8.2). The resulting vector will be noted $\boldsymbol{g}$ and have a size of 80.

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of linearly uncorrelated variables called principal components. The orthogonal principal components are defined as the eigenvectors of the covariance matrix, which is symmetric. In this case, the covariance matrix built from the vectors δp_i has a size of (80, 80), thus the PCA algorithm provides 80 eigenvectors. The vector $\boldsymbol{g}$ is expected to be the eigenvector with the highest eigenvalue.

The PCA procedure was run on all the available *fast samples* periods (about 70000), its outputs are shown in Fig. 5.3 for the channel `00_100-1a`. In the covariance matrix, the negative/positive checkerboard is the result of the square shape of the raw gain, which is the dominant feature. On the right panel, the raw gain appears as the first eigenvalue which is clearly the most significant, and is about 50 times above the eigenvalues noise plateau (values after index 30). This is a good indicator that the extraction of the raw gain has a low degeneracy level with other effects. Finally, the inspection of the first eigenvectors (bottom plot) shows a lot of sinusoidal components. These sinusoids are clearly the result of the residuals 4 K lines variations on the δp_i vectors, which is discussed in Chap. 6. As a first guess, a manual fit for the frequencies of the sinusoidal vectors yields only expected 4 K lines harmonics frequencies.

The PCA procedure has been run with excellent results on most *HFI* detector channels, and the resulting *raw gains* are shown for two channels in Fig. 5.4. There is some variation in the detector response depending on the detector time constant. This is visible for the 857 GHz channel which has a quicker stabilization after the transient. For a comparison purpose, the first time constant of the `00_100-1a` channel is about 10 ms (i.e. about two half periods), while it is only 3.4 ms for the `25_857-1` channel (i.e. about 0.6 half periods). The strong influence of the detectors time constants on the *raw gain* is an interesting result, because it illustrates how the *raw gain* provides constraints for the time transfer function of the detectors. Hence after the first successful extractions of the *raw gain*, a work has beens tarted on the analytical model of the time transfer function of the detectors. The results of this work have already been presented in Chap. 2.

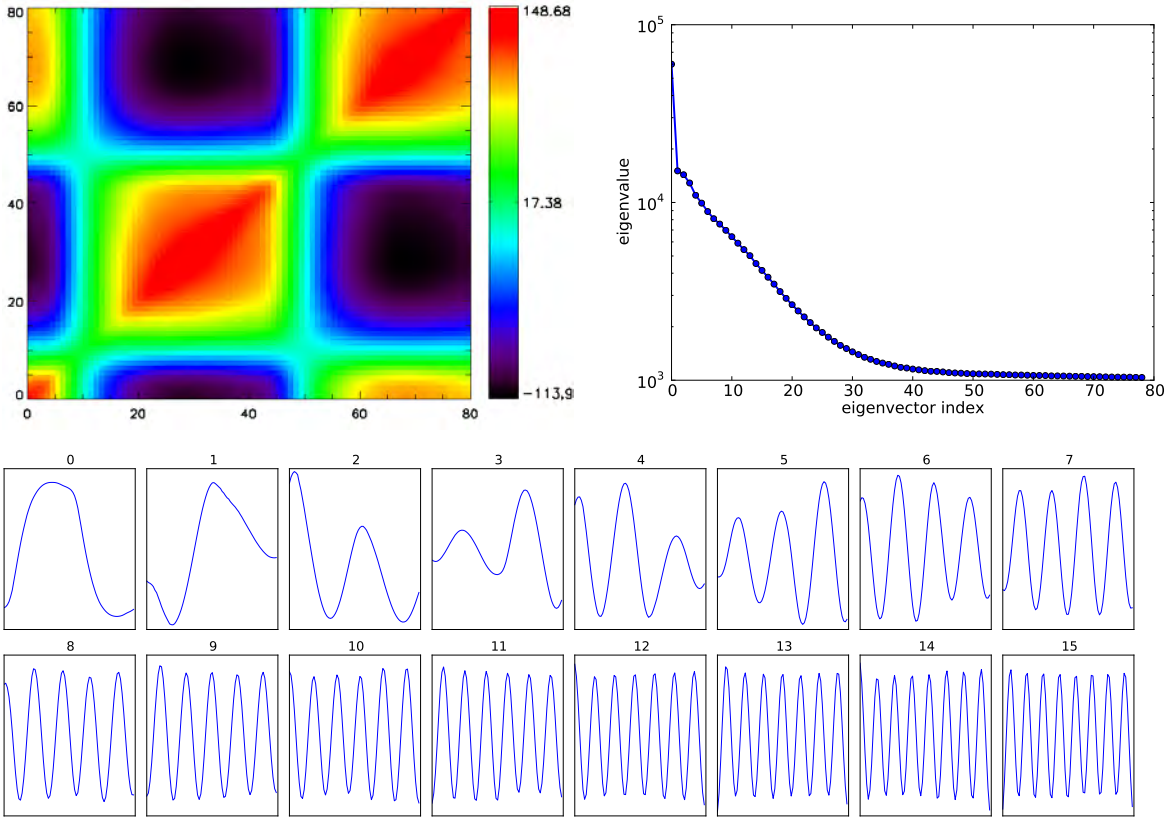


Figure 5.3: Principal Component Analysis outputs for channel `00_100-1a`. *Top left panel*: covariance matrix of *fast samples* period differences δp_i, units are LSB^2. The colorscale is linear. *Top right panel*: eigenvalues yielded by the PCA algorithm and sorted by absolute value. *Bottom panel*: first 16 eigenvectors. The vector at index zero is the *raw gain*.

Inference of the relative error on the *raw gain*

It is important to know the relative error $\varepsilon(\boldsymbol{g})$ on $\boldsymbol{g}$ which is a global parameter to the signal model all over the mission, because it will propagate as a residual bias after ADC nonlinearity correction. $\varepsilon(\boldsymbol{g})$ is defined as

$$\varepsilon(\boldsymbol{g}) = \frac{\delta g}{\max(|g|)}\,, \tag{5.2}$$

where δg account for total errors of the $\boldsymbol{g}$ estimate.

A first estimate of the noise level can be inferred from the data. It can be observed that the maximum signal variation of the *fast samples* period induced by the solar dipole is about 7 LSB. Then it comes

$$\varepsilon(\boldsymbol{g}) \approx \frac{\sqrt{2}\sigma}{7\sqrt{n_p}}\,, \tag{5.3}$$

where n_p is the total number of *fast samples* periods used and σ is the RMS of the *fast samples* noise for the considered channel. This simple estimator yields an average value of 1×10^{-3} on *HFI* detector channels, which is in good agreement with the results obtained from *raw gain* consistency over the mission (see Sect. 4.8.2).

Since the *raw gain* is expected to be very symmetric, it can be used to estimate the bias on the PCA outputs. The expected symmetry has raised a lot of discussion, because the bolometer bias signal has a difference of a few percents in the raise/fall times of the input compensation square wave (see Chap. 2). To characterize this asymmetry simulations

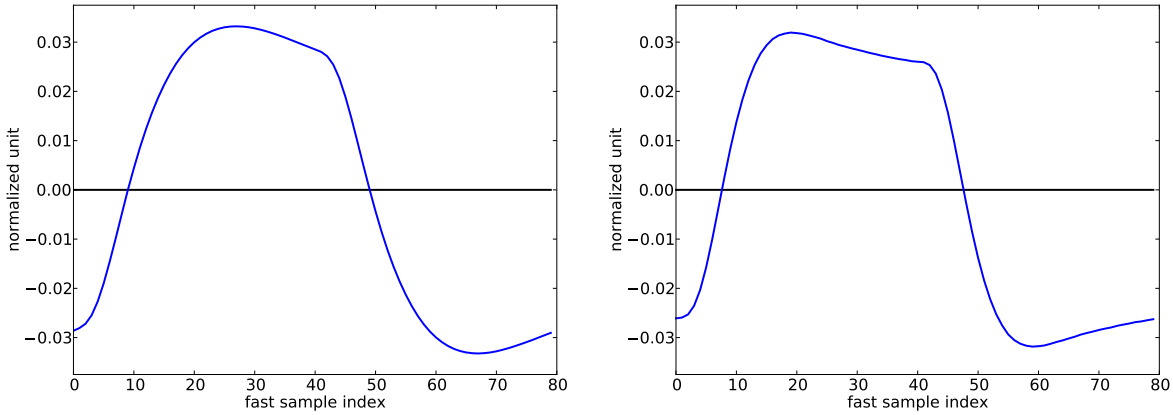


Figure 5.4: Example of *raw gains* estimated with PCA. Left plot is for the channel 00_100-1a. Right plot is for channel 25_857-1.

with `SEB` tool have been run, injecting the same bias tension asymmetry as measured on-ground, yielding a relative asymmetry of $\approx 10^{-5}$ on the measured *raw gains*. Hence it will be considered that the real *raw gain* asymmetry due to raising times is not significant compared to the noise level of the data.

The final estimates derived from asymmetry are shown in Fig. 5.5 for all channels, using $\delta g = \mathrm{RMS}(\boldsymbol{g}(j) + \boldsymbol{g}(j+40), j \in [\![0, 39]\!]$. The relative error is in general below 5×10^{-3}, but is significantly higher than the expected statistical level on most channels. This level is not compatible with the *raw gain* consistency throughout the mission which is much better as seen in Sec. 4.8.2. As a consequence, given the shape of the residual eigenvectors, and neglecting ADC correction errors, the most probable source of asymmetry on the estimated *raw gain* appears to be the leakage of 4 K lines signal in the *raw gain* estimate. There may be also a contribution from glitch tails, and signal dynamic effects (the second order of the analog signal model).

The error on the *raw gain* propagates as a bias in the analog signal model. This bias can be estimated considering the signal dynamic of the dipole which is about 7 LSB around the steady state average signal. The bias is expected to have a maximum level of $\varepsilon(\boldsymbol{g}) \times 7 = 0.035$ LSB. This value is acceptable in the case of the cosmological signal, because the *raw gain* error will not bias significantly the correction, compared to the uncertainties on the ADC characterization.

5.2.4 The *raw constant*

Deglitching

It is necessary to remove *fast samples* periods affected by transient events before computing the *raw constant* and the *raw gain*. This can be done in a simple manner by filtering out the summed squared residuals R_i above the 95 percentile threshold. As the *fast samples* period deformation is very different for glitches or high dynamic signal, these outliers are efficiently filtered out with this method. Some well known events like solar flares and EOL tests are also filtered out by removing the periods in the selected event time slice.

The Fig. 5.6 shows an example of glitch falling on the detector inside the first half period. After the event, the detector is completely out of balance in the *ringing* phase of the glitch.

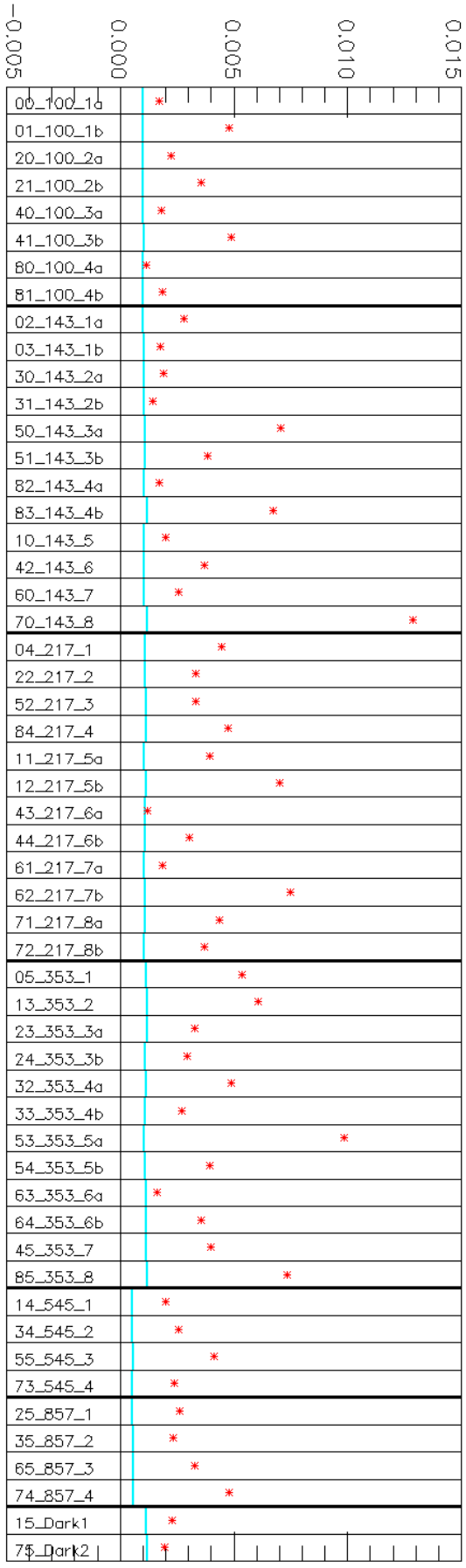


Figure 5.5: Relative error of *raw gains* estimated from asymmetry, with channels sorted by frequency. The cyan line is the minimum error estimate from noise only. The red star is the asymmetry relative error estimate. The *fast samples* have been corrected with the best available INL (version GP41) before running the PCA.

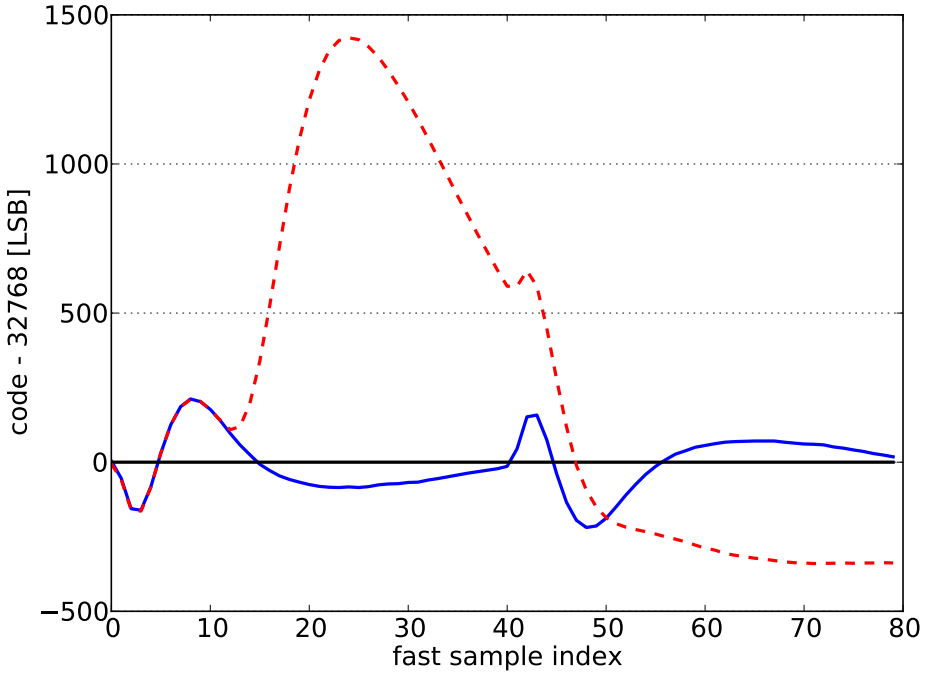


Figure 5.6: Example of glitch on *fast samples* period for channel `00_100-1a`, where the energy deposit effect is visible after index 10. The blue line is a nominal period, and the red dashed line is the following period impacted by a glitch.

Characterization the *raw constant* parameter

There were two known sources of variations of the *raw constant* shape all over the mission, at the time of the production of the *HFI* 2013 ADC correction:

- The sky power;
- The modulation offset.

The modulation offset, also called the *AC offset*, is the average value of the signal on science data before demodulation, it's average value is 32768×40 DSN on science data, and 32768 LSB on *fast samples*. As there are significant leaks from the crosstalk signal of neighbours channels (characterized to be as high as 100 LSB), the *AC offset* variation are not well characterized. In theory it is a constant level, but it can also have complex variations depending on the crosstalk signal.

For the characterization of the *raw constant* drift all over the mission, the sky signal variations have to be removed. These variations on a given *fast samples* period p_i will be estimated with

$$\Delta \boldsymbol{c}(p_i) = \Big(s_d(p_i) - s_d(p_{ref})\Big) \times \boldsymbol{g}\,, \tag{5.4}$$

where p_{ref} is a period chosen as reference, and $s_d(p)$ is calculated with Eq. 4.29. After removal of $\Delta \boldsymbol{c}(p_i)$, the residuals show wavy variations (see Fig. 5.7), relative to the chosen reference. These wavy residuals are not due to *AC offset* variations alone, because they are not symmetric, and symmetry is expected for a steady state signal.

Per ring binning

As the variations of the *fast samples* period over the *HFI* mission cannot be modeled (they are mostly due to 4 K lines degenerating with the $\Delta \boldsymbol{c}(p_i)$ estimate), it has been decided,

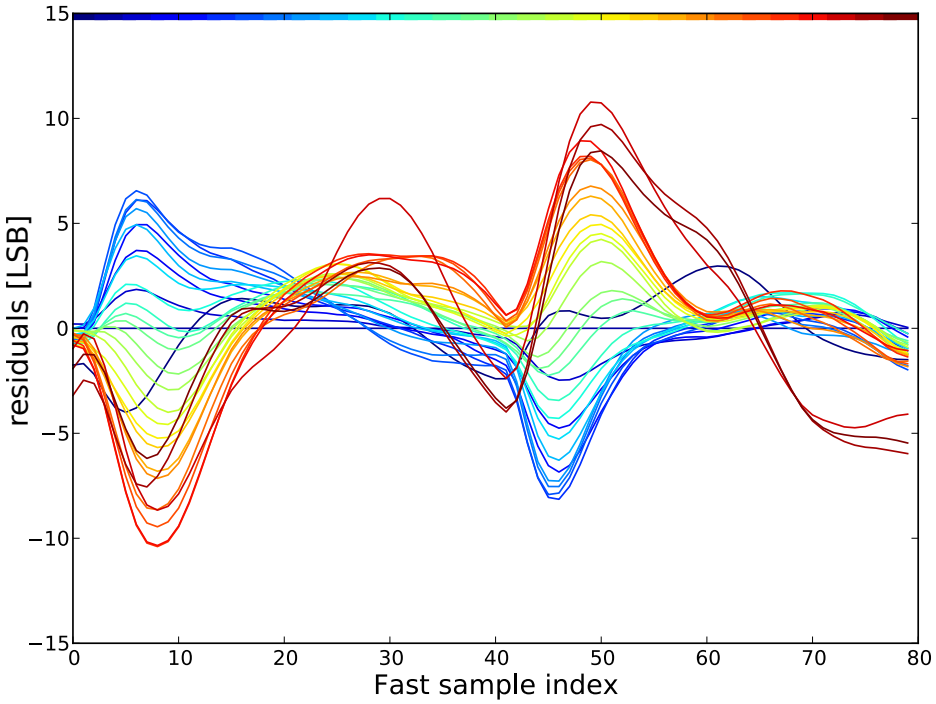


Figure 5.7: *raw constant* drift over the *HFI* mission for the channel `42_143-6`, relative to ring 1000. The upper line show the color code color ranging from black to red over the full mission.

within the ADC team, to use a binning at the ring level (or one stable pointing period). Hence the *raw constant* parameter expression for ring index r is

$$\boldsymbol{c}_r(j) = \langle \{p_i(j)\}_r \rangle \,, \tag{5.5}$$

where the *fast samples* period set $\{p_i\}_r$ includes only periods p_i captured inside ring r, for which R_i is below the 95 percentile threshold. This methodology allows us to extend the timescale of the binning to several rings without changing the code.

The main drawback to the one ring binning, is an increased bias of the correction at ring level due to the residual noise which is about $\approx \sigma/\sqrt{35}$ (0.6 LSB on CMB channels), as there are 35 *fast samples* period per ring on average over the mission. Then to overcome this issue, François Couchot has put a lot of efforts into building a semi-analytical model of the *raw constant*, but he could not achieve a satisfying level of accuracy because of:

- The 4 K lines parasitic signal.
- Crosstalk signal leaks from multiple neighbor channels, having very different shapes;
- The mixing of the complex bolometer response with the balance square;

5.2.5 Iterative estimation of the parameters

The knowledge of the *raw gain* is mandatory for the filtering of the *fast samples* periods outliers, but at the beginning this parameter is unknown. As a consequence, the estimations of $\boldsymbol{g}$, and $\boldsymbol{c}_r$ have to be done with an iterative process. Before proceeding, the ADC correction is done once and for all, using the best available INL estimate.

In practice, 3 steps are necessary :

- a first pass compute a rough estimate of the *raw gain* $\boldsymbol{g}^0$ and the *raw constants* $\boldsymbol{c}_r^0$ from a basic filtering of *fast samples* periods. Using Eq. 4.29 the demodulated signal amplitude of the *fast samples* periods differences $|s_d(\delta p_i)|$, is used to discard periods p_i yielding a value above the 95 percentile. The first estimates $\boldsymbol{g}^0$ and $\boldsymbol{c}_r^0$ are produced using the filtered data set, which allow to compute R_i^0 with Eq. 5.10 ;
- The second pass produce accurate estimates of $\boldsymbol{g}^1$ and $\boldsymbol{c}_r^1$, using the dataset filtered with R_i^0. Then R_i^1 is calculated with $\boldsymbol{g}^1$ and $\boldsymbol{c}_r^1$;
- The final versions $\boldsymbol{g}$ and $\boldsymbol{c}_r$ are produced, as for pass two, with the estimates from the previous iteration. They are used to calculate the final estimate for the RMS of *fast samples* noise σ.

5.3 Digitization with noise

5.3.1 Noise model

The noise at the input of the ADC has to be taken into account in the model, because each science data sample is the sum of 40 noisy *fast samples*. As for the EOL sequence, it can be approximated as a Gaussian white noise component $n_\sigma(t)$ with a standard deviation value σ added to the signal. The remaining noise at frequencies higher than f_{mod} is considered as signal from the ADC point of view, and is not taken into account.

For a constant signal $s(t) = s$ in volt at the input of the ADC, the effect of noise is a local dispersion of ADC output codes around the $\mathcal{A}(s)$, where the ADC transfer function $\mathcal{A}$ has been defined in Sec. 4.7.2. Considering a large number of *fast samples* n_f their average value will converge, so we can write

$$\langle \mathcal{A}(s + n_\sigma(t)) \rangle \underset{n_f \to \infty}{\longrightarrow} \mathcal{A}_\sigma(s)\,, \tag{5.6}$$

where $\mathcal{A}_\sigma$ is the effective transfer function of the ADC for a noisy signal. The expression for $\mathcal{A}_\sigma$ is

$$\begin{aligned} \mathcal{A}_\sigma(s) &= \int_{-\infty}^{+\infty} \mathcal{N}_\sigma(x)\mathcal{A}(s+x)dx \\ &= \int_{-\infty}^{+\infty} \mathcal{N}_\sigma(x)\mathcal{A}(s-x)dx \\ &= (\mathcal{A} * \mathcal{N}_\sigma)(s)\,, \end{aligned} \tag{5.7}$$

where $\mathcal{N}_\sigma$ is the normal PDF with a standard deviation value of σ and $\mathcal{A} * \mathcal{N}_\sigma$ is the convolution product between $\mathcal{A}$ and $\mathcal{N}_\sigma$[1]. The level of the noise is considered to be constant throughout the mission and is about 3.7 LSB for CMB channels and 1.7 LSB for submm channels (for which the relative gain is $\texttt{gamp} = 1/3$), the method to estimate these values is detailed in Sec. 5.3.3.

An example of the effect of the convolution of $\mathcal{A}$ by the noise is shown on Fig. 5.8. This channel has been chosen because it exhibits visible steps at 64 codes boundaries, thus it shows how effective is the reduction of HCF nonlinearity. The convolution has a smoothing

[1] It is faster and much more precise to use the erf function to compute the integral of $\mathcal{N}_\sigma(x)$ over the ranges where the discrete output of $\mathcal{A}$ is constant. In this case a finite $\pm 5 \times \sigma$ window for $\mathcal{N}_\sigma$ covers about 36 codes, so the calculation of $\mathcal{A}_\sigma$ needs about $65536 \times r \times 36$ operations, where r is the resolution of $\mathcal{A}_\sigma$ or the number of points per quantization step. For the erf based solution r=3 is sufficient with a step size of 0.33 LSB compared to $\sigma \approx 3.7$ LSB. For a DFT based method, the discrete step like nature of $\mathcal{A}$ needs a resolution $r = 256$ to avoid ringing effects, leading to $65536 * r * \log2(65536 \times r) = 65536 \times r \times 6144$ operations.

effect making the 64-code jumps become the dominant features. The linear input tension V_{IN} is expressed in LSB units (offset to have $\mathcal{A}(32768) = 32768$) , this is also the case in all the *HFI* data processing code, which allow to convert straightforwardly nonlinear ADC codes values to linear LSB units with $\mathcal{A}^{-1}$ (see Eq. 4.27). Let us stress here, that while this plot is similar to an INL, it is not an INL, for the reason that the X axis domain is linear (V_{IN}). While for an INL, the X axis is the nonlinear code value.

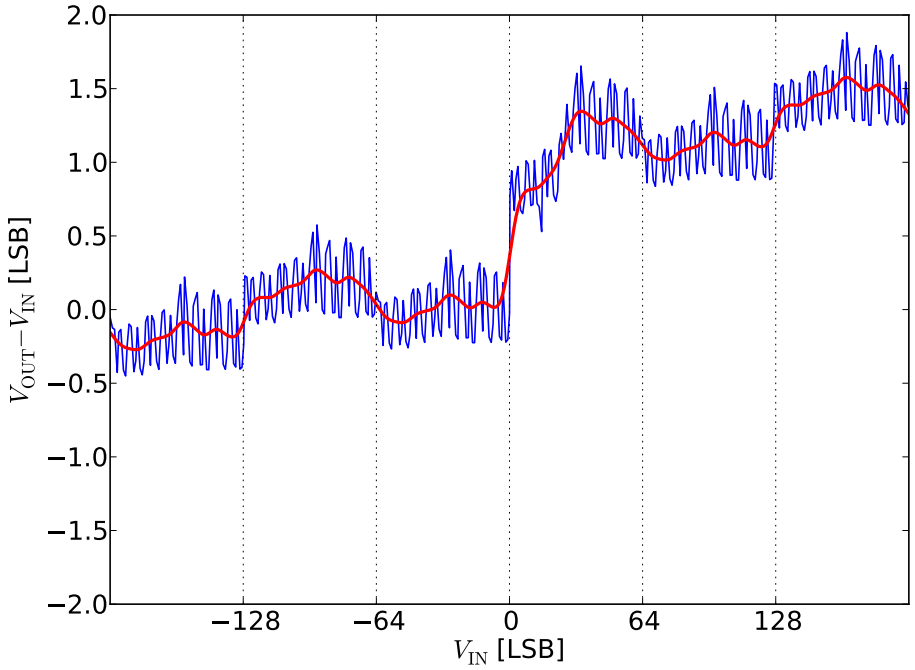


Figure 5.8: Effect of noise on the ADC transfer function for the channel `01_100-1b`. The blue line is for $\mathcal{A}(V_{\mathrm{IN}}) - V_{\mathrm{IN}}$. The red line is for the convolved version $\mathcal{A}_\sigma(V_{\mathrm{IN}}) - V_{\mathrm{IN}}$.

5.3.2 summation of noisy *fast samples*

Now let us consider the DPU summation on the 40 *fast samples* of a half period for a fixed value of the input sky power P_{sky}. The value of a science data sample for one half period of the modulation is given by

$$s(t)|_{P_{\mathrm{sky}}} = \sum_{j=0}^{39} \mathcal{A}\Big(s(j\Delta t) + n_\sigma(t + j\Delta t)\Big)\,, \tag{5.8}$$

where Δt is the time between the digitization of two *fast samples*, and t has been omitted on purpose from $s(j\Delta t)$ to reflect the periodicity of the modulation for a stable input sky power (this will be discussed further in the next section).

By considering that the noise on *fast samples* is uncorrelated, then for a large number of science data samples using Eq. 5.7 yields

$$\underbrace{\left\langle \sum_{j=0}^{39} \overbrace{\mathcal{A}\Big(s(j\Delta t) + n_\sigma(t+j\Delta t)\Big)}^{\text{discrete}} \right\rangle}_{\text{continuous}} = \sum_{j=0}^{39} \left\langle \mathcal{A}\Big(s(j\Delta t) + n_\sigma(t+j\Delta t)\Big) \right\rangle$$

$$\simeq \sum_{j=0}^{39} \mathcal{A}_\sigma(s(j\Delta t)) \tag{5.9}$$

The key point here, is that it is supposed that $s(j\Delta t)$ is known, so an accurate description of the analog input signal before digitization is needed.

5.3.3 Estimation of the *fast samples* noise RMS

The white noise RMS value σ of the *fast samples* signal is estimated from the *fast samples* periods, after signal removal with the model of Eq. 5.1. Its estimated is performed on a per ring basis. A rough estimate is sufficient for the noise level, because it is a second order parameter for the correction (smoothing of INL).

First, the *fast samples* are corrected using $\mathcal{A}^{-1}$ (from Eq. 4.27). Then a *raw constant* vector $\boldsymbol{c}_r$ is calculated from the average of all periods captured within the ring at index r, and a single *raw gain* vector $\boldsymbol{g}$ is used, using the estimation described in Sec. 5.2.3. For each *fast samples* period p_i a single free parameter $P_{\text{sky}}(i)$ is fitted[2]. The squared residuals

$$R_i = \sum_{j=0}^{80} \Big(p_i(j) - (\boldsymbol{c}_{r(i)}(j) + \delta P_{\text{sky}}(i) * \boldsymbol{g}(j))\Big)^2, \tag{5.10}$$

where $\boldsymbol{c}_{r(i)}$ is the *raw constant* vector for ring $r(i)$ built from the average of all *fast samples* periods p_i falling in ring r, and δP_{sky} is the variation from the $\langle P_{\text{sky}} \rangle$ value corresponding to $\boldsymbol{c}_r(i)$. The reason for binning on one ring will be detaild in Sec. 5.2.4.

The value of σ can be inferred from the R_i distribution, by assuming that R_i is described by a χ^2 distribution built with 80 independent Gaussian random variables. In this case there are $n_{\text{free}} = 79$ free parameters, so it is assumed, using the formula for a theoretical χ^2 distribution that

$$\underbrace{\text{median}(R_i)}_{\text{measure}} = \underbrace{\sigma^2}_{\text{unknown}} \times (n_{\text{free}} - 2/3)\,. \tag{5.11}$$

For visual inspection the histogram of R_i (normalized as a PDF) is shown in Fig. 5.9 for the first 100 GHz channel, along with the theoretical χ^2 PDF. The real data distribution is much more compact than the theoretical one. The possible cause for this discrepancy is some correlation in the fit residuals.

The cause of the distribution excessive compactness is investigated with three simulations:

- CASE 1: using white noise only to produce *fast samples* periods;
- CASE 2: using a realistic noise produced by filtering white noise with the analytical model of the *HFI* bandpass filter of the electronics readout. At *fast samples* period level, the noise has a $1/f$ shape which produce some correlation between samples;

[2] Its value is not used as we are only interested in the noise estimation

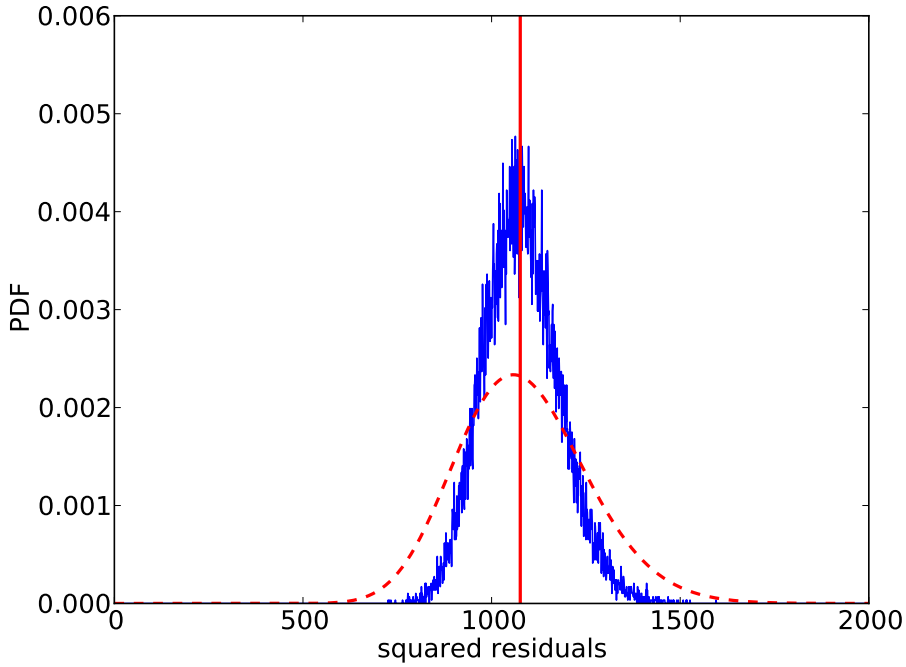


Figure 5.9: PDF of the squared residuals for *fast samples* periods fit on channel `01_100-1b`. The red vertical thick line is the median of the squared residuals. The red dashed line is the theoretical χ^2 distribution for $n_{\text{free}} = 79$ and X axis scaled with σ^2.

- CASE 3: using white noise plus a sinusoid at $2 * f_{\text{mod}}$. This case mimics some intra-ring variations of the 4 K lines signal which is the dominant source of systematics effect on *fast samples* after ADC nonlinearity (see Chap 6).

The resulting squared residuals distributions are compared to the one of the real data in Fig. 5.10. For case 1, the distribution of the squared residuals follows exactly the theoretical χ^2 distribution. For case 2, the distribution widens and moves to the origin, this is consistent with a smaller number of degrees of freedom, but it does not describe the real data. For case 3, an excellent agreement with real data is obtained when tuning by hand the total noise to a slightly lower RMS value of 3.68 LSB. In this case, the white noise RMS value is set to 1.52 LSB and the sinusoid RMS value to 3.35 LSB. The agreement obtained in case 3 is a strong indicator that the short time scale variations of the 4 K lines are an important contributor to the total noise power at *fast samples* periods level. However this component cannot be modeled easily and is considered hereafter as white noise from the point of view of the digitization.

The value of σ estimated for CMB channels is on average 3.7 LSB, while for the submm channels (545 GHz and 857 GHz) it is 1.7 LSB respectively. This difference is due to the relative gain factor of 1/3 applied to the submm channels. Unfortunately this low noise RMS value is an issue for submm channels because the correction will be more susceptible to high code frequency errors in the estimated INLs, as the noise smoothing will be less efficient.

5.4 Science data transfer function

5.4.1 Separation of modulation half periods

The *fast samples* summation on each half period of the modulation, imposes to model separately the response of the readout chain for the positive (parity 0) and negative (parity 1)

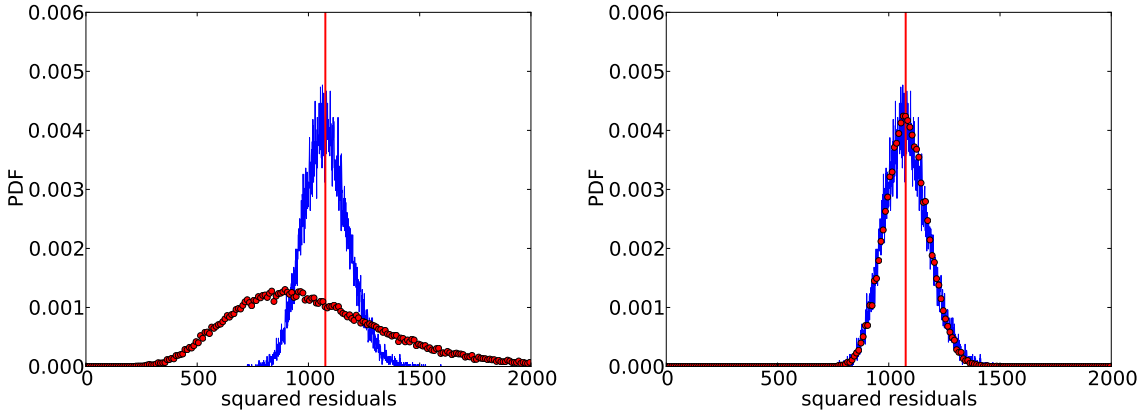


Figure 5.10: PDF of the squared residuals for *fast samples* periods compared with simulations. The color code is the same as for Fig. 5.9. *Left*: (case 2) noise generated using the *HFI* electronic filter. *Right*: (case 3) noise generated with a combination of white noise and a sinusoid (values given in text).

half periods. The Fig. 5.11 shows the $\boldsymbol{c}_r$ and $\boldsymbol{g}$ parameters shifted to the first *fast samples* index of parity 0. For convenience, the parameters for parity 0 will be noted ($\boldsymbol{c}_{0,r}$, $\boldsymbol{g}_0$), and for parity 1 ($\boldsymbol{c}_{1,r}$, $\boldsymbol{g}_1$), respectively. A side effect of the *fast samples* window shift, relative to the half periods boundaries, is that the parity 1 half period is not causal, i.e. one half is visible at the beginning of the period and another one at the end of the period corresponding to a different science data sample. While this is not an issue for the correction method, the channels which are the most impacted by the 4 K lines parasitic wave show significant asymmetry on the *raw constant* parameter.

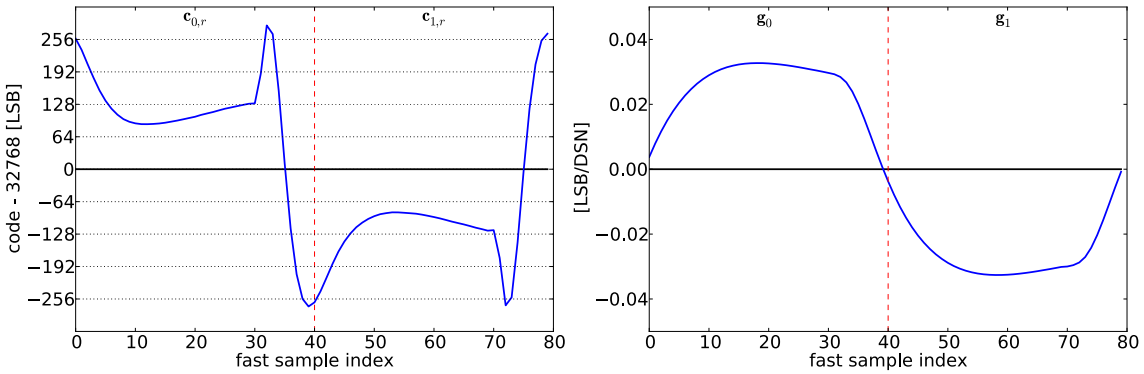


Figure 5.11: Correction parameters shifted to the first *fast samples* integration index for channel `01_100-1b`. *Left panel*: *raw constant* for $r = 1000$. *Right panel*: *raw gain*.

5.4.2 Transfer function expression

Finally, putting all together with Eq. 5.9, the expression for the response of the readout chain with a nonlinear ADC to the input sky power is

$$\mathcal{F}_{p,r}(\langle P_{\rm sky}\rangle + \delta P_{\rm sky}) = \sum_{j=0}^{39} \mathcal{A}_\sigma\Big(\underbrace{\boldsymbol{c}_{p,r}}_{\langle P_{\rm sky}\rangle}(j) + \delta P_{\rm sky}(i) \times \boldsymbol{g}_p(j)\Big)\,. \tag{5.12}$$

This will be referred to as the **science data transfer function** hereafter. It must be noted that this function is increasing monotonously because this is a property of $\mathcal{A}_\sigma$, which results from the convolution of a gaussian with the ADC transfer function which is monotonous. This property is important because it allows us to use the reciprocal of $\mathcal{F}$ for the correction of science data.

An example of the resulting functions $\mathcal{F}_{p,r}$ for both parities are shown in Fig. 5.12. It is visible from the ranges covered by the signal within one ring (the colored vertical bars) that the thermal drift (moving this range on the x-axis) of the detector plates changes completely the nonlinear effect throughout the mission. It results in observations that were difficult to interpret before the ADC issue discovery. However let stress that this function changes also with the pointing period because of the $\boldsymbol{c}_r$ parameter drift in time. Hence **it is not possible to fit a single function for the full mission**, even if it is very smooth.

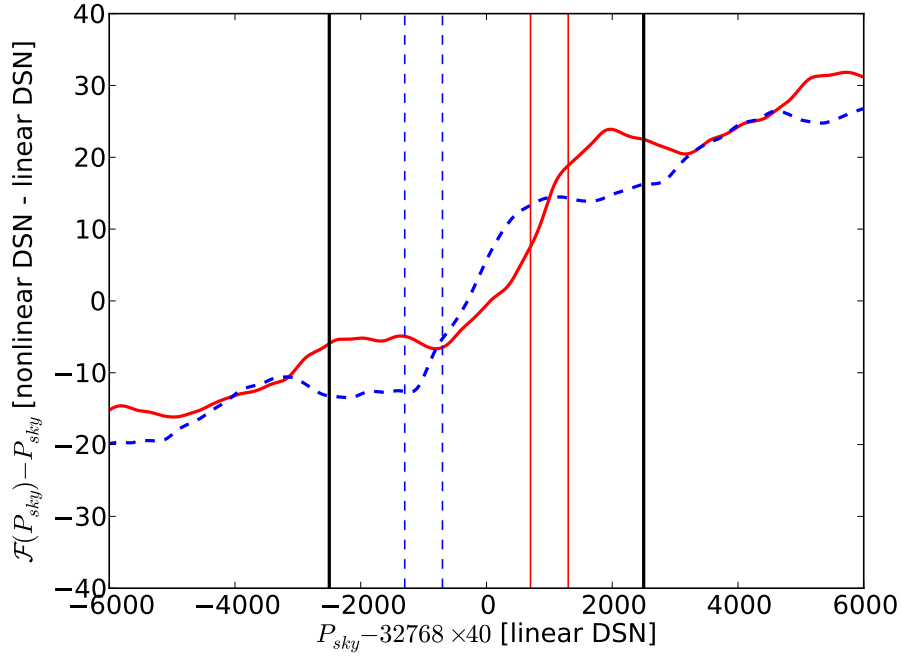


Figure 5.12: Science data transfer function of readout chain with nonlinear ADC for both modulation parities on channel `01_100-1b` (the same as for the Fig. 5.8) at ring $r = 1500$. Red line is for parity 0 ($\mathcal{F}_{0,1500}$), and vertical red lines show the typical range covered by the solar dipole range on one ring. Blue dashed line is for parity 1 ($\mathcal{F}_{1,1500}$) respectively. The black thick vertical lines show the range covered by the full mission due to thermal drift of the detector plate.

5.4.3 Correction of science data

In practice, the function $\mathcal{F}_{p,r}^{-1}$ is applied on the unprocessed science data TOIs, and called the correction function hereafter. For the 27 000 rings of the mission, 54 000 correction functions will be needed, one for each ring and each parity.

5.4.4 Validation of the transfer function with simulation

An important milestone is to check the effectiveness of the correction function on noisy digitized data with simulation. This is done by producing science data samples from *fast samples* periods using the analog signal model and digitizing with the discreteADC transfer function function $\mathcal{A}$, the science data samples values are calculated by summation. The result of the simulation can be compared to the output of $\mathcal{F}(P_{\mathrm{sky}})$.

The simulation steps for a fixed value of $P_{\text{sky}} = \langle P_{\text{sky}} \rangle + \delta P_{\text{sky}}$ are:

- an input sky power variation value δP_{sky} is selected;
- the analog modulation period is generated with $\boldsymbol{p} = \boldsymbol{c} + \delta P_{\text{sky}} \times \boldsymbol{g}$;
- white Gaussian noise is added to *fast samples* yielding $\boldsymbol{p}_\sigma(i) = \boldsymbol{p}(i) + n_\sigma(i)$;
- the analog *fast samples* period is digitized with $\mathcal{A}$;
- nonlinear science data samples (s_0, s_1) are calculated with the sum of each half period. This operation is repeated n_s time to produce the datasets $\{s_0\}$ and $\{s_1\}$. Booth sets will be designed with $\{s_p\}$ using p for the parity hereafter, the same will hold for $\mathcal{F}_p$;
- then $\langle \mathcal{F}_p^{-1}(\{s_p\}) \rangle$ is compared with P_{sky}.

First it has been checked that the simulation is consistent with $\mathcal{F}$, this is performed by injecting a ramp for δP_{sky} at simulation input. The simulation result is shown in Fig. 5.13 and compared to $\mathcal{F}_{0,1000}$ for parity zero. For a direct comparison the noise RMS on the produced samples which is $\sigma \times \sqrt{40} \approx 23$ DSN, is reduced by a factor 1000 by producing $n_s = 10^6$ averaged science data samples for each P_{sky} point (the big red circles). The RMS of the residuals between the simulation and the science data transfer function is on average ≈ 0.02 DSN and is consistent with the expected noise value $\sigma \times \sqrt{40/n_s}$, yielding a perfect agreement between the simulation producing discrete *fast samples* and the science data transfer function. Hence the noise model is validated.

Next, the correction is checked by applying $\mathcal{F}^{-1}$ on the simulated science data samples. The corrected science data samples are averaged and the expected value of the average is the input value for P_{sky}. The result for parity zero is shown in Fig. 5.14. The RMS of the residuals $P_{\text{sky}} - \langle \mathcal{F}_0^{-1}(\{s_0\}) \rangle$ with $n_s = 10^6$ is on average ≈ 0.02 DSN which is in perfect agreement with the expected value $\sigma \times \sqrt{40/n_s}$. The result is the same for the parity one respectively. This an important validation of the correction function, because it introduces **neither bias nor additional noise** to the corrected science data.

Application to simulations at full mission scale

It has been shown with the simulations that for a large number of samples

$$\langle \{s_p\} \rangle \underset{n_s \to \infty}{\longrightarrow} \mathcal{F}_p(P_{\text{sky}})\,. \tag{5.13}$$

This result is of great interest for simulations of ADC nonlinearity effect over the full mission, because $\mathcal{F}$ can be used to produce very quickly science data instead of the calculation of *fast samples* which needs 40 times more CPU resources. The drawback is that these quick simulations do not have the same noise properties and are limited to the study of ADC nonlinearity.

This method will be used in the next section for validating the correction over the full mission.

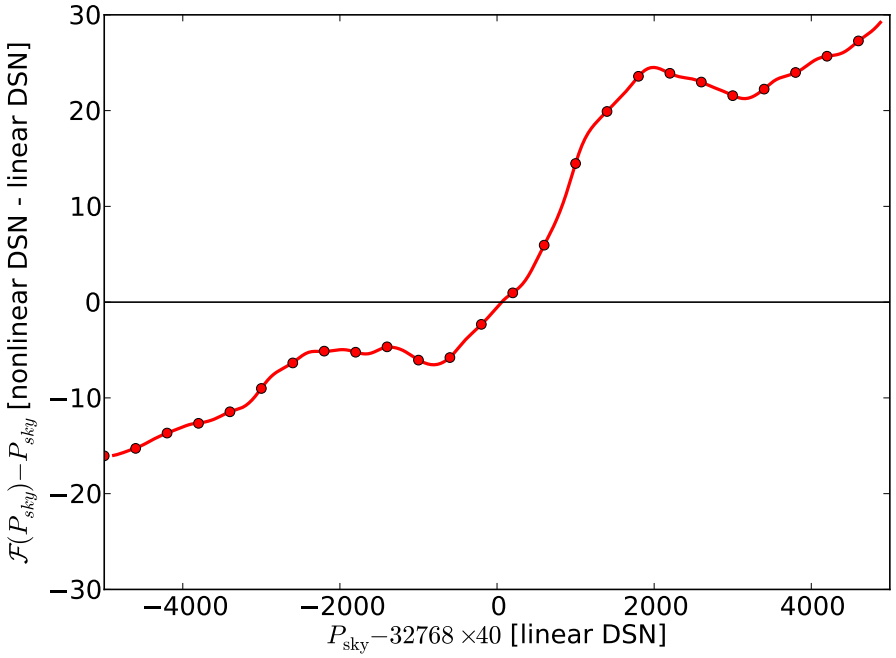


Figure 5.13: Simulation at *fast samples* level of the readout chain with nonlinear ADC, for the channel `01_100-1b` at ring index 1000. The red dots are the averaged output of $n_s = 10^6$ simulated science data samples for parity zero. The red line is the science data transfer function $\mathcal{F}_{0,1000}$. Errors bars are smaller than the dots and are not shown.

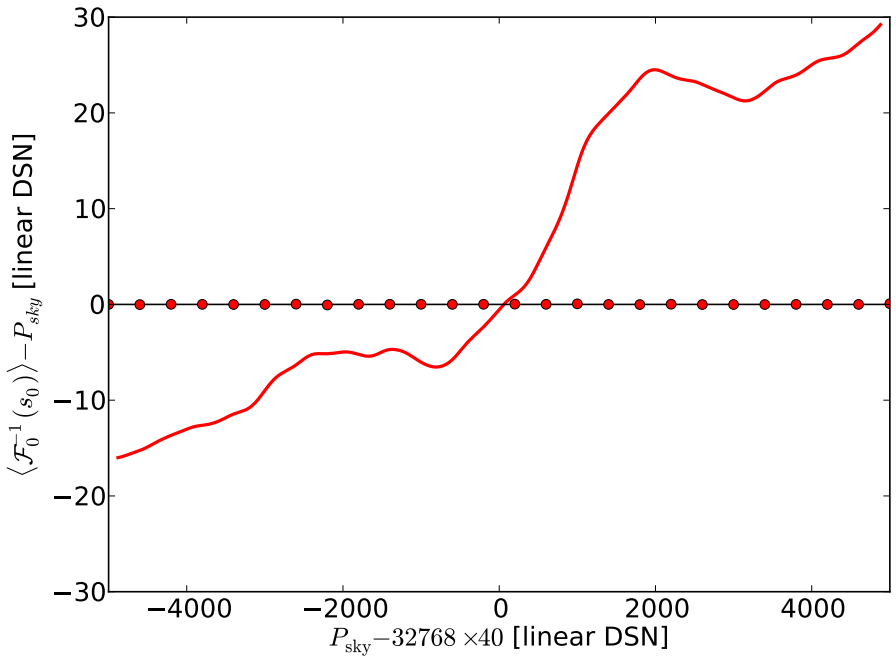


Figure 5.14: Simulation of ADC nonlinearity correction residuals for parity zero when averaging many noisy science data samples. The channel and ring index are the same as for Fig. 5.14, and there is $n_s = 10^6$ science data samples generated for each selected value of $P_{\rm sky}$. Red dots are the ADC nonlinearity correction residuals $P_{\rm sky} - \langle \mathcal{F}^{-1}_{0,1000}(\{s_0\})\rangle$. The red line if for the science data transfer function, which has been left as an eye guide.

5.5 Validation of the correction

The astrophysical signal acquired by the *Planck* satellite is by nature unknown. There is here a fundamental problematic for characterizing the ADC nonlinearity effect on science data, and the effectiveness of the correction defined in the previous section. Hopefully, the very nature of ADC nonlinearity has an effect which is different for both half periods of the modulated analog signal. Therefore this difference can be used to build observables.

First the method for calculating ADC nonlinearity observables is described. Next they will be used to characterize the ADC effect with simulation, and to validate the correction model on real data.

5.5.1 Formalism

Before proceeding, let start with a simple description of the science data acquired by the satellite during it's rotation over the sky. This description will be the reference for the science data TOI and its version binned in phase: the PBR (as described in Sec. 1.3.1).

A stable pointing period is referenced by its ring index r, and during this period of time, the satellite does on average 35 complete rotations over the sky. A complete satellite rotation will produce one sky circle, the source signal is noted $\mathrm{SKY}(\Phi)$ and is considered in DSN units. For each modulated science data sample $s_m(i)$ at index i, the position inside the current scan circle is referenced by it's phase $\Phi(i)$, and stored in the `Phase` TOI.

Using these notations, the modulated signal of the science data TOI can be written

$$s_m(i) = o_{\mathrm{AC}} + (-1)^{p(i)} \times \mathrm{SKY}(\Phi(i)) + N(i)\,, \tag{5.14}$$

where o_{AC} is the modulation offset (about 32768×40) for the current ring, $\mathrm{p}(i)$ returns the parity of the sample at index i (zero or one), and $N(i)$ is the science data noise considering only frequencies below f_{mod}. For the time being, $N(i)$ is supposed white and Gaussian.

After binning in phase of the science data TOI for each parity separately, the two resulting PBRs read

$$\begin{cases} \mathrm{SKY}_0(\Phi(k)) &= o_{\mathrm{AC}} + \mathrm{SKY}(\Phi(k)) + n_0(\Phi(k)) \\ \mathrm{SKY}_1(\Phi(k)) &= o_{\mathrm{AC}} - \mathrm{SKY}(\Phi(k)) + n_1(\Phi(k))\,, \end{cases} \tag{5.15}$$

where n_0 and n_1 are the noise residuals after projection, and k is the bin index of the PBR which contains usually 10800 bins corresponding to the average number of samples per spacecraft rotation.

5.5.2 Production of PBRs with the 1D drizzling algorithm

The 1D drizzling projection

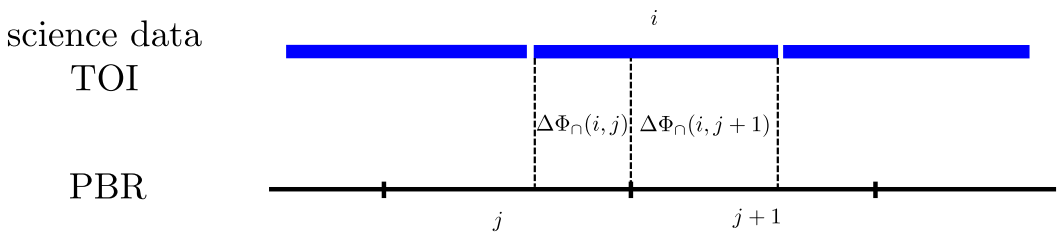


Figure 5.15: The 1D drizzling projection

The 1D drizzling is a custom version of the drizzling (**?**) algorithm used for the linear reconstruction of Hubble deep field images. The goal here is to project samples from the 35

(on average) sky acquisitions of one ring, that are not phase-synchronous, into two phase-synchronous sky circles, one for each parity of the modulation.

Fig. 5.15 shows the projection principle of the 1D drizzling algorithm:

- A science data sample $s_m(i)$ at TOI index i, integrates the sky power over a phase range $\Delta\Phi_{\mathrm{IN}}(i)$;
- The phase range $\Delta\Phi_{\mathrm{IN}}(i)$ intersects with one or more output bins $\mathrm{PBR}(j)$ on the phase ranges $\Delta\Phi_{\cap}(i,j)$.
- The quantity $a_{ij} \times s_m(i)$ is added to each intersected output bin $\mathrm{PBR}(j)$, where $a_{ij} = \Delta\Phi_{\cap}(i,j)/\Delta\Phi_{\mathrm{IN}}(i)$ is the ratio of the input power redistributed on $\mathrm{PBR}(j)$.
- The ratios a_{ij} are added to their corresponding output bin weights $w(j)$.

This projection process is repeated for each non flagged sample of the science data TOI, and after the projection step, the non empty output values $\mathrm{PBR}(j)$ are normalized with the total accumulated weight $w(j)$.

In practice this kind of projection is robust against leaks of signal derivative into the comparison between the PBR built with party zero and one respectively. This property will be required for the estimation of ADC-related relative gain variations between the two PBRs at 10^{-4} level.

Filtering effect of the 1D drizzling

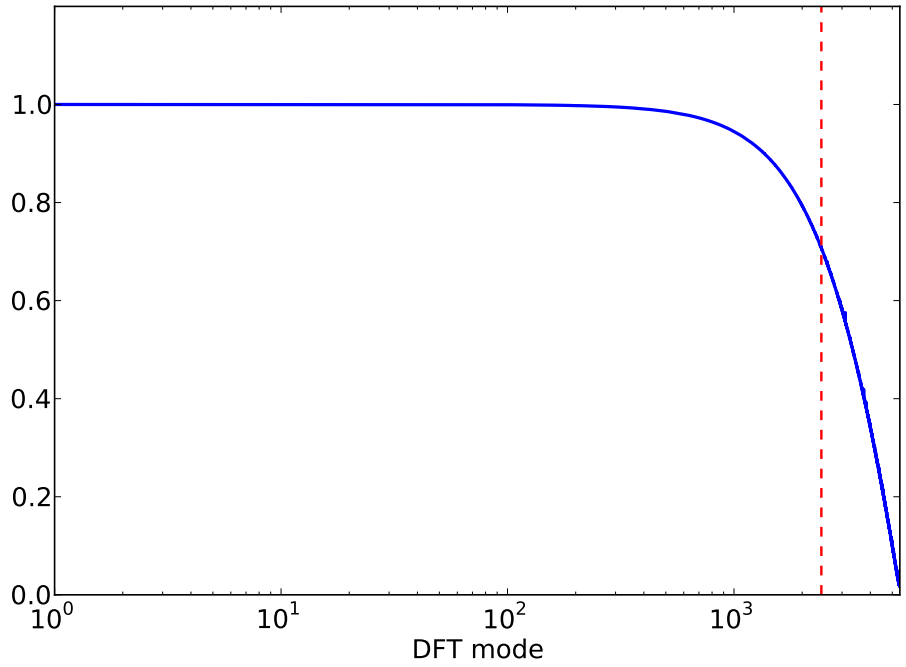


Figure 5.16: Transfer function modulus of the 1D drizzling in the *HFI* configuration. The red vertical dashed line indicates the location of the -3 dB cutoff frequency. DFT mode 1 correspond to one single rotation on the sky or 2π in phase, and the x-axis extends on the right up to half the angular sampling frequency.

Let have a look at the filtering effect, in frequency domain, of the 1D drizzling. This analysis is restricted to the *HFI* case, where input and ouptut bin size are similar in the phase domain, because the general case is much more complex. The transfer function modulus of

the 1D drizzling is shown in Fig. 5.16. It has been produced with a random sky circle $\mathrm{SKY}_{\mathrm{IN}}$, it is then converted to a science data TOI (with a similar resolution in phase) by interpolating the `Phase` TOI, and projected with the 1D drizzling to $\mathrm{SKY}_{\mathrm{OUT}}$, using the same resolution in phase as for $\mathrm{SKY}_{\mathrm{IN}}$. The figure shows the ratio of the discrete Fourier transform of $\mathrm{SKY}_{\mathrm{OUT}}$ over $\mathrm{SKY}_{\mathrm{IN}}$. Hence the 1D drizzling acts as a low pass filter with a cutoff frequency close to 1/4 of the sampling frequency, this is a consequence of the smoothing introduced when splitting science data samples over more than one output bins.

The low pass filter behavior has a positive effect, because it filters out the sky signal in a frequency range mostly dominated by noise. The ADC nonlinearity effect builds harmonics from the dominant signal of the solar dipole, which is the first mode of the PBRs DFT. These harmonics are expected close to their source frequency, so the low pass filtering, with a cutoff frequency over mode 1000, will not affect the dominant ADC nonlinearity residuals.

Half parity PBRs

The 1D drizzling projection is applied separately to science data samples of parity zero and parity one, respectively. The yielded phase synchronous PBRs are shown in Fig. 5.17, and are labelled SKY_0 for parity zero, and SKY_1 for parity one respectively. These PBRs are **modulated**, i.e. the projection does not remove any offset, nor change the signal sign. Because of the averaging of about 35 sky circles, the noise is reduced by a factor $\sqrt{35}$ compared to the unprocessed science data TOI.

The final expression for the drizzled PBRs is derived from Eq. 5.15 and reads

$$\begin{cases} \mathrm{SKY}_0(\Phi) &= o_{\mathrm{AC}} + o_\varepsilon + \mathrm{SKY}(\Phi) + \varepsilon_0(\Phi) + n_0(\Phi) \\ \mathrm{SKY}_1(\Phi) &= o_{\mathrm{AC}} + o_\varepsilon - \mathrm{SKY}(\Phi) + \varepsilon_1(\Phi) + n_1(\Phi)\,, \end{cases} \tag{5.16}$$

where o_ε is the small modulation offset change introduced by ADC nonlinearity, $(\varepsilon_0(\Phi), \varepsilon_1(\Phi))$ are the small ADC nonlinearity effect on the sky signal which is different for both parities. The impact of ADC nonlinearity on the PBR noise is neglected.

In some rare (and well identified) cases, the satellite spin frequency f_{spin} is an even integer multiple of the acquisition frequency f_{acq}. These rings are called *resonnant rings*, and in this case all sky circles are phase synchronous, meaning that SKY_0 and SKY_1 do not have any intersection (perfect interleaving) in the phase domain. For this reason, the *echoic rings* cannot be used to build a reliable ADC nonlinearity observable and are discarded.

5.5.3 The half parity gain observable

The gain difference, introduced by the ADC nonlinearity, between SKY_0 and SKY_1 can be estimated with a linear fit. In this case the model, based on Eq. 5.17, for χ^2 minimization is

$$\mathrm{SKY}_1(\Phi) = g \times \mathrm{SKY}_0(\Phi) + c\,, \tag{5.17}$$

where g is the relative gain parameter, with a value close to -1, and c (discarded after the fit) is an offset corresponding to the thermal drift. The parameter g is expected to be close to -1 and biased by the terms $\varepsilon_0(\Phi)$ and $\varepsilon_1(\Phi)$, while the noise residuals $n_0(\Phi)$ and $n_1(\Phi)$ are expected to average out when the statistic increases. An example of fit result is shown in Fig. 5.18 for a 100 GHz channel, where the signal dynamic is dominated by the solar dipole. At first sight, there is a concern with the noise level of this measurement.

The uncertainty σ_g on g will vary with the ring index, as the solar dipole amplitude and the ring duration changes over the six month period of a full sky survey. It is a value with no unit and homogenous with a relative error because $g \approx 1$. The value of σ_g has been estimated

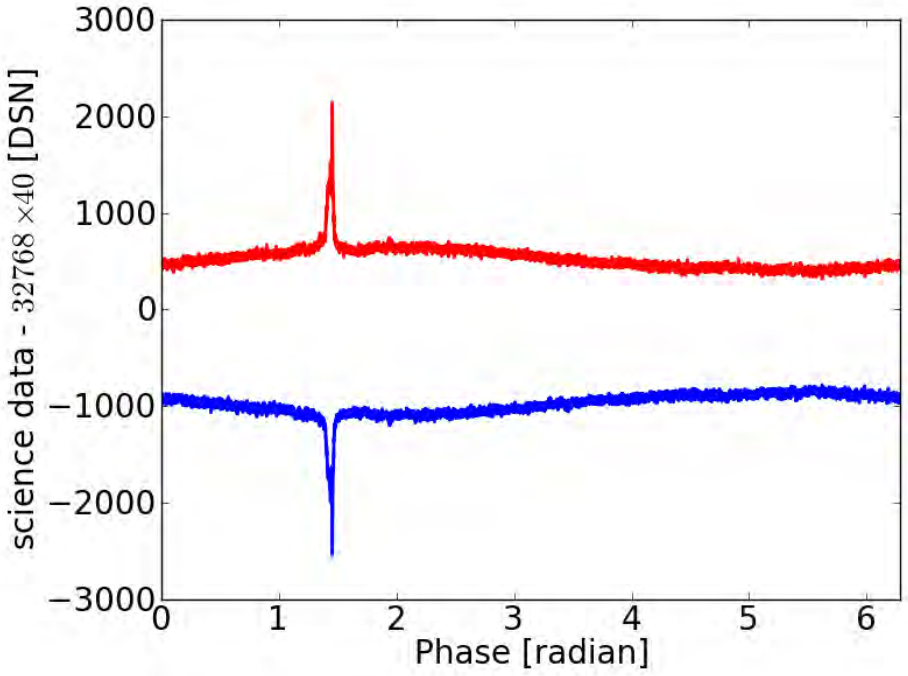


Figure 5.17: Example of modulated PBRs produced with the 1D drizzling projection. The red line is for SKY_0, and the blue line is for SKY_1.

with a Monte Carlo method using the PBR of the sky signal, which yields an average value of $\langle \sigma_g(r) \rangle \approx 2 \times 10^{-3}$.

The result after processing the full mission is shown in Fig. 5.19. Before correction, the observed gain discrepancy between the signal acquired from different parities is at 1% level. To improve the SNR of the half parity gain observable, PBRs are averaged over 128 ring before calculating g, the uncertainty on the observable is reduced down to $\approx 1.7 \times 10^{-4}$.

After application of the correction function $\mathcal{F}_{r,p}^{-1}$ on science data, the reduction of gain discrepancy between SKY_0 and SKY_1 is dramatic! This result has been extremely encouraging for the exploration of the analog model based methodology. Because in this case, the correction of science data, does only count on parameters estimated from other sources of data. i.e. this reduction of the half parity gain observable is a strong validation of the whole electronic readout chain model.

The observables for most channels are shown in Fig. 5.20 for the very first production run of the ADC nonlinearity correction. There is a significant reduction in the observable for the first channels, while for higher frequencies, the correction inputs yet needs to be improved or produced. Another point to be mentioned is the lack of 4 K lines handling in this version, which is the main reason why the observables are worse or poorly reduced on some channels.

5.5.4 The half parity sum observable

This second observable is built by calculating the standard deviation of the sum $\mathrm{SKY}_0 + \mathrm{SKY}_1$. Considering the expression of Eq. 5.16, the sum builds the value $2(o_{\mathrm{AC}} + o_\varepsilon) + \varepsilon_0(\Phi) + \varepsilon_1(\Phi) + n_0(\Phi) + n_1(\Phi)$. As for the half parity gain observable, averaging the PBRs before computing the observable will average out the terms n_0 and n_1.

The result of the processing is shown in Fig. 5.21. It has the main advantage that the expected noise level is pretty much constant over the mission, however it is more difficult to read in terms of gain variations. This is the reason why the half parity gain observable has been preferred to this one, but as it is used in the simulations (see below) this result is

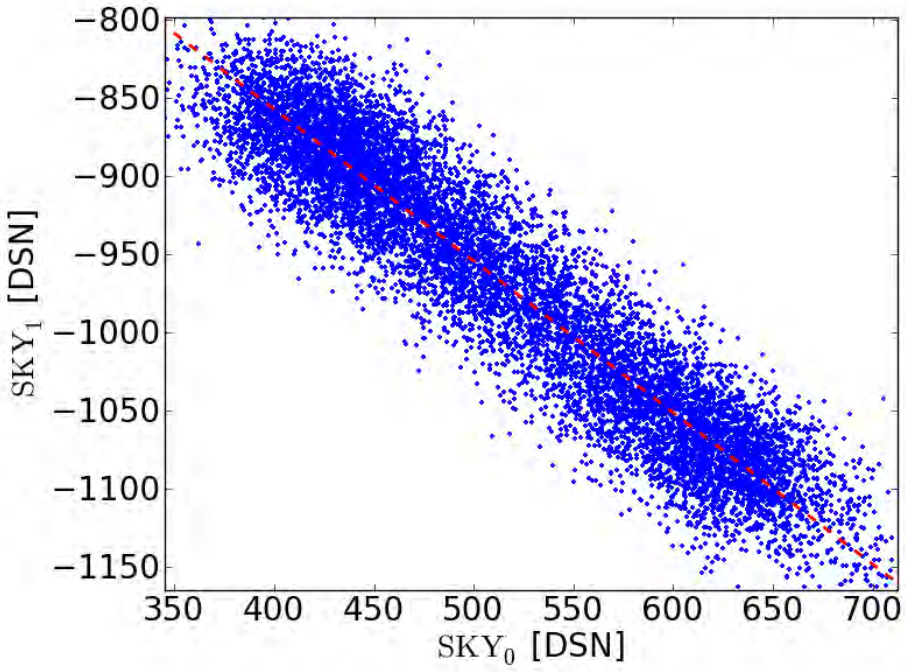


Figure 5.18: Correlation plot between PBRs produced from each modulation parity for ring 1000 of channel `00_100-1a`. The red dashed line is the linear fit of (SKY_0, SKY_1).

presented quickly here. Let note that some local steps in the estimated noise level are due to variations in the ring duration, and are consistent with the per ring observables variations when ADC nonlinearity effect is not dominant.

5.5.5 Impact of ADC nonlinearity correction on `bogopix` gains

The half parity gain observable provides in addition to the estimate of ADC nonlinearity residuals, a biased estimation of gain variations throughout the mission. It is then natural to compare them after ADC nonlinearity correction, to the results obtained from the previous estimates of gain variations obtained with `bogopix` (Tristram et al. 2011).

As stated before, `bogopix` has been developed to correct for the observed gain variations over the mission. The basic algorithm behind the scene is to fit two parameters for each ring: an offset and a gain, in order to minimize the sky survey differences. As `bogopix` compares the sky scanned over the full mission, it is not ADC nonlinearity specific because it is sensitive to other systematic effects like scalar gain variations (due to drift in analog electronic response) for example.

An example of `bogopix` gains is shown in Fig. 5.22. The effect of ADC nonlinearity correction is dramatic on the bogopix gains. This is in agreement with the half parity observables results, and is a strong validation of the transfer function of the readout chain with ADC nonlinearity $\mathcal{F}$ model. The seemingly periodic pattern on the ADC corrected data is not physical, it is a result of the averaging of all surveys by `bogopix`, so a local gain variation will produce periodic residuals in each survey.

Quantitatively the `bogopix` gain variations are around a 1.5% level. It stresses that the local half parity observable underestimate significantly the ADC nonlinearity effect all over the mission. This underestimation effect will be characterized with simulations in sect. 5.5.6. While `bogopix` allows us to compare signal acquired from very different ADC digitization regions, it is not immune itself to underestimation of gain variations.

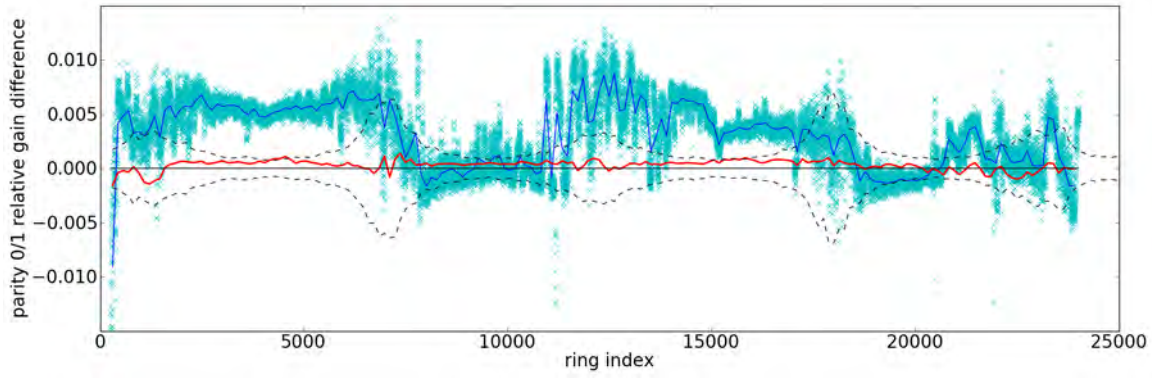


Figure 5.19: ADC observable: the half parity gain relative difference $(g+1)$ over all the mission on channel `00_100-1a`. The cyan crosses are for raw science data. The blue line is calculated using the PBRs averaged over 128 rings. The red line has been calculated after ADC correction with the best available inputs (built with INL version `GP41`), using PBRs averaged over 128 rings. The black dashed line show the 1σ envelope for g in the absence of averaging.

5.5.6 Simulations

Now that the correction function has been tested with success against real data, in this section is presented the simulation setup used to validate this methodology on science data all over the mission. The half parity observables, have to be characterized at the same time.

Simulation setup

The transfer function of the readout chain $\mathcal{F}$ (see Eq. 5.12), has been already validated with simulations. Hence it allows to skip the *fast samples* step and make direct use of science data with $\mathcal{F}$ to introduce the ADC nonlinearity effect.

Construction method for the reference and simulated PBRs:

- A reference demodulated PBR, $\mathrm{SKY_{ref}}$ is calculated for each ring with $\mathrm{SKY_{ref}}(r) = (\mathrm{SKY_0}(r) - \mathrm{SKY_1}(r))/2$;
- The modulation offset for ring r is estimated with $o_{\mathrm{AC}}(r) = (\langle \mathrm{SKY_0}(r) \rangle + \langle \mathrm{SKY_1}(r) \rangle)/2$.
- Two reference PBRs for the two signal parities are produced with $\mathrm{SKY_{ref0}}(r) = o_{\mathrm{AC}}(r) + \mathrm{SKY_{ref}}(r)$ and $\mathrm{SKY_{ref1}}(r) = o_{\mathrm{AC}}(r) - \mathrm{SKY_{ref}}(r)$. These two PBRs are consistent, phase synchronous, and considered to be linear science data. Let stress that the modulation offset is mandatory to be consistent with the *raw constant* used to produce $\mathcal{F}$;
- The reference science data $\mathrm{TOI_{ref}}$ is interpolated using the `Phase` TOI for ring r. For parity zero, samples are interpolated from $\mathrm{SKY_{ref0}}(r)$, and for parity one from $\mathrm{SKY_{ref1}}(r)$ respectively.
- The functions $\mathcal{F}_{r,0}$ and $\mathcal{F}_{r,1}$ are applied on the samples from the corresponding parity, to produce the $\mathrm{TOI_{ADC}}$ with the simulated ADC nonlinearity effect.
- The simulated TOI is drizzled to produce two modulated PBRs with the ADC nonlinearity effect: $\mathrm{SKY_{ADC0}}$ and $\mathrm{SKY_{ADC1}}$.

Note: noise is not added on TOIs in this setup, this effect is already taken into account by the function $\mathcal{F}$, and has been discussed in Sec. 5.4.4.

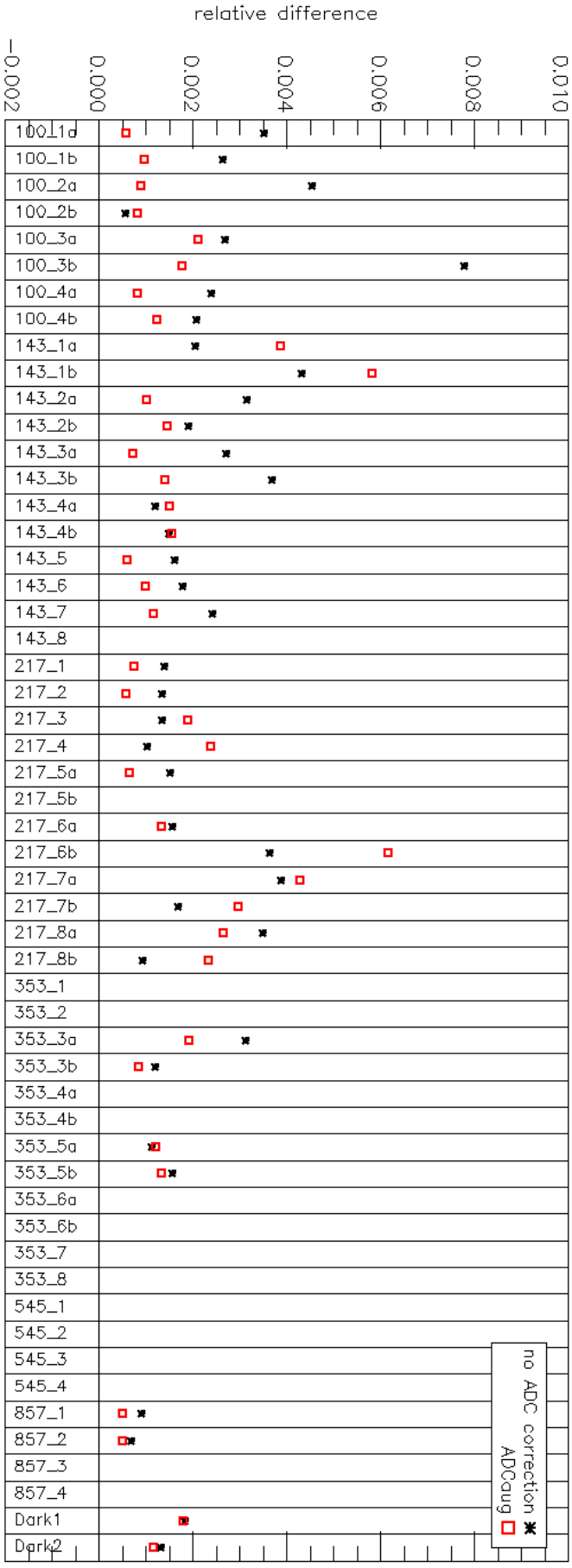


Figure 5.20: Half parity gain observable for all channels. Black stars are for uncorrected science data. Red squares are for the first experimental ADC correction built using INL versions GP12 .

