From the observation of UHECR signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments

Antony Escudie

UNIVERSITE MATIERE BRETAGNE MOLECULES LOIRE ET MATERIAUX

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE MINES-TÉLÉCOM ATLANTIQUE
BRETAGNE PAYS DE LA LOIRE - IMT ATLANTIQUE
COMUE UNIVERSITÉ BRETAGNE LOIRE

ÉCOLE DOCTORALE N°596
Matière, Molécules, Matériaux
Dicipline: *Physique*
Spécialité : *Physique Subatomique et Instrumentation Nucléaire*

Par

Antony ESCUDIE

From the observation of UHECR radio signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments

Thèse présentée et soutenue à Nantes, le 27 Septembre 2019
Unité de recherche : SUBATECH UMR 6457
Thèse N° : 2019IMTA0145

Rapporteurs avant soutenance :

M. Jacob LAMBLIN, MCF Université Grenoble Alpes, LPSC (France)
M. Stéphane CORBEL, Professeur Université de Paris, OBSPM (France)

Composition du Jury :

Président :	**M. Ginés MARTINEZ**	Directeur de Recherche CNRS, SUBATECH (France)
Examinateurs :	**M. Stéphane CORBEL**	Professeur Université de Paris, OBSPM (France)
	M. Krijn DE VRIES	Assistant Professor VUB, IIHE (Belgique)
	M. Jacob LAMBLIN	MCF Université Grenoble Alpes, LPSC (France)
Directeur de thèse :	**M. Benoît REVENU**	Directeur de Recherche CNRS, SUBATECH (France)
Co-encadrant de thèse :	**M. Richard DALLIER**	Maître assistant IMT Atlantique, SUBATECH (France)

Invité :

Co-encadrant de thèse :	**M. Lilian MARTIN**	Chargé de Recherche CNRS, SUBATECH (France)

From the observation of UHECR radio signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments

Antony Escudie

SUBATECH laboratory, CNRS-IN2P3/IMT Atlantique Nantes/Université de Nantes

From the observation of UHECR radio signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments

Le 19 juin 2018, le groupe Astro s'est rendu à l'observatoire de radio-astronomie de Nançay pour une mission de maintenance de ses détecteurs. Il en a profité pour prendre des photos du ciel en pose longue. Sur la photo présentée ci-dessous : au premier plan, une station autonome (antenne Papillon) de l'expérience CODALEMA, suivie par les antennes du radio-télescope NenuFAR, puis au fond à gauche, le miroir primaire du radio-télescope de l'observatoire. L'éclairage de ces objets est complètement naturel venant de la Lune située dans le dos du photogaphe. Au niveau du ciel, au centre de la photo, le plan de notre galaxie, la Voie Lactée.

*©J-B. Jarnoux, T. Lipreau et A. Escudie (NikonD3300 + AF*18 – 55*mm*, *F* /3.5 – 5.6, *pose longue f* /3.5 *25.0s ISO1600* 18.00*mm)*

À Charline.
À ma famille.

"There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after."

J.R.R. Tolkien, The Hobbit

This work was funded by the Région Pays de la Loire, and supported by the Programme National des Hautes Energies of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES.

Acknowledgement

Je commence par remercier le directeur de Subatech, Ginés Martinez, qui m'a permis d'évoluer au sein du laboratoire et d'y développer mes travaux de thèse. Merci d'avoir accepté de faire partie de mon jury de thèse. J'en profite pour remercier l'ensemble des membres de mon jury, Stéphane Corbel, Krijn De Vries, Jacob Lamblin et Ginés Martinez d'avoir accepté d'évaluer mes travaux. Merci particulièrement à Jacob pour nos échanges qui auront permis, j'en suis persuadé, d'améliorer la qualité du manuscrit.

Je remercie mon directeur de thèse, Benoît Revenu. Merci de m'avoir fait connaître ce domaine dès ma première année de master. J'ai pu grâce à toi évoluer au sein de l'équipe Astro pendant près de 3 ans et demi, et y développer mes travaux de thèse dans un environnement scientifique stimulant.

Je remercie mes encadrants scientifiques, Richard Dallier et Lilian Martin. Messieurs, je pense que 328 pages ne seraient pas suffisantes pour vous dire tout ce que j'ai à vous dire (pourtant, entre mes mains, ça parait gros, 328 pages !). Richard, avoir une personne telle que toi dans son encadrement de thèse donne un avantage majeur, et une probabilité de réussite supérieure à 1 (si si, c'est possible). Tu as été d'un soutien moral inconditionnel, qui m'a été précieux sur cette fin de thèse. Merci pour ta disponibilité, ton altruisme et merci de m'avoir accordé ta confiance dès le début de cette aventure.
S'il y a bien un choix que je réitérerais sans hésiter si l'occasion se représentait, c'est celui que j'ai fait en arrivant en stage de master 2. Avec le recul que je peux avoir aujourd'hui, c'est bien ce choix là qui a conditionné toute la suite de l'aventure, et, je crois, sa réussite. Matlab, C++ ou Python ? Néophyte, je choisissais Python, ça sonnait bien. Cela m'a amené à développer une relation thésard-débugueur de code avec Lilian. Relation qui a vite évolué en thésard-conseiller personnel d'orientation de thèse, puis en thésard-encadrant. Tu as été mon maître à penser, mon mentor. Tu me disais souvent "C'est ta thèse, ce n'est qu'une suggestion, tu fais ce que tu veux !". Tu m'as guidé, toujours, dans la bonne direction, avec un objectif bien défini: l'indépendance intellectuelle. Lilian, un grand merci pour ton humanité, tes conseils avisés, ta modestie et ton humour qui m'ont accompagnés tout au long de cette thèse.

J'ai une pensée pour les personnes du groupe Astro avec qui j'ai pu collaborer. Merci à Daniel, ces deux années et demie passées en ta compagnie ont été très enrichissantes. Merci à Didier Charrier et à Florian Gaté pour nos échanges. Merci également aux personnes de Subatech et de l'observatoire de Nançay avec qui j'ai pu échanger avec grand plaisir, de manière formelle ou informelle, de près ou de loin, durant ces trois années : Jean-Luc Beney, Hervé Carduner, Laurent Denis, Marie Germain, Arnaud Guertin, Alain Lecacheux, Frédéric Lefèvre, Rémi Maurice, Louis-Marie Rigallau, Nicolas Thiollière, Frédéric Yiermia ainsi que Noel Servagent, Stéphane Martinez et Dominique Thers avec qui j'ai pu prendre plaisir à enseigner. Merci aux enseignants-chercheurs que j'ai pu croiser pendant mon cursus universitaire, je pense notamment à Virginia De La Mota, Pol-B Gaussiaux, Thierry Gousset, Ferid Haddad, Guy Royer, Taklit Sami...

Je tenais également à remercier les personnes des services techniques et administratifs qui m'ont apporté leur aide lorsque j'en avais besoin, avec une pensée particulière pour

Tanja, Pauline, Mérryl et Séverine !

Évidemment, je me tourne maintenant vers mes camarades de la Section Locale des Doctorants Subatomiques Syndiqués, qui ont supporté mes blagues pourries (entre autres choses...) lors de nos pauses méridiennes quotidiennes. Merci à Chloé, David F, David H aka *il est*, Erwann, Étienne, Florian C, Gabriel, Ophélie, Manu, Quentin, Stéphane. Votre bonne humeur contagieuse me manquera. J'espère sincèrement que l'on restera en contact. Erwann, 6 ans maintenant qu'on se côtoie tous les jours, ces trois années de doctorat en ta compagnie ont été un vrai plaisir. Nos rendez-vous "café" d'avant 8 heures ont eu un impact fort sur mon moral au quotidien. Merci pour ton sens de l'écoute et tes conseils avisés. J'adresse mes salutations aux personnes avec qui j'ai partagé le bureau H125, AdadaH, Fanny, Florian, Grégoire, Loïc et à la nouvelle recrue du mercato estival, Thomas. Bon courage à la relève, Johannès, Maxime, Victor..., l'aventure ne fait que commencer. Mes amitiés à ceux que j'ai croisés, Kevin Micheneau, Alexandre Parent du Châtelet, Flavien Ralite, Alexandre Subercaze, Loïck Virone...

Un grand merci à mes parents, ma sœur, mon beau-frère et mes amis pour leur soutien, tant pendant mon cursus universitaire que pendant ces trois années de thèse. Merci à tous, famille, belle-famille et amis d'être venus en force lors de la soutenance. Bien entendu, merci à mes parents et à ma sœur d'avoir pris en charge l'organisation du pot de thèse, ce qui m'a permis de rester concentrer sur mon affaire !

Ah ! Les gars, merci, merci, merci et encore merci ! Merci à mes deux colocataires qui ont accepté que je partage un bout d'appartement avec eux. Merci d'avoir été présents lors de mes petites baisses de régime sur la fin du marathon. Merci à Jumpy et Nakao, mes deux p'tits chats d'amour !

Enfin, sans ta présence à mes côtés, rien de tout cela n'aurait été possible, et cela n'aurait eu aucun sens. Merci d'avoir toujours été là, contre vents et marées, quoiqu'il arrive. Si toutes les personnes citées précédemment ont joué un rôle dans la construction de mes travaux de thèse, tu en as été les fondations et l'ingrédient qui ont fait que le tout soit plus solide qu'un roc. Ta force et ta positivité ont été et resteront une source d'inspiration. Merci à ma compagne, Charline.

Les rayons cosmiques d'ultra haute énergie pour résoudre les grandes questions de l'Univers

Découverts il y a maintenant plus de cent ans, les rayons cosmiques sont produits par des phénomènes cataclysmiques galactiques, mais aussi venus des confins de l'Univers. Quelles sont leur origines ? Comment sont-ils créés ? Comment sont-ils accélérés ? Autant de questions restant aujourd'hui encore sans réponse, bien que des études récentes tentent de relever ces défis.

Rayons cosmiques, messagers du cosmos

Malgré la découverte des rayons cosmiques par Victor Hess en 1912, il y a donc maintenant plus de cent ans, de nombreuses questions restent aujourd'hui encore sans réponse : que sont les rayons cosmiques, comment sont-ils créés et d'où viennent-ils ? A ce jour, on sait que la Terre est constamment bombardée par ces rayons cosmiques, qui sont des particules chargées d'origine extraterrestre. Les rayons cosmiques sont produits et accélérés par des sources astrophysiques. Ces sources doivent avoir des caractéristiques particulières pour pouvoir produire et accélérer ces particules à de telles énergies, répondant au critère de Hillas qui fait intervenir la taille et le champ magnétique de la source. Les principaux candidats à la création/accélération des rayons cosmiques sont les supernovæ de type II (dernière étape de vie d'une étoile associée à une explosion finale gigantesque), les pulsars qui sont des étoiles à neutrons en rotation rapide et très magnétisées émettant périodiquement un signal électromagnétique, les sursauts gamma (brefs sursauts de rayons gamma de quelques MeV , qui peuvent être produits lors de la coalescence de systèmes binaires d'étoiles à neutrons) et les noyaux actifs de galaxie (trou noir de quelques 10^6 à 10^{10} masses solaires éjectant un faisceau de matière à 99 % de la vitesse de la lumière).

Le flux de rayons cosmiques diminue considérablement avec leur énergie et s'étend sur ~32 ordres de grandeur, pour 12 ordres de grandeur en énergie. Ce comportement remarquablement régulier, unique en physique, pourrait signer un mécanisme de production universel de ces rayons cosmiques. C'est en cela que leur étude est importante, car l'énigme à résoudre se trouve dans ce spectre, et dans sa remarquable régularité. En dessous de 10^{14} eV, il reste possible de les détecter directement grâce à des satellites dédiés tels que l'expérience AMS de la Station Spatiale Internationale ou des expériences en ballon telles que celles réalisées par Victor Hess notamment. Au-delà d'une énergie de 10^{14} eV, il est quasiment impossible de détecter les rayons cosmiques primaires de manière directe, du fait du faible flux arrivant sur Terre : 1 rayon cosmique par m^2 par an à $\sim 3 \times 10^{16}$ eV, tombant à 1 rayon cosmique par km^2 par siècle au-delà de 10^{20} eV.

L'étude des rayons cosmiques d'énergie supérieure à 10^{14} eV se fait de manière indirecte. En effet, lorsqu'un rayon cosmique arrive sur Terre, c'est-à-dire en haut de l'atmosphère, il interagit avec les atomes de celle-ci. Ces interactions peuvent avoir lieu à

des énergies bien plus importantes que celles atteintes en laboratoire. Il est commun de comparer l'énergie des rayons cosmiques les plus énergétiques aux 13 TeV dans le centre de masse en collisions proton-proton produits au Large Hadron Collider (LHC) au CERN : l'énergie atteinte par les rayons cosmiques est 800 fois plus importante dans ce cadre. Nous préférerons ici comparer l'énergie du faisceau de protons du LHC, qui est de 6.5 TeV, avec l'énergie incidente du rayon cosmique le plus énergétique jamais détecté, à savoir 3.2×10^{20} eV, qui est donc des ordres de grandeur plus importante! Ces 50 J correspondent à une balle de tennis allant à 160 km $\cdot$ h^{-1}. La circonférence du LHC est d'environ 27 km. En comparaison, avec le même champ magnétique, il nous faudrait alors construire un accélérateur de particules ayant la circonférence de l'orbite de Mercure, soit 10^7 km, pour atteindre ces énergies. Quoiqu'il en soit, ces énergies ne sont pas reproductibles sur Terre.
Lorsque le rayon cosmique pénètre dans l'atmosphère, la première interaction a lieu à une profondeur d'atmosphère notée X_1 (qui est le *point de première interaction*), engendrant une cascade de particules secondaires chargées communément appelée *gerbe de particules*. Le nombre de particules dans la gerbe va croître jusqu'à la profondeur d'atmosphère $X_{\max}$, puis va décroître jusqu'à l'extinction de la gerbe. Cette gerbe est détectable au sol, et nous fournit des informations sur le rayon cosmique primaire : c'est la détection indirecte. L'objectif est de remonter aux caractéristiques du rayon cosmique primaire ayant engendré la gerbe de particules, donc de déterminer sa direction d'arrivée, sa nature et son énergie. Le challenge expérimental réside dans la capacité à déterminer ces trois quantités avec un niveau de précision suffisant. Cette détection peut se faire durant le développement de la gerbe dans l'atmosphère, via la technique de la fluorescence qui consiste à collecter la lumière de fluorescence émise par la désexcitation de l'azote lors du passage des particules chargées de la gerbe, ou directement au sol, avec des détecteurs dédiés comptant les particules. Cependant, l'arrivée d'une gerbe au sol dépend de plusieurs paramètres, tels que l'énergie du rayon cosmique primaire ou encore son angle d'arrivée. À 10^{18} eV par exemple, on estime le flux des rayons cosmiques à seulement 0.2 particules par km^2 par jour. Aussi, la solution consiste à construire des détecteurs couvrant d'immenses surfaces, telle que l'Observatoire Pierre Auger, en Argentine, qui couvre $\sim$ 3000 km^2 (surface équivalente à l'air urbaine de Nantes et sa campagne).

Il est d'usage d'appeler rayons cosmiques d'ultra-haute énergie (RCUHE) ceux d'énergie supérieure à 10^{17} eV. Les objectifs de l'étude des RCUHE sont de pouvoir caractériser entièrement ces messagers du cosmos :

— Quelles sont les sources capables de créer ces particules primaires (noyaux actifs de galaxie, sursauts gamma) ?
— Quels sont les mécanismes d'accélération (accélération de Fermi par chocs relativistes) ?
— Que sont-ils (proton, noyaux de fer, éléments intermédiaires…) ?

A l'instar d'autres particules telles que les photons issus des collisions partons-partons du LHC qui permettent de sonder la matière [1], les anti-neutrinos qui permettent de sonder les réacteurs nucléaires [2, 3] ou encore les photons gamma issues de radionucléides pour l'imagerie médicale [4], les rayons cosmiques sont eux utilisés comme sondes de la physique des particules et de l'Univers. Une des applications récentes est notamment l'utilisation de muons cosmiques pour sonder la grande pyramide de Khéops [5].

Malgré l'avènement des grands observatoires de rayons cosmiques, tels que l'Observatoire Pierre Auger, qui permettent aujourd'hui à travers l'étude des grandes gerbes atmosphériques de reconstituer les caractéristiques des rayons cosmiques les plus énergétiques de l'Univers, les mécanismes d'accélération, les sources astrophysiques et la partie très énergétique du spectre restent mal connues. Etant des particules chargées, ils sont déviés par les champs magnétiques galactiques et extra-galactiques, rendant

difficile la corrélation entre les directions d'arrivée et les sources ponctuelles dans le ciel.

Afin de mieux contraindre les caractéristiques des rayons cosmiques et fournir des réponses aux questions encore ouvertes aujourd'hui, une grande statistique, c'est-à-dire un grand nombre d'événements détectés, est nécessaire. Il est donc important d'utiliser des méthodes de détection avec un cycle utile optimal. La radio-détection des gerbes atmosphériques peut répondre à ce problème, avec un cycle utile de $\sim$ 100 % (contre seulement $\sim$ 14 % pour la méthode de fluorescence). Développée d'abord dans les années 1960, abandonnée en raison de contraintes technologiques, puis entièrement revue à l'ère du numérique depuis 2002, cette méthode est basée sur le fait que, lors du développement de la gerbe dans l'atmosphère, les particules chargées en mouvement engendrent l'émission d'une impulsion de champ électrique très brève, détectable au sol avec des antennes radio dédiées, sur une large bande de fréquences.

Émission radio-électrique des grandes gerbes atmosphériques

L'émission radio-électrique produite lors du développement de la gerbe est cohérente et a deux origines principales : l'excès de charges négatives dans la gerbe et la variation de courant transverse induite par le champ magnétique terrestre. Ce dernier mécanisme, associé au champ géomagnétique, est appelé *mécanisme géomagnétique*. Les électrons et les positons secondaires présents dans la gerbes sont accélérés dans le champ magnétique terrestre, ce qui conduit à l'émission d'un champ électrique. Cet effet est lié à la force de Lorentz : $\overrightarrow{F} = q\overrightarrow{v} \times \overrightarrow{B}$ où q est la charge de la particule, $\overrightarrow{v}$ sa vitesse et $\overrightarrow{B}$ le vecteur de champ magnétique terrestre. Pour les particules se déplaçant le long de l'axe de la gerbe, les courants résultants seront en moyenne perpendiculaires à l'axe de la gerbe et sont appelés "courants transverses". Ces courants varient selon l'évolution de la gerbe : le nombre de particules secondaires croît pour atteindre un maximum à la profondeur d'atmosphère $X_{\max}$, puis diminue jusqu'à l'extinction de la gerbe. Cette variation temporelle du nombre de particules dans la gerbe conduit à une variation temporelle des courants transverses induisant l'émission d'un champ électrique. Sa polarisation est linéaire, alignée le long de $\overrightarrow{F}$. L'intensité du champ électrique dépend de la direction d'arrivée du rayon cosmique primaire, en particulier de l'angle α entre la direction d'arrivée de la gerbe et la direction du champ géomagnétique.

Au deuxième ordre, le mécanisme d'*excès de charges* négatives vient en complément de la contribution géomagnétique. À haute énergie, les processus de production de paires et de bremsstrahlung dominent. Au fur et à mesure que la gerbe se développe, l'énergie moyenne diminue et d'autres processus apparaissent, tels que la diffusion Compton, la production de rayons delta et l'annihilation des positons avec des électrons atmosphériques libres pré-existants. Cela entraîne un excès de charges négatives de 10 à 20 %. Les électrons d'ionisation sont contenus dans le front de la gerbe, tandis que les ions positifs, beaucoup plus lourds, restent en retrait du front. Au fur et à mesure que la gerbe se développe, la charge négative absolue présente dans la gerbe croît, atteint un maximum et diminue finalement lorsque la gerbe s'éteint. Il existe à nouveau une variation temporelle de cet excès de charge, ce qui entraîne l'émission d'un champ électrique. Contrairement au mécanisme géomagnétique, le vecteur de champ électrique est orienté radialement par rapport à l'axe de la gerbe et son orientation dépend de la position de l'observateur.

Le travail effectué au cours de cette thèse vise à poursuivre les études de ce signal radio résultant de la superposition de ces deux mécanismes et qui se présente sous la forme d'un champ électrique transitoire rapide, d'une durée de quelques dizaines de nanosecondes, pouvant être détecté par des antennes à large bande passante et des systèmes d'acquisition rapide, tels que ceux utilisés par les expériences CODALEMA et

EXTASIS. L'originalité de ce travail réside essentiellement dans l'étendue de la gamme de fréquence étudiée, de 1 à 200 MHz, pour la première fois, sur les mêmes événements détectés.

Des antennes radio pour observer le ciel

Depuis 2002, l'expérience CODALEMA, située au sein de l'Observatoire de radio-astronomie de Nançay, est l'une des expériences pionnières et promotrices de ce renouveau de la radio-détection des grandes gerbes atmosphériques, qui est aujourd'hui adoptée par une multitude d'expériences dans le monde. Couplée à une compréhension de plus en plus sophistiquée des processus impliqués dans l'émission radio permettant ainsi l'utilisation de codes de simulation performants, la radio-détection des gerbes atteint maintenant un niveau de maturité suffisant pour être compétitive avec les méthodes de détection plus traditionnelles. La radio-détection est habituellement effectuée dans une gamme en fréquences restreinte à [20-80] MHz (notée MF dans ce qui suit pour les fréquences moyennes) par des expériences telles que AERA, Tunka-Rex, TREND, Yakutsk ou LOFAR. Cette restriction est principalement due à la radio-diffusion artificielle à basse et moyenne fréquence (bandes AM, FM), ce qui a conduit à choisir des taux d'échantillonnage relativement bas ($\sim 200\ \mathrm{MS \cdot s^{-1}}$) mais suffisants pour échantillonner correctement les signaux sur 60 MHz. Cependant, CODALEMA fonctionne avec un taux d'échantillonnage de $1\ \mathrm{GS \cdot s^{-1}}$, ce qui permet d'étendre les observations jusqu'à la bande [110-200] MHz. La principale limite de la bande de fréquences est alors due à la largeur de bande de l'antenne utilisée, [20-200] MHz pour CODALEMA, appelée ci-après «Extended Medium Frequencies (EMF)». À ce jour, l'expérience CODALEMA est composée de :

— un réseau carré ($0.4 \times 0.4\ \mathrm{km^2}$) de 13 détecteurs de particules (scintillateurs, détecteur de surface),
— un ensemble de 57 stations autonomes, équipées de dipôles croisés, synchronisées par datation GPS, fonctionnant dans la bande EMF, et distribuées sur $1\ \mathrm{km^2}$,
— un réseau compact de 10 antennes à polarisation croisée, disposées en étoile de 150 m d'extension et dont l'acquisition du signal (dans la bande EMF) est déclenchée par les détecteurs de particules.

Le réseau de stations autonomes est auto-déclenché sur les signaux radio, ce qui signifie que chaque station est indépendante. Les transitoires provenant de gerbes ou de toute autre source (bruits, avions...) sont soit stockés sur un disque distant pour une analyse hors ligne, soit envoyés directement à une acquisition (DAQ) centrale capable de construire en ligne l'événement sur la base de plusieurs signaux de station, en respectant plusieurs critères de sélection. Une vérification croisée peut être effectuée hors ligne avec les événements détectés par les détecteurs de particules ou l'un des instruments déclenchés de manière externe par les détecteurs de particules (réseau compact ou antennes basse fréquence).

CODALEMA est aujourd'hui l'expérience hôte d'EXTASIS, un réseau de 7 antennes basse fréquence déclenchées par les détecteurs de particules, qui tire parti des infrastructures existantes. EXTASIS a pour objectif de ré-étudier la bande de fréquences [1-10] MHz, et en particulier d'étudier la contribution dite de "Mort subite", le champ électrique attendu et émis par le front de la gerbe lorsqu'il atteint le sol et que les particules y disparaissent brutalement.

L'émission radio galactique pour mettre tout le monde d'accord

Dans le passé, de nombreux désaccords ont été signalés sur les amplitudes du champ électrique détecté pour une énergie donnée, notamment en raison de la difficulté à étalonner les détecteurs radio alors utilisés. Une part du travail de cette thèse a été de proposer

une méthode de calibration des antennes radio de CODALEMA (stations autonomes et antennes du réseau compact). L'idée a été d'utiliser l'émission radio de la Galaxie comme source externe pour calibrer les antennes. Afin de mener à bien cette calibration, deux choses sont nécessaires : une description de l'antenne et une description de l'émission radio galactique.

Concernant le deuxième point, nous avons comparé trois modèles d'émission radio galactique. Nous avons montré que les modèles étaient en accord entre eux, nous permettant d'en choisir un sans risque de biaiser notre procédure. Notre choix s'est porté sur le Global Sky Model (GSM) de *de Oliveira-Costa et. al.* Nous avons ensuite couplé un modèle d'antenne simulé avec le logiciel NEC avec le modèle d'émission galactique. Nous avons calculé la contribution de chaque partie du ciel, en convoluant la carte du ciel obtenue avec GSM par le lobe de l'antenne en fonction du temps sidéral local (LST). Une simulation du signal mesuré à une heure LST donnée est obtenue en sommant chaque pixel de la carte du ciel convoluée de la réponse d'antenne. Le résultat obtenu est appelé *courbe de transit*, que l'on peut comparer aux données réelles. La calibration est faite par sous-bande de fréquences de largeur 10 MHz et par pas de 5 MHz.

Nous avons montré que, dans le cadre d'une calibration relative entre antennes (en opposition à une calibration absolue), le modèle d'émission galactique ou encore les conditions d'environnement proche de l'antenne (tels que la composition du sol) n'avaient pas d'impact sur les résultats finaux. Nous avons mis en exergue une variation présente sur certaines stations autonomes, que nous avons assimilée en première approximation à une variation saisonnière. Afin de prendre en compte son effet, nous avons décidé de calculer les coefficients de calibration mois par mois, et avons démontré que cela permettait en effet de tenir compte de cette variation et d'en corriger les effets. La simulation du signal obtenue est en accord avec les données détectées par les antennes. Ce constat est d'une importance capitale, car en plus de permettre de calibrer nos antennes, l'utilisation de l'émission radio galactique nous permet de valider le modèle d'antenne utilisé. Ce modèle d'antenne est au centre de la majeure partie des analyses de cette thèse, notamment celle consistant à convoluer les simulations du champ électrique de la gerbe par la réponse d'antenne et à les comparer avec les données réelles pour estimer les paramètres de la gerbe et donc du rayon cosmique primaire.

Avant la calibration, une disparité importante du niveau de signal existait entre les antennes, de l'ordre de 70 µV pour les stations autonomes et de l'ordre de 40 µV pour les antennes du réseau compact. Ces disparités sont largement diminuées après calibration, et sont de l'ordre de 20 µV pour les stations autonomes et de l'ordre de 5 µV pour les antennes du réseau compact. Ces disparités résiduelles sont très inférieures au RMS (moyenne quadratique) du bruit enregistré par les antennes lors d'un événement, ce qui permet une première estimation de la qualité de la calibration.

Nous avons dans un second temps vérifié et quantifié les effets de la calibration, et montré que la calibration est nécessaire pour reconstruire de manière plus précise les paramètres de la gerbe.

Événements cosmiques

La mesure au sol du champ électrique émis lors du développement des gerbes est rendue difficile par le bruit de fond. Dans [20 – 200] MHz, il s'agit d'un bruit de fond galactique prévisible (qui nous a permis de calibrer nos antennes) et d'un bruit de fond parasite dû à des émissions artificielles, des interférences ou encore des décharges électriques d'orages à l'échelle planétaire. Ce bruit de fond est l'une des limitations des détecteurs

autonomes : le déclenchement de l'acquisition sur du bruit entraîne un temps mort important, empêchant la détection sur un événement dit "rayon cosmique". Il est donc impératif de bien comprendre l'environnement de fond pour développer un cadre d'analyse capable de générer un lot d'événements radio le plus pur possible qui puisse être analysé ultérieurement. Lors de ce travail de thèse, différents algorithmes de rejet de bruit ont été explorés afin de développer et de définir un cadre d'analyse complet permettant de reconstruire les caractéristiques des rayons cosmiques primaires. Ils sont basés sur le temps de montée du signal, sur la compacité de l'événement et sur le comportement de la polarisation de l'événement. Nous avons montré que, bien qu'exploratoires, ces trois méthodes couplées permettent un taux de sélection/rejet intéressant (99.96 % des événements enregistrés sont rejetés à juste titre, autrement dit 99.96 % des événements enregistrés sont des événements fortuits ne provenant pas d'une gerbe initiée par un rayon cosmique).

Dans [1.7 – 3.7] MHz (la bande basse fréquence de l'expérience EXTASIS), le bruit de fond est la superposition de trois contributions dominantes : le bruit thermique minimal de l'atmosphère (en d'autres termes, sa luminosité liée à sa température) et les bruits dits atmosphériques, constitués de la contribution des décharges électriques d'orages à l'échelle planétaire, et de tout type de bruits générés par l'homme (émetteurs radio, transitoires anthropiques). Les conditions de bruit atmosphériques entravent sévèrement la détection à basse fréquence. L'une des contributions principales de cette thèse a été l'élaboration d'un cadre d'analyse dédié à la bande basse fréquence, le signal associé étant souvent noyé dans le bruit.

Informations issues des messages du cosmos

Nous avons montré qu'une connaissance avancée des instruments était nécessaire. En effet, pour revenir aux caractéristiques des rayons cosmiques primaires, à savoir l'énergie et la composition, des comparaisons systématiques entre données et simulations doivent être effectuées. La calibration de nos détecteurs radio ainsi que la convolution des simulations par la réponse d'antenne pour les comparer aux données nécessitent une description et une compréhension fines des détecteurs utilisés. L'une des contributions importantes de cette thèse a été la mise au point d'un cadre d'analyse automatique couplé aux algorithmes de calibration des antennes et de réjection du bruit mentionnés précédemment afin de pouvoir reconstruire toutes les propriétés de gerbes avec le signal radio, telles que la direction d'arrivée, l'énergie mais aussi la nature du rayon cosmique primaire. Nous sommes partis d'une méthode de comparaisons de simulations avec les données déjà utilisée au sein du groupe, et l'avons généralisée afin d'exploiter toutes les capacités des instruments de CODALEMA : découplage des polarisations EW et NS des antennes (permet de garder l'information de polarisation et de polarité, qui était perdue avec l'ancienne méthode utilisant la somme quadratique des deux polarisations) et utilisation de toute la plage en fréquence (permet notamment de mieux reconstruire les événements gerbes très inclinés, i.e avec un angle zénithal important). Nous avons démontré que la généralisation de la méthode permettait d'obtenir de meilleurs résultats sur les paramètres reconstruits des gerbes, les incertitudes sur les paramètres obtenus avec la nouvelle méthode devenant comparables à celles obtenues par la technique de fluorescence, par exemple.

Concernant EXTASIS, nous avons observé 25 événements basse fréquence détectés en coïncidence avec les instruments de CODALEMA et nous avons estimé un seuil de détection de $23 \pm 4\ \mu V \cdot m^{-1}$ à partir de comparaisons avec des simulations. Nous avons aussi mis en évidence une forte corrélation entre l'observation de signaux à basse fréquence et la valeur du champ électrique atmosphérique au moment de la détection. Nous

avons montré que la détection du signal de mort subite était très peu probable pour les expériences à basse altitude, telles qu'EXTASIS, et qu'afin d'augmenter les chances de détection de cette contribution, des sites plus haut en altitude tels que l'Observatoire Pierre Auger étaient à privilégier.

Concernant CODALEMA, après sélection et reconstruction des événements, nous nous sommes intéressés à la composition en masse des rayons cosmiques. Le cadre d'analyse mis en place nous a permis d'obtenir une statistique suffisante pour le nombre d'événements reconstruit. Nous avons montré, de manière préliminaire, que les résultats obtenus concernant l'estimation de l'énergie et du paramètre X_{max} pour les événements de CODALEMA étaient en accord avec les résultats obtenus par les autres expériences telles que LOFAR ou Auger. Différents tests ont été menés afin d'asseoir les résultats obtenus, notamment en appliquant différentes coupures/sélections sur le lot d'événements à traiter. Le résultat final obtenu est en accord avec ceux proposés par la communauté, validant ainsi toutes les étapes de notre cadre d'analyse, de la calibration à la généralisation de la méthode de comparaison de simulations avec les données en passant par l'amélioration de la sélection des événements.

Perspectives

Bien que nous ayons, tout au long de ce travail, montré que la technique de la radio détection était capable de nous fournir les renseignements attendus concernant les rayons cosmiques primaires, à savoir leur énergie et leur nature, le tableau n'est pas tout rose. En effet, lorsque nous regardons l'efficacité de la radio détection, ainsi que le pourcentage d'événements reconstructibles, la conclusion est que la technique de radio détection ne peut remplir, à courts termes, toutes les attentes, notamment en ce qui concerne l'augmentation de la statistique. Bien que la technique de la radio détection soit aujourd'hui suffisamment mature pour rivaliser avec les autres techniques en termes de qualité de reconstruction, le constat est le suivant : la radio détection ne semble avoir apporté que peu au domaine des rayons cosmiques. En effet, les autres techniques ont déjà fourni par le passé, et continuent aujourd'hui, à fournir les informations nécessaires et attendues : la composition des rayons cosmiques de l'observatoire Pierre Auger, ou de Telescope Array, est fournie par les télescopes à lumière de fluorescence.

Si la technique de la radio détection des rayons cosmiques semble atteindre ses limites, de nouveaux projets émergent cependant, tels que *AugerPrime* qui vise à instrumenter les 1600 cuves d'Auger avec des antennes radio pour étudier les gerbes inclinées, ou encore le programme clef "rayons cosmiques" de SKA. Parallèlement, la radio détection des neutrinos d'ultra haute énergie est largement investiguée, par les expériences ARA ou ARIANNA en Antarctique, mais aussi par les nouveaux projets tels que RNO et IceCube-Gen2.

De même, un projet au sein du groupe émerge pour la radio détection des gerbes atmosphériques initiées par des rayons gamma. Contrairement aux rayons cosmiques dont les sources ne sont pas identifiées et qui nécessitent donc des instruments couvrant la totalité du ciel, les sources de rayons gamma sont connues. Il devient alors possible, grâce à un radio télescope, de pointer directement vers les sources pour augmenter la sensibilité de détection, ce qui est nécessaire puisque les rayons gamma n'atteignent pas les énergies des rayons cosmiques, et les gerbes produites se développent très haut dans l'atmosphère, n'atteignent que peu le sol et le signal radio associé est plus faible. C'est l'objectif du projet GammuFAR, qui prévoit l'utilisation du radio télescope NenuFAR situé à l'observatoire de radioastronomie de Nançay.

Ce manuscrit traite d'un sujet complexe, tant par la complexité des processus physiques impliqués que par leur pluralité. Il a été construit de manière à ce que le lecteur ait une vision et une compréhension progressive des éléments étudiés au cours de ces trois années. Les études menées au cours de cette thèse et développées tout au long de ce manuscrit se situent à la frontière entre différentes disciplines, telles que la physique des particules, la physique des rayons cosmiques et l'astrophysique. Le manuscrit est divisé en quatre parties. La première partie introduit la notion de rayons cosmiques, donne le cadre théorique minimal à la compréhension du domaine et se conclut sur l'étude des grandes gerbes atmosphériques. La seconde partie du manuscrit se concentre dans un premier temps sur le signal radio émis par les gerbes, détecté par les expériences CODALEMA et EXTASIS dont la configuration expérimentale est discutée. Nous détaillons ensuite la méthode de calibration utilisée, ainsi que les méthodes de sélection des événements détectés. La troisième partie présente les résultats obtenus lors de ces travaux de thèse, notamment les résultats obtenus avec EXTASIS sur la bande basse fréquence [1.7 – 3.7] MHz, mais aussi les résultats obtenus avec CODALEMA sur la composition des rayons cosmiques et les analyses développées pour y arriver. Enfin, la dernière partie comporte une conclusion sur les travaux réalisés, et ouvre sur une discussion critique sur la technique de la radio détection des rayons cosmiques et autres astroparticules.

Contents

List of Figures

List of Tables

Prologue

Preamble

As it is mentioned in the title, these thesis researches deal with the study of the radio signal in [1 – 200] MHz which is emitted by cosmic ray induced air showers, in order to determine their characteristics. Despite the discovery of cosmic rays by Victor Hess in 1912, there are more than one hundred years, many questions remain unanswered today: what are cosmic rays, how are they created and where do they come from? Nowadays, we know that the Earth is constantly bombarded by these cosmic rays. It is a flow of charged particles of ultra-high energy arriving in the Earth's atmosphere, which decreases with the energy of the primary cosmic ray. Thus, the higher the energy, the lower the flow. This observation is presented in the figure 1, presenting the differential flux of the cosmic rays as a function of the energy. The flux extends over

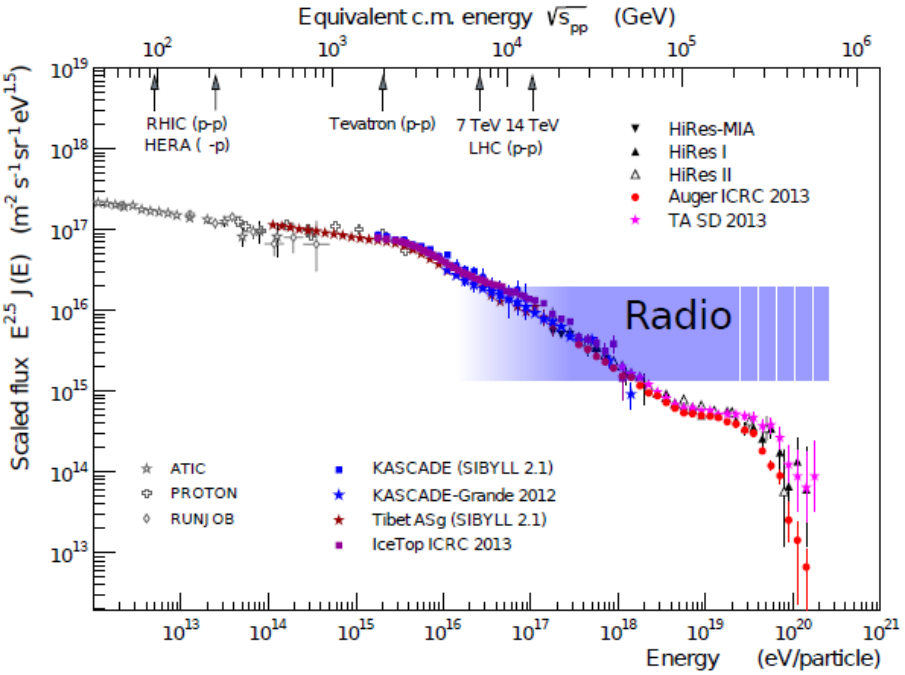


Figure 1 – *Energy spectrum of the highest energy cosmic rays measured by various experiments [6].*

more than 32 order of magnitude, and over about 12 order of magnitude for their energy. This is unique in physics, which could sign a universal production mechanism of these cosmic rays. This flow of cosmic rays decreases in a remarkably regular way according to their energy. It is in this that their study is important, since the understanding of their production mechanisms, their sources, goes through the study of the spectrum. Let us now turn our attention on the units of the flux. It is defined in terms of unit area, unit time, unit solid angle and unit energy. We see that at high energy, the flux drops drastically to only one particle per km^2 per century for a particle of 10^{20} eV. Beyond 10^{14} eV, the cosmic rays can't be detected directly because of their low flux. Below this energy, they can be detected directly thanks to dedicated satellite (such as the AMS experiment installed on the International Space Station [7]) or experiment with stratospheric balloons. For the cosmic rays with an energy above 10^{14} eV, their study will be indirect. Indeed, when

a cosmic ray arrives on Earth, it will interact with the components of our atmosphere. During this interaction, the molecule will "break" and create secondary particles, which in turn will interact with the components of the atmosphere, and generate chain reactions that is called a shower of particles. This shower can be detected during its development in the atmosphere, with the so-called fluorescence technique, which consists in collecting the fluorescent light emitted by the de-excitation of the nitrogen during the passage of the charged particles composing the shower, or directly on the ground, with dedicated detectors counting the particles. However, a shower may not reach the ground. It depends on several parameters, such as the energy of the primary cosmic ray and its angle of incidence. As written above, we expect one particle per square kilometer per year at the ankle. It is easy to understand that it is impossible to wait as long to detect a physical phenomenon. The solution is to build detectors covering huge surfaces, such as the Pierre Auger Observatory, in Argentina, which covers ~ 3000 km^2 (equivalent area of Nantes and its campaign). To date, the most energetic cosmic ray ever detected (Fly's eye [8]) had an energy of 3.2×10^{20} eV, which is 50 J, corresponding to a tennis ball going at 160 km $\cdot$ h^{-1}. The most powerful particle accelerator to date, the Large Hadron Collider (LHC), is able to accelerate particles up to 7 TeV per nucleon. Its circumference is about 27 km. By comparison, and with the same technology, we would then have to build a particle accelerator with the circumference of the orbit of Mercury, which is 10^7 km. Impossible, so. This colossal energy makes cosmic rays a very interesting object of study, because the energies reached are not reproducible in the laboratory. Moreover, the objectives of this study is to fully characterize these messengers of cosmos:

— the type of source which produce this primary particle (active galactic nuclei, gamma-ray burst),

— production mechanism (relativistic shocks by Fermi acceleration),

— energy,

— the nature of the primary particle (proton, iron ...).