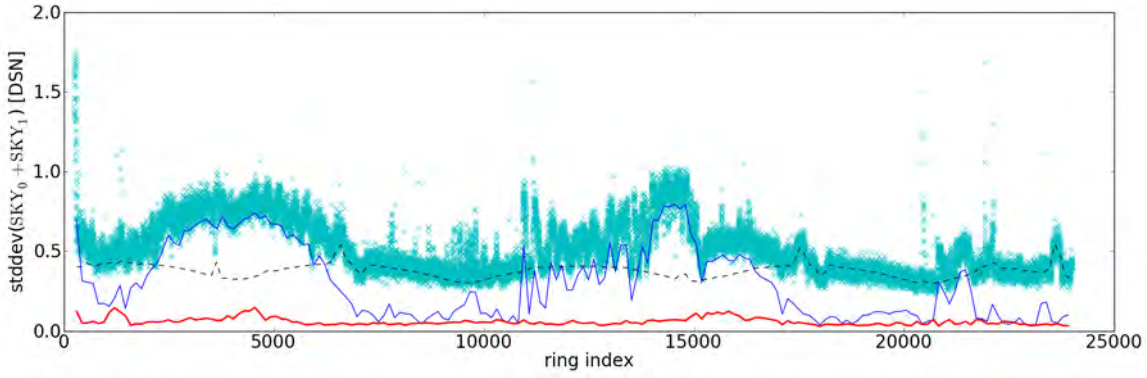


Figure 5.21: ADC observable: the half parity sum standard deviation, calculated all over the mission on channel `00_100-1a`. The legend is the same as for half parity gain in Fig. 5.19.

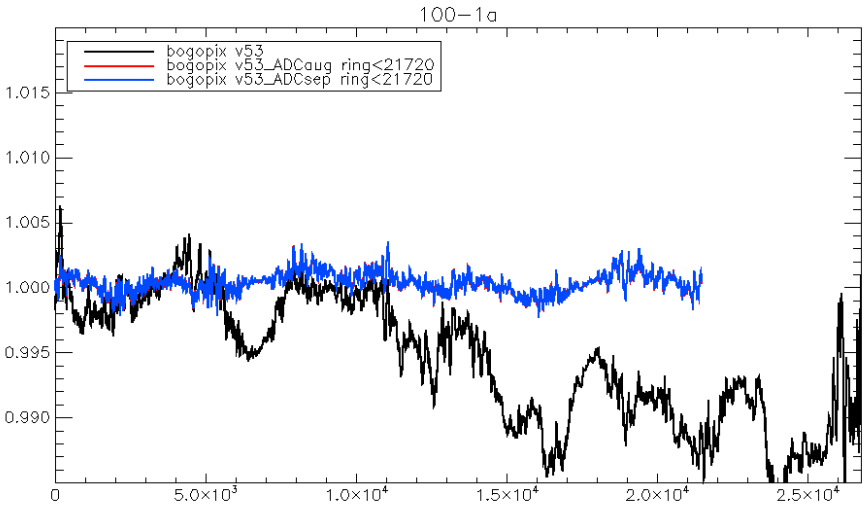


Figure 5.22: Bogopix gains on channel `00_100-1a` calculated from the uncorrected science data (black line), and from science data corrected with two consecutive versions of the ADC nonlinearity correction (red and blue lines). Figure reproduced with the permission of Olivier Perdereau.

The resulting half parity sum observables on real data and on the simulated data are shown in Fig. 5.23, and are discussed in the following sections. All the inputs parameters for ADC nonlinearity, are the same that have been used for the 2013 *Planck* data release.

Nulltest

The first step consist in checking the effect of the forward simulation and the backward correction on the observable. It is done by applying the ADC nonlinearity correction $\mathcal{F}^{-1}$ on the $\mathrm{SKY}_{\mathrm{ADC}}$ PBRs. The resulting observables are the orange line on the right plot, which is not visible because it is flat and near-zero. The red line is for another test related to 4 K lines and will not be discussed here.

This result shows that the 1D drizzling projection for ring folding, interpolation for PBR unrolling to TOI format, and application of $\mathcal{F}$ and $\mathcal{F}^{-1}$, do not introduce any significant discrepancy between the simulated PBRs. A second check has been done by comparing $\mathcal{F}^{-1}(\mathrm{SKY}_{\mathrm{ADC}})$ with $\mathrm{SKY}_{\mathrm{REF}}$, with equivalent results.

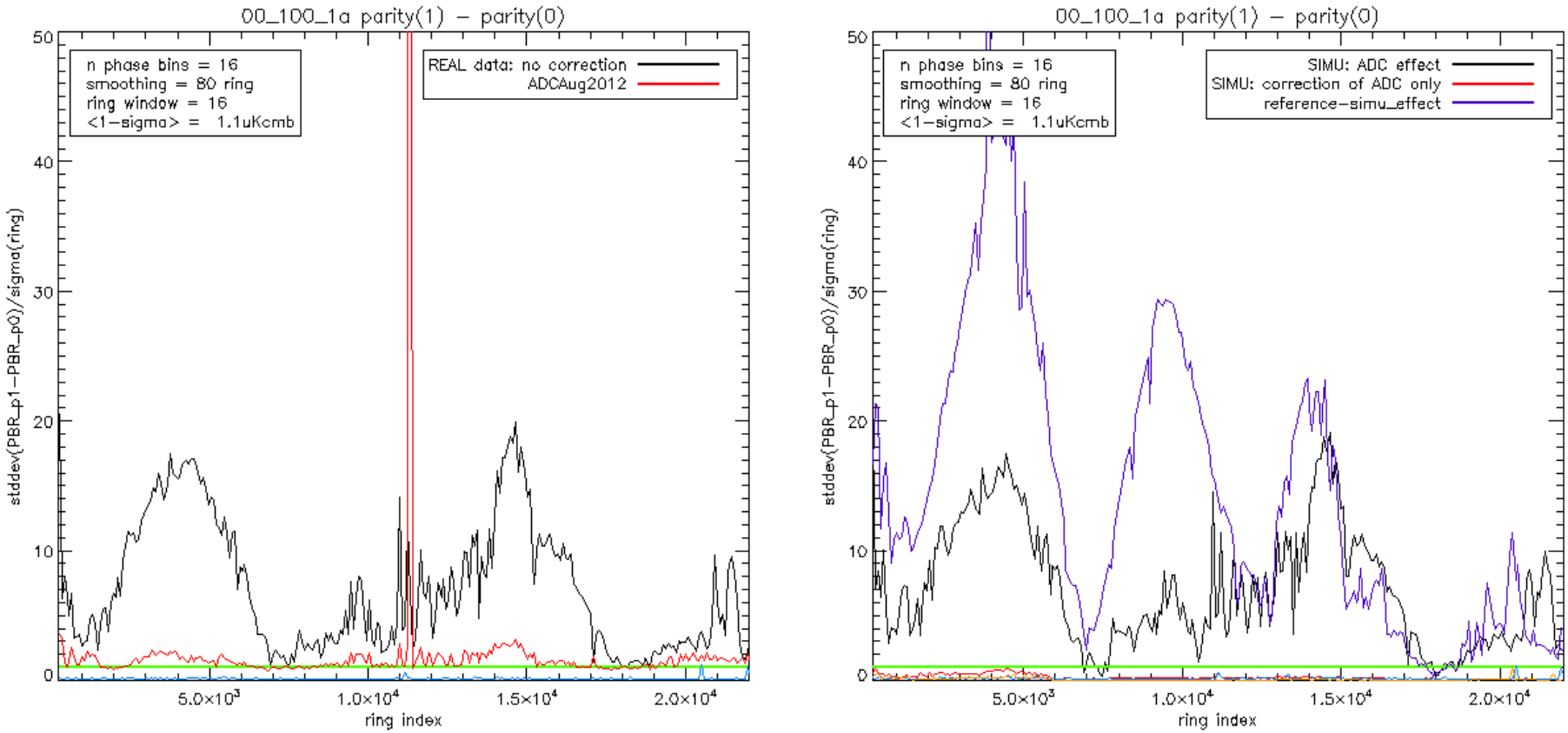


Figure 5.23: Simulation of ADC nonlinearity effect over the full mission, the observables are built with the half parity sum method. The y-axis is normalized with the 1σ estimation for the observable. *Left panel*: (real data for comparison) The black line is the observable before application of the correction, and the red line is the observable after application of the 2013 data release correction. *Right panel*: (simulation) The black line is the observable calculated after application of ADC nonlinearity effect. The orange line which is very flat (as expected) and not visible is for the null-test, i.e after application of the nominal correction on the science data simulated with ADC nonlinearity effect. The red line in this case is for an experimental correction with some nuisance from a small amount of spurious 4 K lines. The green thick line is the average level of noise estimated for the observable.

ADC nonlinearity effect validation

The second step is the validation of the ADC nonlinearity effect introduced with $\mathcal{F}$. In this case the comparison of the observables yielded by the simulated $\mathrm{SKY}_{\mathrm{ADC}}$ (black line on the right) and the raw uncorrected SKY (black line on the left), are in very good agreement. The main difference is the absence of TOI noise in the simulation.

All channels where the 4 K lines level is low, yield a similar result. Hence this test validates the first order model of ADC nonlinearity.

Half parity observable characterization

How much of the ADC nonlinearity residuals are we estimating with the half parity observables? To answer this question, the difference between the reference PBR and the one with ADC nonlinearity effect has been calculated with $d_0 = \mathrm{stddev}(\mathrm{SKY}_{ref0} - \mathrm{SKY}_{ADC0})$, and $d_1 = \mathrm{stddev}(\mathrm{SKY}_{ref1} - \mathrm{SKY}_{ADC1})$. In Fig. 5.23 The purple line is the RMS of d_0 and d_1, it is homogeneous with the half parity sum observable. The observed discrepancy is significantly higher **by a factor two or three** than the observables inferred with parity zero and one.

This result is similar on all the four channels on which the simulation setup has been run. It can be explained by the high level of symmetry of the signal around the ADC central defect. Hence the half parity observables measure only half of the real deformation effect.

This has some important consequences, it means that the ADC nonlinearity *bends* the sky measured from both parities in a similar way. As a consequence, one has to be cautious when taking the average of many detectors as a reference, because the ADC nonlinearity effect is not expected to average out, as it would be for noise.

In practice this very point has not been thoroughly characterized, but raised again in 2015, when the peoples remaining active in the *Planck* Collaboration have been facing the difficulty of measuring reionisation τ biased significantly by large angular scale ADC nonlinearity residuals. To work this issue, the sky maps from all the detectors of a single frequency are averaged to build a reference for the purpose of removing ADC nonlinearity residuals (and others effects).

Frequency domain

The ADC nonlinearity effect in frequency domain is estimated by averaging on the first survey (S1) the DFT modulus of the ADC nonlinearity residuals $R_0 = \mathrm{SKY}_{ref0} - \mathrm{SKY}_{ADC0}$ and $R_1 = \mathrm{SKY}_{ref1} - \mathrm{SKY}_{ADC1}$. Fig. 5.24 shows the quantity $\Big(< |\mathrm{DFT}(R_0)| >_{\mathrm{S1}} + < |\mathrm{DFT}(R_1)| >_{\mathrm{S1}} \Big)/2$. The DFT mode index in the sky phase domain has been converted to a time domain frequency with a 1/60 factor, accounting for the 60 seconds average duration of the PBR.

The ADC nonlinearity residuals from the simulation have globally a $1/f$ slope in the frequency domain, with a steeper fall on the first modes of the CMB channels which is consistent with the main source of signal dynamic on these channels (the solar dipole).

This is an important information, because the ADC nonlinearity effect will impact mainly at large angular scale on the sky. The important scientific objectives that are impacted are:

- The study of the epoch of reionization, and the calculation of τ;
- Gravitational waves detection: this is potentially a major issue because it is one of the primary objectives of the *Planck* mission;
- The low frequency spectrum part of the thermal SZ effect.

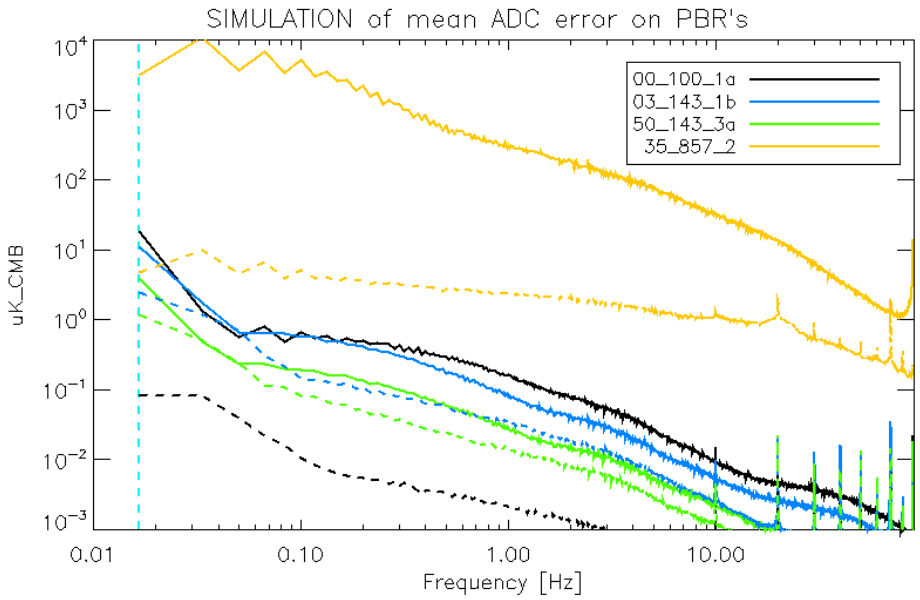


Figure 5.24: Simulation of ADC nonlinearity effect in frequency domain. The thick lines are the Fourier transform modulus, averaged over the first survey, of the difference between the reference PBR and the PBR with ADC nonlinearity effect. The dashed lines are experimental versions comparing the reference PBR with the one including ADC nonlinearity effect plus correction with a nuisance from spurious 4 K lines.

An interesting side effect of the low frequency expression of the ADC nonlinearity, is that it is possible to work with large bins in the phase domain, leading to much smaller PBRs. It has been used PBRs with 64 bins in phase, this format allows to process a full mission in one minute, and provide nice potential for fit procedures and Monte Carlo's. This topic will be discussed further in Chap. 6.

5.6 Conclusions

In this chapter, it has been shown how to correct the *HFI* science data for ADC nonlinearity effect. For this purpose, the bolometers response has been modeled at first order to approximate the analog signal before digitization. A noise model has also be developed to account for the noise effect on the digitization and summation. With these elements, the science data transfer function has been built, and its reciprocal used for the correction of science data. A specific observable of ADC nonlinearity has been studied to infer the quality of the correction on science data, and it has been used with simulations to validate the ADC nonlinearity correction.

The *raw gain* parameter of the first order analog signal model is of particular interest because its shape is in relation with the time constants of the bolometer. It was a work direction that pushed to the writing of Sauvé and Montier (2016). The ADC nonlinearity correction yielded excellent results on science data, but there is some limitations. The results of the PCA analysis used to produce the *raw gain*, its non-symmetry higher than expected, and the analysis of the *fast samples* periods noise were strong indicators that the 4 K lines was a significant component of the analog signal before digitization. Therefore the parasitic signal from the 4 K lines and its impact on the ADC nonlinearity correction had to be characterized. This is covered in details in Chap. 6. Finally, the simulations shows also that the ADC nonlinearity effect has a $1/f$ shape in the frequency domain, so it will be mainly a concern for science at large angular scales. This will be covered in Chap. 7.

Chapter 6

4 K lines

After validating the efficiency of the first ADC nonlinearity corrections, enthusiasm has been quickly cooled by the understanding of the complexity introduced by the 4 K lines, leaked on the detector channels by the 4K stage mechanical cooler. There has been a lots of hints of this problematic already, ranging from the residuals eigenvectors seen in the PCA outputs of the raw gain estimation, to the *raw constant* variability over the mission and, last but not least, their omnipresence on the science data. Their impact on the ADC nonlinearity correction has been definitively acted after the observation that the correction was failing mostly on channels with a high level of 4 K lines.

This parasitic signal adds periodic waves over the analog signal before digitization. It is a problem because its periodicity extends over several *fast samples* periods, making the analog signal drift significantly from the model described in Chap. 5. Hence this has to be taken into account in the analog signal model to build an accurate version of the science data transfer function. However, the summation of *fast samples* in the DPU, destroys not only information by mixing nonlinear ADC codes, but also folds in frequency the numerous harmonics of 4 K lines with frequency higher than the modulation frequency. The summation is again a major issue making the estimation of the unfolded 4 K lines harmonics very challenging. That estimation effectively required much more work than for the steps leading to the initial ADC nonlinearity correction.

Sec. 6.1 introduces the topic by the characterization of the 4 K lines from science data TOIs to the *fast samples* level. Then the modification of the analog signal model and the new expression for the science data transfer function is described in Sec. 6.2, and the the impact of 4 K lines on the science data transfer function is characterized with simulations. The first estimation of the unfolded 4 K lines harmonics is detailed in Sec. 6.3. It has been done with a method based on Lagrange multipliers, this estimation has been used for the *Planck* 2013 data release. However this method presents some limitations that prevents the correction to reach an accuracy level better than $\approx 10^{-3}$ of the solar dipole amplitude. To overcome these limitations a work was started on a new estimation method (see Sect. 6.4) based on a Monte Carlo Markov Chain algorithm. This work has been merged with a similar work from Guillaume Patanchon, to produce an updated model and new unfolded 4 K lines harmonics estimations, allowing to reach an accuracy level close to $\approx 10^{-4}$ of the solar dipole amplitude.

The correction parameters given in Sec. 6.3 and including the 4 K lines in the analog signal model are the ones that have been used for the ADC nonlinearity correction of the *Planck* 2015 data release.

6.1 Characterization of 4 K lines

6.1.1 Origin

The source of 4 K lines is well known and identified as the 4K cooler mechanical compressors of the cryogenic chain of the *Planck* spacecraft. The 4K cooler system is based on the Joule

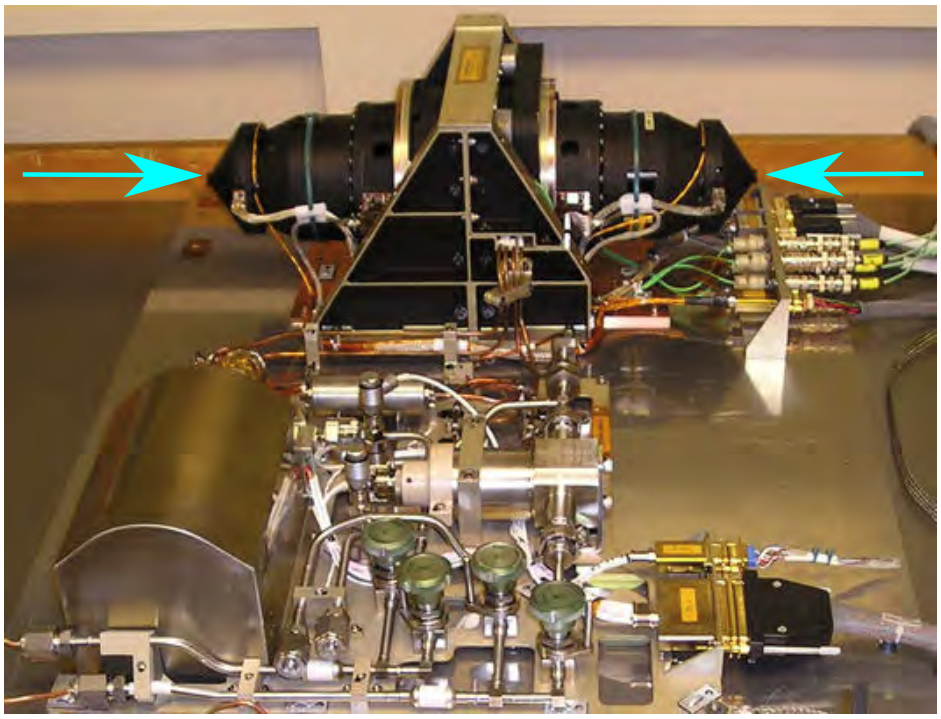


Figure 6.1: The 4K mechanical compressors (aka the root of evil) and gas cleaning equipment, prior to integration to the spacecraft. The cyan arrows points the extremities of the two compressors. Image credit: **?**.

Thompson effect and works with a closed helium circuit, this subsystem is described in details in **?**, and its integration within the spacecraft in **?**. It is designed with the requirement to provide a heat lift of 9 mW at a temperature of 4.2 K. The Fig. 6.1 shows the two mechanical compressors of the 4K cooler mounted in symmetrical positions to cancel transfer of momentum to the spacecraft.

To avoid propagation of microphonic noise that could affect science data in various ways, an active electronic vibration damping system has been added: the Vibration Control System (VCS). The VCS operates on the first seven unfolded harmonics of the base 40 Hz frequency of the compressors (40, 80, 120, 160, 200, 240 and 280 Hz), by injecting an opposite compensation force to the one measured by force transducers. During the ground tests (**?**), some specific microphonic lines could be observed above the detector noise, that have been found to originate from the 4K cooler system, but this weak microphonic contribution disappear when the VCS is activated. However some electromagnetic interferences originating from the electronics of this system still propagates to science data, this is what is called the 4 K lines.

6.1.2 Impact on ADC nonlinearity observables

The impact of 4 K lines on ADC nonlinearity correction can be estimated by looking to the correlation between the 4 K lines level and the ADC nonlinearity observables built from science data over the full *HFI* mission. A global 4 K lines level estimator on a given channel can be built by calculating the RMS value of one 4 K lines harmonic over the mission, and by summing the result obtained with the four harmonics visible on science data. The ADC nonlinearity residual level on a given channel, can be estimated by taking the RMS value of the half parity gain observable minus one over the mission.

The correlation plots between the 4 K lines level and ADC nonlinearity residual level are shown in Fig. 6.2. It is clear by looking at the result after ADC nonlinearity correction that

the 4 K lines introduce a significant bias. If they are not taken into account in the correction model, a clear correlation appears (bottom panel) between the level of ADC nonlinearity observable and the global level of 4 K lines, albeit the ADC nonlinearity correction reduces significantly the residual level. For some channels, the ADC nonlinearity residual level is even higher after correction, this is due to the bias introduced by the 4 K lines to the *raw constant* and will be discussed further using simulations in Sec. 6.2.3.

To understand what's happening quantitatively, let estimate the 4 K lines level on *fast samples*. For a given DFT modulus of 20 DSN, the real amplitude of the sinusoid is twice, i.e. 40 DSN. A rough estimation of the signal before DPU summation translates this value to a 1 LSB sinusoid at *fast samples* level. The exact value is obtained with the DPU summation transfer function given in Appendix H. The order of the parasitic signal amplitude is highly significant compared to the ADC central defect of one LSB. In practice, there is no *HFI* channel which is free from the 4 K lines parasitic signal. Hence it is mandatory to take this systematic into account for the ADC nonlinearity correction.

6.1.3 4 K lines on science data

Spectral distribution on TOIs

The 4 K lines signal on science data is generally below the noise level and cannot be seen by direct observation, but shows up as four **sharp lines** in frequency domain. An example of DFT of the raw demodulated science data is shown in Fig. 6.3. After deglitching, performed by replacing glitch flags with phase interpolated PBR of the sky, the spectra of the demodulated signal shows the 4 K lines at the apparent frequencies of 10 Hz, 30 Hz, 50 Hz and 70 Hz.

These lines are the folded versions of the analog signal before digitization and demodulation, the effective frequencies on science data before demodulation are 20 Hz**,** 40 Hz**,** 60 Hz **and** 80 Hz. In practice, the DPU summation folds all the harmonics present on *fast samples* above the modulation frequency $f_{\mathrm{mod}} \approx 90Hz$. The DPU summation effect is detailed in Appendix H. From now, when dealing with science data, it will be referred only to the frequencies on the modulated signal.

Estimation of 4 K lines with the 18 samples chunks method

On science data, 4 K lines can be described with one complex Fourier coefficient for each of the four folded harmonics frequencies, with DFT based methods. In the *HFI* DPC pipeline, the `cutfreq` software performs this task with the goal of cleaning science data from 4 K lines. However this software cannot be used in the ADC nonlinearity correction context for several reasons:

- `cutfreq` runs on demodulated and time deconvolved science data in Watts, providing outputs incompatible with the raw modulated science data in DSN which is mandatory for the ADC nonlinearity analysis;
- `cutfreq` suffers from a design flaw because it replaces glitch flagged data by interpolated PBR which is not cleaned from 4 K lines and re-injects spurious 4 K lines power in the final result. It is incompatible with the purpose of estimating unfolded harmonics for a correction, and also builds spurious wings around 4 K lines folded harmonics on the signal spectra. The re-injection problem is shown in Fig. 6.4.

To provide estimations of 4 K lines compatible with the objective of using them as constraints for the ADC nonlinearity correction, it was used chunk of data which size is a multiple of the 4 K lines periodicity (9 samples) and the modulation periodicity (2 samples). The algorithm works as follows:

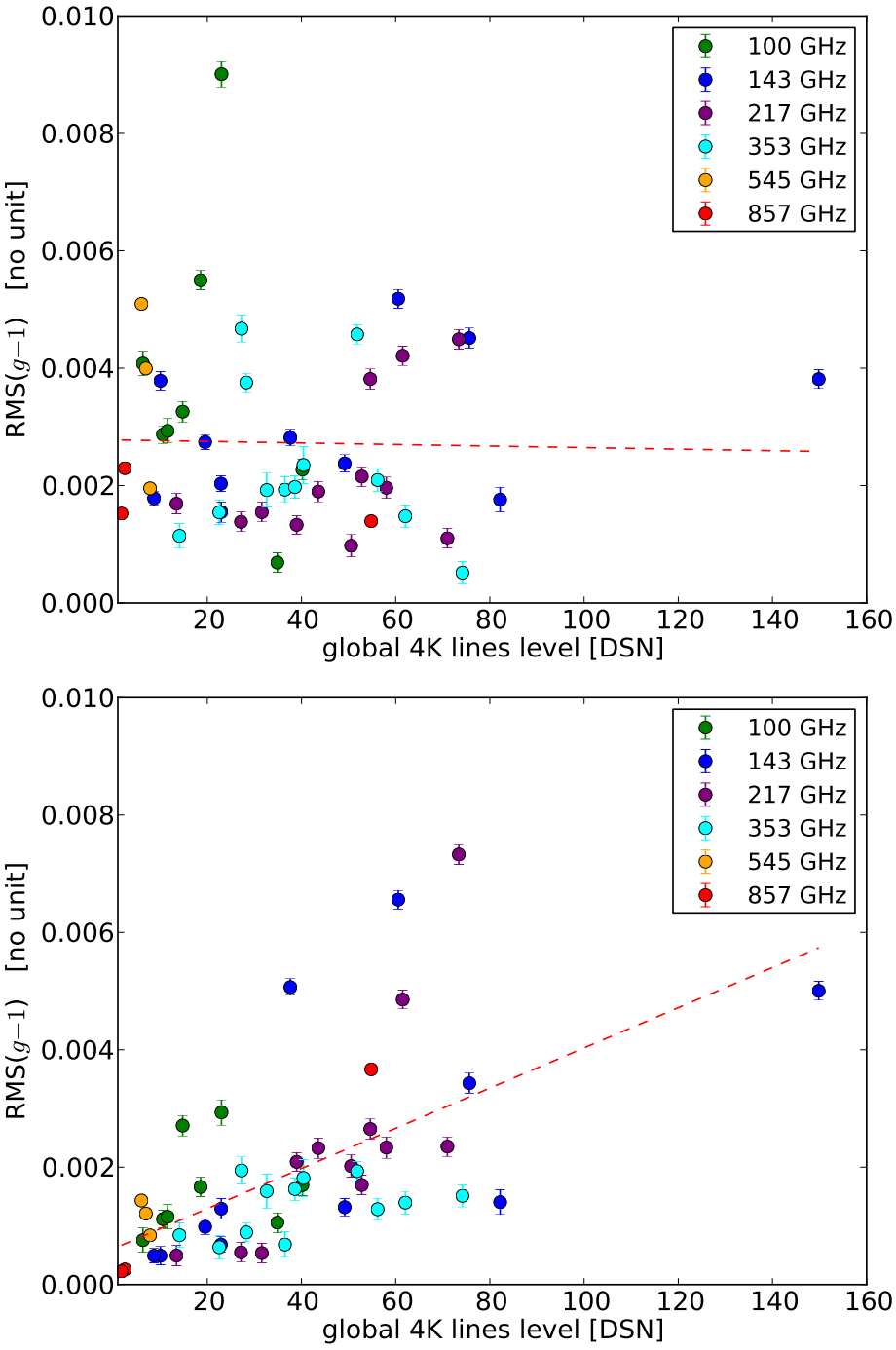


Figure 6.2: Correlation plots between the 4 K lines level and ADC nonlinearity observables (half parity gain), for the 50 channels where a correction is available. *Upper panel*: raw science data. Error bars are the one σ extent for the 128 ring bins used to build observables. The red dashed line shows a straight line fit. *Bottom panel*: ADC nonlinearity corrected data using the best available INL (`GP41`) but without any handling of 4 K lines. Same legend as for upper panel.

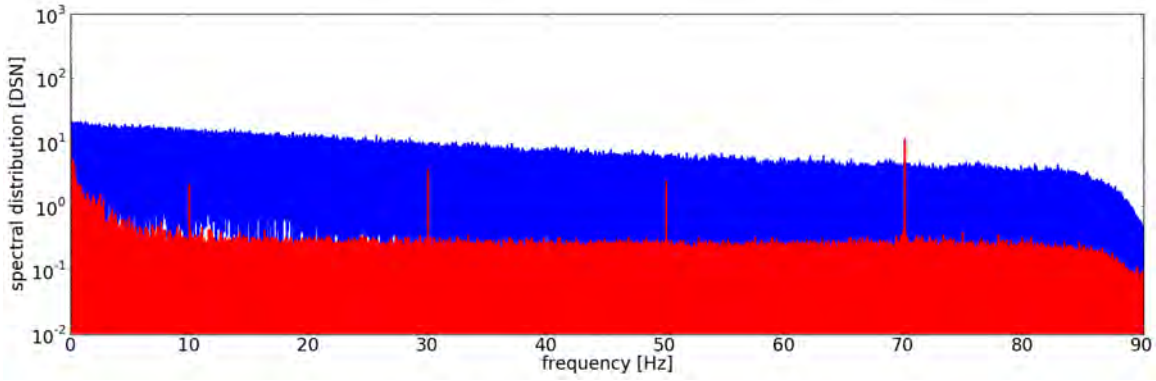


Figure 6.3: 4 K lines on science data TOI spectra for ring 10000 on channel `02_143-1a`. The blue line is the DFT of raw demodulated science data. The red line is the DFT of deglitched demodulated science data.

- TOI MAPPING TO CHUNKS: the science data TOI is mapped to 18 samples using the TOI index modulo 18. The TOI index zero is the reference as the first sample of the fist chunk;
- CHUNKS FLAGGING: all chunks containing at least one invalid sample are discarded (glitch, unstable pointing flag). In addition to the invalid data filtering, chunks with bright signal are also discarded to avoid leaks of signal derivative in the estimation; the brightness of one sample is determined if the corresponding PBR location is twice larger than the standard deviation of the PBR.
- CALCULATION OF GLOBAL PERIOD: all valid chunks of 18 samples are averaged to produce one single chunk containing two 4 K lines periods plus the modulation signal;
- ESTIMATION OF FOLDED HARMONICS: a DFT is run on the global period to get 18 complex Fourier coefficients. The 4 K lines Fourier coefficients are in increasing frequency order at indices (2, 4, 6, 8). Using 18 samples provides in addition four Fourier coefficients used as bias estimators at indices (1, 3, 5, 7), where nothing is expected but noise.

An example of the resulting DFT coefficients is shown in Fig. 6.5. The folded 20 Hz line is typically the strongest as it can be seen in the time domain period on which it correspond to the dominant sinusoidal shape.

This methods averages out the dipole signal, leading to a direct measurement of 4 K lines folded harmonics. In practice, the bias estimators values are consistent with the noise level of about 0.2 DSN all over the mission. Simulations built from a reference PBR and interpolated science data TOIs plus noise, injected 4 K lines, and using real data glitch flags, gave results consistent with the expected noise level. This validates the estimation method.

The ADC nonlinearity effect could be a concern for the estimated 4 K lines coefficient. In practice it is not the case because its effect is only about 1% of the folded harmonics amplitude. This level is completely insignificant compared to the noise level of the estimation. It is also a second order term in the context of the ADC nonlinearity correction.

Drift over the mission

Using the previous estimation, the Fig. 6.6 show the variation in amplitude and phase of the 4 K lines folded harmonics throughout mission for one typical channel. Their amplitude

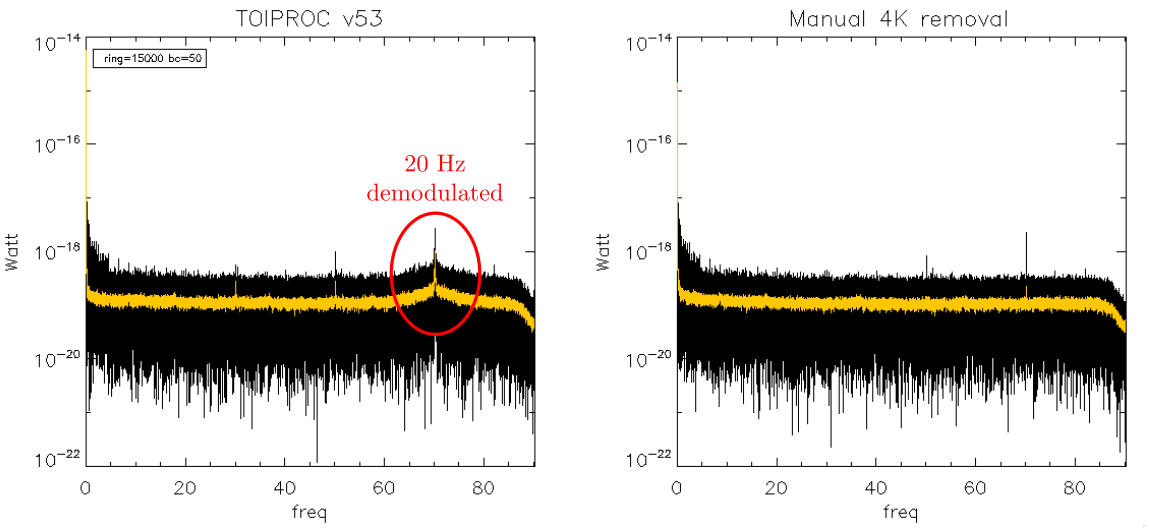


Figure 6.4: Difference between the 4 K lines removal by `cutfreq` and a manual version at ring 10000 of channel `50_143-3a`. Left plot show the demodulated science data DFT after processing by `cutfreq` (20 Hz line appears at 70 Hz), where the black line is the raw DFT and yellow line is the DFT smoothed with a light bowcar average. Right plot is an early test of manual 4 K lines removal taking care of not re-injecting 4 K lines power from PBR.

drift slowly in time, with a general increasing trend over the mission for the 20 Hz channel. Curiously, the three other frequencies do not show the same increasing trend. No valid explanation has been found for this behavior. The phase also changes slowly with several discrete jumps, which may be attributed to the discrete nature of the VCS compensation. There is a specific event at ring 544, when the *HFI* DPU had to be reboot, at this moment the synchronization with the 4K cooler has been lost, leading to a visible phase jump. However, the analog signal did not change really.

At the beginning of the mission, the telemetry transponder of the spacecraft was periodically set on for transmission purposes. It had an unexpectedly high impact on the thermal behavior of the 4K cooling stage visible on the 4 K lines level of Fig. 6.6 over the first ≈ 5000 rings. The 4 K lines amplitude is modulated by the periods of activity of the transponder. To avoid the modulation effect and potential harm to science data, it has been set on permanently after ring 5252 (beginning of full sky survey 2).

To get an overview of the full picture, the 4 K lines amplitudes are displayed for all channels with the same y axis scale in Fig. 6.7. The 20 Hz line is clearly dominant on most channels, this fact is important because it will allow us to make some simplifications when estimating 4 K lines folded harmonics on analog signal for ADC nonlinearity correction. There is a general pattern with an increasing trend for the 20 Hz line. This pattern looks like a generic template that can be used to describe the 4 K lines signal. However there is a wide variety of significant local deviations that will be explored in details in Sec. 6.1.3.

The 545 and 857 GHz channels typically have a very low level of 4 K lines, this can be attributed to the 1/3 gain factor set in the REU for these channels. Hence the probable location of injection of the 4 K lines signal is before application of the signal gain factor, which makes the part of the electronic readout chain including the JFET box the most probable culprit. The injection cannot occur at detector level, because they would be heated by that and the observed 4 K lines frequencies are incompatible with the expected signal product between the 90 Hz modulation harmonics and the base 4K cooler harmonic at 40 Hz.

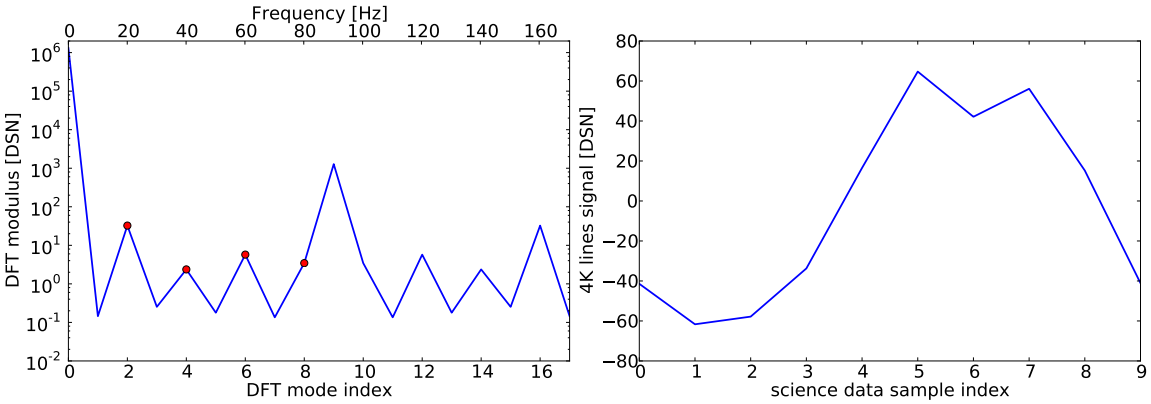


Figure 6.5: Estimation of 4 K lines with the 18 samples chunks method for the ring 20000 of channel `02_143-1a`. Left plot is the DFT modulus of the global period, the red dots marks the locations of the 4 K lines Fourier coefficients. Right plot is the corresponding time domain 4 K lines period on science data.

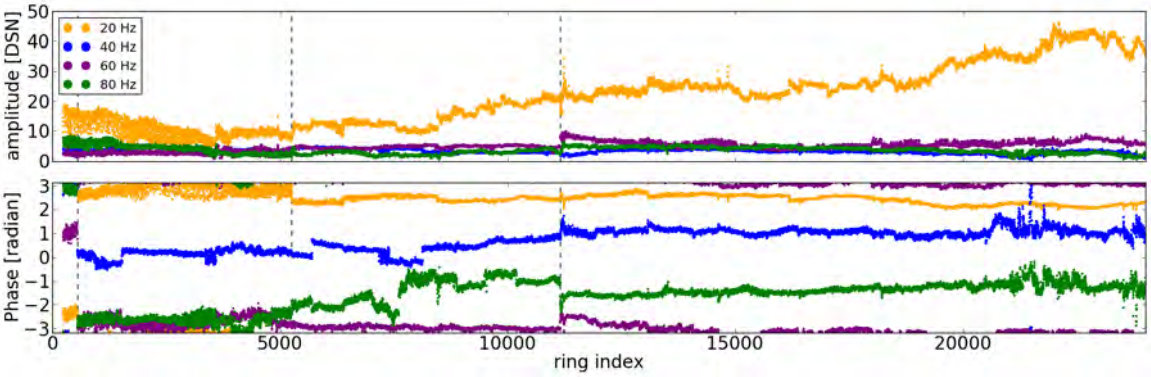


Figure 6.6: 4 K lines drift throughout the mission for channel `02_143-1a`. The upper plot is the DFT modulo of 4 K lines estimated on a per ring basis. The bottom plot is the DFT coefficients phase, relative to science data TOI index zero. Black vertical dashed lines indicate the DPU reboot at ring 544, the setting of the transponder always on at ring 5252, and the sorption cooler switchover at ring 11149.

Sub-ring variations

To investigate the stability of 4 K lines at sub-ring level, the 18 samples chunks method has been applied at sky circle level (10800 samples, i.e. one spacecraft rotation over the sky), this is the minimal usable time scale where the sky signal averaging approximation holds. This timescale of one minute increases the time resolution by a factor 35 on average, compared to the ring version that is the typical timescale in the *HFI* DPC pipeline. However doing so increases significantly the noise level to about 1.4 DSN, which is too high to distinguish small variations of noise. The noise problem can still be solved by averaging the measurements for all channels, as the purpose here is to do qualitative analysis. The final noise level on the averaged 4 K lines estimations at circle level is about 0.2 DSN, which is equivalent to the single channel estimation at ring level.

The Fig. 6.8 shows the averaged 4 K lines estimations, in a range of rings where the transponder on/off cycling modulates the 4 K lines amplitude. The 24 hour cycling of the transponder is most visible on the 20 Hz line, but also on the 60 Hz and 80 Hz lines. The 60 Hz line shows a negative correlation in the modulation shape, which can be explained by the folding of several *fast samples* lines on the science data at the same frequency. Thus the

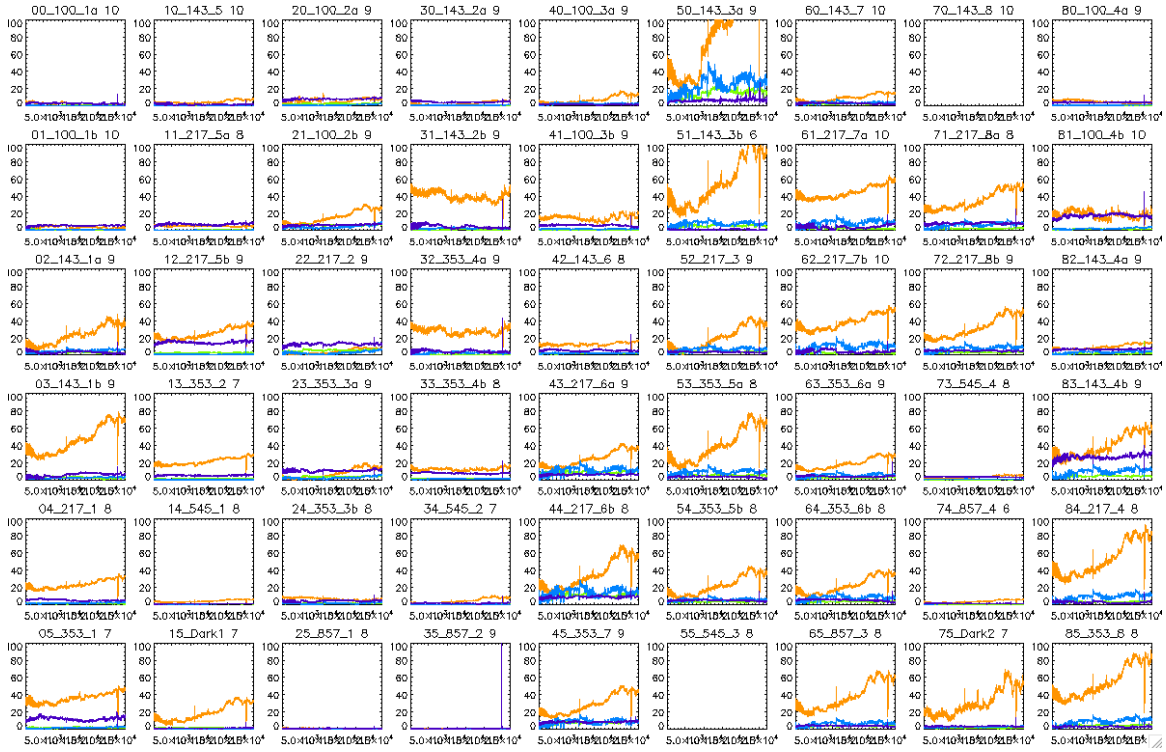


Figure 6.7: 4 K lines amplitude drift in time on science data for all channels. The column correspond to the electronic belt index from 0 to 8, and the row correspond to the channel index from 0 to 5. The color code for folded harmonic frequencies is the same as for Fig. 6.6.

phase distribution and relative amplitude modulation of these lines can produce a change of sign. There is still a noisy feature at all frequencies that disappears when averaging the estimations at ring level. Albeit this feature is above noise level it looks like noise, but may also be residual sky signal feature.

To understand the origin of the high level of subring variations, let look at the averaged 4 K lines estimations further in the mission in Fig. 6.9, two interesting features can be observed now. The first one is stringent on the 20 Hz line where a 20 minutes periodic modulation is visible, it is about 10% of the 20 Hz line amplitude. It is probably the same feature visible in the transponder modulated area, but the modulation here presents a higher amplitude. This aspect will be discussed in Sec. 6.1.3. The second feature is discrete level jumps on the 60 Hz line, which are distinct from the periodic local spikes present at approximately one ring interval. This has been called 60 Hz *popcorn*, in reference to the discrete level science data signal jumps visible on several channel called *popcorn* (but of supposedly different nature). The 60 Hz *popcorn* has been observed and confirmed by the Grenoble team in charge of the 4 K lines removal with `cutfreq`.

This analysis demonstrates clearly that 4 K lines exhibit a significant dynamic at lower timescale than the ring level. To go one step further, it is usefull to look at the frequency spectra of the 4 K lines amplitudes. This processing is possible because the 18 samples chunks method is robust against the spacecraft unstable pointing periods areas which otherwise are flagged out for sky maps production. Still a few sky circles are missing due to complete absence of data, the 4 K lines amplitudes at these locations are set to zero. The result, shown in Fig. 6.10 presents a 1/f shape with a noise plateau reached only for the 40 Hz and the 80 Hz lines. It is clearly noticeable that there are a lot of lines at shorter time scale than the typical 45 minutes ring duration. Most of these have been attributed to the sorption cooler bed switching which is discussed in Sec. 6.1.3.

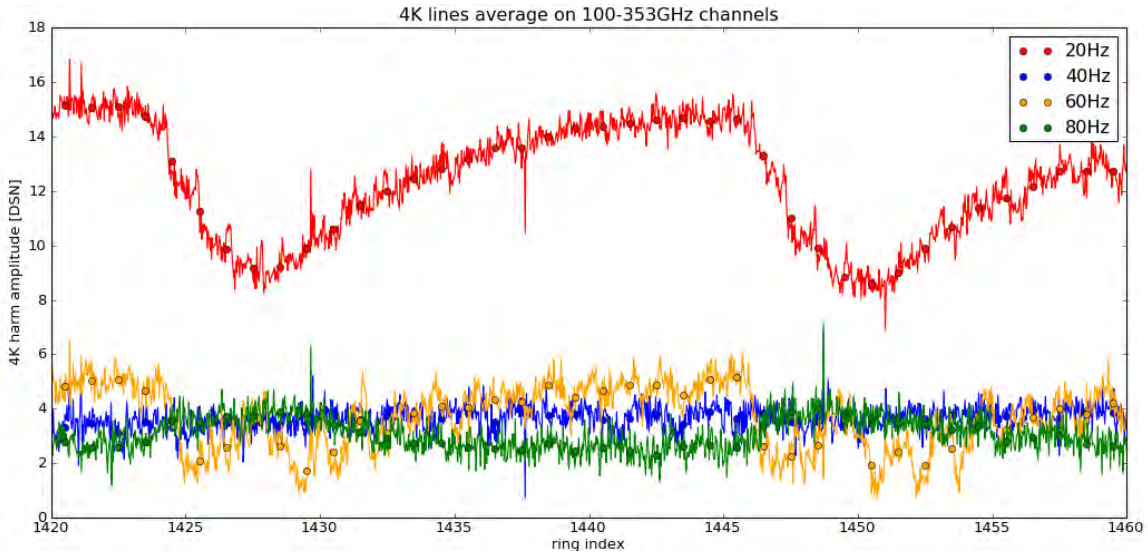


Figure 6.8: 4 K lines at sky circle time scale averaged on all channels in the ring range 1420 to 1460. The colored lines correspond to the one circle time scale, while the dots with the same color are the ring averaged versions.

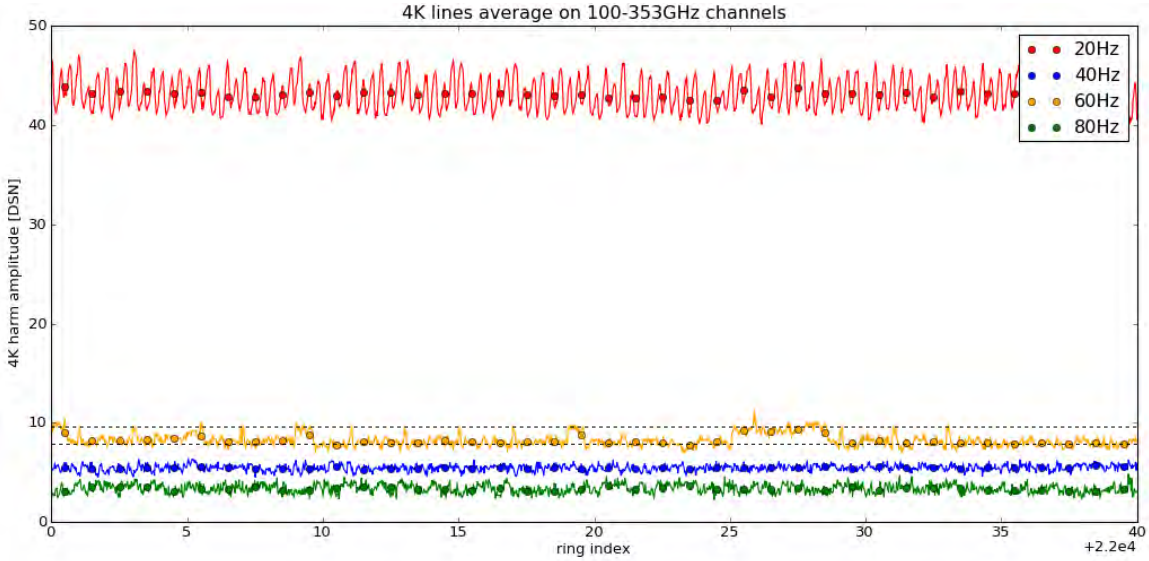


Figure 6.9: 4 K lines at sky circle time scale averaged on all channels in the ring range 22000 to 22040. Same legend as for Fig. 6.8. The horizontal dashed lines are guides to the eyes showing the two discrete levels of 60 Hz *popcorn*.

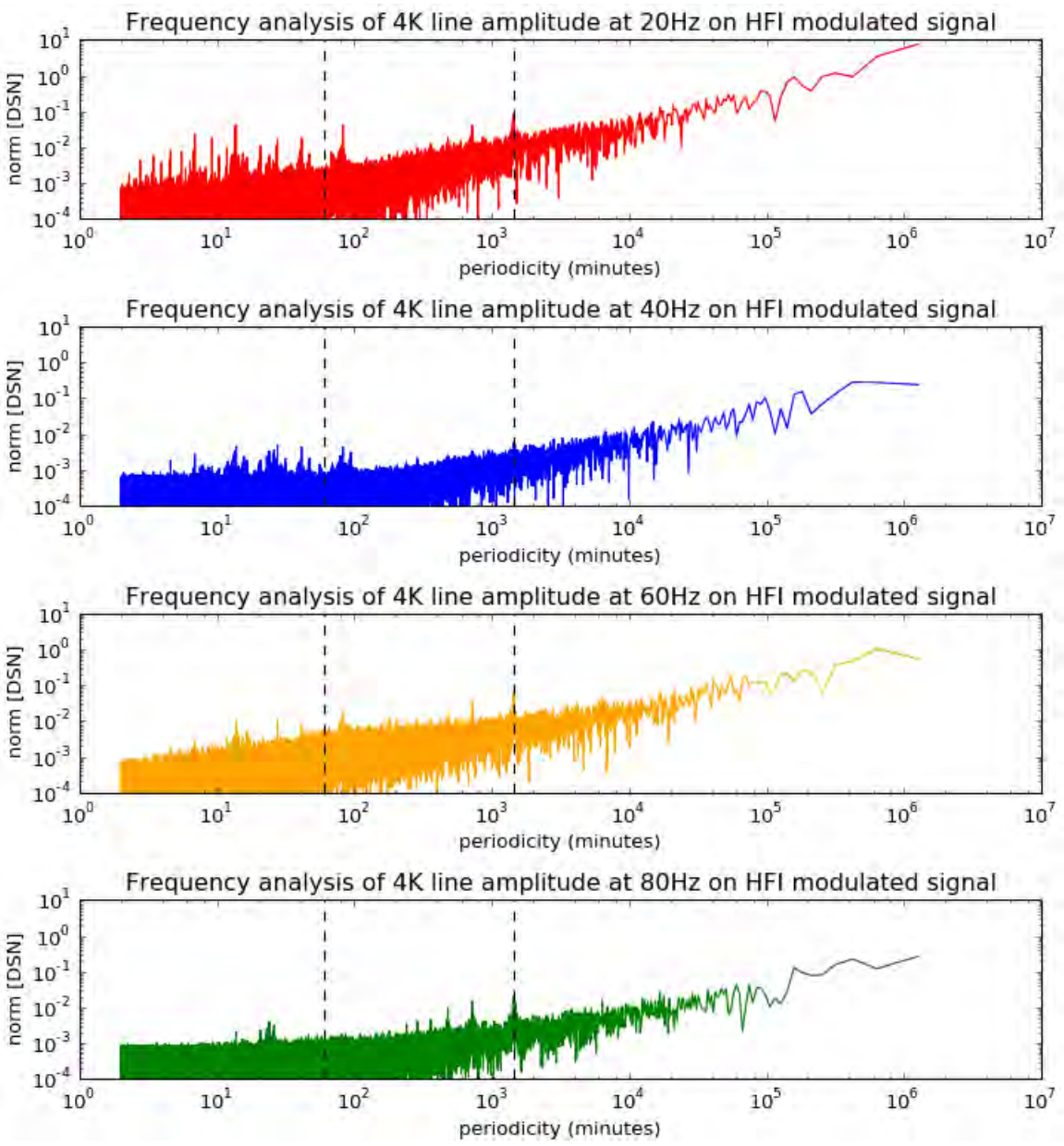


Figure 6.10: DFT of 4 K lines amplitudes all over the mission at one minute time resolution from the all channels averaged estimation. The X axis is the **period** of the DFT mode. Two vertical dashed lines indicate the one hour and 24 hours locations.

Cross channel correlations

As stated before, the 4 K lines variations seem to be the same on many channels (see Fig. 6.7). This behavior can be quantified by inspection of the Pearson's correlation coefficient between the 4 K lines estimations of two channels. However there are several jumps in the 4 K lines level which are associated to single channels and artificially decrease the correlation coefficient. Thus to overcome the jumps issue, a separate computation has been performed on each sky survey.

The data used for the correlation is the 4 K lines folded harmonic Fourier coefficient estimated at ring level. The Pearson's correlation coefficient is run separately on the modulus and the phase, and for each folded frequency. The results are presented in matrix format in Fig. 6.11, 6.12, 6.13 and 6.14 for survey 1 to 4, respectively. Each matrix is organized as follows: the bottom left pixel is the first channel of belt 0, and pixels follows the channels numbering in ascending order. There are 12 belts presented here as the 4 K lines folded harmonics can also be estimated on electronic belts 9, A, B. The three last belts provide valuable information on the propagation of this spurious signal because they are not bolometer detectors.

A clear geographic localization of the signal do appear around belt five for the 60 Hz folded harmonic, which is the frequency with the highest level of correlation between channels. The correlation result is confirmed by **?** with a similar method (Using the demodulated signal for which the folded harmonic frequency is 30 Hz). Unfortunately for the ADC nonlinearity correction, the 20 Hz line which is the strongest and most impacting, shows much fewer correlation. There is in general a wide dispersion between channels and depending also on the period of time of the analysis. These results tend to indicate that the 4 K lines variation have to be estimated on a per channel basis, which renders the work on that effect more difficult because a global template does not show up.

Modulation by the sorption cooler bed switching

The 4K cooler helium flow is pre-cooled at 18 K by the *LFI* 20 K sorption cooler. The *LFI* 20 K sorption cooler is described in details in **?**, and its integration within *LFI* in **?**. it is based on the Joule-Thompson effect and uses a closed hydrogen circuit. The compression of helium occurs by heating one by one each of the six sorption beds which desorb helium at high pressure, a schematic of the system within the spacecraft is shown in Fig. 6.15. The sorption cooler has two basic periods: the first related to the individual bed switching (varying between $\tau_{\rm bed} \approx 1000$ s at the beginning of the cold mission, down to $\tau_{\rm bed} \approx 500$ s at the EOL); the second related to the whole cooler cycle which is $6 \times \tau_{\rm bed}$. The sorption cooler operations modulate the temperature at the pre-cooling junction with *HFI* with variations of order $\approx$ 1K (**?**).

To see the effect of bed switching on the 4 K lines signal, the the sorption cooler housekeepings were investigated. These housekeepings are available in the *HFI* DMC groups `Hs_GRP-1664` at the rate of one sample per 10 s. The one minute timescale 4 K lines averaged estimations were used to compare their variations to the sorption cooler housekeepings, and found a very good correlation (≈ 0.8 for the Pearson correlation coefficient) in the 22000 to 22040 ring range with the pressure sensor labelled `SM050540` on Fig. 6.15. An overplot of the sensor and the 4 K lines averaged estimations is shown in Fig. 6.16, showing the high correlation between 4 K lines modulation and the sorption cooler operations. The related schematics and list of sensors are provided in Appendix F.

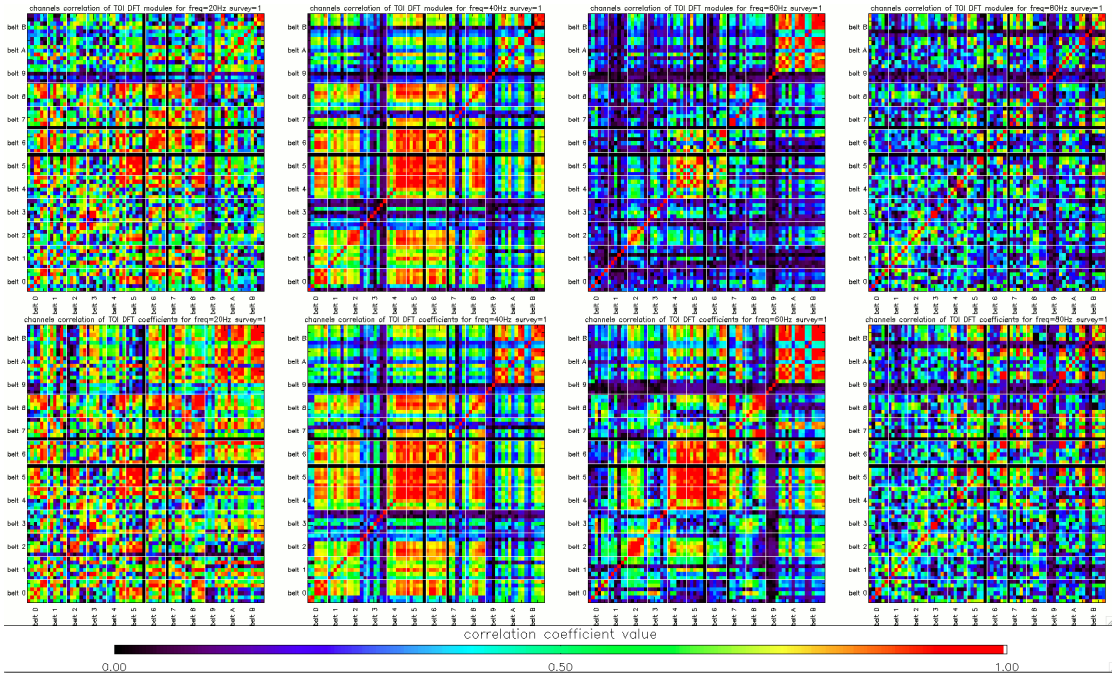


Figure 6.11: 4 K lines 72x72 correlation matrices for survey 1 displaying the Pearson coefficient calculated from the variations of amplitude or phase between two channels. Each matrix is displayed with the channel `00` at the bottom left position, and the channels are sorted by increasing `belt-channel` identifier value. *Upper row*: amplitude. *Bottom row*: phase. *From left to right*: 20, 40, 60, 80 Hz lines.

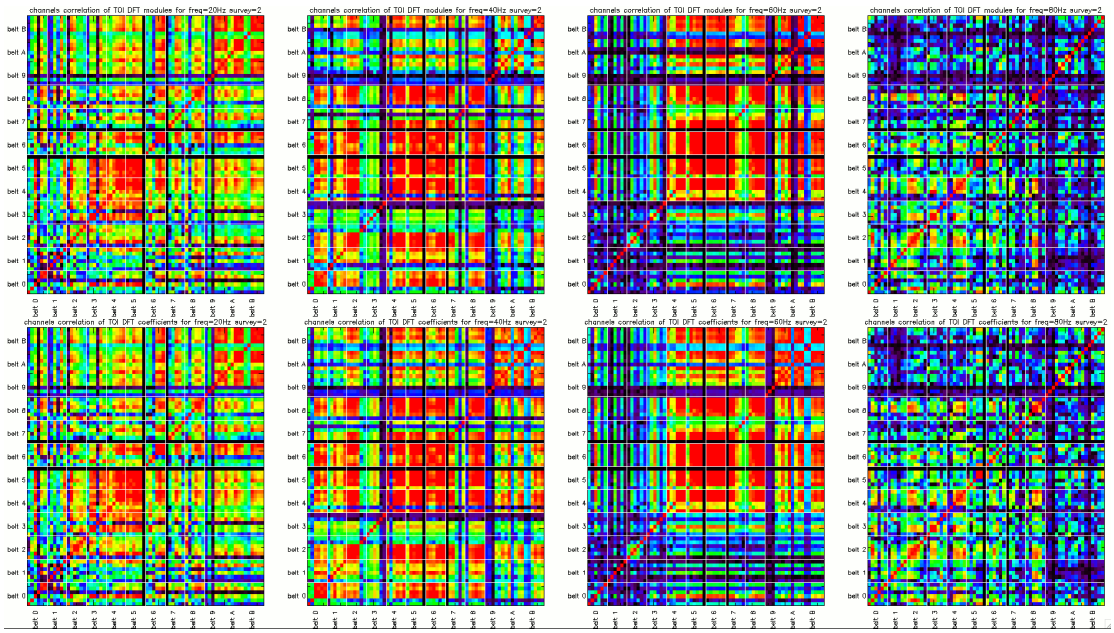


Figure 6.12: 4 K lines correlation matrices for survey 2. The presentation is the same as for Fig. 6.11.

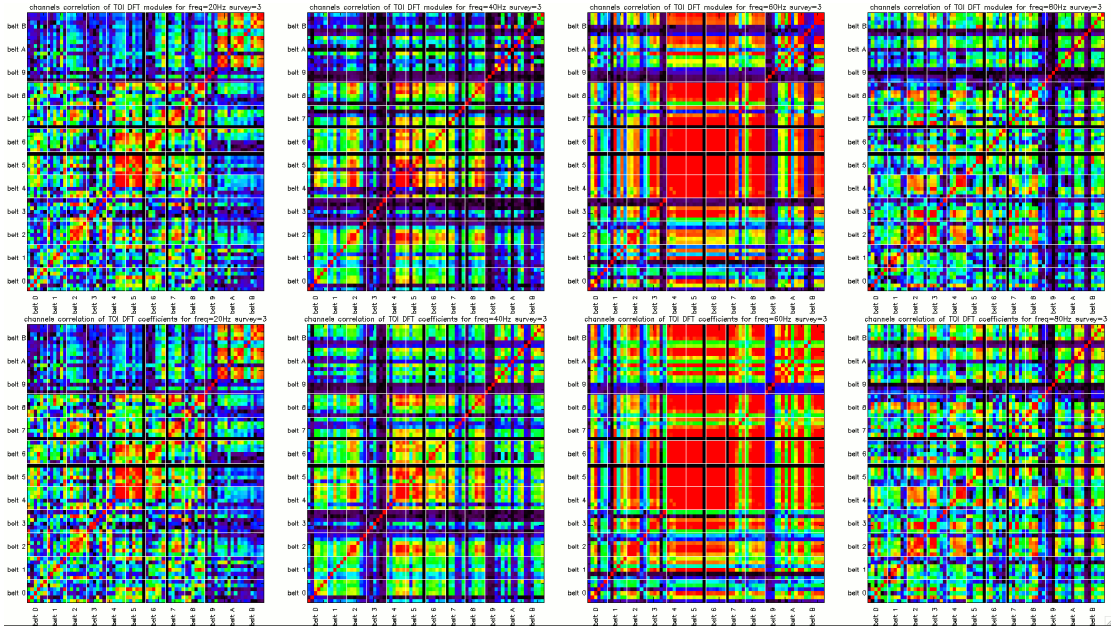


Figure 6.13: 4 K lines correlation matrices for survey 3 The presentation is the same as for Fig. 6.11.

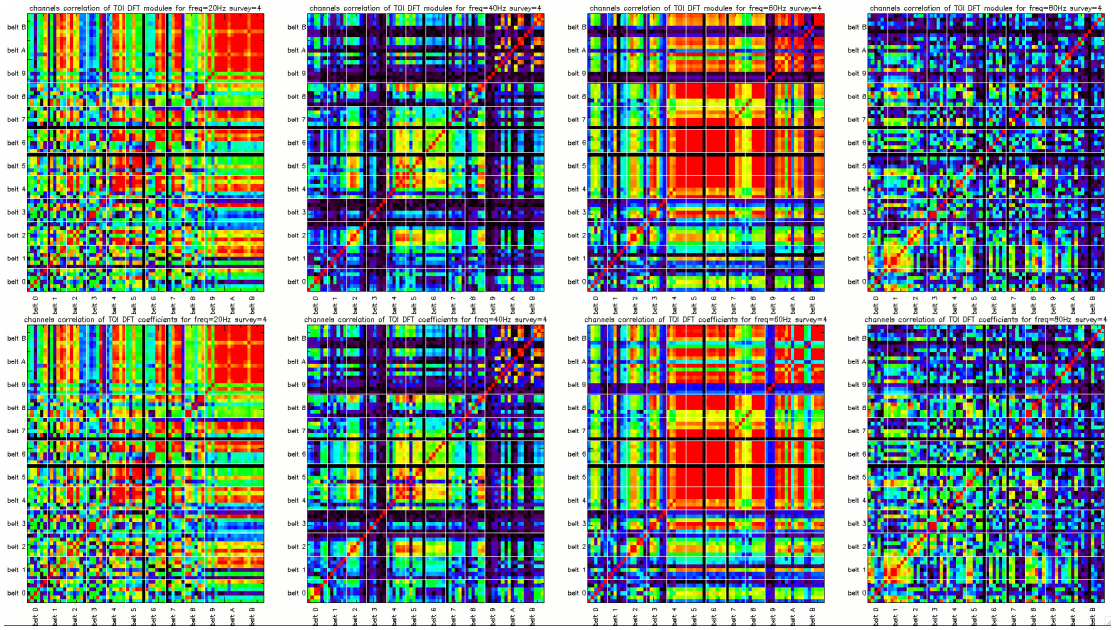


Figure 6.14: 4 K lines correlation matrices for survey 4 The presentation is the same as for Fig. 6.11.

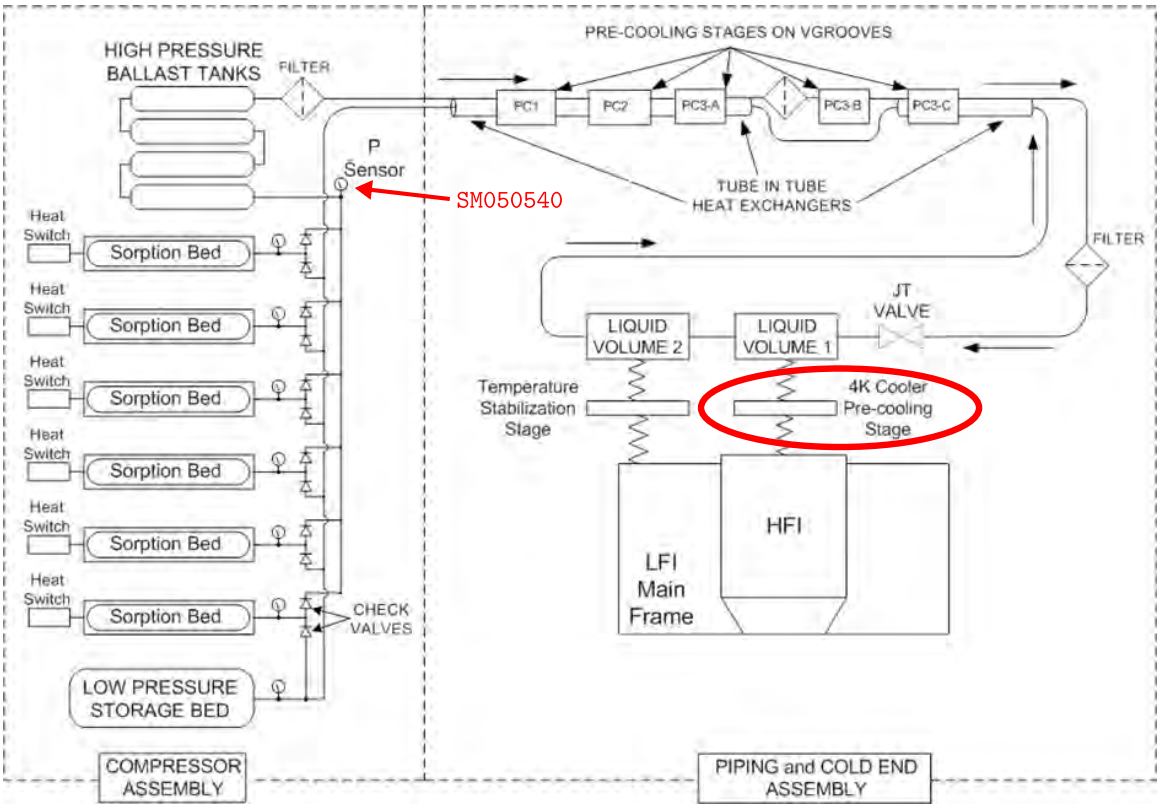


Figure 6.15: Schematic of the *LFI* 20 K sorption cooler. The location where the *HFI* pre-cooling occurs is circled in red, and the P7 pressure sensor labelled `SM050540` is pointed by a red arrow. Image credit: **?**

The consistency of the time variations of τ_{bed} was checked with the 4 K lines one minute timescale estimations after a discussion on this topic with Charles Lawrence[1]. The algorithm used is rather rough: DFT of 4 K lines variations are calculated on a one week basis, then the first mode with modulus above 3 times the standard deviation of the DFT modulus is selected. The Fig. 6.17 shows the extracted period values in minutes from the weekly analysis. These results are perfectly in line with the expected variations of τ_{bed} which is slowly tuned down as the sorption beds ages with a reset at the moment of the sorption cooler switchover with the redundant unit. A few outliers are supposed to be caused by the roughness of the method.

The global impact on the ADC nonlinearity correction of the 4 K lines modulation by the sorption cooler operation is difficult to estimate. There is a modulation at the *fast samples* level of the 4 K lines unfolded harmonics, thus modifying the shape of the *raw constant* parameter at τ_{bed} timescale. Even if it is possible to estimate accurately the 4 K lines folded harmonics at ring level, the unfolded harmonics modulated by the sorption cooler are a different and much more complex problematic. For the time being, the *HFI* ADC nonlinearity correction does not take into account these variations, but they may be a limiting factor for the accuracy of the correction. Because of others priorities, lack of time and resources, no simulation of this effect has been done, but it **still needs to be quantified**.

[1] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, U.S.A.

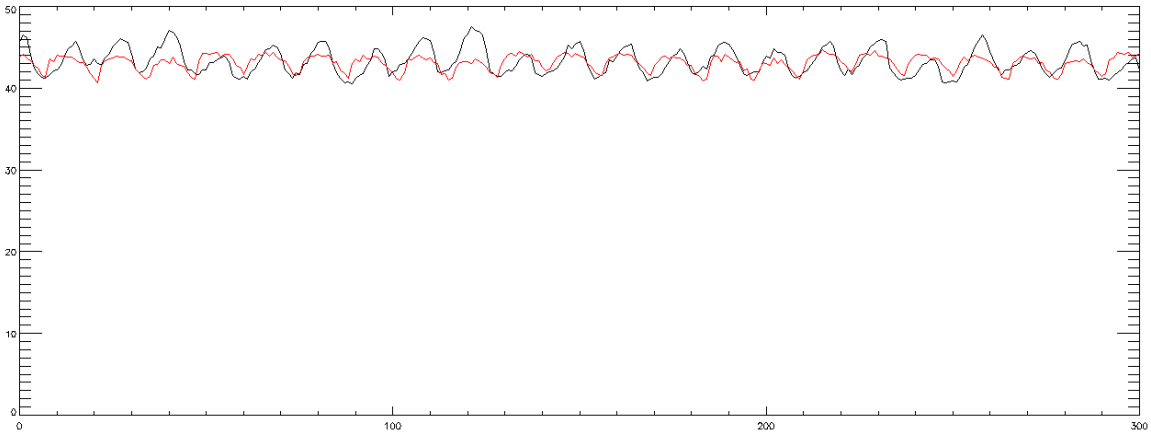


Figure 6.16: Comparison of sorption cooler helium pressure and 4 K lines variations in the ring range 22000 to 22040. The black line is the sorption cooler pressure housekeeping `SM050540`. The red line is the scaled 20 Hz all channels averaged estimation at one minute timescale.

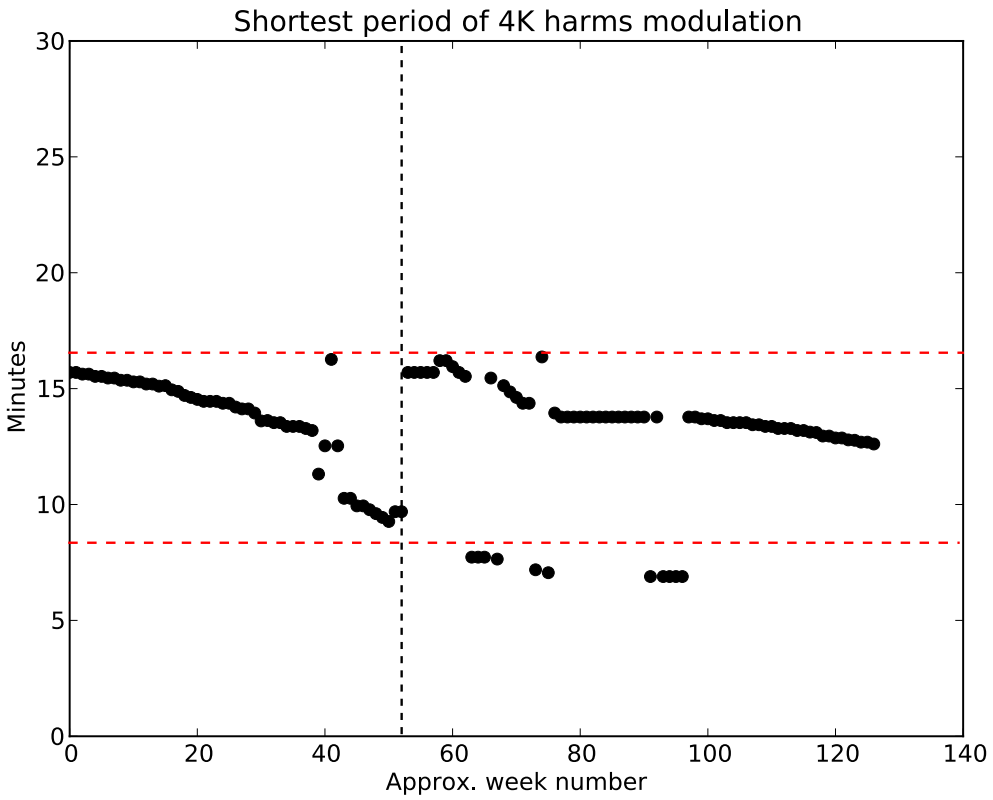


Figure 6.17: Estimated periodicity of the 4 K lines shortest modulation. The horizontal red dashed lines indicates the 1000 s and 500 s sorption cooler beds switching period at the beginning and the end of *HFI* mission. The horizontal black dashes indicates the time of the sorption cooler switchover.