During the development of the shower, several phenomena occur, and namely the creation and the emission of an electric field, which can be detected at ground with dedicated radio antennas like those of the CODALEMA and EXTASIS experiments. The experiment is located at Nançay, in the Cher department in central France, within the Nançay Radioastronomy Observatory. It covers 1 km^2, allowing us to study the cosmic ray from 10^{16} eV to 10^{19} eV, either the high energy part of the spectrum. The CODALEMA experiment was at the beginning of the 2000's a demonstrator, aiming at prove the feasibility of the radiodetection. CODALEMA has demonstrated the feasibility of the radio detection method, and highlighted the geomagnetic and charge excess mechanisms as sources of the radio emission. The radiodetection of EAS, often falsely named the radio detection of cosmic rays, is used by several experiments now. Indeed, the radio signal emitted by the EAS is bright, and contains all the pieces of information needed to answer the question previously mentioned. For instance, the signal amplitude is directly proportional to the energy of the cosmic ray. The timing of the signal in the different antennas allows to reconstruct the direction of arrival. Moreover, contrary to the fluorescence method which has a duty cycle of 14 %, the radiodetection has a duty cycle of ~ 100 %. This means that the radio method can always be used, except during storm, and thus increases the statistics of detection. Another advantage of the radio method is the use of low cost detector, namely less than 3,000 euros for one antenna of the CODALEMA experiment. The low cost of these detectors could permit to cover large area to allow a high energy study of the cosmic rays, such as the GRAND experiment which provides for the installation of about 200,000 antennas over 200,000 km^2.

Organization of the manuscript

This manuscript, entitled *From the observation of Ultra High Energy Cosmic Ray (UHECR) radio signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments*, deals with a complex subject, as much by the complexity of the involved physical processes as by their plurality. It was constructed in a way that the reader has a view and a progressive comprehension of the studied elements during these three years. The studies conducted during this thesis and developed throughout this manuscript are at the boundary between different disciplines, such as particle physics, cosmic ray physics and astrophysics. As the reader will be able to note, many notions are addressed in the first chapters in order to set a clear framework, I hope, for the rest of the manuscript. Since most of these notions are all linked, the greatest difficulty has been to split the chapters so that they are more or less self-supporting. Which has not always been the case. Many cross-references are present, to help the reader find the necessary information in the right place.

The manuscript is divided in four parts:

— The first part aims to introduce the important notions of the field of cosmic rays. A historical introduction seemed necessary for the reader to grasp the complexity of the field, and the time it took to reach this maturity. We then give the basic theoretical framework for the field of cosmic rays, ranging from their discovery to the necessary process for their creation. An overview of the cosmic ray domain is then presented, based on the results of the major experiments available at the time of writing of this manuscript. The work of this thesis being on the UHECR, the notion of Extensive Air Shower (EAS) is mandatory, and discussed in chapter 2, in which we deals with the relevant electromagnetic processes for shower development and with the phenomenology and physics of EAS. The different EAS detection methods are then discussed.

— The second parts focuses on the radio mechanism occurring during the EAS development in the atmosphere. Chapter 3 aims to gather all the pieces of information related to the radio emission, from the physics of radio transient emission mechanisms to the modern radio experiments, passing by the simulation of the electric field emitted by EAS and the detection of the radio signal in the digital era. Chapter 4 presents the instrumental setups of the CODALEMA and EXTASIS experiments. Several results are also discussed in this chapter for instruments that can not by themselves be part of a chapter in their own right, such as the three-fold antenna. Chapter 5 deals with the cross calibration of the CODALEMA radio detectors using the Galactic radio emission, in order to have a more accurate radio measurement and to be able to go back to the characteristics of the primary cosmic rays (which is the purpose of the chapters 7 and 9), namely the energy and the composition. We then, in chapter 6, deals with exploratory methods to improve the selectivity of the radio events, especially in order to have a batch of events as clean as possible for the further works described latter, and with the necessary developments of analysis methods for transient search in EXTASIS signals.

— The third part is dedicated to the analysis. Chapter 7 presents the comparison method between data and simulations used to reconstruct the primary cosmic ray parameters. Chapter 8 gathers all the results obtained for the EXTASIS experiment, part of it have been published in [9]. We report the observation of 25 air showers detected in coincidence by CODALEMA and EXTASIS, and estimate a detection threshold of 23 ± 4 $\mu V \cdot m^{-1}$ from comparisons with detailed simulations. We also report a strong correlation of the low-frequency signal observation with the

atmospheric electric field. This part ends with chapter 9 dedicated to the results obtained on the estimation of the primary cosmic ray parameters and presents a proposition of mass composition. It will notably discuss the reconstruction of events with particle detectors, comparing the results obtained with the radio method and we will give for the first time a proposition of mass composition with the CODALEMA experiment.

— The last part begins with chapter 10 which compiles all the results and observations obtained during this thesis. We propose a discussion centered on the obtained results, in particular of the efficiency of the radio detection method to detect the EAS and reconstruct the parameters of the primary cosmic rays. Chapter 11 is a comprehensive summary of the work done and presented in this manuscript.

Various appendices are also present at the end of the manuscript, especially in order not to overload the chapters. They are intended to clarify certain points, to specify the conventions used and to share certain areas of work to be exploited:

— Appendix A defines the convention used throughout the manuscript, namely for the type of filters used and for the Power Spectrum Densities (PSD).

— Appendix B presents the Hierarchical Equal Area isoLatitude Pixelisation, comporting a set of functions for astrophysics and cosmology data representation, which is used to produce several figures of chapters 3, 5 and 6.

— Appendix C presents the information coming from the Global Data Assimilation System, which is used to have a more precise description of the atmosphere at the time of ours cosmic-ray detections and for a location close to the experiment site.

— Appendix D is a review of the pioneer works on the low-frequency range carried out in the 70's and up to the 90's. It is composed of a short review where the most relevant pieces of information are summarized in a table, and of a complete review giving information on the experiment site, the instrumental setup and a personal commentary for each paper.

— Appendix E deals with the exploitation of the particular pattern which appears from time to time in the CODALEMA data: a part of a Cherenkov ring. This particularity can be exploited for further improvements.

— Appendix F gives the necessary elements for calculating the cosmic ray flux for a given experiment.

Part I

Visitors from elsewhere !

"To infinity and beyond!"
Toy Story, Buzz

CHAPTER 1

Ultra High Energy Cosmic Rays: From the discovery to the modern observations

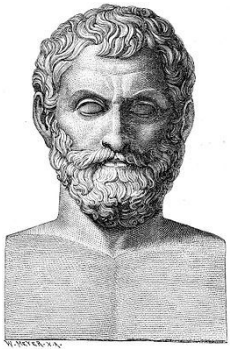
Outline

Cosmic rays are commonly defined as relativistic atomic nuclei travelling through the Universe, and possibly arriving on Earth. The distribution of their flux according to their energy is one of the most remarkable and questioning in the modern physics, spreading over 32 orders of magnitude in flux and more than 13 orders of magnitude in energy (from a few hundred MeV up to more than 10^{20} eV for the most energetic ones observed so far, and that are not explained by any identified physical process). According to their energy, they constitute, just like electromagnetic radiation, a single source of information on phenomena of galactic (up to $\simeq 10^{15}$ eV) and extragalactic origin (above 10^{15} eV). However, the important interactions with the galactic and extragalactic magnetic fields complicate the interpretation of the collected data in order to determine their source and their nature.
This chapter is in no case a chapter on high energy physics. Nevertheless, it aims to gather all essential elements on cosmic ray physics to the understanding of the following manuscript. For that, it seems to me appropriate to begin from the beginning, to the origin of scientific discoveries leading to the cosmic ray study. We thus redraw the history of this kind of physics, starting from the origin of the electromagnetism until the discovery of cosmic rays [10, 11].

1.1. The beginning

1.1.1. The signs of the electromagnetism

The electromagnetism is frequently associated to well known names, such as Charles-Augustin de Coulomb (1736-1806) or James Clerk Maxwell (1831-1879), but its history begins long before. The first observations of the effect of electromagnetism had been carried out during antiquity, when the great thinkers of that time observed the attraction of small pieces of iron by magnet stone (magnetite) or the attraction of straws by a piece of polished amber. This distant action was first experimented by Thalès de Milet (around 625-547 before J.-C.), see figure 1.1.(a). The magnetite has its name from its geographical



(a) *Thalès de Milet.*

(b) *Magnetite stone [10].*

Figure 1.1 *– Thalès de Millet and a magnetite stone.*

origin, the region of Magnesia, in Anatolia (Asia Minor) (see figure 1.1.(b)). In some Platon's writings (around 428-348 before J.-C.) is mentioned the magnetic behaviour of piece of iron in contact with a magnet. An attempt to explain these behaviours will be reported later by Démocrite of Abdère, with the atomistic theory proposing a concept of the universe constituted of atoms and vacuum. It is worth noting that two effects were distinguished: the magnetite which only attracts pieces of iron, and polished amber which attracts mustard seed for instance. Using the first effect, the first compass was devised with a spoon in magnetite placed on a polished surface.

1.1.2. Science of magnetism

This is only in the 13th century that the science of magnetism begin with Pierre de Maricourt who provided the properties of attraction-repulsion of magnetic poles in the first treaty on the properties of magnet in 1629. The explanation of the magnetic declination of a compass (angle formed by the deviation of the needle relative to the geographical north) will be provided only around the 15th century by Europeans. The first map of the magnetic declinations was established around 1700 by Edmund Halley (1656-1742). Halley explained that the magnetic declination varied with time due to the fact that Earth was constituted of magnetic concentric layers in rotation relative to each other, leading to a terrestrial magnetic field (this magnetic field is also at the origin of polar aurora). The measure of the longitude by this information was thus abandoned in favour of the inclination, which recounts the fact that the needle is heading parallel relative to the terrestrial field on a non-horizontal plan. It was discovered by Georg Hartmann in 1544, but the first demonstration of the phenomenon was carried out by Norman (see figure 1.2 left).

The goal of this demonstration is to test the hypothesis of an attractive point. The needle and the cork are then immersed in water. The apparatus does not show any movement, only the needle orientates. The hypothesis of an attractive point is so rejected,

and the correlation of the orientation of the needle and the terrestrial magnetism is shown. William Gilbert (1544-1603), English scientist and physician, was the first to gather the ideas, namely those of Maricourt and Norman, and brought a fresh pair of eyes on the domain. He wrote what one can consider as the first book on magnetism, *De Magnete*, see figure 1.2 right.

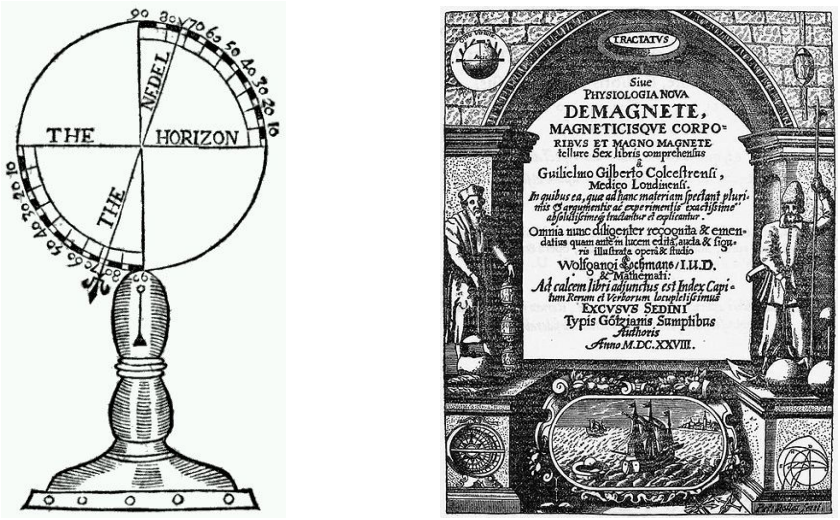


Figure 1.2 *– Left: Norman's experiment [10]. Right: De Magnete, first book on magnetism.*

Bringing a fresh pair of eyes, he brought into opposition with Maricourt's idea, who considers as candidate the sky as the cause of the orientation of the needle of the compass. For his part, and using the well known hypothesis (heliocentricity theory) brought by Copernic (1473-1543), he considered that Earth was the cause of this orientation. He attributed thus to the declination phenomenon the irregularity of the terrestrial surface, while for the inclination phenomenon the cause came from the Earth's magnetization. It was from him that came the following convention: the tip of the needle heading towards the geographical north is named the south pole of the needle, attracted by the magnetic north pole, nearby the geographical north pole. In some of his writings, one can find the fact that the magnetic behaviour of a magnetic piece disappeared by heating it, what we called nowadays the lost of permanent magnetization when the piece of metal passes the Curie's point.

1.1.3. Then came electricity

Many advances were carried out in the electricity domain, but the first remarkable at this time was the Leyden jar (see figure 1.3). It is in 1746, in Leyden in Netherland, that Petrus van Musschenbroek (1692-1786) tried to catch electricity in a water bottle. Obviously (what we can say with our knowledge of today, but certainly not with those at this time), he received a violent discharge when he tried to retrieve his bottle. The first capacitor was born. Stephen Gray (1666-1736) distinguished two types of bodies during experiments on the electricity transmission. There are the bodies allowing electricity propagation, such as metals, and those which do not allow it, such as glass. However, the notion of conductor/insulating materials was evoked only around 1740 by John Theophilus Desaguliers (1683-1744). One considers nowadays that Gray was the precursor of the discovery of property of the transmission of electricity, therefore specific to all bodies. These works have been continued by Charles-François de Citernay Dufay (1698-1739). He showed that non electrifiable bodies by frictions can be electrifiable by contact. He postulated as principle that electricity passes from a body to another. The first one must be charged and the second one neutral. The neu-

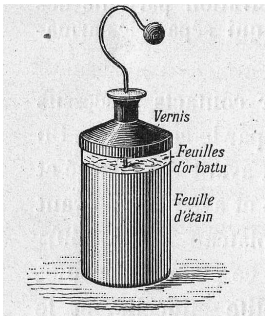


Figure 1.3 *– Leyden jar.*

tral body is then attracted by the first body. When electricity is transmitted to the neutral body, the two bodies repeal each other. The clergyman Nollet (1700-1770) thought that the observed phenomena were linked to a fluid into and out of the charged bodies. Here, it is a question of the unique fluid theory, supported by Benjamin Franklin (1706-1790) (see figure 1.4), inventor of the lightning rod. This theory postulates that electricity is a fluid constituting bodies (same quantity of fluid whatever the body), and that the vector of exchange is the friction, passing the fluid from a body to the other. The body which emits the fluid becomes negatively charged, and the body which receives the fluid becomes positively charged. But this fluid theory does not stop here. It already arrived to all of us, upon withdrawal of clothes, to note the apparition of sparks. This is what happened to Robert Symmer (1707-1763), upon withdrawal of his stockings. Therefore, based on the statements of Franklin, he postulated the existence of two fluids, permitting to forget what object emits or receives the fluid. Both theories of unique fluid and of double fluids coexist in the 18th century.

Figure 1.4 *– Benjamin Franklin (1706-1790).*

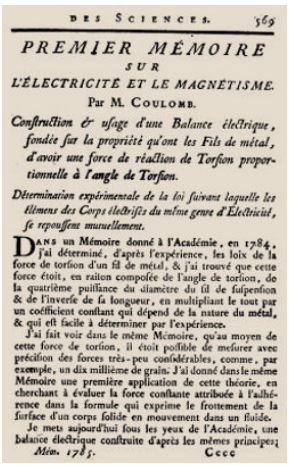
1.1.4. Basis of Electrostatics

Thanks to his experiments related in his series of memories with the torsion pendulum shown in figure 1.5, the french physicist Charles-Augustin Coulomb states in 1785 his law on the electrostatic interaction between two charged bodies : «*La force répulsive de deux petits globes électrisés de la même nature d'électricité est en raison inverse du carré de la distance du centre des deux globes.*»

$$\vec{F}_{1\to 2} = k\frac{q_1 q_2}{r^2}\vec{F} \tag{1.1}$$

where this force $\vec{F}_{1\to 2}$ is in Newton, noted N, q_1 et q_2 are the charges of two bodies stated in Coulomb, noted C and r is the distance between the two bodies in metre. The prefactor k, equal to $8.9875517873681764 \times 10^9$ N $\cdot$ m^2 $\cdot$ C^{-2}, commonly called Coulomb constant, will be determined later by Heaviside (1850-1925), and is defined from the vacuum permittivity ϵ_0. The torsion pendulum is composed, at its stand, of a large cylinder which contains angular graduations. The upper storey is composed of a cylinder with a lower diameter, which contains the torsion thread, whose extremities are equipped with a horizontal axis which ends by a small charged sphere. In the base of the cylinder, at the same altitude, is fixed a secondary charged sphere. In this way, a balance is established between the two spheres, and so between the electrostatic forces involved. The intensity of the force can be determined, and the relative position of spheres permits to check the inverse dependence with the square distance of the force of Coulomb. However,

(a) *Charles-Augustin Coulomb (1736-1806).*



DES SCIENCES. 569

PREMIER MÉMOIRE
SUR
L'ÉLECTRICITÉ ET LE MAGNÉTISME.
Par M. COULOMB.

Construction & usage d'une Balance électrique, fondée sur la propriété qu'ont les Fils de métal, d'avoir une force de réaction de Torsion proportionnelle à l'angle de Torsion.

Détermination expérimentale de la loi suivant laquelle les élémens des Corps électrisés du même genre d'Electricité, se repoussent mutuellement.

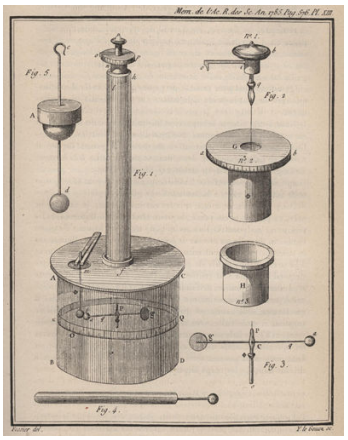
DANS un Mémoire donné à l'Académie, en 1784, j'ai déterminé, d'après l'expérience, les loix de la force de torsion d'un fil de métal, & j'ai trouvé que cette force étoit, en raison composée de l'angle de torsion, de la quatrième puissance du diamètre du fil de suspension & de l'inverse de sa longueur, en multipliant le tout par un coefficient constant qui dépend de la nature du métal, & qui est facile à déterminer par l'expérience.

J'ai fait voir dans le même Mémoire, qu'au moyen de cette force de torsion, il étoit possible de mesurer avec précision des forces très-peu considérables, comme, par exemple, un dix millième de grain. J'ai donné dans le même Mémoire une première application de cette théorie, en cherchant à évaluer la force constante attribuée à l'adhérence dans la formule qui exprime le frottement de la surface d'un corps solide en mouvement dans un fluide.

Je mets aujourd'hui sous les yeux de l'Académie, une balance électrique construite d'après les mêmes principes

Mém. 1785. Cccc

(b) *First memory of the series of Coulomb.*



(c) *Drawing of the torsion pendulum of Coulomb.*

Figure 1.5 *– Experiments of Coulomb [10, 11].*

during this handling, Coulomb notices that both spheres discharged over time. Several hypothesis came to his mind: he knows that the ambient air is neutral, but proposes nevertheless an explanation to this discharge of the spheres. From his point of view, other infinitely small bodies (since non-visible to the naked eye) may enter in collision with the charged spheres, inducing thus the discharge. In hindsight that we have, we can see here the first link with the cosmic radiation, ionising the air molecules leading thus to the discharge of the charged spheres of the torsion pendulum. The notion of current has been highlighted by Ebenezer Kinnersley (1711-1778) who proposed an experiment. The direction of current is thus defined as going from the body which has the most of fluid to the body which has the less of fluid. In 1800, Alessandro Volta (1745-1827) invents the first electric cell (see figure 1.6). It is a piling up of different metals (electrodes, ini-

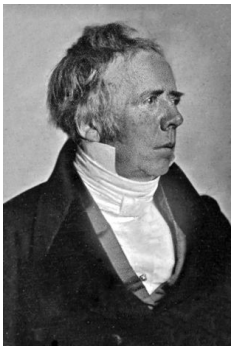
Figure 1.6 *– Alessandro Volta (1745-1827) and the first electric cell.*

tially zinc-copper discs) separated by a conductor medium (electrolyte, initially moist cardboard) creating a continuous electric energy. In 1801, he presented his invention to the French institute : the law of tensions is born. The physicists at this time have now a continuous current source, and the previous experiments, doomed to failure due to the lack of this continuous source, will have the opportunity to lead to concrete results,

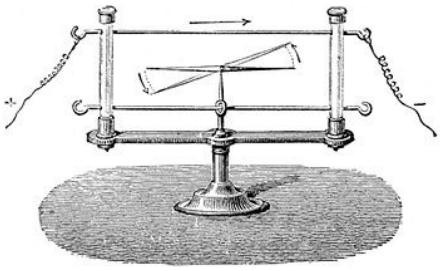
including the magnetization of iron rod under the action of lightning (Franklin's works), or the orientation of magnetized needle during the discharge of the Leyden jar.

1.1.5. Birth of electromagnetism

The first to exploit this scientific potential is the physics teacher at the Copenhagen university, Hans Christian Oersted (1777-1851). Giving a course during the 1819 winter on the calorific effect of the voltaic cell, he notes that the needle of a compass situated near the cell starts to move. The deviation is maximal when the conductor wire is placed alongside of the magnetized needle, which is perpendicularly oriented to the wire. When there is an inversion of the two poles of the cell, the orientation of the needle is done in the opposite direction: this is the demonstration that a current circulating in a wire creates a magnetic field which is perpendicular to the wire. Let us recall that at this time, the two fluids model is adopted. His explanation by the fluids is just as follows: «*the negative electric matter describes a spiral on the right and acts on the north pole while the positive electric matter has a movement in the inverse direction and has the property to act on the south pole*»(see figure 1.7). Another adept of this theory will play an important role: Andre-



(a) *Hans Christian Oersted (1777-1851).*



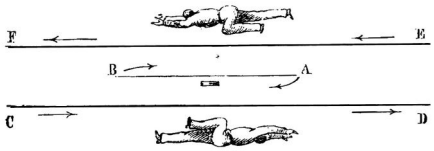
(b) *Apparatus.*

Figure 1.7 *– Oersted's experiment.*

Marie Ampère (1775-1836). Based on the previous assumptions, he will be the first to deal with the conventional direction of the current. He conserves the concept of positive and negative electricity, but indicates that it is only a convention. Indeed, knowing the direction of flow of one of the two fluids, the direction of the flow of the second fluid can be deducted. The chosen convention by Ampère is the positive fluid. It is worth noticing that what is described here by Ampere is not the direction of flow of the current but the direction of the currents themselves, and the right-hand rule appears. Depending on the direction of the current, this rule permits to understand the deviation of the needle of a compass: «*Pour définir la direction du courant relativement à l'aiguille, concevons un observateur placé dans le courant de manière que la direction de ses pieds à sa tête soit celle du courant, et que sa face soit tournée vers l'aiguille ; on voit alors que dans toutes les expériences rapportées ci-dessus le pôle austral de l'aiguille aimantée est porté à la gauche de l'observateur ainsi placé*»(see figure 1.8). Using a galvanometer, invented by Johann Schweigger (1779-1857) but named after Luigi Galvani (1737-1798), Ampère measures the intensity of this electric current. The resulting mathematical theorem is well known: «*la circulation, le long d'un circuit fermé, du champ magnétique engendré par une distribution de courant, est égale à la somme algébrique des courants qui traversent la surface définie par le circuit orienté, multipliée par la perméabilité du vide*». Internationally translated by:

$$\oint \mathbf{B}.\mathrm{d}\ell = \mu_0 \sum I_{traversant} \tag{1.2}$$

(a) *Andre-Marie Ampère (1775-1836).*

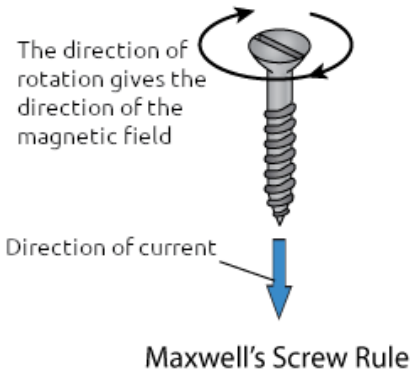


(b) *Right-hand rule.*

Figure 1.8 *– Ampère's man.*

with **B** the magnetic field, ℓ is the path element on the closed contour, μ_0 is the vacuum permeability and the last term is the sum of currents which pass through the formed surface. Despite a high acceptance of the two fluids theory, England remained cautious and faithful to Franklin, then adept of the one fluid theory. He will find an important ally in the person of James Clerk Maxwell (1831-1879), using this theory for his interpretation of electromagnetic phenomena [12, 13, 14]: «*If a right handed cork screw is assumed to be held along the conductor, and the screw is rotated such that it moves in the direction of the current, direction of magnetic field is same as that of the rotation of the screw*», also known as the rule of «tire-bouchon», see figure 1.9). So far, the electric theory was described by one or two

(a) *James Clerk Maxwell (1831-1879).*



(b) *Maxwell's screw rule.*

Figure 1.9 *– James Clerk Maxwell and his screw rule.*

fluids depending on the chosen description. But this explanation ends in 1897 with the discovery of the electron by Joseph John Thomson (1856-1940) and Jean Baptiste Perrin (1870-1942). They demonstrate that cathode (electrode output current in an electrolyser) rays are composed of «corpuscles of electricity», that we will call electron. This discover is capital, and supports the initial concept of Franklin: the unitary theory. However, Franklin had erred, Thomson and Perrin observed that grains which compose the cathode rays were negatively charged: «*Au lieu de considérer, comme le faisait cet auteur, le fluide électrique comme étant de l'électricité positive, nous le considérons comme de l'électricité négative... Un corps chargé positivement est un corps qui a perdu une partie de ses corpuscules*».

The first great usage of all these discoveries previously described comes to the english physicist Peter Barlow (1776-1862), whose the wheel named after his name can be assimilated to the first electric machine. The design of the machine shown in figure 1.10 is as follows: a wheel, composed of many metallic jagged peaks which are immersed in mercury, is traversed by an electric current. A magnet in horseshoe-shape is disposed near the recipient, creating a magnetic field. When a jagged peak is near the magnet, it suffers the electromagnetic force of Laplace who determined the analytic expression of the current-magnet force, which sets it in motion, perpendicularly to the axe of the wheel. This movement lead to the rotation of the wheel, another jagged peak arrived in the field of the magnet, and the mechanism repeats as long as an electric current circulates, and for that is used the electric cell of Galvani and Volta. Although we can not deal with an electric motor (the engine power which is delivered is vanishingly small), it is here a revolution, because a continuous generator of electricity can set a wheel in movement.

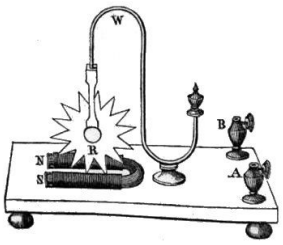


Figure 1.10 *– Barlow's wheel.*

1.2. From electromagnetism to cosmic ray discovery

1.2.1. Discovery of an extraterrestrial radiation

As said in section 1.1.4, Charles-Augustin de Coulomb suspected the presence of infinitely small charged particles in the air. His observations were confirmed in 1835 by Micheal Faraday. William Crookes, in 1879, noticed that the rate of discharge in the air measured with an electrometer (also known as electroscope, see figure 1.11) decreases with atmospheric pressure, and conclude that these discharges were due to the air conductivity, and thus to the air ionization. At this time, it was a really difficult question to determine what was the cause of the air ionization and so of the atmosphere ionization. With the work and the discovery of radioactivity by Pierre and Marie Curie in 1896, and then the discovery of polonium in 1896, the main idea was that the ground and its constituents were the first main source of the atmosphere ionization. In 1909, Theodur Wulf

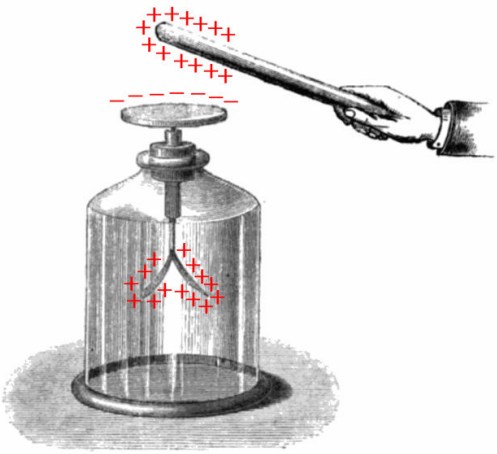
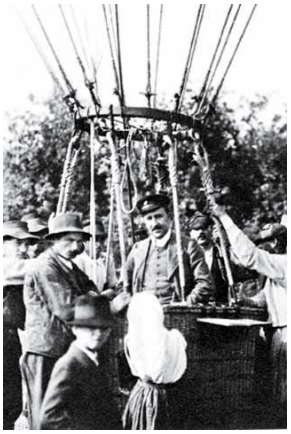
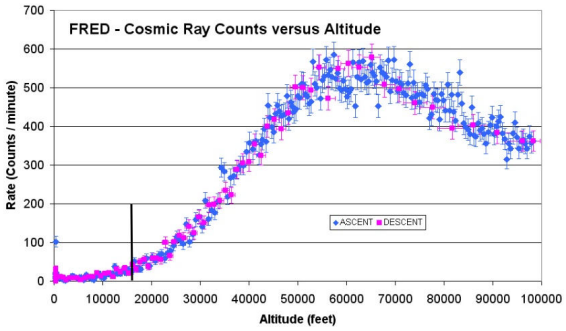


Figure 1.11 *– Illustration of an electroscope. The operation of an electroscope is quite the same as the one of the torsion pendulum of Coulomb. A disk terminal is attached to the top of the device, and the gold leaves are installed inside a recipient. When the top disk terminal is charged, the two leaves received the same sign of charge and repeal. If there is no external action on the disk terminal or the two leaves, they should be kept apart in an inverted 'V'.*

built a transportable electroscope. He measured the air ionization at the bottom and the top of Eiffel Tower: the ionization rate decreases with altitude, from 6 ions by cm^3 to 3.5 ions by cm^3. However, the hypothesis was that the ionization was due to gamma rays coming from the ground (via natural radioactivity), and the measurements made by Wulf were in contradiction with what was expected, namely a faster decay of the ionization rate with the altitude. May the source be from above? This question raises time of ballooning experiment. The first one was in 1911 with Victor Hess (see figure 1.12.(a)). The experiment aimed to measure the possible variation of the radioactivity up to 1,300 meters (4,265 feet, first flight). He needed seven flights to obtain correct results, bringing him to near 5,200 meters (~17,000 feet). Figure 1.12.(b) shows the compiled data of cosmic ray count versus the altitude from the FRED balloon from the LaACES program. The highest altitude reached by Hess is indicated by the vertical black line. The ionization rate is decreasing up to 1 km and increases with altitude beyond this point. Hess organized



(a) *Victor Hess after take-off in 1912 [15].*



(b) *Cosmic ray count versus altitude. The vertical black line represents the highest altitude reached by Hess (17,000 feet, ~ 5200 m). Extracted from [16]. Courtesy T. Gregory Guzik, Louisiana State University, LaACES program.*

Figure 1.12 *– Victor Hess after take-off and ionization rate measured by LaACES program.*

flights during a total solar eclipse, which led him to assume that the ionization came from above, but also that this ionization was not coming from the sun. These observations were confirmed by Werner Holhörster in 1913, his flights reaching 9,300 meters (30,511 feet).

From that time, it took 20 years, and many experiments especially led by Kolhörster, Millikan, Skobeltzyn and Compton to show and admit that this very energetic radiation is of cosmic origin. At that time, the state of the art was as follows. The measured cosmic radiation does not depend on:

— the Sun,
— the Milky Way,
— other regions of the Universe

and presents a constant intensity and seems to come from all the directions of the sky.

1.2.2. Cosmic ray?

The common definition of cosmic rays is that they are charged particles with an extraterrestrial origin that reach the Earth. The majority of cosmic ray particles are well described by this definition, because they mainly consist of charged nuclei. This is a common belief that cosmic rays can include all stable nuclei [17]. Moreover, the definition

also includes the flux of electrons, but does not include the neutral particles. What do we want to learn about cosmic rays? The main astrophysical question is the origin of this extra solar system matter. Some cosmic rays reaching the Earth have a very high energy: how does galactic cosmic ray matter get accelerated to such high energies? How do the particles propagate from source(s) to Earth? Are there signatures of any exotic physics? As we can see with these questions, cosmic ray physics is a cross-disciplinary area including astrophysics, high energy particle physics and plasma physics. Figure 1.13 shows a part of the energy spectrum of cosmic rays. The entire spectrum covers 12

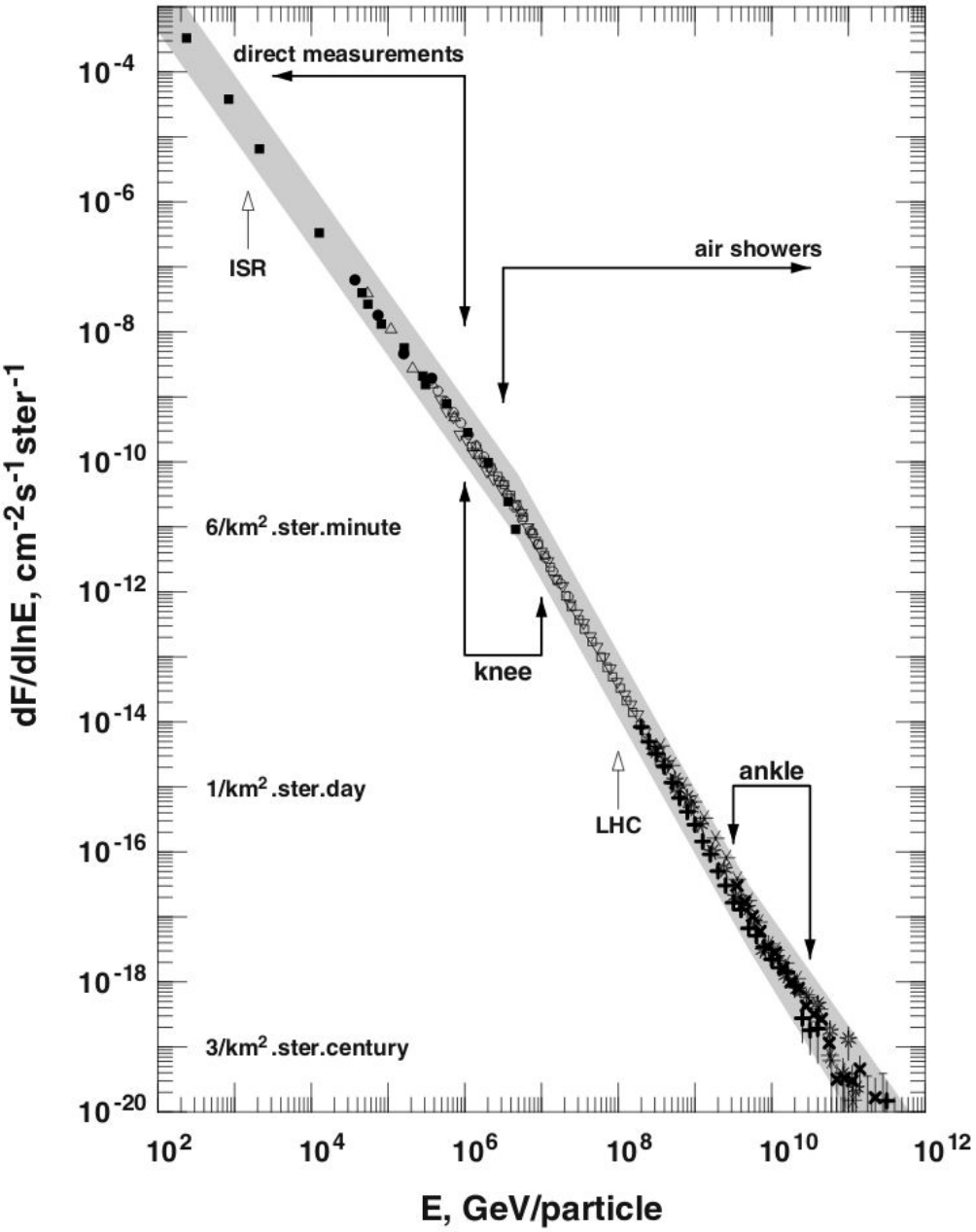


Figure 1.13 *– Energy spectrum of all cosmic ray nuclei [17].*

decades in energy and 32 decades in flux. This is a well-known fact that the particle flux decreases rapidly with the energy. The spectrum follows a power law $\mathrm{d}N/\mathrm{d}E \propto E^{-3}$, -3 corresponding to an approximate value of the spectral index. The energy is given in GeV and the flux in $\mathrm{cm}^{-2} \cdot \mathrm{s}^{-1} \cdot \mathrm{sr}^{-1}$. The spectrum presents two breaks, where the slope on spectral index changes, called knee and ankle. One can note that the flux of cosmic rays at 10^{11} GeV is sixteen orders of magnitude smaller than at 10^2 GeV. At 10^7 GeV, we have six particles per square kilometre per minute per steradian which enter in the atmosphere,

the number falling down to 3 particles per square kilometre per century per steradian at 10^{11} GeV. This drop of the flux is responsible for the difficulties encountered in the progress of cosmic ray field, and is responsible in the study of the secondary charged particles produced by the interaction of the primary cosmic ray with the components of atmosphere.