Sequence	Begin	End
CPV 1	2009/07/03 22:00:00	2009/07/04 23:59:59
CPV 2	2009/07/12 00:00:00	2009/07/12 23:59:59

Table 6.1: dates of the *fast samples* CPV sequences

6.1.4 4 K lines on *fast samples*

The 4 K lines window problematic

For the ADC nonlinearity correction, the knowledge of the full period for the unfolded 4 K lines is necessary. Its inclusion in the signal model will be detailed in Sec. 6.2, but there is a big issue preventing to access the full period all over the mission.

The 4K cooler system is phase-synced with the DPU such as the base 40 Hz frequency of the two compressors reads $f_{4K} = 4f_{mod}/9$. This design choice makes the 4 K lines removal on science data easier. However the *fast samples* period capture occurs every 101 seconds (see Sec. 1.2.6), which is in practice every 18288 science data samples. The problem comes from the fact that $18288 = 127 \times 18 \times 8$, is a multiple of 18. As a consequence, the *fast samples* period is phase locked to the same exact window viewing only 2/9 of the unfolded 4 K lines period. Hence the signal estimation of the full 4 K lines period cannot be made from direct observation: end of the may-have-been-short story!

It has been debated before the mission if the period capture would be always the same window or not. May the phase locking have been be relaxed there would have been some other issues:

- the 4 K lines signal mixes with sky power, it would have impacted the estimation of the *raw gain*;
- the one minute time scale variability (see Sec. 6.1.3) of 4 K lines would have rendered the reconstruction of the full period from the non causal chunks of 2/9 of the full period very difficult.

The CPV *fast samples* sequences

In the CPV phase, three dedicated sequences have been run in the *fast samples* telemetry mode, already described in Sec. 4.7.1. These sequences ran much more smoothly that those executed at EOL, and the two first, labelled CPV 1 and CPV 2, yielded data usable for the unfolded 4 K lines analysis. The great advantage of these sequences is that they provide several minutes of contiguous *fast samples* acquisition, which allows us to make direct estimation of 4 K lines with DFT based methods.

The CPV 1 sequence provided on average 1 minute of contiguous data per channel, and the CPV 2 sequence had better stability providing on average 5 minutes of data per channel. The date of the acquisition of both sequences is given in Tab. 6.1. An example of signal from the second Dark bolometer is shown in Fig. 6.18, the signal dynamic is high because at this moment the compensation square is not optimized. After removal of the average modulation period the unfolded 4 K lines signal shows up, and a glitch is visible after index 2400. Using this dark bolometer channel has the visual advantage that the modulation signal is constant because of the absence of sky power, thus making easier to check the unfolded 4 K lines signal by eye.

Estimation of unfolded 4 K lines harmonics : the forrest

The CPV 1 & 2 sequences are a treasure, because they have been the only opportunity in all the *HFI* mission to have an accurate estimation of the unfolded 4 K lines, but they also

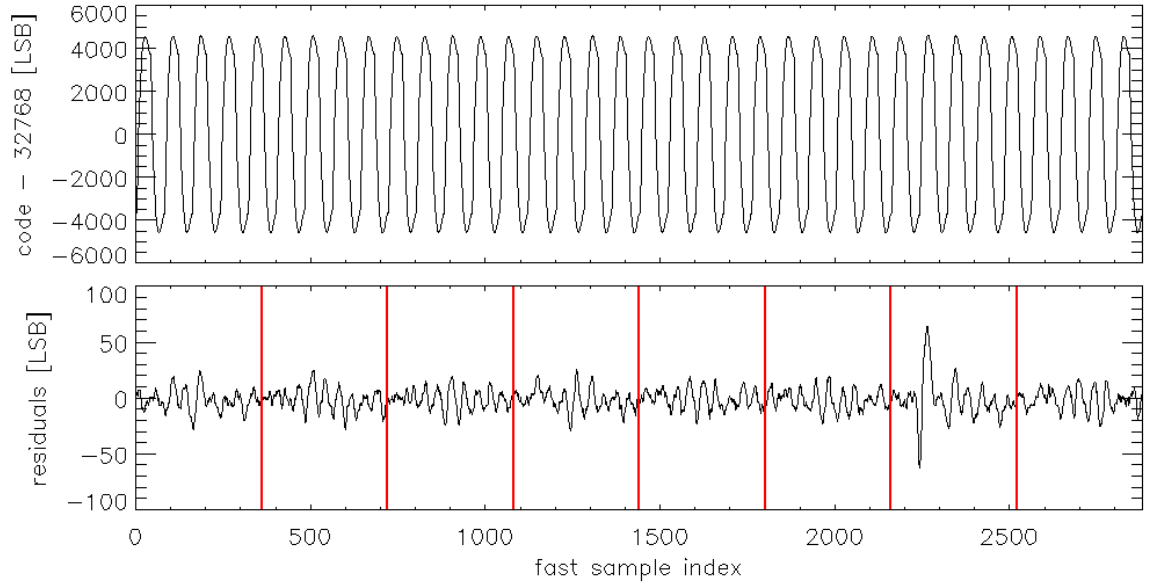


Figure 6.18: Example of contiguous *fast samples* periods from the CPV2 sequence. The upper plot is the raw signal, and the bottom plot show the 4 K lines residuals after removal of the average modulation period. The red vertical lines show the boundary between two 4 K lines periods of 40×9 *fast samples*. A glitch is visible in the penultimate period.

provide valuable informations about glitches and the electronic noise spectral distribution. For the time being glitches are noise for the estimation of unfolded 4 K lines and need to be removed. The algorithm used is similar to the 18 chunks based method described in Sec. 6.1.3, with an adaptation to flag glitch manually (while for the *HFI* mission the DPC glitch flags are used):

- GLITCH FLAGGING: First the science data TOI is computed from *fast samples*. Then a baseline is subtracted, it is produced with a moving boxcar median of 101 samples. Each half period which science data values is over 3 sigma of the noise is flagged out, along with the previous one and the 10 following half periods;
- HALF PERIOD MAPPING TO CHUNKS: As for the science data based version, the *fast samples* are mapped to chunk of 18×40 consecutive *fast samples*, corresponding to two periods as shown in the bottom plot of Fig. 6.18;
- CALCULATION OF GLOBAL PERIOD: all chunks of 18 half periods are averaged (without flagged data) to produce one single period of 18×40 *fast samples* containing two 4 K lines periods plus the modulation signal;
- ESTIMATION OF UNFOLDED HARMONICS: a DFT is run on the global period, the Fourier coefficients of the unfolded 4 K lines harmonics are present at indices multiples of 2×40.

With the CPV datasets, as many as 180 unfolded 4 K lines harmonics frequencies are accessible, instead of the four visible on science data, but the effective number is reduced to 160 because of the degenearcy with the modulation signal at frequencies multiples of $f_{\rm acq}$. This is still a lot!

There is in fact a forrest of harmonics as can be seen in Fig. 6.19 for a detector with an intermediate level of 4 K lines. The Fourier coefficient of the unfolded 4 K lines are estimated using the bootstrap method (**?**), and the estimated errors are in good agreement with the noise level (at location where there is no vertical bars).

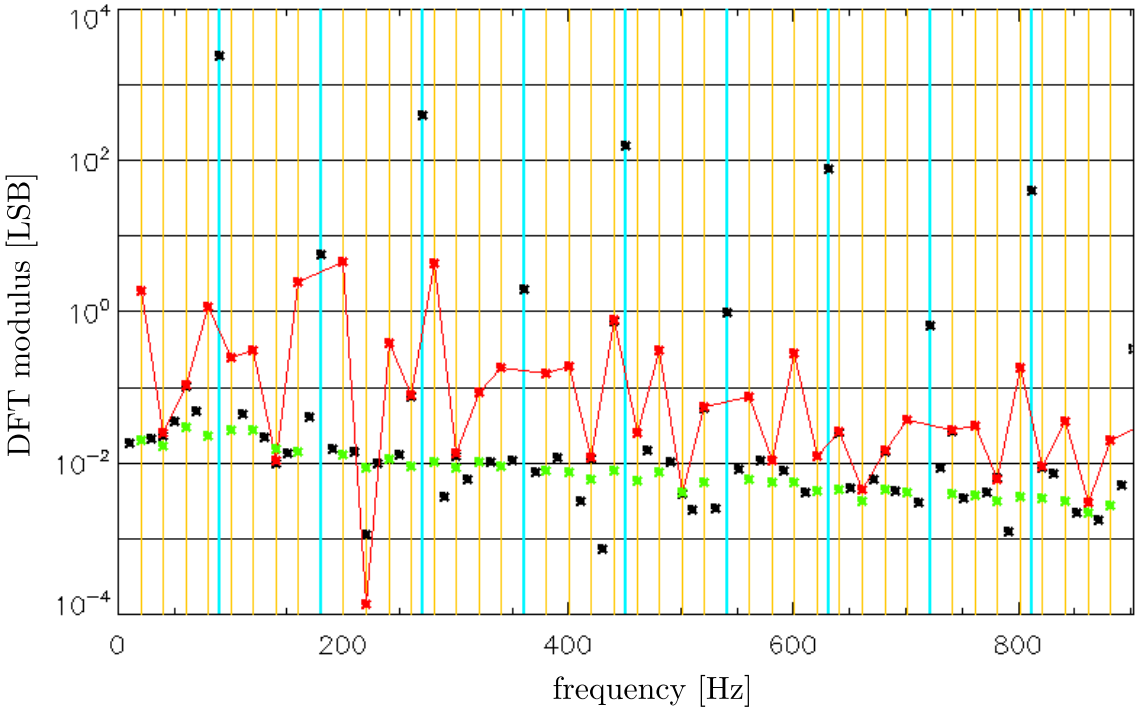


Figure 6.19: Spectral distribution of the unfolded 4 K lines during the CPV 2 sequence for channel `05_353-1`. Black dots are the DFT modulus of the *fast samples* global period. Reds dots are the modulus of the estimated 4 K lines Fourier coefficients. The green dots are the bootstrap estimation of the error for the unfolded 4 K lines harmonics. Eye guides are provided for 4 K lines harmonics frequencies (orange vertical lines) and the modulation harmonics(cyan vertical lines).

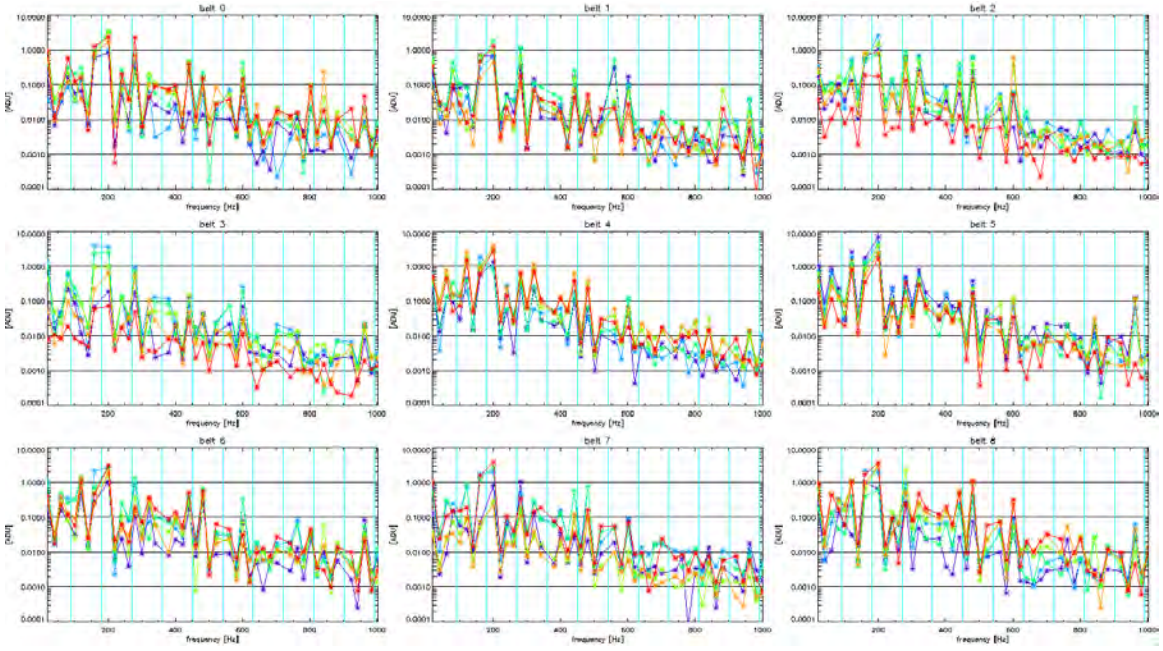


Figure 6.20: Spectral distribution of the unfolded 4 K lines by belt during the CPV 2 sequence. The Y axis is in LSB units. For each belt the 6 channels color index is 0:purple, 1:blue, 2:greenblue, 3:lightgreen, 4:orange, 5:red.

The spectral distribution of the noise is in agreement with the expected response of the *HFI* readout bandpass filter centered on the modulation frequency at 90 Hz with a $1/f$ slope after this frequency. The unfolded 4 K lines spectral distribution tends to follow also this trend, but disagrees about the peak location. This is stringent on the synthetic view for all channels in Fig. 6.20. In contrast with the noise, **the unfolded 4 K lines peak frequency is about 200 Hz**, this point will reveal to be critical for the per ring estimation presented in Sec. 6.3. The reason for the 200 Hz peak frequency is not clear, but most probably due to the shape of their source.

The selection of the harmonics which are significant for ADC nonlinearity correction is problematic because:

1. the level of significance is difficult to estimate, in theory any signal causing a displacement on the ADC scale greater that the INL goal of 1e-2 LSB is significant (see Sec 4.2.3).

2. the spectral distribution changes significantly over the mission, and after the CPV sequences it is not possible to make a direct estimation of the unfolded 4 K lines harmonics.

There is here a real challenge because an estimate of the unfolded 4 K lines harmonics is necessary to build the analog signal model and the *raw constant* parameter for the science data **all over the mission at ring level**.

Time domain drift over CPV 1 and CPV 2

The time domain periods of the unfolded 4 K lines are reconstructed from the frequency estimates calculated in the previous section.

The matrix unfolded 4 K lines periods are provided in Fig. 6.21 for all detector channels plus the fine time resolution channels from belt 9, A, and B. There is a very good agreement between CPV 1 (p1) and CPV 2 (p2) unfolded 4 K lines periods, this is not a fit but a result of phase-syncing between channels.

The 4 K lines signal is rather stable between CPV 1 and CPV 2 acquired one day apart. However it can be seen by eye a wide dispersion between the shapes on all channels,.

Fig. 6.22 presents the absolute periods variation for all channels, the error bars have been inferred from bootstrap results and are much smaller that the variations except for the channel `92_Ther_PID1N_100mK` affected by saturation. It was not intended first to be a correlation plot, but in fact clear tendencies do appear:

- the signal increase given by the fitted line is $\approx 21\%$ and is on average ≈ 0.4 LSB for RMS(p1-p2). This is one order above the significance level for ADC nonlinearity correction;

- the variations observed on these datasets is strongly correlated to the unfolded 4 K lines signal strength. This indicates that either the power of the source does vary, or the transfer function of its chain of transmission does vary.

The modulation by the transponder is a candidate for the observed variations of the 4 K lines signal strength. However it is not the only one, because after the transponder has been set off (at the end of the first survey), the 4 K lines signal continues to exhibit important variations.

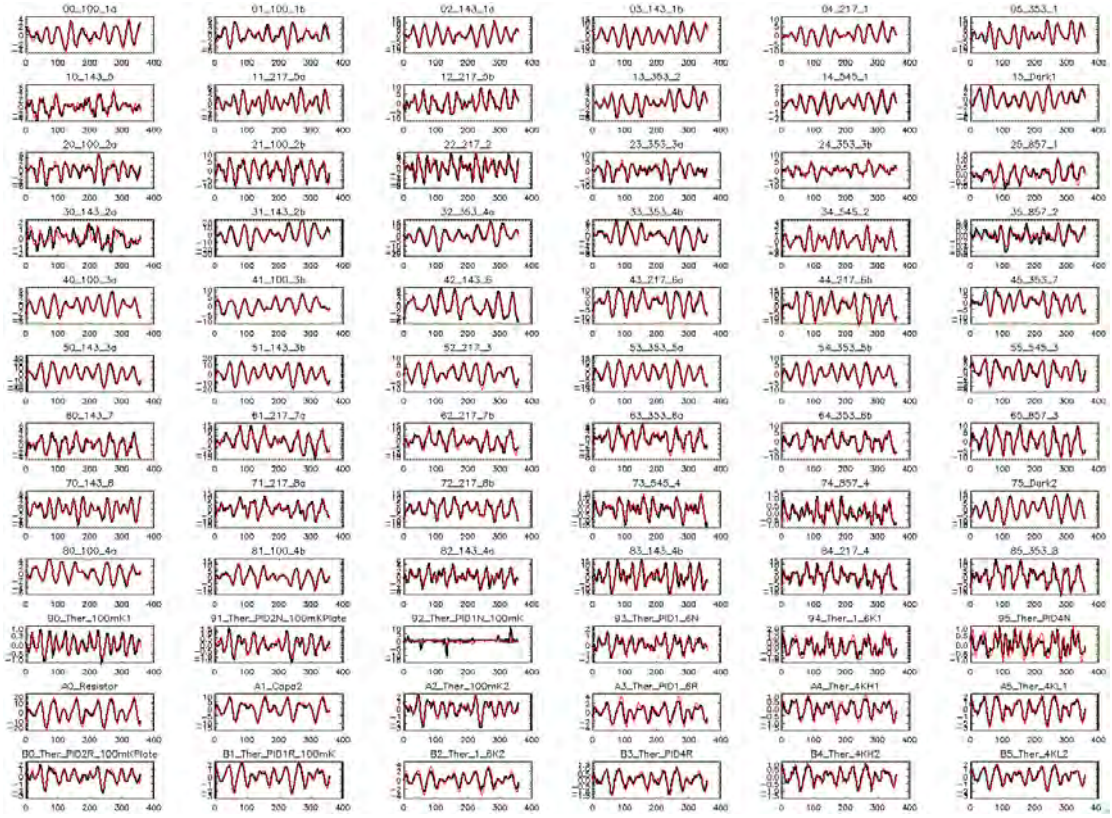


Figure 6.21: Reconstructed time domain 4 K lines periods for CPV 1 (black line) and CPV 2 (red line). The column is the channel index from 0 to 5, and the row is the electronic belt index from 0 to B ($B = 11$ in octal). All periods are shown with the same time reference relative to the 4 K lines periodicity.

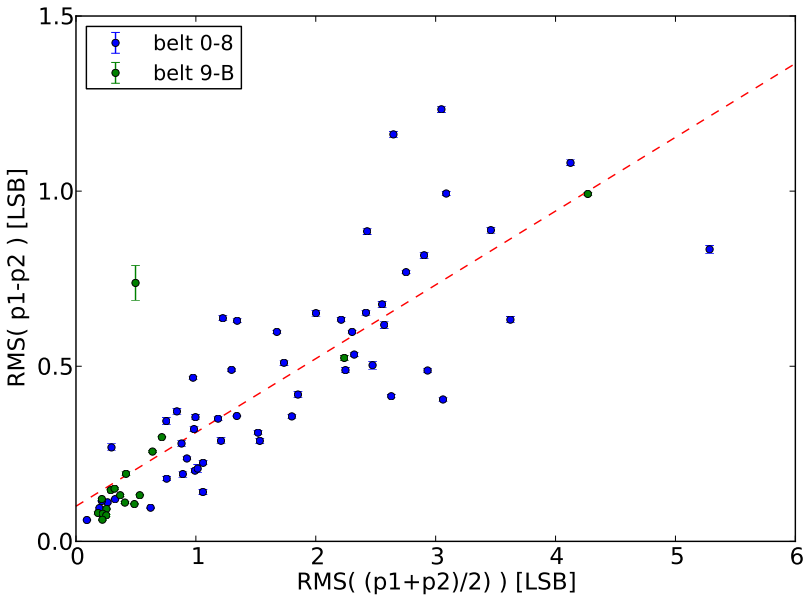


Figure 6.22: Variation of the time domain unfolded 4 K lines period between CPV 1 and CPV 2, as a function of unfolded 4 K lines RMS. Error bars are the one sigma extent. All the 72 fine time resolution channels are present including belt 9,A,B. The CPV 1 unfolded time domain period is labelled $p1$ and for CPV 2 it is labelled $p2$ respectively.

Spectral drift over CPV 1 and CPV 2

To check if it is possible to build a model for the spectral distribution of unfolded 4 K lines, the variations of this distribution have compared between CPV 1 and CPV 2.

The relative variation of Fourier coefficients modulus over CPV 1 and CPV 2 are shown in Fig 6.23. A filtering is applied to select only relevant harmonics obeying the two criteria:

- the modulus must be above 5 time the level of the noise in a 5 Hz window around the harmonic frequency;
- the absolute level of the harmonic must also be above 0.05 LSB.

No specific pattern is visible for the spectral drift, even for the belt five which has the strongest parasitics from the 4 K lines all over the mission. The variations on a frequency basis appear to be rather random, except for the 440 Hz harmonic (22×20 Hz) which increase significantly between both sequences for most channels. A potential origin for the apparent dispersion of the spectral drifts is the modulation by the sorption cooler bed switching, but this option has not been checked, it requires a dedicated work that could not be scheduled.

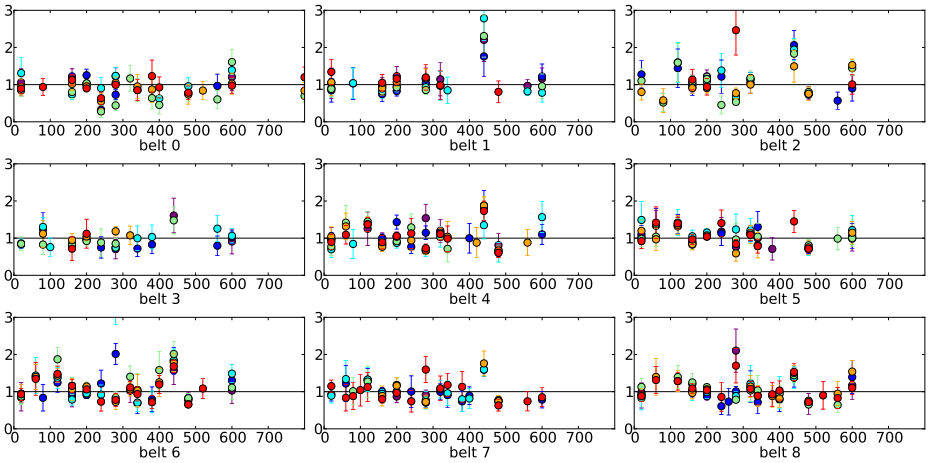


Figure 6.23: Spectral drift over CPV 1 and CPV 2 in from zero to 800 Hz. Each data point is the relative variation of the Fourier coefficients (should be 1.0 for no variation), and error bars are the one sigma extent. Channels are grouped by belt with the same color index as for Fig. 6.20.

6.2 Inclusion of 4 K lines in the ADC nonlinearity correction

6.2.1 Updated analog signal model

The unfolded 4 K lines signal $h(t)$ can be described concisely from a minimal set H_{4K} containing the indices of the most significant harmonics as multiples of the base 20 Hz frequency. With this notation, the signal $h(t)$ reads

$$h(t) = \Re\left(\sum_{k \in H_{4K}} z_k e^{jk\omega_{4K}t}\right), \tag{6.1}$$

where z_k is the Fourier coefficient of the unfolded 4 K lines harmonics at frequency kf_{4K}, $\omega_{4K} = 2\pi f_{4K}$ is the base angular frequency of 4 K lines and $\Re(z)$ stands for real part of the complex number z.

After inclusion of the unfolded 4 K lines signal $h(t)$, the analog signal $s(t)$ given in Eq. 5.1 which has been described in Sec. 5.2, reads

$$s(t)|_{P_{\mathrm{sky}}+\delta P_{\mathrm{sky}}} = \overline{s}(t) + h(t) + \delta P_{\mathrm{sky}} \times g(t) + \mathcal{O}(\delta {P_{\mathrm{sky}}}^2)\,. \tag{6.2}$$

6.2.2 Updated science data transfer function

The transfer function expression cannot be described any more with only two half periods. In contrast, the periodicity of $\overline{s}(t) + h(t)$ extends over 18 half periods. To take into account $h(t)$, a new vector $\boldsymbol{h}$ of size 9×40 *fast samples* is introduced, its time reference is taken for the science data TOI index 0.

Now the updated expression of Eq. 5.12 for the science data transfer function is

$$\mathcal{F}_{p,k,r}(P_{\mathrm{sky}}) = \sum_{j=0}^{39} \mathcal{A}_\sigma \left(\underbrace{\boldsymbol{c}_{p,r}}_{\langle P_{\mathrm{sky}} \rangle}(j) + \boldsymbol{h}_r(k \times 40 + j) + \delta P_{\mathrm{sky}} \times \boldsymbol{g}_p \right), \tag{6.3}$$

where:

- $p = i \bmod 2$ is the parity of the science data sample at TOI index i;
- $k = i \bmod 9$ is the relative index of the science data sample in the 4 K lines period;
- r is the ring index for which the science data transfer function parameters are estimated;
- $\boldsymbol{c}_{r,p}$ is the *raw constant* vector as described in Sec. 5.4;
- $\boldsymbol{g}_{r,p}$ is the *raw gain* vector as described in Sec. 5.4;
- $\boldsymbol{h}_r$ is the unfolded 4 K lines period vector of 9×40 *fast samples*.

With the introduction of the unfolded 4 K lines signal there are now as many as 18 science data transfer functions for each ring. In practice a total of $18 \times 27\,000$ functions needs to be calculated to correct for the ADC nonlinearity over the full *HFI* mission on every channel.

6.2.3 Simulations

Simulation of the 4 K lines effect on correction

For a better understanding of the effect of 4 K lines on the ADC nonlinearity correction, a simulation was run over the full mission range. The goal is to distinguish the effect of 4 K lines alone from the bias introduced on the *raw constant* parameter which is the general case of interest. The simulation is run on noiseless PBR data processed with the science data transfer function as it has already been shown to be equivalent to the science data TOI processing in Sec. 5.4.4.

- a reference demodulated PBR $\mathrm{SKY}_{\mathrm{ref}}$ is generated by drizzling (see Sec. 5.5.2) of the demodulated science data after cleaning of the folded 4 K lines. The modulation offset o is estimated from the average value of science data (in DSN units);
- two modulated reference PBRs are produced with $\mathrm{SKY}_0 = o + \mathrm{SKY}_{\mathrm{ref}}$ and $\mathrm{SKY}_1 = o - \mathrm{SKY}_{\mathrm{ref}}$;
- the 18 PBRs with ADC nonlinearity effect are produced by application of the updated (Eq. 6.3) science data transfer function yielding $\mathrm{SKY}^{\mathrm{4K}}_{p,k} = \mathcal{F}_{p,k}\,(\mathrm{SKY}_{\mathrm{p}})$;

- the 18 PBRs are ADC corrected using the correction version 4 K lines (Eq. 5.12) which reads $\mathrm{SKY}^{\mathrm{C}}_{p,k} = \mathcal{F}_p^{-1}(\mathrm{SKY}^{\mathrm{4K}}_{p,k})$.

The half parity gain observable can be calculated from the 18 PBRs version by averaging the nine PBRs of each parity.

Input parameters (INL, unbiased *raw constant* and unfolded 4 K lines) have been taken from the 2013 data release estimates and are considered to be realistic.

Results are shown in Fig. 6.24 for a channel which is known to have a medium level of 4 K lines and for which the ADC nonlinearity observables (half parity gain) are significantly worse after correction. The feature of interest is the seemingly small effect of 4 K lines Indeed, in both cases, simulation and correction, the result is not far from the version without 4 K lines.

The explanation of this apparently small impact lies in the degeneracy between the unfolded 4 K lines signal and the *raw constant* parameter. For the simulations, the *raw constant* parameter is unchanged but unfolded 4 K lines harmonics are added or removed from the analog signal model. For real data the *raw constant* parameter can only be estimated with a residual bias due to the 4 K lines. This bias is present in the 18 versions of the science data transfer function, hence producing a cumulatively higher effect.

Two important points arise from this analysis:

1. 4 K lines real impact on the science data transfer function is of second order;
2. the *raw constant* parameter estimation is critical, or in other word it must be cleaned from 4 K lines to the best possible extent.

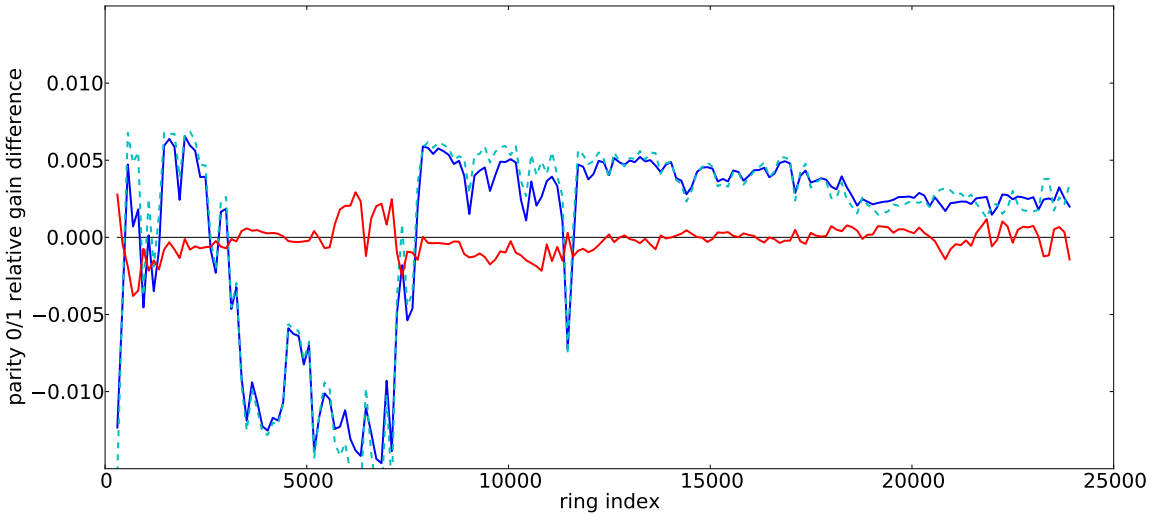


Figure 6.24: Simulation of 4 K lines impact on ADC nonlinearity using half parity gain observable. The blue line is the observables for the ADC nonlinearity effect including 4 K lines on $\mathrm{SKY}^{\mathrm{4K}}$ PBRs. The red line is for $\mathrm{SKY}^{\mathrm{4K}}$ corrected without handling of 4 K lines. The cyan dashed line has been added to show the simulation of ADC nonlinearity effect without inclusion of 4 K lines.

Susceptibility of ADC nonlinearity to 4 K lines harmonics

In order to select the unfolded 4 K lines harmonics presenting the highest impact on ADC nonlinearity, simulations were run to quantify their impact on the half parity gain observables. The susceptibility coefficient $S_g(k)$ of the half parity gain observable to unfolded

harmonic k is defined from the observables over the full mission as follow

$$S_g(k)|_{\Re(z_k)} = \frac{\delta RMS(g(r)-1)}{\Re(\delta z_k)}$$
$$S_g(k)|_{\Im(z_k)} = \frac{\delta RMS(g(r)-1)}{\Im(\delta z_k)}\,, \tag{6.4}$$

where $g(r)$ is the half parity gain for ring r as defined in Eq. 5.18. The harmonic susceptibility coefficient $S_g(k)$ is taken as the quadratic mean of real and imaginary versions.

The results for all channels are shown in Fig. 6.25. It can be noted that the susceptibility spectra shows a peak for the harmonics around the multiples of f_{acq} on most channels. This is expected as frequencies close to harmonics of f_{mod} tends to be in synchronization with the modulation, thus degenerate more strongly with the *raw constant* parameter.

The unfolded 4 K lines harmonics at **160 Hz and 200 Hz** are of particular interest for three reasons: first the ADC nonlinearity is on more susceptible to these frequency on most channels (see orange dots in Fig. 6.25), next they are among the strongest of the CPV sequences for most channels, and finally they also fold at 20 Hz which is the strongest line on science data for most channels. Hence they will be included into the restricted set $H_{4\text{K}}$ of harmonics estimated for science data correction.

6.3 Estimation of parameters for 2013 data release

To produce the parameters for the updated version of the science data transfer function, the 4 K lines signal must be estimated on a per ring basis along with the *raw constant* parameter. There were two main issues: the first one is that both 4 K lines and the *raw constant* signals are degenerate because of the 2/9 window visibility of the captured *fast samples* period (described in Sec. 6.1.4). The second one is the destruction of the unfolded 4 K lines harmonics forrest by the DPU summation.

In this section, it is detailed how the degeneracy can be broken between the 4 K lines signal and the *fast samples* period. Additional constraints from the CPV are needed for this solution. Finally the estimates of the unfolded 4 K lines harmonics are calculated with a constrained fit using the Lagrange multipliers method. The constraints for this fit comes from the folded 4 K lines on science data.

The parameters produced with the method described below have been the ones used to correct for the ADC nonlinearity of the *Planck* 2013 data relase.

6.3.1 Breaking degeneracy with the *half parity sum*

As a consequence of the *fast samples* capture window detailed in Sec. 6.1.4, there is a strong degeneracy between the unfolded 4 K lines signal and the average modulation shape aka the *raw constant*. This problem has to be handled in order to provide constraints for the estimation of unfolded 4 K lines.

Hopefully it is possible to take advantage of the symmetry of the *raw constant* to break this degeneracy, because the *fast samples* period is exactly one modulation period in length. Let write first the *raw constant* from its Fourier coefficients

$$c(i) = \sum_{k=-n}^{n} m_k e^{j2\pi ki/80}\,, \quad i \in [\![0, 79]\!]\,. \tag{6.5}$$

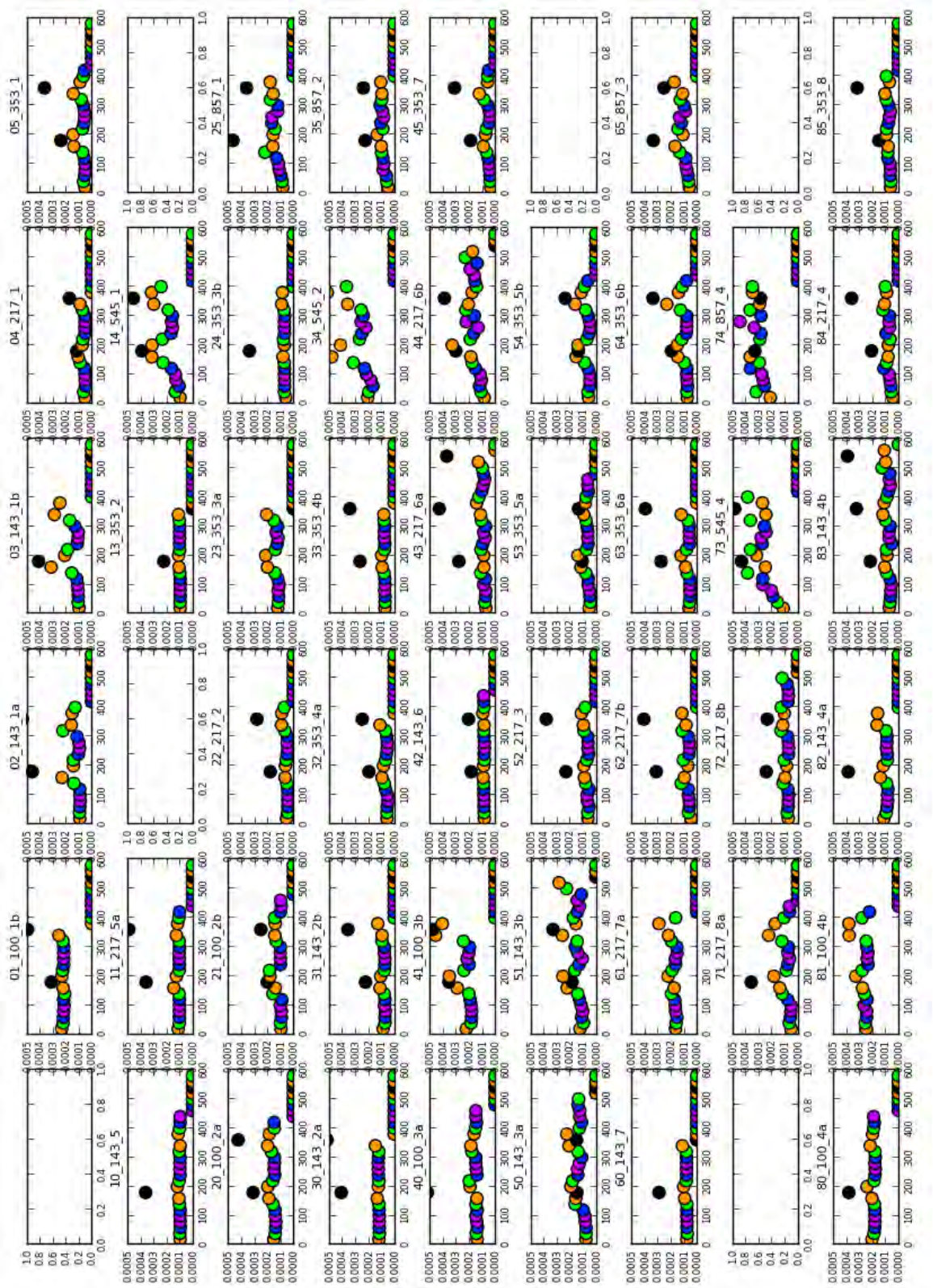


Figure 6.25: Spectrum of ADC nonlinearity susceptibility S_g to unfolded 4 K lines in LSB^{-1} units. The x-axis ranges from 0 to 600 Hz and ticks are displayed with increments of 100 Hz. The y-axis ranges from 0.0 to 5×10^{-4} and has no unit. The color code for the folding frequency on science data is: orange folds to 20 Hz, green folds to 40 Hz, blue folds to 60 Hz and purple folds to 80 Hz. Black dots are for harmonics multiple of the modulation frequency (180 Hz, 360 Hz, ...).

Taking the sum of the first and the second half of the *raw constant* reads

$$\begin{aligned} c(i) + c(i+40) &= \sum_{k=-n}^{n} m_k \left(e^{j2\pi ki/80} + e^{j2\pi k(i+40)/80} \right) \\ &= \sum_{k=n}^{n} m_k e^{j2\pi ki/80} \underbrace{\left(1 + e^{2\pi k/2}\right)}_{\in \{0,2\}}, \end{aligned} \tag{6.6}$$

where n is the finite number of harmonics for the decomposition. The rightmost factor value is two for even harmonics, and zero for odd harmonics, making these one vanish. Let note here that m_0 is the modulation offset real value usually noted o.

The sum of the first and second half of the *raw constant* will be called the *half parity sum* hereafter. It has the property of cancelling the even harmonics of the modulation, but not the unfolded 4 K lines harmonics for which the 20 Hz base period is $1/2f_{4\mathrm{K}} = 2/9f_{\mathrm{mod}}$. The unfolded 4 K lines harmonics index multiples of 9 (180 Hz) which are also multiples of the acquisition frequency, are considered to be taken into account by the *raw constant* term itself, and will not be considered any more. A specific analysis of the effect of the sorption cooler bed modulation would be useful at these frequencies, but it has not been performed yet for the same reasons as the study of the sorption cooler impact on ADC nonlinearity.

6.3.2 Using even modulation harmonics from CPV

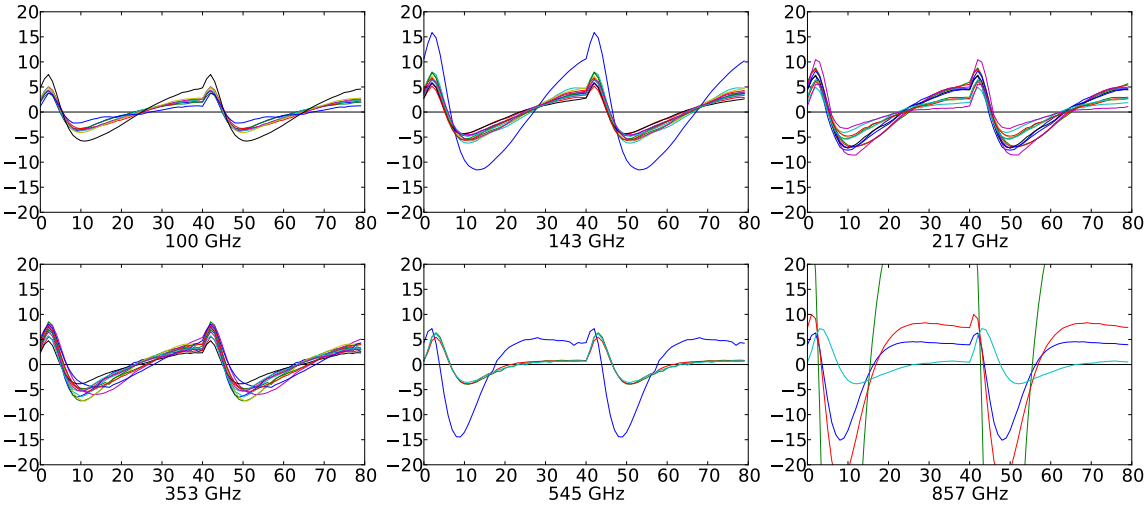


Figure 6.26: Even modulation harmonics templates from CPV for all channels and grouped by detector optical frequency. The average value (the modulation offset) has been subtracted.

The even modulation harmonics signal completely degenerate with the unfolded 4 K lines, as a consequence the *fast samples* period acquired during the mission cannot be used reliably, but the degeneracy can be broken with the datasets from the CPV 1 and CPV 2 sequences. The spectral distribution of the modulation harmonics can be seen in the previous spectra in Fig. 6.19, where the even modulation harmonics are under 10^1 LSB. They can be estimated straightforwardly as a byproduct of chunk 18 method described in Sec. 6.1.4, which contain the modulation plus 4 K lines information at the same time, allowing to separate both.

The time domain templates of even modulation harmonics for all channels are shown in Fig. 6.26. There is a lot of dispersion between channels, confirming that individual templates are required.

A critical point is that the `ibias` parameter, setting the bias current intensity for the detector channels, has been changed by a few digits after the CPV sequences and before the beginning of the nominal mission. This new setting makes the signal of the even modulation harmonics drifting slightly from the template that has been calculated with the CPV *fast samples*. As it strongly degenerates with the 4 K lines (significantly stronger at the beginning of the mission than during the CPV sequences) it is impossible to make new estimates during the cold mission with the currently available methods, but the impact if about 1 LSB.

The consequence of the drift from CPV templates is a bias for the estimation of unfolded 4 K lines harmonics. Hence the strongest the 4 K lines signal, the better the estimation of the harmonics coefficients will be.

6.3.3 Parameters estimation using the Lagrange multipliers method

It is possible to estimate the *raw constant* parameter along with the unfolded 4 K lines harmonics using the *fast samples* periods captured during the mission and the science data folded 4 K lines estimates. To do so a least square minimization with constraints is performed using the Lagrange multipliers method.

Let us consider a ring index r for a fixed detector channel. The data that has to be modeled is the half parity sum vector noted $\hat{p}$ with a size of 40 *fast samples*. All *fast samples* data is corrected with the function $\mathcal{A}^{-1}$ and is in LSB units. Using the half parity sum properties, Eq. 6.2 leads to

$$\hat{p} = 2 \times o_{\text{AC}} + 2 \times \boldsymbol{m}_e + \boldsymbol{y}\,, \tag{6.7}$$

where o_{AC} is the modulation offset, $\boldsymbol{m}_e$ is the 40 *fast samples* vector of the even modulation harmonics and $\boldsymbol{y}$ is the half parity sum vector of the unfolded 4 K lines signal for the captured *fast samples* period. The parameter o_{AC} is estimated from the average value of the nonlinear science data TOI, and is further refined by using the average value of the science data TOI after ADC nonlinearity correction. The fixed parameter $\boldsymbol{m}_e$ is obtained from the *fast samples* data of CPV sequences.

The free parameters of the problem are the Fourier coefficients of the unfolded 4 K lines harmonics noted $\boldsymbol{x}$. The size of $\boldsymbol{x}$ is twice the size of the selection of unfolded 4 K lines harmonics H_{4K} because the solving is done in the real domain, therefore the real and imaginary part are dissociated. Using matrix notation $\boldsymbol{y}$ can be calculated from its harmonics coefficients with

$$\boldsymbol{y} = \boldsymbol{F}\boldsymbol{x}\,, \tag{6.8}$$

where $\boldsymbol{F}$ is the matrix of the linear function building the half parity sum from the 4 K lines. The matrix $\boldsymbol{F}$ is built with sin and cos functions.

The constraints come from the science data folded version of 4 K lines noted $\boldsymbol{y}'$. It contains eight elements for the real and imaginary part of the four science data Fourier coefficients. Using matrix notation $\boldsymbol{y}'$ can be written as the output of a linear function with

$$\boldsymbol{y}' = \boldsymbol{G}\boldsymbol{x}\,, \tag{6.9}$$

where $\boldsymbol{G}$ is the DPU summation function of the unfolded 4 K lines harmonics. The analytical expression for the transfer function of the DPU summation is described in Appendix H, and the abacus of its modulus is shown in Fig. 6.27. This abacus has really been used a lot for the comparison of relative harmonics impact on science data.

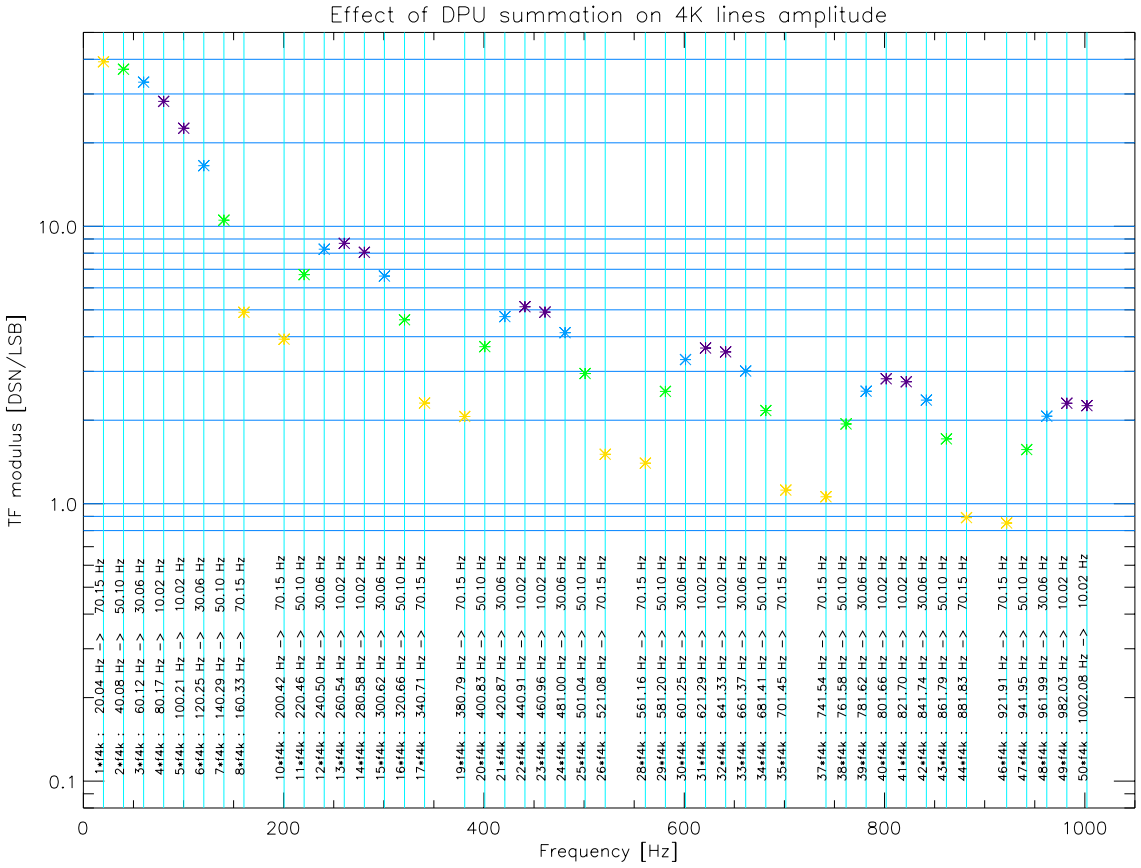


Figure 6.27: Abacus of the DPU summation transfer function in frequency domain. Color code for folded frequency on modulated science data is 20 Hz:yellow, 40 Hz: green, 60 Hz:blue, 80 Hz:purple. For each harmonic the frequency value at each side of the arrow is for `unfolded modulated -> folded demodulated` signal.

In the general case the estimate $\hat{\boldsymbol{x}}$ found by least squares minimization is

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{F}^{\mathbf{T}}\boldsymbol{W}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^{\mathbf{T}}\boldsymbol{W}\boldsymbol{y}\,, \tag{6.10}$$

where $\boldsymbol{W} = \boldsymbol{I}_d/(2\sigma^2/n_p)$ is the inverse of the covariance matrix of $\boldsymbol{y}$, $\boldsymbol{I}_d$ is the 8×8 identity matrix, n_p is the number of *fast samples* period for ring r and σ is the RMS of the noise on the *fast samples*. The factor two in the denominator of $\boldsymbol{W}$ takes into account the increase in the noise due to the *half parity sum*.

Introduction of the science data TOI constraints can be done by using the elegant matrix transcription of the Lagrange multipliers method described in **?**, and **?** (see p. 5, 13 and 44). Finally the solving boils down to the system

$$\left(\frac{\hat{\boldsymbol{x}}}{\boldsymbol{\lambda}}\right) = \left(\begin{array}{c|c} \boldsymbol{F}^{\mathbf{T}}\boldsymbol{W}\boldsymbol{F} & \boldsymbol{G}^{\mathbf{T}} \\ \hline \boldsymbol{G} & 0 \end{array}\right)^{-1}\left(\frac{\boldsymbol{F}^{\mathbf{T}}\boldsymbol{W}\boldsymbol{y}}{\boldsymbol{y}'}\right)\,, \tag{6.11}$$

where $\boldsymbol{\lambda}$ is an unused output of the method which contains the Lagrange multipliers factors for the eight science data TOI constraints.

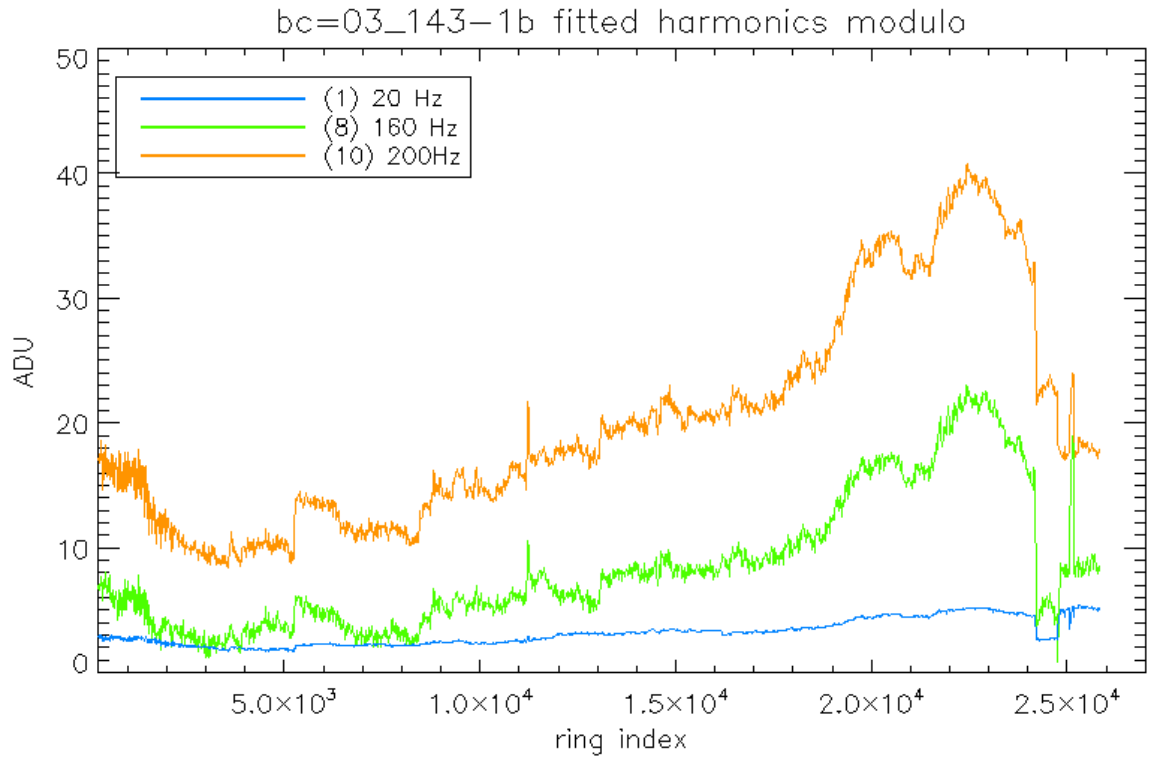


Figure 6.28: Estimates of unfolded 4 K lines harmonics for channel 03_143-2b in LSB units.

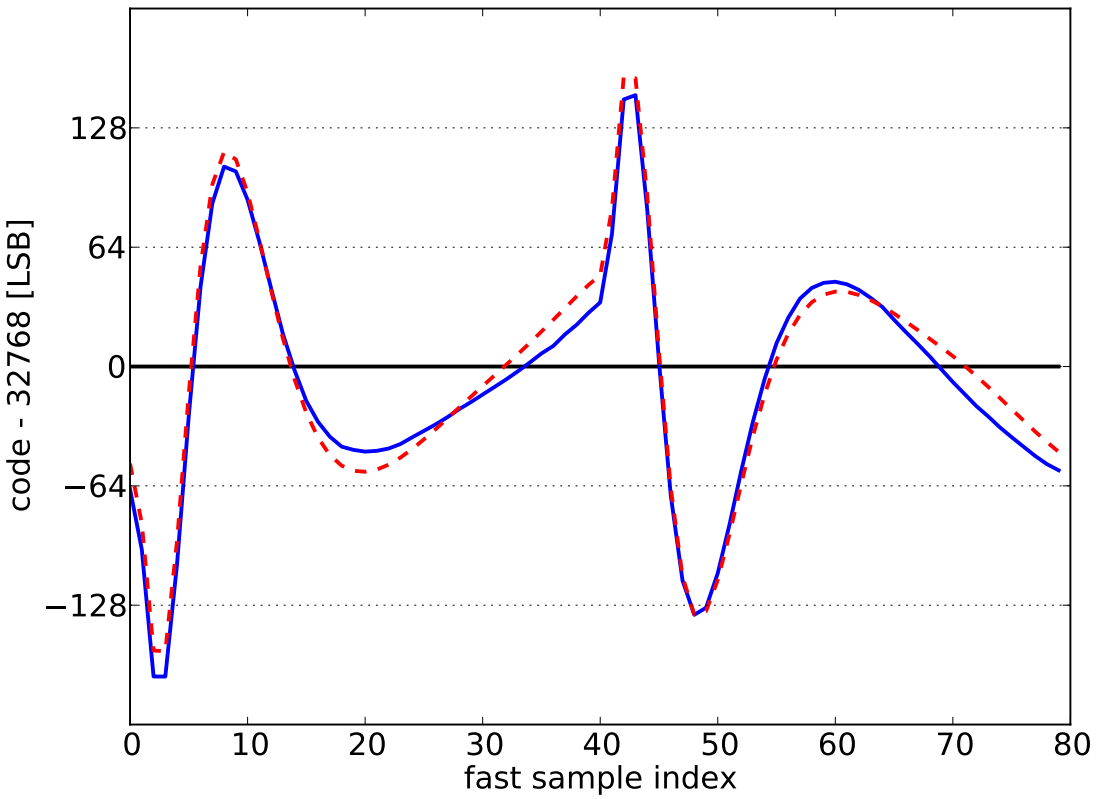


Figure 6.29: Removal of 4 K lines from the averaged *fast samples* period for channel 03_143-2b at ring index 17500. The blue line is before removal, and the red dashed line is the cleaned *raw constant* parameter.

6.3.4 2013 correction results

For this run, the set of unfolded 4 K lines harmonics indices fitted has been reduced to

$$H_{4\mathrm{K}} = \{1, 8, 10\}\,, \tag{6.12}$$

corresponding to 20 Hz, 160 Hz and 200 Hz all folding to 20 Hz. There are two main motivations for this selection: First the conditioning of the square matrix used to solve for $\boldsymbol{x}$ grows quickly over 1000 for more than three harmonics (6 free parameters), thus limiting the number of harmonics available for the fit; The second motivation is that the 20 Hz folded harmonic is the strongest on science data (see Fig. 6.7). As the two strongest harmonics in the CPV sequences are 20 Hz and 160 Hz for most channels they are selected by default. Finally the DPU summation amplifies the relative weight of the 20 Hz harmonic by a factor 10 in comparison to the 160 Hz and 200 Hz (see the first three yellow dots on the abacus in Fig. 6.27), thus making its inclusion mandatory.

An example of fitted unfolded harms is shown in Fig. 6.28. The relative amplitude of the harmonics is in good agreement with the observations from the CPV sequences. An unexpected step is visible just before ring 25000 (after the end of the 5^{th} survey). This step is bounded by two successive DPU reboots, in practice the amplitude variation is not real thus demonstrating that the fit procedure is imperfect. Other jumps are real variations after events like the sorption cooler switchover and solar flares.

The *raw constant* parameter $\boldsymbol{c}_r$ is obtained by *cleaning* the *fast samples* averaged period from the fitted 4 K lines signal. An example of the result is shown in Fig. 6.29. The effect of 4 K lines appears as a significant departure from the averaged *fast samples* period $\boldsymbol{p}$. It demonstrates visually why the impact on the science data transfer function is so strong.

Quantification of the improvement

Finally the averaged half-parity gain observables after correction (version of *Planck* 2013 data release) are shown in Fig. 6.30 for all channels, with the same scale and presentation as in Fig. 6.2. The first striking observation is a **significant reduction of the ADC nonlinearity observables and their dispersion** for all channels, at the exception of a few. The correlation between the 4 K lines strength and the ADC nonlinearity observables has mostly disappeared. These results are a strong validation of the ADC nonlinearity correction accuracy, since the observables are completely uncorrelated from the data used to build the correction.

The average value of half-parity gain observables after correction is around 10^{-3}, with many channels close to 10^{-4}, which is the expected noise limit. However a large number of channels still exhibit observables significantly higher than the expectation for two main reasons:

- the even harmonics template from the CPV is biased because of the change of the `ibias` parameter changed after the CPV sequences;
- the fit procedure is not enough constrained to estimate more than 3 harmonics, while the 4 K lines forest counts many.

The Lagrange multipliers method produces excellent results for the correction of ADC nonlinearity. This can be explained by the fact that it performs well in the removal of the 4 K lines spurious signal from the *raw constant*, which has been shown to be a critical parameter of the correction (see Sec. 6.2.3).

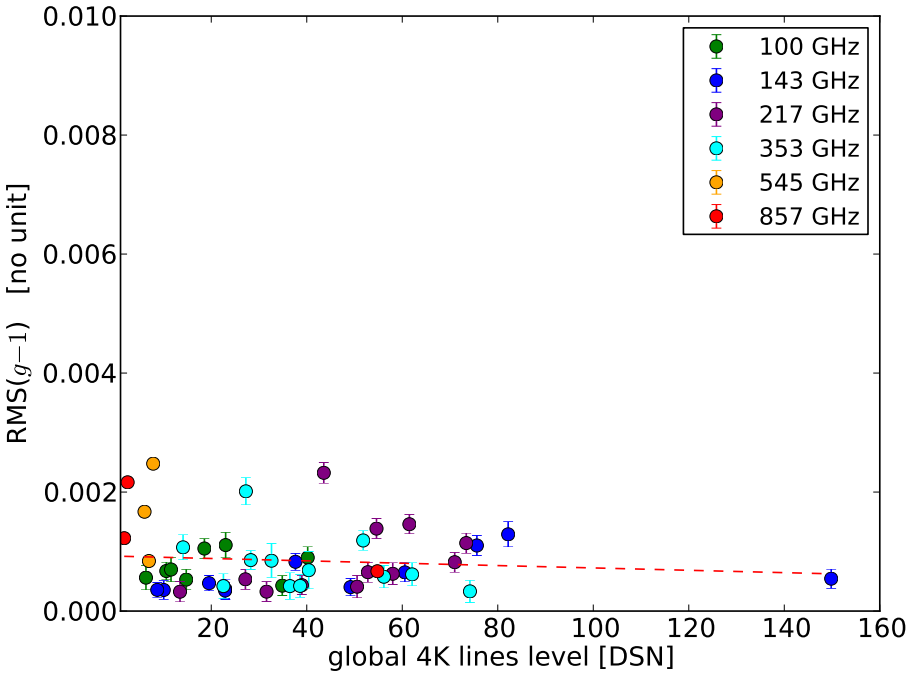


Figure 6.30: ADC nonlinearity observables for the *Planck* 2013 data release. The correlation plots uses the same presentation as for Fig. 6.2. The science data transfer function is parametrized with the *raw constant* and the unfolded 4 K lines used for the *Planck* 2013 data release.

Analysis with the half parity 9-gain observable

The values provided by the half parity gain observable are a little simplistic when considering 4 K lines. For the analysis of 2013 correction results including the handling of 4 K lines it must be refined. In fact it has been seen in Sec. 6.2.2 that there are 18 different versions of the science data transfer function, thus as many (slightly) different version of the sky because of the ADC nonlinearity. Let see how the half parity gain observable described in Sec. 5.5.3 can be extended to provide useful information on the ADC nonlinearity introduced by 4 K lines.

The 18 different PBRs corresponding to each configuration of the science data transfer function are noted $\{\mathrm{SKY}_k, k \in [\![0, 17]\!]\}$. They are produced by drizzling science data samples for which the value of the TOI index modulo 18 is k. A new observable can be built by doing the comparison between two consecutive PBRs and rewriting Eq. 5.17 as

$$\mathrm{SKY}_{2k+1} = g_k \times \mathrm{SKY}_{2k} + c \quad k \in [\![0, 8]\!]\,, \tag{6.13}$$

where $g_k \approx 1$ is the relative gain parameter between SKY_{2k} and SKY_{2k+1}. This set of nine relative gain estimations is called the **half parity 9-gain observable**.

The matrix of observables for all detectors are presented in Fig. 6.32 for uncorrected science data, and in Fig. 6.33 for science data processed with the 2013 ADC correction. Several cases can be distinguished by comparison of the observables after correction:

- THE GOOD: for many channels, the correction with the 4 K lines handling brings improvement for either g and g_k observables (for example the `03_143-1b`). On these channels there is a high probability that the unfolded 4 K lines harmonics fit is of good quality. Thus the correction can be considered reliable.

- THE BAD: for several channels, the g observable is better after correction, but the g_k observables are still bad or worse after the correction (see for example the channel

50_143-3a). In this case, it is probable that the *raw constant* parameter estimated from the fitted unfolded 4 K lines harmonics is not biased too much. However the harmonics estimates are most probably too far from the reality, as seen with the high dispersion of the g_k observables.

- THE UGLY: a few channels are worse after correction for both observables kind, this is the case for several 545 GHz and 857 GHz channels, the worst case is the channel 25_857-1. The reason has been found in 2015 when working on the new correction. It comes from the even harmonic template m_e which changed dramatically between the CPV sequences and the *HFI* mission.

It is highly instructive to have a look at the 9-gain observables correlation plots with the 4 K lines level. The evolution is shown between the science data before and after ADC nonlinearity correction of 2013 in Fig. 6.31. After ADC nonlinearity correction of 2013, the correlation with 4 K lines level is still there, albeit it has disappeared for the half parity gain observable. While the observables values are more compact and the effect is lowered by the correction, it is probably mostly due to the first order impact of the INL and mostly cleared *raw constant*. This is clearly a demonstrating that 4 K lines effect is not completely handled. A lot of effort will be put in this direction for the next version of the correction, because these residuals indicates clearly that 4 K lines are a limiting factor for the correction. The impact on science at large angular scale will be discussed in the next chapter.

6.4 Alternative parameters estimation for 2015 correction

The excellent results of the 2013 ADC nonlinearity correction on science data and the analysis of detector response are strongly in favor of the signal model used to produce the science data transfer function. However the ADC nonlinearity residuals estimated with the observables still introduce a bias at $\approx 10^{-3}$ level relative to the solar dipole amplitude. This is yet one order of magnitude above the target value of 10^{-4} (as stated in Sec. 4.2.3 for science at large angular scales. Starting from the hypothesis that the signal model is valid, this section presents the work leading to a new set of parameters for an improved ADC nonlinearity correction. A new constraint will be introduced when fitting the parameters by the inclusion of the ADC nonlinearity observables on science data.

First a characterization of the model parameters is performed with a Metropolis Hastings [2] (MH) algorithm which is a Markov Chain Monte Carlo (MCMC) method. A comparison with similar work performed by Guillaume Patanchon resulted in a significant simplification by considering the stability of the *raw constant* parameter throughout the mission. Then it will be detailed how the new set of parameters is produced by merging these studies which gave excellent results on science data.

6.4.1 Characterization of science data transfer function parameters

This section describes the characterization of the unfolded 4 K lines parameter $\boldsymbol{x}$. The unfolded 4 K lines harmonics most probable parameter values are estimated with a Metropolis Hastings algorithm.

Setup of the MH algorithm

Basically the purpose of the Metropolis Hastingsalgorithm (e.g. see **?**) is to approximate the probability distribution $\mathcal{H}(\boldsymbol{x})$ of the parameter $\boldsymbol{x}$. In this case $\boldsymbol{x}$ is the unfolded 4 K lines

[2] See: https://en.wikipedia.org/wiki/Metropolis-Hastings_algorithm

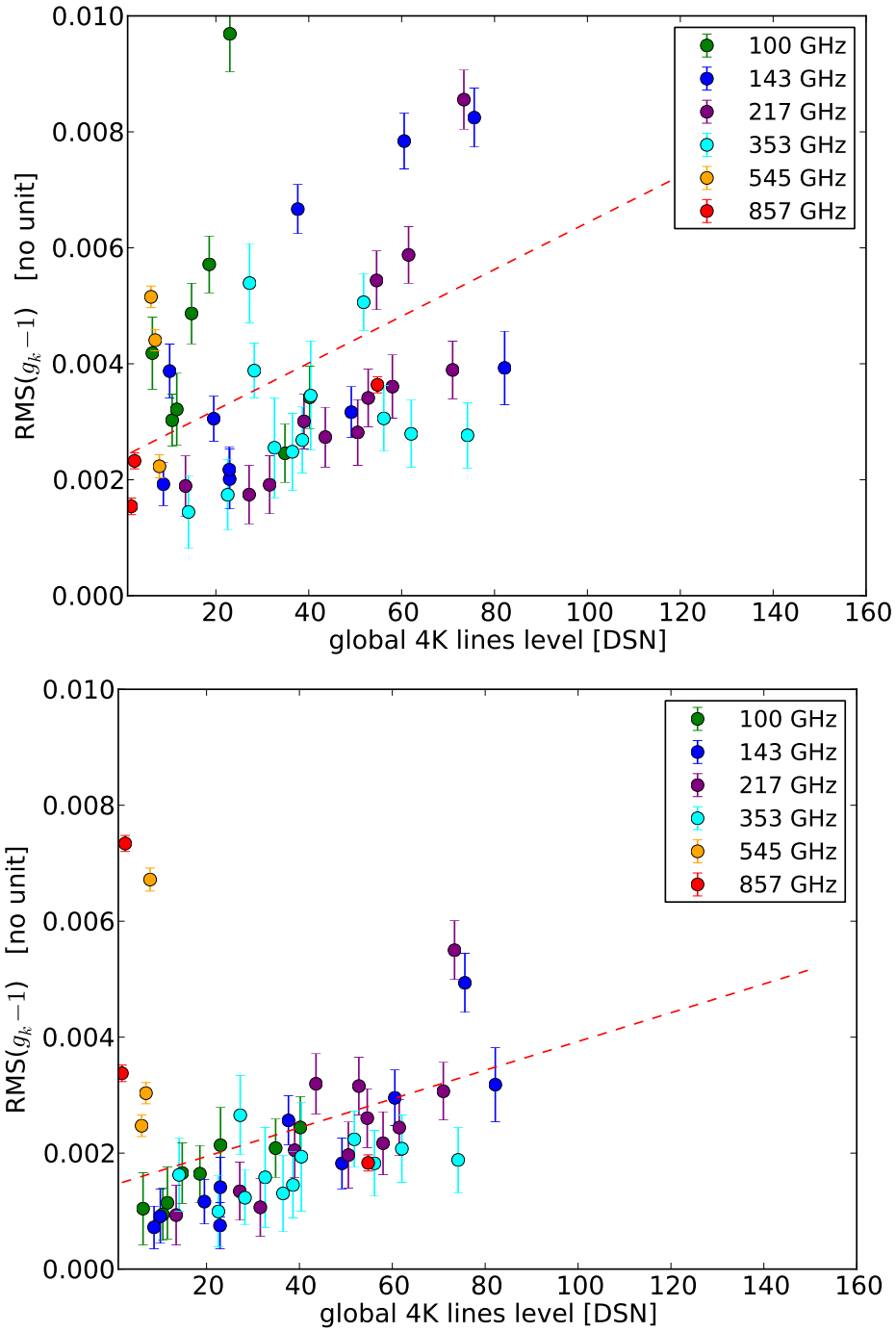


Figure 6.31: ADC nonlinearity 9-gain observables. The correlation plots uses the same presentation as for Fig. 6.2, and errors bars are three times the ones for the half parity gain observable. *Upper panel*: without ADC nonlinearity correction. *Bottom panel*: after application of the ADC nonlinearity correction of the *Planck* 2013 data release.

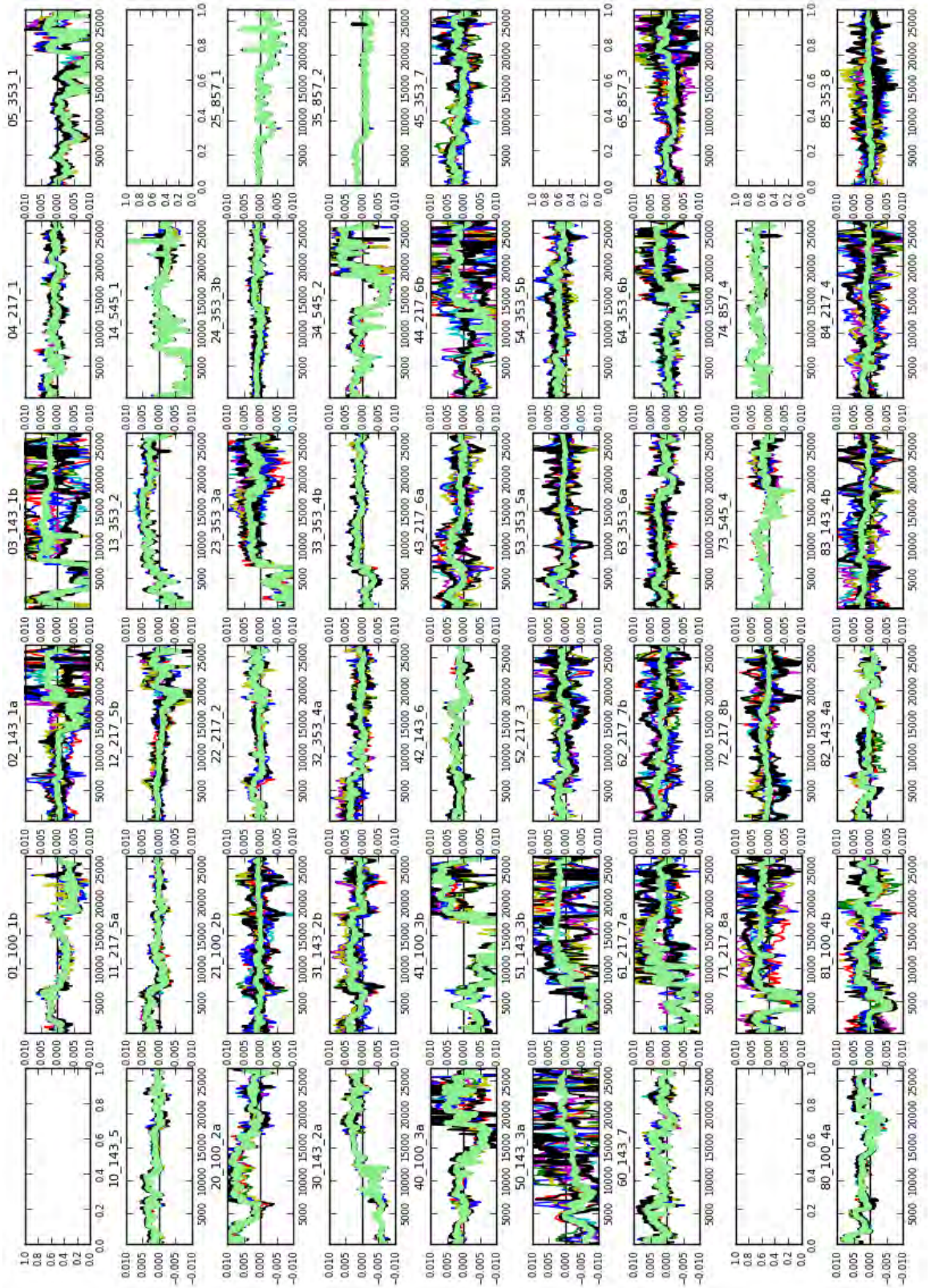


Figure 6.32: Half parity 9-gain observables for all *HFI* detector channels (no correction). The thick green line is the half parity gain, while the 9 thin lines with different colors are the half parity 9-gain. Before calculating the observables, the PBRs are averaged over 128 rings to reduce the noise. A few channels are missing: because either some data was not available (`00`, `15`, `75`) or the channels are too badly affected by random telegraph noise aka popcorn (`55`, `70`).