1.2.3. Cosmic ray creation and acceleration

In this section, we decide to go straight to the point. This is why we will not develop the *top-down* models. These models are based on the decay of supermassive particles or ultra high-energy process producing cosmic rays that are, directly from their creation, at ultra-high energy. At the time of writing this manuscript, we know that this type of models is discarded since approximately ten years by data from large experiments on cosmic rays such as the Pierre Auger Observatory [18]. This is the reason why we will only deal with the *bottom-up* models, where particles at low energy are accelerated by a series of astrophysical phenomena, for instance by astrophysics sources.

To understand well the cosmic ray creation processes, the characteristics of the potential sources have to be known, that is to say their size and magnetic field. These characteristics are sketched in figure 1.14. It is also known as Hillas diagram, coming from the physicist Hillas, who has defined the minimum requirement to be a possible acceleration site:

$$\left(\frac{B}{G}\right)\left(\frac{R}{pc}\right) > \frac{0.2}{\beta_{sh} Z}\left(\frac{E}{10^{20}\,\mathrm{eV}}\right) \tag{1.3}$$

where B is the magnetic field of the object, R is its dimension, Z is the charge of the charged particle (the cosmic ray) and β is the reduced velocity of the phenomenon which will accelerate the charged particle (known as a shock wave, see section 1.2.3.2 for more details) and E is the energy acquired by the charged particle. This diagram represents the

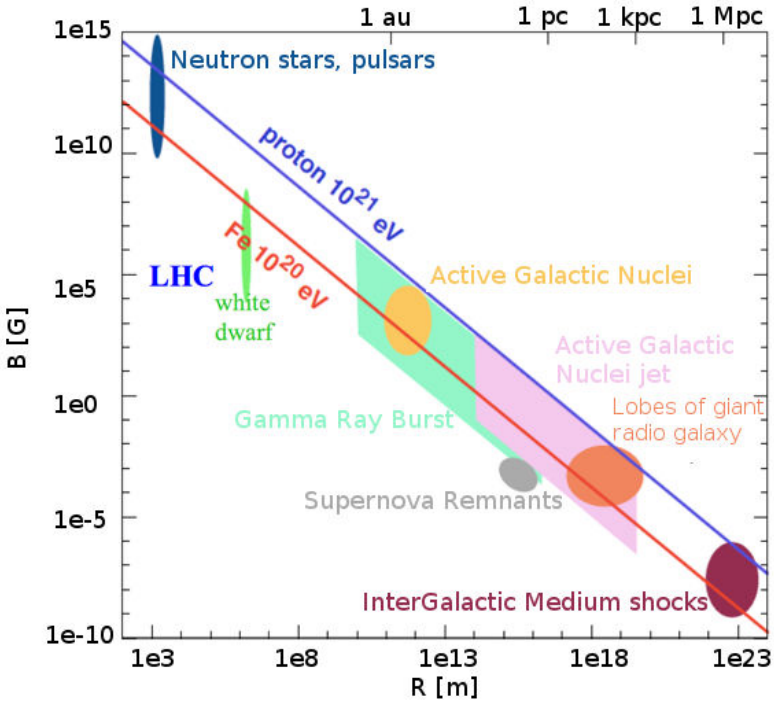


Figure 1.14 *– The Hillas diagram. Adapted from [19]. See text for details.*

different types of sources (magnetic field in G in function of the radius of the astrophysical

object in m). The blue line is the upper limit to accelerate a proton at 10^{21} eV, the red one is the upper limit to accelerate iron nuclei at 10^{20} eV. To be a possible source of acceleration, astrophysical objects have to meet some criteria [17, 20]:

— Maintain the accelerated particle in the acceleration site for a sufficient time.
— Provide enough energy to the acceleration.
— Have density and power enough to weight in the cosmic ray flux.
— Fulfil several conditions on the energy gain and loss of the accelerated particle.

Below 10^2 GeV, their origin is expected to be galactic. For example, a 30 GeV proton in a magnetic field of 6 µG has a gyro radius around 0.5×10^{-5} pc (the diameter of the Milky Way is ~ 40 kpc and its thickness is equal to ~ 200 pc). Above 10^{10} GeV, their origin is expected to be extragalactic. A 3×10^{12} GeV proton in a magnetic field of 3 µG has a gyro radius around 1 Mpc. For example, Andromeda galaxy is ~ 0.65 Mpc from Earth. The gyro radius, also known as Larmor radius, is defined in SI units as:

$$r_g = \frac{m v_\perp}{|q|B} \tag{1.4}$$

where m is the mass of the charged particle, $v_\perp$ is the component of the velocity perpendicular to the direction of the magnetic field, q is the electric charge of the charged particle, and B is the strength of the magnetic field. Considering the Hillas criteria (see equation 1.3), r_g has to be smaller than the radius of the accelerator site otherwise the charged particle escapes from the source. The first break in the spectrum, the so-called first knee, is at around 3×10^6 GeV, and applying the Hillas criteria, we obtain for proton $E_{max} \simeq 3.2 \times 10^6$ GeV considering a magnetic field of 1 µG and 1 pc for the radius of the accelerator site (corresponding to galactic supernova remnants). For iron nuclei, $E_{max} \simeq 10^8$ GeV. Accompanying this break, there is a change in the spectral index. Below the knee, number of particles is divided by 50 when the energy is multiplied by ten, and above the knee this factor is about 100. Beyond this energy, galactic charged particles become rare (around 1 particle per square metre per year), and we observe a second break in the spectrum: the so-called second knee [21]. Around 10^9 GeV, charged particles start to escape our galaxy, due to their gyro radius which becomes larger than the size of the Milky Way: this is the so-called ankle region. Beyond this point, detected cosmic rays are supposed to come from extra galactic objects.

1.2.3.1. Galactic and extra-galactic sources

Several candidates meet the needed criteria to accelerate cosmic rays at ultra-high energy, such as type II supernovae, pulsars and shock accelerations in supernova remnants for the galactic sources and gamma-ray bursts (GRB, brief outbursts of MeV gamma-rays, which can be originated from the merger of binary neutron stars such as for the detection of the gravitational wave GW170817 event associated to the gamma ray burst GRB 170817A detected by the LIGO collaboration [22]) plus AGN for extragalactic sources. An AGN consists in a black hole of about 10^6 to 10^{10} solar masses ejecting a matter beam at 99 % of the speed of light. A magnetic field at the centre of an AGN could reach 5 G in volumes of linear dimension of 0.02 pc. However, the energy losses in very dense radiation field in the centre of an AGN is very large. This means that proton escaping from the center of an AGN will probably loss a lot of energy. Nevertheless, several sources are able to accelerate protons above 10^{20} eV such as neutron stars, AGN, lobes of giant radio galaxies (so-called hot spots which are the termination shock of the jet in its propagation in extragalactic media), gigaparsec shocks in the intergalactic medium (from structure formation due to gravitational attraction), GRB, as it is shown on the Hillas diagram.

As I wrote previously, to understand well the cosmic ray creation processes, the characteristics of the potential sources have to be known, but we also need to understand the acceleration mechanisms which permit cosmic rays to reach those huge energies.

There are several acceleration mechanisms, such as one-shot and diffusive accelerator (see [23] for more details), but the Fermi mechanisms, particularly the first order mechanism, are the most considered. In the following, I describe the Fermi mechanisms, by beginning with the second order first, which was historically the first proposed by Fermi to understand the problem of the origin of the cosmic ray.

1.2.3.2. Fermi acceleration

In 1949, Enrico Fermi proposed a model for acceleration mechanism of cosmic rays [24], known as the second-order Fermi acceleration. It explains the acceleration of relativistic charged particles by means of their collision with interstellar clouds. These clouds can be assimilated to magnetic mirrors with moving in random directions, and the charged particles are reflected by them (see figures 1.15 and 1.16 left). Let us consider a charged particle with a velocity vector $\overrightarrow{v}_{\text{in}}$ interacting with a magnetic cloud having a velocity vector $\overrightarrow{V}$. The charged particle encounters the magnetic cloud with a pitch angle $\theta_{\text{in}} = \left(\widehat{\overrightarrow{v}_{\text{in}}, \overrightarrow{V}}\right)$ measured in the galactic frame. The pitch angle of the charged particle after the interaction is $\theta'_{\text{out}} = \left(\widehat{\overrightarrow{v}_{\text{out}}, \overrightarrow{V}}\right)$. Here, the prime symbol stands for the quantities expressed in the magnetic cloud frame. One can write:

$$E'_{\text{in}} = \gamma E_{\text{in}}(1 - \beta \cos\theta_{\text{in}}) \tag{1.5}$$

$$E_{\text{out}} = \gamma E'_{\text{out}}(1 + \beta \cos\theta'_{\text{out}}) \tag{1.6}$$

As shown in figure 1.15, inside the magnetic cloud, the charged particle does not undergo a change of its energy, meaning that $E'_{\text{in}} = E'_{\text{out}}$. This gives:

$$E_{\text{out}} = \gamma^2 E_{\text{in}}(1 - \beta \cos\theta_{\text{in}})(1 + \beta \cos\theta'_{\text{out}}) \tag{1.7}$$

$$\frac{\Delta E}{E} = \frac{\beta(\cos\theta'_{\text{out}} - \cos\theta_{\text{in}}) + \beta^2(1 - \cos\theta'_{\text{out}}\cos\theta_{\text{in}})}{1 - \beta^2} \tag{1.8}$$

We obtain the mean energy gain for a charged particle after interacting with the magnetic field by calculating the mean of the cosine of the angle of entry and exit of the charged particle. We make the hypothesis that the direction of the charged particle after diffusion in the cloud is isotropic, giving $\langle\cos\theta'_{\text{out}}\rangle = 0$. Concerning $\langle\cos\theta_{\text{in}}\rangle$, its distribution is non-uniform and depends on the relative velocity of the charged particle and the magnetic cloud, as depicted in figure 1.15. The mean number of encountered magnetic clouds by the charged particle during Δt is proportional to $(v - V\cos\theta)\Delta t$. From that, $\langle\cos\theta_{\text{in}}\rangle$ is distributed according to the following statistical law $P(\cos\theta_{\text{in}}) \propto (v - V\cos\theta_{\text{in}})$, and assuming that the charged particle is relativistic:

$$\langle\cos\theta_{\text{in}}\rangle = \frac{\int_{-1}^{1}\cos\theta_{\text{in}}(v - V\cos\theta_{\text{in}})\mathrm{d}(\cos\theta_{\text{in}})}{\int_{-1}^{1}(v - V\cos\theta_{\text{in}})\mathrm{d}(\cos\theta_{\text{in}})} = \frac{-2V/3}{2v} \backsimeq -\frac{1}{3}\beta \tag{1.9}$$

with $\beta = \frac{V}{c}$. Using Equation 1.8, we obtain:

$$\langle\frac{\Delta E}{E}\rangle = \frac{4}{3}\left(\frac{V}{c}\right)^2 \tag{1.10}$$

where V is the speed of the cloud and c is the speed of light and also the velocity of the particle (since the particle is relativistic). The energy gain is proportional to the second term $\left(\frac{V}{c}\right)^2$. Moreover, we add that the number of head-on collisions is greater than head-tail following collisions in the galactic frame, leading, by mean, to an acceleration of

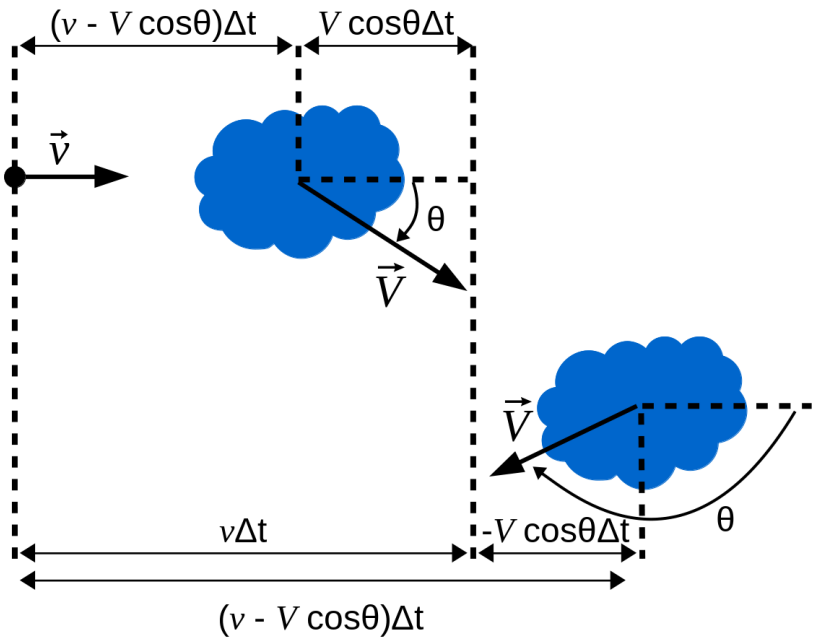


Figure 1.15 *– Sketch for the calculation of the stochastic distributions of the pitch angle of a charged particle in a magnetic cloud.*

the charged particle. By solving diffusion-loss equations in the steady state, the energy spectrum for such acceleration mechanism can be expressed as:

$$N(E)\mathrm{d}E = A \times E^{1+\frac{1}{\alpha\tau_{escape}}}\,\mathrm{d}E \tag{1.11}$$

where A is a constant. However, this mechanism cannot explain the spectral index of ~ 2.7, which is determined by the factor $\alpha\tau_{escape}$. The second order mechanism of Fermi acceleration has another limitation: the typical velocity of a magnetic cloud is $\beta \backsimeq 10^{-4}$, meaning that the mean relative energy gain is around 10^{-8} per collision. Considering the typical distance between two magnetic clouds, estimated to be equal to ~ 1 pc, a charged particle needs one billion years to triple its energy (the age of cosmic ray is estimated to 10^7 years [25]).

To overcome this problem, Bell introduced in 1978 the first-order Fermi mechanism, aiming to obtain an energy gain linear with $\frac{v}{c}$, condition for what the acceleration process is more effective at high v. A high v occurs when the relativistic charged particles collide with strong shock waves (which replace the magnetic cloud of the second order mechanism) reaching supersonic velocities, noted β_{sh}, like in supernova explosions or AGN for instance. As shown in figure 1.16 right, the relativistic charged particle can cross the strong shock waves on both sides: the undisturbed side noted upstream and the shocked side noted downstream. Proceeding and reasoning in the same way as for the second-order Fermi mechanism, the prime symbol stands for the quantities expressed in the downstream frame, according to the equation 1.5. So, the charged particle is first in the upstream with an energy E_{in}. It crosses the shock wave and passes in the downstream with an energy E'_{in}. It diffuses in the downstream and has a certain probability to cross again the shock wave, and therefore is back in the upstream side with an energy E_{out}. As for the case of the second order mechanism, $E'_{\mathrm{in}} = E'_{\mathrm{out}}$ and the equation 1.6 is always valid, but with $\beta = \beta_{\mathrm{rel}}$ corresponding to the relative velocity of the downstream compared to the upstream. This process can occur several times due to the magnetic irregularities near the shock wave inducing magnetic scattering processes. We obtain the mean energy gain for a charged particle after interacting with the shock wave by calculating the mean

of the cosine of the angle of entry and exit of the charged particle. We assume that the angular distribution of the charged particles in upstream is isotropic. By such similar calculations than for the second order mechanism, we obtain:

$$\langle \cos\theta_{\text{in}} \rangle = -2/3 \tag{1.12}$$

$$\langle \cos\theta'_{\text{out}} \rangle = 2/3 \tag{1.13}$$

with θ_{in} the cross angle from upstream to downstream and θ'_{out} the cross angle from downstream to upstream. The main difference with equation 1.9 is that the pitch angle is now independent of the charged particle and shock velocities. Using the equations 1.8 and 1.12, we obtain an average energy of:

$$\langle \frac{\Delta E}{E} \rangle = \frac{4}{3}\beta_{\text{rel}} \tag{1.14}$$

We see that the charged particle has an energy gain proportional to the relative velocity of the downstream compared to the upstream. Using the shock velocity ratio noted $r = \beta_{\text{u}}/\beta_{\text{d}}$, we obtain:

$$\langle \frac{\Delta E}{E} \rangle = \frac{4}{3}\frac{r-1}{r}\beta_{\text{sh}} \tag{1.15}$$

and see that the charged particle has an energy gain proportional to the velocity of the shock wave. The energy spectrum obtained for this mechanism is:

$$N(E)\mathrm{d}E = A \times E^{-2}\mathrm{d}E \tag{1.16}$$

The typical velocity of a shock wave from a supernova gives $\beta_{\text{sh}} \backsimeq 10^{-2}$, which is a million time greater than the β velocity of the magnetic cloud of the second order of Fermi mechanism. Moreover, considering the typical time between two interactions of the charged particle and the shock wave, the acceleration time is estimated to $\sim 10^6$ s ($\sim$ 12 days), to be compared to one billion years in the case of the second order mechanism. Finally, the spectral index of equation 1.16 is obtained for strong shock wave ($r \to 4$), and is close to the value of 2.7 experimentally obtained.

Using equation 1.3 and adjusting the parameters of the astrophysics sources, we can estimate which sources are candidates to accelerate cosmic rays to ultra-high energy, as depicted in figure 1.14.

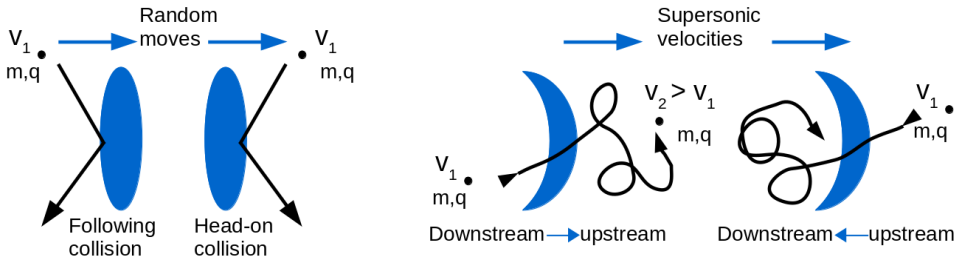


Figure 1.16 *– Left, second-order Fermi acceleration mechanism. Right, first-order Fermi acceleration mechanism.*

1.2.4. Flux suppression: End of cosmic ray spectrum

In figure 1.17, showing the redressed cosmic ray spectrum (the flux is multiplied by $E^{2.6}$), the results of HiRes1 and 2, AugerSD and Telescope Array at the end of the energy scale are in a very good agreement with the energy cut-off theory proposed by Greisen, Zatsepin and Kuzmin, as explained below. In contrast, Agasa experiment observed an

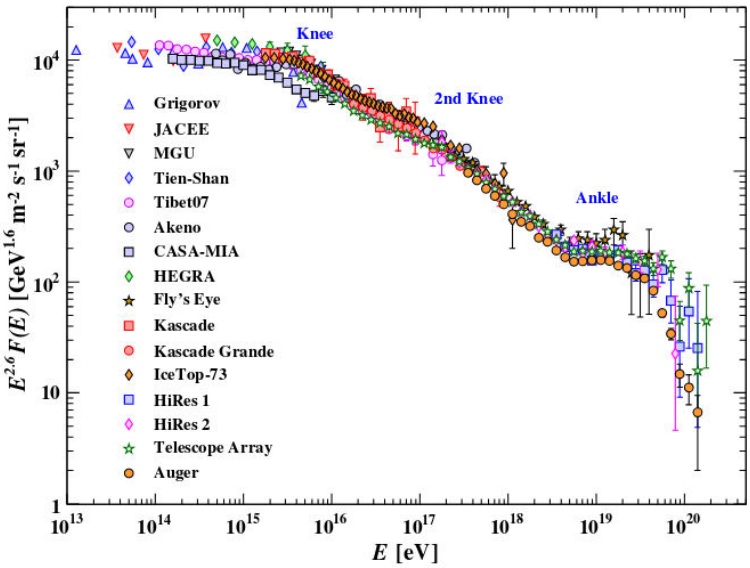


Figure 1.17 *– Redressed cosmic ray spectrum. Adapted from [28].*

increase of the flux above 10^{20} eV [26]. Nowadays, the existence of a cut-off in the cosmic ray flux is well established [27], but not clearly understood. The first observation of a shower with an energy of 10^{20} eV was carried out in 1963 by the Volcano Ranch array, and reported by John Linsley. We recall that the gyro radius of a 10^{20} eV proton is of the same order than the Milky Way radius, and so this kind of protons have to be accelerated by more powerful astrophysical objects than the Milky Way. Today, a serious candidate to explain this cut-off is the Greisen-Zatsepin-Kuzmin (GZK) (physicists at the origin of this prediction [29, 30]) limit, but it could be also due to the limit of the acceleration mechanism or an extinction of sources.

1.2.4.1. GZK limit and cosmic microwave background

The cosmic microwave background (noted CMB) was discovered by Penzias and Wilson, and is the strongest proof of the big-bang and nucleosynthesis. It has a blackbody spectrum of 2.75 ± 0.002 K. Figure 1.18 shows the energy density of the CMB in units $\text{eV} \cdot \text{cm}^{-3}$. The CMB is universal and isotropic, and because of the expansion of the Universe, its temperature increases with the redshift z (increase in wavelength of the electromagnetic radiation). The main idea of the GZK limit is the interaction of the cosmic ray with the cosmic microwave background. Indeed, in [30], Greisen estimates that in addition to the photoproduction interactions, lower energy protons will interact with the CMB producing e^+e^- pairs and this pair creation would peak at 3×10^{19} eV, the so-called GZK limit. The GZK process for protons is as follows:

$$p\,(\text{E} > 10^{19.5}\,\text{eV}) + \gamma_{CMB} \rightarrow \Delta^+ \rightarrow \pi^+ + n \tag{1.17}$$

The produced neutron will decay by β^- in a proton giving the emission of a muon-anti-neutrino, and the pion will decay in μ^+ with an emission of muon-neutrino, and the μ^+ will decay in a positron giving a muon-anti-neutrino and an electron-neutrino. As a reference, a proton of 10^{20} eV loses as an average 12 % of its energy every 6 Mpc by interacting with the photons of the CMB. From that, we deduce that 90 % of the flux of protons at 60 EeV should come from distances smaller than 200 Mpc. For protons of 80 EeV, their sources should lie at a maximum distance of 90 Mpc. As a reminder, the Milky Way has a radius of 20 kpc and a thickness of 200 – 300 pc. The closest galaxy is the

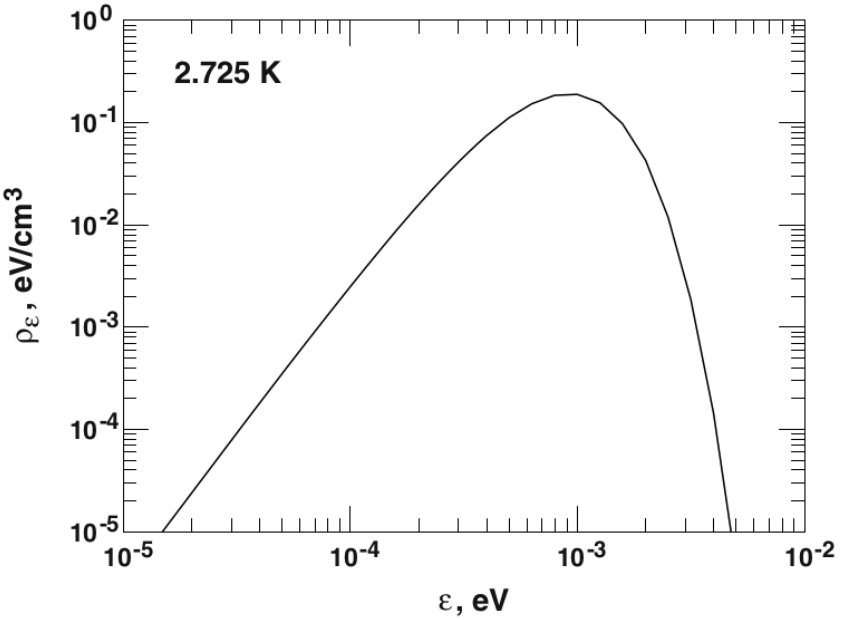


Figure 1.18 *– Energy density of the microwave background. Adapted from [17].*

Andromeda galaxy, situated at 780 ± 17 kpc.

The GZK limit also exists for heavy nuclei, which lose energy by photodisintegration:

$$A_N + \gamma_{CMB} \rightarrow A_{N-1} + n \tag{1.18}$$

The produced neutron will decay β^- in a proton giving the emission of a muon-anti-neutrino. With the same consideration, we can show that the limits in which the sources of heavy nuclei cosmic rays (iron nuclei for instance) should lie are approximately the same than for protons, and that the limits in which the sources of intermediate nuclei cosmic rays (silicon nuclei for instance) should lie are really smaller. With these considerations, we deduce that only heavy nuclei such as iron nuclei can be produced by sources at consistent distance.

1.3. Anisotropy and chemical composition of cosmic rays

Both anisotropy and composition plus the study on the cosmic-ray spectrum are mandatory to elucidate, decipher their origin and nature. It is worth noticing that the composition of the cosmic rays is also important for the understanding of the propagation and acceleration processes described in section 1.2.3, and thus for the sources. For instance, at low energy (from few MeV up to TeV), the composition of the cosmic-rays are interesting for the study of the dark matter components or for constraining parameters of super-symmetric theory [31, 32]. At higher energy, the composition of cosmic-rays is not well known, and is currently actively studied. We present in this section the current results on the anisotropy and the composition of cosmic rays. The update of the results presented during the *2019 International Cosmic Ray Conference* are not included and can be found in [33, 34, 35].

1.3.1. Anisotropy in the arrival directions of cosmic rays

1.3.1.1. Strong anisotropy at large angular scales – Auger results

As mentioned in [36], the Pierre Auger Observatory reported in 2017 a large-scale anisotropy in the arrival direction of cosmic rays. As we will see in the next chapter,

above 10^{14} eV, the cosmic ray arriving in the atmosphere will induce an extensive air-shower (EAS) which can reach the ground. During its development, some phenomena occur, as described in the next section, allowing us to reconstruct the arrival direction of the EAS, and thus the arrival direction of the primary cosmic ray. The Pierre Auger Observatory [37] is located in Argentina, and is the largest cosmic ray observatory in the world (covering an area of ~ 3000 km^2). It is an hybrid observatory, gathering different observation techniques, such as particle detectors (1,600 water-Cherenkov detectors), fluorescence telescopes, muon counters and radio antennas. As a reminder, cosmic rays are charged particles with an extraterrestrial origin. From that definition, we understand that the localisation of the sources of cosmic rays is difficult. Indeed, the cosmic ray, once produced by an astrophysical object, will interact with the magnetic fields present in the inter and extra-galactic medium and will be deflected before reaching the atmosphere of Earth. The deflection is linked to the gyro-radius r_g which is:

$$r_g = \frac{E}{BZ}\,[\text{pc}] \tag{1.19}$$

where E is the energy of the cosmic ray, B is the magnetic field present in the medium and Z the atomic number of the cosmic ray. It is worth noticing that for cosmic rays with an energy above 8 EeV, the gyro-radius of a proton is larger than the Milky Way dimensions for typical values of magnetic field (µG). This seems to indicate a probable extra-galactic origin. Indeed, there is no excess in the arrival of cosmic ray in the direction of the galactic center. Moreover, by considering the magnetic rigidity explained as:

$$R = r_g \times B \tag{1.20}$$

where r_g is the gyro-radius and B the magnetic field, we see that the highest energy cosmic rays could have a magnetic rigidity sufficient to compensate the magnetic deflection. For example, protons ($Z = 1$) with an energy above 5×10^{19} eV might point to their sources. To search for an anisotropy in the arrival directions, an analysis of non-uniformities in the distribution of the arrival directions of the event is made. By a Rayleigh analysis in right ascension (astronomical coordinate) for cosmic rays with an energy above 8 EeV, the Pierre Auger Observatory found [36] an anisotropy at large angular scales, incompatible at a level of 5.2σ with an isotropic distribution, as shown in figure 1.19. This anisotropy (233°,−13°) is far away from the galactic center (0°,0°). This signs an extra-galactic origin for these cosmic rays. In figure 1.19, the observed dipole at Earth does not coincide with the dipole measured by the 2 Micron All-Sky Redshift Survey (2MRS) [38] (55° away from the dipole measured by the 2MRS). For an extra-galactic observer, the direction of the dipole observed at Earth is deflected by the magnetic field present in the galaxy [36]. The arrow denotes the prediction of the deflection of a dipolar distribution of cosmic rays in the same direction than the 2MRS dipole (i.e extra-galactic) by a model of galactic magnetic field. This prediction is consistent with the dipole measured on Earth.

1.3.1.2. Moderate anisotropy at intermediate angular scales – Auger results

As explained in [39], the Pierre Auger Observatory also found an anisotropy at intermediate angular scales, hinting an excess of arrivals from nearby sources. The bulk is to study if the directions corresponding to excess of cosmic rays can be associated with extragalactic sources detected by Fermi-LAT [40] (4 catalogues), namely the AGN and the starburst galaxies (SBG). They reported that an isotropy of cosmic-rays is disfavoured with a 4σ confidence level for the SBG after corrections, see figure 1.20 left. There are also hints of excess of cosmic ray in the direction of the AGN Centaurus A above 58 EeV where 19 events over 203 have been seen, when only 6 were expected, as shown in figure 1.20 right.

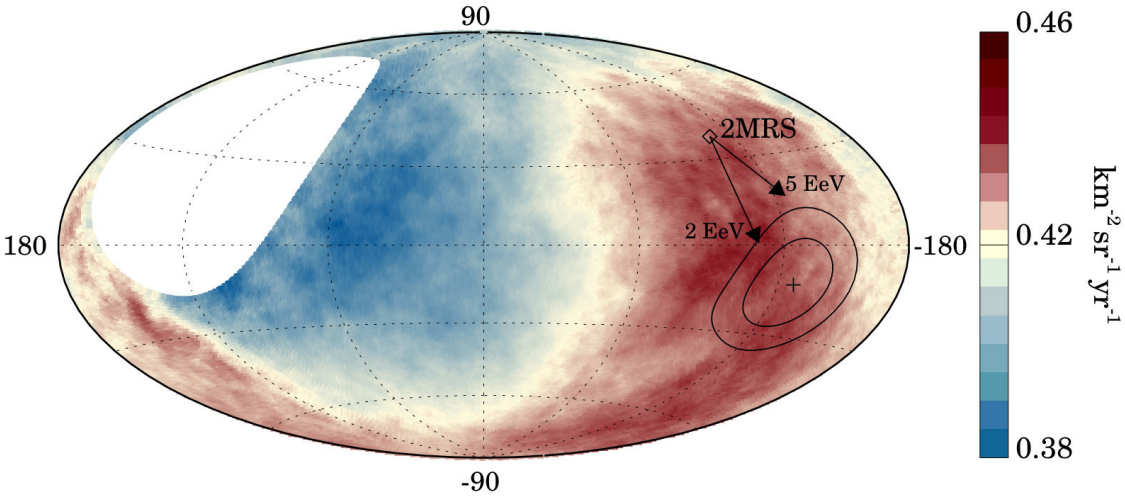


Figure 1.19 *– Distribution of the arrival directions of the cosmic rays above* 8 EeV *in Galactic coordinates [36]. The galactic center is at the center of the figure, The cross indicates the measured dipole direction at Earth, the contours represent respectively the confidence levels of* 68 % *and* 95 %.

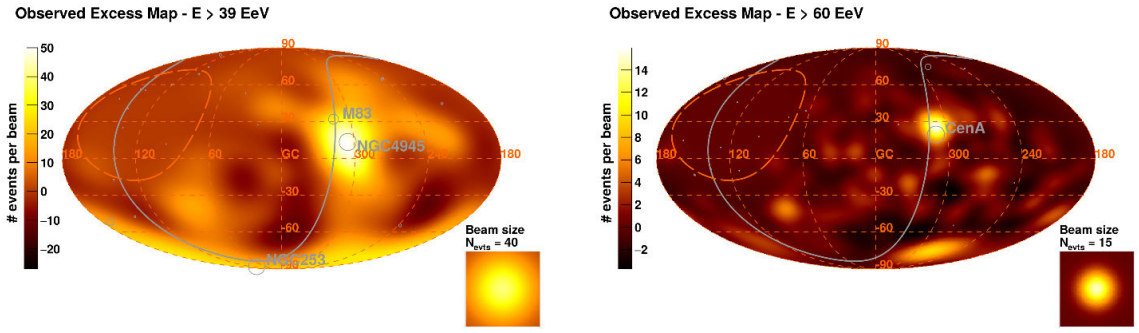


Figure 1.20 *– Observed excess maps, for the best-fit parameters obtained with SBGs above 39 EeV (left) and AGNs above 60 EeV (right). The color scale indicates the number of events per smearing beam. The supergalactic plane is shown as a solid gray line. An orange dashed line delimits the field of view of the array. From [39].*

1.3.1.3. Anisotropy at intermediate angular scales – Telescope Array results

Telescope Array (TA) [41] is the largest hybrid cosmic ray observatory of the northern hemisphere, and is in operation since 2007. It is located in Utha, USA. It covers an area of $\sim$ 700 km^2 and it is composed of 507 scintillator detectors and 36 fluorescence telescopes dispatched on 3 sites: Middle Drum, Black Rock and Long Ridge. TA also reports an intermediate-scale anisotropy [42] of cosmic-rays above 15.8 EeV. As for Auger, they search correlations between the directions of arrival of the cosmic-rays and the Large-Scale Structure (LSS) model and the isotropic model. For an energy above 57 EeV, the directions of arrival of the cosmic-rays (9 years of data, 72 events) are compatible with the LSS model, and incompatible with the isotropic model. For the 72 events, and at an angular scale of 25°, there is an indication of a hotspot with a 5σ level of significance, as shown in figure 1.21. They also tried to find correlation with the SBG catalogue as Auger by applying the Auger best-fit parameters on the TA data. There is no significant excess so far.

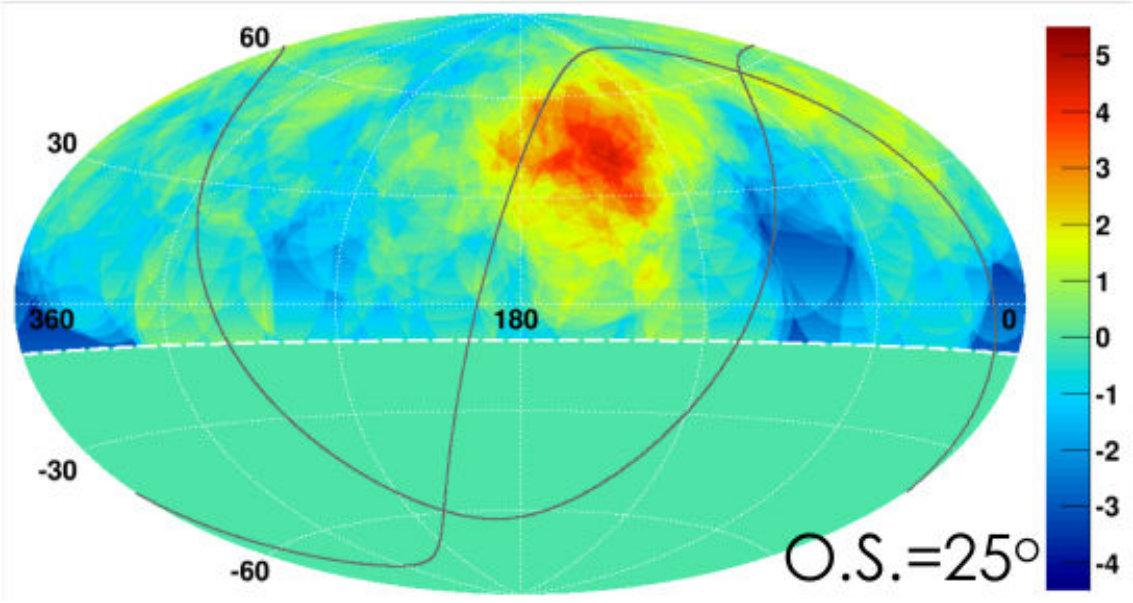


Figure 1.21 *– A significance map of UHECR events. The events are smoothed by* 25° *oversampling radius circle (O.S.) which was determined to maximize the significance. From [43].*

1.3.2. Composition of cosmic rays arriving on Earth

1.3.2.1. Auger composition

By statistical studies, the mass composition of cosmic-rays arriving on Earth can be deduced from the shower maximum $X_{\max}$ (see Chapter 2 for a complete description of the shower phenomenon and the definition of $X_{\max}$), shown in figure 2.2. The results of the Pierre Auger Observatory are shown in figure 1.22 left. The figure compares the mean values of the $X_{\max}$ distributions measured by the surface detectors of the Pierre Auger Observatory with the predicted $X_{\max}$ distributions [44]. The predictions for iron nuclei and protons are represented by the blue and red lines respectively, and for two hadronic interaction models. Data indicate heavier composition above 3 EeV. This increase in mass composition at 3 EeV corresponds to the ankle region where the slope of the energy spectrum changes from −3.3 to −2.5 (see figure 1.22 right). From that, and due to the rigidity expressed in equation 1.19, particle astronomy will be complicated for heavy particles such as iron nuclei beyond the GZK cutoff, while it should have been possible for lighter particles such as protons (protons with an energy above 5×10^{19} eV might point to their sources, as said in section 1.3).