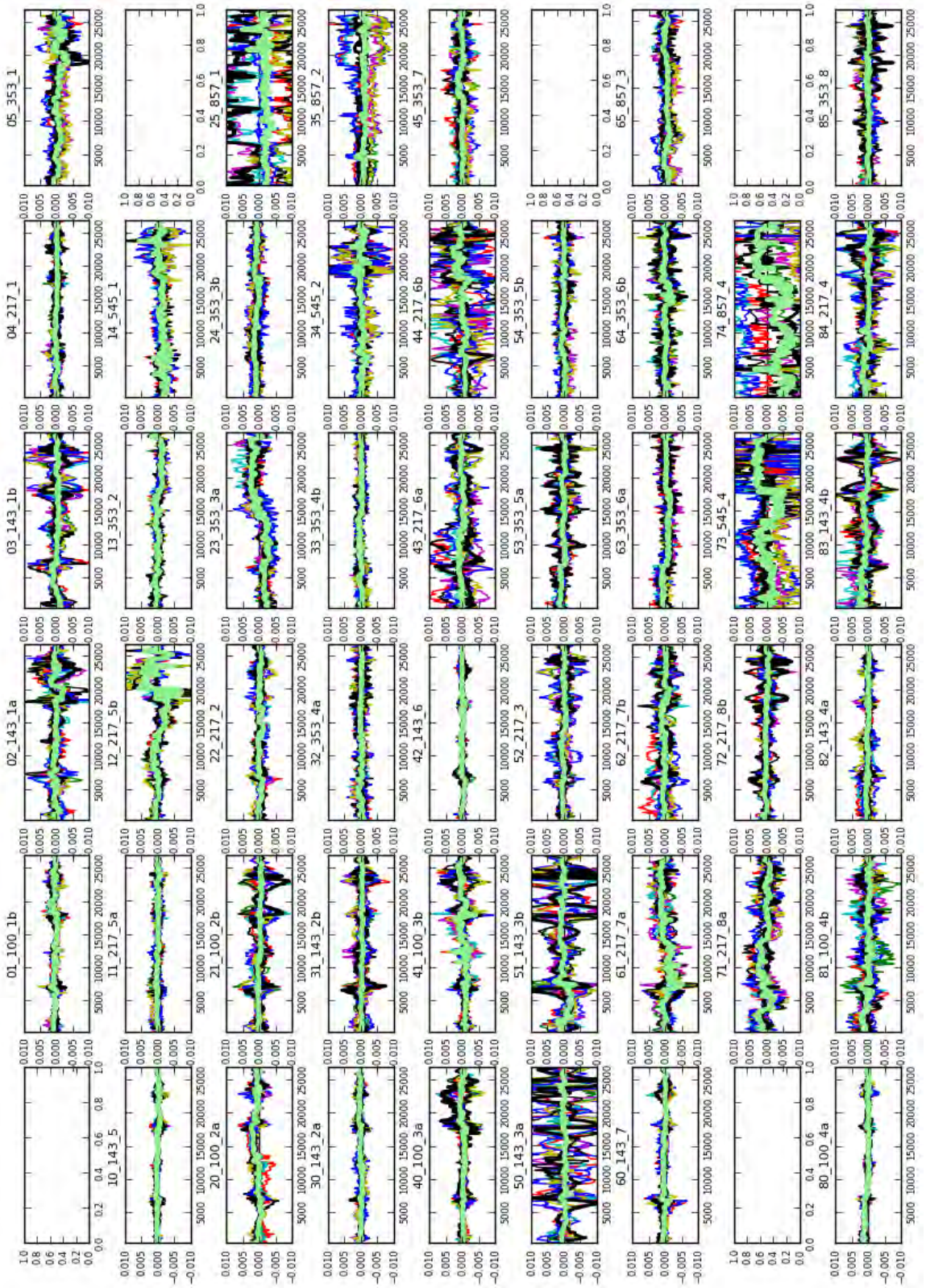


Figure 6.33: Half parity 9-gain observables for all *HFI* detector channels (2013 correction). The presentation is the same as for Fig. 6.32.

harmonic parameters as defined in Sec. 6.3.3, and the distribution $\mathcal{H}(\boldsymbol{x})$ is considered for a given ring index.

The algorithm requires a likelihood function $f(\boldsymbol{x})$ proportional to $\mathcal{H}(\boldsymbol{x})$ to build the target distribution. This likelihood function can be built by using the *half parity sum* residuals, and the two kind of half parity gain observables for which the expected value is 1. Putting all three data elements together, the error function at ring index r is

$$R(\boldsymbol{x})|_r = \underbrace{\sum_{j=0}^{39} \frac{(\hat{\boldsymbol{p}}(j) - (2 \times o + 2 \times \boldsymbol{m}_e))^2}{2\sigma^2}}_{\text{fast samples}} + \underbrace{\frac{(g-1)^2}{\sigma_g^2} + \sum_{k=0}^{9} \frac{(g_k - 1)^2}{3\sigma_g^2}}_{\text{science data}}, \tag{6.14}$$

where σ_g^2 is the variance of the half parity gain observable g, and g_k is the the half parity 9-gain. The value of σ_g as a function of r has been estimated with a Monte Carlo method, for an example see Fig. 5.19 in Sec. 5.5.3. The likelihood function is defined as

$$f(\boldsymbol{x}) = e^{-R(\boldsymbol{x})}, \tag{6.15}$$

with the assumption that $f(\boldsymbol{x}) \propto \mathcal{H}(\boldsymbol{x})$.

The target PDF $\mathcal{H}(\boldsymbol{x})$ for $\boldsymbol{x} \in \Omega$ is built by random exploration with an iterative procedure that will explore preferably the locations of maximum likelihood. An initial value $\boldsymbol{x}^0$ is drawn from the subset Ω^{fold} which contains all the possible values for $\boldsymbol{x}$ folding to the 4 K lines science data Fourier coefficients $\boldsymbol{y}'$. In other words $\boldsymbol{x}$ is a solution of Eq. 6.9. The steps for generating a Markov chain transition $\boldsymbol{x}^i \to \boldsymbol{x}^{i+1}$ are:

1. Candidate generation: a candidate $\boldsymbol{x}^{i+1}$ is drawn from Ω^{fold} with the PDF $g(\boldsymbol{x}^{i+1}|\boldsymbol{x}^i)$ which is described in Sec. 6.4.1.

2. acceptance ratio: the candidate $\boldsymbol{x}^{i+1}$ is checked against the acceptance ratio $\alpha = \mathcal{H}(\boldsymbol{x}^{i+1})/\mathcal{H}(\boldsymbol{x}^i)$. As f is proportional to $\mathcal{H}$, the acceptance ratio is calculated with $\alpha = f(\boldsymbol{x}^{i+1})/f(\boldsymbol{x}^i)$;

3. acceptation or rejection: if $\alpha \geq 1$, the new candidate is more likely than $\boldsymbol{x}$ and is automatically accepted. Otherwise it is only accepted if $\alpha > u$, where u is drawn from $\mathcal{U}(0,1)$ the uniform distribution on $(0,1)$. In case of rejection the algorithm returns to Candidate generation.

4. add to chain: if $\boldsymbol{x}^{i+1}$ is accepted, it is added to the chain and a new candidate is generated.

In theory the distribution of the elements of the chain converges toward $\mathcal{H}(\boldsymbol{x})$.

If the candidate generation is done carefully the chain transition $\boldsymbol{x}^i \to \boldsymbol{x}^{i+1}$ is reversible by satisfying the condition

$$\mathcal{H}(\boldsymbol{x}^{i+1}) g(\boldsymbol{x}^i|\boldsymbol{x}^{i+1}) = \mathcal{H}(\boldsymbol{x}^i) g(\boldsymbol{x}^{i+1}|\boldsymbol{x}^i). \tag{6.16}$$

However it cannot be strictly true when enforcing the constraint of folded 4 K lines harmonics in Eq. 6.9. This is due to the degeneracy between $\boldsymbol{x}$ and the estimations made from science data (the folded harmonics $\boldsymbol{y}'$ and the modulation offset o). In practice the analysis of the MCMCs in the next section shows that this degeneracy is very low and the condition in Eq. 6.16 will be considered to be fulfilled.

Generation of transitions with folding constraints

In an early version transitions were generated with relaxed priors (constraints on the parameters), using for candidate transitions the PDF

$$g(\boldsymbol{x}^{i+1}(j)\,|\,\boldsymbol{x}^{i}(j)) = \mathcal{N}_{\sigma_j}(\boldsymbol{x}^{i+1}(j) - \boldsymbol{x}^{i}(j))\,, \tag{6.17}$$

where $j \in H_{4\mathrm{K}}$ is the harmonic index and $\mathcal{N}_{\sigma_j}$ is the normal distribution centered on zero with a standard deviation σ_j set to the CPV 2 harmonic amplitude.

This first approach led to a global acceptation ratio lower that 10^{-3}. A correct acceptation ratio should be between 0.2 and 0.5, meaning that the algorithm was unable to explore Ω^{fold} in a reasonable amount of time.

To overcome this issue a new version was drawing only transitions falling into Ω^{fold}. The principle is to start with $\Delta\boldsymbol{x}_1^i$ drawn from $g(\boldsymbol{x}^{i+1}|\boldsymbol{x}^i)$. Now the problem to solve is that in the general case $G(\boldsymbol{x}^i + \Delta\boldsymbol{x}_1^i) \neq \boldsymbol{y}'$, i.e. the transition does not fall into Ω^{fold}. The solution adopted is to complement the first transition with an additional one $\Delta\boldsymbol{x}_c^i$ that cancels the change in the folded value, meaning that it must fulfill the relation

$$G(\Delta\boldsymbol{x}_c^i) = -G(\Delta\boldsymbol{x}_1^i)\,. \tag{6.18}$$

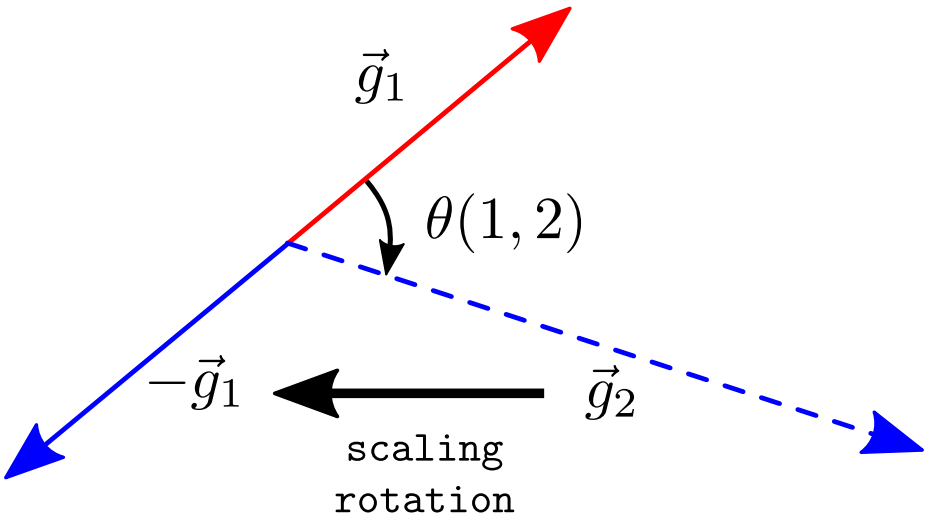


Figure 6.34: Schematic view of the operations producing the folded vector complement for the transition. For one given folding frequency, the vector $-\vec{g}_1$ is obtained by scaling and rotating the harmonics folding to $\vec{g}_2$.

The requirements for $\Delta\boldsymbol{x}_c^i$ can be met easily by applying scaling and rotation on a new transition $\Delta\boldsymbol{x}_2^i$ generated with the same PDF as $\Delta\boldsymbol{x}_1^i$. The next steps are run separately on each of the four folding frequencies. Let us consider the subset of harmonics folding to 20 Hz and use the vector notation $\vec{g}_1 = G(\Delta\boldsymbol{x}_1^i)|_{20\,\mathrm{Hz}}$ and $\vec{g}_2 = G(\Delta\boldsymbol{x}_2^i)|_{20\,\mathrm{Hz}}$ ($(\vec{g}_1, \vec{g}_2) \in \mathbb{C}^2$). The following operations are applied on $\Delta\boldsymbol{x}_2^i|_{20\mathrm{Hz}}$ (as shown in Fig. 6.34):

- NORMALIZATION: the modulus of the vector subset is adjusted by applying the factor $\|\vec{g}_1\|/\|\vec{g}_2\|$ to the harmonics folding to $\vec{g}_2$;
- ANTI-ALIGNMENT: if $\theta(1,2)$ is the angle between $\vec{g}_1$ and $\vec{g}_2$, then the two associated vectors can be anti-aligned by applying the rotation $e^{j\pi-\theta(1,2)}$ to each of the harmonics of $\Delta\,\boldsymbol{x}_2^i|_{20\,\mathrm{Hz}}$. For the negative folding frequencies $\nu \in [2*n*f_{\mathrm{mod}}, (2*n+1)*f_{\mathrm{mod}}]$, the sign has to be inverted to obtain the correct anti-alignement of both folded vectors.

The transition complement $\Delta \boldsymbol{x}_c^i$ is obtained by applying the same operations with the three remaining folding frequencies (40 Hz, 60 Hz and 80 Hz). Finally the expression for the new chain candidate is

$$\boldsymbol{x}^{i+1} = \boldsymbol{x}^i + \frac{\Delta \boldsymbol{x}_1^i + \Delta \boldsymbol{x}_c^i}{\sqrt{2}}, \tag{6.19}$$

where the $\sqrt{2}$ in the denominator is a normalization factor accounting for the summation of two variables drawn from g.

It is important to note that while the folded sum is zero, the harmonic sum $\Delta \boldsymbol{x}_1^i + \Delta \boldsymbol{x}_c^i$ is not zero in the general case. An example histogram of the transitions is shown in Fig. 6.35, the amplitudes of the generated Fourier coefficients follow the expected χ^2 distribution and the phases follow a random uniform distribution as expected. For all the transitions the sum of the folded harmonics is zero.

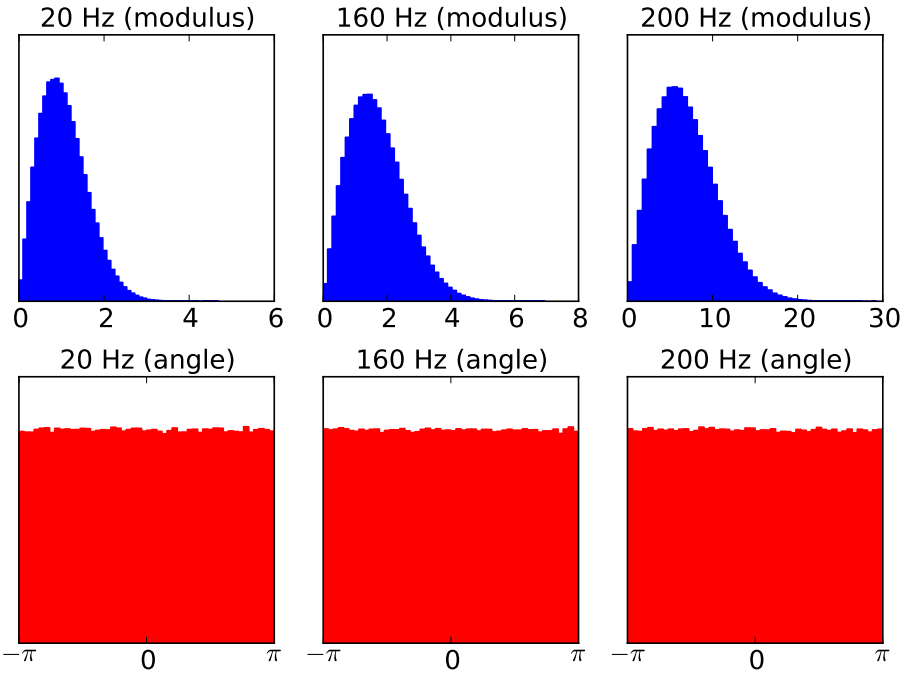


Figure 6.35: PDF of the generated harmonics for $\Delta \boldsymbol{x}_1^i + \Delta \boldsymbol{x}_c^i$ folding to 20 Hz, with relative values for σ_j taken from the CPV 2 estimates on channel `50_143-3a`.

As Ω^{fold} is a hyperplane of the parameter space Ω, an additional prior must be set on the maximum length of the harmonics to avoid potential drift to infinity. It is implemented in a simple way at the rejection step.

Analysis of the MCMCs

The result of the MCMC analysis is detailed below for several configurations.

Burn in: A typical MCMC produced with the MH algorithm starting from a random point far from the $\mathcal{H}(\boldsymbol{x})$ peak location and then first elements of the chain must be discarded before the MCMC stabilize into $\mathcal{H}(\boldsymbol{x})$. This is called the *burn in* period, an example is shown in Fig. 6.36. With this setup a typical *burn in* duration is 200 chain elements, but for safety 1000 elements are dropped. As can be seen in the figure, the χ^2 value obtained by normalizing the residuals $R(\boldsymbol{x})$ with the number of data samples (40+1+9) is about 4 instead of being close to one. This behaviour is due to the half parity 9-gain observable g_k that is significantly higher than its expected noise value. It is an important result showing that **something does not properly work with either the 4 K lines fitting or the**

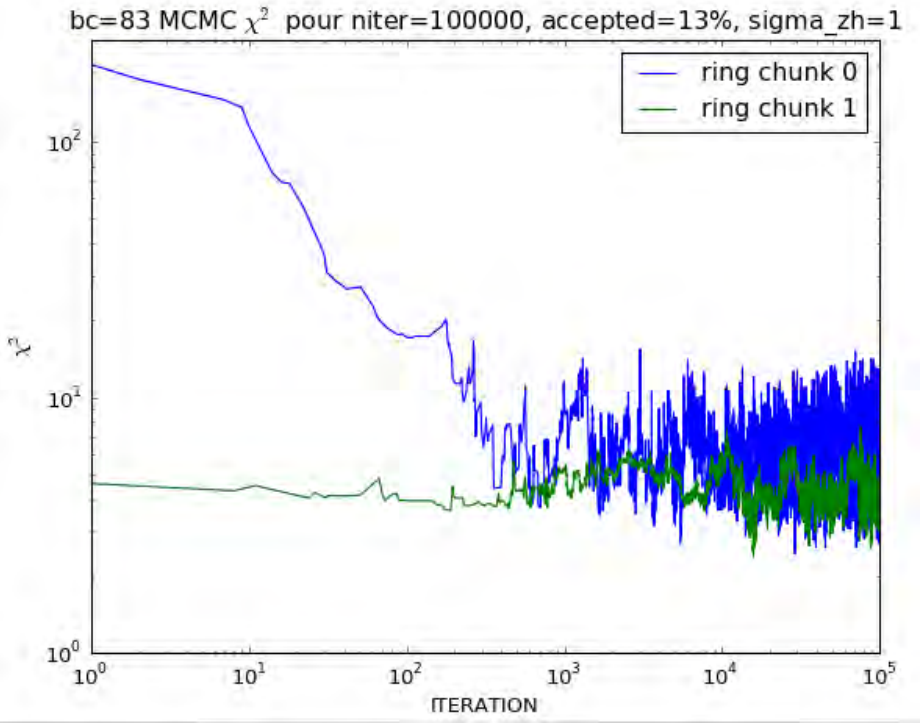


Figure 6.36: Burn in period of the MCMC. Two examples of run are shown starting from different random seeds, the second one (green line) starts clearly in a location where the chain is already stable. The Y axis is the output value of the likelihood function $f(\boldsymbol{x})$.

signal model itself. Potential candidates for the bad behavior of g_k are the 4 K lines popcorn (sudden jumps in the 4 K lines signal on timescale lower than ring level) and their modulation by the sorption cooler bed switching.

PRIOR ON HARMONIC AMPLITUDE: In the first runs a limit is set on harmonic amplitudes at 5 times the CPV 2 estimate, see Fig. 6.37 for an example. This constraint allows for a stable convergence of the chain over several runs with different initial random seeds. The 80 Hz harmonic of the example (the rightmost) is special in the sense that there is only one harmonic folding at this frequency, thus it shows only some variation in a 10^{-2} LSB range. The variations of the harmonic with only one folding frequency is representative of the impact of re-estimating the science data folded value, which is hopefully insignificant. However the constraint in the amplitude size is clearly reached for several frequencies for which the signal power is lower. It can be seen in the 2D likelihood (real, imaginary) from the distribution shape forming a circle. Hence it indicates that this frequency provides no information on $\mathcal{H}(\boldsymbol{x})$, and that most probably the overall result is biased.

RELAXING THE PRIOR ON HARMONIC AMPLITUDES: The *circle shape* of some distributions seen in the previous run shows clearly that the limits on the harmonic amplitudes are too tight. This time the MH algorithm was run with a relaxed prior on the amplitude (max = 100 LSB), an example of which is shown in Fig. 6.38. With this setup there is no convergence of the chain to a unique distribution for $\mathcal{H}(\boldsymbol{x})$ over several runs. Moreover for each run the distributions of harmonics are multimodal as it can be seen in the figure. Altough the demonstration channel has the strongest level of harmonics, this demonstrates that with this model and setup a unique solution cannot be found with too many harmonics, because of degeneracies between them.

LIMITED NUMBER OF HARMONICS: Results are much better with a 3:2:1:1 selection of harmonics with 3 folding to 20 Hz and two folding to 80 Hz. Fig. 6.39 shows an example run for a channel chosen because the level of folded harmonics at 20 Hz is equivalent to the level

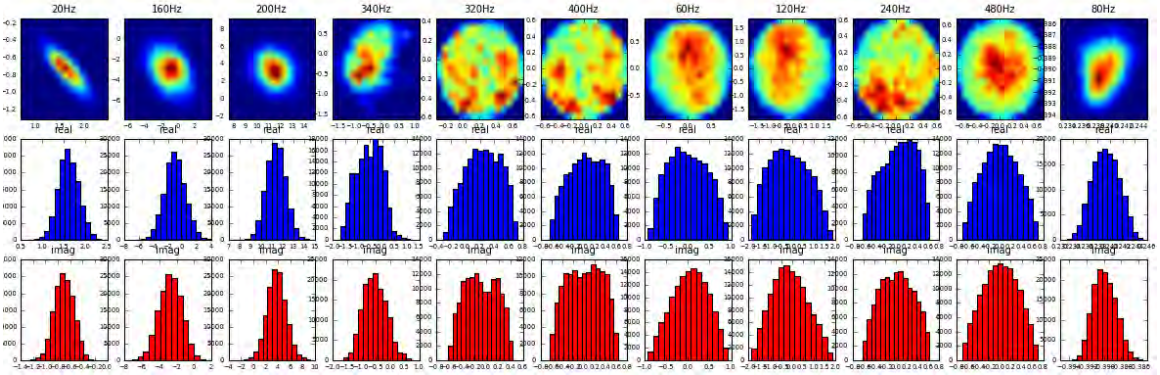


Figure 6.37: Likelihood of harmonic values with tight prior on amplitudes for ring 240 of channel `50_143-3a`. The likelihood of each fitted harmonic is shown in a 2D plot (upper row) with color scale ranging from blue (zero) to red at peak. The second and third row show the corresponding histogram of the elements of the Markov chain for the real and imaginary part respectively. The harmonic frequencies are in order (grouped by folding value): [20, 160, 200, 340]→20 Hz, [320, 400]→40 Hz, [60, 120, 240,480]→60 Hz, [80]→80 Hz.

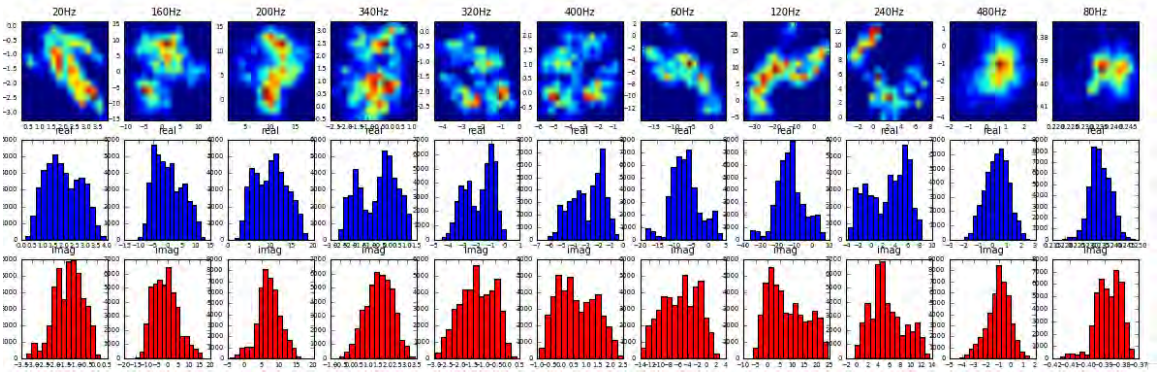


Figure 6.38: Likelihood of harmonics values with relaxed prior on amplitudes. The ring, channel, and frequencies are the same as for Fig. 6.37.

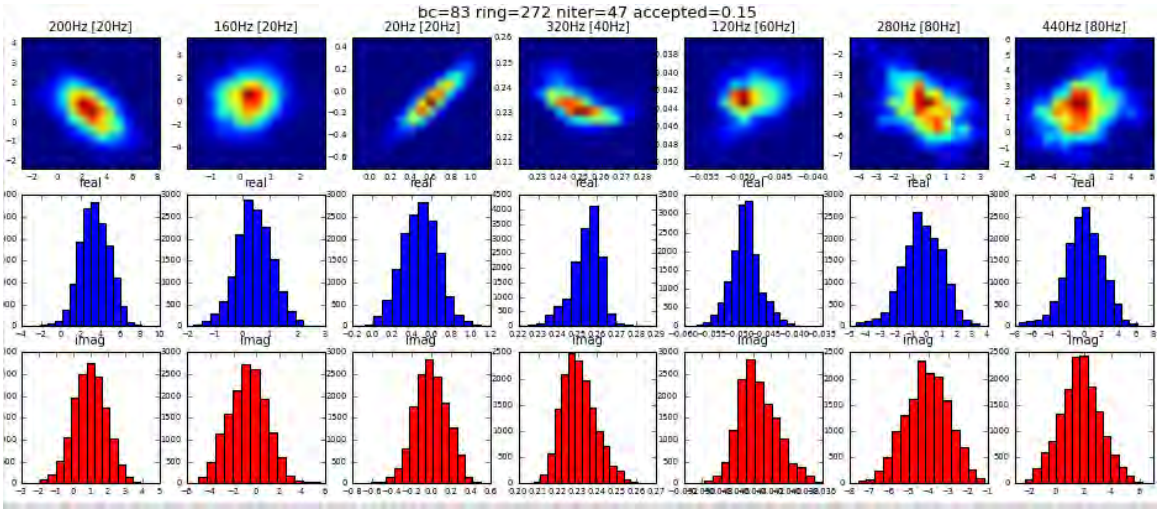


Figure 6.39: Likelihood with only 7 harmonics. The harmonic frequency are in order (grouped by folding value): [200, 160, 20]→20 Hz, [320]→40 Hz, [120]→60 Hz, [280,440]→80 Hz

at 80 Hz. It appears as a good candidate for checking if an improved handling of a second folding frequency yields improvements. With this limited set of harmonics the MCMCs do converge to the same distributions, but the solutions found do not result in a significant improvement in the ADC nonlinearity observables over the mission. However the analysis of *fast samples* periods residuals after cleaning the estimated 4 K lines harmonics reveals a striking consistency over the mission, this consistency is discussed in the next section.

Stability of the *raw constant* over the mission

The MCMCs performed with a limited set of harmonics all share the same property: the half parity sum residuals (left part of Eq. 6.14) for the value of $\boldsymbol{x}$ with the highest likelihood are very consistent over the mission. This is shown for three channels on Fig. 6.40. For the channel `25_857-1` the big spike at the beginning is due to a dramatic variation of the even modulation harmonics m_e between the CPV sequences and the mission. This variation degenerates with the unfolded 4 K lines which explains the relaxation-like shape of the residuals.

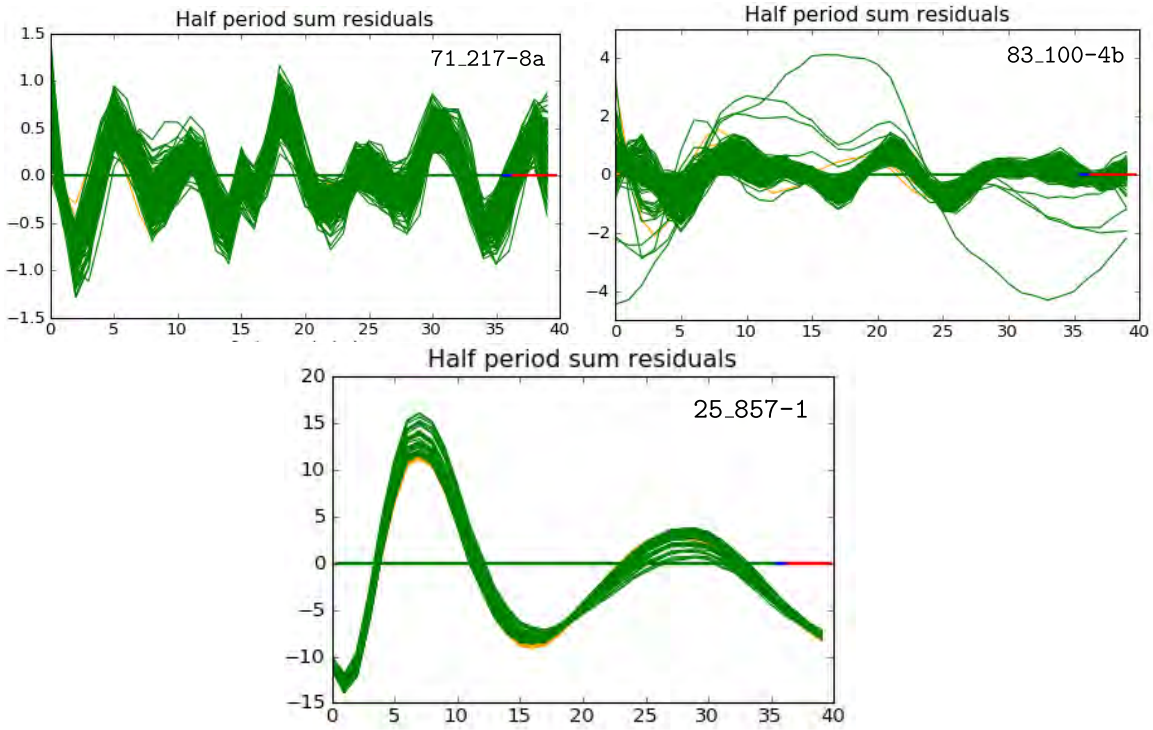


Figure 6.40: Consistency of half period sum residuals over the mission. For each channel, consecutive half-periods sum residuals are averaged and it is the 100 resulting averaged half-periods which are displayed

The consistency of the *raw constant* over the mission can still be seen in the *fast samples* period itself, albeit there is a bit more dispersion, it is clearly dominated by the 4 K lines signal. An example is shown on Fig 6.41 from the *raw constant* resulting after removal of the estimated 4 K lines. At the *fast samples* period level the sky power variation term $\delta P_{\text{sky}} \times \boldsymbol{g}$ and modulation offset o must also be estimated and subtracted over the entire mission to achieve this consistency. This consistency has not been seen before because of a residual level of unfolded 4 K lines which was too high.

This result is a strong argument in favor of the **stability of the *raw constant* term**, and this is the base hypothesis for the generation of a new parameter set. It is detailed in the next section.

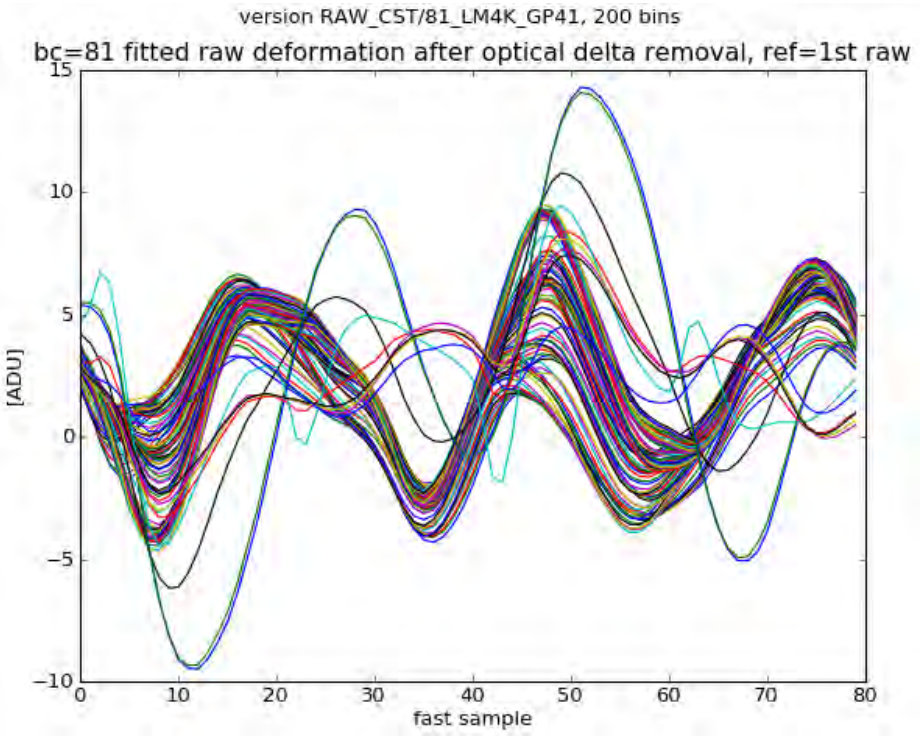


Figure 6.41: Consistency of the *raw constant* over the mission for the channel `81_100-4a` after removal of the 4 K lines and the sky power variations $\delta P_{\rm sky} \times \boldsymbol{g}$ over the mission.

6.4.2 The global *raw constant* method

Global *raw constant* model

The consistency of the residuals of the *half parity sum* are strong indicators of the stability of the modulation shape over the mission, but the modulation offset has also to be taken into account. It is made the hypothesis that the modulation offset o_r for ring r, albeit varying all over the mission, is a constant at the *fast samples* period level. Hence the *raw constant* parameter can be written

$$\boldsymbol{c}_r = o_r + \overline{\boldsymbol{c}}\,, \tag{6.20}$$

where $\overline{\boldsymbol{c}}$ is called the global *raw constant* parameter.

The expression for the science data transfer function remains the same as for Eq. 6.3.

Estimation of the global *raw constant* parameter

The global *raw constant* parameter $\overline{\boldsymbol{c}}$ is estimated by minimizing the half parity gain observables all over the mission. It does not rely any more on the biased template from CPV, thus the *fast samples* are not used for this purpose but are still used in the estimation of 4 K lines independently.

The residuals function of $\overline{\boldsymbol{c}}$ on science data for a given ring reads

$$R(\overline{\boldsymbol{c}})|_{r,\boldsymbol{x}} = \frac{(g-1)^2}{\sigma_g} + \sum_{k=0}^{9} \frac{(g_k-1)^2}{3\sigma_g^2}\,, \tag{6.21}$$

where the unfolded 4 K lines harmonics $\boldsymbol{x}$ are fixed (see below). The quantity that is to be minimized is $\sum_r R(\overline{\boldsymbol{c}})^2$ for rings over the first five *HFI* surveys and not flagged as bad or resonant. To allow flexible modification of the shape of $\overline{\boldsymbol{c}}$ without getting too much non

physical, the first 13 modulation harmonics are used, corresponding to 26 free parameters in Fourier space. In practice the fit starts with the initial value $\overline{\boldsymbol{c}}^0 = \langle \boldsymbol{c}_r - o_r \rangle$ and converges in five iterations with an implementation based on `mpfit`[3].

It has already been discussed in Sec. 6.2.3 with the full mission simulation, that the 4 K lines harmonics are a second order parameter of the science data transfer function. Hence the estimation of the harmonics and the global *raw constant* do not need necessarily to be done at the same time. Even with an imperfect set of harmonics the correction is much better with a correct *raw constant*. Based on this assertion two sets of parameters have been generated (see Tab. 6.2), labeled `globcst-h3` and `globcst-merged`. For the first set the global *raw constant* is estimated using for $\boldsymbol{x}$ with the the unfolded 4 K lines harmonics estimates of 2013 (three per channel, hence the "h3"), while the second set uses new estimates built by Guillaume Patanchon in mid 2015.

	`globcst-h3`	`globcst-merged`
INL	GP41	GP41
harms set $H_{4\mathrm{K}}$	$\underbrace{1:8:10}_{\to 20\,\mathrm{Hz}}$	$\underbrace{1:8:10}_{\to 20\,\mathrm{Hz}}, \underbrace{6}_{\to 60\,\mathrm{Hz}}, \underbrace{4:14}_{\to 80\,\mathrm{Hz}}$
channels processed	ALL	02, 03, 30, 31, 50, 51, 81, 82, 83

Table 6.2: Global *raw constant* parameter sets produced in 2015. The harms set indicates the unfolded harmonics indices that are estimated on a per ring basis.

Let say a few words about the harmonics estimated by Guillaume. His method does also a minimization of half parity observables but does not rely on the biased templates from CPV. It uses a custom gradient descent algorithm with a marginalization on the folded 4 K lines values. The parameters produced with his processing does a better job than the 2013 correction to reduce the half parity gain discrepancy all over the mission. However the result produced is slightly non physical because the modulation offsets calculated with the resulting *raw constants* drift from the science data estimations by about 0.5 LSB. The so called merged parameters dataset has been produced with his harmonics estimates but using the global *raw constant* produced starting from the *raw constant* he produced. Doing so improve significantly the half parity gain observables.

Results

The results obtained with both global *raw constant* parameter sets are very similar and provide a dramatic reduction in the ADC nonlinearity measured with the half parity gain, and a moderate improvement for the half parity 9-gain observable. An example is shown in Fig. 6.43 for a channel on which the 4 K lines handling was critical. After application of the ADC nonlinearity correction using the global *raw constant* parameter set, the observables are very close to the estimated noise level.

The correlation plots between the 4 K lines level and the averaged half parity observables provide valuable information. Booth parameter sets are gathered in Fig. 6.44, their respective performance in the reduction of observables are very close. The half parity gain values are very compact and all close to the expected noise value, but not below, which is a good point because it would otherwise indicates that some over-fit happened.

The half parity 9-gain correlation shows a higher level of correlation between the scattering of the 18 skies and the 4 K lines level than the 2013 data release (see Fig. 6.31). This point deserves a specific discussion. At the time when this manuscript is written, the 18

[3] `mpfit` perform Levenberg-Marquardt least-squares minimization, based on MINPACK-1. The version used is the python implementation available from `https://code.google.com/archive/p/astrolibpy/downloads`.

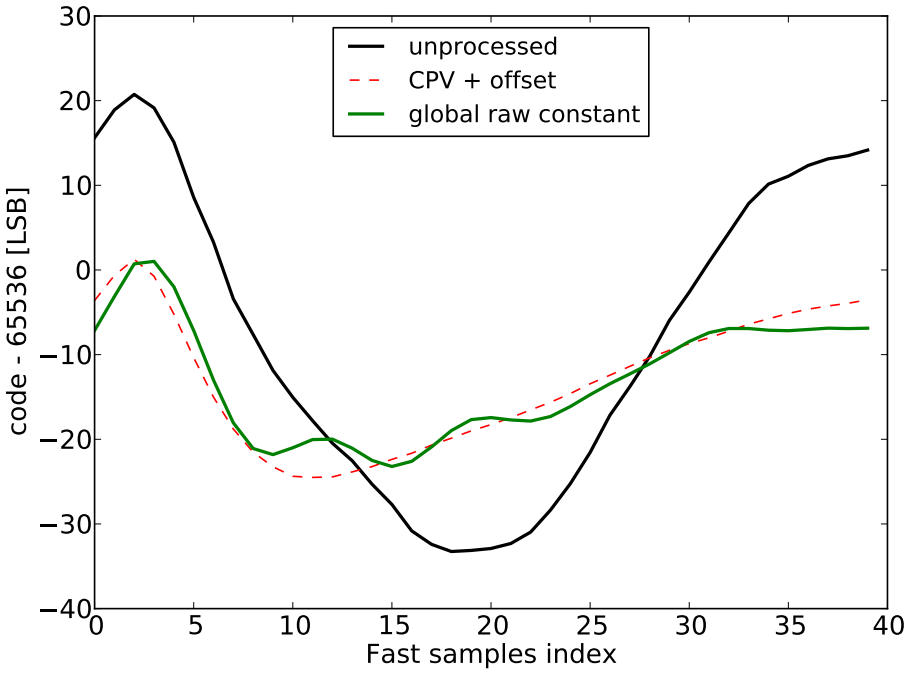


Figure 6.42: *half parity sum* of the global *raw constant* (version `globcst-merged`) for channel `03_143-1b` with the offset of ring 240 (in green). The black line show the *half parity sum* of the averaged *fast samples* period without processing and polluted with 4 K lines. The dashed red line is the CPV template $m_e + o_{240}$ for comparison.

skies scattering has never been fully understood, and the best results obtained so far shows correlation between g_k and the 4 K lines level which is perfectly in line with the noise level of the observable. The very late occurrence of the global *raw constant* parameter set has limited my ability to investigate this point, but there is strong evidences that **the signal model has to be refined**. The modulation of the 4 K lines amplitude by the sorption cooler bed switching on a 15 minutes basis is very likely to be the cause of the limitation in the ADC nonlinearity correction efficiency.

6.5 Conclusions and perspectives

Inclusion of 4 K lines in the analog signal model brings a significant improvement to the ADC nonlinearity correction on most channels, but also worsen the observables for several channels. Using the half parity 9-gain observables it can be seen that the remaining ADC nonlinearity residuals are correlated to the 4 K lines level.

The latest ADC nonlinearity correction in its 2015 version shows a dramatic improvement of the ADC nonlinearity observables, but did not improve significantly the cross survey differences (jacknives in *Planck* nomenclature). This discrepancy between different survey can be due to uncertainty on INL, mostly the 64 code just after midscale. It can also be a consequence of an inaccurate analog signal model, which is physically motivated by the difficulty in reducing the half parity 9-gain observables.

A specific characterization of the impact of the sorption cooler bed switching should be done, to provide some answers to the question: “Can we reduce ADC nonlinearity residuals to noise level in order to reduce the bias on the sky at large angular scales?”.

Here are a few directions to do this:

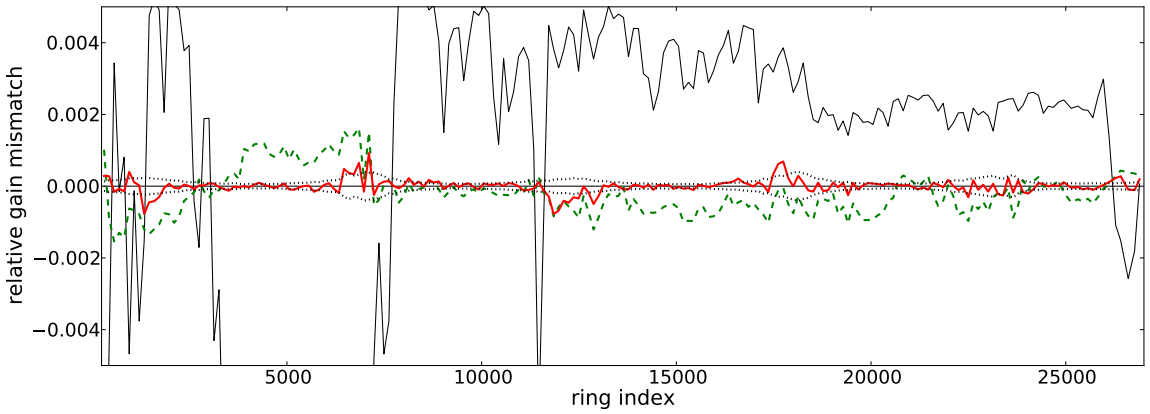


Figure 6.43: ADC nonlinearity residuals after correction with the global *raw constant* parameter set. The half parity gain observables are plotted with a red line for the global *raw constant* parameter set (version `globcst-merged`), with the dashed green line for the 2013 data release and the black line is for science data without correction. All observables are calculated with 128 rings binned PBRs and the black thin dashed line is the one σ extent.

1. CHARACTERIZATION ON SCIENCE DATA: A quick analysis can be made from science data by correlating the ADC nonlinearity effect with the temperature of the 4 K cooler. For example samples can be selected by their location within the up/down cycle of the τ_{bed} period.

2. CHARACTERIZATION WITH SIMULATIONS: A simulation of the modulation by the sorption cooler must be done at *fast samples* level. This tool exists and is maintained by Sylvain Mottet under the name of `DNLSIM`. It has been proposed to run a simulation with this tool but it could not be produced due to lack of time and resources. The setup should be to inject harmonics from the CPV sequence which are very well known, and apply a modulation using the 4 K stage temperature. Then the science data correction can be done with the average value of the injected harmonics and compared to the input sky.

3. CHARACTERIZATION WITH CPV DATA: The CPV data is more difficult to use because of the short acquisitions timescale, but still the CPV 2 offers several minutes of data for each channel which can be compared to the variations of the 4 K stage temperature. A second step could be to check if the cross channels correlation is better by taking the temperature of the 4 K stage into account, which may provide constraints for a model of variations.

4. TRANSFER FUNCTION UPDATE: the variation of the analog signal model can be difficult to implement or very CPU intensive in the case of ADC nonlinearity correction of real data. Hopefully it is very easy to implement a first order Taylor expansion of $\mathcal{F}|_{\boldsymbol{x}}$, because it is simple to develop an analytical expression for $\partial\mathcal{A}_\sigma(s)/\partial s$. Hence the calculation of $\partial\mathcal{F}/\partial\boldsymbol{x}$ can be done quickly, and allow the correction of real data in decent computing time.

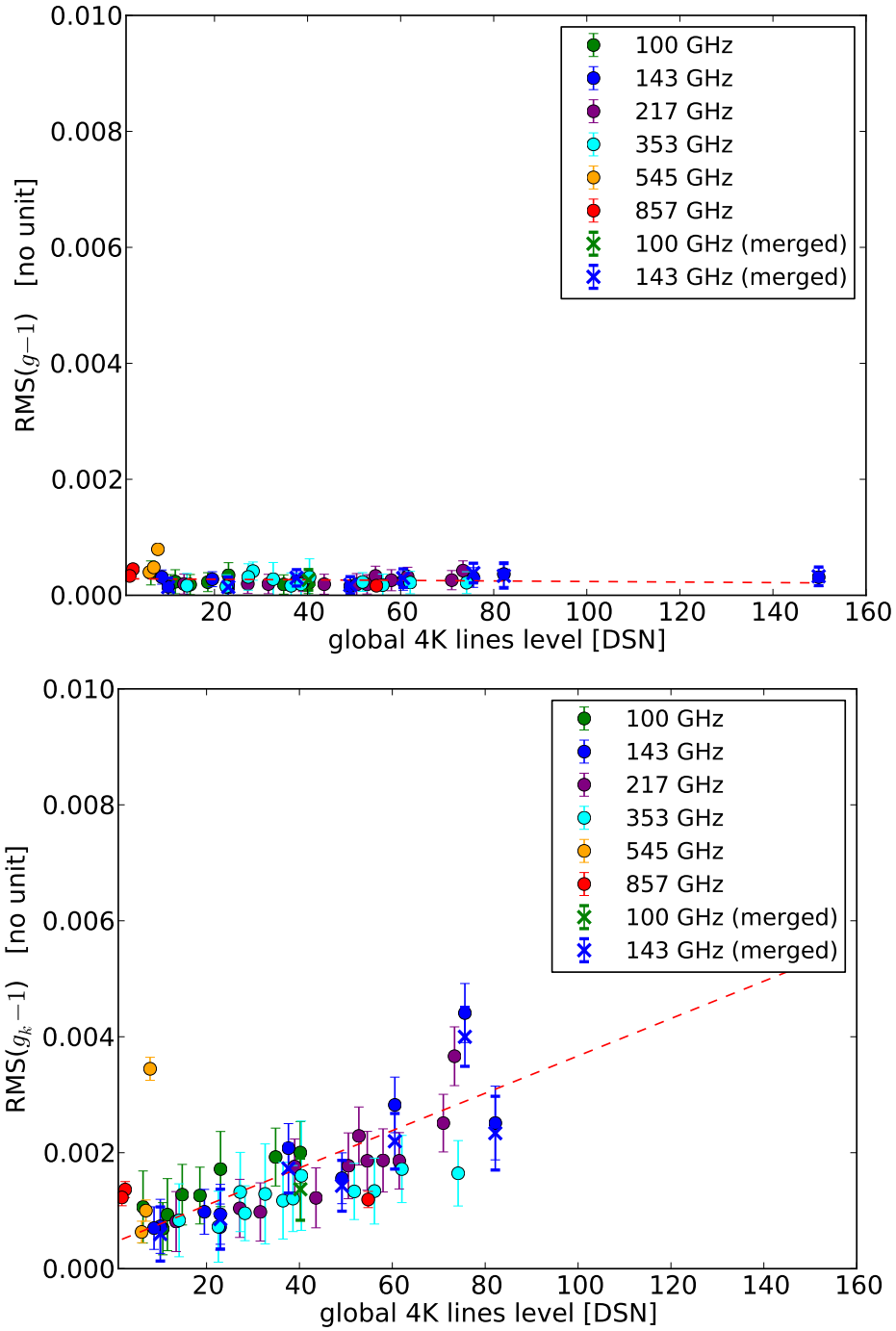


Figure 6.44: ADC nonlinearity observables calculated with the 2015 parameter set. The correlation plots uses the same presentation as for Fig. 6.2. Median value of the half parity observables is $\approx 2.3 \times 10^{-4}$ which is very close from the estimated noise level. *upper panel*: half parity gain observables. *Bottom panel*: half parity 9-gain observables.

Part III

Propagation of systematics on science

Chapter 7

Propagation of ADC nonlinearity to science

This chapter describes how the ADC nonlinearity systematic affects the scientific analysis of *HFI* data. It focuses mainly the thermal dust polarized emission and the CMB power spectrum. As stated before, an important issue is the absence of calibration sources that can be used to assess this effect on *HFI* science data at the required precision level, which is about 10^{-4} of the solar dipole amplitude. Nevertheless it is possible to infer reliable estimations of the effect with simulations using the model developed in Chap. 5. This model was implemented in the end-to-end simulations (see Sec. 5.4 in **?**), but different approaches are discussed below. To complement these simulations with observables from real data, the methodology based on the half parity observables described in Sec. 5.5.3. will be extended to full sky maps produced with a single parity of the modulation.

In Sec. 7.1 a simulation is used to study the impact of ADC nonlinearity on Galactic polarization at 353 GHz. Next, in Sec. 7.2 the half parity based methodology is used to infer how ADC nonlinearity residuals affect the CMB angular power spectra.

Some parts of the work presented below have been included in several sections of (**?**).

7.1 ADC impact on Galactic polarization

The simulation described below has been performed to infer quantitative estimation of the impact of the ADC on polarization angle at 353 GHz. This was for the paper on polarized thermal emission from Galactic dust (Planck Collaboration Int. XIX 2015).

7.1.1 Simulation

(I, Q, U) maps of polarization Stokes parameters were produced without ADC nonlinearity correction by the DPC. The principle of the simulation was to produce maps of ADC nonlinearity residuals $(\Delta I_{\rm ADC}, \Delta Q_{\rm ADC}, \Delta U_{\rm ADC})$ that could be used to infer ADC effect on polarization fraction and polarization angle.

The functional diagram of the simulation is shown in Fig. 7.1. As the ADC nonlinearity effect change with time, the simulation must be performed at the relevant TOI level. For this purpose the `Desire`[1] simulation software was used to produce the reference $\mathrm{TOI}_{\rm REF}$ science data TOIs of the simulation. The *Planck* sky model used for the Full Focal Plane version 6 simulations (FFP6) was used as the input of `Desire`. The ADC nonlinearity effect was applied using the function $\mathcal{F}$ in Eq. 5.12 to produce $\mathrm{TOI}_{\rm ADC}$. As stated in Sec. 1.4, a minimum of three detectors with different orientations are needed to produce the Stokes parameters, therefore the four detectors of the *HFI* 353 GHz detector set 1 (referred to as Detset-1= { `23_353-3a`, `24_353-3b`, `53_353-5a`, `54_353-5b`}) were used. Next, the two

[1] `Desire` is a tool developed at IRAP and written in `C++` dedicated to the simulation of *HFI* TOIs and housekeepings over the full mission

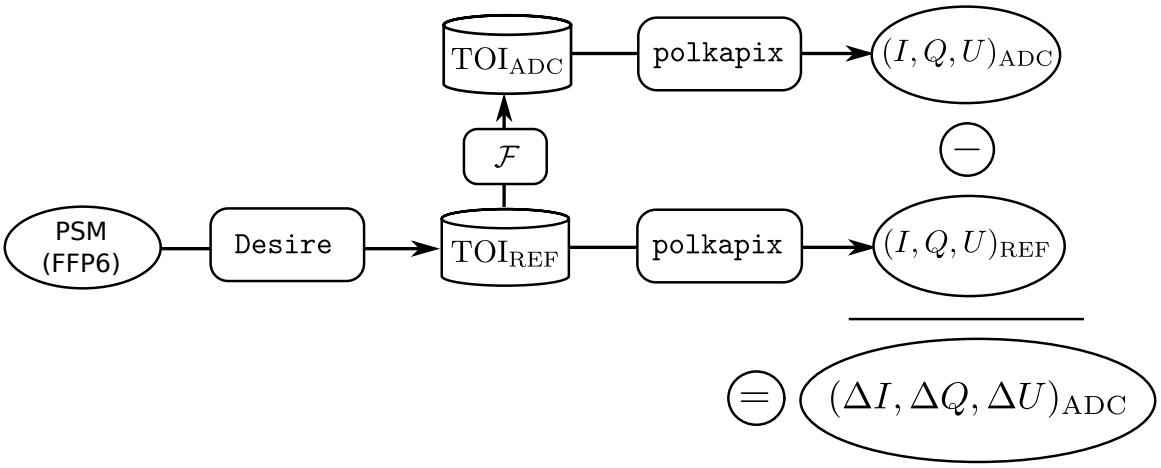


Figure 7.1: Functional diagram for the simulation of ADC nonlinearity residuals in polarization

sets of TOIs were processed with the official DPC software `polkapix` (Tristram et al. 2011) performing the destriping operation (removal of monopole) and the projection to sky maps. This processing yields two sets of Stokes parameter maps: $(I,Q,U)_{\rm REF}$ shown in Fig. 7.2 and $(I,Q,U)_{\rm ADC}$ shown in Fig. 7.3. From these maps the ADC nonlinearity residuals in polarization are calculated with

$$\begin{aligned} \Delta Q_{\rm ADC} &= Q_{\rm ADC} - Q_{\rm REF} \\ \Delta U_{\rm ADC} &= U_{\rm ADC} - U_{\rm REF}\,. \end{aligned} \tag{7.1}$$

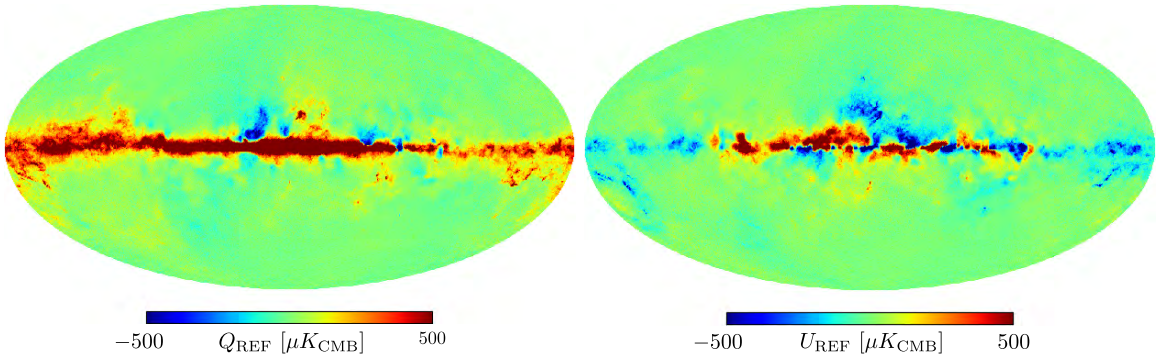


Figure 7.2: Simulation of polarization maps at 353 GHz for Detset-1. A 1° smoothing is applied on the Mollweide projection. The input sky of the simulation was the planck sky model used for the FFP6.

Two different patterns can be seen on maps $\Delta Q_{\rm ADC}$ and $\Delta U_{\rm ADC}$ depending on the emission of the component involved:

- the thermal dust signal in the Galactic plane has a high dynamic range which cross quickly the regions of the ADC yielding different gains. Thus the resulting pattern follows the Galactic plane shape with sharp sign changes in the Galactic plane;
- the solar dipole produce typically a different pattern over the sky because it has a smaller dynamic range than the dust. It is also dependent on the time of its acquisition, as the thermal variations of the bolometer plate over the mission change significantly the location of its acquisition on the ADC scale. This can be seen on the maps as diffuse patterns following the scanning strategy, and will project at low multipoles on the angular power spectrum.

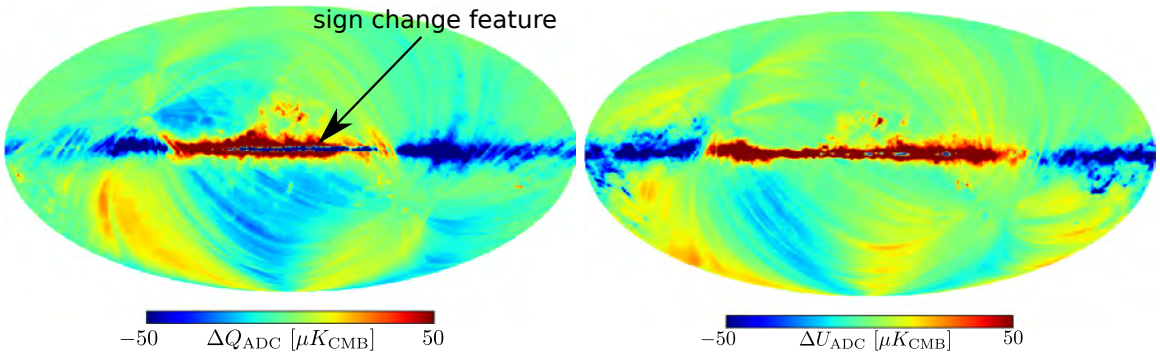


Figure 7.3: Simulation of ADC nonlinearity residuals on polarization maps at 353 GHz for Detset-1. A 1° smoothing is applied on the Mollweide projection. The sign change feature visible as a thin blue line at the center of the galactic plane is typical of ADC nonlinearity which produces step-like gain variation when the intensity of signal change

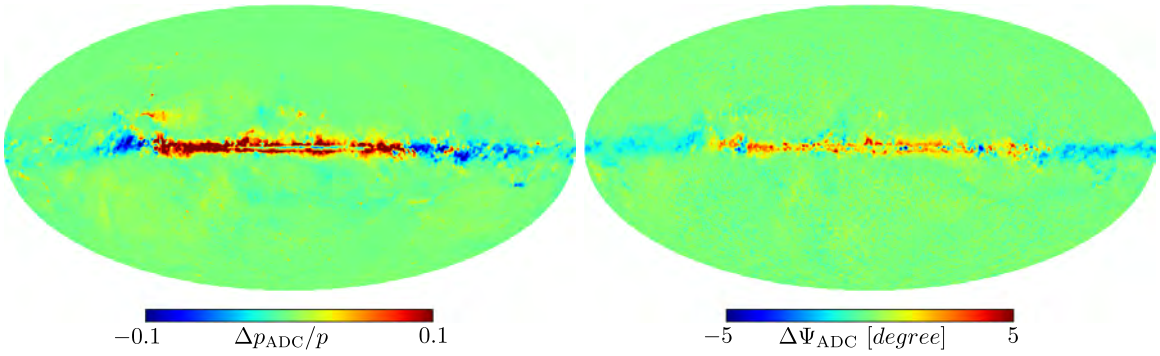


Figure 7.4: Simulation of ADC nonlinearity effect on polarization fraction (left) and angle (right) at 353 GHz for Detset-1. A 1° smoothing is applied on the Mollweide projection.

The propagation of ADC nonlinearity residuals to the polarization fraction p and angle Ψ can be calculated from Eq. 1.5, leading to the expression

$$\begin{aligned}\Delta p_{\rm ADC}/p &= \frac{\sqrt{Q_{\rm ADC}^2+U_{\rm ADC}^2}-\sqrt{Q_{\rm REF}^2+U_{\rm REF}^2}}{\sqrt{Q_{\rm REF}^2+U_{\rm REF}^2}}\\ \Delta\Psi_{\rm ADC} &= \frac{1}{2}\arctan\left(\frac{U_{\rm ADC}}{Q_{\rm ADC}}\right)-\frac{1}{2}\arctan\left(\frac{U_{\rm REF}}{Q_{\rm REF}}\right).\end{aligned} \tag{7.2}$$

Maps of the bias on these parameters are shown in Fig. 7.4, where it can be seen that the ADC nonlinearity biases by about 10% the polarization fraction p and by about 5° the polarization angle Ψ. Most of the effect is concentrated on the Galactic plane.

7.1.2 Degeneracy with the bandpass mismatch (BPM) effect

HFI bolometers are calibrated using the CMB orbital dipole, but each foreground component such as thermal dust has a different spectral density over the bandpass of the corresponding optical filter. Therefore a transmission correction coefficient (also called color correction) has to be applied on the foreground polarized emission when subtracted from the CMB polarized emission. As the optical filters of *HFI* detectors have a slightly different spectral response from each other, the correction is individual to a channel. This systematic effect, called the BPM, cause leakage of intensity into polarization (see Sec. 7.3.1 in Planck Collaboration

VIII 2016). In addition, BPM also exist between bolometer pairs of a given PSB because the path of the light reaching each bolometer through the common PSB filter is slightly different, due to the different location of the bolometers in the integrating cavity.

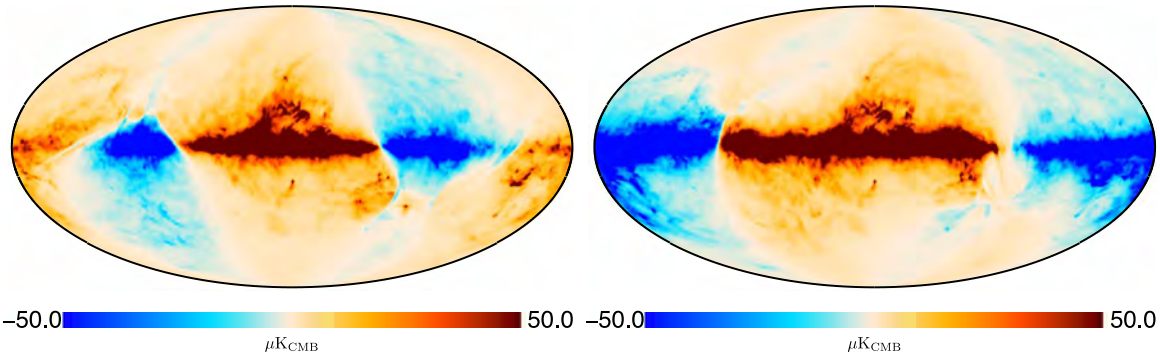


Figure 7.5: Dust leakage correction maps from ground-based measurements on 353 GHz channels. *Left*: leakage of intensity into Q. *Right*: leakage of intensity into U. Image credit: (Planck Collaboration VIII 2016)

It can be seen on the 353 GHz BPM maps of dust leakage into polarization in Fig. 7.5 that this systematic effect is strongly degenerated with the ADC nonlinearity effect being at a similar level, when comparing with Fig. 7.3. This can be explained by the ADC nonlinearity behaving like a gain variation which depend on signal intensity. The signal of the CMB signal can be acquired in a location on the ADC scale with a different gain than for the much brighter dust this will result in an effect similar to the BPM. Hence **BPM cannot be fitted simply on in-flight data in the presence of ADC nonlinearity residuals**, the sign change feature visible in the $\Delta Q_{\rm ADC}$ map in Fig. 7.3 is a good demonstration of this complexity.

Close to the galactic plane, and in particular in the inner regions, the magnetic field direction inferred from dust polarization is expected to be statistically aligned with the Galactic plane, and in practice the distribution of the polarization angles tends to achieve this property in the Galactic plane after application of BPM correction. As reported in Sec. 2.2 of Planck Collaboration Int. XIX (2015), it was observed that the early ADC nonlinearity correction of 2013 combined to BPM correction roughly improved this alignment.

7.2 ADC impact on CMB science

This section describes how the ADC nonlinearity impact the scientific analysis of the CMB angular power spectrum.

The inference of ADC nonlinearity residuals is performed with half-parity maps which are detailed in Sec. 7.2.1. These maps produced with a single parity of the modulation have specific noise properties which are studied to assess reliable estimations of the ADC impact on the TT power spectrum at low frequency. Next the analysis of the angular power spectra produced from these maps is performed in Sec. 7.2.2.

Some angular power spectrum calculated using half parity maps were provided content to (**?**, Sec. 5.2 in).

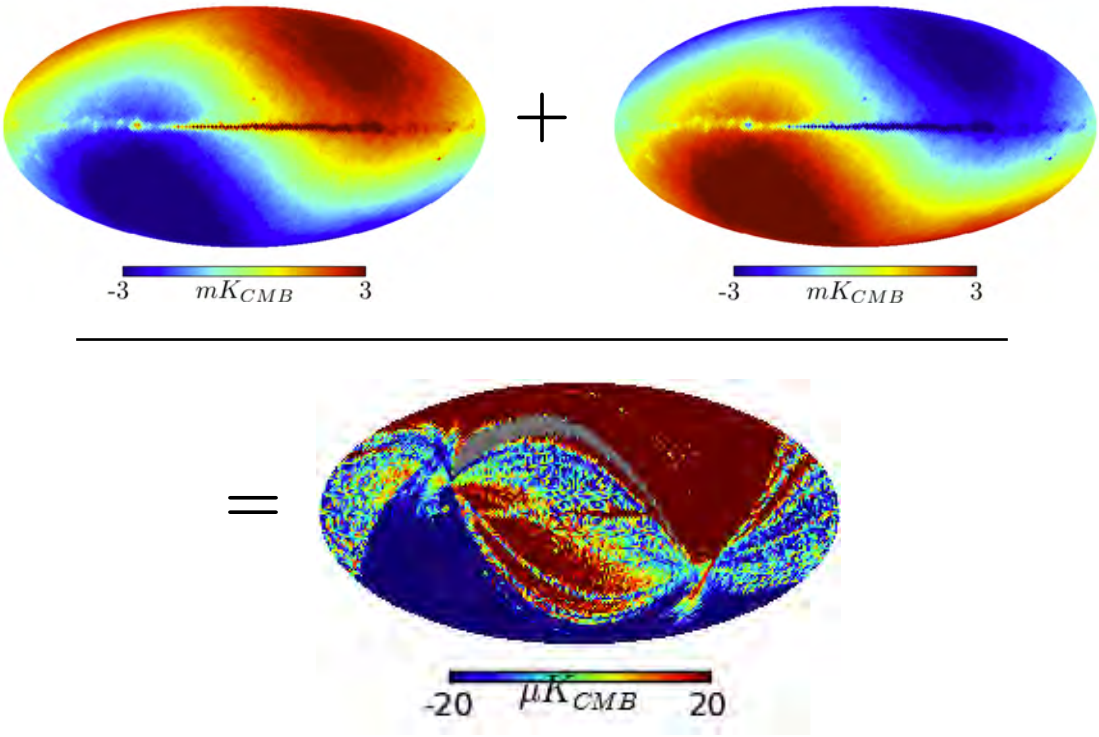


Figure 7.6: Half-parity difference map principle. The sum of maps produced from samples of single parity cancels the sky signal, leaving only noise and residuals expected to be caused dominantly by the ADC nonlinearity

7.2.1 Half-parity methodology

Principle

The target was to measure the ADC nonlinearity residuals on the CMB angular power spectrum. For this purpose, the half-parity gain observable developed in Sec. 5.5.3 were not sufficient to asses the ADC impact on the sky after the complex TOI processing. Instead of comparing PBRs from each parity the half-parity principle was extended to full sky maps. Fig. 7.6 show synthetically how this is was done: full sky MAP_0 and MAP_1 are produced with samples from single parity, then a Half-Parity Difference map is produced with

$$\mathrm{HPD} = \mathrm{MAP}_0 + \mathrm{MAP}_1 \,, \tag{7.3}$$

where both maps have different sign for the sky signal because of the modulation. Finally the angular power spectrum in intensity (TT spectrum) is calculated from the HPD map for analysis.

Production of HPD maps

The full *HFI* DPC pipeline cannot be used because of the selection of one sample over two. The resulting science data is down-sampled by a factor two, this would require a complete adaptation of several software including the time deconvolution and the deglitcher. Instead a pragmatic approach has been used working only with raw modulated TOI of science data. The following steps are performed:

- ADC NONLINEARITY CORRECTION: it is applied optionally on raw unprocessed science data TOIs;

- GLITCH FLAGGING: the official glitch flags produced by the deglitcher software are applied on TOIs. The effect is to drop the glitch head where most of the signal is present and to leave the low frequency tail present with *long* glitches. These tails are low frequency noise which is canceled by the half-parity processing, thus they are considered negligible.

- PROJECTION TO MAP: the `polkapix` software is run on the corresponding samples selected by parity. It produces the MAP_0 and MAP_1 sky maps corresponding to the science data signal from parity 0 and 1, respectively.

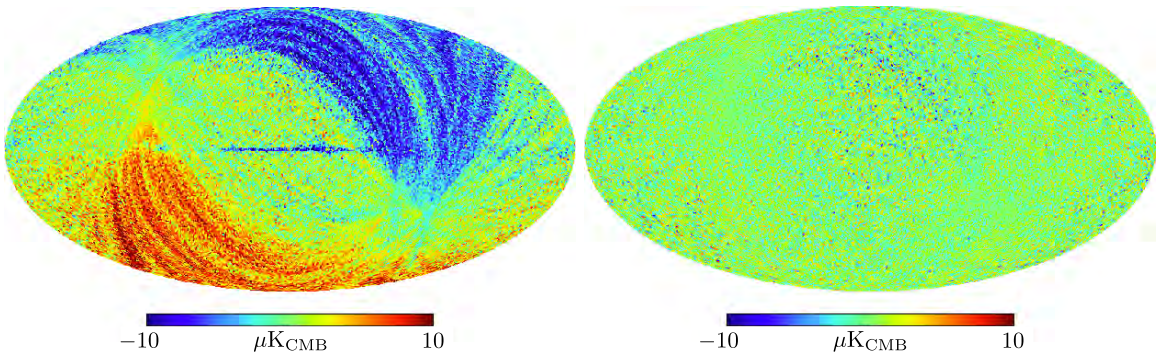


Figure 7.7: HPD maps for channel `10_143-5` with resolution downgraded by a a factor 8 (`HEALPix, nside`=256) to improve SNR of ADC nonlinearity residuals. *Left*: map produced from RAW data. *Right*: map produced from science data TOI processed with the ADC nonlinearity correction of the *Planck* 2015 data release

Fig. 7.7 shows an example of HPD maps for channel `10_143-5` before and after the ADC nonlinearity correction of the *Planck* 2015 data release. Before correction, there are large stripes following the scanning strategy which are caused by the solar dipole dynamic range over the ADC scale. After ADC nonlinearity correction, the stripes have nearly disappeared, this result is a strong validation of the ADC nonlinearity correction as it is based on a physically motivated model which is completely unrelated to the observables provided by the HPD maps.

Characterization of the ADC effect on the TT spectrum with a simulation

The angular power spectrum is calculated with `anafast`[2]. Fig. 7.8 shows the angular power auto-spectrum (TT spectrum) of the two HPD temperature maps previously shown in Fig. 7.7, There are several features of interest on this spectrum which are:

- a $1/\ell$ component at low multipole;

- a drastic reduction by one order of magnitude of the $1/\ell$ component after ADC nonlinearity correction;

- an "unexpected raising" of the spectrum above $\ell = 100$. This feature will be studied in Sec. 7.2.1.

The $1/\ell$ component is mostly expected from the ADC nonlinearity effect, but there is also a suspicious raising of the power spectrum above $\ell = 100$. To characterize accurately the effect of ADC nonlinearity over the TT spectrum, a simulation was run with the same

[2] The `anafast` software is included in the `HEALPix` package, it performs quick spherical harmonics decomposition and calculation of angular power spectra

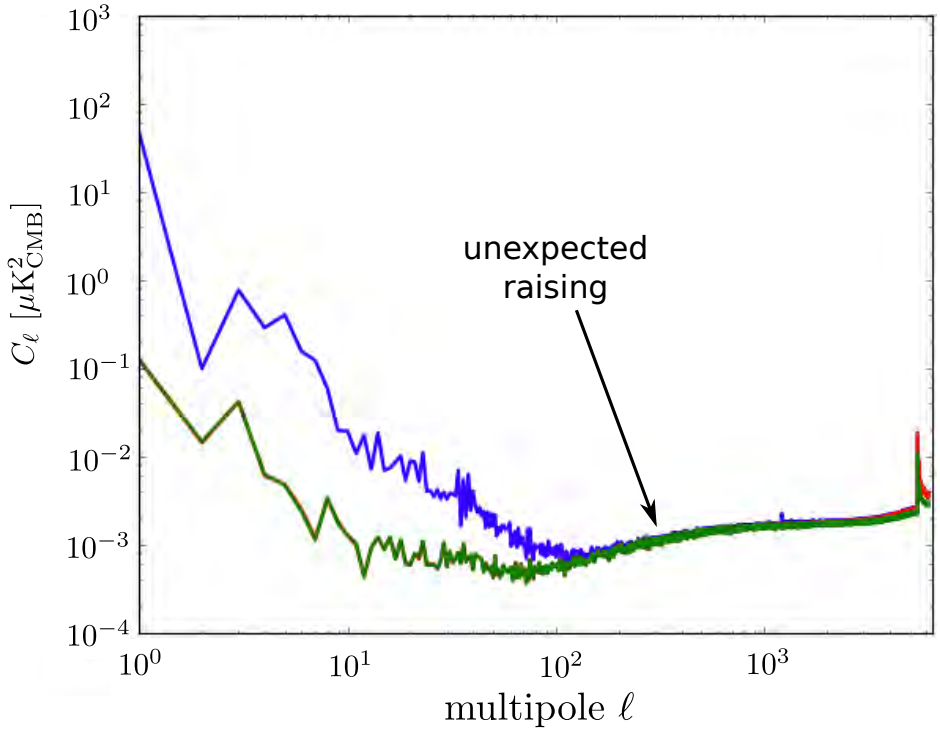


Figure 7.8: TT power spectrum calculated with the same HPD maps as in Fig. 7.7. The blue line is for HPD built from raw science data. The green line is for HPD built from science data processed with the ADC nonlinearity correction of the *Planck* 2015 data release. There is a small spike visible at $\ell = 1200$ which is due to the 20 Hz 4 K lines, and the spike at $\ell = 5400$ is a leakage of the signal modulation

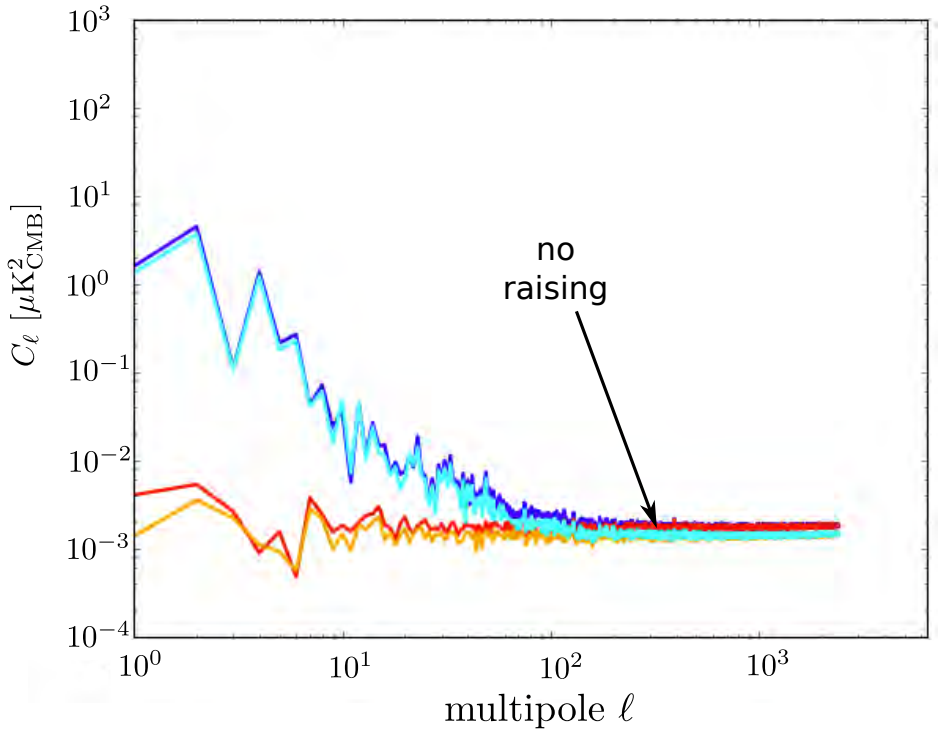


Figure 7.9: TT power spectrum calculated with **simulated** HPD maps. The dark-blue line is for the data with ADC nonlinearity effect. The red line for the nominal simulation. The cyan/orange lines are for TT spectra computed after masking 15% of the bright Galactic plane.

processing pipeline and glitch flags as for the real data. The pipeline was the same as the one shown in Fig. 7.1, with the addition of parity selection at the `polkapix` step. A $1/f$ noise component is present in the noise generated by `Desire` on science data TOIS, so it can be checked how it projects on the TT spectrum.