1.3.2.2. TA composition

The results for mass composition measurements made by TA are shown in figure 1.23. The coloured points represent the estimation made with the QGSJET hadronic model for different nuclei. BR/LR hybrid is for measurements made by Black Rock and Long Ridge fluorescence detector (FD) stations plus surface array, BR/LR/MD stereo is for measurements made by FD-only using multiple coincidences of 2 or more FD stations, MD hybrid is for measurements made by the Middle Drum FD station plus surface array. Below 10^{19} eV, FD measurements are consistent with a light composition (proton or He). For this energy range, TA has a sufficient statistics to make a confident measurement of the cosmic-ray composition. However, above 10^{19} eV, the flux decreases leading to the fall of the statistics. Moreover, a zenith angle dependence appears in the energy spectrum, due to the maximum atmospheric mass overburden which limits the penetration of the

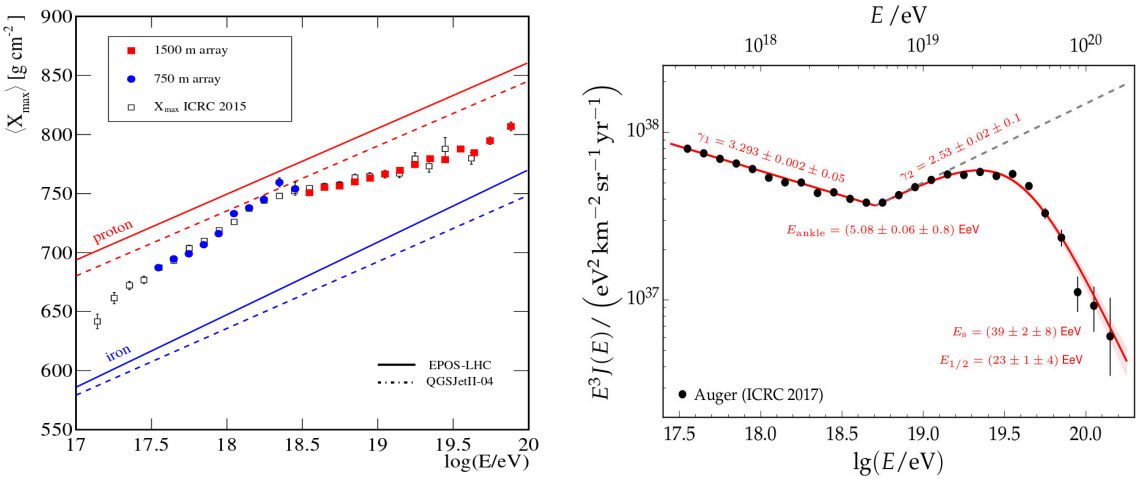


Figure 1.22 *– Left: Evolution of* $\langle X_{\max} \rangle$ *as a function of energy. The figure compares the mean values of the* $X_{\max}$ *distributions measured by the surface detectors of the Pierre Auger Observatory with the predicted* $X_{\max}$ *distributions. The blue points are the measurement of the* 750 m *spacing array and the red squares the measurements of the* 1500 m *spacing array [45]. The predictions for iron nuclei and protons are represented by the blue and red lines respectively. From [44]. Right: Redressed cosmic ray spectrum. From [46].*

light in the telescope. There is no hint for the composition since TA can not distinguish proton, helium and nitrogen, as depicted in figure 1.23, presenting the evolution of $\langle X_{\max} \rangle$ as a function of energy.

1.3.3. Compatibility of Auger and TA results on energy spectrum and composition

Since ICRC 2010 in Nagoya, Japan, a working group aims to compare the results obtained by the Pierre Auger and Telescope Array Observatories. The latest results are promising. Indeed, recent reports on the energy spectrum [27, 48, 49] establish that the measurements of the two observatories are in good agreement after a rescaling of the energies measured with the surface detectors. However, after a restriction to a common declination band ($-15.7\,° < \delta < 24.8\,°$), statistical differences are still present. There is no declination dependence for the Auger spectrum, while a 3.9σ declination dependence for TA spectrum is present.
Concerning the latest results on the composition [50], compatibility tests were carried out and showed that there is a very good agreement between the $X_{\max}$ measured by the two experiments. The $X_{\max}$ measurement by TA is also compatible to a mixed composition as measured by Auger. However, the TA data are also consistent with a light composition above $10^{18.2}$ eV. A larger set of TA data is needed to get a final result on the composition [50, 51, 52].

***

In this chapter, we partly study the theory of cosmic rays. As mentioned in the preamble, the objective was to gather the main current knowledge on cosmic rays, from their discovery to the main results known at the time of writing this manuscript. In the next chapter, we will present the detection methods used for the study of ultra-high energy cosmic rays. Indeed, above 10^6 GeV, the cosmic ray flux becomes very low (1 particle per

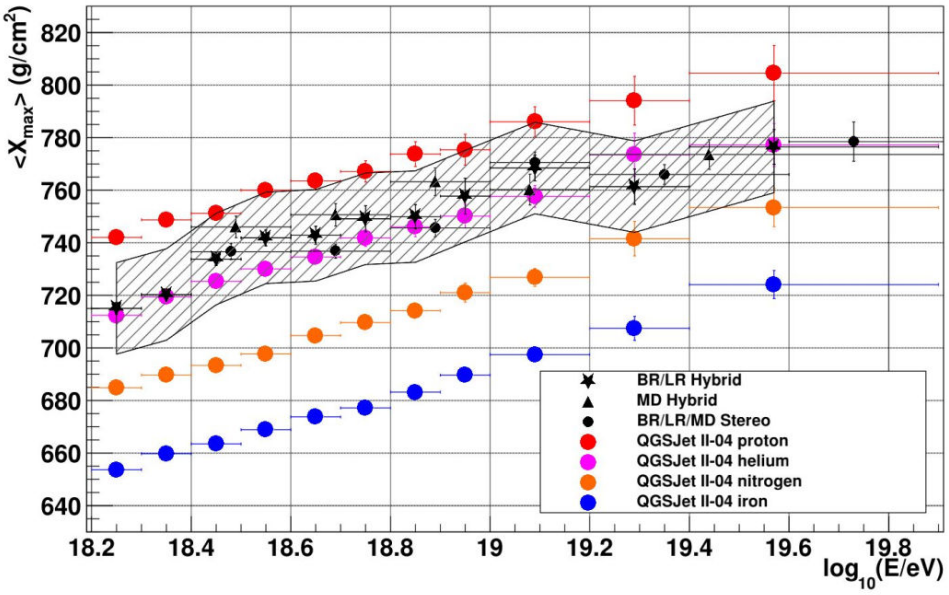


Figure 1.23 *– Evolution of* $\langle X_{\max} \rangle$ *as a function of energy. The coloured points not in black represent the estimation made with the QGSJET hadronic model for different nuclei. BR/LR hybrid is for Black Rock and Long Ridge FD stations plus surface array, BR/LR/MD stereo is for FD-only measurement using multiple coincidences of 2 or more FD stations, MD hybrid uses the Middle Drum FD station and surface array. The systematic uncertainty on the data is* $17\ \mathrm{g \cdot cm^{-2}}$ *(black band). Taken from [47].*

square metre per year), and the study of cosmic rays is possible through the detection of associated air showers, i.e through indirect detection. In the following chapters, the part of cosmic ray spectrum which is studied is above 10^6 GeV. Chapter 2 is dedicated to the Extensive Air Showers (EAS) phenomenon, which is at the centre of the preoccupations of the studies carried out during this PhD thesis.

Indeed, the indirect study of cosmic rays has the main consequence of losing sight of the object of study: the cosmic ray. Thus, it is not the primary cosmic ray that is studied, but the EAS induced by the primary cosmic ray after interacting with the components of the atmosphere. It must be kept in mind that the ultimate goal is to characterize the primary cosmic ray that generated the EAS in order to answer these questions: what is the origin of this extra solar system matter? How does cosmic ray matter get accelerated to such high energies? How do the particles propagate from source(s) to Earth? Are there signatures of any exotic physics? The EAS thus becomes the main object of study, to allow the characterization of the primary cosmic ray. But how? This is the purpose of the second chapter.

"By Toutatis, the sky will fall on our heads!"

Asterix – The Gaul

CHAPTER 2

Cosmic Ray Induced Extensive Air Showers

Outline

Parts of this chapter have been presented in URSI-France 2019 conference, and published in [53].

Ultra-high energy cosmic ray (UHECR) physics refers to the experimental study of cosmic rays above 10^{14} eV. Above this limit, direct measurements of cosmic rays with balloons or satellites are limited by their small collecting area regarding the extremely low flux of particles (down to 1 per km^2 per century above 10^{20} eV). Instead, one has to detect the cascade of secondary particles, called extensive air shower (EAS), which follows the interaction of the cosmic rays with the Earth's atmosphere and spread at ground over very large areas. The interaction of a cosmic ray with a nucleus of the atmosphere can occur at energies much higher than that which one can reach in the laboratory (up to 800 times higher than the 13 TeV in the center of mass in proton-proton collisions in the LHC at CERN), making them also specially interesting to refine our view on the standard model of particle physics. The primary particle characteristics (energy, arrival direction, nature) are then indirectly inferred from the EAS measurements: detected ground particles, shower front arrival time at the detector locations and the longitudinal profile of the shower which is the distribution of the number of particles in the shower over time along the shower axis. The experimental challenge stands in the ability to determine these 3 quantities with a sufficient level of accuracy [54]. EAS are commonly observed by ground-based particle detectors and by telescopes that observe the fluorescent light emitted by the atmospheric di-nitrogen molecules excited by the passage of the cascade.

2.1. Extensive air showers

2.1.1. Discovery of extensive air shower

As shown in figure 1.13, above 10^6 GeV, the cosmic ray flux is very low (1 particle per square metre per year), and their direct detection becomes improbable. Hopefully, high energy charged particles can interact with the atmosphere components and generate billions of secondary charged particles called "air showers". In the 30's, Skobelstyn, who employed for the first time the term *shower particles*, observed the arrival of cosmic ray clusters. In 1932, Rossi observed simultaneous discharges in two Geiger counters separated by several meters. He concluded that there was a low probability that it was by chance coincidence, and so that groups of particles arrived at the same time in the apparatus. Rossi decided to use a cloud chamber equipped with different lead layers. The shower intensity (number of particles in the shower) first increases with the lead density to reach a maximum and then decreases as the shower dies out, as shown in figure 2.1. Anderson proposed a first interpretation of the shower composition. Armed with the

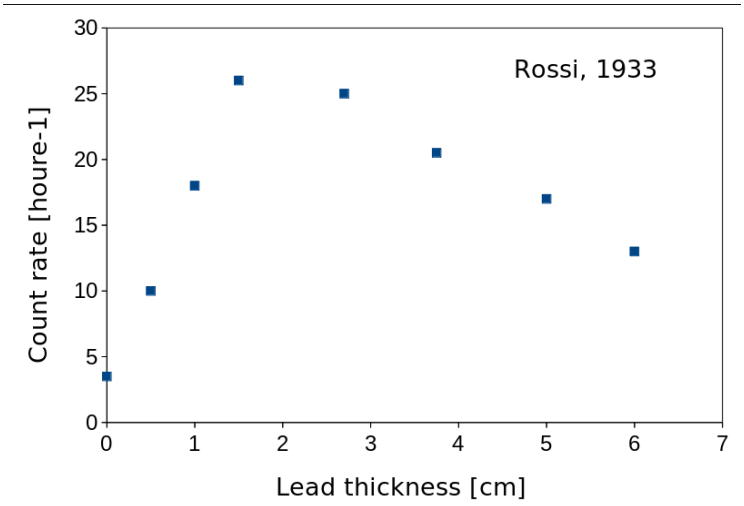


Figure 2.1 *– Counting rate of cosmic rays in function of lead thickness. The counting rate first grows with the lead density to reach a maximum and then decreases as the shower dies out. A large thickness of lead corresponds to a low atmosphere altitude. Data extracted from [55].*

discovery of the positron in 1932, he argued that a chain reaction occurred: the e^+e^- pair creates a very energetic photon by Bremsstrahlung (see section 2.1.2.1), which creates a new e^+e^- pair by pair creation process (see section 2.1.2.2). This interpretation will be confirmed by Pierre Auger and collaborators in the end of the 30's [56], first at the Observatoire du Pic du Midi and then at the scientific station of Jungfraujoch, where two distant apparatus have both detected particle clusters at the same time. He showed that the coincidence rate increases, at fixed time window, when the detectors are closer. He concluded that these particle clusters came from the same event: a shower of particles.

2.1.2. Relevant electromagnetic processes for shower development

We focus in this section on the relevant electromagnetic processes for shower development in the atmosphere: Bremsstrahlung by electrons and positrons, pair production by photons, Coulomb scattering, energy loss of electrons by ionization and Compton processes. We consider only high energy particles with an energy above 1 GeV, and we only discuss the electromagnetic interactions of particles, particularly for electrons, positrons and photons which compose the electromagnetic component of showers [17, 57, 58].

2.1.2.1. Bremsstrahlung

Bremsstrahlung, also known as braking radiation or Coulomb bremsstrahlung (in contrast to magnetic bremsstrahlung), is a radiation emitted by the deceleration of a charged particle. The energy losses can be expressed as:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{N_A}{A}\int_0^{E-mc^2} \sigma_{br}(E,k)k\mathrm{d}k \tag{2.1}$$

where N_A and A are respectively the Avogadro's number and the mass number, m and E are respectively the mass and the energy of the charged particle, r_e is the classical radius of the electron ($r_e = \frac{e^2}{4\pi\epsilon_0 m_e c^2} = 2.81794$ fm) and k the energy of the emitted photon. The bremsstrahlung cross section is expressed as:

$$\sigma_{br} = \frac{4Z^2\alpha r_e^2}{k}F(E,k) \tag{2.2}$$

where Z is the number of charges, α is the fine-structure constant and the function F is relative to the screening of the atomic electrons. The energy loss after integration is given by:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{4N_A Z}{A}\alpha r_e^2 E\left[\ln 184.15Z^{-1/3} + 1/18\right] \tag{2.3}$$

The last term (1/18) represents the different interactions of the projectile with the fields of the atomic electrons.

2.1.2.2. Creation of electron-positron pairs

Physically, the e^+e^- pair creation is the opposite of the Bremsstrahlung, and is also known as Coulomb pair production (in contrast to magnetic pair production). This is the creation of an electron and a positron by the interaction of a wave of energy or photon (packet of energy) with a nucleus. The energy threshold for the pair creation is equal to 1.052 MeV (corresponding to the sum of the rest mass energies of the electron-positron pair). The cross section of pair production in air is given by:

$$\sigma_{air}(k,E) = \sigma_{br}(k,E)\frac{E^2}{k^2} = \frac{4Z^2\alpha r_e^2}{k}G(k,E) \tag{2.4}$$

with k the energy of the incident photon (and $k = h\nu$), E the energy of the e^- or the e^+. The function G is relative to the screening of the atomic electrons. By integrating over energy, and considering no screening effect, the total pair production cross section in air can be expressed as:

$$\sigma_{air}(k) = \int_{mc^2}^{k-mc^2} \sigma_{pp}(E,k)\mathrm{d}E = 4Z^2\alpha r_e^2\frac{\ln(191Z^{-1/3})}{9} - \frac{1}{54} \tag{2.5}$$

where, as for Bremsstrahlung, the last factor (1/54) represents pair production in the fields of the atomic electrons. In the case of EAS, the medium is the atmosphere, and the cross section is equal to $\sim$ 500 mb.

2.1.2.3. Coulomb scattering

We have seen with the equation 1.1 that the force between two point charges is the founding principles of the electromagnetism. The Coulomb scattering was experimented by Ernest Rutherford, who bombarded gold nuclei with alpha particles: this is an example of "elastic scattering" because the energy and velocity of the deflected particle are the same as for the projectile particle. The deviation angle of the alpha particle is given by:

$$\tan\frac{\theta}{2} = \frac{zZe^2}{Mv^2 b} \tag{2.6}$$

where θ is the deviation angle between the initial trajectory and the final trajectory of the alpha particle, b the impact parameter, v and z are respectively the velocity and the atomic number of the projectile and Z, M are respectively the atomic number and the atomic mass of gold nucleus of the target, e is the charge of the electron. If we consider an electron as projectile, the differential cross-section can be expressed as:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{b}{\sin\theta}\frac{\mathrm{d}b}{\mathrm{d}\theta} \tag{2.7}$$

When the projectile is deflected, an energy transfer occurs between the projectile and the target. One can conclude that, with this process, the energy and direction (impulsion) of the projectile are modified. This equation is only valid for point like, non relativistic projectiles. A correction for a relativistic projectile is given by Mott, adding the term $1-\beta^2\sin^2\frac{\theta}{2}$ to the numerator of the equation 2.7. This process is the main one responsible for the lateral spread of the shower front disk (the thin particle disk which evolves during the shower development). One has to know that in a dense medium, 10^7 to 10^8 scattering events per g $\cdot$ cm^{-2} are occurring. The mean square scattering angle in an infinitesimal layer of thickness dx [cm] can be expressed as:

$$\langle\theta^2_{\mathrm{scatt}}\rangle\mathrm{d}x = 16\pi N_A\frac{Z^2}{A}r_e^2\frac{m_e^2c^4}{p^2\beta^2c^2}\ln\left(184.15Z^{-1/3}\right)\mathrm{d}x \tag{2.8}$$

where p and β are respectively the momentum and velocity of the projectile, N_A the Avogadro's number, Z and A are respectively the atomic and mass number of the scattering center.

2.1.2.4. Ionization loss

This process is well described by the Bethe formula. A charged particle crossing a medium loses energy by ionization and excitation of the atoms of the medium. These processes are present whenever a charged particle is propagating in a medium. This energy loss per unit of column depth can be expressed in MeV $\cdot$ cm^{-2} by integrating the Bethe and Bloch formula to the maximum energy loss W:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = -\frac{N_AZ}{A}\frac{2\pi(ze^2)^2}{Mv^2}\left[\ln\frac{2Mv^2\gamma^2W}{I^2} - 2\beta^2\right] \tag{2.9}$$

where Z is the atomic number of the medium, A its mass number, I its average ionization potential (equal to 80.5 eV in air), N_A the Avogadro's number, ze the projectile charge, v its velocity, M its mass. γ and β are respectively the Lorentz factor (energy) and velocity of the projectile. At a low energy, the energy losses decrease when energy increases till a minimum value called the *minimum ionization*.

2.1.2.5. Compton scattering

This process is named after its discoverer Arthur Holly Compton. This is the inelastic scattering of a photon by a charged particle, principally atomic electrons. The photon transfers a fraction of its energy to an atomic electron. Compton scattering and absorption are relevant at low energy, below the critical energy (defined in the next section). This process is important at the later stage of the shower development, far from the shower axis, because the number of low energy photons beyond the depth of maximum development (noted X_{max}, see section 2.1.3 for the definition) is important, producing many low energy electrons by this process. If we consider a photon of energy k, the differential cross-section can be expressed as:

$$\frac{\mathrm{d}\sigma_C}{\mathrm{d}k\mathrm{d}k'} = 2\pi r_e^2\frac{1}{k'}\frac{1}{q}\left[1+\left(\frac{k'}{k}\right)^2 - \frac{2(q+1)}{q^2} + \frac{1+2q}{q^2}\frac{k'}{k} + \frac{1}{q^2}\frac{k}{k'}\right]\ \mathrm{cm^2sr^{-1}/electron} \tag{2.10}$$

where k' is the final photon energy and q the primary photon energy in units of electron mass.

2.1.2.6. Cherenkov light

This is P. Cherenkov who discovered this process (Nobel prize in 1958 for the interpretation of this process). When a charged particle moves in a medium with a velocity higher than the velocity of light in this medium, a radiation is emitted: this is the Cherenkov radiation. In other words, a threshold energy is setted for this emission, depending on the value of the refractive index in the considered medium (at sea level, $n_{air} = 1.0003$ and the threshold energy for electrons is ~ 21 MeV). This radiation is emitted in the form of light (photon in UV range) on a cone around the projectile trajectory with an opening angle given by:

$$\cos\theta = \frac{1}{\beta n} + q \tag{2.11}$$

with $q = \frac{\hbar k}{2p}\left(1 - \frac{1}{n^2}\right)$ a quantum correction factor with a small influence (where $\hbar k$ and p are the momenta of the photon and the moving secondary charged particles), $\beta = v/c$ with v being the velocity of the particle. The opening angle is maximum when the projectile moves at the speed of light in the medium ($\beta = 1$). The opening angle is directly inversely proportional to the refractive index of the considered medium, whatever the value of β. The number of emitted photons per unit pathlength can be expressed as:

$$\frac{\mathrm{d}N}{\mathrm{d}L} = z^2 \frac{\alpha}{\hbar c}\left[1 - \frac{1}{\beta^2 n^2}\right] \text{ [photon/cm]} \tag{2.12}$$

where α is the fine-structure constant, z the charge of the particle and n the index of refraction of the medium.