Fig. 7.9 shows the TT spectrum yielded by the simulation. First it can be noticed that the $1/\ell$ component in the ADC nonlinearity case has a shape and amplitude in very good agreement with the one observed on real data, while the spectrum in the ADC free case is very flat. This leads to the conclusions that:

- the HPD processing cancels efficiently the signal and the correlated noise component in the science data;
- the ADC nonlinearity effect is expected to show-up as a $1/\ell$ component at low frequency on then HPD TT spectrum.

Additionally the impact of the high dynamic range of the signal in the Galactic plane has been checked by masking 15% of the map in the brightest region of the Galactic plane and including point sources. The masking resulted in about 15% decrease in the power of the TT spectrum, thus demonstrating that there was no significant signal leak into the HPD map at this frequency[3].

However the unexpected raising feature remains unexplained by this approach and needs further characterization to fully understand the noise profile of the TT spectrum produced with HPD maps.

Multi-component Noise model for science data

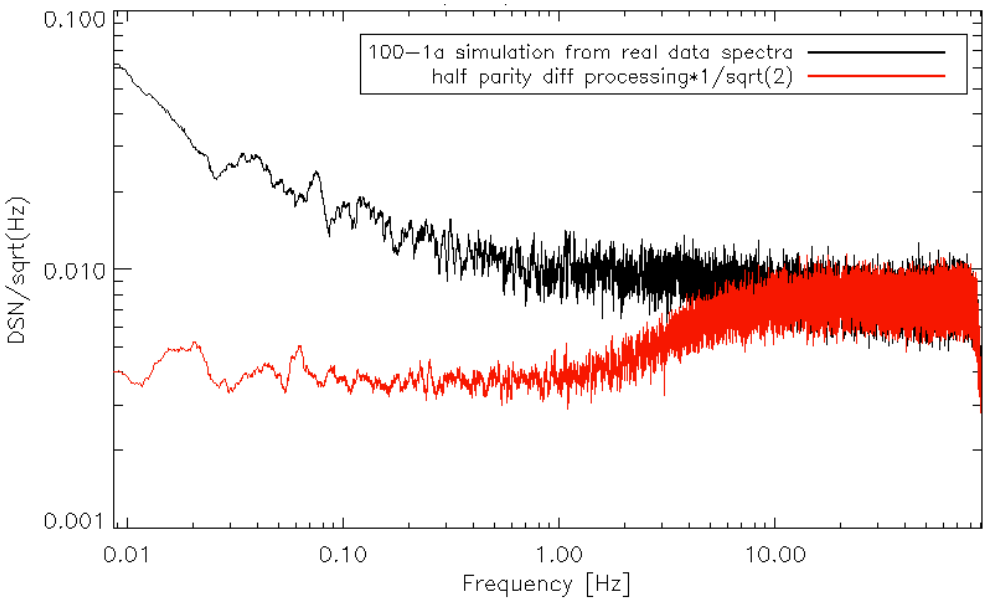


Figure 7.10: Simulation of the HPD noise spectrum at TOI level. The black line is the TOI noise spectrum simulated from the real data spectrum of channel `00_100-1a`. The red line is the spectrum of the HPD processing after application of a factor $1/\sqrt{2}$ to take into account the increase in the noise level due to HPD processing.

To understand the raising feature on the TT spectrum, it has been characterized first in time domain with a simple simulation. Then the flat noise plateau origin could be understood with a modeling of the noise on science data.

[3] In practice leaks were observed on real data at frequencies >= 353 GHz for $\ell > 10^3$

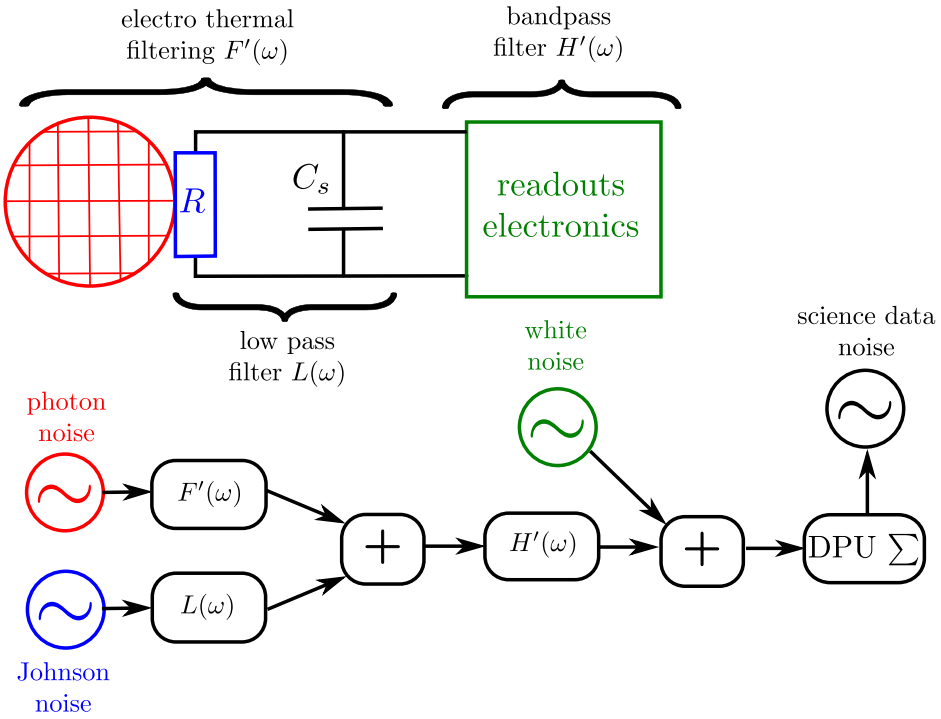


Figure 7.11: Functional diagram for the multi-component noise model of science data. The upper part is a schematic view of the elements relevant to the noise generation and filtering, it shows in red the detector grid receiving the optical power, the bolometer impedance R along with the stray capacitance C_s.

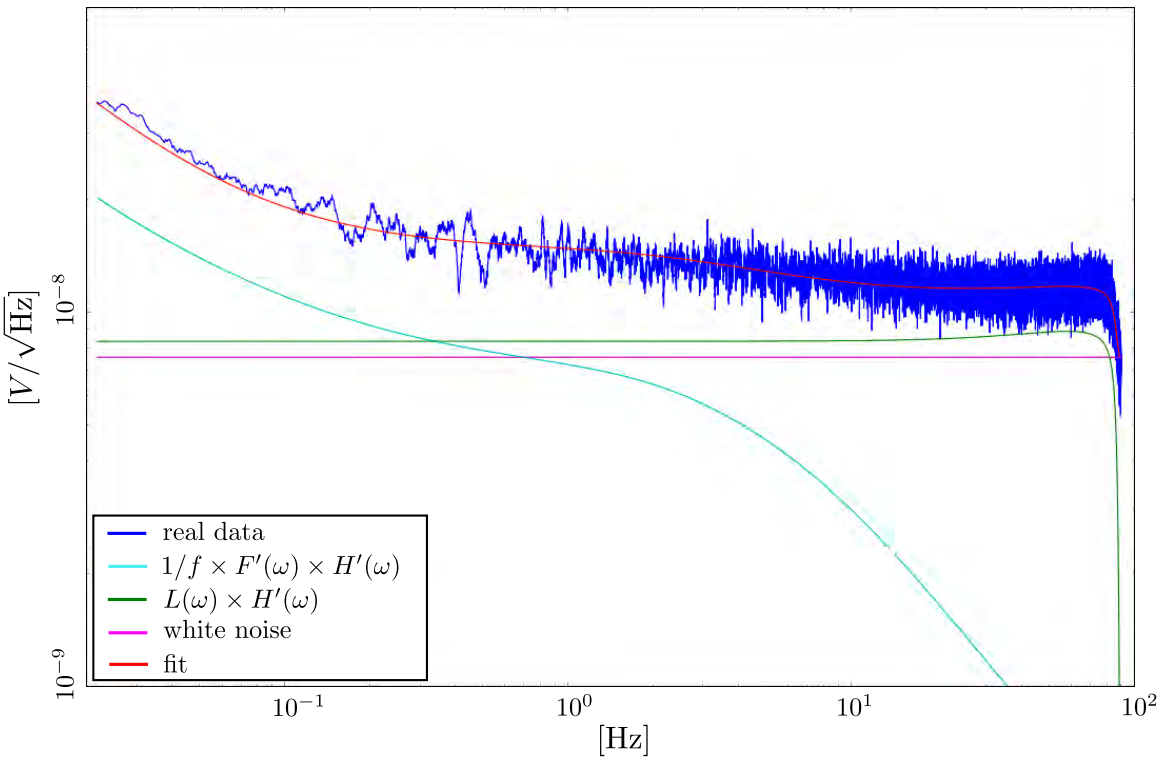


Figure 7.12: Validation of the multi-component noise model on demodulated science data TOI for ring 12 000 of channel `00_100-1a`. The noise components are displayed after scaling with their fitted level. The effect of the summation by the DPU is implicitly included.

Fig. 7.10 shows a noise spectrum (the black line) generated from a realistic noise template in frequency[4]. The HPD processing is performed on the ring of noise generated from the template, yielding a spectrum with a raising feature very similar in shape and location to the one observed on the TT spectrum of real data HPD map. The correspondence between frequency and the multipole ℓ is calculated by applying a factor 60 on the frequency accounting for the approximative 60 s duration of the full spacecraft rotation over the sky.

The flat noise plateau at low frequency could be explained by doing an accurate multi-component modeling of the noise on science data. The noise model used in the end-to-end simulations (see Sec. 5.4 in **?**) was not relevant as it was developed for the science data after the full TOI processing and it is the unprocessed data which is used for the HPD maps. Fig. 7.11 shows the functional diagram for the multi-component noise model on the raw science data. There are three dominant sources of noise which have been identified:

- a $1/f$ component accounting for white photon noise plus thermal variations. This component is filtered by the time transfer function noted $F'(\omega)H(\omega)$, using similar notations as in Sec. 2.1;
- the Johnson noise of the bolometer impedance is filtered by the low pass filter constituted with the stray capacitance C_s and the electronics readout bandpass filter, which are noted $L(\omega)H'(\omega)$;
- a white noise component is present which is not filtered by the electronics readout bandpass filter $H'(\omega)$. This noise component is the one responsible for the low ℓ plateau and its most probable source is the ADC chip.

This model performed very well in reproducing the spectrum of real data noise as shown in Fig. 7.12, with better performances than the one implemented in end-to-end simulations.

Finally the flat noise plateau at low multipole must be understood as the white noise component produced at the end of the electronics readout chain. And the raising feature is the photon noise component filtered by the bolometer and the bandpass filter of the electronics readout.

Comparison with half-ring difference (HRD) map

The noise level on *HFI* maps can be inferred reliably with half-ring maps which are part of the official data release. Therefore TT spectra produced with both methods were compared to assess the level of ADC nonlinearity residuals above the expected noise level on the *Planck* 2015 data release.

Half-ring processing is very similar to the half-parity one, the main difference is that the selection is performed on the first or second half of a ring. Two maps are produced with each selection and their difference cancels the sky signal but also the ADC nonlinearity effect which is the same on the first and second half of the ring. Hence the TT spectra produced with HRD maps contains only noise . Fig. 7.13 shows the comparison between both processing. It can be seen that in the case of the 2015 data release with ADC nonlinearity correction, this systematic effect has becomes significant only for the first multipoles. The deconvolution step is not applied for the HPD processing, but it does not change significantly the spectrum in the multipole range relevant for ADC nonlinearity as shown in the bottom panel of Fig. 7.13.

[4] Noise is extracted from science data by removing the sky signal with the 1D drizzled PBR. The noise template is produced by performing this operation over all the rings of survey 1 and averaging the DFT modulus for each ring

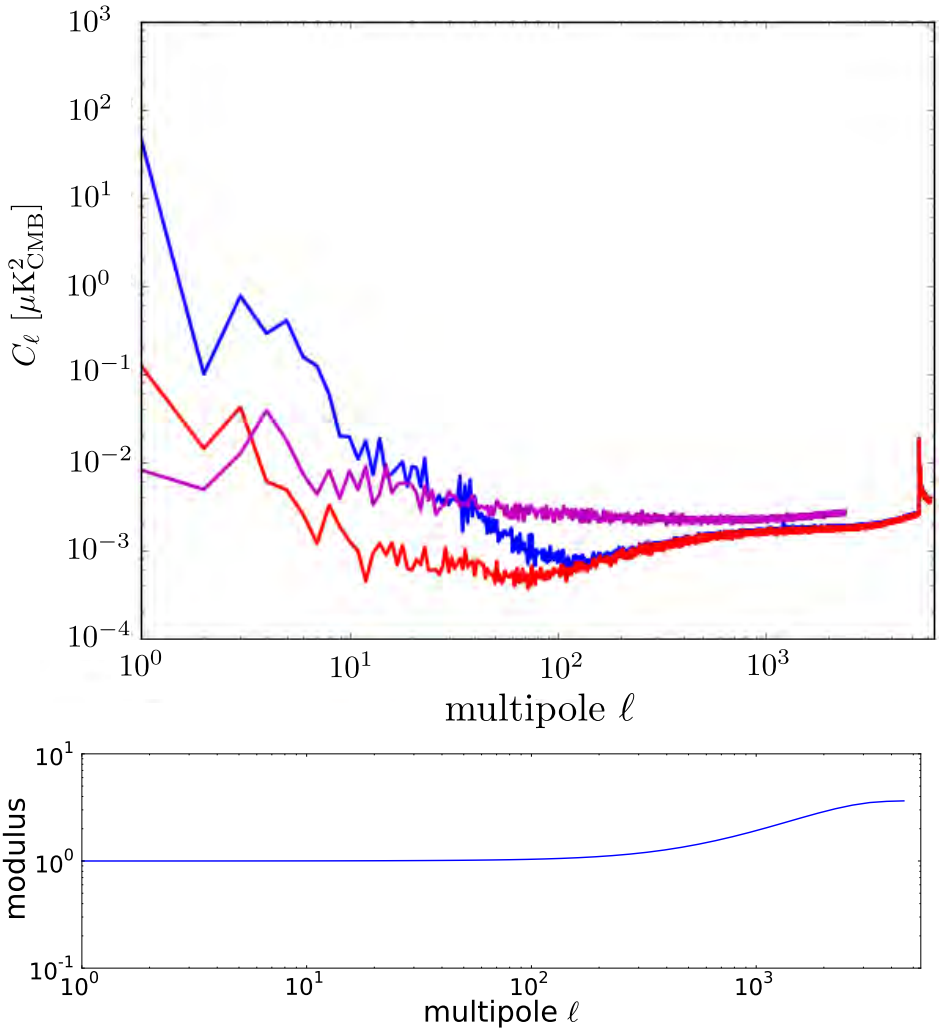


Figure 7.13: *Top*: Comparison of the HPD map TT spectrum with the one of the HRD map on channel `10_143-5`. The blue and red lines are the same spectra as in Fig. 7.8. The purple line is the TT spectrum of the half-ring difference map, which is not affected by the ADC nonlinearity and provides a good ruler to check how much of the ADC nonlinearity residuals are above the expected noise level. *Bottom*: profile in multipole domain of the deconvolution function including the regularization low pass filter which is applied on the science data in the case of the HRD processing, but not in the case of the HPD processing.

7.2.2 Analysis of ADC nonlinearity impact on science

Once the HPD map methodology has been strongly validated, it is used to infer the real impact of the ADC nonlinearity residuals after correction on the scientific analysis. The topics covered below are the photometric calibration, the TT spectrum and the EE and BB spectra in polarization.

Photometric calibration

The photometric calibration of the *Planck* 2015 CMB channels is performed with the orbital dipole (whose amplitude is about 10%) of the solar dipole), appearing at frequency 1/60 Hz in time domain and corresponding approximately to multipole $\ell = 1$ in spherical harmonics domain. Thus an estimate of the photometric calibration error caused by ADC nonlinearity can be obtained by comparing the power of the TT spectrum calculated with HPD maps to the dipole amplitude.

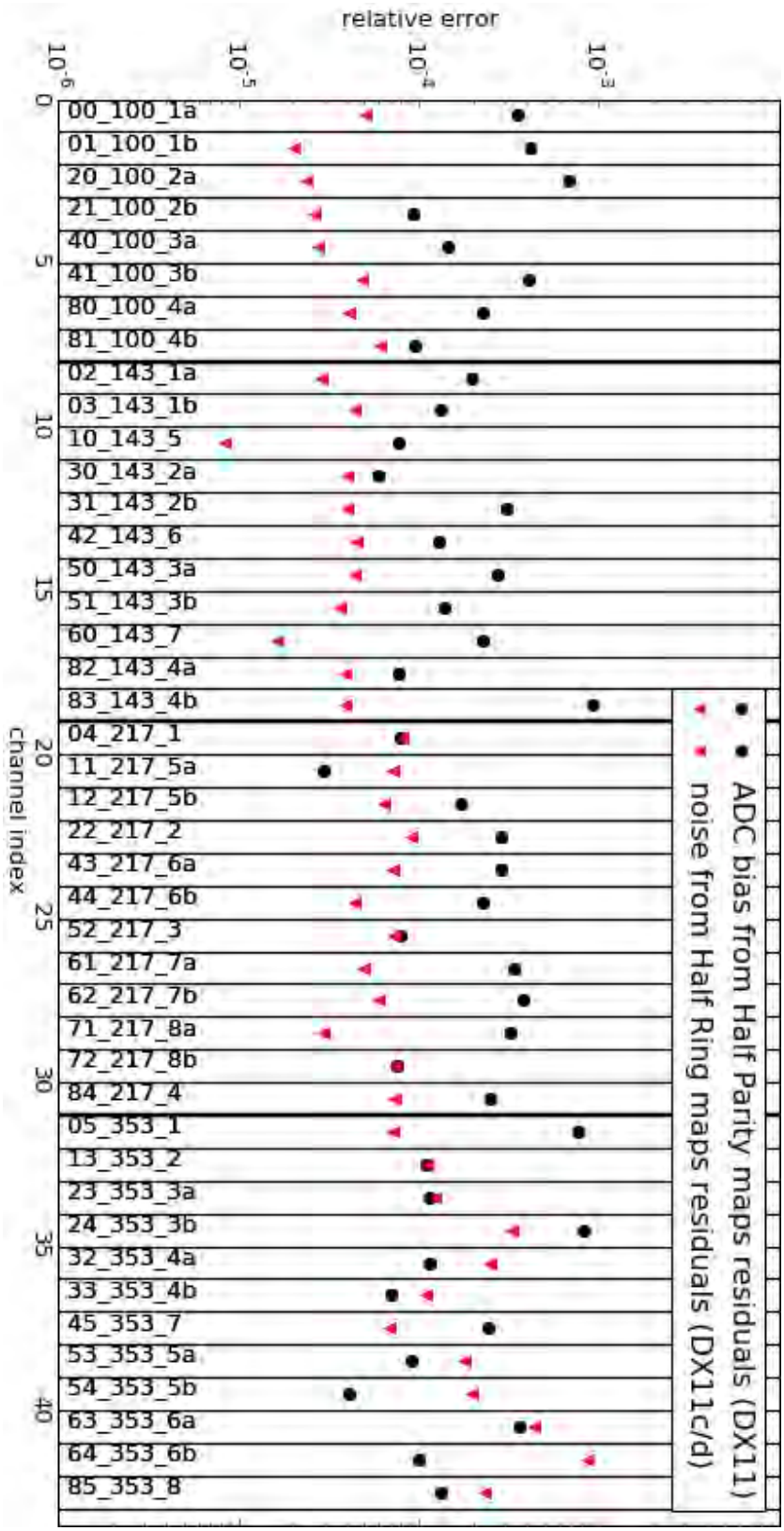


Figure 7.14: ADC nonlinearity error inferred from the *TT* spectrum of HPD maps for all *HFI* channels sorted by frequency after application of the correction used for the *Planck* 2015 data release. The black dot the power at multipole $\ell = 1$ relative to the one of the solar dipole. The red dot is the noise power of HRD maps relative to the one of the solar dipole amplitude, and provides the target for the reduction of the ADC nonlinearity effect.

Fig. 7.14 shows the ratio between the power at multipole $\ell = 1$ of the HPD TT spectrum for all *HFI* channels to the solar dipole amplitude for the ADC corrected science data of the *Planck* 2015 release. The relative error inferred from HPD maps is generally about five times the noise level estimated from the TT spectrum of maps. This level is significant and the following elements must be taken into account:

- the HPD processing underestimates the ADC nonlinearity level by about a factor two;
- the ADC nonlinearity residuals do not average out as the noise when combining data from several channels, as there is some level of correlation of this effect between channels. Hence to reduce this systematic effect the method chosen cannot rely on the average of the signal from several channels.

Impact on the TT spectrum

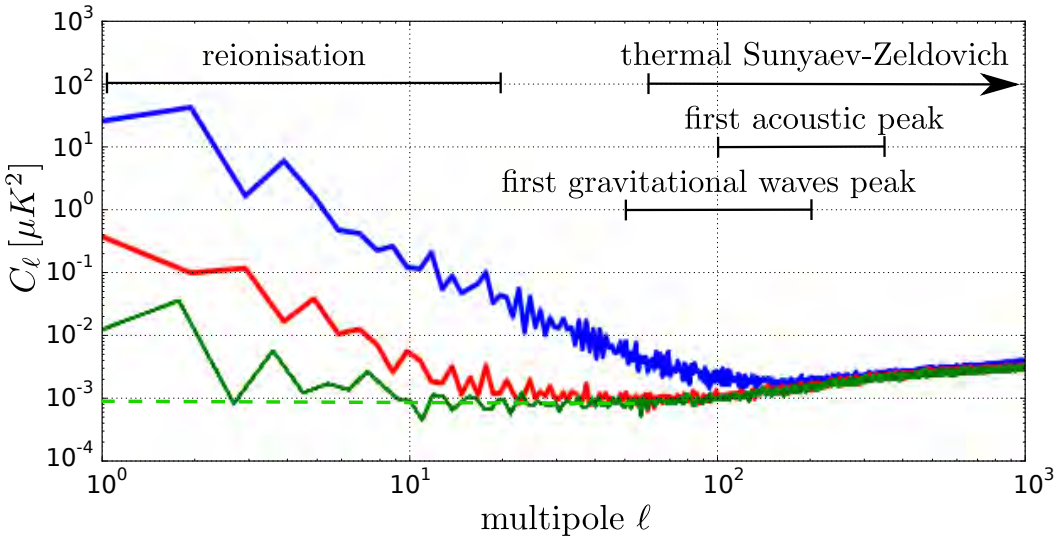


Figure 7.15: Impact of ADC nonlinearity on the scientific analysis of the TT spectrum, the power spectra are calculated from HPD maps of channel `03_143-1b` for different versions of the ADC nonlinearity correction. The blue line is for raw uncorrected data. The red line is for the 2015 data release with ADC nonlinearity correction. The green line is for the 2015 alternative ADC nonlinearity correction detailed in Sec. 6.4. The dashed green line is the expected noise plateau at low multipole for the spectrum of the HPD map.

Fig 7.15 shows an annotated TT spectrum of HPD maps which is representative of the different level of correction for the ADC nonlinearity effect. It can be seen that in the case of the 2015 data release, the multipole range above $\ell = 20$ relevant to the study of the CMB TT spectrum and the tSZ effect is no more affected significantly by this systematic effect. Hence the ADC nonlinearity correction for the *Planck* 2015 data release can be considered relatively "safe" for the scientific analysis of the the angular power spectrum in temperature.

7.2.3 Impact in polarization

Considering the *Planck* 2015 data release, in the multipole range critical for the study of reionisation, the $1/\ell$ residuals of ADC nonlinearity are two orders of magnitude above the expected noise level. Albeit this level is significant and the inter-calibration between channels is also significantly affected, it does not allow to draw definitive conclusions in the case of the different and complex processing of polarization data. Thus, an analysis has been performed from the HPD maps to estimate the ADC nonlinearity impact on the analysis of E and

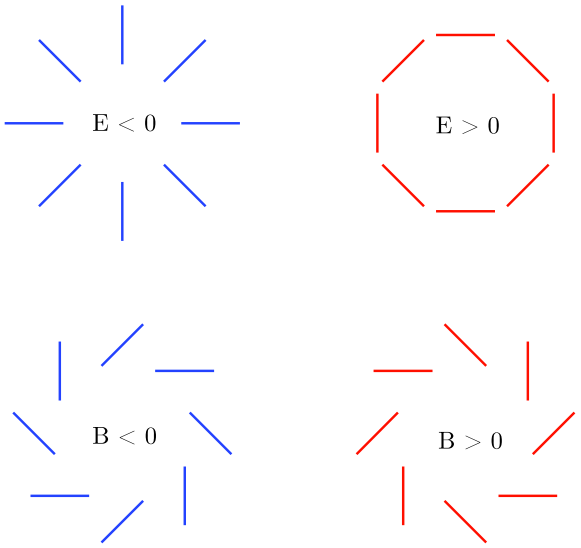


Figure 7.16: Polarization can be decomposed into E and B-modes. The former are radial or tangential with no preferred handedness, similar to an electric field. Like magnetic fields, B-modes do have handedness. Image credit: **?**

B modes in polarization for the *Planck* 2015 data release. The production of the ADC nonlinearity E and B modes residuals is described below.

All channels at 143 GHz have been processed together to produce (I_0, Q_0, U_0) maps for parity zero and (I_1, Q_1, U_1) maps for parity 1, respectively. Specific attention was put on the channel inter-calibration which is mandatory for calculating the Stokes parameter. When this analysis was done, the *Planck* 2015 data release was not yet out and the inter-calibration coefficients used for the *HFI* internal `v61` release candidate were used. The inter-calibration was done after application of the ADC nonlinearity correction used in the *Planck* 2015 data release and is then relevant for this analysis. The stokes coefficients of ADC nonlinearity residuals are calculated with

$$\begin{aligned} \Delta I_{\mathrm{ADC}} &= I_0 - I_1 \\ \Delta Q_{\mathrm{ADC}} &= Q_0 - Q_1 \\ \Delta U_{\mathrm{ADC}} &= U_0 - U_1 \,. \end{aligned} \tag{7.4}$$

Let us note again that the deconvolution cannot be performed when there is a selection by parity, hopefully the filtering introduced by the deconvolution is not significant in the multipole range $\ell \in [0, 20]$ which is critical to this analysis.

Polarization can be decomposed into E and B-modes corresponding to specific orientations of the polarization as shown in Fig. 7.16. This decomposition is performed with spin ± 2 spherical harmonics (**?**) using the `anafast` software, which produces the EE and BB angular power auto spectra. To avoid leaks of features from the bright Galactic plane into the EE and BB spectra, 15% of the sky in the Galactic plane region plus point sources are masked, which results mainly in a drop by 15% of the power measured.

Fig. 7.17 shows the TT, EE and BB spectrum of ADC nonlinearity residuals in polarization. The same $1/\ell$ pattern appears at low multipole on the three spectra. Its relative importance on the TT spectra is not worrying with respect to the noise level estimated with HRD maps, but it is significantly above noise for the EE and BB spectra. For the BB spectrum, the multipoles range that is affected by ADC nonlinearity is below $\ell = 30$, which is probably safe for the analysis of B-modes either from the gravitational waves perturbations or the gravitational lensing. The EE spectrum is more problematic because it

is the $\ell \in [2, 20]$ multipole range which is relevant to the analysis of the reionization and this part of the spectrum is clearly biased by residuals of the ADC nonlinearity effect. This demonstrates clearly that a better handling of the ADC nonlinearity effect is required. The alternative ADC nonlinearity correction studied in 2015 is a good candidate, but it did not reduce as expected the differences between survey acquired one year apart, therefore it has not been used in the 2016 likelihood for constraining reionisation.

7.3 Conclusions

It has been studied how ADC nonlinearity affects quantitatively the Galactic polarization of the thermal dust at 353 GHz. This effect produces a bias at the level of about 10% of the polarization fraction p and about 5° on the polarization angle Ψ. As this is an estimate for the effect before ADC nonlinearity correction, it is expected that the residual bias in the case of the *Planck* 2015 data release is much less important. However it was observed that the ADC nonlinearity effect degenerate with the bandpass mismatch leakage effect, thus a great attention should be taken when performing fit-like procedure of bandpass mismatch.

The impact of ADC nonlinearity on the scientific analysis based on the TT spectrum of the *Planck* 2015 data release was performed with the half-parity maps methodology for which the noise spectrum has been studied with care. It results from this analysis that the ADC nonlinearity leaves residuals with a $1/\ell$ shape in the low multipole $\ell \in [1, 20]$ range which can affect the photometric calibration at a level below 10^{-3} of the solar dipole amplitude for the *Planck* 2015 data release. The same methodology has been applied to the EE and BB spectrum in polarization for the whole set of available 143 GHz channels with similar residuals. While the BB spectrum analysis is not significantly impacted by ADC nonlinearity, this effect is more problematic on the EE spectrum for the study of the epoch of reionization with residuals that are more than one order of magnitude above the noise at the lowest multipoles. While the ADC nonlinearity has been significantly reduced for the *Planck* 2015 data release, it is still expected to affect the determination of the optical depth parameter τ of the reionization for this data release.

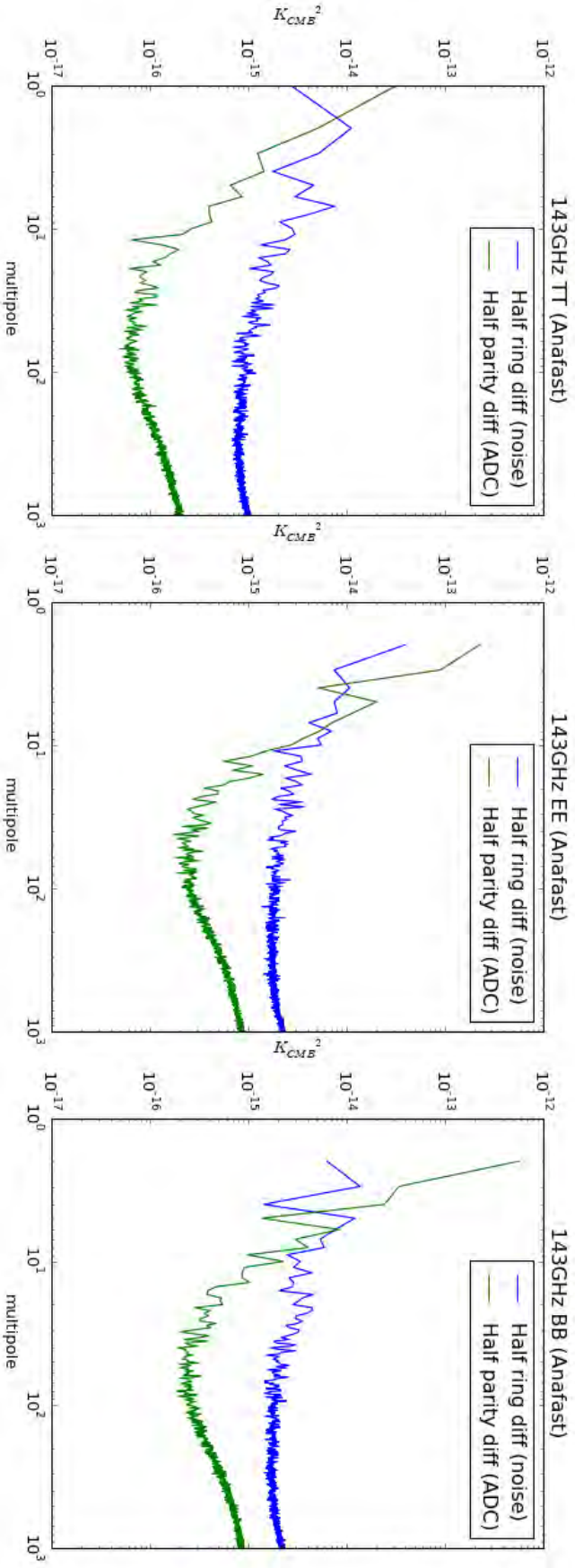


Figure 7.17: TT EE and BB spectra (from left to right) of ADC nonlinearity residuals estimated from HPD maps using all available 143 GHz channels. The blue line is the spectrum of the HRD maps, and the green line is the spectra of the HPD maps, respectively.

Chapter 8

Preliminary study for the detection of primordial matter-antimatter annihilation with Planck data

In the modern view of the Universe, laboratory physics and astrophysical observations indicate that baryons are dominant over anti-baryons. Three main cases can be considered to explain the baryon asymmetry. In the first one, "initial conditions" before the inflation leads to the observed asymmetry. The second one, which is the favored hypothesis today, relies on an asymmetry in the physics of baryons and anti-baryons favoring slightly the former. The last one considers a zero baryon charge at large scales with local fluctuations throughout the Universe, allowing some anti-matter domains to exist. In this chapter, it is the later that is considered, i.e. if some antimatter (AM) domains survived at the epoch of recombination, their annihilation at the interface with matter would have produced a characteristic signature on the *Planck* map of thermal Sunyaev-Zeldovich effect (tSZ-map). Therefore the unprecedented precision of the y parameter at full sky coverage in this map can be used to constrain further the AM domains hypothesis. It will be shown that the *Planck* scanning strategy is a limiting factor to the detection of the annihilation signature, and the impact of the leakage of the scanning strategy in the tSZ-map will be studied.

In Sec. 8.1 is discussed how the *Planck* tSZ-map can be used to perform the detection of the energy released by matter-AM annihilation at the epoch of recombination. Then in Sec. 8.2 the leakage of the *Planck* scanning strategy in the tSZ-map is studied.

This analysis has been used for the quantitative estimation of scanning strategy systematics on the tSZ-map published in Sec. 4.1 of Planck Collaboration XXII (2016).

8.1 Introduction

8.1.1 Context

Gamma-ray observations in the 100 MeV domain (e.g. Nolan et al. 2012; von Ballmoos 2014) have virtually excluded the existence of substantial amounts of close-by baryonic antimatter, constraining its fraction to a few times 10^5 to 10^9 in structures as large as galaxy clusters. As a consequence, a universal matter-antimatter symmetric Universe seems to be incompatible with theory (e.g. Cohen et al. 1998) on the grounds of existing MeV data (Weidenspointner and Varendorff 2001). The observed dominance of matter today is supposed to originate from a Charge Parity (CP) violation during the baryogenesis in the early Universe (e.g. Cohen and Kaplan 1987), leading to a slight excess in favor of baryons. Since this epoch, most of the anti-baryons have annihilated with baryons generating cascades of secondary particles and gamma photons. However in a CP invariant Universe, the baryon charge could have been distributed non-homogeneously throughout the Universe and some local domains of matter and AM could exist. In the later case, if the AM domains (called Anti-Baryonic

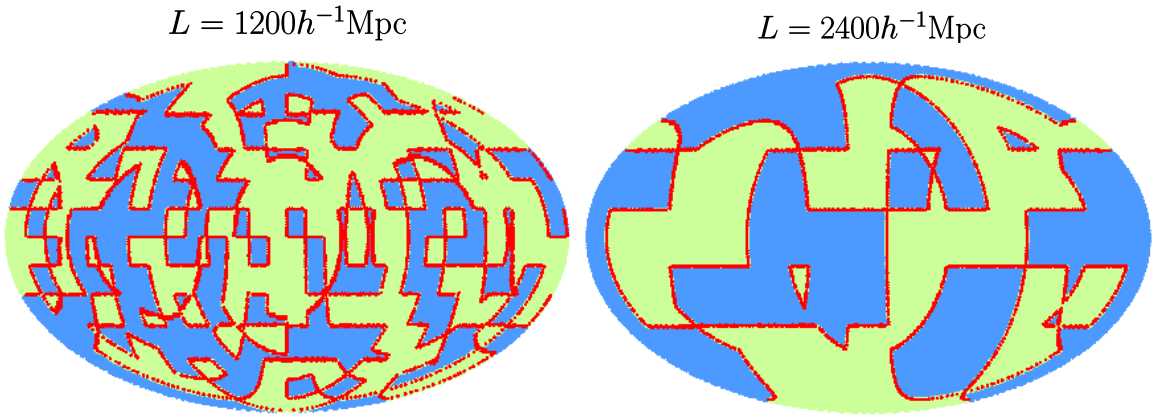


Figure 8.1: Realizations of distortions caused by matter-AM annihilation on the CMB for two cubic domain sizes L. Domains of opposite baryon asymmetry are shown in contrasting shadings, with the interfaces between domains highlighted in red. The spectral distortion of the CMB is confined to these interfaces, which appear as long "ribbons". Image credit: (Kinney et al. 1997)

Clouds (ABC) hereafter), did survive up to the recombination they may have left an imprint on the CMB. Naselsky and Chiang (2004) considers the impact of small ABC annihilation with matter as a delay of the recombination, involving modifications of the angular power spectrum of anisotropies in temperature and polarization. A second approach which is the subject of the systematics analysis detailed in the next sections, consist in searching for local thermal variations caused by annihilation which can be observed on the tSZ-map.

Let us consider the later approach where baryogenesis produces regions of very inhomogeneous baryon charge and ABC have survived up to the epoch of recombination. The inevitable annihilation at the M-AM interface generates at the end of the decay chain neutrinos, about 200 MeV photons, and electrons and positrons. These electrons and positrons are the dominant source of heat transfer to the CMB photons through inverse Compton scattering, yielding distortions of its blackbody spectrum. This scenario has been studied by Kinney et al. (1997) with the conclusions that the spectrum distortion of the CMB at the interface between ABC and matter should produce tSZ effect at level of $y \approx 1 \times 10^{-6}$. The annihilation interface should produce ribbons on the tSZ-map as shown in Fig. 8.1 of width $\theta_W \approx 6'$ and a characteristic separation $\theta_L \approx 1°(L/100h^{-1}\text{Mpc})$ set by the domain size L. A similar approach has been undertaken by Cohen and De Rujula (1997) with slightly different conclusions but a similar tSZ magnitude of $y \approx 1 \times 10^{-6}$. In both cases the observational data from full-sky surveys available at the epoch did not reach the required sensitivity level on y to check this hypothesis.

Observational constraints on the domain size L are also provided by gamma-ray astronomy. Considering the emission model developed by Cohen et al. (1998) (See Sec. 5), the proton-antiproton annihilation at the interface of ABC clouds produces gamma photons primarily from π^0 pions at the end of the decay chain. Today, the resulting redshifted relic spectrum of these photons is expected to peak at about 70 MeV. However the observed gamma spectrum, as shown in Fig. 8.2, is not compatible with small values of $L \leq 1$ Gpc. As a consequence the inhomogeneous baryon charge hypothesis is disfavored by the gamma-ray observations when considering typical ABC sizes much smaller than the observable Universe. Hence, in addition to the gamma spectrum, the new observational data from the recent *Planck* tSZ map is expected to provide improved constraints for the baryogenesis model.

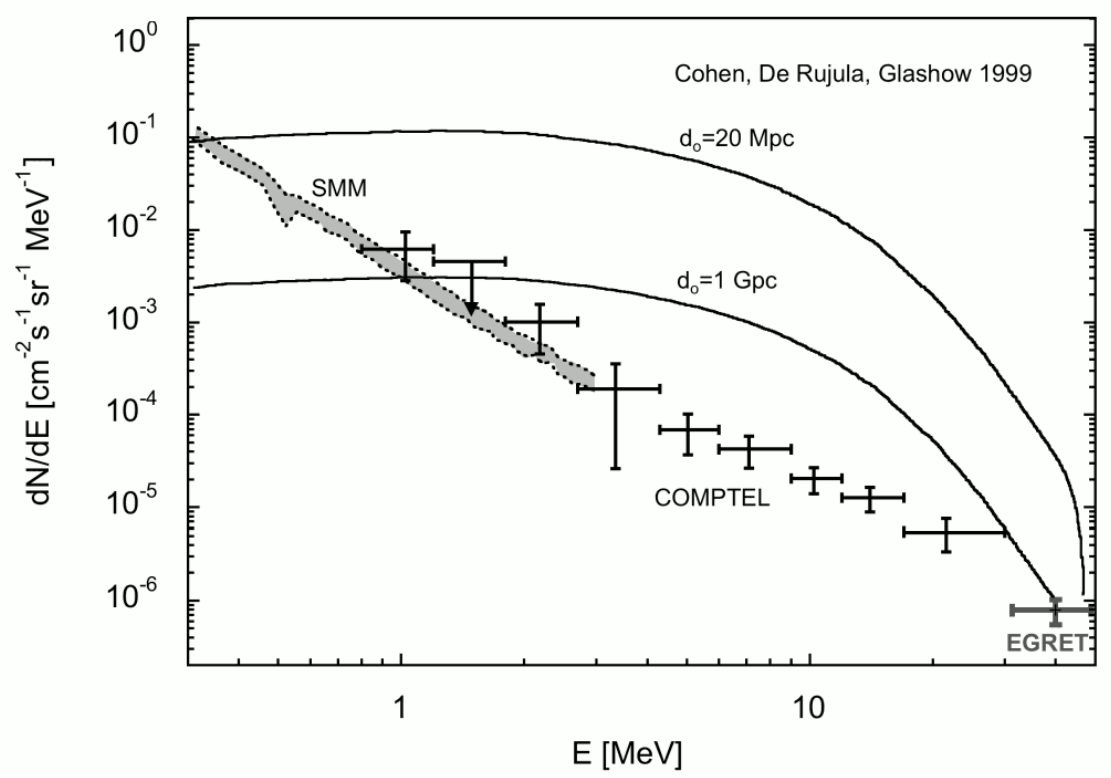


Figure 8.2: Model spectra for the redshifted ABC gamma emission described in Cohen et al. (1998) for two domain size noted d_0 here, They are compared to gamma observations with the CGRO/COMPTEL data from Weidenspointner and Varendorff (2001). Image credit: von Ballmoos (2014)

8.1.2 The *Planck* tSZ-map

The *Planck* tSZ-map is available from the PLA in two versions: `MILCA` (Modified Internal Linear Combination Algorithm, Hurier et al. 2013) and `NILC` (Needlet Independent Linear Combination, Remazeilles et al. 2011) depending on the component separation method used. Both maps are produced with a $10'$ resolution and are very consistent at the exception of a small amount of leakage from the Galactic foreground emission with different patterns. The `MILCA` version of the full sky view tSZ-map is shown in Fig. 8.3.

The noise RMS of the y parameter in both maps is about 1×10^{-6}, which is just at the level required for the detection of the M-AM annihilation interface. The angular resolution is also very close to the theoretical value of $\theta_W \approx 6'$. Therefore the *Planck* tSZ-map is just at the limit of detection required to check the hypothesis of inhomogeneous baryogenesis with surviving ABC at the time of recombination. The expected ribbon-like shape of the annihilation interface should provide sufficient SNR ratio to probe for this kind of structure on the tSZ-map. However there is an issue with the scanning strategy which produces ribbon-like structures in the tSZ-map as shown in Fig. 8.4. These structures degenerate with the expected shape of the annihilation interface. The best suited tool for the detection of annihilation interface is some kind of ribbon-shaped wavelet. Thus the leakage of the scanning strategy would impact significant the power spectrum obtained with a wavelet based method. Hence the leakage of the scanning strategy into the tSZ-map have to be characterized prior to attempt any detection of annihilation interface.

8.2 Characterization of the scanning strategy leakage

To characterize the leakage in the tSZ-map, a custom filtering in the spherical harmonics domain space has been specifically designed. This custom filtering is detailed bellow, then it is applied on the tSZ-map, and finally the observations are validated with simulations.

8.2.1 Angular power spectrum split in the co/cross-scan directions

To estimate the scanning strategy leakage in the co-scan direction, a custom split of the spherical harmonics decomposition is studied. This approach takes the advantage from the fact that the angular power spectrum and the tools used to compute spherical harmonics decomposition are well known, widely spread and quick to run.

The large circles of the *Planck* scanning strategy are roughly intersecting at the ecliptic poles of the sky in the nominal Galactic coordinates. If the map is rotated in ecliptic coordinates then the scan circles intersect approximately the poles which are the axis of symmetry in the referential of the spherical harmonics. This property can be put in relation with the pattern of the spherical harmonics which is aligned in the co-scan direction for $|m| = \ell$ (i.e. follows line of same longitude) and in the cross-scan direction for $m = 0$ (i.e follows line of same latitude). Fig. 8.7 shows the spherical harmonics for $\ell \in [0, 5]$ and positive values of m[1]. Therefore two subsets can be defined:

- the co-scan harmonics for $|m| > (l + 1)/2$ which are dominantly aligned in the scan direction;
- the cross-scan harmonics for $|m| < (l + 1)/2$ which are dominantly orthogonal to the scan direction.

[1] The spherical harmonic harmonic function $Y_{\ell(-m)}$ is identical to $Y_{\ell m}$ after a rotation around the axis of symmetry, thus has the same alignment properties with respect to the scanning strategy in ecliptic coordinates. The harmonics corresponding to negative values of m are visible in Fig. 1.2

These two subsets define the co-scan angular power spectrum $C_\ell^{\parallel}$, and the cross-scan angular power spectrum $C_\ell^{\perp}$. Their expression is derived from Eq.1.3, leading to

$$\begin{aligned} C_\ell^{\parallel} &= \left\langle |a_{\ell m}|^2 \right\rangle \Big|_{m>(\ell+1)/2} \\ C_\ell^{\perp} &= \left\langle |a_{\ell m}|^2 \right\rangle \Big|_{m<(\ell+1)/2} . \end{aligned} \tag{8.1}$$

8.2.2 Application of the co/cross-scan harmonic decomposition to the tSZ-map

Fig. 8.7 show the $C_\ell^{\parallel}$ and $C_\ell^{\perp}$ Power spectra of the tSZ map after application of a mask over 57% of the sky including point sources and the Galactic plane to avoid contamination by residuals of the component separation algorithm. An excess of power is observed in the scan direction at all angular scales with a median value of $\sqrt{C_\ell^{\parallel}/C_\ell^{\perp}}$ being about 14% and 16% for the `NILC` and `MILCA` map. These values are consistent with the RMS based estimates provided in Planck Collaboration XXII (2016). The sky mask has been found to have no impact on the ratio of co-scan power over cross-scan power for multipoles above $\ell = 100$ which are the target of this study.

The harmonic decomposition in two halves has been applied on the tSZ-map by Juan Macías-Pérez, leading to the reconstructed co/cross-scan maps shown in Fig. 8.7. Most of the structure from the scanning strategy is visible in the co-scan map. There is also some ringing visible around the clusters in the reconstructed map which has been found to be due to the co/cross-scan decomposition.

8.2.3 Characterization of the scanning strategy leakage with simulations

To characterize further the co/cross-scan power spectrum results, I repeated the same analysis as in Chap. 7 on noise-only simulations using the same scanning strategy.

The noise-only TOI were generated with `Desire` over the full mission, then destriped and projected to full-sky maps in `HEALPix` format with `polkapix`. Two setup were used for the noise generation:

- white uncorrelated Gaussian noise only;
- white uncorrelated Gaussian noise plus a $1/f$ component with a knee frequency of 0.13 Hz. The setup of channel `00_100-1a` has been used with a NEP of 1.4×10^{-19} W fitted on real data[2].

Fig. 8.8 shows the co/cross-spectra calculated on the simulated maps of noise. In the white noise only case there is an excess of power in the co-scan power spectrum which disappear if the map is normalized with the hit count, i.e. by multiplying the signal by the square root of the number of TOI data samples falling in each pixel. Hence validating that the co/cross-spectra are identical with white Gaussian noise. This also shows that there is a structure in the spectrum of the noise, it has been found to be independent of the destriping step and depends only of the hit statistic in each pixel. The excess of power is significantly higher in the $1/f$ noise case which includes also residual signal from low frequency noise. In the $1/f$ noise case the co/cross-spectra cannot be made more consistent after normalizing with the hit count, which is expected from the random signal of the $1/f$ noise. With `polkapix`, the $1/f$ noise generates inevitably some residual stripes (for an illustration see Fig. 21 in Tristram et al. 2011).

[2] These values are found in the DMC database at the location `IMO:HFI:DET:Phot_Pixel Name="00_100_1a":NoiseAndSyst:FlightNoise`

8.3 Conclusions and perspectives

The scanning strategy leakage in the tSZ-map has been studied in the context of the detection of ribbon-like structures caused by M-AM annihilation at the epoch of recombination. The *Planck* tSZ-map map is theoretically at the limit of the precision requirements, of $y \approx 1 \times 10^{-6}$ and a resolution of about $6'$, for a detection. An excess has been observed of about 15% of power on the angular power spectrums in the co-scan direction. This excess is caused by the scanning strategy which structures the statistics of the noise over the 2D sphere, and by the residuals of the $1/f$ noise.

The scanning strategy leakage on the power spectrum from the white and $1/f$ component can be modeled with Monte-Carlo simulations, thus it is expected their residuals can be constrained efficiently on the angular power spectrum of the tSZ-map. This is valid for the detection of M-AM annihilation at the power spectrum level, however for localized detections with a wavelet-based analysis, the $1/f$ co-scan residuals and the structure of the noise are limiting factors that have to be taken into account. The next step would be to perform a wavelet based analysis using directional and steerable wavelets with ribbon-like shape (e.g. Wiaux et al. 2008; Leistedt et al. 2013).

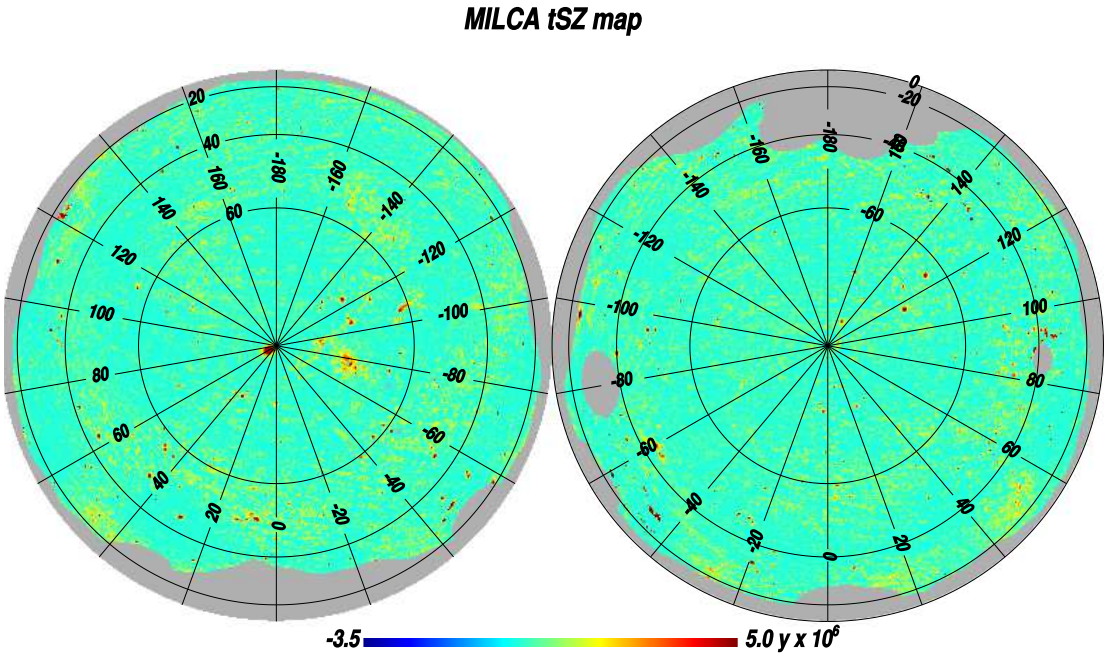


Figure 8.3: *Planck* map of the Compton parameter produced with the `MILCA` (2015) component separation algorithm in orthographic projection. Image credit: Planck Collaboration XXII (2016)

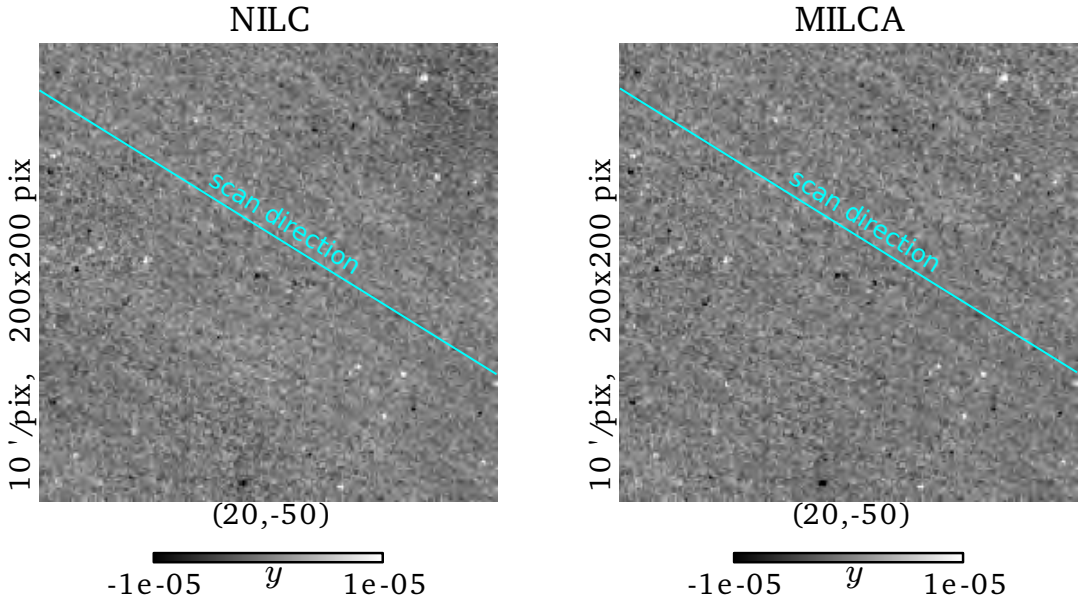


Figure 8.4: Leakage of the scanning strategy in the tSZ-map. The map in `HEALPix` format is projected in gnomonic view at $10'$ resolution and centered at $(-20°, -50°)$ in Galactic coordinates. The greyscale colormap improves the visibility of the scanning strategy leakage. *Left*: map produced with the NILC (PLA 2015) component separation algorithm. *Right*: map produced with the MILCA (PLA 2015) component separation algorithm

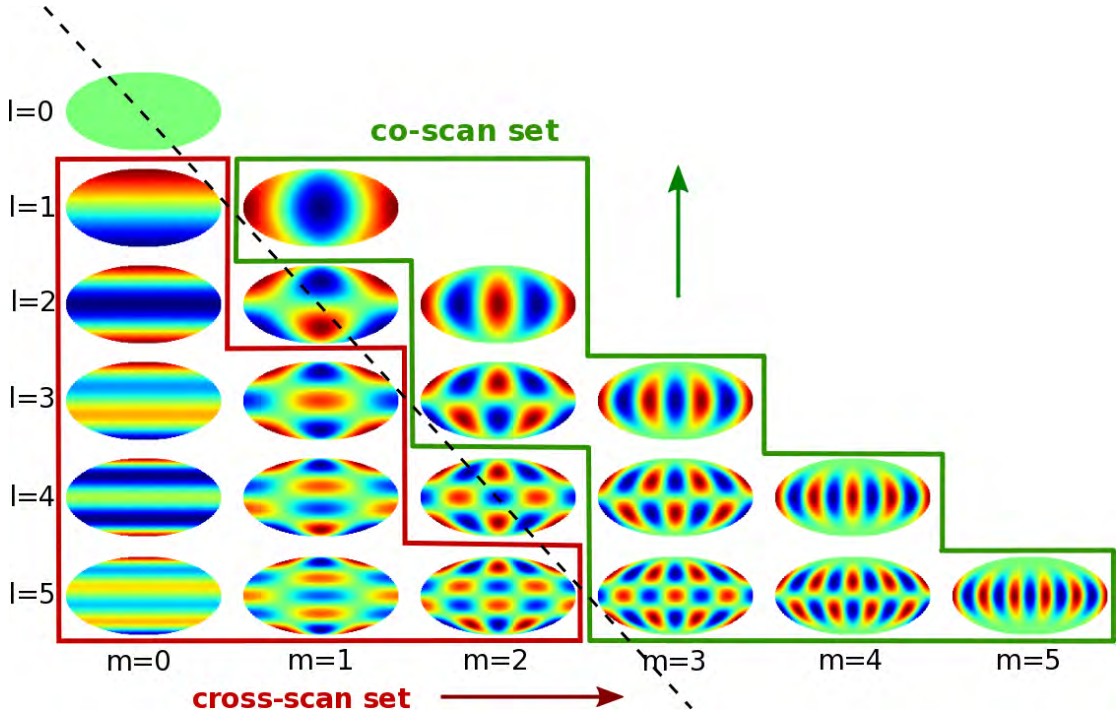


Figure 8.5: Spherical harmonics selection for co-scan / cross-scan directions

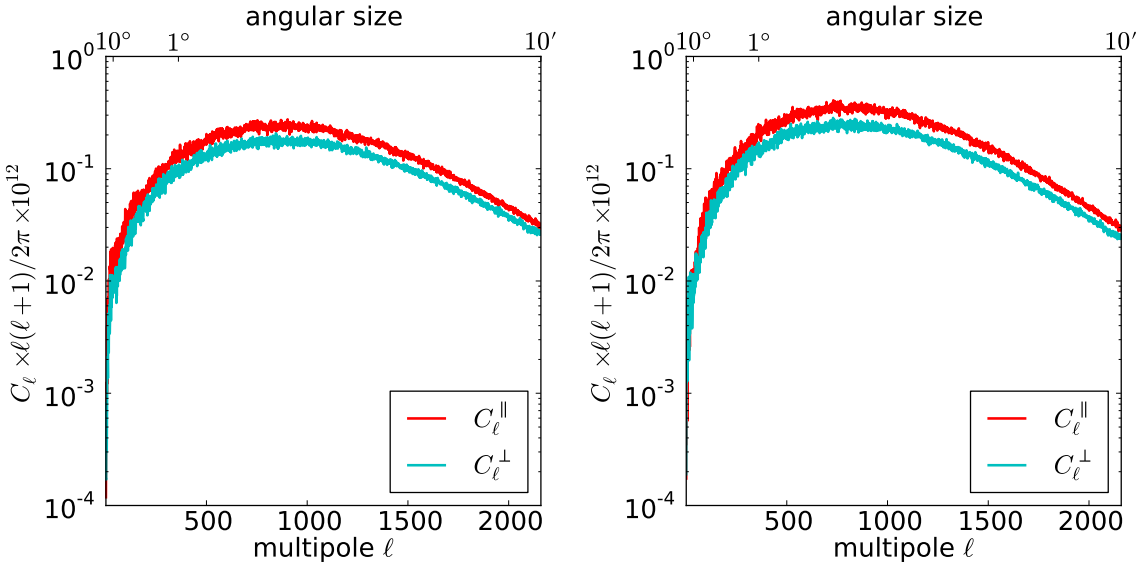


Figure 8.6: Co/cross-scan angular power spectra of the tSZ-map. The Galactic plane and point sources are masked, excluding 57% of the sky (`COM_Mask_Compton-SZMap_2048_R2.00.fits`, PLA 2015) prior to the calculation of the angular power spectra. *Left*: `NILC`. *Right*: `MILCA`

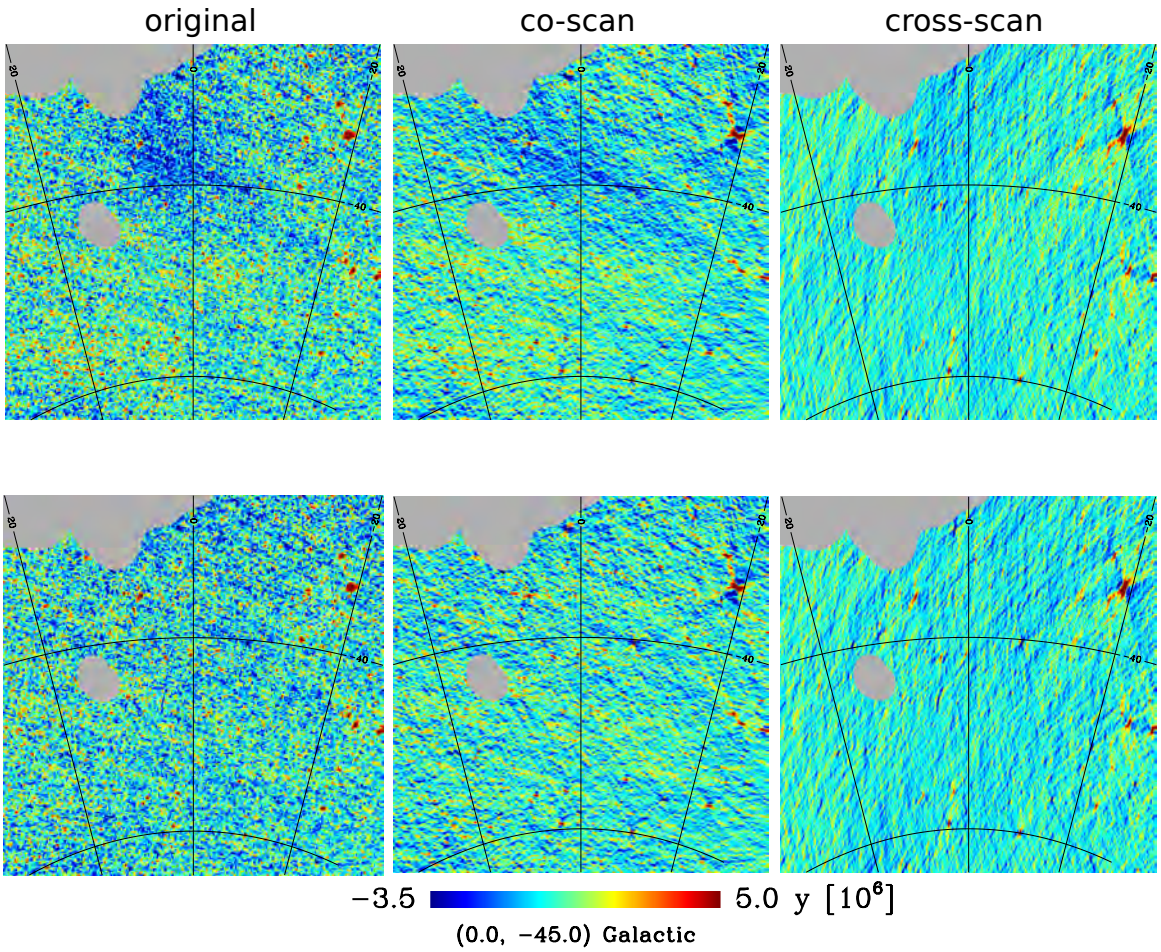


Figure 8.7: Co/cross scan harmonic decomposition in a small region of the tSZ-map centered at $(0°, -45°)$ in Galactic coordinates. *From left to right*: original map, co-scan only harmonics, cross-scan only harmonics. *Top*: `NILC` map. *Bottom*: `MILCA` map. Image credit: Planck Collaboration XXII (2016)

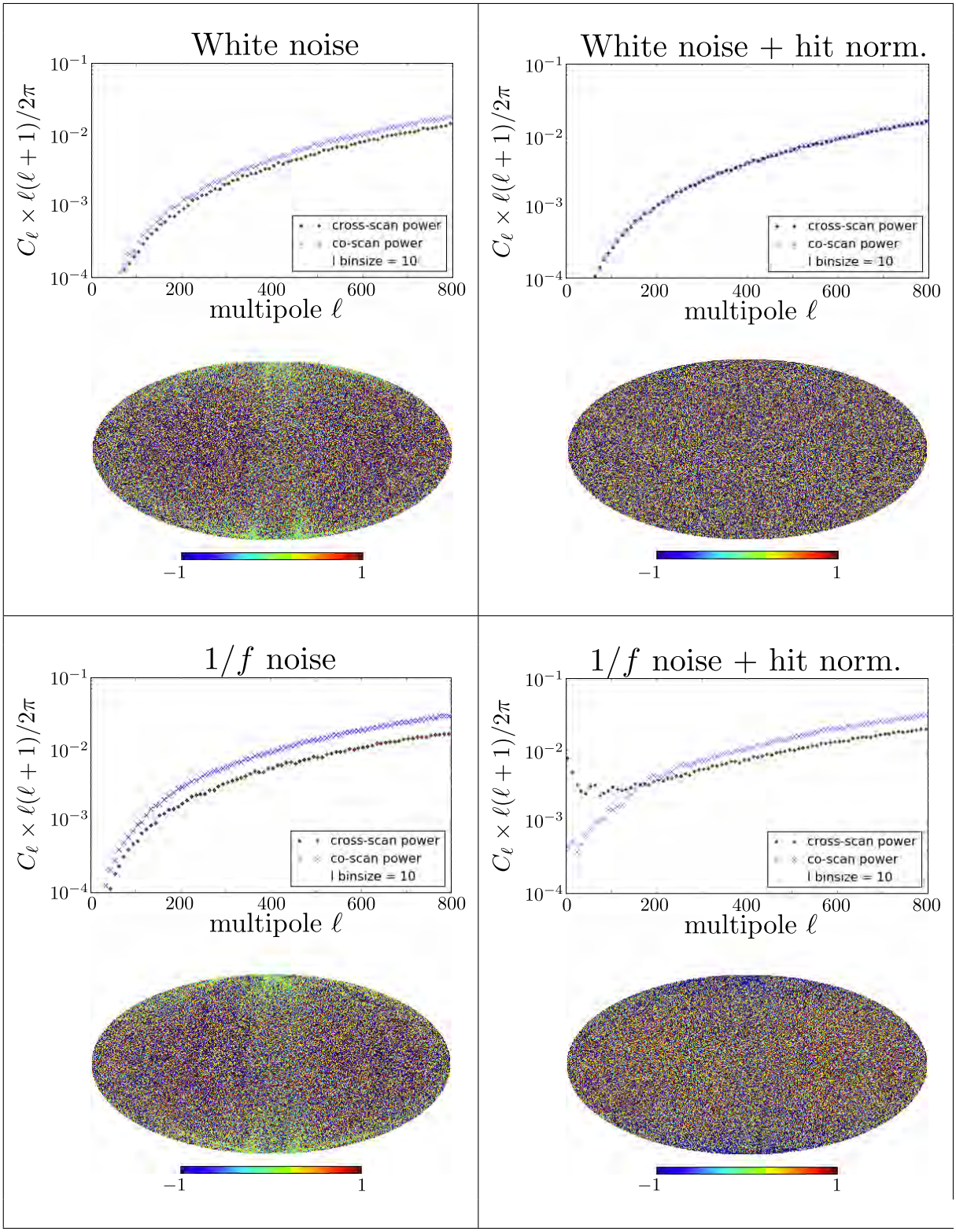


Figure 8.8: Co/cross-scan power spectra of noise only simulations over the full mission. Each panel shows the power spectra above and the Mollweide projection of the destriped map below. *Top*: white noise case. *Bottom*: $1/f$ noise case. *Left*: Nominal simulation. *Right*: The sky map of noise is normalized with the hit count in each pixel.

Conclusions et perspectives

Ce travail de thèse a commencé dans le contexte de la fin de mission du satellite *Planck*, au moment où l'implication de l'ADC dans les variations de gains observées venait juste d'être découverte. Il restait moins d'un an avant la décommission de *Planck* et la perte définitive des possibilités de communication avec la chaîne de lecture de l'électronique de bord qui était nécessaire pour collecter les données vitales pour la correction des effets de non-linéarité de l'ADC. Cela a été le commencement d'un long parcours, où était impliqué le travail avec une large collaboration constituée de personnes de nombreuses origines. Ce qui a mené non seulement à l'acquisition de nombreuses connaissances dans les domaines scientifiques et techniques nécessaires pour la compréhension d'un télescope spatial très complexe, mais aussi à de très riches interactions humaines et une aventure personnelle unique. Passons maintenant aux conclusions de ce travail.

La première partie est dédiée à la caractérisation et à la modélisation de la fonction de transfert temporelle des détecteur bolométriques d'*HFI*. Avec comme motivation, à l'origine, l'excès observé dans la réponse temporelle à basse fréquence des détecteurs. Cet excès, observé durant les campagnes d'étalonnage au sol, a été particulièrement compliqué à étalonner précisément une fois en vol. Ceci parce que les constantes de temps impliquées sont particulièrement longues, de l'ordre de 50 ms à une minute. Au sein de la *Tau Tiger Team* l'approche empirique qui a été adoptée pour construire le modèle labellisé LFER de la fonction de transfert temporelle pouvait reproduire la réponse temporelle des détecteurs mais ne fournissait pas de contraintes à basse fréquence. La problématique principale était que la réponse temporelle des bolomètres alimentés en courant modulé était approximée avec la théorie développée pour l'alimentation avec un courant continu, celle-ci ne décrivait pas assez précisément la réponse observée pour pouvoir atteindre le niveau de précision de l'ordre de 10^{-3} requis pour *HFI*. Et donc, pour fournir des contraintes sur la fonction de transfert temporelle, un modèle analytique de la réponse temporelle des bolomètres polarisés avec un courant carré a été développé à partir du modèle simple de bolomètre. Les résultats ont été publiés dans Sauvé and Montier (2016). Il est apparu que l'approche classique par passage dans le domaine de Fourier n'était pas utilisable dans le cas de la version linéarisée des équations différentielles qui décrivent l'équilibre électro-thermal des bolomètres. Il a donc fallu développer une méthode basée sur une décomposition spécifique des signaux en fréquence pour pouvoir calculer une solution matricielle avec l'algèbre linéaire. Le modèle analytique obtenu a été validé en le comparant à une simulation réalisée avec le logiciel `SEB` qui réalise l'intégration des équation différentielle décrivant le système avec la méthode de Runge-Kutta à l'ordre 4. L'accord obtenu entre le modèle analytique et la simulation était de l'ordre de 10^{-3} pour un détecteur à 100 GHz. Il a été montré cependant que la version de base du modèle analytique avec un seul composant thermique n'était pas suffisante pour reproduire correctement l'estimation de la fonction de transfert temporelle faite avec le modèle LFER. Un second composant thermique a été ajouté au modèle analytique. Avec seulement 5 paramètres thermiques libres, il a pu reproduire la fonction de transfert temporelle utilisée pour la déconvolution des données du catalogue *Planck* 2015, avec un accord à 2.7σ de l'estimation sur son incertitude. Cette nouvelle version était très proche de l'objectif et décrivait correctement la constante de temps principale du détecteur autour de ≈ 10 ms, avec une seconde constante de temps autour de ≈ 50 s. La réponse temporelle spécifique aux bolomètres en régime alternatif imposait au modèle LFER l'utilisation de plus de cinq composant thermiques,

alors que dans le cas du modèle analytique, deux composants arrivent à un résultat qui est pratiquement équivalent. Ces très bon résultats fournissent plusieurs pistes d'applications directes à ce modèle, en commençant par de nouvelles contraintes sur la reproduction des composantes basse fréquence de la réponse temporelle des détecteurs de l'instrument *HFI*. Ce qui nécessite cependant un travail complémentaire pour adapter la procédure d'estimation de la fonction de réponse angulaire de l'optique. Le modèle analytique peut également reproduire le *raw gain* qui est un nouvel observable inexploité de la fonction de transfert temporelle. Cette propriété est particulièrement intéressante pour améliorer la correction de la non-linéarité des ADC, car cela fourni un modèle du *raw gain* qui est un paramètre critique de cette correction.

Une approche complémentaire à la modélisation analytique a été l'utilisation des *glitch* qui sont la réponse des détecteurs liée au dépôt d'énergie d'une particule qui les traverse. Habituellement les *glitch* sont une nuisance qui cause des pertes de données, mais ils peuvent également être utilisés pour contribuer à la validation d'une fonction de transfert temporelle candidate. Dans le cas présenté ici, ce raisonnement a été poussé plus loin. Une méthode, basée sur la géométrie Euclidienne, a été développée pour reconstruire directement la réponse impulsionnelle des détecteurs avec les *glitch*. La problématique principale était la présence de plusieurs *familles* de *glitch* correspondant aux différentes parties du bolomètre où se faisais le dépôt d'énergie. La méthode a été adaptée pour séparer efficacement les familles dominantes. Il a été observé que la la réponse impulsionnelle de la famille de *glitch* labellisée *short* était celle qui reproduisait le mieux la fonction de transfert temporelle estimée avec le modèle LFER. A ce moment le modèle LFER était en compétition avec un autre modèle nommé JH10. Ce résultat a donc été un argument discriminant en faveur du modèle LFER qui a été utilisé depuis pour la déconvolution temporelle des données scientifiques ayant servi à produire les catalogues *Planck* 2013 et 2015. Plus tard la méthode de reconstruction dans sa version définitive a permit de reproduire la fonction de transfert temporelle estimée avec le modèle LFER avec une précision de 1%. Et donc, bien que la réponse impulsionnelle des *glitch* soit un outil très puissant pour l'analyse de la réponse thermique des bolomètres, elle est limitée par des biais comme la proximité de plusieurs familles de *glitch short* et par la difficulté de normalisation par l'énergie des particules. Cependant il a été démontré que le paramètre $\tau_{\mathrm{H}p}$ correspondant au filtrage par l'électronique de lecture des détecteurs était bien contraint par la réponse impulsionnelle des *glitch*, mais cela dépend du modèle qui est utilisé pour reproduire la réponse thermique des bolomètres. Dans ce cas la présence de plusieurs familles de *glitch* correspondant à des réponses thermiques différentes aidait à améliorer cette contrainte. Une extension possible de ce travail consisterait à utiliser le modèle analytique présenté ci-dessus pour reproduire la réponse thermique, ce qui fournirait de bien meilleures contraintes sur le filtrage effectué par l'électronique de lecture des détecteurs. De plus le bruit non-gaussien généré par les queues des familles de *glitch* avec une relaxation très lente (labellisés *glitch long*) pourrait être enlevé plus efficacement avec une combinaison du modèle analytique utilisé pour reproduire ces glitches.

Dans la seconde partie de ce travail, les effets de la non linéarité introduite par l'ADC sur les données scientifiques sont caractérisés et corrigés. Cet effet systématique s'est d'abord manifesté sous forme de variations apparentes du gain au cours du temps de l'ordre de 1.5%, et s'est révélé particulièrement difficile à gérer. La première urgence était de faire une caractérisation précise des composant pour les 52 canaux bolométriques de *HFI* avant le décommissionnemment du satellite. Deux campagnes d'acquisition ont été menées au sol avec des composants de rechange, ceci en vue de collecter les informations nécessaires à la constitution d'un modèle paramétrique des défauts de l'ADC utilisé dans *HFI*. En parallèle, Maxwell Technologies a fourni le modèle `SPICE` du composant. L'absence d'un fichier a empêché l'utilisation du modèle pour lancer des simulations afin de reproduire les défauts des composants à bord du satellite, mais son étude a fourni de précieuses informations sur

l'architecture du composant. Les données fournies par les deux campagnes d'acquisition au sol ont montré la présence d'un motif pseudo-périodique sur 64 codes consécutifs de l'ADC qui a été corroboré par l'étude du modèle `SPICE`. Le modèle paramétrique a donc été abandonné au profit d'une approche empirique basée sur ce qui a été nommé le *motif à 64 codes* avec en supplément une estimations des sauts qui apparaissent entre deux motifs. Il faut noter cependant que les 64 codes juste après la mi-échelle de l'ADC n'obéissent pas au motif et leurs défauts doivent être estimés séparément. Toutes ces informations ont été prises en compte pour configurer ce qui a été appelé *la séquence EOL* et qui a consisté à étalonner les défauts des composants du satellite avec le bruit Gaussien des *fast samples*. Puis l'estimation de ces défauts a été menée à bien avec une méthode de maximisation de vraisemblance. Finalement la précision obtenue a été d'environ $\approx 3 \times 10^{-3}$ sur la taille relative des pas de quantification des ADC obéissant au motif à 64 codes, très près de l'objectif qui était de $\approx 1 \times 10^{-3}$. Cependant les 64 codes juste après la mi-échelle ne bénéficient pas de la statistique utilisable avec le motif et sont un facteur limitant pour la correction de l'effet de non-linéarité de l'ADC. Ceci pourrait être facile à améliorer, en l'absence d'autres sources de biais comme les lignes 4 K, car on pourrait mieux contraindre ces 64 codes en se servant des résidus de l'effet de non-linéarité de l'ADC sur les données scientifiques. Une autre possibilité pour mieux contraindre ces 64 codes serait d'utiliser les données scientifiques acquises pendant les 10 mois de la séquence EOL, durant lesquels *HFI* n'étant plus à 0.1 K, à mesuré uniquement du bruit qui a été insuffisamment exploité.

L'estimation précise des défauts des ADC en vol était un vrai défi, mais cette information seule ne permettait pas d'effectuer la correction de l'effet de non-linéarité sur les données, à cause de la sommation des *fast samples* effectuée dans le DPU pour produire les données scientifiques. Cette étape détruit l'information critique sur la valeur exacte des codes de l'ADC qui est la seule utilisable avec les données de caractérisation du composant. En conséquence la valeur de ces codes devait être estimée en devinant quel était le signal analogique qui avait servi à les produire. Ce qui nécessitait une modélisation très précise de la réponse de la chaîne de l'électronique de lecture à la puissance optique reçue par le bolomètre. Ce modèle combiné avec l'estimation précise des défauts de l'ADC a permit de produire la fonction de correction pour les données scientifiques. La validation de cette correction nécessitait une estimation du niveau des résidus de l'effet de non-linéarité de l'ADC sur les données scientifiques. Des observables spécifiques ont été développés en comparant les échantillons des données scientifiques acquis avec la première demi-période de la modulation avec ceux de la seconde demi-période, respectivement. Ceux-ci sont qualifié d'observables par *demi-parité* et sont sensibles uniquement à l'effet de non-linéarité de l'ADC car c'est la seule source connue de différence, hormis le bruit, entre les ciels scannés par chacune des deux demi-périodes de la modulation. En les utilisant il a été observé que la première version de la correction réduisait de façon significative les variations de gain dues à l'ADC qui pouvaient descendre à 0.1% pour certains canaux. En complément, des simulations de l'effet de non-linéarité de l'ADC étaient en très bon accord avec les observations sur les données scientifiques, ce qui validait le modèle de correction et la méthodologie basée sur les observables par demi-parité. Cependant il a été trouvé que cette correction était limitée par une source de signal parasite dont l'origine était le compresseur mécanique de l'étage à 4 K, signal qui a été surnommé lignes 4 K. Ce signal nécessitait une caractérisation spécifique afin de l'inclure dans le modèle de correction.

Les lignes 4 K apparaissent sur les données scientifiques sous la forme de 4 lignes très fines dans le domaine des fréquences, mais en réalité il y a une "forêt" de lignes sur le signal des *fast samples* avant la sommation par le DPU qui produit les échantillons des données scientifiques. Une fois de plus l'étape de la sommation détruit des informations vitales en repliant les harmoniques au dessus de la fréquence de modulation. L'inclusion des lignes 4 K dans le modèle de signal analogique nécessite la connaissance de sa forme dans le domaine temporel

dont la périodicité s'étend sur neuf demi périodes de la modulation. Mais ces données étaient inaccessibles pendant la mission. Ce qui a eu pour conséquence que le signal périodique des lignes 4 K pouvait seulement être estimé à partir des observables qui sont les quatre versions repliées des lignes 4 K visibles sur les données scientifiques et la période des *fast samples* capturée toutes les 101 secondes. L'analyse des séquences continues de *fast samples* capturées pendant la dernière phase de commissionnage (CPV) du satellite n'a pas montré de corrélation significative entre les canaux qui pouvait être utilisée pour modéliser l'évolution des lignes 4 K durant la mission. Pour cette raison, la solution de repli a été l'estimation des trois harmoniques dominantes sur les *fast samples* en utilisant les observables disponibles. Cette solution a nécessité le calcul pour chaque période où le pointage est stable de 18 fonctions de corrections, une pour chaque configuration des échantillons de données scientifiques relative aux lignes 4 K et à la modulation. Ce qui a amélioré de façon significative la correction des données scientifiques en abaissant le niveau des variations de gain autour de 10^{-3} pour la majorité des canaux. Mais ce n'était pas encore suffisant pour atteindre l'objectif de 10^{-4}. Une nouvelle solution a été développée à partir d'une analyse MCMC de l'espace des solutions pour les lignes 4 K en incluant les observables par demi-parité dans la fonction de vraisemblance. Ce qui a montré que le modèle de signal analogique pouvait être simplifié en utilisant un signal de modulation (i.e la *raw constant*) identique sur toute la mission. En partant de cette constatation il était possible d'estimer une seule *raw constant* en minimisant les observables par demi-parité sur toute la mission. Ce qui, combiné avec une estimation des lignes 4 K sur les *fast samples* fournie par Guillaume Patanchon, a permit de réduire les variations de gain à un niveau proche de 10^{-4}. Cependant les différences *annuelles* entre les cartes du ciel prises à un an d'intervalle n'ont pas montré d'amélioration significative en comparaison de la solution précédente. Pour cette raison une approche empirique (`SRoll`) est utilisée depuis 2016 pour réduire les variations observées entre les cartes du ciel vues par des canaux différents. Une des raisons possibles pour l'absence d'amélioration sur les différences annuelles est une variation réelle du gain de l'électronique de lecture des détecteurs. Il est également envisagé que la procédure d'estimation des défauts de l'ADC capture une fraction de la distribution du bruit qui a servi à les estimer, et une approche différente pour cette estimation a été proposée. Il reste également à caractériser l'impact de la modulation significative des lignes 4 K induite par les opérations de l'étage à 20 K qui ont une périodicité typique de 15 min.

Dans la troisième partie de ce travail, l'impact des effets systématique sur l'analyse scientifique des cartes *HFI* est étudié. Dans le cadre de la préparation du papier (Planck Collaboration Int. XIX 2015), la propagation de l'effet de non linéarité de l'ADC sur l'emission polarisée de la poussière Galactique a été étudié avec une simulation. Le modèle de signal analogique et la caractérisation de l'ADC décris précédemment dans la seconde partie de ce travail ont été utilisés pour produire la simulation. Il a été déterminé que le biais introduit sur la fraction de lumière polarisée est de l'ordre de 10% et que celui sur l'angle de polarisation est de l'ordre de 5° dans le plan Galactique, de plus les résidus de l'effet de non linéarité de l'ADC sur les cartes de paramètres de Stokes dégénèrent fortement avec l'incertitude sur la correction de couleur (bandpass mismatch). Une étude similaire nécessite d'être refaite sur les cartes du catalogue *Planck* 2015, et une attention particulière doit être portée dans le cas de l'incertitude sur la correction de couleur à cause de la dégénérescence avec l'effet de l'ADC. Ensuite, dans le cadre de l'analyse du spectre de puissance angulaire du CMB, les résidus causés par la non linéarité de l'ADC été inférés à partir des données scientifiques de *HFI*. Pour cela une méthodologie basée sur des cartes par demi-parité a été développée et validée avec des simulations. La somme des deux cartes de demi-parité contenant principalement du bruit et ces résidus, elle permet l'analyse des résidus de l'ADC sur le spectre de puissance angulaire. Il a été ensuite observé que ces résidus se manifestent par un excès de puissance en $1/\ell$ à bas multipôles pour $\ell \in [1, 20]$, et que la correction utilisée pour le

catalogue de données *Planck* 2015 est suffisante pour ne pas biaiser l'analyse du spectre de puissance angulaire en température du CMB pour $\ell > 20$. Cependant le biais résiduel sur le multipôle $\ell = 1$ est de l'ordre de cinq fois le niveau de bruit attendu, et il est donc significatif pour l'étalonnage photométrique des détecteurs. Dans le cas de la polarisation, l'analyse des cartes par demi-parité à 143 GHz a montré que le résidus en $1/\ell$ est problématique pour l'analyse du spectre *EE* sur lequel il apparait un ordre de magnitude au dessus du bruit. Il apparait donc qu'il est nécessaire d'améliorer la correction de l'effet de non-linéarité de l'ADC pour l'étude de l'époque de la réionisation.

Finalement une étude préliminaire a été menée dans le cadre de la détection de l'annihilation de domaines résiduels d'antimatière à l'époque de la recombinaison. En théorie, l'interface d'annihilation entre les domaines de matière et d'antimatière laisse une signature sur la carte d'effet tSZ de *Planck* à un niveau d'environ $y \approx 1 \times 10^{-6}$ sous forme de rubans ayant une largeur angulaire d'environ $6'$. Cet effet est à la limite de détection sur la carte d'effet tSZ dont le RMS du bruit est d'environ 1×10^{-6} et la résolution de $10'$. Il se trouve que la stratégie de balayage de *Planck* créé des fuites sur cette carte qui dégénèrent avec la signature recherchée. L'étude de ces fuites a été réalisée avec une partition choisie de la base des harmoniques sphériques qui sélectionne les harmoniques sphériques orientés majoritairement dans la direction du balayage et dans celle qui est orthogonale. Il a été ensuite mesuré un excès médian de signal de 14% (16%) via le spectre de puissance angulaire dans la direction du balayage pour la carte produite avec l'algorithme `NILC` (`MILCA`) à toutes les échelles angulaires. La stratégie de balayage donne une structure en 2 dimensions au bruit sur la sphère dont l'impact sur le spectre de puissance angulaire a été caractérisé avec des simulations. Cette structure dois donc être prise en compte lors d'une recherche de signature d'annihilation à la limite du seuil de détection. Cette recherche peut être faite avec des ondelettes par exemple mais une détection sera plus significative dans des direction orthogonales à la stratégie de balayage.

Au moment ou ce manuscrit est rédigé, il y a encore des résidus à grande échelle angulaire, sur les différences entre les cartes du ciel construites avec des balayages effectués à un an d'intervalle (voir la Sec. 6.2.4 dans Planck Collaboration VIII 2016). Ces résidus concernent le catalogue *Planck* 2015 et sont problématiques pour l'analyse scientifique à grande échelle angulaire, principalement l'étude de l'époque de la réionisation. L'ADC est toujours pointé comme la source dominante de ces différences, et l'analyse des lignes 4 K indique que leur modulation par les opérations de l'étage à 20 K est potentiellement le facteur limitant dont il faut pousser l'analyse. En ce qui concerne l'étalonnage de la fonction de transfert temporelle à basse fréquence , le modèle analytique qui a été développé dans ce but est supposé l'améliorer significativement, mais nécessite encore du travail pour atteindre son objectif. La leçon à retenir sur le compresseur mécanique, c'est que l'isolation contre les interférences électromagnétiques doit être réalisée avec autant de soins que pour les vibrations mécaniques. Une autre leçon à retenir pour éviter un autre "problème d'ADC" c'est que ces composants doivent être caractérisés au sol de façon très précise et systématique avec une méthode de type rampe. L'étape de la sommation est également problématique et une solution simple pour éviter des conséquences néfastes est d'appliquer la correction d'ADC en vol avant cette sommation. Malgrés tout, il ne faut pas oublier que *Planck* a été une mission couronnée de succès, et que ce n'est pas la dernière des missions cosmologiques. Plusieurs nouvelles missions sont prévues pour analyser plus précisément la multitude d'informations que recèle la polarisation du CMB et sont très attendues comme LiteBIRD.

Conclusions and perspectives

This thesis work started in the context of the end of mission of the *Planck* spacecraft, when the implication of the ADC in the gain variation observed on *HFI* science data was just being discovered. There was less than one year left before the decommissioning of *Planck* and the definitive loss of communication capabilities with the on-board electronics readout needed to get the precious elements for the correction of the ADC nonlinearity effect. This was the start of a long run, and this process involved the work with a large international collaboration of people from many different origin. It led not only to the gathering of a vast amount of knowledge in the various scientific and technical aspects involved in understanding a complex space telescope, but also to rich human interactions and a unique personal adventure. Let us now conclude this new drop poured in the ocean of knowledge.

The first part of this work is dedicated to the characterization and the modeling of the time transfer function of *HFI* bolometric detectors. It has been motivated at the origin by the challenging excess response of the detectors at low frequency. This excess observed on-ground during the calibration campaigns has been extremely difficult to calibrate in-flight because the involved "apparent" time constants were very long, from 50 ms up to one minute scale. The empirical approach used within the *Tau Tiger Team* to build the so-called LFER version of the time transfer function for the deconvolution of science data could fit the detector temporal response but did not provide constraints at low frequency. The main issue was that the bolometers temporal response under AC biasing was approximated with DC biasing theory which is not accurate enough for the ($\approx 10^{-3}$) relative level required by *HFI*. Hence to provide strong constraints for the time transfer function, an analytical modeling of the bolometers temporal response under AC biasing was developed from the simple bolometer model, and the results published in Sauvé and Montier (2016). It appeared that the classical signal processing approach in Fourier domain to the linearised differential equations describing the electro-thermal equilibrium of bolometers was not usable in the case of AC biasing. Therefore a specific method was developed with a custom decomposition of the signals in frequency domain, and the problem could be solved with matrix-based linear algebra. The analytical model built with this method has been validated by comparison with a simulation computed with the `SEB` software doing Runge-Kutta order 4 integration of the differential equations describing the system. The relative agreement between the analytical model and the simulation was about 10^{-3} for a 100 GHz detector. It was shown that a single thermal component is not enough to describe accurately the time transfer function estimated in-flight with the empirical model. The analytical model was extended with a second component which was capable to fit the in-flight time transfer function estimation with a much better agreement. Using only 5 free thermal parameters the fit was within 2.7σ of the estimated error for the empirical model used in the *Planck* 2015 data release. This new version was very close to the goal and described accurately the detector dominant time constant at ≈ 10 ms with the addition of a very long time constant component at ≈ 50 s. The specific temporal response of bolometers under AC biasing could be approximated by the empirical LFER model with more than five thermal components, while the analytical model could achieve a similar result with only two thermal components. These excellent results provide several direct applications to this model, starting with new constraints for fitting the LFER component of the *HFI* detectors, but it requires probably to add a third thermal component and also some work with the beams estimation pipeline. The analytical

model has also the ability to fit the *raw gain* observable of detectors which is a completely new and unused constraint for the time transfer function. It can also be used to improve the ADC nonlinearity correction by providing a model of the *raw gain* which is a critical parameter to this correction.

A complementary approach to the analytical modeling of the bolometer response was to use the glitches caused by cosmic rays energy deposit on the detectors. Glitches are usually a nuisance which cause loss of data, but they can also be used to check the validity of an existing estimation of the time transfer function. In the case presented here the idea was pushed further. A method based on Euclidean geometry was developed to reconstruct directly the impulse response of the detectors from glitches. The main problematic was with the existence of several glitch *families* corresponding to different physical locations of the energy deposit. Therefore the reconstruction method was adapted to separate efficiently the dominant families. It was found that the so-called *short* family produced the impulse response with the best agreement compared to the LFER model estimation of the in-flight time transfer function. At this moment two time transfer function models were in competition, the other one being the so-called JH10. This result has been a strong argument in favor of the LFER model which has eventually been adopted for the time deconvolution of *HFI* science data in the 2013 and 2015 data releases. Later, using the definitive version of the impulse response reconstruction method, the time transfer function estimation from *short* glitches was found to be in agreement within a 1% level of the in-flight time transfer function fitted against the LFER model. Albeit being a powerful tool for the analysis of the thermal response of the bolometers, the impulse responses reconstructed with glitches are limited by systematic effects such as closeness between several *short* families and normalization by the energy of the particles. However it has been shown that the τ_{HP} electronics readout filtering parameter is well constrained, but it depends on the model used to fit the thermal response of the bolometer. In this case the presence of several glitch families corresponding to different thermal responses helps to provide better constraints. A possible extension of this work would be to use the analytical model developed previously in conjunction with impulse responses from glitches. The analytical model is more constraining for the thermal response of the bolometers, thus it would help to provide much better constraints on the electronics readout filtering. Additionally the non-Gaussian noise generated at some extent by the glitch families with slow thermal relaxation (the so-called *long* family) could be removed more efficiently with a combination of the analytical model fitted against them.

In the second part of this work, the ADC nonlinearity effect on science data is characterized and corrected for. This systematic effect could be observed as gain variations in time of science data at about 1.5% level and revealed to be extremely challenging to deal with. The first urgency was to make an accurate characterization of the ADC chip for all the 52 bolometers channels of *HFI* before the spacecraft decommissionning. Two acquisition ground campaigns were run with spare ADC chips to collect information that could be used to build a parametric model for the defects of the ADC used in *HFI*. In addition, Maxwell Technologies provided us with the `SPICE` model of the ADC chip. Albeit a missing file in the model prevented to run simulations for testing fits of in-flight components defects, it provided critical information on the architecture of the component. The data collected with both on-ground campaigns exhibited a pseudo-periodic pattern over 64 consecutive ADC codes which was confirmed by the `SPICE` model analysis. Therefore it was decided to drop the parametric model in favor of an empirical approach based on the so-called 64-code pattern plus some specific estimations at the jump location between two regions of the ADC obeying the pattern. At the notable exception of the 64 codes just after midscale of the ADC which were found to have a different pattern and their defects had to be estimated separately. All these informations were taken into account to tune the so-called EOL sequence which consisted in the acquisition of the Gaussian noise of *fast samples* to calibrate ADC defects on-board for

the 10 remaining monthes. This data was used to estimate the quantization step size of every ADC codes in the range covered by the signal during the mission by the mean of a maximum likelihood method. The final estimation of ADC nonlinearity achieved a relative accuracy of about $\approx 3 \times 10^{-3}$ on the quantization steps size for the 64-code pattern, very close to the target requirement of $\approx 1 \times 10^{-3}$. However the 64 codes after midscale that cannot benefit from the 64-code pattern are a limiting factor to the overall result attainable with the final ADC nonlinearity correction. This issue, in the absence of other sources of bias such as the 4 K lines, would be easy to work because these 64 codes could be constrained by using ADC nonlinearity residuals on the science data. Another possibility to better constrain these 64 codes would be to use the science data acquired during the 10 monthes of the EOL sequence, when *HFI* above 0.1 K measured only noise which has been insufficiently exploited.

The accurate characterization of the ADC in-flight was a great challenge, but this information alone could not be used to correct for ADC nonlinearity on science data because of the summation in the DPU of the *fast samples* to produce science data. This step destroys the critical information of exact ADC codes values that are the only data relevant to the in-flight ADC defects estimation. As a consequence these codes values had to be estimated by guessing which was the analog signal producing them. It required an accurate modeling of the electronics readout chain response to the sky power at the input of the bolometer. This model combined to the accurate estimation of ADC defects was used to produce the correction for science data. The validation of this correction required an estimation of the level of ADC nonlinearity residuals on science data. For this purpose ADC nonlinearity observable were developed by comparing the signal from both modulation parities which are acquired in different regions of the ADC. These so-called *half parity* observables were very specific to the ADC nonlinearity effect because the ADC is the only known source of discrepancy, excepted for the noise, between the skies scanned with each parity. Using them it was found that the first version of the correction reduced significantly the initial gain variations down to a level of about 0.1% for some channels. Further simulations of the ADC nonlinearity effect were in very good agreement with the observation on science data, thus validating the correction model and the half parity observables methodology. However this correction was found to be limited by a source of spurious signal originating from the 4 K stage mechanical cooler and this signal was called the 4 K lines. Hence it required a specific characterization to be included in the correction model.

The 4 K lines appear on science data as four sharp lines in the frequency domain, but in practice there is a forest of harmonics on the *fast samples* signal before the summation by the DPU producing science data samples. Once again the summation step destroys critical information by folding the harmonics above the modulation frequency. The inclusion of this signal in the analog signal model requires the knowledge of its time domain shape which extends over nine half periods of the modulation. The main problem is that only contiguous acquisition of *fast samples* on a given channel can provide the full knowledge of the 4 K lines period, but this data is not available during the mission. As a consequence the 4 K lines period can only be estimated from the available observables which are the science data folded versions and one *fast samples* period captured each 101 seconds. The analysis of the two contiguous *fast samples* acquisition sequences captured during the CPV, showed no significant correlations between channels that could be used to model the 4 K lines throughout the mission. Hence a fall-back solution was to estimate the three dominant 4 K lines on the *fast samples* with the available observables. This solution required to calculate for every stable pointing period 18 correction functions for each configuration of the science data samples relative to the 4 K lines and the modulation. It provided a significant improvement to the correction of science data, bringing down the residual gain variations level on most channels below 10^{-3}. This was not sufficient to reach the target requirement for gain variations of 10^{-4}. A new solution was developed starting with a MCMC analysis of the solution space for

4 K lines and including half parity gain observables in the likelihood function. It was shown that the analog signal model could be simplified by using the same modulation shape (aka the *raw constant*) throughout the whole mission. With this consideration it was possible to fit a single *raw constant* by minimizing the half parity observables over the mission. Combined with an estimation of the 4 K lines provided by Guillaume Patanchon it was possible to correct the gain variations down to a level close to 10^{-4}. However the yearly difference of full-sky surveys scanned one year apart showed no significant improvement over the previous correction, and for this reason an empirical approach (`SRoll`) is used since 2016 to reduce the differences between sky maps from different detector channels. One possible reason for the absence of improvement on the yearly differences is real gain variations of the electronics readout chain. It is also possible that the likelihood method used to estimate ADC defects leaks the noise distribution in the estimates, and a different method for this estimation has been suggested. It also remains to characterize the impact of the significant modulation of the 4 K lines by the 20 K stage sorption cooler operations which has a typical time constant of 15 minute.

In the third part of this work, the impact of systematic effects on the scientific analysis of *HFI* maps is studied. For the (Planck Collaboration Int. XIX 2015) paper, the propagation of ADC nonlinearity to the polarized emission of the Galactic thermal dust has been studied with a simulation. The analog signal model and the ADC characterization described in the second part of this work have been used to produce the simulation. It has been estimated that the bias on the polarized light fraction was about 10% and the one on the polarization angle was about 5° in the Galactic plane. It has also been seen that the ADC nonlinearity residuals in the maps of Stokes parameters degenerate strongly with the bandpass mismatch (BPM) leakage. A similar study should be performed on the maps of the *Planck* 2015 data release and a special care should be taken in the case of BPM leakage because of the degeneracy. Next, in the case of the analysis of the angular power spectrum of the CMB, the residuals from the ADC nonlinearity effect were inferred from the *HFI* science data. For this purpose a method based on half-parity maps has been developed and validated with simulations. The sum of the two half-parity maps contains mainly noise and ADC nonlinearity residuals, allowing the analysis of ADC nonlinearity residuals on the angular power spectrum. It was observed that these residuals introduce an excess of power with a $1/\ell$ shape at low multipole for $\ell \in [1, 20]$. Thus with the correction used for the *Planck* 2015 data release, the ADC nonlinearity residuals are not expected to bias significantly the TT spectrum of the CMB for $\ell > 20$. However the residual bias at multipole $\ell = 1$ has been estimated to be about five time the expected noise level, therefore it is significant for the photometric calibration of detectors. In the case of polarization, the analysis of half-parity maps at 143 GHz showed that the $1/\ell$ residual is problematic for the analysis of the EE spectrum on which it appears one order of magnitude above the expected noise level. Hence it is necessary to improve the ADC nonlinearity correction for studying the reionization epoch.

Finally a preliminary study was performed for the detection of the annihilation between surviving domains of antimatter at the recombination epoch. In theory, the annihilation interface between matter and antimatter domains leaves an imprint on the *Planck* tSZ map at a level of about $y \approx 1 \times 10^{-6}$ visible as ribbons with an angular width of about $6'$. This effect is at the detection limit on the *Planck* tSZ map whose RMS of noise is about 1×10^{-6} with a $10'$ resolution. It happens that the *Planck* scanning strategy leakage into this map is degenerate with the expected annihilation signature. The study of this leakage was performed with a specifically taylored partition of the spherical harmonics which selects those dominantly oriented in the co-scan and the cross-scan directions. It was measured a median excess of signal of 14% (16%) in the co-scan direction of the angular power spectrum for the map produced with the `NILC` (`MILCA`) algorithm. The scanning strategy gives a 2 dimension structure to the noise on the sphere whose impact on the angular power spectrum

has been characterized with simulations. This structure has to be taken into account when looking for the signature of annihilation at the limit of the detection level. The search for a signature may be performed with a wavelet based analysis, but a detection would be more significant in the cross-scan direction.

At the time when this manuscript is written, there are still large angular scale residuals in the difference of *HFI* sky surveys maps scanned one year apart (see Sec. 6.2.4 in Planck Collaboration VIII 2016). These residuals found in the *Planck* 2015 data release are supposed to be mainly due to ADC nonlinearity. This is yet an issue for science at large angular scale, mainly the study of the epoch of reionisation. The analysis of 4 K lines indicates that their modulation by the 20 K stage cooler operation is potentially the sticky point that need to be further explored. For the time transfer function calibration at low frequency, the analytical model developed for this purpose is expected to improve this aspect significantly, but requires some work to achieve this objective. The lesson to be learned about the 4 K mechanical compressor is that EMC interferences must be shielded with as much care as for mechanical vibrations. Another lesson to be learned to avoid a new "ADC nonlinearity issue" is that ADC chips should always be accurately characterized on-ground with a ramp-like method. The summation step is also problematic and an easy fix to avoid worst case issues is to apply the ADC nonlinearity correction in-flight before the summation. Finally *Planck* was a great and successful mission, but this is not the last of the cosmological missions. Several new experiments are much awaited and will analyse more precisely the information that the CMB polarization holds such as LiteBIRD.

Appendix A

Building the analytical response in frequency domain of AC biased bolometers

This article was published in Experimental Astronomy (Sauvé and Montier 2016)

Experimental Astronomy manuscript No.
(will be inserted by the editor)

Building the analytical response in frequency domain of AC biased bolometers

Application to *Planck/HFI*

Alexandre Sauvé · Ludovic Montier

the date of receipt and acceptance should be inserted later

Planck

Abstract CONTEXT: Bolometers are high sensitivity detector commonly used in Infrared astronomy. The *HFI* instrument of the *Planck* satellite makes extensive use of them, but after the satellite launch two electronic related problems revealed critical. First an unexpected excess response of detectors at low optical excitation frequency for $\nu < 1$ Hz, and secondly the Analog To digital Converter (ADC) component had been insufficiently characterized on-ground. These two problems require an exquisite knowledge of detector response. However bolometers have highly nonlinear characteristics, coming from their electrical and thermal coupling making them very difficult to modelize.

GOAL: We present a method to build the analytical transfer function in frequency domain which describe the voltage response of an Alternative Current (AC) biased bolometer to optical excitation, based on the standard bolometer model. This model is built using the setup of the *Planck/HFI* instrument and offers the major improvement of being based on a physical model rather than the currently in use had-hoc model based on Direct Current (DC) bolometer theory.

METHOD: The analytical transfer function expression will be presented in matrix form. For this purpose, we build linearized versions of the bolometer electro thermal equilibrium. And a custom description of signals in frequency is used to solve the problem with linear algebra. The model performances is validated using time domain simulations.

RESULTS: The provided expression is suitable for calibration and data processing. It can also be used to provide constraints for fitting optical transfer function using real data from steady state electronic response and optical response. The

A. Sauvé
IRAP (CNRS)
9, avenue du Colonel Roche - BP 44346 - 31028 Toulouse Cedex 4 - FRANCE
Tel : +33 5 61 55 66 66
Fax : +33 5 61 55 86 92
E-mail: asauve@gmail.com

L. Montier
IRAP (CNRS)
E-mail: ludovic.montier@irap.omp.eu

accurate description of electronic response can also be used to improve the ADC nonlinearity correction for quickly varying optical signals.

Keywords Planck, HFI, bolometer, method, analytical model, transfer function

Contents

1 Introduction

Bolometers are high sensitivity thermal detectors commonly used in astronomy in the domain of infrared to sub-millimeter wavelengths. They are basically semi conductor thermometers, connected to a heat sink, which impedance vary with temperature. AC biased bolometers have been extensively used for the last ten years in balloon borne and space experiments as in the *Planck* satellite, mainly for their good performances in regard to low frequency $1/f$ noise. However AC biased bolometers detectors are still described based on DC theory by Holmes et al. (2008). Some work has been done by Catalano et al. (2010) for optimizing the AC biasing of bolometers in the case of the *Planck/HFI* instrument, but without describing the shape of the electronic response.

The *Planck* experiment (Tauber et al., 2010), designed to observe the Cosmic Microwave Backgroud (CMB), reached an unprecedented sensitivity with $\frac{\Delta T}{T}$ better than 10^{-5} for the CMB anisotropies observation. With respect to this objective, two problems related to electronics have revealed as critical after the satellite launch: the low frequency excess response (LFER) and Analog to Digital Converter (ADC) nonlinearity.

First the detectors response to optical excitation exhibited an excess response at low frequencies for $\nu < 1Hz$ (Planck Collaboration et al., 2015). The main culprit for this excess is intermediate components in the thermal path to the heat sink operating at 100 mK. These components produce also specific response to energy deposit from particles (Catalano et al., 2014). The thermal model have been extended with a chain of order 1 low pass filters to build an ad-hoc transfer function model Planck Collaboration (2014) which described well the detector response at first order. However the last version of the model needs to fit up to seven thermal components (Planck Collaboration et al., 2015).

The second in flight issue is the ADC nonlinearity which has been insufficiently characterized on ground. This systematic effect becomes very difficult to correct in the *Planck/HFI* case, because signal is averaged onboard over 40 samples of the modulation half period, before being sent to the ground. A very good knowledge

of time domain signal at 40 times the modulation frequency is then required to apply the ADC nonlinearity correction, Currently an empirical model based on the hypothesis of slowly varying signal is in use by Planck Collaboration et al. (2015), with limited performances in the case of bright and quickly varying signal.

In order to address these very demanding objectives, the present article describes, an analytical model built in frequency domain for the voltage response to an optical excitation for an AC biased bolometer. This model is based on the physical model of the bolometer and has a selectable bandwidth limit. The only free parameter is the optical excitation angular frequency ω.

We will first describe in Sect. 2 the electrical and thermal model of the *Planck/HFI* bolometer detectors. Then in Sect. 3 we show how to build a suitable linearized version of the electro-thermal equilibrium equations. The solving of the equilibrium will be done in frequency domain and involve convolution of frequency vectors. To make the solving possible, In Sect. 4 a custom frequency representation of signals and matrix formalism is presented. Finally in Sect. 5 the model performances will be compared to the time domain simulations computed with the Simulation for the Electronic of Bolometers (`SEB`) tool, used in the *Planck* consortium since 2007.

2 Bolometer model

In this section, we will describe the *Planck/HFI* detector model, which involves the electronic design and the thermal model. The NTD-Ge semi-conductor bolometers used have a negative thermal response. Which means their impedance decrease when their temperature increase. So when temperature raise from incoming optical radiative power the dissipated heat from Joule effect decrease. In this case, the *electro-thermal coupling* helps the system to reach a stable equilibrium. The coupled electronic and thermal equations will be used in Sect. 3 to build the linear response.

2.1 Bias circuit

The *Planck/HFI* bias circuit is presented in Fig. 1 which is an excerpt from the readout chain described in *Planck* Pre-Lauch paper (Lamarre et al., 2010, Sect. 4) and in Montier (2005).

The differential equation describing the voltage of the bolometer is

$$V(t) + R(t)(C_{eq} + C_s)\frac{dV(t)}{dt} = R(t)C_{eq}\frac{dV_{\mathrm{bias}}(t)}{dt}\,, \tag{2.1}$$

where V_{bias} is the input bias voltage, $V(t)$ is the bolometer measured output voltage, $R(t)$ is the bolometer real impedance, C_1 and C_2 are the bias capacitances, with $C_{eq} = C_1C_2/(C_1+C_2)$, and C_s is a stray capacitance. All quantitative values are referenced in Table 1.

Using values from Table 1, $C_s \simeq 60C_{eq}$. The very high value of C_s relative to C_{eq} has a large impact on full circuit design and frequency response as described in Piat (2000, Sect. 4.3). The stray capacitance occurs from the length of cable between the detector and the JFet Box of the electronic readout chain of *Planck/HFI* and it applies a low pass filter on the bias current I. It is a design

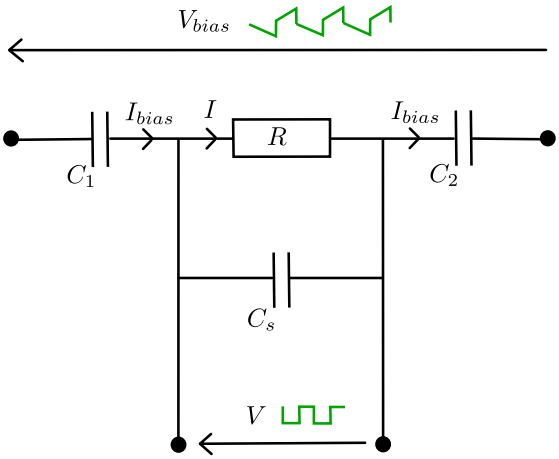


Fig. 1 Bolometer bias circuit of *Planck/HFI*. The schematic is reduced to the minimal set involved in the description of electro-thermal coupling. R is the bolometer impedance. C_1 and C_2 are the bias capacitances, doing the derivation operation on the bias voltage. C_s is the stray capacitance coming from cables length. $V_{\rm bias}$ is the circuit bias voltage composed of a triangle wave plus a square wave added to compensate for C_s. I_{bias} is the bias current provided to the $R+C_s$ couple. I is the current flowing through the bolometer, it is designed to be as constant in absolute value as possible, to mimic a DC bias. And V is the output bolometer voltage.

goal to mimic DC bias (Lamarre et al., 2010) with a square shape for I. It allows to have a joule effect as stable as possible and then reduce variations of R which maximize the linearity of the response. This is why the bias voltage $V_{\rm bias}$ is composed of a triangle plus a *compensation* square wave. The square wave allow to compensate for C_s by stabilizing quickly the current through the bolometer. The major drawback from the square wave addition is a frequency shape in $1/f$, while the triangle wave alone has a frequency shape in $1/f^2$. As an effect, the number of harmonics needed to describe the signal will be significantly increased, as for the computational complexity.

2.2 Thermal model

The thermal equilibrium of the bolometer is described by the standard theoretical bolometer model. We will use the conventions proposed by Piat et al. (2006). The thermal equilibrium, as represented in Fig. 2, can be written

$$C_{th}\frac{dT}{dt} = P_J + P_{opt} - P_{sink}, \tag{2.2}$$

where C_{th} is the bolometer heat capacity, T is the bolometer temperature, $P_J = \frac{V^2}{R}$ is the power dissipated through the bolometer by Joule effect, P_{sink} is the power dissipated through the heat sink, and P_{opt} is the optical radiative power falling on the bolometer.

The bolometer heat capacity is given as a function of temperature by (Mather, 1984)

$$C_{th}(T) = C_0 T^{\gamma}\,, \tag{2.3}$$

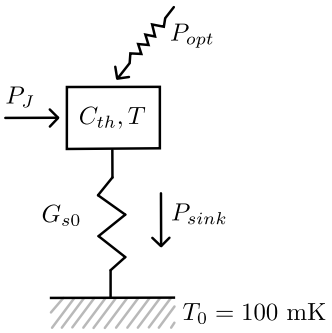


Fig. 2 Standard bolometer thermal model. The bolometer itself is represented by the central box with heat capacity C_{th} and temperature T. P_{opt} is the total optical radiative power falling on the detector. P_J is the Joule effect power produced by bias current. P_{sink} is the power flowing out of the bolometer via the thermal link with static thermal conductance G_{s0}. The heat sink at temperature T_0 is supposed to have an infinite thermal capacity.

where γ and C_0 are experimental measurements obtained during calibration campaigns at Jet Propulsion Laboratory (Holmes et al., 2008). Numerical values are referenced in Table 2.

In the absence of electric field, the variation of bolometer impedance can be expressed as a function of its temperature from the simple bolometer model

$$R(T) = R_G e^{\sqrt{\frac{T_G}{T}}}, \tag{2.4}$$

where R_G and T_G are experimental measurements.

The power dissipated through the heat sink is a function of bolometer temperature, it can be written as (Mather, 1982)

$$P_{sink}(T) = \frac{G_{s0}}{T_{ref}^{\beta}(\beta+1)}(T^{\beta+1} - T_0^{\beta+1}), \tag{2.5}$$

with β, G_{s0} the static thermal conductance, being experimental measurements. T_{ref} is a reference temperature and the heat sink temperature is fixed to $T_0 = 100mK$.

3 Linearizing differential equations

In this section, we describe how electrical and thermal equilibriums can be rewritten to obtain a linear system of two equations of V and R only. We make a clear distinction between steady state and optical excitation signals because we need to solve for the two cases separately. For this purpose Taylor expansions are used to reach a target goal of 10^{-3} relative precision. This is a reasonable objective, because as we will see in Sect. 5.2, the simulations show the response to an optical excitation is linear at 10^{-5} level.

These linearized forms will be used in Sect. 4 where we describe a method to solve the problem in frequency domain.

3.1 Signals decomposition

Let start by making a clear distinction between signals in steady state, and the ones generated by optical excitation. We make the hypothesis that the incoming optical radiative power is the sum of a constant bacground $\langle P_{opt} \rangle$ and a small monofrequency optical excitation term $\widetilde{P}_{opt}$.

The bolometer response to $\widetilde{P}_{opt}$ is linear to a good approximation, so we make the hypothesis that we can write its voltage as

$$V(t) = \langle V \rangle + \overline{V}(t) + \widetilde{V}(t) + \mathcal{O}(\widetilde{V}^2)\,, \tag{3.1}$$

using the notations

- $\langle V \rangle$ for the average value of V (in this case $\langle V \rangle = 0$)
- $\overline{V}$ for the steady state modulation harmonics
- $\widetilde{V}$ for the harmonics appearing in response to $\widetilde{P}_{opt}$

The second order terms in $\mathcal{O}(\widetilde{V}^2)$ can be neglected, see Sect. 5.1 for quantitative values.

The same hypothesis and notations will be applied to $T(t)$, $R(t)$, $C_{th}(t)$, $P_{sink}(t)$ and $P_J(t)$. These notations allow to distinguish easily the order of the terms in the equations as we do have in practice the generic relation $\langle T \rangle \gg \overline{T} \gg \widetilde{T}$. Numerical values will be detaild in Sect. 5.1. Consequently, from now, all terms containing a product of two optical excitation component like $\widetilde{V}\widetilde{R}$ will be considered as negligible, and nonlinear.

3.2 Electrical equilibrium

The electrical equilibrium is already a function of V and R. Using previously defined notations, we separate the steady state terms from the optical excitation terms. Then Eq. (2.1) reads

$$\overline{V} + (\langle R \rangle + \overline{R})(C_{eq} + C_s)\dot{\overline{V}} = (\langle R \rangle + \overline{R})C_{eq}\dot{V}_{bias}\,, \tag{3.2}$$

and

$$\widetilde{V} + (C_{eq} + C_s)\left((\langle R \rangle + \overline{R})\dot{\widetilde{V}} + \widetilde{R}\dot{\overline{V}}\right) = \widetilde{R}C_{eq}\dot{V}_{bias} + \mathcal{O}(\widetilde{R}\widetilde{V})\,. \tag{3.3}$$

3.3 Thermal equilibrium

The thermal equilibrium Eq. (2.2) is a function of T, V and R. In this section, we rewrite its terms as linearized versions of V and R only. To do so we first express T as a function of R.

3.3.1 Expressions of T from R

From Eq. (2.4) the value of T is

$$T(R) = \frac{T_g}{\left(\ln R - \ln R_g\right)^2} \tag{3.4}$$

As motivated by simulation analysis (see Sect. 5.1), the time derivative of optical excitation component has 15% error when using 1st order Taylor expansion, so we will provide coefficients obtained from Eq. (2.4) for the two first orders which reduces error to 0.2%. These coefficients reads

$$\begin{aligned} \left.\frac{\partial T}{\partial R}\right|_{\langle T\rangle} &= a_{1T} = -\frac{2T_g}{\langle R\rangle\left(\ln\langle R\rangle - lnR_g\right)^3}\,, \\ \frac{1}{2!}\left.\frac{\partial^2 T}{\partial R^2}\right|_{\langle T\rangle} &= a_{2T} = \frac{T_g\left(\ln\langle R\rangle - lnR_g + 3\right)}{R^2\left(\ln\langle R\rangle - \ln R_g\right)^4}\,, \end{aligned} \tag{3.5}$$

so we can write

$$T(\langle R\rangle + \delta R) = \langle T\rangle + a_{1T}\delta R + a_{2T}\delta R^2 + \mathcal{O}(\delta R^3) \tag{3.6}$$

For steady state plus optical excitation component $\delta R = \overline{R} + \widetilde{R}$, then we can write the Taylor expansion of temperature variations as

$$T(\langle R\rangle + \delta R) = \langle T\rangle + \underbrace{a_{1T}\overline{R} + a_{2T}\overline{R}^2}_{\overline{T}} + \underbrace{a_{1T}\widetilde{R} + 2a_{2T}\overline{R}\widetilde{R}}_{\widetilde{T}} + \mathcal{O}(\delta R^3)\,, \tag{3.7}$$

where the $a_{2T}\widetilde{R}^2$ term has been neglected in the expression.

And finally the time derivatives of $\overline{T}$ and $\widetilde{T}$, which we will need for linear expression of Eq. (2.2), can be written from the derivation of Eq. (3.7) over time as

$$\begin{aligned} \dot{\overline{T}} &= a_{1T}\dot{\overline{R}} + 2a_{2T}\dot{\overline{R}}\,\overline{R} + \mathcal{O}(\overline{R}^3)\,, \\ \dot{\widetilde{T}} &= a_{1T}\dot{\widetilde{R}} + 2a_{2T}\left(\dot{\overline{R}}\widetilde{R} + \overline{R}\,\dot{\widetilde{R}}\right) + \mathcal{O}(\widetilde{R}^3)\,. \end{aligned} \tag{3.8}$$

3.3.2 Heat quantity variations

We now build the linear version of heat capacity C_{th} defined in Eq. (2.3) as a function of R. Its first order Taylor expansion, as a function of T, is

$$C_{th}(\langle T\rangle + \delta T) = \langle C_{th}\rangle + \underbrace{\left.\frac{\partial C_{th}}{\partial T}\right|_{\langle T\rangle}}_{a_{1C}}\delta T + \mathcal{O}(\delta T^2)\,, \tag{3.9}$$

with $\langle C_{th}\rangle = C_0\langle T\rangle^\gamma$ and $a_{1C} = \gamma C_0\langle T\rangle^{\gamma-1}$.

To build the expression as a function of R at first order, we can replace δT by $a_{1T}\delta R$ using Eq. (3.5). Then the 1st order Taylor expansion becomes

$$C_{th}(\langle T\rangle + \delta T) = \langle C_{th}\rangle + a_{1C}\delta R + \mathcal{O}(\delta R^2)\,, \tag{3.10}$$

with $a_{1C} = \left(\gamma C_0 \langle T \rangle^{\gamma-1}\right) a_{1T}$. Using $\delta R = \overline{R} + \widetilde{R}$ we can express C_{th} for steady state plus an optical excitation as

$$C_{th}(\langle T \rangle + \delta T) = \langle C_{th} \rangle + \underbrace{a_{1C}\overline{R}}_{\overline{C}_{th}} + \underbrace{a_{1C}\widetilde{R}}_{\widetilde{C}_{th}} + \mathcal{O}(\widetilde{R}^2)\,. \tag{3.11}$$

Now the linear version of $C_{th}\dot{T}$ term of Eq. (2.2) reads

$$C_{th}\dot{T} = \left(\overline{C}_{th} + \widetilde{C}_{th}\right)\left(\dot{\overline{T}} + \dot{\widetilde{T}}\right)\,. \tag{3.12}$$

We can separate the steady state only term from the product. Using Eq. (3.8) the result as a function of R reads

$$\begin{aligned}\overline{\left(C_{th}\dot{T}\right)} &= \overline{C}_{th}\,\dot{\overline{T}} + \mathcal{O}(\overline{R}^5) \\ &= \left(\langle C_{th} \rangle + a_{1C}\overline{R}\right)\left(a_{1T}\dot{\overline{R}} + 2a_{2T}\dot{\overline{R}}\,\overline{R}\right) + \mathcal{O}(\overline{R}^5) \\ &= \langle C_{th} \rangle a_{1T}\dot{\overline{R}} + \mathcal{O}(\overline{R}^2)\,.\end{aligned} \tag{3.13}$$

where the $\overline{R}^2$ terms have been neglected. The term for an optical excitation is respectively

$$\begin{aligned}\widetilde{\left(C_{th}\dot{T}\right)} &= \overline{C}_{th}\dot{\widetilde{T}} + \widetilde{C}_{th}\dot{\overline{T}} + \mathcal{O}(\widetilde{C}_{th}\widetilde{T}) \\ &= \left(\langle C_{th} \rangle + a_{1C}\overline{R}\right)\left(a_{1T}\dot{\widetilde{R}} + 2a_{2T}\left(\dot{\widetilde{R}}\widetilde{R} + \overline{R}\dot{\widetilde{R}}\right)\right) + a_{1C}\widetilde{R}\left(a_{1T}\dot{\overline{R}} + 2a_{2T}\dot{\overline{R}}\,\overline{R}\right) + \mathcal{O}(\widetilde{R}^2)\end{aligned} \tag{3.14}$$

3.3.3 Heat sink dissipated power

To build the linear expression of the power dissipated trough the heat sink as a function of R, we will consider the first order Taylor expansion of Eq. (2.5) given by

$$P_{sink}(\langle T \rangle + \delta T) = \langle P_{sink} \rangle + \underbrace{\left.\frac{\partial P_{sink}}{\partial T}\right|_{\langle T \rangle}}_{G_d}\delta T + \mathcal{O}(\delta T^2)\,, \tag{3.15}$$

with

$$G_d = G_{s0}\left(\frac{T}{T_{ref}}\right)^{\beta}\,, \tag{3.16}$$

where G_d is the dynamic thermal conductance.

This expression can be written as a function of R, by replacing δT using Eq. (3.7) as

$$P_{sink}(\langle T \rangle + \delta T) = \langle P_{sink} \rangle + \underbrace{G_d a_{1T}\overline{R}}_{\overline{P}_{sink}} + \underbrace{G_d\left(a_{1T}\widetilde{R} + 2a_{2T}\overline{R}\widetilde{R}\right)}_{\widetilde{P}_{sink}} + \mathcal{O}(\widetilde{R}^2)\,. \tag{3.17}$$

3.3.4 Joule effect

By definition

$$P_J = \frac{(\overline{V} + \widetilde{V})^2}{\langle R \rangle + \overline{R} + \widetilde{R}} \,. \tag{3.18}$$

The denominator is of the form $\frac{1}{(x+\epsilon)}$, with $x = \langle R \rangle$ and $\epsilon = \overline{R} + \widetilde{R}$. Its first order Taylor expansion is $\frac{1}{x} - \frac{\epsilon}{x^2}$, which leads to

$$P_J = \underbrace{\frac{\overline{V}^2}{\langle R \rangle} - \frac{\overline{V}^2 \overline{R}}{\langle R \rangle^2}}_{\overline{P}_J} + \underbrace{2\frac{\overline{V}\widetilde{V}}{\langle R \rangle} - \frac{\overline{V}^2 \widetilde{R}}{\langle R \rangle^2} - \frac{2\overline{V}\,\overline{R}\widetilde{V}}{\langle R \rangle^2}}_{\widetilde{P}_J} + \mathcal{O}(\widetilde{R}^2) \,. \tag{3.19}$$

4 Solving in frequency domain

In the following section, we build the optical transfer function describing the detector voltage response to an optical excitation at angular frequency ω. To build the solution in frequency domain, a matrix formalism is developed to work on signals projected in custom Fourier domain basis. The solving is done separately for steady state only first, and then for optical excitation case.

4.1 Matrix formalism

Now that we have linear version of differential equations we show how we can use linear algebra to solve the problem.

For the discrete description of signals in frequency domain, the number of considered modulation harmonics will be fixed to n. The vectors in frequency domain for real signals need to be of size $2n + 1$ including complex conjugates for negative frequencies. Best values for n will be discussed in Sect. 5.3. Steady state signals are described first, signals resulting from optical excitation must be handled differently. After the custom frequency vector basis are defined, the time derivative of signals can be expressed in a simple form.

4.1.1 Steady state frequency vectors

Let consider a steady state signals, $\overline{S}(t)$. It is composed only of modulations harmonics and we can write using complex notation

$$\overline{S}(t) = \sum_{k=-n}^{n} \bar{\mathbf{s}}_k e^{jk\Omega_m t} \,, \tag{4.1}$$

where j is the imaginary unit, and $\Omega_m = 2\pi f_{mod}$ is the modulation angular frequency. The steady state vector resulting from the projection of $S(t)$ in Fourier domain is noted $\bar{\mathbf{s}}$. As $\overline{S}(t)$ is a real valued signal, we do have the relation $\bar{\mathbf{s}}_{-k} = \bar{\mathbf{s}}_k^*$, where z^* is the conjugate of z. This representation is homogenous with the outputs

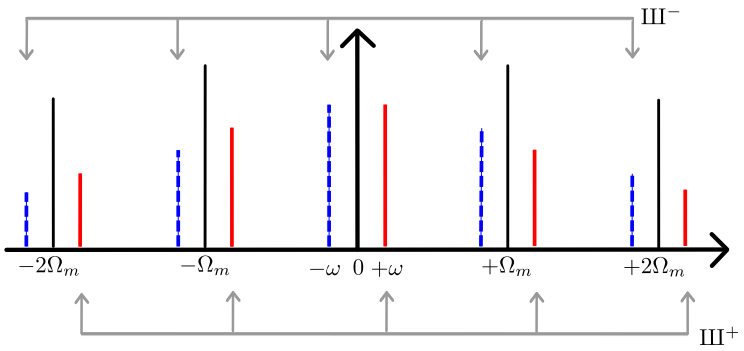


Fig. 3 Schematic for decomposition of a signal $\widetilde{S}_2(t)$ resulting from an optical excitation at frequency ω, with $n = 2$. The red solid lines represent the harmonics on the frequency support Ш^+. While the blue dashed lines represent the harmonics on the frequency support Ш^-. The size difference around each modulation frequency emphasize that considered modulated harmonics result from the system response, and not from the modulation of a single frequency signal.

of the fast Fourier transform algorithm. We note Ш the discrete frequency support of $\overline{\mathbf{s}}$ defining the Fourier basis of size $2n + 1$ for steady state signals description.

The key point to solve electrical and thermal equilibrium is to represent the product of time domain signals. Hopefully the convolution theorem states that a product in time domain is equivalent to circular convolution product in frequency domain. Then let consider the product of two steady state signals, $\overline{S}_1(t)$ and $\overline{S}_2(t)$, with corresponding steady state vectors $\overline{\mathbf{s_1}}$ and $\overline{\mathbf{s_2}}$. Their discrete circular convolution product can be written in matrix notation using the circulant matrix of left vector as

$$\overline{\mathbf{s_1}} \otimes \overline{\mathbf{s_2}} = \boldsymbol{\mathcal{C}}(\mathbf{s_1})\overline{\mathbf{s_2}} = \boldsymbol{\mathcal{C}}(\overline{\mathbf{s_2}})\overline{\mathbf{s_1}}\,, \tag{4.2}$$

where $\boldsymbol{\mathcal{C}}(\overline{\mathbf{s_1}})\overline{\mathbf{s_2}}$ is the matrix product of the circulant matrix $\boldsymbol{\mathcal{C}}(\mathbf{s_1})$ and vector $\overline{\mathbf{s_2}}$. The result of the product is also in Ш.

4.1.2 Optical excitation vectors

Let consider now a product $\overline{S}_1(t)\widetilde{S}_2(t)$. The signal $\widetilde{S}_2(t)$ results from the system response to an optical excitation at angular frequency ω. $\widetilde{S}_2(t)$ is not stricto sensu a modulated signal, because the electro-thermal equilibrium does not build in the form $\overline{S(t)} * \cos(\omega t + \varphi)$. However its harmonics appears at the same frequencies, as shown in Fig. 3. $\widetilde{S}_2(t)$ can be written

$$\widetilde{S}_2(t) = \sum_{k=-n}^{n} \left(\widetilde{\mathbf{s}}^{+}_{\mathbf{2}\ k}\, e^{j(k\Omega_m+\omega)t} + \widetilde{\mathbf{s}}^{-}_{\mathbf{2}\ k}\, e^{j(k\Omega_m-\omega)t} \right) . \tag{4.3}$$

The frequency step between $\widetilde{S}_2(t)$ harmonics is not homogenous. Then the convolution product in frequency cannot be written with a single vector for $\widetilde{S}_2(t)$ as for steady state only case. A classic DFT based method could be used to compute the convolution product. However for an exact computation it would require using a frequency step which is a multiple of Ω_m and ω. This would render the computing

extremely CPU intensive, or infeasible. Because the solving step would involve inversion of huge matrix which is an $\mathcal{O}(n^3)$ operation. So we have to use another solution and split the problem in two.

From the coeficients defined in Eq. 4.3, we can write $\widetilde{S}_2(t) = \widetilde{S}_2^+(t) + \widetilde{S}_2^-(t)$. The product $\overline{S}_1(t)\widetilde{S}_2(t)$ can be written in frequency domain as

$$\overline{S}_1(\nu) \otimes \widetilde{S}_2(\nu) = \overline{S}_1(\nu) \otimes \widetilde{S}_2^+(\nu) + \overline{S}_1(\nu) \otimes \widetilde{S}_2^-(\nu)\,. \tag{4.4}$$

All terms in the right part of the equation have a frequency step of Ω_m. The frequency support for $\widetilde{S}_2^+$ and $\widetilde{S}_2^-$ are noted $\text{Ш}^+ = [\![k\Omega_m + \omega]\!]$ and $\text{Ш}^- = [\![k\Omega_m - \omega]\!]$ with $k \in [\![-n \cdots n]\!]$.

With this partitioning in frequency, the convolution product can be computed in the same way as for steady state, but separately for $\widetilde{\mathbf{s}}_\mathbf{2}^+$ and $\widetilde{\mathbf{s}}_\mathbf{2}^-$. Considering that $\overline{\mathbf{s_1}} \otimes \widetilde{\mathbf{s}}_\mathbf{2}^+$ builds only terms at frequencies $p\Omega_m + (q\Omega_m + \omega) = (p+q)\Omega_m + \omega$, with $p \in [\![-n \cdots n]\!]$, and $q \in [\![-n \cdots n]\!]$. Then it comes

$$\overline{\mathbf{s_1}} \otimes \widetilde{\mathbf{s}}_\mathbf{2}^+ \in \text{Ш}^+\,. \tag{4.5}$$

Respectively

$$\overline{\mathbf{s_1}} \otimes \widetilde{\mathbf{s}}_\mathbf{2}^- \in \text{Ш}^-\,. \tag{4.6}$$

As a consequence, the terms in Ш^+ and Ш^- have to be manipulated separately, when working in frequency domain.

As $\widetilde{S}(t)$ is a real valued signal, we do have by construction $\widetilde{\mathbf{s}}_{\mathbf{2}\ k}^+ = \widetilde{\mathbf{s}}_{\mathbf{2}\ -k}^{-\ *}$. From this consideration the computing of solutions is needed only for $\widetilde{\mathbf{s}}_\mathbf{2}^+$ vectors, as the $\widetilde{\mathbf{s}}_\mathbf{2}^-$ version can be built from it.

4.1.3 Time derivative of vectors

The time domain derivative of a signal can be written from its DFT vector, considering that $\frac{d}{dt}e^{jk\Omega_m t} = jk\Omega_m e^{jk\Omega_m t}$. Then the time derivative operation can be represented using a complex square diagonal matrix of size $2n+1$, and can be written

$$\frac{d}{dt}\overline{\mathbf{s}} = \mathbf{D}\overline{\mathbf{s}}\,, \tag{4.7}$$

with

$$\mathbf{D} = \begin{pmatrix} -jn\Omega_m & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & jn\Omega_m \end{pmatrix}. \tag{4.8}$$

Time domain derivatives for optical excitation vectors $\widetilde{\mathbf{s}}^+$ and $\widetilde{\mathbf{s}}^-$ respectively, are defined with matrix $\mathbf{D}^+$ and $\mathbf{D}^-$ respectively. Their diagonal elements are $j(k\Omega_m + \omega)$ and $j(k\Omega_m - \omega)$ respectively, for $k \in [\![-n \cdots n]\!]$.

4.1.4 indexing of vectors

The $2n+1$ vector elements are referenced by their harmonic index varying monotonically from $-n$ to n. The harmonic index 0 will be considered to be at center of vector. In practice many DFT implementations store negative elements at the end of the matrix. This is completely equivalent for the expressions presented in this article.

4.2 Steady state solving

The steady state is characterized by $\overline{V}$ and $\overline{R}$, their values must be computed first as they are needed for solving the optical excitation expression. For steady state, the electrical equilibrium Eq. (3.2) reads in matrix form

$$\overline{\mathbf{v}} = (\mathbf{I_d} + (C_s + C_{eq})\boldsymbol{\mathcal{C}}(\langle R\rangle + \overline{\mathbf{r}})\mathbf{D})^{-1}\, C_{eq}\boldsymbol{\mathcal{C}}(\langle R\rangle + \overline{\mathbf{r}})\mathbf{D}\mathbf{V}_{\text{bias}}\,. \tag{4.9}$$

where $\mathbf{I_d}$ is the identity matrix.

Impedance variations $\overline{R}$ are of order 10^{-3} of $\langle R\rangle$ (see Sect. 5.1). So the electrical equilibrium behave at first order as a static impedance circuit and the vector $\overline{\mathbf{v}}$ can be computed directly from Eq. 4.9 by setting $\overline{\mathbf{r}}$ to zero. The value of $\langle R\rangle$ is supposed to be known. This is for real life detectors a parameter generally obtained by direct measurement.

The value of the vector $\overline{\mathbf{r}}$ is obtained from thermal equilibrium Eq. (2.2). With steady state notations thermal equilibrium reads

$$\overline{\left(C_{th}\dot{T}\right)} = \langle P_J\rangle + \overline{P}_J - \langle P_{sink}\rangle - \overline{P}_{sink}\,. \tag{4.10}$$

Using expressions defined in Eq. (3.13), (3.17) and (3.19) the matrix form is

$$\overline{\mathbf{r}} = \left(\langle C_{th}\rangle a_{1T}\mathbf{D} + \frac{\boldsymbol{\mathcal{C}}(\boldsymbol{\mathcal{C}}(\overline{\mathbf{v}})\overline{\mathbf{v}})}{\langle R\rangle^2} + G_d a_{1T}\mathbf{I_d}\right)^{-1}\left(\frac{\boldsymbol{\mathcal{C}}(\overline{\mathbf{v}})\overline{\mathbf{v}}}{\langle R\rangle} - \langle P_{sink}\rangle + \langle P_{opt}\rangle\right)\,, \tag{4.11}$$

where the value of $P_{sink}(\langle T\rangle)$ is obtained using Eq. (2.5) and (3.4). Let's stress here that adding a scalar value like $\langle P_{opt}\rangle$ to a steady state vector is equivalent to adding the value to the harmonic index 0 of the vector.

Once the vector $\overline{\mathbf{r}}$ is computed, it can be used to update the value of $\overline{\mathbf{v}}$ by using Eq. (4.9). This will improve the precision of $\overline{\mathbf{v}}$ by an order.

A second iteration for $\overline{\mathbf{r}}$ is not usefull because the value used for $\overline{C}_{th}\,\dot{\overline{T}}$ in the final version of Eq. (3.13) cancels second order terms in $\overline{R}^2$. So it would not bring better precision.

4.3 Optical excitation solving

Electrical equilibrium equation 4.12 reads in matrix form

$$\begin{cases}\widetilde{\mathbf{r}}^{+} = \mathbf{E}^{+}\widetilde{\mathbf{v}}^{+}\\ \widetilde{\mathbf{r}}^{-} = \mathbf{E}^{-}\widetilde{\mathbf{v}}^{-}\,,\end{cases} \tag{4.12}$$

with

$$\begin{cases}\mathbf{E}^{+} = \mathbf{E1}\left(\mathbf{I_d} + (C_{eq} + C_s)\langle R\rangle\mathbf{D}^{+}\right)\\ \mathbf{E}^{-} = \mathbf{E1}\left(\mathbf{I_d} + (C_{eq} + C_s)\langle R\rangle\mathbf{D}^{-}\right)\\ \mathbf{E1} = \{C_{eq}\boldsymbol{\mathcal{C}}(\mathbf{D}\mathbf{V}_{\text{bias}}) - (C_{eq} + C_s)\boldsymbol{\mathcal{C}}(\mathbf{D}\overline{\mathbf{v}}))\}^{-1}\,.\end{cases} \tag{4.13}$$

The thermal equilibrium Eq. (2.2) reads using optical excitation notations

$$\widetilde{\left(C_{th}\dot{T}\right)} = \widetilde{P}_J - \widetilde{P}_{sink} + \widetilde{P}_{opt}\,. \tag{4.14}$$

Which can be written in matrix form using Eq. (3.14), (3.19), (3.17) and (4.12). From this expression the analytical transfer function appears in the form of two $2n+1$ square matrices noted $\boldsymbol{\mathcal{F}}^+$ and $\boldsymbol{\mathcal{F}}^-$ in

$$\begin{cases} \widetilde{\mathbf{v}}^+ = \underbrace{\left[\left(\mathbf{C_1} + \mathbf{C_2}\mathbf{D}^+ - \mathbf{J_2} + \mathbf{S}\right)\mathbf{E}^+ - \mathbf{J_1}\right]^{-1}}_{\boldsymbol{\mathcal{F}}^+} \widetilde{\mathbf{p}}^+ \\ \widetilde{\mathbf{v}}^- = \underbrace{\left[\left(\mathbf{C_1} + \mathbf{C_2}\mathbf{D}^- - \mathbf{J_2} + \mathbf{S}\right)\mathbf{E}^- - \mathbf{J_1}\right]^{-1}}_{\boldsymbol{\mathcal{F}}^-} \widetilde{\mathbf{p}}^-, \end{cases} \tag{4.15}$$

with

$$\begin{cases} \mathbf{C_1} = \boldsymbol{\mathcal{C}}\left[\{(2\langle C_{th}\rangle a_{2T} + a_{1C}a_{1T})\mathbf{I_d} + 2a_{1C}a_{2T}\boldsymbol{\mathcal{C}}(\langle R\rangle + \bar{\mathbf{r}})\}\,\mathbf{D}\bar{\mathbf{r}} + 2a_{1C}a_{2T}\boldsymbol{\mathcal{C}}(\mathbf{D}\bar{\mathbf{r}})\bar{\mathbf{r}}\right] \\ \mathbf{C_2} = \boldsymbol{\mathcal{C}}\left[\langle C_{th}\rangle a_{1T} + \{(2\langle C_{th}\rangle a_{2T} + a_{1C}a_{1T})\mathbf{I_d} + 2a_{1C}a_{2T}\boldsymbol{\mathcal{C}}(\bar{\mathbf{r}})\}\,\bar{\mathbf{r}}\right] \\ \mathbf{J_1} = \frac{2}{\langle R\rangle}\boldsymbol{\mathcal{C}}(\bar{\mathbf{v}}) - \frac{2}{\langle R\rangle^2}\boldsymbol{\mathcal{C}}(\bar{\mathbf{v}})\bar{\mathbf{r}} \\ \mathbf{J_2} = -\frac{1}{\langle R\rangle^2}\boldsymbol{\mathcal{C}}(\boldsymbol{\mathcal{C}}(\bar{\mathbf{v}})\bar{\mathbf{v}}) \\ \mathbf{S} = G_d\boldsymbol{\mathcal{C}}\left(a_{1T} + 2a_{2T}\bar{\mathbf{r}}\right) \end{cases} \tag{4.16}$$

And where $\widetilde{\mathbf{p}}^+$ and $\widetilde{\mathbf{p}}^-$ describe the, real valued, optical excitation signal at angular frequency ω. Considering $\widetilde{P}_{opt}(t) = ze^{j\omega t} + z^*e^{-j\omega t}$, then $\widetilde{\mathbf{p}}^+$ will only have one nonzero coefficient at modulation harmonic index 0 with value z, respectively $\widetilde{\mathbf{p}}^-$ will only have one nonzero coefficient at the same index with value z^*. **C1**, **C2** are heat quantity variation matrices. **J1**, **J2** are Joule effect matrices. And **S** is heat sink dissipated power matrix. The expression is built in a way to make clearly visible which matrices depends on ω (with a $+$ or $-$ sign as exponent) from the ones that can be computed once and for all.

5 Results

In this section we will present the simulation tool used to build the time domain signal response to optical excitation. Then transfer function results will be validated from the simulation results. And finally we will discuss the transfer function shape in the data processing low sampling frequency case.

5.1 Simulation setup

The tool used to build realistic simulations conform to *Planck/HFI* readout signal is the Simulation for Electronic and Bolometers (`SEB`) which has been developed at IRAP. It is an IDL implementation of the electrical and thermal differential equations of the bolometer with the bias circuit presented in Sect. 2.1. The numerical integration is performed using finite differences with Runge Kuta order 4 method with 10000 points per modulation period. Simulations have been computed with the setup of the 1st 100GHz channel of *Planck/HFI* (`00_100-1a`) with the numerical values presented in Table 1 and Table 2 of Appendix C. Simulation

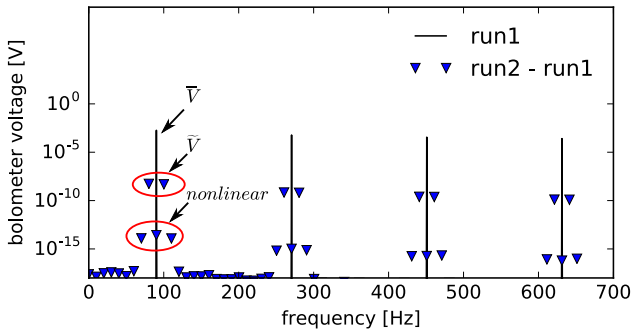


Fig. 4 Bolometer optical excitation linearity. Black lines are bolometer voltage for run1, only the modulation harmonics are visible. Blue triangles are the excitation residuals of bolometer voltage for run2 - run1. Optical excitation signal $\widetilde{V}$ appear at frequencies $(2k+1)\Omega_m \pm \omega$. Nonlinear excitation harmonics $\mathcal{O}(\widetilde{V}^2)$ are visible about 5 orders bellow $\widetilde{V}$ at frequencies $(2k+1)\Omega_m \pm 2\omega$. A modulation nonlinear response is visible at the same level as for $\mathcal{O}(\widetilde{V}^2)$ at frequencies $(2k+1)\Omega_m$.

outputs have been checked with two other simulations tools : `SIMHFI`[1] and `SHDet` [2], yielding a very good agreement between all approaches.

The simulation setup uses experimental measurements from the *Planck/HFI* first 100GHz channel. Electrical and thermal parameters are given in Table 1 and Table 2 respectively which are provided in Appendices (Sect. C).

Timelines of two seconds are produced, and 1 second of data is discarded to allow the steady state to stabilize at numerical precision. The nominal bias voltage is built with a triangle wave of amplitude *vtri* (not peak to peak) plus a square wave of amplitude *vsqu*. Additionally a linear slope of 4.9% of the modulation period is added on the square wave at up and down location to mimic real signal raise and fall time.

Two runs of `SEB` have been defined :

1. the reference for steady state with a constant optical background $P_{opt} = \langle P_{opt} \rangle$;
2. the response to an optical excitation with $P_{opt} = \langle P_{opt} \rangle + \widetilde{P}_{opt}$, where $\widetilde{P}_{opt}$ is a sin wave of amplitude $9,6593.10^{-18}$W (about 3% of CMB dipole from Doppler effect due to solar system motion) and angular frequency $\omega = \Omega_m/18$ which is close to 10Hz;

see Table 3 in Appendix D for amplitudes of $\widetilde{R}$, $\widetilde{T}$ and $\widetilde{V}$ signals relative to their average and steady state values.

5.2 Validation of response linearity

First we check the linearity of the system response to an optical excitation. The value chosen for ω in run 1 and 2 allow for the optical period to cover exactly

[1] `SIMHFI` is a simulation tool developed with the `LabView` software by R. V. Sudiwala at Cardiff University

[2] `SHDet` is a fast simulation tool developed with the `C` programming language by S. R. Hildebrandt at Jet Propulsion Laboratory

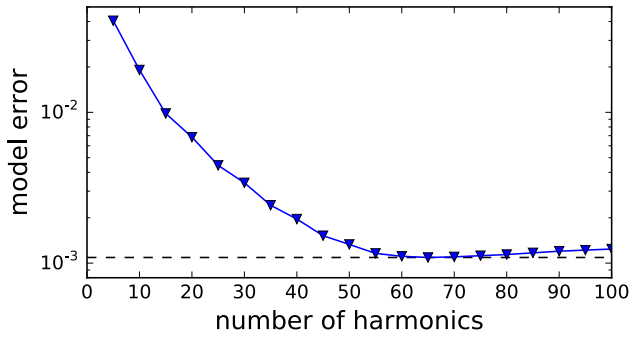


Fig. 5 Convergency of the analytical transfer function time domain residuals as a function of n. Blue triangles are the relative error of the analytical transfer function output compared to the run2 − run1 reference signal. The black dashed line is drawn at the minimal error value which is about 10^{-3}.

18 modulation periods, so the frequencies of interest are not aliased. Simulation results are shown in frequency in Fig. 4. The optical response $\widetilde{V} + \mathcal{O}(\widetilde{V}^2)$ is produced with the output voltage of run2−run1 and is about 5 orders bellow the main modulation harmonics. Optical excitation harmonics appear at $(2k+1)\Omega_m \pm \omega$. Nonlinear response components $\mathcal{O}(\widetilde{V}^2)$ appear at $(2k+1)\Omega_m \pm 2\omega$ and are about 5 orders bellow $\widetilde{V}$.

So the theoretical bolometer model performs very well in the simulation setup with nonlinear behavior at 10^{-5} level of optical excitation response. This result is in good agreement with inflight results from planet crossings estimated at 10^{-4} level, this topic is discussed in section 3.4 of Planck Collaboration (2014).

5.3 Transfer function performances

Next, we check the convergency of the analytical model as a function of the number of modulation harmonics n. When solving the system of equations in matrix form, there is a competition between the number of harmonics increasing precision, and the matrix conditioning increasing the systematic error due to the finite frequency support and the $1/f$ signal shape. As a consequence the error as a function of n should reach a plateau then increase again.

The model error is computed in time domain with the expression $\sigma(\widetilde{V}_{\mathcal{F}} - \widetilde{V}_{\rm simu})/\sigma(\widetilde{V}_{\rm simu})$, where $\widetilde{V}_{\mathcal{F}}$ is the optical response computed from the analytical model and $\widetilde{V}_{\rm simu}$ is computed from simulation data run2 − run1 . The convergency as a function of n is shown in Fig. 5, and the optical response in time domain is shown in Fig.6 for the optimal value of n. The lowest relative error for $\widetilde{\mathbf{v}}^+$ is 10^{-3} obtained with 65 harmonics. This result is in very good agreement with the target error objective and the measured error from linear version of equilibrium equations as shown in Table 3. Depending on the target goal, as few as 15 harmonics are necessary to reach a 1% precision objective.

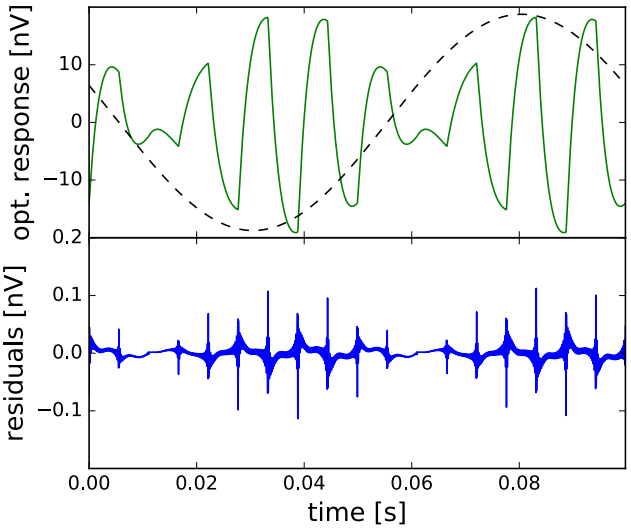


Fig. 6 Time domain response to the 10Hz optical excitation. The upper plot green line is the $\widetilde{V} + \mathcal{O}(\widetilde{V}^2)$ signal obtained from run2-run1, and referenced as $\widetilde{V}_{\rm simu}$ in text. The dashed black sinusoid is the scaled input optical excitation. The bottom plot show the residual error between the simulation and analytical transfer function output and referenced as $\widetilde{V}_{\boldsymbol{\mathcal{F}}} - \widetilde{V}_{\rm simu}$ in text.

5.4 Response for $w \to 0$

For $\omega = 0$, we call *steady state gain* the response to a small change δP_{opt} in the constant optical load $\langle P_{opt} \rangle$ falling on the detector. We will see in this section how it can be inferred from the analytical model matrix expression, and in which frequency domain it can be used for slowly varying signal model.

The *steady state gain* is an observable of particular importance :

- it provides a calibration source given the knowledge of bolometer thermal and electrical parameters;
- it can provide constraints on the bolometer optical response;
- it allows us to build a first order model of electronic response shape for slowly varying signal.

The later has been used for the ADC nonlinearity correction of *Planck/HFI* data (Planck Collaboration et al., 2015).

Starting from the matrix expression Eq. (4.15), $\widetilde{\mathbf{v}}^+ = \boldsymbol{\mathcal{F}}^+(\omega)\widetilde{\mathbf{p}}^+$ and $\widetilde{\mathbf{v}}^- = \boldsymbol{\mathcal{F}}^-(\omega)\widetilde{\mathbf{p}}^-$. We notice that only coefficient of $\widetilde{\mathbf{p}}^+$ and $\widetilde{\mathbf{p}}^-$ at harmonic index zero are non null. Then only one column of the $2n+1$ square matrices $\boldsymbol{\mathcal{F}}^+$ and $\boldsymbol{\mathcal{F}}^-$ at harmonic index 0 are needed to describe the optical response. Expanding the time domain expression of $(\widetilde{\mathbf{v}}^+, \widetilde{\mathbf{v}}^-)$ reads

$$\widetilde{V}(t)\Big|_{\omega} = p \sum_{k=-n}^{n} \boldsymbol{\mathcal{F}}^+_{k,0} e^{j(k\Omega_m+\omega)t} + p^* \sum_{k=-n}^{n} \boldsymbol{\mathcal{F}}^+_{k,0} e^{j(k\Omega_m-\omega)t} \,. \tag{5.1}$$

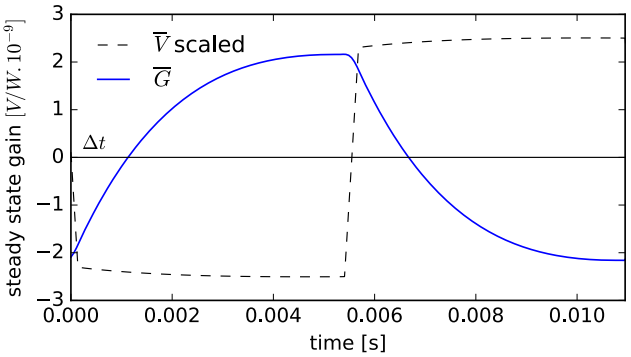


Fig. 7 steady state gain $\overline{G}(t)$ for $\omega = 0$ built using run1 configuration and $n = 35$. One period is shown. Δt is the location when $\overline{G}(t)$ changes sign and is a tuning parameter of the integrated version of the transfer function, see Appendix A.

We have $\boldsymbol{\mathcal{F}}^{+}(0) = \boldsymbol{\mathcal{F}}^{-}(0)$, and also $\boldsymbol{\mathcal{F}}^{+}_{k,0} = \boldsymbol{\mathcal{F}}^{-*}_{-k,0}$ by construction property of optical excitation vectors as seen in Sect. 4.1.2. Then we can write $\widetilde{V}(t)$ as a product

$$\begin{aligned}\widetilde{V}(t)\Big|_{\omega=0} &= (pe^{j\omega t} + p^* e^{-j\omega t}) \sum_{-n}^{n} \left(\boldsymbol{\mathcal{F}}^{+}_{k,0} e^{jk\Omega_m t} + \boldsymbol{\mathcal{F}}^{+*}_{k,0} e^{-jk\Omega_m t} \right) \\ &= \widetilde{P}_{opt}(t)\overline{G}(t)\,, \end{aligned} \tag{5.2}$$

where $\overline{G}(t)$ is the *steady state gain* with the same periodicity as modulation, and $\overline{\mathbf{g}}_k = \boldsymbol{\mathcal{F}}^{+}_{k,0}(0) + \boldsymbol{\mathcal{F}}^{-}_{k,0}(0) = 2\boldsymbol{\mathcal{F}}^{+}_{k,0}(0)$. The *steady state gain* period is shown for the run1 setup in Fig. 7. It appears here that $\overline{G}(t)$ is very different in shape from $\overline{V}(t)$, also the half period signs are opposite because the impedance variations are negative for a positive change in temperature of the bolometer.

We have seen that for $\omega = 0$ we have $\widetilde{V}(t) = \widetilde{P}_{opt}(t)\overline{G}(t)$. This is the expression used for slowly varying signal model. Now we want to characterize its robustness for $\omega \to 0$. The values of $\boldsymbol{\mathcal{F}}^{+}(\omega)$ and $\boldsymbol{\mathcal{F}}^{-}(\omega)$ are different when $\omega > 0$ and lead to an expression who drifts from the product of two time domains signals. As $\overline{G}(t)$ is built using the first column of $\boldsymbol{\mathcal{F}}^{+}(0)$, a simple heuristic is to compare $\overline{\mathbf{g}}$ with the first columns of $\boldsymbol{\mathcal{F}}^{+}(\omega)$ and $\boldsymbol{\mathcal{F}}^{-}(\omega)$. The Parseval's theorem allowing us to switch the comparison to frequency domain.

We estimate the slowly varying signal hypothesis error with

$$\Delta\overline{G}(\omega) = \sqrt{\frac{\sum_{k=-n}^{n} \left| \overline{\mathbf{g}}_k - \boldsymbol{\mathcal{F}}^{+}_{k,0}(\omega) - \boldsymbol{\mathcal{F}}^{-}_{k,0}(\omega) \right|^2}{\sum_{k=-n}^{n} |\overline{\mathbf{g}}_k|^2}}\,. \tag{5.3}$$

The result is shown in Fig. 8. The slowly varying signal hypothesis $\widetilde{V}(t) = \widetilde{P}_{opt}(t)\overline{G}(t)$ exhibits less than 1% estimated error for $\omega < 2Hz$.

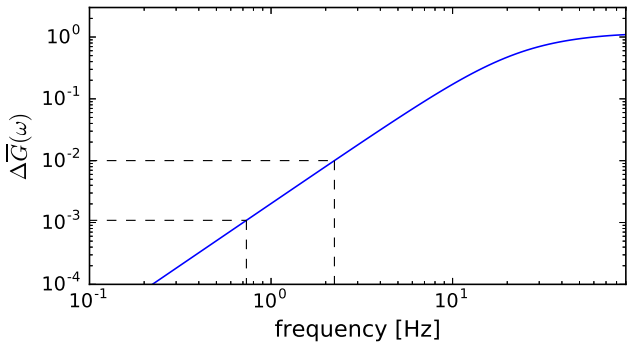


Fig. 8 Slowly varying signal hypothesis error as a function of optical frequency. The dashed lines indicate locations where the estimated error level is 1% and 0.1%. The frequency range is bounded on the right at modulation frequency.

5.5 Integrated version

We have inspected the performances of the analytical model of the transfer function performances at high frequencies. Now we check it's behavior in the *Planck/HFI* data processing common use case. The 80 samples per modulation period are summed over each half period before being send to earth and demodulated. The integrated version $\mathcal{F}_\Sigma(\omega)$, including summation on 40 samples and demodulation, is extensively described in Appendix *A*. The output is shown in Fig. 9 for the modulation frequency range.

For numerical validation, the summation and demodulation operation has been run on the output of `SEB` for run1 and run2, and is shown as red dots on Fig. 9. With an average difference of 1.9×10^{-3} between the simulation and the analytical expression of $\mathcal{F}_\Sigma$ the outputs is perfectly in the range of the linear approximation used.

For comparison purposes, the integrated version $\tau_{b\Sigma}(\omega)$ and $\tau_{e\Sigma}(\omega)$ of two low pass filters is also shown. The generic expression of $\tau_\Sigma(\omega)$ is detailed in appendix A. In the litterature (Chanin and Torre, 1984; Grannan et al., 1997) the bolometer physical time constant is under DC bias $\tau_b = \frac{\langle C_{th}\rangle}{G_d}$ and the effective time constant, taking into account the heat flow from Joule effect, is $\tau_e = \frac{\langle C_{th}\rangle}{G_d - \alpha P_j}$ with the dimensionless temperature coefficient $\alpha = \frac{1}{R}\frac{\partial R}{\partial T}$ and in our case, using Eq. (3.5), $\alpha = \frac{1}{\langle R\rangle a_{1T}}$.

The AC biased version is significantly different from a low pass filter. The $\mathcal{F}_\Sigma$ modulus behaves like τ_b at low frequency and like τ_e near the modulation frequency, and the variations of the angle are also more important than for low pass filters. The observed shape of the analytical model which is a feature specific to AC biasing could explain why the *Planck/HFI* time transfer function is built with as many as 7 low pass filters (Planck Collaboration et al., 2015) for some channels.

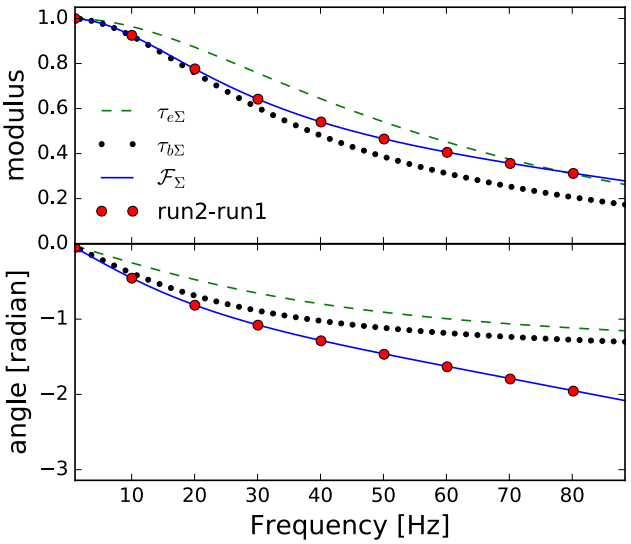


Fig. 9 Integrated version $\mathcal{F}_\Sigma$ of the transfer function in the modulation frequency range. Upper plot is the normalized modulus at 0 frequency, bottom plot is for the angle in radian. Plain blue line is for $\mathcal{F}_\Sigma$, red dots are the values computed using (run1,run2) for validation (average relative error of $1,9.10^{-3}$), the dashed green curve is for $\tau_e = 4.1ms$ and the black dots are for $\tau_b = 6.5ms$

6 Conclusions

We have designed a method to build the analytical expression in frequency domain of the AC biased bolometers response. Starting from the simple bolometer model used in *Planck/HFI* and the bias circuit of the detector, we provided linearized version of the electro thermal equilibrium. The target objective is 0.1% relative precision in presence of a monofrequency optical excitation. The system of equation is solved using linear algebra using a custom and compact Fourier basis designed specifically for this purpose. The method has been applied and tested with the case of the simple bolometer model used in *Planck/HFI*.

The analytical model performances have been characterized using the time domain simulation tool developed at *IRAP*. The reached accuracy is 0.1% relative to the time domain simulation reference, when using $n = 65$ harmonics. The solution is built using square matrices of size $2 * n + 1$ considering a bias signal with n harmonics. An accuracy of 1% can be reached using only $n = 15$ harmonics.

The proposed analytical model is suitable for deconvolution of real data, as only classical matrix inversion tools are needed. We also show how the analytical transfer function can be used to build the *steady state gain* in time domain. This observable provides a first order model of the electronic response, with an accuracy of 1% for optical excitation frequencies less than 2Hz. The *steady state gain* observable can also be used as a calibration source.

Using the presented matrix formalism, the matrix expression of the transfer function allows us to adapt it to different electronic or thermal models. A direct extension is to add several thermal components. We provide an example

with one thermal component. While the *Planck/HFI* low frequency components of time transfer function still need some improvement when writing this article. The generic form of the proposed model is designed as a tool which can be used to get better constraints through the measured optical response and the completely new addition of the *steady state gain*. Finally, the analytical transfer function model allow to build a signal model with high resolution in time domain. This property is needed to build an improved ADC nonlinearity correction for *Planck/HFI*, see Sauvé et al. (2016) (SPIE proceeding in preparation).

Appendices

A Integrated transfer function

This appendix describes the integration/downsampling of the electronic signal as done by the onboard Data Processing Unit of *Planck/HFI*. The corresponding integrated version of the transfer function $\mathcal{F}_\Sigma(\omega)$ will be built from the matrix formalism described in this paper. The filtering applied by the readout amplification chain and the rejection filter is described in the *Planck/HFI* timeresponse paper by Planck Collaboration (2014) and will not be considered here.

The data sent to earth at rate $2f_{mod} = 180.3737$ Hz is the sum of $N = 40$ samples per modulation half period. And the real onboard data acquisition frequency is $2f_{mod} * N = 7214.948$ Hz before the Data Processing Unit make the summation. We will start by applying the summation process to a sinusoidal wave described by

$$s(t) = e^{j\omega t}\,, \tag{A.1}$$

where t has the same origin as the filtered optical signal.

The sample i acquired at time $t_i = \frac{i}{f_{acq}}$ is

$$S_\omega[i] = \sum_{k=0}^{N-1} e^{j\omega\left(\frac{i+k/N}{f_{acq}}+\Delta t\right)} \tag{A.2}$$

$$= e^{j\omega\left(\frac{i}{f_{acq}}+\Delta t\right)} \sum_{k=0}^{N-1} e^{j\omega\frac{k}{Nf_{acq}}}\,, \tag{A.3}$$

where the time delay term Δt is a tunable parameter (referenced as S_{phase} by Planck Collaboration (2014)) allowing to maximize $\mathcal{F}_\Sigma$ gain by adjusting the integration range on the *steady state gain* period. As can be seen on Fig. 7 it has a phase advance of about 1/8 of a modulation period.

Considering the sum of the numbers in a geometric progression

$$\sum_{k=0}^{N-1} x^k = \frac{1-x^N}{1-x}\,, \tag{A.4}$$

then with $x = e^{j\frac{\omega}{Nf_{acq}}}$ the expression of $S_\omega[i]$ can be rewritten

$$S_\omega[i] = e^{j\omega\left(\frac{i}{f_{acq}}+\Delta t\right)}\,\frac{1-e^{j\frac{\omega}{f_{acq}}}}{1-e^{j\frac{\omega}{Nf_{acq}}}}\,. \tag{A.5}$$

We use the Euler relation $\sin(x) = (e^{jx} - e^{-jx})/(2j)$ to simplify the geometric progression sum

$$S_\omega[i] = e^{j\omega\left(\frac{i}{f_{acq}}+\Delta t\right)}\,\frac{e^{j\frac{\omega}{2f_{acq}}}\left(e^{-j\frac{\omega}{2f_{acq}}}-e^{j\frac{\omega}{2f_{acq}}}\right)}{e^{\frac{\omega}{2Nf_{acq}}}\left(e^{-j\frac{\omega}{2Nf_{acq}}}-e^{j\frac{\omega}{2Nf_{acq}}}\right)}$$

$$= e^{j\omega\left(\frac{i}{f_{acq}}+\frac{N-1}{2Nf_{acq}}+\Delta t\right)}\,\frac{\sin\left(\frac{\omega}{2f_{acq}}\right)}{\sin\left(\frac{\omega}{2Nf_{acq}}\right)}\,. \tag{A.6}$$

The term with a sinus denominator is continuous for $\omega \to 0$ and it's limit is N, which can be also found by using the summation method on a constant signal of value 1 for $\omega = 0$. The fraction is also continuous for $\omega = kN\Omega_m, k \in \mathbb{N}$, but these values are out of the frequency range of interest because they are cut off by the onboard anti aliasing electronic rejection filter.

The sample capture starts at time t_i, so the integration transfer function reads

$$\Sigma(\omega) = e^{j\omega\Delta t_I}\,\frac{\sin\left(\frac{\omega}{2f_{acq}}\right)}{\sin\left(\frac{\omega}{2Nf_{acq}}\right)}\,, \tag{A.7}$$

with $\Delta t_I = \frac{N-1}{2Nf_{acq}} + \Delta t$.

The following step involve the folding of bolometer voltage response built using $\boldsymbol{\mathcal{F}}^+$ and $\boldsymbol{\mathcal{F}}^-$ matrices. The modulated linear response appears at odd modulation harmonics. We will consider only these significant harmonics at angular frequency $(2*p+1)\Omega_m$, and we will use the notation $\Omega_k^+ = k\Omega_m + \omega$.

Let consider the complex optical excitation power on the bolometer

$$\widetilde{P}(t) = e^{j\omega t}\,. \tag{A.8}$$

then, using the matrix formalism developped in Sect. 4.1 only one vector is needed to represent it with $\widetilde{\mathbf{p}}^{+} = [1, 0\cdots]$. And the complex output voltage is

$$\widetilde{V}(t) = \sum_{\substack{p\geq -\frac{n+1}{2} \\ k=2p+1}}^{p\leq \frac{n-1}{2}} \boldsymbol{\mathcal{F}}^{+}_{k,0} e^{j\Omega^{+}_{k}t}\,. \tag{A.9}$$

The acquisition of summed samples at time t_i for odd modulation harmonics folds the signals so $ze^{j\Omega^{+}_{k}t_i} = z^{*}e^{j(\Omega_m-\omega)t_i}$. The folded version of output voltage at f_{acq} sampling frequency is

$$\widetilde{V}(t_i) = \sum_{\substack{p\geq -\frac{n+1}{2} \\ k=2p+1}}^{p\leq \frac{n-1}{2}} \boldsymbol{\mathcal{F}}^{+*}_{k,0} \Sigma^{*}(\Omega^{+}_{k}) e^{j(\Omega_m-\omega)t_i}\,. \tag{A.10}$$

The signal is demodulated by applying a $e^{-j\Omega_m t}$ factor and taking the conjugate of the result to get a positive frequency, so we have

$$\widetilde{V}_d(t_i) = \sum_{\substack{p\geq -\frac{n+1}{2} \\ k=2p+1}}^{p\leq \frac{n-1}{2}} \boldsymbol{\mathcal{F}}^{+}_{k,0} \Sigma(\Omega^{+}_{k}) e^{j\omega t_i}\,. \tag{A.11}$$

The signal described with this expression integrates N samples from time t_i and has a time offset of Δt_I appearing in Eq. (A.7) compared to input signal. The time offset can be corrected to obtain the final causal version of the integrated transfer function

$$\mathcal{F}_{\Sigma}(\omega) = \sum_{\substack{p\geq -\frac{n+1}{2} \\ k=2p+1}}^{p\leq \frac{n-1}{2}} \boldsymbol{\mathcal{F}}^{+}_{k,0} \Sigma(\Omega^{+}_{k}) e^{-j\omega\Delta t_I}\,. \tag{A.12}$$

The bolometer under DC current bias behaves as an order 1 low pass filter which can be written

$$T(w) = \frac{1}{1+j\omega\tau}\,, \tag{A.13}$$

where τ is the time constant of the filter. For comparison purposes we can apply the sampling summation process on $T(w)$. There is no demodulation in this case and we have $\tau_{\Sigma}(\omega) = \Sigma(\omega)T(\omega)e^{-j\omega\Delta t_I}$. As there is no modulated harmonics the time offset Δt_I cancels with the one in $\Sigma(\omega)$ so we have

$$\tau_{\Sigma}(\omega) = T(\omega)\frac{\sin\left(\frac{\omega}{2f_{acq}}\right)}{\sin\left(\frac{\omega}{2Nf_{acq}}\right)}\,. \tag{A.14}$$

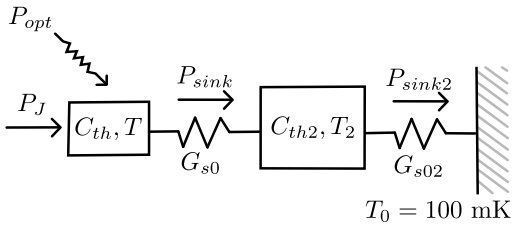


Fig. 10 Extended thermal model with an intermediate heat capacity between the bolometer and the heat sink. The bolometer itself is represented by the left component with thermal capacity C_{th} and temperature T. P_J is the Joule effect thermal power produced by bias current. P_{opt} is the total incoming radiative power falling on the detector. P_{sink} is the thermal power flowing out of the bolometer to the second component via a thermal link with static thermal conductance G_{s0}. C_{th2} and T_2 are the thermal capacity and temperature of the second thermal component. P_{sink2} is the thermal power flowing from the second component via a thermal link with static thermal conductance G_{s02}. T_0 is the heat sink temperature with supposed infinite thermal capacity.

B Adding a thermal component

We have seen how to compute the bolometer response with only one thermal component. As it has been seen in *Planck/HFI* with the low frequency excess response (Planck Collaboration et al., 2015), the detector behaves as if there are several thermal component on the thermal path between the bolometer and heat sink, which alter significantly the filtering of optical signal. To represent a more realistic system, we will detail an example of adding one thermal component to the expression of $\boldsymbol{\mathcal{F}}^+$ and $\boldsymbol{\mathcal{F}}^-$.

B.1 Extended thermal model

We will use a simple thermal architecture as presented in Fig. 10, by adding a single component with heat capacity C_{th2} and tempearture T_2 between the bolometer and the heat sink. The new component is connected to the heat sink via a link with thermal conductance G_{s02}.

To keep things simple we make the hypothesis that all physical characteristics are well known, and that for steady state $\langle R\rangle$, $\langle T\rangle$ and $\langle T_2\rangle$ are also known. In practice, a simple way to obtain their values is with a fit. If we notice that in steady state $\langle P_{sink}\rangle - \langle P_{sink2}\rangle = 0$ and that $\langle P_J\rangle + \langle P_{opt}\rangle - \langle P_{sink}\rangle = 0$. Setting as constraint $T > T_2 > T_0$ we can use the heat flow equilibrium to fit $\langle T\rangle$ anf $\langle T_2\rangle$. A commonly used tool as `mpfit`[3] converge in 5 iterations with a 10^{-5} relative difference for the stoppping criteria. This is a very quick operation considering an $\mathcal{O}$ (n) complexity coming from the $\langle P_J\rangle$ term which needs the computation of voltage harmonics at first order.

Using the same notations as we did in Sect. 3.3.3, the heat flow at the output of the bolometer can be written at first order

$$P_{sink} = \langle P_{sink}\rangle + G_{d12}\overline{T}_1 - G_{d21}\overline{T}_2 + \mathcal{O}(\overline{T}^2) + \mathcal{O}(\overline{T}_2^2)\,, \tag{B.1}$$

[3] `mpfit` is a tool for non linear least squares fitting developed by Craig Markwardt and based on the MINPACK-1 software. `https://www.physics.wisc.edu/~craigm/idl/cmpfit.html`

with

$$\begin{aligned} G_{d12} &= \left.\frac{\partial P_{sink}}{\partial T}\right|_{T=\langle T\rangle} &&= \frac{G_{s01}}{T_{ref}^{\beta}}\langle T\rangle^{\beta} \\ G_{d21} &= -\left.\frac{\partial P_{sink}}{\partial T_2}\right|_{T_2=\langle T_2\rangle} &&= \frac{G_{s01}}{T_{ref}^{\beta}}\langle T_2\rangle^{\beta}\,. \end{aligned} \tag{B.2}$$

The second thermal component on the thermal path can be characterized at first order by

$$\begin{aligned} C_{th2} &= C_{02}T_2^{\gamma_2} = \langle C_{th2}\rangle + \mathcal{O}(\overline{T}_2) \\ P_{sink2} &= \langle P_{sink}\rangle + G_{d2}\overline{T}_2 + \mathcal{O}(\overline{T}_2^2)\,, \end{aligned} \tag{B.3}$$

with

$$G_{d2} = \frac{\partial P_{sink2}}{\partial T_2} = \frac{G_{s02}\langle T_2\rangle^{\beta_2}}{T_{ref}^{\beta_2}}\,, \tag{B.4}$$

where $\langle C_{th2}\rangle$,G_{s02} and β_2 are free parameters. We neglect $\langle C_{th2}\rangle$ variations for the following developments.

And the thermal equilibrium for the second component reads at first order

$$\begin{aligned} C_{th2}\dot{T_2} &= P_{sink} - P_{sink2} \\ \langle C_{th2}\rangle\dot{\overline{T}}_2 &= G_{d12}\overline{T} - (G_{d21}+G_{d2})\overline{T}_2 + \mathcal{O}(\overline{T})\,. \end{aligned} \tag{B.5}$$

B.2 Extended steady state

The first iteration for $\overline{\mathbf{v}}$ is the same as in Sect. 4.2 because we only need to know $\langle R\rangle$.

With the second component, its steady state temperature vector $\overline{\mathbf{t}}_2$ is needed for the thermal equilibrium of the bolometer. $\overline{\mathbf{t}}_2$ can be expressed from $\overline{\mathbf{r}}$ using the thermal equilibrium Eq. (B.5), then it comes

$$\begin{aligned} \overline{\mathbf{t}}_2 &= \mathbf{X}_2\overline{\mathbf{r}} \\ \mathbf{X}_2 &= \left[\langle C_{th2}\rangle\mathbf{D} + (G_{d21}+G_{d2})\mathbf{I_d}\right]^{-1} G_{d12}a_{1T}\,, \end{aligned} \tag{B.6}$$

where $\overline{\mathbf{t}}_2$ is the intermediate steady state vector for $\overline{T}_2$.

And finally the expression of $\overline{\mathbf{r}}$ can be written from the single thermal component version Eq. (4.11) by updating the P_{sink} term

$$\overline{\mathbf{r}} = \left[\langle C_{th}\rangle a_{1T}\mathbf{D} + \frac{\boldsymbol{\mathcal{C}}(\boldsymbol{\mathcal{C}}(\overline{\mathbf{v}}).dot(\overline{\mathbf{v}}))}{\langle R\rangle^2} + G_{d12}a_{1T}\mathbf{I_d} - G_{d21}\mathbf{X}_2\right]^{-1}\left[\frac{\boldsymbol{\mathcal{C}}(\overline{\mathbf{v}})\overline{\mathbf{v}}}{\langle R\rangle} - \langle P_{sink}\rangle + \langle P_{opt}\rangle\right], \tag{B.7}$$

Once $\overline{\mathbf{r}}$ is obtained, a second iteration can be done as in Sect. 4.2 to get $\overline{\mathbf{v}}$ with a better precision.

B.3 Extended transfer function

To compute the optical excitation response with a second thermal component, we will make some (optional) simplifications on the second component. The main hypothesis is that the additional thermal component add a small thermal feedback from $\widetilde{T}_2$ to the heat quantity variations on the bolometer itself. Then we will use the first order expression for heat quantity variation, and with the expression of $\widetilde{T}$ from Eq. 3.7 the thermal equilibrium will read

$$\langle C_{th2}\rangle \dot{\widetilde{T}}_2 = G_{d12}\left(a_{1T}\widetilde{R} + 2a_{2T}\overline{R}\widetilde{R}\right) - \left(G_{d21}+G_{d2}\right)\widetilde{T}_2 + \mathcal{O}(\widetilde{R}^2)\,. \tag{B.8}$$

From which we can write the vector version of $\widetilde{T}_2$ as

$$\begin{cases} \widetilde{\mathbf{t}}_2^{+} = \mathbf{Y}_2^{+}\widetilde{\mathbf{r}}^{+} \\ \widetilde{\mathbf{t}}_2^{-} = \mathbf{Y}_2^{-}\widetilde{\mathbf{r}}^{-}\,, \end{cases} \tag{B.9}$$

with

$$\begin{cases} \mathbf{Y}_2^{+} = \left[\langle C_{th2}\rangle \mathbf{D}^{+} + (G_{d21}+G_{d2})\mathbf{I_d}\right]^{-1} G_{d12}\left[a_{1T}\mathbf{I_d} + 2a_{2T}\boldsymbol{\mathcal{C}}(\overline{\mathbf{r}})\right] \\ \mathbf{Y}_2^{-} = \left[\langle C_{th2}\rangle \mathbf{D}^{-} + (G_{d21}+G_{d2})\mathbf{I_d}\right]^{-1} G_{d12}\left[a_{1T}\mathbf{I_d} + 2a_{2T}\boldsymbol{\mathcal{C}}(\overline{\mathbf{r}})\right]\,. \end{cases} \tag{B.10}$$

Finally the new transfer function matrices expressions, $\boldsymbol{\mathcal{F}}_2^{+}$ and $\boldsymbol{\mathcal{F}}_2^{-}$, can be written from Eq. (4.15) and Eq. (4.16) by replacing the P_{sink} component $\mathbf{S}$ with a new version taking into account the new thermal path

$$\begin{cases} \boldsymbol{\mathcal{F}}_2^{+} = \left[\left(\mathbf{C_1} + \mathbf{C_2}\mathbf{D}^{+} - \mathbf{J_2} + \mathbf{S_2^{+}}\right)\mathbf{E}^{+} - \mathbf{J_1}\right]^{-1} \\ \boldsymbol{\mathcal{F}}_2^{-} = \left[\left(\mathbf{C_1} + \mathbf{C_2}\mathbf{D}^{-} - \mathbf{J_2} + \mathbf{S_2^{-}}\right)\mathbf{E}^{-} - \mathbf{J_1}\right]^{-1}\,, \end{cases} \tag{B.11}$$

with

$$\begin{cases} \mathbf{S_2^{+}} = G_{d12}\boldsymbol{\mathcal{C}}\left(a_{1T} + 2a_{2T}\overline{\mathbf{r}}\right) - G_{d21}\mathbf{Y}_2^{+} \\ \mathbf{S_2^{-}} = G_{d12}\boldsymbol{\mathcal{C}}\left(a_{1T} + 2a_{2T}\overline{\mathbf{r}}\right) - G_{d21}\mathbf{Y}_2^{-}\,. \end{cases} \tag{B.12}$$

C Simulation setup

Table 1 Electrical parameters of the first 100GHz *Planck/HFI* channel

Name	Value	Unit	Description
f_{mod}	90.18685	Hz	modulation frequency
$vtri$	0.62399	V	triangle wave amplitude
$vsqu$	0.17796	V	square wave amplitude
C_0	4.886×10^{-12}	F	bias capacitance 0
C_1	4.711×10^{-12}	F	bias capacitance 1
C_s	148.8×10^{-12}	F	stray capacitance
$\langle R\rangle$	11.0×10^{6}	Ω	bolometer average impedance

Table 2 Thermal Parameters of the first 100GHz *Planck/HFI* channel

Name	Value	Unit	Description
C_0	22.47 pF	J/K	heat capacity coefficient
γ	1.91	Adimentional	heat capacity temperature exponent
R_G	57.46	Ω	bolometer impedance a reference temperature
T_G	11.18	K	reference temperature for bolometer impedance
β	1.3	Adimentional	thermal conductance temperature exponent
G_{s0}	4.533×10^{-11}	W/K	static thermal conductance
T_{ref}	0.1	K	thermal conductance reference temperature
T_0	0.1	K	heat sink temperature
$\langle T \rangle$		K	bolometer average temperature
$\langle P_{opt} \rangle$	4.5484×10^{-13}	W	optical power average

D Numerical results

Table 3 Relative signal amplitude for steady state signals and optical excitation signals. Values are obtained from run2-run1 simulation signals.

X	$\sigma(\overline{X})/\langle X \rangle$	$\sigma(\widetilde{X})/\sigma(\overline{X})$
V		3.5e-06
T	5.5e-04	1.5e-03
R	3.3e-03	1.5e-03

References

W. A. Holmes, J. J. Bock, B. P. Crill, T. C. Koch, W. C. Jones, A. E. Lange, and C. G. Paine. Initial test results on bolometers for the Planck high frequency instrument. Appl. Opt., 47:5996–6008, November 2008. doi: 10.1364/AO.47.005996.

A. Catalano, A. Coulais, and J.-M. Lamarre. Analytical approach to optimizing alternating current biasing of bolometers. Appl. Opt., 49:5938, November 2010. doi: 10.1364/AO.49.005938.

J. A. Tauber, N. Mandolesi, J.-L. Puget, T. Banos, M. Bersanelli, F. R. Bouchet, R. C. Butler, J. Charra, G. Crone, J. Dodsworth, and et al. Planck pre-launch status: The Planck mission. A&A, 520:A1, September 2010. doi: 10.1051/0004-6361/200912983.

Planck Collaboration, R. Adam, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, and et al. Planck 2015 results. VII. HFI TOI and beam processing. *ArXiv e-prints*, February 2015.

A. Catalano, P. Ade, Y. Atik, et al. Impact of particles on the Planck HFI detectors: Ground-based measurements and physical interpretation. A&A, 569:A88, September 2014. doi: 10.1051/0004-6361/201423868.

Planck Collaboration. Planck 2013 results. VII. HFI time response and beams. A&A, 571:A7, November 2014. doi: 10.1051/0004-6361/201321535.

J.-M. Lamarre, J.-L. Puget, P. A. R. Ade, et al. Planck pre-launch status: The HFI instrument, from specification to actual performance. A&A, 520:A9, September 2010. doi: 10.1051/0004-6361/200912975.

Ludovic Montier. *Planck : De l'étalonnage de l'instrument à l'étude des poussières galactiques et intergalactiques*. Theses, Université Toulouse III, Paul Sabatier - Toulouse, September 2005.

Michel Piat. *Contributions à la définition des besoins scientifiques et des solutions instrumentales du projet Planck-HFI*. Theses, Université Paris Sud - Paris XI, October 2000. URL `https://tel.archives-ouvertes.fr/tel-00004038`.

Michel Piat, Jean-Pierre Torre, Eric Bréelle, Alain Coulais, Adam Woodcraft, Warren Holmes, and Rashmi Sudiwala. Modeling of planck-high frequency instrument bolometers using non-linear effects in the thermometers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559(2):588 – 590, 2006. ISSN 0168-9002. doi: http://dx.doi.org/10.1016/j.nima.2005.12.076. URL `http://www.sciencedirect.com/science/article/pii/S0168900205025027`. Proceedings of the 11th International Workshop on Low Temperature DetectorsLTD-1111th International Workshop on Low Temperature Detectors.

John C. Mather. Bolometers: ultimate sensitivity, optimization, and amplifier coupling. *Appl. Opt.*, 23(4):584–588, Feb 1984. doi: 10.1364/AO.23.000584. URL `http://ao.osa.org/abstract.cfm?URI=ao-23-4-584`.

John C. Mather. Bolometer noise: nonequilibrium theory. *Appl. Opt.*, 21(6): 1125–1129, Mar 1982. doi: 10.1364/AO.21.001125. URL `http://ao.osa.org/abstract.cfm?URI=ao-21-6-1125`.

G. Chanin and J. P. Torre. Electrothermal model for ideal semiconductor bolometers. *J. Opt. Soc. Am. A*, 1(4):412–419, Apr 1984. doi: 10.1364/JOSAA.1.000412. URL `http://josaa.osa.org/abstract.cfm?URI=josaa-1-4-412`.

S. M. Grannan, P. L. Richards, and M. K. Hase. Numerical optimization of bolometric infrared detectors including optical loading, amplifier noise, and electrical nonlinearities. *International Journal of Infrared and Millimeter Waves*, 18(2):319–340, 1997. ISSN 1572-9559. doi: 10.1007/BF02677923. URL `http://dx.doi.org/10.1007/BF02677923`.

A. Sauvé, Couchot F., Patanchon G., and Montier L. Inflight characterization and correction of the planck hfi analog to digital converter nonlinearity. In *SPIE Astronomical Telescopes + Instrumentation (AS16)*, volume 9914-119 of *Proc. SPIE*, 2016.

Appendix B

SPIE *Astronomical Telescope + Instrumentation* proceeding

Inflight characterization and correction of Planck/HFI Analog to Digital Converter nonlinearity

Sauvé, A.[a], Couchot, F.[b], Patanchon, G.[c], and Montier, L.[a]

[a]IRAP, Institut de Recherche en Astrophysique et Planétologie, CNRS, 9 Avenue du Colonel Roche, BP 44346, 31028 Toulouse, France
[b]LAL, Laboratoire de l'Accélerateur Linéaire, CNRS & Université Paris 11, Bâtiment 200, 91898 Orsay Cedex, France
[c]APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

ABSTRACT

The *Planck* Satellite launched in 2009 was targeted to observe the anisotropies of the Cosmic Microwave Background (CMB) to an unprecedented sensitivity. While the Analog to Digital Converter of the HFI (High Frequency Instrument) readout electronics had not been properly characterized on ground, it has been shown to add a systematic nonlinearity effect up to 2% of the cosmological signal. This was a limiting factor for CMB science at large angular scale. We will present the inflight analysis and method used to characterize and correct this effect down to 0.05% level. We also discuss how to avoid this kind of complex issue for future missions.

Keywords: ADC, Nonlinearity, bolometer, infrared, Planck, HFI, CMB

1. INTRODUCTION

The *Planck* satellite, launched in 2009, was designed to reach an unprecedented sensitivity to the signal of the cosmic microwave background (CMB) in the submillimetre and millimetre wavelengths over the whole sky. These performances have been made possible thanks to an extremely low level of noise and a perfect control of the systematic effects. In the case of the *Planck* High Frequency Instrument (HFI), we had to face an unexpected systematic to reach this goal : a component of the readout electronic chain, the analogic to digital converter (ADC), was not properly characterized on ground before launch, and exhibited an unexpected nonlinear behaviour, responsible of a variable gain of the instrument over the mission. This effect was identified relatively late, after more than one year of mission.[1]

The quality of the *Planck* data in intensity and polarisation has been finally reached thanks to a great effort to identify the origin of this variable gain, and to correct the data from the ADC nonlinearity, based on an inflight characterisation of the ADC components for the 54 detector channels. While a first version of the ADC correction has been included in the *Planck* 2013 data release, a new version required a lot of work to obtain more than a decade in noise reduction and gain stability. We present in this paper how the *Planck* team succeeded in characterizing the ADC nonlinearity in flight and applied the appropriate correction to the *Planck* data.

The paper is organized as follows: we first present the data acquisition chain of the *Planck/HFI* instrument[2] in Sect. 2. The method of inflight characterization of the ADC nonlinearity is described in Sect, 3, while the building of the correction is detailed in Sect. 4. We quantify in Sect. 5 the efficiency of this correction and its impact on *Planck* science. We finally provide suggestions to avoid this kind of issue for future missions in Sect. 6, before concluding in Sect. 7.

2. THE DATA ACQUISITION CHAIN OF HFI

In this section, we describe the main elements of the data acquisition chain of HFI.

Further author information: (Send correspondence to A. Sauveé and G. Patanchon)
A. Sauvé: E-mail: asauve.irap@gmail.com, Telephone: +33 (0) 6 64 82 77 21
G. Patanchon: E-main: patanchon@apc.univ-paris7.fr, Telephone: +33 (0) 1 57 27 60 87

Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VIII,
edited by Wayne S. Holland, Jonas Zmuidzinas, Proc. of SPIE Vol. 9914, 99143E

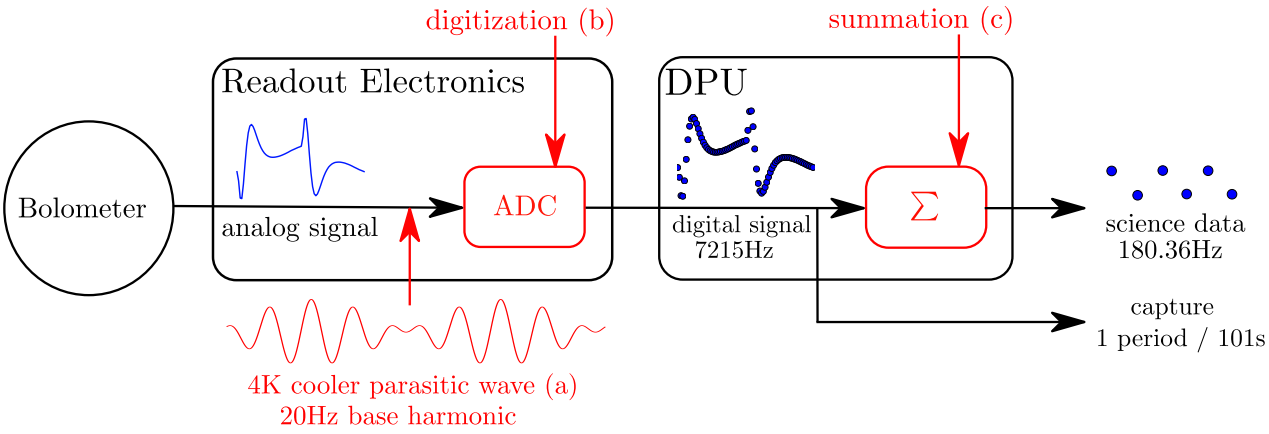


Figure 1. Simplified diagram of inflight signal processing. Parasitic wave (a) from 4K stage mechanical cooler leaks on the signal through the readout electronics chain, then the digitization by the ADC (b) introduce the nonlinear effect, and the signal summation (b) in the DPU cause loss of information.

2.1 Processing of detector signal

The HFI detectors are biased with AC current at modulation frequency $f_{mod} = 90.18$Hz. The circuit is designed to produce a current flow as square as possible through the detector to mimic a DC current bias in order to optimize the linearity of the detector response to sky signal. The bias current intensity can be set on a per channel basis with the `ibias`* parameter. The square signal at the output of the detector presents a high dynamic, which has to be reduced to allow the capture of the bright signal from galactic plane and planets (used for calibration[1]). This high dynamic signal is balanced by adding a square signal in phase opposition, yielding the analog signal (see Fig.1), which is designed to be as flat as possible.

A parasitic wave (see (a) on Fig.1) from the 4K cryogenic stage mechanical compressor adds 20Hz harmonics to the analog signal, before digitization. This parasitic signal is seen on science data and known as 4K *lines*.

Also to be exhaustive, it is worth to mention here that electrical crosstalk from other channels are also sources of parasitic signal on a given detector channel. While this electrical crosstalk is not significant for science data, its amplitude is of order of the balanced signal and as a consequence renders an analytical modeling of the balanced signal shape nearly impossible.

2.2 Digitization by the ADC : the *fast samples*

The analog signal, mixed with the parasitic 4K *lines* is sampled at a rate of $80 f_{mod} = 7215$Hz, yielding the digital signal (see (b) on Fig.1) counting 80 samples per modulation period. This signal is also known as *fast samples* and cannot be sent on earth, because of the limited bandwidth for all the 72 high data rate channels of the HFI instrument. However every 101 seconds, and for each detector, a period of 80 *fast samples* is captured before summation by the Data Processing Unit (DPU) for further data analysis.

It is important to note here that the 4K mechanical compressor runs in synchronization with the modulation signal, so that the acquisition of the *fast samples* of a given detector channel falls always on the same phase of the 4K *lines* signal during the whole mission. Furthermore, since the 4K *lines* base harmonic frequency is $f_{4K} = 2f_{mod}/9$, only 2/9 of the 4K *lines* period is captureed in the *fast samples* window for a given detector. So that we miss 7/9 of the 4K *lines* period, which has to be approximated. As the full knowledge of the 4K *lines* period is mandatory for science data correction, two estimation methods have been developed (see Sec. 4).

The digitization is the step where the ADC nonlinearity effect is introduced, while the balancing sets the analog signal location exactly at the center of the ADC scale, around code 32768 on a digitization scale counting 65536 codes. As a consequence, the nonlinearity of the ADC has a maximum impact.

* channels parameters are digital values that can be set in the range $[\![0, 4095]\!]$

2.3 Signal summation

The last step involved in the signal processing is the half period summation (see (c) on Fig.1). Each science data sample is defined as the sum of the 40 *fast samples* of its corresponding half period, yielding the science data signal sampled at $2f_{mod}$ frequency and modulated at f_{mod} frequency. The science data signal has the interesting property that one sample over two is captured in very different regions of the ADC, and then comparing samples from both parities can be used to build an observable on ADC nonlinearity residuals (see Sec. 5).

However the signal summation leads to a critical loss of information:

- Loss of ADC code value: the summation mixes the nonlinear codes[†] at the output of the ADC;
- Loss of 4K lines real signal: the spectral distribution of 4K harmonics extends at much higher frequencies than f_{mod}, and the summed modulated signal has only 4 visible harmonics at {20Hz, 40Hz, 60Hz, 80Hz} resulting from the folding of many higher frequencies harmonics. As a consequence it is not possible to estimate the real 4K signal directly from science data.

It is clear that the knowledge of the ADC nonlinearity alone is not sufficient to correct for its effect on the science data. A model describing the bolometer response to sky signal along with 4K *lines* is necessary to reconstruct the signal as seen by the ADC.

3. INFLIGHT CHARACTERIZATION OF ADC NONLINEARITY

As the problematic of the ADC has been discovered after the satellite launch, a specific method had to be designed to characterize the chip nonlinearity inflight for each channel. Usually the characterization is done using well known linear ramp input, but in the HFI case we had to perform the reconstruction of the ADC transfer function $\mathcal{A}$ with small overlapping gaussians.

In this section we describe the sequence run to acquire the *fast samples* for ADC nonlinearity characterization, and the model used to describe the digitization of the analog signal by the ADC. After that, we describe how the ADC nonlinearity is estimated with a maximum likelihood method.

3.1 Formalism

The ADC converts an input analog signal $s(t)$ in volt, over a fixed range $[0, V_{\mathrm{REF}}]$, into an integer value κ (a code). Its transfer function can be written

$$\begin{cases} \mathcal{A}(s(t)) = \kappa \\ s(t) \in [v(\kappa), v(\kappa+1)[\,, \end{cases} \tag{1}$$

with output values changing only for discrete locations $v(\kappa)$ called the Code Transition Points (CTP). The ADC effect can also be described as a quantization of the input voltage scale into 2^n quantization steps $Q(\kappa) = v(\kappa+1) - v(\kappa)$, with $n = 16$ in the case of HFI. The unit for the code value is the Least Significant Bit (LSB) considering an ideal quantization step $Q = 1\mathrm{LSB}$.

The nonlinearity of the chip can be characterized with the transfer function $\mathcal{A}$, and can be described only with the locations of the CTP. Another common representation is the Differential NonLinearity (DNL) which is defined as the deviation of each quantization step size from the average value $Q = V_{\mathrm{REF}}/2^n$, and it reads $\mathrm{DNL}(\kappa) = Q(\kappa)/Q - 1$. For the characterization of large scale nonlinearity it is preferable to use the Integral NonLinearity (INL), which is the drift from a linear behavior and is calculated with $\mathrm{INL}(\kappa) = \sum_{j=0}^{\kappa} \mathrm{DNL}(j)$.

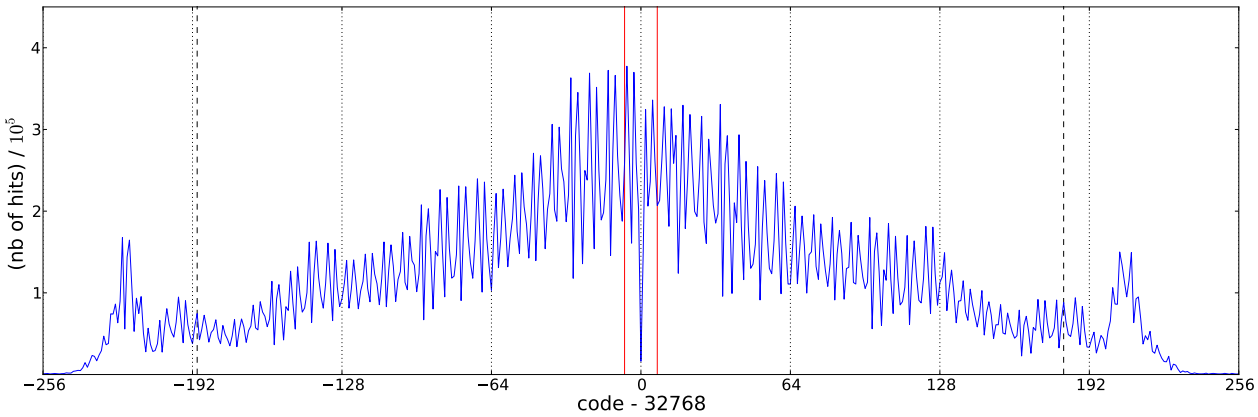


Figure 2. Histogram of the ADC code occurrences of the *fast samples*, computed over the full HFI science mission on channel `00_100-1a`. Black vertical dashed lines are the 5% and 95% percentiles boundaries. Red vertical lines represent the maximum solar dipole amplitude, about 14 LSB peak to peak. The high frequency spikes are caused by the DNL, the central main defect, at code 32768, is clearly visible as a poorly populated bin. The histogram extension is much wider than the typical solar dipole because of thermal drift during the mission. It is mostly symmetric, because the modulated signal extension is symmetric around midscale.

3.2 ADC nonlinearity evidence

The ADC nonlinearity effect on HFI data has been detected in mid 2010, but was first interpreted as a variable gain over the mission and corrected with a gain factor per stable pointing period. It was only in December 2011 that the ADC was identified as the dominant source for the apparent variable gain, as illustrated in Fig. 2. The solar dipole maximum peak to peak amplitude is shown on the figure, as it is the main source of cosmological signal dynamic for CMB channels (photometric calibration relies on 10% of solar dipole amplitude). The ADC central defect (visible as a nearly empty bin) is about 7% in size compared to the solar dipole peak to peak amplitude, so this figure shows why the ADC nonlinearity had such an important impact. Hopefully, the *fast samples* summation smooth a bit the central defect impact, by averaging codes digitized from 40 slightly different analog signal locations.

The thermal drift of the focal plane over the mission is the main cause of the visible spread on the histogram, and this effect vary very slowly. So if the ADC was ideal, each code would have the same quantization step size Q and the histogram should be very smooth. As the thermal drift is by nature poorly characterized, it cannot be used as a source of calibration for the DNL. This is why after HFI End Of Life (EOL) for scientific mission, a specific sequence had to be designed to characterize the ADC nonlinearity on the 54 detectors channels in warm state after the HFI instrument End Of Life (EOL).

3.3 EOL sequence

The EOL sequence consists in sending to the satellite a batch of `ibias` commands to change the detector bias current value, so the analog signal can cover the ADC scale used by the HFI mission (± 512 codes around midscale). During the EOL sequence, 20 `ibias` values are used, starting from zero, and the electronic readout gain is set to its maximum value so that the Root Mean Square (RMS) of the noise can reach the maximum possible value of 20 LSB. This sequence has been running for ten monthes (see Fig. 3) and *fast samples* periods have been captured at the same rate as during the HFI mission.

Two strong hypothesis must be made so the *fast samples* acquired in EOL sequence can be used as a source of calibration for the ADC nonlinearity:

[†] A code is the integer value in the range $[\![0, 65535]\!]$ resulting from the conversion of the analog input signal by the ADC.

- STABILITY OF THE DNL IN TIME: The data acquisition is performed after HFI EOL, but the REU hosting the chip is in the 300K stage of the satellite. As the REU temperature is stable enough over the mission (mesured variations are of order 1K), we consider that the ADC nonlinear features are stable all over the mission.

- GAUSSIAN NOISE: The noise is expected to be gaussian and stationary, in other words the 1-σ value of the noise is expected to be constant at least in each stable `ibias` acquisition period. This includes the fact that additional bias from the 4 Kelvin mechanical compressor is non-significant, and that electronic signal shape is constant to a very good precision. Also the warm detectors are expected to have non significant sky power response, because at temperature greater than 1K they a very low impedance (about 1kΩ compared to the average 12MΩ at the nominal mission temperature of 100mK)

Some additional constraints can be added from, the on-ground acquisitions campaigns, run in early 2012. The analysis of two fully characterized ADCs shows that their DNL has a 64 code periodicity. This periodicity has an exception for the first 64 codes located just after the midscale code, so they have to be estimated independently.

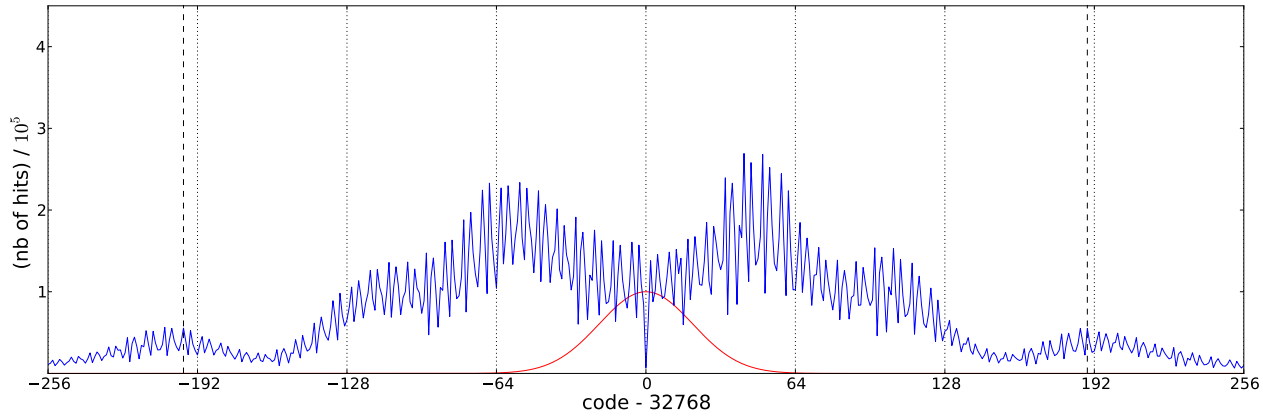


Figure 3. Histogram of the ADC code occurrences of the *fast samples*, computed over the 10 monthes of the EOL sequence duration. Black vertical dashed lines are the 5% and 95% percentiles boundaries. Red line shows the typical scale of the gaussian noise distribution. The histogram scale is the same as for Fig.2.

3.4 Signal model for EOL sequence

To build a model of input gaussian noise processed by the ADC, let consider data sample $d_i(t)$ acquired at time t. The index i maps for a fixed `ibias` value and a given phase index in the modulation period of one target detector channel. These samples are the result of the digitization of the analog input signal by the ADC such as $d_i(t) = \mathcal{A}(s_i(t))$, and are captured at times separated by 101 seconds by the DPU. The corresponding dataset is noted $\{d_i(t)\}$. Using the hypothesis that the input signal noise is gaussian, $s_i(t)$ is described by the normal probability density function reading

$$p_i\left(s_i(t)|\{\mu_i,\sigma_i\}\right) = \frac{1}{\sqrt{2\pi}\sigma_i}\, e^{-\frac{(s_i(t)-\mu_i)^2}{2\sigma_i^2}}\,, \tag{2}$$

where μ_i and σ_i are the average and standard deviation in volt of $s_i(t)$.

Using these notations, the probability of the measured dataset $\{d_i(t)\}$ is

$$q_i(\{d_i(t)\} = \mathrm{K}|\theta) \int_{v(\mathrm{K})}^{v(\mathrm{K}+1)} p_i(x|\{\mu_i,\sigma_i\})dx\,, \tag{3}$$

with $\theta = \{\mu_i, \sigma_i, v(\mathrm{K})\}$ being the unknown parameter set of size 65537.

Using the mean and standard deviation of the dataset as measured by the ADC, an initial (biased) estimate (μ_i^0, σ_i^0) can be calculated. As these parameters degenerate with the code transition points $\{v(\mathrm{K})\}$, they must be estimated together. The standard deviation of the noise σ_i is degenerated with local variation of the ADC transfer function gain, and the mean of the input signal μ_i is degenerated with a local distortion of the ADC transfer function similar to the gauss function derivative. Hopefully these degeneracies can be partially broken with

- The overlapping of the $\{d_i(t)\}$ datasets;
- The 64 code periodicity constraint (excepted for the first 64 codes located just after midscale).

A global degeneracy between the values of σ_i is fixed by scaling the average value of $Q(\mathrm{K})$ to 1LSB.

3.5 Likelihood function

Let consider the set $\{h_i(\mathrm{K})\}$, which is the histogram of dataset $\{d_i(t)\}$ containing N_i elements. We assume that data samples are independent from each others, so there is no loss of information when using the histogram instead of the the dataset.

To build an estimator for the unknown parameters θ, we can derive a likelihood function from the histograms. For this purpose the histogram expected value can be put in relation with its previously defined probability as

$$\mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta] = \mathrm{N_i q_i}(\mathrm{K}|\theta)\,. \tag{4}$$

The likelihood function of an histogram is described by a multinomial distribution that we approximate with a gaussian function, that is justified by the fact that the dataset used contains a large statistic. Given this approximation, the global log likelihood function for all the histograms can be written

$$\Phi(\theta) = \sum_i \sum_{\mathrm{K}} \frac{\left(h_i(\mathrm{K}) - \mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta]\right)^2}{2\mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta]} + \log \mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta] \tag{5}$$

A second approximation will be needed for the expected value of histograms hit count. To do this, we note that the value of σ_i is about 20 LSB. So the integral of the very smooth probability density function on the short range $[v(\mathrm{K}), v(\mathrm{K}+1)]$ (about 1 LSB) can be approximated with the trapezoidal rule. Then the first order development of Eq. 4 can be written

$$\mathrm{E}[h_i(\mathrm{K})|\theta] \simeq N_i Q(\mathrm{K}) p_i \left(\frac{v(\mathrm{K}) + v(\mathrm{K}+1)}{2} \middle| \{\mu_i, \sigma_i\} \right)\,, \tag{6}$$

with $Q(\mathrm{K}) = v(\mathrm{K}+1) - v(\mathrm{K})$ the quantization step size of code K.

3.6 Free parameters

There is two categories of parameters, the problem to solve itself described by the global code transition points $\{v(\mathrm{K})\}$, and the local small gaussians described by $\{\mu_i, \sigma_i\}$ which are used for calibration. It is equivalent to use quantization step size $Q(\mathrm{K}) = v(\mathrm{K}+1) - v(\mathrm{K})$ instead of code transition point, booth notations will be used.

The initial problems brings a part of the solution: as the low signal dynamic of the cosmological signal over the HFI mission is concentrated on ± 512 codes around midscale, the full scale 2^{16} code transition points of the ADC does not need to be determined. This holds for the bright signal of galaxy and planets where the ADC effect is considered negligible. Also the hypothesis of the 64 codes periodicity reduces the parameters number, because each period can be described with 62 quantization step size, and 2 more parameters for the transitions located each 64 codes. However the 64 codes just after midscale do not follow the 64-code periodicity and have to be described separately with 62 parameters.

For the noise sources, each `ibias` measurements, is described by 80 values of μ_i corresponding to the shape of the electronic signal period, all of them sharing the same value for σ_i. Hence for the 20 different values of `ibias` used, 20 * (80) + 20 parameters are needed.

Finally to describe $\{v(\mathrm{K})\}$ we need 62 parameters for the pattern periodicity, 62 more for the codes after midscale that does not follow the periodicity, and 32 for the transitions at each period boundary in the range ± 512 codes around midscale. Also to describe the input gaussians $\{\mu_i, \sigma_i\}$, 1600 parameters are needed for the locations and 20 parameters for the standard deviation of noise, considering large period of time where the noise is supposed to be constant.

3.7 DNL reconstruction method

The number of free parameters is too high to work on a global analytical solution. So the reconstruction process used is a three steps iterative process, and the estimates computed at iteration index j will be noted with j as exponent.

At the beginning of the process, the ADC is supposed linear with $v^0(\mathrm{K}) = \mathrm{K}$, and the initial parameters for the noise are initialized with their biased estimates, σ_i^0 and μ_i^0. The iteration steps are:

1. ESTIMATION OF $\{\mu_i^j\}$: The position of the average electronic signal position μ_i^j is estimated using the intermediate parameters $\{v_i^j(\mathrm{K})\}$ which are the local CTP derived from $\{d_i(t)\}$ independently. The local CTP are estimated by ensuring $\mathrm{E}[\mathrm{h_i}(\mathrm{K})|\theta_\mathrm{i}^\mathrm{j}] = \mathrm{h_i}(\mathrm{K})$ (using Eq. 4), and using the global solution $\{v^{j-1}(\mathrm{K})\}$ obtained from the previous iteration. To do so, for each dataset $\{d_i(t)\}$, the difference between the local estimate $\{v_i^j(\mathrm{K})\}$ and the previous global estimate $\{v^{j-1}(\mathrm{K})\}$, is minimized iteratively. The expression to minimize is

$$\Phi_i = \sum_{\mathrm{K}} \left(Q_i^j(\mathrm{K}) - Q^{j-1}(\mathrm{K}) \right)^2 p_i \left(\frac{v_i^j(\mathrm{K}) + v_i^j(\mathrm{K}+1)}{2} \middle| \{\mu_i^j, \sigma_i^{j-1}\} \right) , \tag{7}$$

   where Φ_i is a criterion derived from Eq. (2, 5, 6), $Q_i^j(\mathrm{K}) = v_i^j(\mathrm{K}+1) - v_i^j(\mathrm{K})$, $Q^{j-1}(\mathrm{K}) = v^{j-1}(\mathrm{K}+1) - v^{j-1}(\mathrm{K})$. In practice, 100 iteration of a custom MCMC code are needed at this step.

2. ESTIMATION OF $\{v^j(\mathrm{K})\}$: The global estimate for code transition points is obtained by merging the local estimates $\{\{v_i^j(\mathrm{K})\}\}$ of step 1. In practice the retained effective solution is a linear combination of the local estimates weighted by the local statistics, and the global quantization step size reads

$$\begin{cases} Q^j(\mathrm{K}) = \left(\sum_i \mathrm{W}_i(\mathrm{K}) \right)^{-1} \left(\sum_i \mathrm{W}_i(\mathrm{K}) Q_i^j(\mathrm{K}) \right) \\ \mathrm{W}_i(\mathrm{K}) = h_i(\mathrm{K}) \, , \end{cases} \tag{8}$$

   where the expression for the local solution weight $\mathrm{W}_i(\mathrm{K})$ can be derived using again Eq. 5, and the first order approximation is equivalent to use the histogram statistics. The 64 code periodicity is enforced at the end of this step by averaging the quantization steps $Q_i^j(\mathrm{K} \bmod 64)$ for the 62 codes between two transitions.

3. ESTIMATION OF $\{\sigma_i^j\}$: The value of the local standard deviations is adjusted by minimizing the discrepancy between all the local $\{\{v_i^j(\mathrm{K})\}\}$ found at step 2. As stated before, changing the value of σ_i is equivalent to a local gain change, which can be written using quantization step size as $Q_i^j(\mathrm{K}) = \frac{\sigma_i^j}{\sigma_i^{j-1}} Q_i^{j-1}(\mathrm{K})$. Then the values of each σ_i^j independently is adjusted with a factor α_i^j to minimize the expression

$$f(\alpha_i^j) = \sum_{\mathrm{K}} w_i(\mathrm{K}) \left(Q_i^j(\mathrm{K}) - \alpha_i^j Q^j(\mathrm{K}) \right)^2 , \tag{9}$$

and using as weight $w_i(\mathrm{K}) = h_i(\mathrm{K})$, for $\frac{\partial f}{\partial \alpha_i^j} = 0$ it raises the solution

$$\begin{cases} \sigma_i^j &= \alpha_i^j \sigma_i^{j-1} \\ \alpha_i^j &= \dfrac{\sum_{\mathrm{K}} w_i(\mathrm{K}) Q_i^j(\mathrm{K}) Q^j(\mathrm{K})}{\sum_{\mathrm{K}} w_i(\mathrm{K}) Q(\mathrm{K})^2} \end{cases} \tag{10}$$

A total of 6 iterations with these three steps is sufficient for the convergency of parameters. The final output of the reconstruction method is presented in Fig 9. The high frequency defects (thick hairy features) tend to average out with noise, while the jumps each 64 code do not, the most visible being the central one which is about 1 LSB in size. Also the 64 codes periodicity constraint helps to reduce the random walk visible as a black line on the left intermediate INL. The relative error $\varepsilon(\mathrm{K})$ for each quantization step can be estimated as for an histogram with $1/\sqrt{\sum_i h_i(\mathrm{K})}$, but the 64 code periodicity allow to get a better estimate with $1/\sqrt{\sum_{\mathrm{K}} \sum_i h_i(\mathrm{K})/64}$, and the final relative error estimate is about 10^{-3} LSB. But in practice the error is dominated by local deviations from the 64 periodicity model.

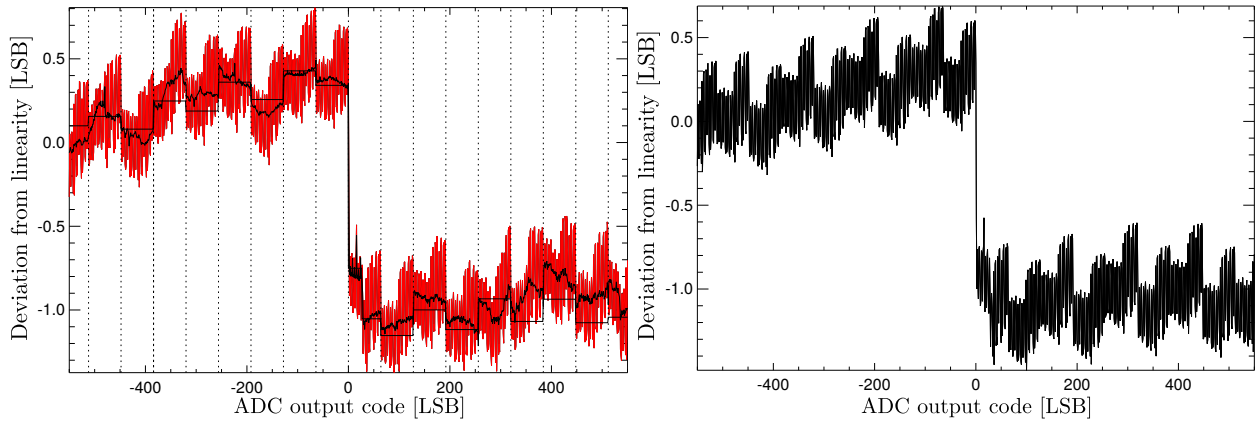


Figure 4. Reconstructed INL for the ADC of the channel `00_100-1a`, the 32768 value has been subtracted from the output code value. Left figure is taken from an intermediate step, when the 64 codes periodicity is not enforced, the black line is the deviation from the averaged 64 periodic pattern. Right figure is the final result after convergency of the iterative procedure.

4. CORRECTION OF SCIENCE DATA

As seen in Sec. 2.3, a science data sample is the result of a sum over 40 *fast samples* digitized by the ADC. This summation represents a major issue since ADC characterization can be applied only on the *fast samples* which remain unknown.

We describe in this section the global methodology applied to correct the HFI science data for the ADC nonlinearity, providing details about the two methods used to build the 2013 and 2015 data releases, respectively.

4.1 Methodology

The basic idea behind the correction of science data consist in estimating which *fast samples* have been summed to product one science data sample. Hence for building the correction function for science data we need a model for the analog signal response to sky power before digitization, and another model to describe the digitization of noisy signal with summation by the DPU. Combining these elements yields the global chain response to input sky power $\mathcal{F}(P_{sky}(t)) = d(t)$, where $P_{sky}(t)$ is the input sky power and $d(t)$ is the output science data sample.

4.1.1 Analog signal model

Let start with the analog signal $s(t)$ at the input of the ADC, just before digitization. Making a first order approximation, considering a constant input sky power, it is given by

$$s(t)|_{P_{sky}+\delta P_{sky}} = \overline{s}(t) + h(t) + \delta P_{sky} \times g(t) + \mathcal{O}(\delta {P_{sky}}^2)\,, \tag{11}$$

where $\overline{s}(t) = s(t)|_{P_{sky}}$, known as the *raw constant* (see Fig. 5 on the right), is the steady state modulation shape for the constant sky power P_{sky}, $h(t)$ is the 4K *lines* signal with an exact periodicity of 4.5 modulation periods, and $g(t) = \partial s(t)/\partial P_{sky}|_{P_{sky}}$, known as the *raw gain* (see Fig. 5 on the left) in *Planck* Collaboration, is the steady state response to a change δP_{sky} in the constant input sky power. In practice, this approximation is used by considering a slowly varying input sky power having a fixed value inside the modulation half period. It has been shown with an analytical model[3] that the slowly varying signal hypothesis still holds for a 2Hz pulsation of input sky power. Then as the main source of sky signal dynamic used for calibration (the solar dipole) is acquired at 0.016Hz, this approximation is considered valid.

4.1.2 Estimation of the *raw gain*

The *raw gain* is considered to be constant all over the mission. This hypothesis allow to extract the *raw gain* with a Principal Component Analysis (PCA) method on the captured *fast samples* periods. First the signal of the *fast samples* is ADC corrected using $\mathcal{A}^{-1}$, allowing to work with linear *fast samples*. However the 4K *lines* signal varies slowly during the mission and degenerates with the detector response if not taken into account. We reduce the impact of the 4K *lines* signal by taking the difference of two consecutive periods. As the capture time of the period is always at the same phase of the 4K *lines* signal the difference of two consecutive periods can be approximated using Eq. 11 with

$$\begin{aligned} s(t+\Delta T) - s(t) &= \epsilon(h(t)) + (P_{sky}(t+\Delta T) - P_{sky}(t))g(t) \\ \Delta s(t) &\simeq \Delta P_{sky} g(t) \end{aligned} \tag{12}$$

where ΔT is the period of 101 seconds between the capture of two consecutive *fast samples* period on their target detector, and the residual 4K *lines* term $\epsilon(h(t))$ is considered negligible. Finally, the *raw gain* is estimated by taking the primary eigenvector of the PCA algorithm run with the $\Delta s(t)$ vectors as input. The example of *raw gain* shown on the right of Fig. 5 is produced with the PCA algorithm.

The estimated error on the *raw gain*, by mean of a jacknife algorithm, is less than 1% on most channel. Which is sufficient for the correction of the cosmological signal dynamic.

4.1.3 ADC noise model for signal summation

The noise at the input of the ADC has also to be taken into account because of the DPU signal summation. As for the inflight measurements can be approximated as a white gaussian noise component $n_\sigma(t)$ added over the signal. The remaining noise at frequency higher than f_{mod} is considered as signal from the ADC point of view, and is not taken into account.

For a constant signal $s(t) = s$, the effect of noise is a local averaging of ADC codes (also called noise dithering), so the effective transfer function of the ADC for a noisy signal, taking the average value of many samples, is

$$\mathcal{A}_\sigma(s) = (\mathcal{A} * \sigma)(s)\,, \tag{13}$$

where $\mathcal{A} * \sigma$ is the convolution product between the ADC transfer function $\mathcal{A}$ and the normal probability density function with σ as standard deviation. The standard deviation of the noise is considered to be constant for all the mission and is about 3.7LSB for CMB channels and 1.7LSB for submm channels ($\texttt{gamp} = 1/3$).

Now let consider the DPU summation on the 40 samples of a half period. For white gaussian noise, samples are uncorrelated, it allows to approximate the effect of the noise in the digitization plus summation process with

$$\sum_{j=0}^{39} \mathcal{A}\Big(s(t+j\Delta t) + n_\sigma(t+j\Delta t)\Big) \simeq \sum_{j=0}^{39} \mathcal{A}_\sigma(s(t+j\Delta t)) \tag{14}$$

where Δt is the time between the digitization of two *fast samples*.

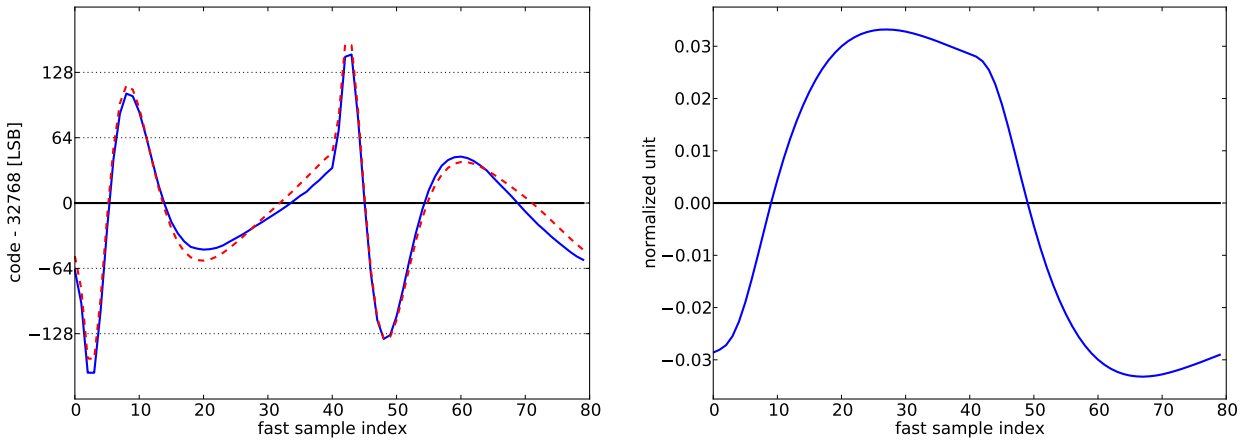


Figure 5. Left: the *raw constant* for the channel `03_143-1b` at ring index 17500, the blue line is for the averaged *fast samples* during the stable pointing period, and the red dashed line is after 4K *lines* removal. Right: the normalized *raw gain* for the same channel. The data acquired at midscale of the ADC (code 32768) will be the most affected by the nonlinearity, so the visible displacement induced by 4K *lines* affect the nonlinear response and need to be taken into account for a correction of ADC nonlinearity.

4.1.4 Science data correction function

Finally the correction has to be built considering 18 samples configuration, because the 4K *lines* signal $h(t)$ has a periodicity of 4.5 modulation periods. Then we need to build 18 different correction functions, but the main problem here is that the *fast samples* are only viewed through the one period window captured each 101 seconds, and $h(t)$ will degenerate with the steady state signal $\overline{s}(t)$. For this reason, the *raw constant* signal has to be estimated along with the 4K *lines*, and the estimation method will be described in the next two section.

Putting all together, we consider a stable pointing period, indexed by its ring[‡] number r, where $\overline{s}(t)$ and $h(t)$ are supposed to be stationary. Then the detector chain response function, for a science data sample at index i, can be approximated on a half period basis using Eq. 14 with

$$\mathcal{F}_{r,p,k}(P_{sky}(i)) = \sum_{j=0}^{39} \mathcal{A}_\sigma\Big(\mathbf{c}_{r,p}(j) + \mathbf{h}_{r,k}(j) + P_{sky}(i) \times \mathbf{g}_p(j)\Big)\,, \tag{15}$$

where $\mathbf{c}_{r,p}$ is the vector for the *raw constant* estimate of $\overline{s}(t)$ at parity index $p = i \bmod 2$, $\mathbf{h}_{r,k}$ is the vector estimate for the $4K$ lines signal $h(t)$ at 4K *lines* period index index $k = i \bmod 9$, and $\mathbf{g}_p$ is the global *raw gain* vector for $g(t)$ at parity p.

4.2 2013 parameters version : per ring estimation

For the 2013 data release,[4] the correction inputs are built for each stable pointing period on on a per ring basis. First the 4K *lines* signal $h(t)$ is estimated using a constrained least square method, then the *raw constant* is calculated by removing the estimate of $h(t)$ from $s(t)$.

Let consider a given pointing period, $\mathbf{p}$ is the average of the $nc \simeq 30$ *fast samples* periods captured during this stable pointing period at times $\{t_{nc}\}$ and ADC corrected with $\mathcal{A}^{-1}$. For the estimation of $h(t)$, it is possible to cancel the steady state signal by taking the sum $\mathbf{q}(i) = \mathbf{p}(i) + \mathbf{p}(i+40)$ of the two half periods. This operation cancels the f_{mod} odd harmonics and double the even f_{mod} harmonics in $\mathbf{q}$ including the zero mode (signal offset). The even harmonics can also be removed by using a template $\mathbf{q}_{\mathrm{CPV}}$ built from *fast samples*

[‡] *Planck* scans the same circle on the sky about 45 times at rate of one rotation per minute, the set of 45 circles is called a ring.

only acquisition made during the inflight Calibration and Performance Verification (CPV) phase. Now, we can construct the data vector used for estimation of $h(t)$ as

$$\begin{aligned}\mathbf{y}(i) &= \mathbf{q}(i) - \mathbf{q}_{\mathrm{CPV}}(i) - 2\,\widetilde{o} + \frac{\sqrt{2}}{\sqrt{nc}} n_\sigma(t_i) \\ &= h(t_i) + h(t_i + 40\Delta t) + \frac{\sqrt{2}}{\sqrt{nc}} n_\sigma(t_i)\,, \end{aligned} \tag{16}$$

where $\widetilde{o}$ is the signal offset estimation built using the average of science data.

The 4K *lines* signal can be written as a sum of harmonics $h(t) = \sum_k a_k \cos(k\omega_{4K} t) + b_k \sin(k\omega_{4K} t)$, with $\omega_{4K} = 2\pi f_{4K}$ and a_k, b_k are the unknown parameters. If we call $\mathbf{x}$ the vector of free parameters, we can build the matrix $\mathbf{F}$ of the linear function such as $\mathbf{Fx} = \mathbf{y}$. In the general case the solution is found for $\mathbf{x} = \left(\mathbf{F^T W F}\right)^{-1} \mathbf{F^T W y}$, where $\mathbf{W} = \mathbf{I_d}/(2\sigma^2/nc)$ is the inverse of the covariance matrix of $\mathbf{y}$. However this problem is highly underdetermined, because the *fast samples* period sees only 2/9 of the full 4K *lines* window, and 4K *lines* have a wide spectral distribution.

To add more constraints we use the science data estimates of the 4K *lines*. As for the *fast samples*, we can build the matrix $\mathbf{G}$ of the linear function such as $\mathbf{Gx} = \mathbf{y'}$, where $\mathbf{y'}$ is the constraints vector with eight values for the (cos, sin) parameters of the 4 folded frequencies of 4K *lines* on science data. Using the V. Blobel[5] elegant matrix transcription of the Lagrange multipliers least square with constraint methods, the new system to solve is

$$\left(\begin{array}{c|c} \mathbf{F^T W F} & \mathbf{G^T} \\ \hline \mathbf{G} & 0 \end{array}\right) \left(\begin{array}{c} \mathbf{x} \\ \hline \boldsymbol{\lambda} \end{array}\right) = \left(\begin{array}{c} \mathbf{F^T W y} \\ \hline \mathbf{y'} \end{array}\right)\,, \tag{17}$$

where $\boldsymbol{\lambda}$ is an unused output of the method.

In practice, the dominant folded 4K *lines* frequency on modulated science data is 20Hz, and a careful analysis of CPV data and *fast samples* allowed the selection of harmonic at index $\{1, 8, 10\}$ as the dominant non folded 20Hz candidates. So the chosen set of harmonics is $k \in \{1, 2, 3, 4, 8, 10\}$. This method give very good results on most channels but has some caveats:

- It does not work well with channels having a low or moderate amount of 4K *lines*;
- If more than one folded frequency is significant, there are too many unknown parameters to be solved;
- Electronic bias parameters have been set to new values after the CPV phase, so the CPV template has changed significantly for several channels.

4.3 2015 parameters version : global *raw constant*

The 4K *lines* problematic has not been solved completely with the 2013 version (see Sec. 5.2), and the ADC correction leave some significant residuals at large angular scale on the sky maps. These residuals are critical for the science that can be done with *Planck* data concerning the gravitational waves detection and the epoch or reionisation. For this reason a new algorithm has been developed. A careful analysis of result from the 2013 correction gave indications that the shape of the steady state signal $s(t)$ has been very constant during the mission (excepted from degenerated leaks caused by the 4K *lines*). Introducing this new hypothesis, the two major changes in the 2015 version of parameters estimation are the use of a global *raw constant* for the full mission, and the use of a custom MCMC algorithm to estimates 4K *lines* on each stable pointing period.

First, the computing of the global *raw constant* relies on estimation of corrected data quality, this is done by comparing the gain of science data for parity 0 and parity1. This comparison is calculated from Phase Binned Rings (PBR) which are 1-dimension projection of science data samples using the sky phase φ provided by the onboard star tracker. Using an order one polynomial fit, $(\mathrm{PBR}_0, \mathrm{PBR}_1)$ from booth parities are fitted with

$$\mathrm{PBR}_1(\mathrm{r}, \varphi) = -(1 + \varepsilon(\mathrm{r}))\mathrm{PBR}_0(\mathrm{r}, \varphi) + \mathrm{c}\,, \tag{18}$$

where $\varepsilon(r)$ is an estimate of the small gain mismatch due to the ADC nonlinearity, and c is a constant due to signal modulation modulation. In practice PBRs are built with 64 bins, because ADC nonlinearity has an impact only on the low frequencies of input signal, and this data compression speeds up considerably the computations by reducing the amount of data to process by a factor 3000.

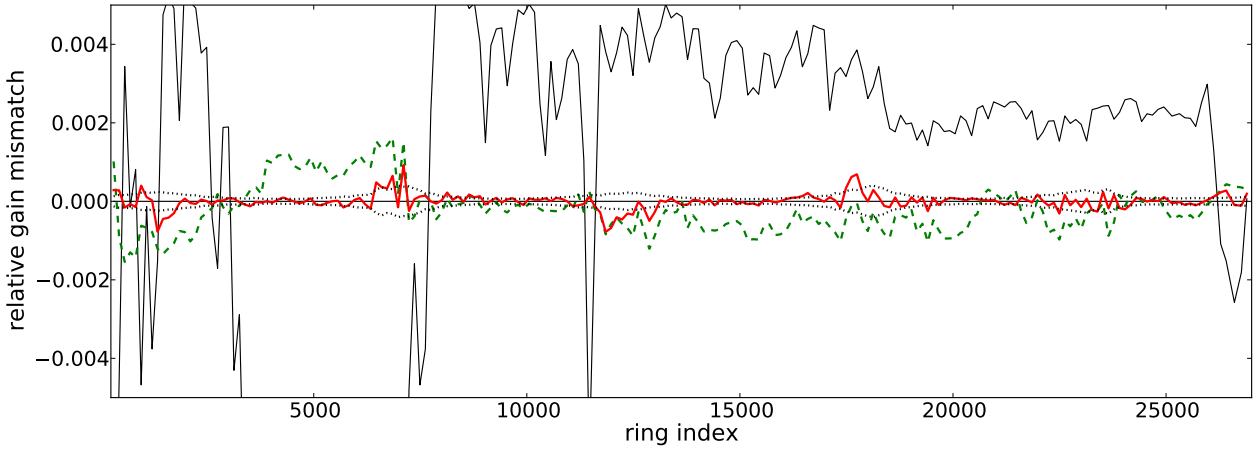


Figure 6. ADC nonlinearity observable from the relative gain mismatch $\varepsilon(r)$ between science data from parity 0 compared to parity 1, for the channel `03_143-1b`. The X axis is the stable pointing period index (ring) over the HFI mission, and the Y axis is the relative gain mismatch. Black thin line is for science data without correction (the full extent range is from -0.015 to 0.01), green dashed line is for 2013 correction and red thick line is for 2015 correction. To reduce the observable noise, the data is binned over 128 rings periods. The black thin dots envelope is the error estimation for the binned observable, the variations in amplitude of the envelope are due to different ring durations and variations in solar dipole amplitude.

To build the global *raw constant* (80 free parameters), the quantity $\varepsilon(r)$ is minimized with an iterative procedure. The signal model used for correction of PBRs is derived from Eq. 15 with

$$\mathcal{F}_{r,p,k}(P_{sky}(i)) = \sum_{j=0}^{39} \mathcal{A}_\sigma\Big((\widetilde{o}_r + \mathbf{c}_p(j)) + \mathbf{h}_{r,k}(j) + P_{sky}(i) \times \mathbf{g}_p(j)\Big)\,, \tag{19}$$

where $\mathbf{c}_p$ is the global *raw constant* vector at parity p, and $\widetilde{o}_r$ is the signal offset re-estimated after each correction iteration.

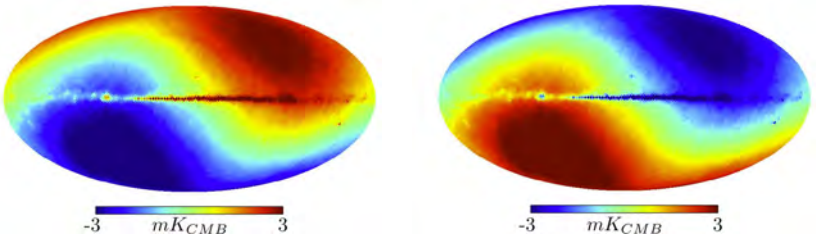
5. CORRECTION EFFICIENCY ON SCIENCE DATA

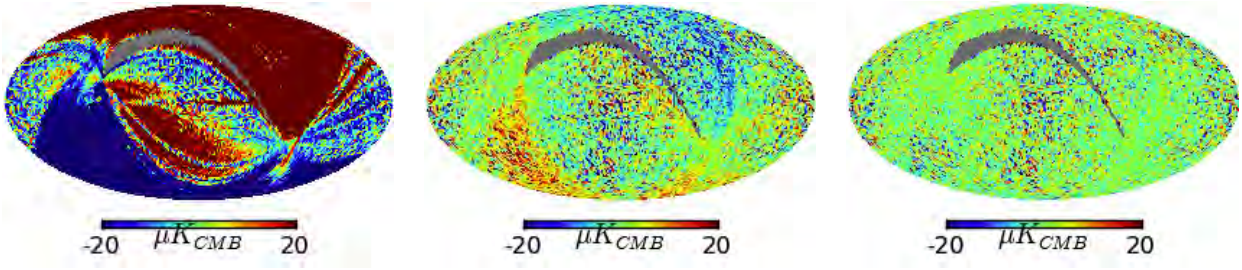
In this section we describe how we can build sky maps of ADC nonlinearity observables. We show how they can be used to make quantitative estimation of the impact of the ADC nonlinearity on science, and how efficient is the correction.

5.1 ADC residuals on sky with *half parity difference maps*

We have seen in Sec. 4.3 that the ADC nonlinearity effect is different for the two half periods of the modulation, because the analog signal is digitized by different regions of the ADC. To use this property, let consider the full sky *half parity map* SKY_0 built with samples from parity 0 (first half period) only, and SKY_1 built with samples from parity 1 (second half period) respectively (see Fig. 7). The sky signal at location Φ can be approximated with

$$\begin{cases} \mathrm{SKY}_0(\Phi) & = O_0 + S(\Phi) + n(\Phi) + w_0(\Phi) + a_0(\Phi) \\ \mathrm{SKY}_1(\Phi) & = O_1 - S(\Phi) - n(\Phi) + w_1(\Phi) + a_1(\Phi)\,, \end{cases} \tag{20}$$





multipoles, up to $\ell \sim 100$. The noise shape of these spectra is very different from the one of the official maps, because of the specific processing of the signal (one sample over two). As a consequence the *Planck* noise level cannot be taken as a reference, however the absolute level of ADC nonlinearity residuals can be compared to the source of photometric calibration as a rough impact estimator.

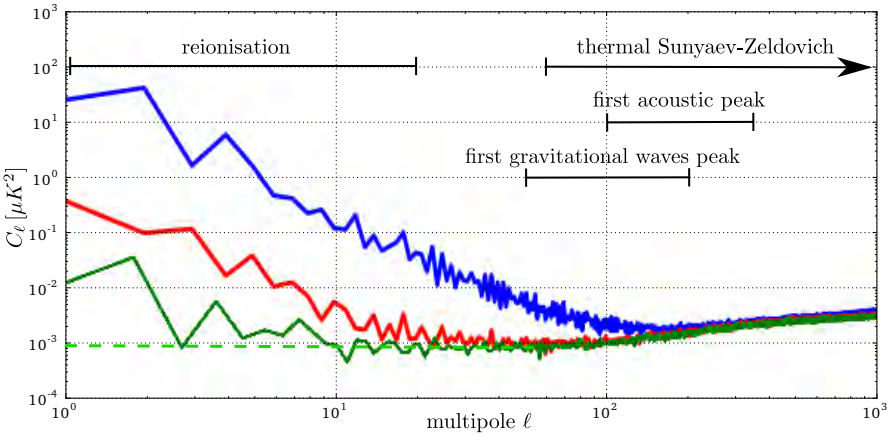


Figure 9. *TT* spectrum of ADC nonlinearity residuals built from half parity difference maps. Blue line (top) is uncorrected data. Red line (middle) is for 2013 correction. Green line (bottom) is for 2015 correction. The additional lightgreen dashed line indicates the expected level of noise, when no ADC nonlinearity is present.

The photometric calibration source for CMB channels is the orbital dipole[8] (10% of the solar dipole, i.e. an amplitude of about 300μK), which is a time variable source of calibration depending only of the satellite velocity with respect to the solar system, The orbital dipole is degenerated with the solar dipole in time domain for a given stable pointing period, so the ADC nonlinearity residuals at low frequency have a direct impact on the photometric calibration. On the example channel (see Fig. 9), a direct comparizon between the 2015 best case residuals (with first mode of the *TT* spectrum), and the orbital dipole amplitude, indicates that ADC nonlinearity bias is about 10^{-3} of the photometric calibration source. This value is mitigated by the averaging of all the channels for a given frequency, but the reduction does not behave as $\sqrt{n_{detectors}}$ as it would do with gaussian noise.

One of the major challenges of modern cosmology is the detection of B-mode polarization anisotropies in the CMB. The analysis of the *TT* spectrum of the *half parity difference maps* gives us a good indication of the ADC nonlinearity contamination for this detection, since in the multipole range where a detection is expected (around $\ell = 80$), the 2013 correction has already removed almost any significant residuals. However, the HFI polarized detectors are calibrated independently from each other and can suffer from the low multipole ADC nonlinearity residuals. So this intercalibration must be done very carefully otherwise the intensity can leak into the (Q,U) stokes parameters and build up some spurious B-modes. A possible solution against this kind of leak is to use the CMB first acoustic peak as a photometric calibration source, as it is in the ADC safe area of the spectrum (after, at least 2013, ADC correction).

The most problematic issue remains for the study of reionsation epoch, in the low multipole range, where there is the highest level of ADC nonlinearity residuals. The results are channel dependent because of the 4K *lines* which are the limiting factor in the correction efficiency (and mostly the reason for the effort put into the 2015 correction).

6. POSSIBLE SOLUTIONS FOR FUTURE MISSIONS

Some future missions, like the QUBIC[9] experiment will make use of sample summation, or similar processing to increase the Signal-to-Noise Ratio (SNR) with exquisite SNR requirements. We discuss in this section some possible design to avoid loss of information necessary for nonlinearity correction.

6.1 Increasing signal dynamic

A trivial solution is to increase the signal dynamic, compared to the ADC defects in LSB units. The Signal-to-Noise Ratio (SNR) for the quantization noise is generally calculated in decibels with the formula $SNR = 6.02n + 1.76\text{dB}$, where n is the number of bits of the ADC. But this theoretical number holds only in the case of an ideal component with the quantization noise averaging out, and considering the full input scale of the component. For a non-ideal component, a useful figure to characterize bias level is the maximum value of the INL, or in other words the maximum transfer function drift from linearity in LSB units. In this case the Signal-to-Bias Ratio, which does not average out with statistics, would read

$$SBR = 20\log_{10}\frac{S}{\text{INL}}\text{dB}\,, \tag{21}$$

where S is the considered signal dynamic converted in LSB units. So in the case of HFI, with a best case signal dynamic of 14LSB, for the solar dipole, and an inflight measured INL value about 1LSB, we have SBR = 23dB. If we would like to reach a requirement SBR= 80dB we would need at least 10 more bits with the same INL value of 1LSB. It is possible to discard non significant bits in the onboard DPU, but as the accuracy requirement also increase with new instruments, discarding too much bits in the acquired samples may not be possible.

Increasing moderately the steady state signal amplitude is also useful. In the case of HFI, the signal balancing was a success into making the resulting analog signal very flat (70% of the *fast samples* for a half period cover only 30LSB). But a flat signal can concentrate signal acquisition on a potentially important ADC defect. Changing the design of the balancing to obtain for example a triangular shape for the acquired signal would be a good solution to ensure acquisition is done on different ADC areas, the result would be a reduced bias level by averaging effect.

6.2 Using analog summation

The digital summation implemented in the DPU (label (c) on Fig. 1) is designed to improve the SNR of the science data. But another function has been implemented in the DPU with the `nblank` parameter, still not used in the HFI mission, to allow the removal of `nblank` *fast samples* at transient locations. If we forget about the `nblank` option, the digital summation function can be replaced with an equivalent analog integrator in the readout chain. The square signal of detectors would integrate into a triangle wave, and the acquisition could be done near the peak location once for each half period at twice the modulation frequency. This design does not destroy information after digitization, so nonlinear correction of the ADC component can be done a posteriori.

The main problem with analog integrators is their potential signal instability, so they have to be designed carefully. They also bring some additional noise as each analog component do.

6.3 DPU correction

It is possible to make the the ADC nonlinearity correction in the DPU. This method require to characterize each ADC and to provide this information to the DPU as data tables, one value for each ADC code being sufficient. Also, to avoid floating point operations in the FPGA, the tables must be built with significantly more bits than the ADC output. The main advantage is that complex digital processing is still available with this design, and it has a better SNR than with additional analog component.

For this solution, the hypothesis is that the nonlinear characteristics of the ADC are stable over the mission. Otherwise a runtime characterization must be implemented with a perfectly known calibration source like a ramp signal.

7. CONCLUSIONS

We have described how the Planck/HFI instrument is affected by the nonlinearity of the Analog to Digital Converter (ADC) component of the readout electronics. While this nonlinearity had not been properly characterised before launch, its impact on science data has been enhanced by a combination of circumstances: the small dynamic of the cosmological signal ; the location of the balanced signal at mid-scale of the entire ADC range, i.e., exactly on top of the strongest defect of the ADC ; and the summation of the digitizsed signal on the first and second modulation periods to produce the science signal, and destroying the knowledge of the full modulated signal. To handle this problematic and build an appropriate correction function, two steps have been necessary:

- The ADC chips nonlinearity had to be characterized inflight for each of the 54 detector channels. This was done with a specific EOL sequence on the last ten months of the mission to acquire statistics on the component range relevant for the mission cosmological signal acquisition. Then we have developed a likelihood method to reconstruct each ADC transfer function with a precision of about 10^{-3} LSB on a large code range.
- To correct science data for ADC nonlinearity, a model of the analog signal before digitization by the ADC had to be developed. This model allow us to revert the signal summation operation of the DPU, so that the ADC characterization can be used with this model to build the global science data correction.

In order to estimate the ADC correction efficiency on the final maps, we have also developed a specific method consisting in building two sets of all-sky maps containing only samples from the first or second half period of the modulation, respectively. The difference of these two sets of maps allows us to remove all low frequency signals and to conserve only ADC nonlinearity residuals plus white noise. The analysis of the temperature (TT) power spectrum built from *half parity difference maps* is very powerful to characterize the impact of ADC nonlinearity in the angular frequency domain. The ADC residuals impact significantly the TT spectrum in the first 100 multipoles when no correction is applied. This range is reduced to the first 30 multipoles with the correction included in the 2013 data release, and to the first 10 multipoles with the 2015 correction. The remaining residuals level on CMB channels is about 0.05% of the solar dipole in time domain, and 0.1% of the main photometric calibration source (the orbital dipole) on the TT spectrum. This level seems safe when analysing the CMB TT spectrum, but remains a possible issue when considering polarisation signals, especially for B-mode or reionisaiton analyses performed at low multipoles.

Finally we have presented how such a large impact from the nonlinearity of the ADC component can be avoided, and inflight ADC correction using tables built from ground calibration seems to be the most promising solution.

ACKNOWLEDGMENTS

The work presented in this paper has been realized inside the *Planck* team responsible of the data processing.

The development of *Planck* has been supported by: ESA; CNES and CNRS/INSU-IN2P3-INP (France); ASI, CNR, and INAF (Italy); NASA and DoE (USA); STFC and UKSA (UK); CSIC, MICINN and JA (Spain); Tekes, AoF and CSC (Finland); DLR and MPG (Germany); CSA (Canada); DTU Space (Denmark); SER/SSO (Switzerland); RCN (Norway); SFI (Ireland); FCT/MCTES (Portugal); and PRACE (EU). A description of the *Planck* Collaboration and a list of its members, including the technical or scientific activities in which they have been involved, can be found at http://www.rssd.esa.int/index.php?project=PLANCK&page=PlanckCollaboration.

REFERENCES

[1] Planck Collaboration, Adam, R., Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., et al., "Planck 2015 results. VII. HFI TOI and beam processing," *ArXiv e-prints* (Feb. 2015).

[2] Lamarre, J.-M., Puget, J.-L., Ade, P. A. R., et al., "Planck pre-launch status: The HFI instrument, from specification to actual performance," A&A **520**, A9 (Sept. 2010).

[3] Sauvé, A. and Montier, L., "Building the analytical response in frequency domain of AC biased bolometers Application to Planck/HFI," *ArXiv e-prints* (May 2016).

[4] Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., et al., "Planck 2013 results. VI. High Frequency Instrument data processing," A&A **571**, A6 (Nov. 2014).

[5] Blobel, V. and Kleinwort, C., "A New Method for the High-Precision Alignment of Track Detectors," *ArXiv High Energy Physics - Experiment e-prints* (Aug. 2002).

[6] Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M., "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere," ApJ **622**, 759–771 (Apr. 2005).

[7] Tristram, M., Filliard, C., Perdereau, O., Plaszczynski, S., Stompor, R., and Touze, F., "Iterative destriping and photometric calibration for Planck-HFI, polarized, multi-detector map-making," A&A **534**, A88 (Oct. 2011).

[8] Planck Collaboration, Adam, R., Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., et al., "Planck 2015 results. VIII. High Frequency Instrument data processing: Calibration and maps," *ArXiv e-prints* (Feb. 2015).

[9] Qubic Collaboration, Battistelli, E., Baú, A., Bennett, D., Bergé, L., Bernard, J.-P., de Bernardis, P., Bordier, G., Bounab, A., Bréelle, É., et al., "QUBIC: The QU bolometric interferometer for cosmology," *Astroparticle Physics* **34**, 705–716 (Apr. 2011).

Appendix C

HFI ADC chip specifications design sheet

7809 LPRP

Analog-to-Digital Converter

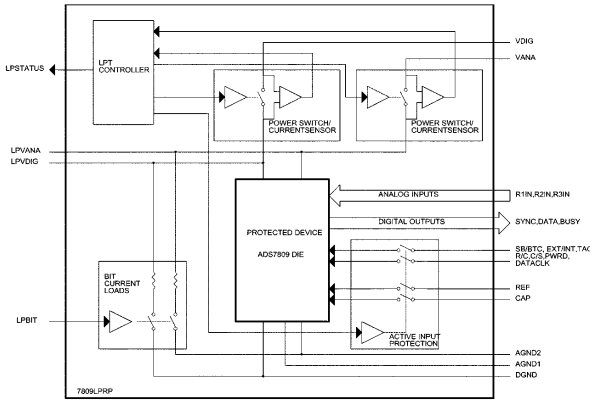
- 16-bit organization
- 24 pin RAD-PAK® flat package
- 100 kHz min sampling rate
- ±10 V and 0 V to 5 V input range
- Advanced CMOS technology
- DNL: 16-bits "No Missing Codes"
- 86dB min SINAD with 20 kHz input
- Single +5 V supply operation
- Utilizes internal or external reference
- Serial output
- Power dissipation: 132 mW max

Space Electronics Analog-to-Digital Converter

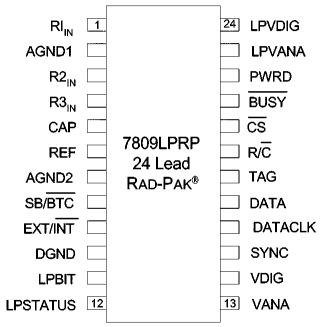
Space Electronics' 7809LPRP is a radiation-hardened, high speed 16-bit analog-to-digital converter features up to 100 kilorad (Si) total dose tolerance dependent upon orbit. The device is radiation-hardened at the microcircuit level using SEi's RAD-PAK® packaging technology and is latchup protected by Space Electronics' Latchup Protection Technology (LPT™). The device is a 24 pin, 16-bit sampling analog-to-digital converter that uses state-of-the-art CMOS structures. The 7809LPRP contains a 16-bit capacitor based SAR A/D with S/H, reference, clock, interface for microprocessor use, and serial output drivers. It is specified at a 100 kHz sampling rate, and guaranteed over the full temperature range. Laser-trimmed scaling resistors provide various input ranges including ±10V and 0 to 5 V, while the innovative design allows operation from a single +5 V supply, with power dissipation of under 132 mW. In addition, the 7809LPRP is available with packaging and screening up to Class S.

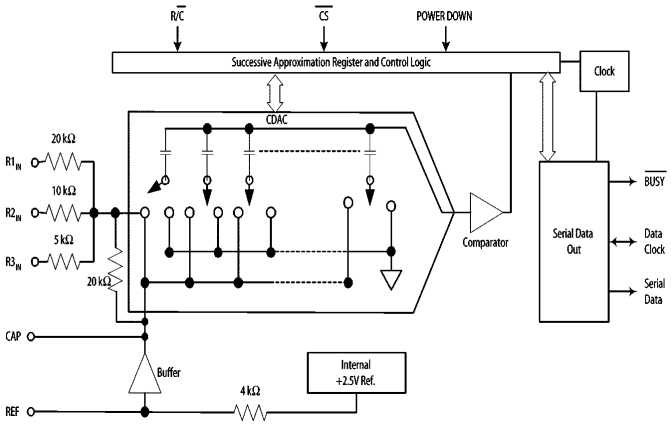
LPT™ Operation

Latchup Protection Technology (LPT™) automatically detects an increase in the supply current of the 7809LPRP converter due to a single event effect and internally cycles the power to the converter off, then on, which restores the steady state operation of the device. The LPTTM circuitry consists of two power switch and current sensor blocks, an LPTTM controller block, a BIT cur- rent load block, and an active input protection block.



Analog-to-Digital Converter





Absolute Maximum Ratings

PARAMETER	SYMBOL	MIN	MAX	UNIT
Analog Inputs	R1IN	-25	+25	V
	R2IN	-25	+25	V
	R3IN	-25	+25	V
	CAP	VANA +0.3	AGND2 -0.3	V
	REF[1]			
Ground Voltage Differences	DGND, AGND2	-0.3	+0.3	V
VANA		--	7	V
VDIG		--	7	V
VDIG to VANA		--	+0.3	V
Specified Performance		-40	85	°C
Digital Inputs		-0.3	VDIG +0.3	V
Derated Performance Temperature	TA	-55	+125	°C
Storage Temperature	TSTG	-65	+150	°C

1. Indefinate short to AGND2, momentarily short to VANA

Accuracy Specifications

PARAMETER	PARAMETER	MIN	TYP	MAX	UNIT
Integral Linearity Error		--	--	±3	LSB[1]
Differential Linearity Error		--	--	+3, -2	LSB
No Missing Codes[2]		15	--	--	Bits
Transition Noise[3]		--	1.3	--	LSB
Full Scale Error[4,5]		--	--	±0.5	%
Full Scale Error Drift		--	±7	--	ppm/°C
Bipolar Zero Error[4]		--	--	±10	mV
Bipolar Zero Error Drift		--	±2	--	ppm/°C
Unipolar Zero Error[4]		--	--	±3	mV
Unipolar Zero Error Drift		--	±2	--	ppm/°C
Power Supply Sensitivity	4.8 V<VS<5.25V	--	--	±8	LSB

1. LSB stands for Least Significant Bit. One LSB is equal to 305 uV. 2. Not tested. 3. Typical rms noise at worst case transitions and temperatures. 4. Measured with various fixed transistors.

Space Electronics
A Maxwell Technologies Company
9244 Balboa Avenue, San Diego, CA 92103
tel: 858-503-3300
fax: 858-503-3301
e-mail: sales@spaceelectronics.com

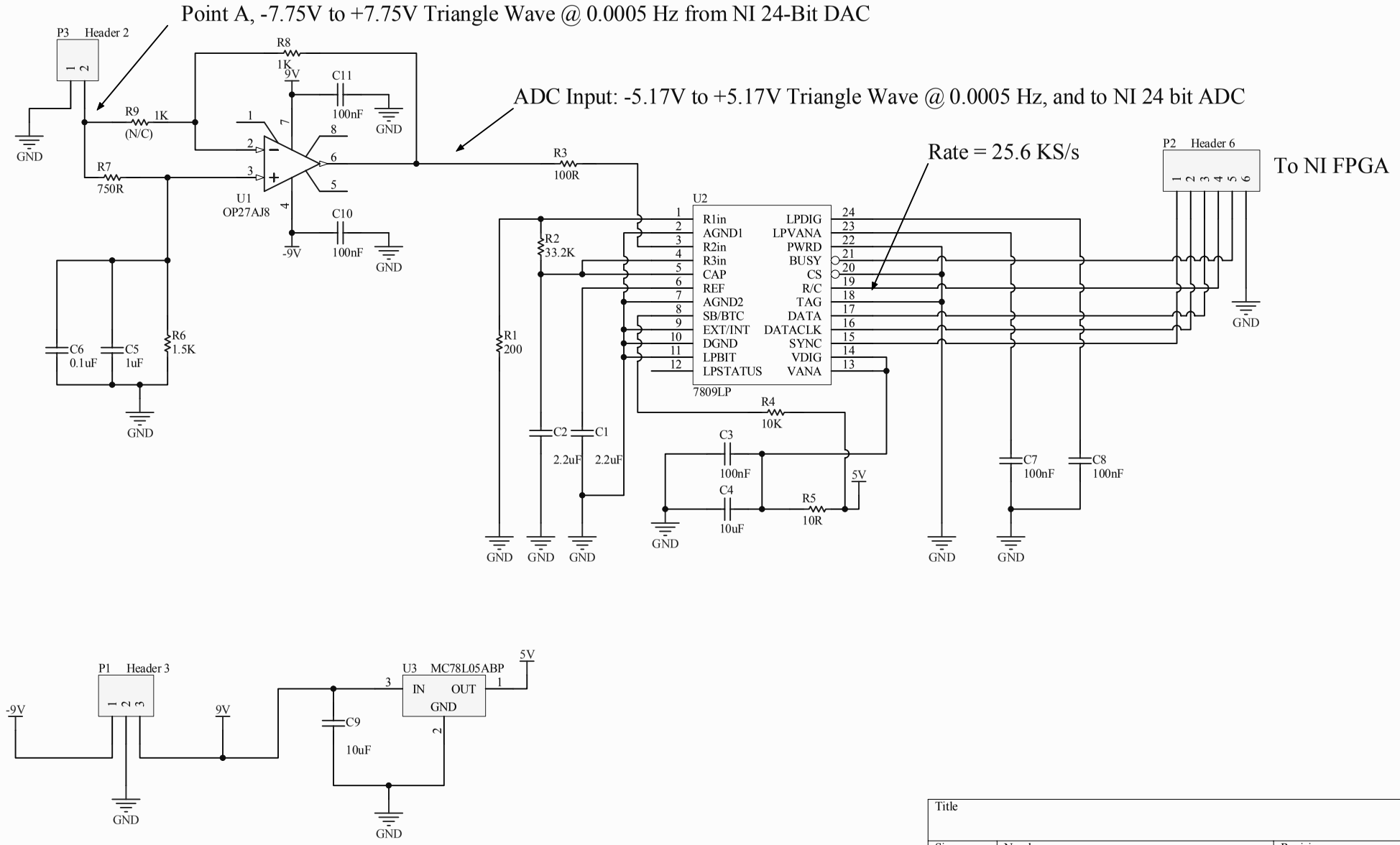
Appendix D

Orsay ground calibration datasets

Channel	Dataset Label	Nb of raw periods	Additional parameters
FS60	B-60-I0	4382409	`vbal`=0
FS62	B-62-I0	560027	`vbal`=0
FS64	B-64-I0	601358	`vbal`=0
FS65	B-65-I0	523880	`vbal`=0
FS60	B-60-I1-V2	5050590	`ibias`=1, `vbal`=2
FS60	B-60-I1-V3	188720	`ibias`=1, `vbal`=3 (Shortened acquisition)
FS60	C-60-ISTEP	900 periods / ibias step	`ibias` from 0 to 49

Appendix E

Schematics of the Cardif setup

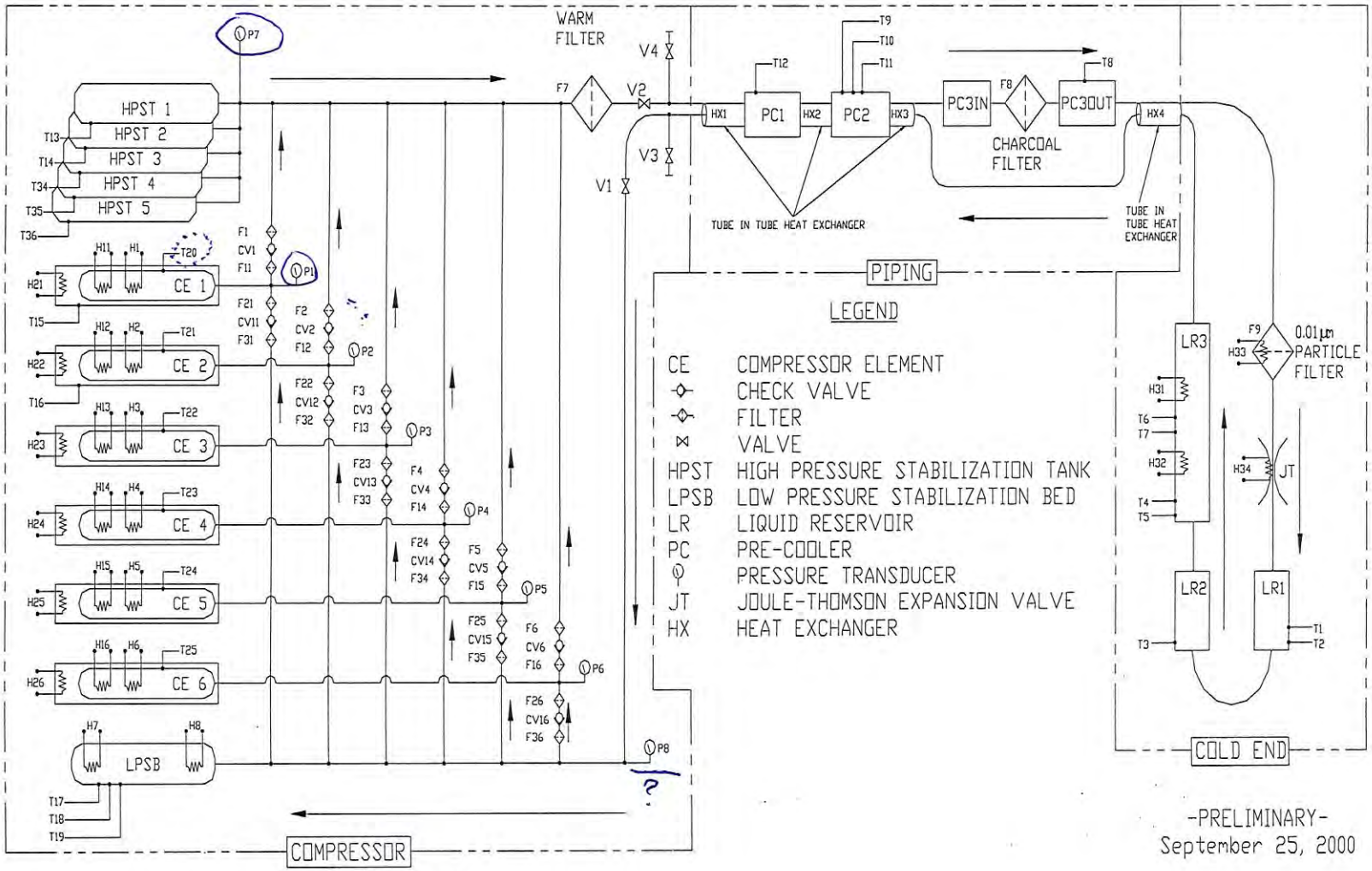


Title		
Size A4	Number	Revision
Date: 14/04/2012		Sheet of
File: C:\Users\..\Planck_ADC_7809LP.SchDoc		Drawn By:

Appendix F

Sorption cooler schematic and sensors

16. Annex 11: Sorption Cooler short description



(Pages=95)

LPSB Low Pression stabilization Bed
HPST

11 ANNEX F: LIST OF SENSORS

Device Number (Hex)	Item	Location	Function	Type	Nominal / Redundant / Single	Parameter name in the MIB	Calibration curve type	Calibration type condition 0 = no condition 1 = cooler + electronics 2 = cooler only
0x1	T1	LR1	Temperature	Cernox	Nominal to T2	SM018540	Numerical	1
0x2	T2	LR1	Temperature	Cernox	Redundant to T1	SM074540		
0x3	T3	LR2	Temperature	Cernox	Nominal to T4	SM019540		
0x4	T4	GROUND	Temperature	Cernox	Redundant to T3	SM075540		
0x8	T5	LR3	Temperature	Cernox	Nominal to T6	SM020540		
0x9	T6	LR3	Temperature	Cernox	Redundant to T5	SM076540		
0x5	T7	JT	Temperature	Cernox	Nominal to T30	SM021540		
0x10	T8	PC3C	Temperature	Cernox	Single	SM022540		
0x11	T9	PC3B	Temperature	Cernox	Single	SM023540		
0x12	T10	PC3A	Temperature	Cernox	Single	SM024540		
0x18	T11	PC2	Temperature	Cernox	Single	SM025540		
0x20	T12	PC1	Temperature	Cernox	Single	SM026540		
0x28	T13	HPST1	Temperature	Prt	Single	SM027540		
0x29	T14	HPST2	Temperature	Prt	Single	SM028540		
0x30	T15	CE1 Housing	Temperature	Prt	Single	SM029540		
0x31	T16	CE2 Housing	Temperature	Prt	Single	SM030540		
0x2a	T17	LPSB	Temperature	Prt	Single	SM031540		
0x2b	T18	LPSB	Temperature	Prt	Single	SM032540		
0x40	T20	CE1	Temperature	Ktc	Single	SM034540	The same numerical one	0
0x41	T21	CE2	Temperature	Ktc	Single	SM035540		
0x42	T22	CE3	Temperature	Ktc	Single	SM036540		
0x43	T23	CE4	Temperature	Ktc	Single	SM037540		
0x44	T24	CE5	Temperature	Ktc	Single	SM038540		
0x45	T25	CE6	Temperature	Ktc	Single	SM039540		
0x32	T26	CE3 Housing	Temperature	Prt	Single	SM040540	Numerical	1
0x33	T27	CE4 Housing	Temperature	Prt	Single	SM041540		
0x34	T28	CE5 Housing	Temperature	Prt	Single	SM042540		
0x35	T29	CE6 Housing	Temperature	Prt	Single	SM043540		
0x6	T30	J-T	Temperature	Cernox	Redundant to T7	SM077540		
0x48	P1	Comp Elem 1	Pressure		Single	SM044540	Numerical (with only 2 points)	2
0x49	P2	Comp Elem 2	Pressure		Single	SM045540		
0x4a	P3	Comp Elem 3	Pressure		Single	SM046540		
0x4b	P4	Comp Elem 4	Pressure		Single	SM047540		
0x4c	P5	Comp Elem 5	Pressure		Single	SM048540		
0x4d	P6	Comp Elem 6	Pressure		Single	SM049540		
0x4e	P7	HPST Heater	Pressure		Single	SM050540		
0x4f	P8	LPSB Heater	Pressure		Single	SM051540		
0x50	U1	SCE After the 28V converter	Voltage		Single	SM052540	Polynomial	0
0x51	U2	SCE After the 12V converter	Voltage		Single	SM053540		
0x52	U3	SCE After the 5V converter	Voltage		Single	SM054540		
0x53	U4	SCE After the +12V converter	Voltage		Single	SM055540		
0x54	U5	SCE After the –12V converter	Voltage		Single	SM056540		
0x55	U6	SCE tension 31V	Voltage		Single	SM101540		
0x56	I1	SCE current LPSB H7	Intensity		Single	SM058540		
0x58	I2	SCC courant H21	Intensity		Single	SM059540		
0x59	I3	SCC courant H22	Intensity		Single	SM060540		
0x5a	I4	SCC courant H23	Intensity		Single	SM061540		
0x5b	I5	SCC courant H24	Intensity		Single	SM062540		
0x5c	I6	SCC courant H25	Intensity		Single	SM063540		
0x5d	I7	SCCcourant H26	Intensity		Single	SM064540		
0x5e	I8	SCC courant H31	Intensity		Single	SM065540		
0x5f	I9	SCC courant H32	Intensity		Single	SM066540		
0x65	TE1	SCE Cold Face	SCE Temperature		Single	SM071540	The same numerical one (with only 2 points)	0
0x66	TE2	SCE Warmest part of PCB	SCE Temperature		Single	SM072540		
0x67	TE3	SCE Warm face	SCE Temperature		Single	SM073540		
0x07	CAL1	Test	Calibration		Single	SM069540	N/A	
0x2c	CAL2	Test	Calibration		Single	SM070540		
0x68	VGAIN	Test	Voltage		Single	SM067540		
0x70	VOFF	Test	Voltage		Single	SM068540		

Figure 92: List of sensors

sortie du mode RUN.

Appendix G

DMC objects Location

Reminder: DMC stands for Database Management Center

Group	Object	Data Type	Comments
Sa_HFI_C_Bolo	HFI_BC_C	INT32	Science data TOI
Sa_HFI_C_Bolo	HFI_BC_C	DOUBLE	Phase TOI
Sa_HFI_BC _RG	HFI_BC_RG	INT32	Phase TOI
Sa_HFI_BC _RG	TIMES_OF_SAMPLES	INT64	*fast samples* timestamps

Table G.1: DMC objects location for detectors channels. The BC label is meant to be replaced by the 2 characters channel code.

Appendix H

Transfer function of the DPU summation

This appendix describes the exact analytical version of the transfer function for the DPU summation which associates to the Fourier coefficient an unfolded analog signal harmonic, its corresponding Fourier coefficient on science data. Only the unfolded 4 K lines harmonic frequencies will be considered.

This version is different from the one used in the official model of the time transfer function of *HFI* electronics, because the later model the summation as an integration with a cardinal sin function, which is slightly inaccurate.

H.1 Parasitic wave description

Let's consider a sinusoidal parasitic wave measured by the *HFI* readout electronics on *fast samples* data as the real part of

$$s_h(t) = a_h e^{j\omega_h t}\,. \tag{H.1}$$

With $\omega_h = 2\pi f_{\rm acq} h/9$ being the angular frequency of the considered 4K harmonic, and h the integer multiple of the 20Hz harmonic ($f_{4K}/2 = 2f_{\rm acq}/9$).

The 4K cooler harmonics are synced by design with the *HFI* signal modulation. Hence it is simpler and more accurate to make a variable change and then use the scientific sample index $i = t.f_{\rm acq}$ instead of time, leading to

$$s_h(i) = a_h e^{j\frac{\omega_h}{f_{\rm acq}}i}\,. \tag{H.2}$$

This expression can be simplified further by defining

$$\overline{w_h} = \frac{2\pi h}{9} = \frac{\omega_h}{f_{\rm acq}}\,, \tag{H.3}$$

which express the angular frequency in radian/samples units.

And finally the wave expression reads

$$s_h(i) = a_h e^{j\overline{w_h}i}\,. \tag{H.4}$$

H.2 Raw samples integration effect

The scientific signal is built by summation of *fast samples* data by chunk of $N = 40$ consecutive values for each half period and reads

$$\begin{aligned} S_h[i] &= a_h \sum_{k=0}^{N-1} s(i + \frac{k}{N} + \frac{\varphi_0}{\overline{w_h}}) \\ &= a_h \sum_{k=0}^{N-1} e^{j\left(\overline{w_h}(i+\frac{k}{N})+\varphi_0\right)} \\ &= a_h e^{j(\overline{w_h}i+\varphi_0)} \sum_{k=0}^{N-1} \left(e^{j\frac{\overline{w_h}}{N}}\right)^k \end{aligned} \tag{H.5}$$

with $i \in \mathbb{N}^*$ and φ_0 in [radian] units being the acquisition reference for $S_h[0]$. The `sphase` electronic parameter is contained in φ_0 and will not be referenced any more.

The sum can be reduced to a polynomial fraction using the expression of a geometric series (see: Wikipedia/Geometric_progression) and becomes

$$\sum_{k=0}^{N-1} x^k = \frac{1-x^N}{1-x}\,. \tag{H.6}$$

Setting $x = e^{j\frac{\overline{w_h}}{N}}$ in Eq. H.6, Eq. H.5 rewrites to

$$S_h[i] = a_h e^{j(\overline{w_h}i+\varphi_0)} \frac{1-e^{j\overline{w_h}}}{1-e^{j\frac{\overline{w_h}}{N}}}\,. \tag{H.7}$$

Next we want to use the euler relation $sin(x) = (e^{jx} - e^{-jx})/(2j)$ to simplify the right part of Eq. H.7 which becomes

$$\begin{aligned} S_h[i] &= a_h e^{j(\overline{w_h}i+\varphi_0)} \frac{e^{\frac{\overline{w_h}}{2}}\left(e^{-\frac{\overline{w_h}}{2}} - e^{\frac{\overline{w_h}}{2}}\right)}{e^{\frac{\overline{w_h}}{2N}}\left(e^{-\frac{\overline{w_h}}{2N}} - e^{\frac{\overline{w_h}}{2N}}\right)} \\ &= a_h e^{j(\overline{w_h}i+\varphi_0)} e^{j\overline{w_h}\frac{N-1}{2N}} \left(\frac{-2jsin(\frac{\overline{w_h}}{2})}{-2jsin(\frac{\overline{w_h}}{2N})}\right) \\ &= a_h e^{j\left(\overline{w_h}(i+\frac{N-1}{2N})+\varphi_0\right)} \frac{sin(\frac{\overline{w_h}}{2})}{sin(\frac{\overline{w_h}}{2N})} \end{aligned} \tag{H.8}$$

H.3 Raw samples integration TF for 4K lines

From Eq. H.8 we can write the transfer function as

$$\begin{cases} H(j\overline{w_h}) = \frac{\mathcal{F}(S_h[i])}{\mathcal{F}(s_h[i])} = e^{j\left(\overline{w_h}\frac{N-1}{2N}+\varphi_0\right)} \frac{sin(\frac{\overline{w_h}}{2})}{sin(\frac{\overline{w_h}}{2N})} \\ \overline{w_h} = \frac{2\pi h}{9}, h \in \mathbb{N} \\ N = 40 \end{cases}\,. \tag{H.9}$$

The transfer function modulus is shown in Fig. 6.27 for the unfolded 4 K lines frequencies.

The poles of this function are located at multiples of acquisition frequency when $\frac{h}{9N} \in \mathbb{N}$. Theses frequencies are however not considered, since the result of the integration is known to be zero.

The zeros are also located at multiples of acquisition frequency when $\frac{h}{9} \in \mathbb{N}$, and are not relevant for 4K lines characterization since theses harmonics are fully described by the *raw constant*.

Bibliography

B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6):061102, Feb. 2016. doi: 10.1103/PhysRevLett.116.061102.

H. Abdi and L. J. Williams. Principal component analysis. 2(4):433–459, 2010. doi: 10.1002/wics.101.

W. R. Bennett. Spectra of quantized signals. *Bell Systems Technical Journal*, 27(3):446–472, July 1948. URL `http://www.alcatel-lucent.com/bstj/vol27-1948/articles/bstj27-3-446.pdf`.

A. Benoît, P. Ade, A. Amblard, R. Ansari, E. Aubourg, J. Bartlett, J.-P. Bernard, R. S. Bhatia, A. Blanchard, J. J. Bock, A. Boscaleri, F. R. Bouchet, A. Bourrachot, P. Camus, F. Couchot, P. de Bernardis, J. Delabrouille, F.-X. Désert, O. Doré, M. Douspis, L. Dumoulin, X. Dupac, P. Filliatre, K. Ganga, F. Gannaway, B. Gautier, M. Giard, Y. Giraud-Héraud, R. Gispert, L. Guglielmi, J.-C. Hamilton, S. Hanany, S. Henrot-Versillé, V. V. Hristov, J. Kaplan, G. Lagache, J.-M. Lamarre, A. E. Lange, K. Madet, B. Maffei, D. Marrone, S. Masi, J. A. Murphy, F. Naraghi, F. Nati, G. Perrin, M. Piat, J.-L. Puget, D. Santos, R. V. Sudiwala, J.-C. Vanel, D. Vibert, E. Wakui, and D. Yvon. Archeops: a high resolution, large sky coverage balloon experiment for mapping cosmic microwave background anisotropies. *Astroparticle Physics*, 17:101–124, May 2002. doi: 10.1016/S0927-6505(01)00141-4.

M. Bersanelli, N. Mandolesi, R. C. Butler, A. Mennella, F. Villa, B. Aja, E. Artal, E. Artina, C. Baccigalupi, M. Balasini, G. Baldan, A. Banday, P. Bastia, P. Battaglia, T. Bernardino, E. Blackhurst, L. Boschini, C. Burigana, G. Cafagna, B. Cappellini, F. Cavaliere, F. Colombo, G. Crone, F. Cuttaia, O. D'Arcangelo, L. Danese, R. D. Davies, R. J. Davis, L. de Angelis, G. C. de Gasperis, L. de La Fuente, A. de Rosa, G. de Zotti, M. C. Falvella, F. Ferrari, R. Ferretti, L. Figini, S. Fogliani, C. Franceschet, E. Franceschi, T. Gaier, S. Garavaglia, F. Gomez, K. Gorski, A. Gregorio, P. Guzzi, J. M. Herreros, S. R. Hildebrandt, R. Hoyland, N. Hughes, M. Janssen, P. Jukkala, D. Kettle, V. H. Kilpiä, M. Laaninen, P. M. Lapolla, C. R. Lawrence, D. Lawson, J. P. Leahy, R. Leonardi, P. Leutenegger, S. Levin, P. B. Lilje, S. R. Lowe, P. M. Lubin, D. Maino, M. Malaspina, M. Maris, J. Marti-Canales, E. Martinez-Gonzalez, A. Mediavilla, P. Meinhold, M. Miccolis, G. Morgante, P. Natoli, R. Nesti, L. Pagan, C. Paine, B. Partridge, J. P. Pascual, F. Pasian, D. Pearson, M. Pecora, F. Perrotta, P. Platania, M. Pospieszalski, T. Poutanen, M. Prina, R. Rebolo, N. Roddis, J. A. Rubiño-Martin, M. J. Salmon, M. Sandri, M. Seiffert, R. Silvestri, A. Simonetto, P. Sjoman, G. F. Smoot, C. Sozzi, L. Stringhetti, E. Taddei, J. Tauber, L. Terenzi, M. Tomasi, J. Tuovinen, L. Valenziano, J. Varis, N. Vittorio, L. A. Wade, A. Wilkinson, F. Winder, A. Zacchei, and A. Zonca. *Planck* pre-launch status: Design and description of the Low Frequency Instrument. A&A, 520:A4, Sept. 2010. doi: 10.1051/0004-6361/200912853.

P. Bhandari, M. Prina, R. C. Bowman, C. Paine, D. Pearson, and A. Nash. Sorption coolers using a continuous cycle to produce 20 K for the Planck flight mission. *Cryogenics*, 44: 395–401, June 2004. doi: 10.1016/j.cryogenics.2004.02.022.

BICEP2/Keck Array and Planck Collaborations. Joint Analysis of BICEP2/Keck Array and Planck Data. Phys. Rev. Lett., 114(10):101301, Mar. 2015. doi: 10.1103/PhysRevLett.114.101301.

V. Blobel. Least Square Methods. 2003. URL `http://graduierten-kurse.physi.uni-heidelberg.de/WiSe2003/lectures/blobel/export/lsqu/master.pdf`.

V. Blobel and C. Kleinwort. A New Method for the High-Precision Alignment of Track Detectors. *ArXiv High Energy Physics - Experiment e-prints*, Aug. 2002.

P. Bobik, G. Boella, M. J. Boschini, C. Consolandi, S. D. Torre, M. Gervasi, D. Grandi, K. Kudela, S. Pensotti, P. G. Rancoita, and M. Tacconi. Systematic investigation of solar modulation of galactic protons for solar cycle 23 using a monte carlo approach with particle drift effects and latitudinal dependence. *The Astrophysical Journal*, 745(2):132, 2012. URL `http://stacks.iop.org/0004-637X/745/i=2/a=132`.

B. Cabral and L. C. Leedom. Imaging vectors using line integral deconvolution. In *Special Interest Group on GRAPHics and Interactive Techniques Proceedings.*, Special Interest Group on GRAPHics and Interactive Techniques Proceedings., pages 263–270, 1993.

J. E. Carlstrom, G. P. Holder, and E. D. Reese. Cosmology with the Sunyaev-Zel'dovich Effect. ARA&A, 40:643–680, 2002. doi: 10.1146/annurev.astro.40.060401.093803.

J. E. Carlstrom, P. A. R. Ade, K. A. Aird, B. A. Benson, L. E. Bleem, S. Busetti, C. L. Chang, E. Chauvin, H.-M. Cho, T. M. Crawford, A. T. Crites, et al. The 10 Meter South Pole Telescope. PASP, 123:568–581, May 2011. doi: 10.1086/659879.

A. Catalano, P. Ade, Y. Atik, A. Benoit, E. Bréele, J. J. Bock, P. Camus, M. Chabot, M. Charra, B. P. Crill, N. Coron, A. Coulais, F.-X. Désert, L. Fauvet, Y. Giraud-Héraud, O. Guillaudin, W. Holmes, W. C. Jones, J.-M. Lamarre, J. Macías-Pérez, M. Martinez, A. Miniussi, A. Monfardini, F. Pajot, G. Patanchon, A. Pelissier, M. Piat, J.-L. Puget, C. Renault, C. Rosset, D. Santos, A. Sauvé, L. D. Spencer, and R. Sudiwala. Impact of particles on the Planck HFI detectors: Ground-based measurements and physical interpretation. A&A, 569:A88, Sept. 2014. doi: 10.1051/0004-6361/201423868.

S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335, 1995. ISSN 00031305. URL `http://www.jstor.org/stable/2684568`.

A. G. Cohen and A. De Rujula. Scars on the CBR? *ArXiv Astrophysics e-prints*, Sept. 1997.

A. G. Cohen and D. B. Kaplan. Thermodynamic generation of the baryon asymmetry. *Physics Letters B*, 199(2):251 – 258, 1987. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/0370-2693(87)91369-4. URL `http://www.sciencedirect.com/science/article/pii/0370269387913694`.

A. G. Cohen, A. De Rújula, and S. L. Glashow. A Matter-Antimatter Universe? ApJ, 495: 539–549, Mar. 1998. doi: 10.1086/305328.

P. Coppi. How do we know antimatter is absent? In *this SLAC Summer Institute, available at `http://www-conf.slac.stanford.edu/ssi/2004/lec_notes/Coppi/default.htm`*, 2004.

B. P. Crill, P. A. R. Ade, D. R. Artusa, R. S. Bhatia, J. J. Bock, A. Boscaleri, P. Cardoni, S. E. Church, K. Coble, P. de Bernardis, G. de Troia, P. Farese, K. M. Ganga, M. Giacometti, C. V. Haynes, E. Hivon, V. V. Hristov, A. Iacoangeli, W. C. Jones, A. E. Lange, L. Martinis, S. Masi, P. V. Mason, P. D. Mauskopf, L. Miglio, T. Montroy, C. B. Netterfield, C. G. Paine, E. Pascale, F. Piacentini, G. Polenta, F. Pongetti, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna, and A. D. Turner. BOOMERANG: A Balloon-borne Millimeter-Wave Telescope and Total Power Receiver for Mapping Anisotropy in the Cosmic Microwave Background. ApJS, 148:527–541, Oct. 2003. doi: 10.1086/376894.

P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, G. De Gasperis, G. De Troia, P. C. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. V. Hristov, A. Iacoangeli, A. H. Jaffe, A. E. Lange, L. Martinis, S. Masi, P. Mason, P. D. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, F. Pongetti, S. Prunet, S. Rao, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna, and N. Vittorio. Detection of anisotropy in the Cosmic Microwave Background at horizon and sub-horizon scales with the BOOMERanG experiment. *ArXiv Astrophysics e-prints*, Nov. 2000.

J. Delabrouille, K. M. Gorski, and E. Hivon. Circular scans for cosmic microwave background anisotropy observation and analysis. MNRAS, 298:445–450, Aug. 1998. doi: 10.1046/j.1365-8711.1998.01646.x.

B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 01 1979. doi: 10.1214/aos/1176344552. URL `http://dx.doi.org/10.1214/aos/1176344552`.

A. Fruchter and R. N. Hook. Novel image reconstruction method applied to deep Hubble space telescope images. In A. G. Tescher, editor, *Applications of Digital Image Processing XX*, volume 3164 of Proc. SPIE, pages 120–125, Oct. 1997.

S. Gaertner, A. Benoît, M. Piat, M. Giard, J. M. Lamarre, R. Pons, and G. Serra. A New Readout Electronic for the Planck Surveyor Bolometric Instrument. In A. Wilson, editor, *The Far Infrared and Submillimetre Universe.*, volume 401 of *ESA Special Publication*, page 369, Aug. 1997.

K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ, 622:759–771, Apr. 2005. doi: 10.1086/427976.

L. Grego, J. E. Carlstrom, E. D. Reese, G. P. Holder, W. L. Holzapfel, M. K. Joy, J. J. Mohr, and S. Patel. Galaxy Cluster Gas Mass Fractions from Sunyaev-Zeldovich Effect Measurements: Constraints on Ω_M. ApJ, 552:2–14, May 2001. doi: 10.1086/320443.

A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan. 1981. doi: 10.1103/PhysRevD.23.347. URL `http://link.aps.org/doi/10.1103/PhysRevD.23.347`.

S. Hanany, P. Ade, A. Balbi, J. Bock, J. Borrill, A. Boscaleri, P. de Bernardis, P. G. Ferreira, V. V. Hristov, A. H. Jaffe, A. E. Lange, A. T. Lee, P. D. Mauskopf, C. B. Netterfield, S. Oh, E. Pascale, B. Rabii, P. L. Richards, G. F. Smoot, R. Stompor, C. D. Winant, and J. H. P. Wu. MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10 arcminutes to 5 degrees. ApJ, 545:L5–L9, Dec. 2000. doi: 10.1086/317322.

C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 MHz all-sky continuum survey. II - The atlas of contour maps. A&AS, 47:1, Jan. 1982.

W. A. Holmes, J. J. Bock, B. P. Crill, T. C. Koch, W. C. Jones, A. E. Lange, and C. G. Paine. Initial test results on bolometers for the Planck high frequency instrument. Appl. Opt., 47:5996–6008, Nov. 2008. doi: 10.1364/AO.47.005996.

G. Hurier et al. MILCA, a modified internal linear combination algorithm to extract astrophysical emissions from multifrequency sky maps. A&A, 558:A118, Oct. 2013. doi: 10.1051/0004-6361/201321891.

W. H. Kinney, E. W. Kolb, and M. S. Turner. Ribbons on the Cosmic Background Radiation Sky: A Powerful Test of a Baryon Symmetric Universe. *Physical Review Letters*, 79:2620–2623, Oct. 1997. doi: 10.1103/PhysRevLett.79.2620.

E. Komatsu and U. Seljak. The Sunyaev-Zel'dovich angular power spectrum as a probe of cosmological parameters. MNRAS, 336:1256–1270, Nov. 2002. doi: 10.1046/j.1365-8711.2002.05889.x.

L. M. Krauss, S. Dodelson, and S. Meyer. Primordial Gravitational Waves and Cosmology. *Science*, 328:989, May 2010. doi: 10.1126/science.1179541.

J. Lamarre, J. Puget, F. Bouchet, P. Ade, A. Benoit, J. Bernard, J. Bock, P. D. Bernardis, J. Charra, F. Couchot, J. Delabrouille, G. Efstathiou, M. Giard, G. Guyot, A. Lange, B. Maffei, A. Murphy, F. Pajot, M. Piat, I. Ristorcelli, D. Santos, R. Sudiwala, J. Sygnet, J. Torre, V. Yurchenko, and D. Yvon. The planck high frequency instrument, a third generation {CMB} experiment, and a full sky submillimeter survey. *New Astronomy Reviews*, 47(11 - 12):1017 – 1024, 2003. ISSN 1387-6473. doi: http://dx.doi.org/10.1016/j.newar.2003.09.006. URL `http://www.sciencedirect.com/science/article/pii/S1387647303002239`. Proceedings of the Workshop on The Cosmic Microwave Background Radiation and its Polarization.

J. Lamarre, J. Puget, P. A. R. Ade, F. Bouchet, G. Guyot, A. E. Lange, F. Pajot, A. Arondel, K. Benabed, J. Beney, A. Benoît, J. Bernard, R. Bhatia, Y. Blanc, J. J. Bock, E. Bréelle, T. W. Bradshaw, P. Camus, A. Catalano, J. Charra, M. Charra, S. E. Church, F. Couchot, A. Coulais, B. P. Crill, M. R. Crook, K. Dassas, P. de Bernardis, J. Delabrouille, P. de Marcillac, J. Delouis, F. Désert, C. Dumesnil, X. Dupac, G. Efstathiou, P. Eng, C. Evesque, J. Fourmond, K. Ganga, M. Giard, R. Gispert, L. Guglielmi, J. Haissinski, S. Henrot-Versillé, E. Hivon, W. A. Holmes, W. C. Jones, T. C. Koch, H. Lagardère, P. Lami, J. Landé, B. Leriche, C. Leroy, Y. Longval, J. F. Macías-Pérez, T. Maciaszek, B. Maffei, B. Mansoux, C. Marty, S. Masi, C. Mercier, M. Miville-Deschênes, A. Moneti, L. Montier, J. A. Murphy, J. Narbonne, M. Nexon, C. G. Paine, J. Pahn, O. Perdereau, F. Piacentini, M. Piat, S. Plaszczynski, E. Pointecouteau, R. Pons, N. Ponthieu, S. Prunet, D. Rambaud, G. Recouvreur, C. Renault, I. Ristorcelli, C. Rosset, D. Santos, G. Savini, G. Serra, P. Stassi, R. V. Sudiwala, J. Sygnet, J. A. Tauber, J. Torre, M. Tristram, L. Vibert, A. Woodcraft, V. Yurchenko, and D. Yvon. *Planck* pre-launch status: The HFI instrument, from specification to actual performance. A&A, 520:A9, Sept. 2010. doi: 10.1051/0004-6361/200912975.

A. Lazarian. Tracing magnetic fields with aligned grains. J. Quant. Spec. Radiat. Transf., 106:225–256, July 2007. doi: 10.1016/j.jqsrt.2007.01.038.

B. Leistedt, J. D. McEwen, P. Vandergheynst, and Y. Wiaux. S2LET: A code to perform fast wavelet analysis on the sphere. A&A, 558:A128, Oct. 2013. doi: 10.1051/0004-6361/201220729.

J. F. Macías-Pérez, G. Lagache, B. Maffei, K. Ganga, A. Bourrachot, P. Ade, A. Amblard, R. Ansari, E. Aubourg, J. Aumont, S. Bargot, J. Bartlett, A. Benoît, J.-P. Bernard, R. Bhatia, A. Blanchard, J. J. Bock, A. Boscaleri, F. R. Bouchet, P. Camus, J.-F. Cardoso, F. Couchot, P. de Bernardis, J. Delabrouille, F.-X. Désert, O. Doré, M. Douspis, L. Dumoulin, X. Dupac, P. Filliatre, P. Fosalba, F. Gannaway, B. Gautier, M. Giard, Y. Giraud-Héraud, R. Gispert, L. Guglielmi, J.-C. Hamilton, S. Hanany, S. Henrot-Versillé, V. Hristov, J. Kaplan, J.-M. Lamarre, A. E. Lange, K. Madet, C. Magneville, D. P. Marrone, S. Masi, F. Mayet, J. A. Murphy, F. Naraghi, F. Nati, G. Patanchon, O. Perdereau, G. Perrin, S. Plaszczynski, M. Piat, N. Ponthieu, S. Prunet, J.-L. Puget, C. Renault, C. Rosset, D. Santos, A. Starobinsky, I. Strukov, R. V. Sudiwala, R. Teyssier, M. Tristram, C. Tucker, J.-C. Vanel, D. Vibert, E. Wakui, and D. Yvon. Archeops in-flight performance, data processing, and map making. A&A, 467:1313–1344, June 2007. doi: 10.1051/0004-6361:20065258.

J. C. Mather. Bolometer noise: nonequilibrium theory. *Appl. Opt.*, 21(6):1125–1129, Mar. 1982. doi: 10.1364/AO.21.001125. URL `http://ao.osa.org/abstract.cfm?URI=ao-21-6-1125`.

J. C. Mather, E. S. Cheng, D. A. Cottingham, R. E. Eplee, Jr., D. J. Fixsen, T. Hewagama, R. B. Isaacman, K. A. Jensen, S. S. Meyer, P. D. Noerdlinger, S. M. Read, L. P. Rosen, R. A. Shafer, E. L. Wright, C. L. Bennett, N. W. Boggess, M. G. Hauser, T. Kelsall, S. H. Moseley, Jr., R. F. Silverberg, G. F. Smoot, R. Weiss, and D. T. Wilkinson. Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument. ApJ, 420:439–444, Jan. 1994. doi: 10.1086/173574.

L. Michaeli, P. Michalko, and J. Šaliga. Unified adc nonlinearity error model for sar adc. *Measurement*, 41(2):198–204, 2008. doi: 10.1016/j.measurement.2006.10.004.

A. Mohammadzadeh, H. Evans, P. Nieminen, E. Daly, P. Vuilleumier, P. Buhler, C. Eggel, W. Hajdas, N. Schlumpf, A. Zehnder, J. Schneider, and R. Fear. The esa standard radiation environment monitor program first results from PROBA-I and INTEGRAL. *IEEE Transactions on Nuclear Science*, 50:2272–2277, Dec. 2003. doi: 10.1109/TNS.2003.821796.

P. D. Naselsky and L.-Y. Chiang. Antimatter from cosmological baryogenesis and the anisotropies and polarization of CMB radiation. Phys. Rev. D, 69(12):123518, June 2004. doi: 10.1103/PhysRevD.69.123518.

P. L. Nolan, A. A. Abdo, M. Ackermann, M. Ajello, A. Allafort, E. Antolini, W. B. Atwood, M. Axelsson, L. Baldini, J. Ballet, and et al. Fermi Large Area Telescope Second Source Catalog. ApJS, 199:31, Apr. 2012. doi: 10.1088/0067-0049/199/2/31.

F. Pajot, P. A. R. Ade, J. Beney, E. Bréelle, D. Broszkiewicz, P. Camus, C. Carabétian, A. Catalano, A. Chardin, M. Charra, J. Charra, R. Cizeron, F. Couchot, A. Coulais, B. P. Crill, K. Dassas, J. Daubin, P. de Bernardis, P. de Marcillac, J. Delouis, F. Désert, P. Duret, P. Eng, C. Evesque, J. Fourmond, S. François, M. Giard, Y. Giraud-Héraud, L. Guglielmi, G. Guyot, J. Haissinski, S. Henrot-Versillé, V. Hervier, W. Holmes, W. C. Jones, J. Lamarre, P. Lami, A. E. Lange, M. Lefebvre, B. Leriche, C. Leroy, J. Macias-Perez, T. Maciaszek, B. Maffei, A. Mahendran, B. Mansoux, C. Marty, S. Masi, C. Mercier, M. Miville-Deschenes, L. Montier, C. Nicolas, F. Noviello, O. Perdereau, F. Piacentini, M. Piat, S. Plaszczynski, E. Pointecouteau, R. Pons, N. Ponthieu, J. Puget, D. Rambaud, C. Renault, J. Renault, C. Rioux, I. Ristorcelli, C. Rosset, G. Savini, R. Sudiwala, J. Torre, M. Tristram, D. Vallée, M. Veneziani, and D. Yvon. *Planck* pre-launch status: HFI ground calibration. A&A, 520:A10, Sept. 2010. doi: 10.1051/0004-6361/200913203.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142:419–421, July 1965. doi: 10.1086/148307.

Planck HFI Core Team. *Planck* early results. VI. The High Frequency Instrument data processing. A&A, 536:A6, 2011. doi: 10.1051/0004-6361/201116462.

Planck Collaboration II. *Planck* early results. II. The thermal performance of *Planck*. A&A, 536:A2, 2011. doi: 10.1051/0004-6361/201116486.

Planck Collaboration VI. *Planck* 2013 results. VI. High Frequency Instrument data processing. A&A, 571:A6, 2014. doi: 10.1051/0004-6361/201321570.

Planck Collaboration VII. *Planck* 2013 results. VII. HFI time response and beams. A&A, 571:A7, 2014. doi: 10.1051/0004-6361/201321535.

Planck Collaboration VIII. *Planck* 2013 results. VIII. HFI photometric calibration and mapmaking. A&A, 571:A8, 2014. doi: 10.1051/0004-6361/201321538.

Planck Collaboration X. *Planck* 2013 results. X. HFI energetic particle effects: characterization, removal, and simulation. A&A, 571:A10, 2014. doi: 10.1051/0004-6361/201321577.

Planck Collaboration XII. *Planck* 2013 results. XII. Diffuse component separation. A&A, 571:A12, 2014. doi: 10.1051/0004-6361/201321580.

Planck Collaboration XX. *Planck* 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts. A&A, 571:A20, 2014. doi: 10.1051/0004-6361/201321521.

Planck Collaboration I. *Planck* 2015 results. I. Overview of products and results. A&A*, in press*, 2016.

Planck Collaboration VII. *Planck* 2015 results. VII. High Frequency Instrument data processing: Time-ordered information and beam processing. A&A*, in press*, 2016.

Planck Collaboration VIII. *Planck* 2015 results. VIII. High Frequency Instrument data processing: Calibration and maps. A&A*, in press*, 2016.

Planck Collaboration IX. *Planck* 2015 results. IX. Diffuse component separation: CMB maps. A&A*, in press*, 2016.

Planck Collaboration X. *Planck* 2015 results. X. Diffuse component separation: Foreground maps. A&A*, in press*, 2016.

Planck Collaboration XIII. *Planck* 2015 results. XIII. Cosmological parameters. A&A*, in press*, 2016.

Planck Collaboration XV. *Planck* 2015 results. XV. Gravitational lensing. A&A*, in press*, 2016.

Planck Collaboration XXII. *Planck* 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect. A&A*, in press*, 2016.

Planck Collaboration XXIV. *Planck* 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. A&A*, in press*, 2016.

Planck Collaboration XXVII. *Planck* 2015 results. XXVII. The Second Planck Catalogue of Sunyaev-Zeldovich Sources. A&A*, in press*, 2016.

Planck Collaboration Int. XIX. *Planck* intermediate results. XIX. An overview of the polarized thermal emission from Galactic dust. A&A, 576:A104, 2015. doi: 10.1051/0004-6361/201424082.

Planck Collaboration Int. XXXVIII. *Planck* intermediate results. XXXVIII. *E*- and *B*-modes of dust polarization from the magnetized filamentary structure of the interstellar medium. A&A, 586:A141, 2016. doi: 10.1051/0004-6361/201526506.

Planck Collaboration Int. XLVI. *Planck* intermediate results. XLVI. Reduction of large-scale systematic effects in HFI polarization maps and estimation of the reionization optical depth. A&A, *submitted*, 2016.

Planck Collaboration Int. XLVII. *Planck* intermediate results. XLVII. Constraints on reionization history. A&A, *submitted*, 2016.

Qubic Collaboration, E. Battistelli, A. Baú, D. Bennett, L. Bergé, J.-P. Bernard, P. de Bernardis, G. Bordier, A. Bounab, É. Bréelle, et al. QUBIC: The QU bolometric interferometer for cosmology. *Astroparticle Physics*, 34:705–716, Apr. 2011. doi: 10.1016/j.astropartphys.2011.01.012.

M. Remazeilles et al. Cmb and sz effect separation with constrained internal linear combinations. *Monthly Notices of the Royal Astronomical Society*, 410(4):2481–2487, 2011. doi: 10.1111/j.1365-2966.2010.17624.x. URL `http://mnras.oxfordjournals.org/content/410/4/2481.abstract`.

A. Sauvé and L. Montier. Building the analytical response in frequency domain of ac biased bolometers. *Experimental Astronomy*, pages 1–26, 2016. ISSN 1572-9508. doi: 10.1007/s10686-016-9514-1. URL `http://dx.doi.org/10.1007/s10686-016-9514-1`.

A. Sauvé, F. Couchot, G. Patanchon, and L. Montier. Inflight characterization and correction of planck/hfi analog to digital converter nonlinearity. volume 9914, pages 99143E–99143E–17, 2016. doi: 10.1117/12.2232671. URL `http://dx.doi.org/10.1117/12.2232671`.

S. Scull, B. Jones, T. Bradshaw, A. Orlowksa, and C. Jewell. *Design and Development of a 4K Mechanical Cooler*, pages 513–519. Springer US, Boston, MA, 2002. ISBN 978-0-306-47090-5. doi: 10.1007/0-306-47090-X_61. URL `http://dx.doi.org/10.1007/0-306-47090-X_61`.

G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss, and D. T. Wilkinson. Structure in the COBE differential microwave radiometer first-year maps. ApJ, 396:L1–L5, Sept. 1992. doi: 10.1086/186504.

D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. ApJS, 170:377–408, June 2007. doi: 10.1086/513700.

R. A. Sunyaev and I. B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. ARA&A, 18:537–560, 1980a. doi: 10.1146/annurev.aa.18.090180.002541.

R. A. Sunyaev and I. B. Zeldovich. The velocity of clusters of galaxies relative to the microwave background - The possibility of its measurement. MNRAS, 190:413–420, Feb. 1980b. doi: 10.1093/mnras/190.3.413.

J. A. Tauber, H. U. Norgaard-Nielsen, P. A. R. Ade, J. Amiri Parian, T. Banos, M. Bersanelli, C. Burigana, A. Chamballu, D. de Chambure, P. R. Christensen, O. Corre, A. Cozzani, B. Crill, G. Crone, O. D'Arcangelo, R. Daddato, D. Doyle, D. Dubruel, G. Forma, R. Hills, K. Huffenberger, A. H. Jaffe, N. Jessen, P. Kletzkine, J. M. Lamarre, J. P. Leahy, Y. Longval, P. de Maagt, B. Maffei, N. Mandolesi, J. Martí-Canales, A. Martín-Polegre, P. Martin, L. Mendes, J. A. Murphy, P. Nielsen, F. Noviello, M. Paquay, T. Peacocke, N. Ponthieu, K. Pontoppidan, I. Ristorcelli, J. Riti, L. Rolo, C. Rosset, M. Sandri, G. Savini, R. Sudiwala, M. Tristram, L. Valenziano, M. van der Vorst, K. van't Klooster, F. Villa, and V. Yurchenko. *Planck* pre-launch status: The optical system. A&A, 520:A2, Sept. 2010. doi: 10.1051/0004-6361/200912911.

R. J. Thornton, P. A. R. Ade, S. Aiola, F. E. Angile, M. Amiri, J. A. Beall, D. T. Becker, H. Cho, S. K. Choi, P. Corlies, K. P. Coughlin, R. Datta, et al. The Atacama Cosmology Telescope: The polarization-sensitive ACTPol instrument. *ArXiv e-prints*, May 2016.

M. Tristram, C. Filliard, O. Perdereau, S. Plaszczynski, R. Stompor, and F. Touze. Iterative destriping and photometric calibration for Planck-HFI, polarized, multi-detector mapmaking. A&A, 534:A88, Oct. 2011. doi: 10.1051/0004-6361/201116871.

J. W. Tukey. Bias and confidence in not-quite large samples (abstract). *The Annals of Mathematical Statistics*, 29(2):614–614, June 1958. ISSN 0003-4851. doi: http://dx.doi.org/10.1214/aoms/1177706647. URL `http://projecteuclid.org/euclid.aoms/1177706647`.

P. von Ballmoos. Antimatter in the universe: constraints from gamma-ray astronomy. *Hyperfine Interactions*, 228(1):91–100, 2014. ISSN 1572-9540. doi: 10.1007/s10751-014-1024-9. URL `http://dx.doi.org/10.1007/s10751-014-1024-9`.

G. Weidenspointner and M. Varendorff. *The Universe in Gamma Rays*, pages 339–366. Springer Berlin Heidelberg, 2001. ISBN 978-3-642-08745-5. doi: 10.1007/978-3-662-04593-0. URL `http://dx.doi.org/10.1007/978-3-662-04593-0`.

Y. Wiaux, J. D. McEwen, P. Vandergheynst, and O. Blanc. Exact reconstruction with directional wavelets on the sphere. MNRAS, 388:770–788, Aug. 2008. doi: 10.1111/j.1365-2966.2008.13448.x.

H. Yan and A. Lazarian. Polarization of Absorption Lines as a Diagnostics of Circumstellar, Interstellar, and Intergalactic Magnetic Fields: Fine-Structure Atoms. ApJ, 653:1292–1313, Dec. 2006. doi: 10.1086/508704.

M. Zaldarriaga and U. c. v. Seljak. All-sky analysis of polarization in the microwave background. *Phys. Rev. D*, 55:1830–1840, Feb 1997. doi: 10.1103/PhysRevD.55.1830. URL `http://link.aps.org/doi/10.1103/PhysRevD.55.1830`.

Résumé

Université Paul Sabatier
Sciences de l'Univers de l'Environnement et de l'Espace

Docteur en philosophie

Caractérisation des effets systématiques de l'instrument *Planck/HFI*, propagation et impact sur les données scientifiques

par Alexandre Sauvé

Planck est un satellite de l'ESA lancé en 2009, il représente la troisième génération d'observatoires spatiaux dans l'ère de la cosmologie de précision. Sa mission était de cartographier le fond diffus cosmologique (CMB) avec une extrême précision ($\Delta T/T < 10^{-5}$). Le niveau de précision requis nécessite un niveau de contrôle extrêmement élevé des effets systématiques introduits par les instruments embarqués. Il se trouve qu'une combinaison inattendue d'éléments a conduit à une amplification de l'impact des effets de non linéarité introduits par le composant de numérisation des données scientifiques de l'instrument *HFI*, conduisant à l'introduction d'un des effets systématiques les plus difficiles à maîtriser. Ce manuscrit présente le travail qui a conduit à la caractérisation et à la correction de cette non-linéarité.

Tout d'abord la modélisation de la réponse thermique complexe des détecteurs bolométriques sous intensité modulée est présentée. Puis la caractérisation de la réponse des détecteurs et de l'électronique de lecture est réalisée via l'utilisation du signal produit par l'impact des particules cosmiques sur les détecteurs. Dans un deuxième temps, les étapes menant à la caractérisation de précision des composants de numérisation sont détaillée. Pour pouvoir corriger l'effet de non-linéarité sur les données scientifiques, la chaîne complète de l'électronique de lecture est modélisée en prenant en compte la réponse des détecteurs sous intensité modulée et les effets non-linéaires du composant de numérisation avec le bruit. En plus de cela, il a fallu tenir compte du signal parasite complexe généré par le compresseur de l'étage de refroidissement cryogénique à 4 K, son inclusion dans la correction est détaillée. Ceci a mené à une correction qui a réduit l'impact de cet effet d'un ordre de grandeur sur le catalogue des données Planck de 2015. Finalement une étude est menée pour mesurer l'impact de cet effet de non-linéarité sur l'analyse scientifique de la Galaxie en polarisation et sur la cosmologie à travers le spectre de puissance angulaire du CMB. Une étude préliminaire est menée pour la détection de l'annihilation de nuages d'antimatière survivant à l'époque de la recombinaison en utilisant la carte d'effet Sunyaev-Zeldovich de *Planck*, et comment cette détection est affectée par la stratégie de scan de *Planck*.