2.1.3. Phenomenology and physics of extensive air showers

Nowadays, it's a well known fact that EAS are initiated by a series of interactions of a high energetic primary cosmic ray hadron with the atmosphere constituent (N_2, O_2 and Ar). The secondary charged particles propagate in the atmosphere, forming a hadronic cascade propagating longitudinally along the primary cosmic ray trajectory (what we will call the shower axis). A lateral dispersion of secondary particles occurs, due to their transverse momentum generated by the different processes discussed in section 2.1.2. During the interactions of the primary particle with a nucleus of the atmosphere, some pions π^0 and $\pi^\pm$ are created. The particles π^0 decay in two γ which produce e^- and e^+ by pair production. These e^-/e^+ could create γ by Bremsstrahlung, and the e^+ annihilate when thermalized by the medium. These γ decay in a pair e^+–e^-. This is the electromagnetic component, containing 90 % of the primary particle energy [6], that can extend to large lateral dispesion ($\sim$ km) depending on the primary cosmic ray energy. Moreover, muons and neutrinos resulting from the decay of charged pions create another component, the muonic component. The hadronic component composed of pions continues to rise until the probability that pions decay in $\mu^\pm$ is higher than the probability of a new interaction in the medium. From that time, the number of particles in the shower decreases until its extinction, corresponding to the time where the number of particles in the shower is zero. It is worth noticing that this phenomenon is in reality more complex because of the characteristics of the atmosphere (change of the density for instance). During the shower development, the secondary particles generate processes such as Cherenkov, fluorescence and radio. A simple illustration of EAS is shown in figure 2.2.(a).

A sketch of the geometry of an EAS is shown in figure 2.2.(b), with the point of first interaction X_1, the point of inflection X_{inf} where the production of secondary particles in the shower is maximum and the point of depth of maximum development X_{max}. As

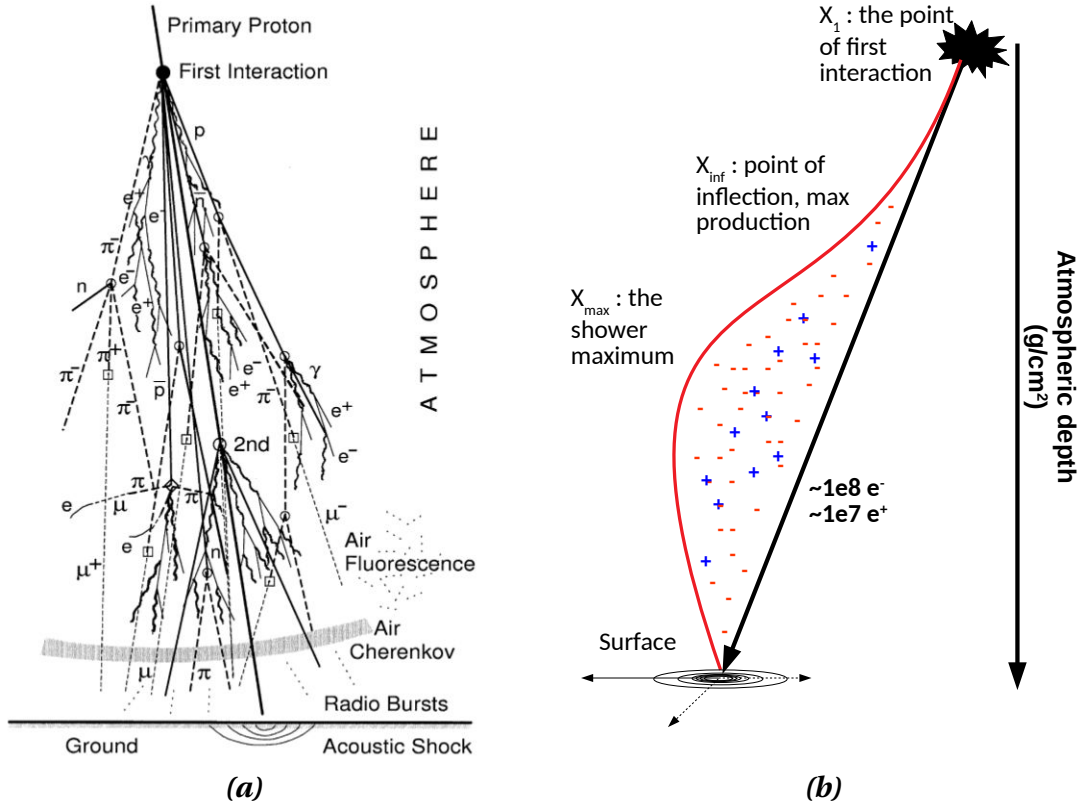


Figure 2.2 *– (a): Schematic view of the longitudinal and lateral development of an extensive air shower. Adapted from [58]. (b): Sketch of the geometry of an EAS, with the parameters X_1 (the first point interaction), X_{inf} (the point of inflection where the production of particles in the shower is maximal) and X_{max} (the atmospheric depth corresponding to the maximum of particles in the shower). The order of magnitude of the number of electrons and positrons is indicated for a primary proton with an energy of* 10^{17} eV*. The atmospheric depth (noted X) is calculated by integration of the density of air from the first point of interaction X_1 to the wanted point X_2 (units in* $\text{g} \cdot \text{cm}^{-2}$*).*

a reminder, parameter X_1 corresponds to the atmospheric depth of the first interaction point, where the chain reaction begins. As previously described, the number of particle increases until X_{max}. From this point, the available energy in the shower front is insufficient and the number of particles absorbed in the atmosphere is larger than the number of produced particles, and the number of particles in the shower decreases till the shower extinction.

The shower development depends on the primary cosmic ray characteristics: its energy, its arrival direction and its nature. If some conditions are fulfilled, the shower front (as a reminder, the thin particle disk which evolves during the shower development) can reach the ground. Let us consider first the primary cosmic ray energy. For a very energetic primary, the shower front could reach the ground, and the shower maximum occurs near the sea level. Inversely, a low energetic primary induces a small shower with a maximum of development reached at higher altitudes in the atmosphere (see figure 2.3) and only muons and neutrinos can reach the ground. One can see in figure 2.3 that the higher the primary energy, the higher the X_{max}, as indicated by the dash-dot line. If we now consider the nature of the primary (see figure 2.4), one can see that the point of first interaction X_1 (see figure 2.2.(b)) is different depending on the primary cosmic ray, because the cross-section depends on the mass number A and on the energy E. More-

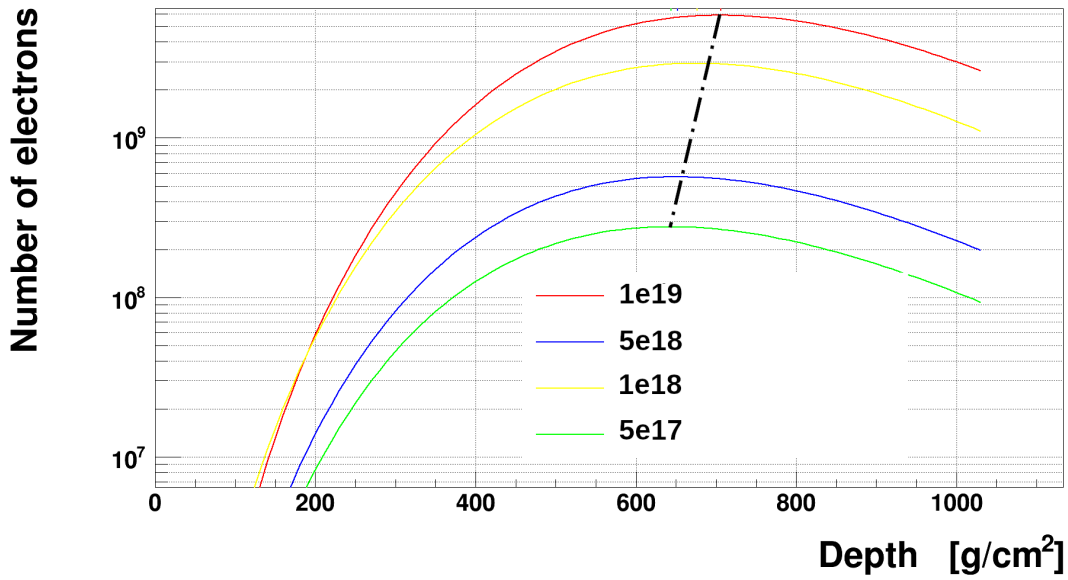


Figure 2.3 *– Longitudinal profile of four EAS induced by iron nucleus for different primary energies:* 5×10^{17} eV, 1×10^{18} eV, 5×10^{18} eV *and* 1×10^{19} eV. *The longitudinal profile is the distribution of the number of particles in the shower over the atmospheric depth along the shower axis. The higher the primary energy, the higher the* $X_{\max}$*, indicated by the dash-dot line. Obtained with the Monte Carlo code CONEX [59].*

over, one can see that the lateral dispersion is very different for the three primaries with an energy of 10^5 GeV. Here are presented the secondary particles that have an energy above 10 GeV. If we considerer all the secondary particles, the γ-induced air shower has a dispersion radius, defined as the lateral dispersion measured from the shower axis at sea level, of 100 m, the proton-induced air shower has a dispersion radius of 500 m and the iron-induced air shower has a dispersion radius of 1,000 m. With a careful attention, one could see that the depth of maximum development of the shower $X_{\max}$ (see figure 2.2.(b)) is also different for the three primaries, leading to the fact that the shower maximum is sensitive to the primary mass.

We are going to discuss first about the simplified Heitler model [60, 61] which, applied to the electromagnetic component, could described the evolution of the longitudinal profile and give us observables for the identification of the primary particle. In this model, the shower is composed of the same type of particles with a radiation length noted $\lambda_{radiation}$. After N steps, the particle number is $N_n = 2^n$ and their individual energy is E_0/N_n [62]. The development stops when this energy goes below the threshold where the loss of energy by Bremsstrahlung is similar to the loss of energy by ionization. This is called the critical energy ($E_c^\gamma = 80$ MeV in the air, where γ refers to the photons which are part of the electromagnetic component). The properties of the electromagnetic component are:

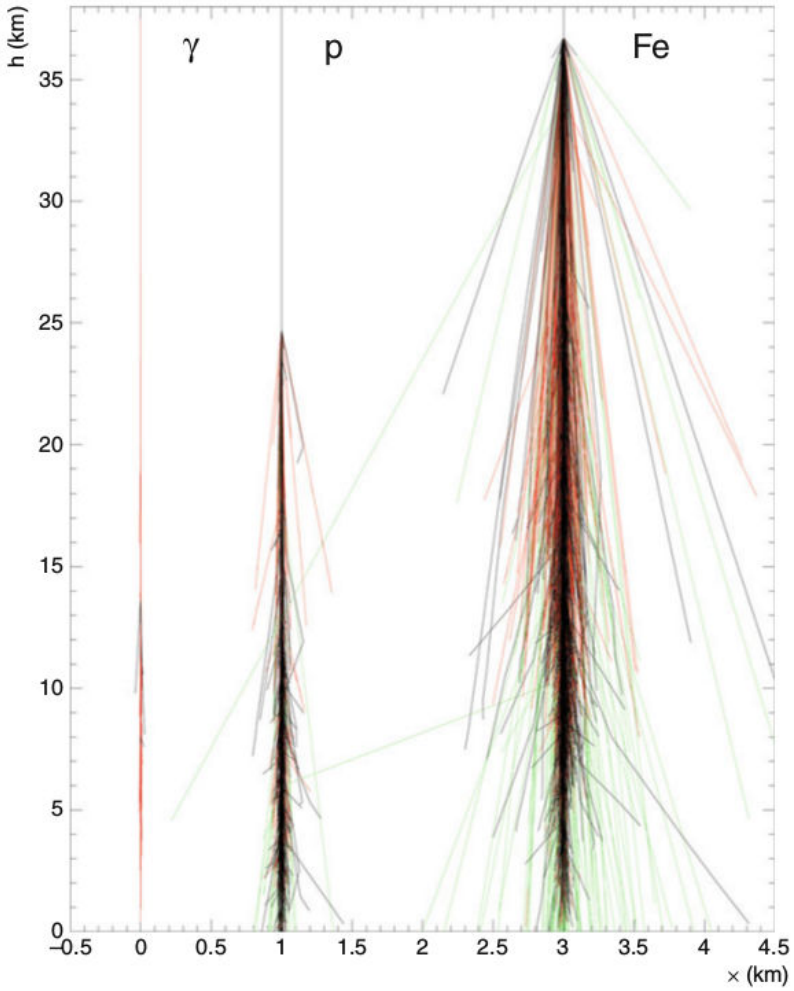


Figure 2.4 *– Extensive air shower development for different primary cosmic rays. The three primaries are simulated with an energy of* 10^5 GeV*. Red component is for electromagnetic, green component for muons and black for hadrons. Adapted from [58].*

— The number of particles at the maximum of the development of the air shower is proportional to the energy of the primary cosmic ray E:

$$N_{\max} = \frac{E}{E_c^{\gamma}}, \tag{2.13}$$

with $E_c^{\gamma} = 80\,\text{MeV}$.

— The evolution of $X_{\max}$ is a logarithmic function of the energy:

$$X_{\max} = X_1 + \lambda_r \ln \frac{E}{E_c^{\gamma}} = X_1 + \lambda_{radiation} \ln N_{\max}, \tag{2.14}$$

with $\lambda_{radiation} = 37\,\text{g}\cdot\text{cm}^{-2}$ in air, and where X_1 is the point of first interaction driven by the cross section of the primary particle. This point is important: indeed, we do not have a lab measurement of cross-sections above 10^{18} eV, and thus cosmic rays permit to measure cross-section at energies unreachable with the Large Hadron Collider (LHC), for instance.

This model presents some limitations: the number of particles at $X_{\max}$ is overesti-

mated (by a factor 2 or 3), and also the ratio of the number of electrons and photons (by a factor 12) [61, 58]. However, despite these limitations and the fact that the hadronic component is different and more complex than the electromagnetic one, the Heitler model could be applied to this hadronic component, the observables $N_{\max}$ and $X_{\max}$ being then expressed differently.

The shape of the shower front particle can be described by a semi hyperbole or a semi parabola centred on the shower axis. The density of the particles in the shower front is variable, being more dense near the shower axis [63]. Its thickness is less than 1 m near the axis, and a few meters away from the axis. The lateral profile will be discussed next. We can write, for a nucleus of atomic number A:

$$X^A_{max} = X_1 \ln\left[\frac{2(1-K_{el})E}{A(\langle m\rangle/3)\epsilon_0}\right] + \lambda_N(E) = X^P_{max} - X_1 \ln A \tag{2.15}$$

describing the interaction of a nucleus of atomic number A with an energy E_0 in the atmosphere at the depth λ_N. In this interaction, it loses $(1-K_{el})$ of its initial energy and produces $\langle m\rangle$ secondary pions. A third of them decay into two photons and supply the electromagnetic component previously described. From these equations, we can draw some conclusions:

- Showers initiated by iron nuclei (heavy) develop earlier in the atmosphere compared to those initiated by a proton (light), as shown in figure 2.5.
- We expect to find more muons in a shower initiated by a heavy nucleus: the ratio N_{e-}/N_μ is an important observable to identify showers with same energy initiated by different nuclei: this is part of the upgrade of the Pierre Auger Observatory, *AugerPrime*, by installing plastic scintillator on top of each water-Cherenkov tank and underground muon detectors ([64, 65, 66]).
- The point of first interaction of the primary particle with atmospheric nuclei, X_1, is the first source of uncertainty.

The number of particles in the shower front can be described by the Gaisser-Hillas function [67]:

$$N(X) = N_{\max}\left(\frac{X-X_1}{X_{\max}-X_1}\right)^{\frac{X_{\max}-X_1}{\lambda}} \exp\left(\frac{X_{\max}-X}{\lambda}\right) \tag{2.16}$$

with $N(X)$ the number of particles at the depth X in $\mathrm{g\cdot cm^{-2}}$, $N_{\max}$ the maximum number of particles and λ the attenuation length of the secondary particles in atmosphere equal to $70\ \mathrm{g\cdot cm^{-2}}$. We can also describe the longitudinal profile with the Greisen-Iljina-Linsley parametrization [68], using the concept of the age of the shower noted s:

$$N(E,A,t) = \frac{E}{E_l}\exp\left(\frac{X-X_1}{X_0}(1-2\ln s) - a - b\left(\ln\frac{E}{E_c} - \ln A\right)\right) \tag{2.17}$$

where A is the mass of the primary, E the energy of the primary, $X_0 = 36.7\ \mathrm{g\cdot cm^{-2}}$ is the radiation length of electrons in air, $E_l = 1.45\times10^3$ MeV is a normalization factor, $E_c = 81$ MeV is the critical energy (ionization and Bremsstrahlung losses are equal), $b = 0.76$ is the elongation rate and $a = 1.7$ is an offset parameter.

We have seen that the number of particles in an EAS is proportional to the energy of the primary cosmic ray, and that the shower maximum shifts with the logarithm of the energy as shown by equation 2.15. Figure 2.5 presents CONEX, an EAS simulation

program [59], simulations of the longitudinal shower profile of 10 vertical protons (red) and 10 vertical irons (blue) induced EAS at 10^{17} eV (top) and at 10^{19} eV (bottom) made with the EPOS LHC hadronic interactions model [69]. The vertical plain line stands for the atmosphere depth corresponding to the altitude of CODALEMA (an experiment dedicated to the observation of the radio signal emitted by EAS, widely described in the following chapters), the vertical dashed line stands for the atmosphere depth corresponding to the altitude of Auger. The energy dependence of the depth of shower maximum is clearly visible: for a given primary and a given zenith angle, the higher the energy is, the larger the $X_{\max}$ is, meaning that the shower maximum is reached later for the highest energies. It is worth noting that there is a clear separation of the protons and iron nuclei for the $X_{\max}$ value, and that there is a large shower-to-shower dispersion of $X_{\max}$ values for primary protons mainly due to stochastic interaction processes occurring during the shower development. All of this permit to determine statistically the composition of cosmic rays, as previsouly described in section 1.3.2.1 page 26 and as discussed in chapter 7 page 159.

The shower can also be described by the integral of the lateral distribution of the particles, which is related to the height of the shower. The lateral distribution is the superposition of the three components previously mentioned (electromagnetic, hadronic, muonic). The lateral dispersion of the electromagnetic component is due to the Compton scattering and Coulomb scattering. The model used to describe it was proposed by Nishimura, Kamata and Greisen (NKG) [70]. It is a parameterization which takes into account the electromagnetic and hadronic components. It depends on the steps of development of the shower s and on the Molière radius r_M (can be assimilated to the radius of a cylinder which contains ~ 90 % of the deposited energy of the shower):

$$\rho(r, s, N_e) = \frac{C_1 N_e}{2\pi r_M^2} \left(\frac{r}{r_M}\right)^{s-2} \left(1 + \frac{r}{r_M}\right)^{s-4.5} \left(1 + C_2 \left(\frac{r}{r_M}\right)^{\delta}\right) [m^{-2}] \tag{2.18}$$

where r is the distance to the air shower axis, N_e is the number of particles at ground and C_1 is a normalization factor which depends on the so-called "age" of the shower s (this is not really the age of the shower, but its stage of development). The age of the shower is correlated to the atmospheric depth, as shown in figure 2.6. In order to obtain the size of the shower, i.e the total number of charged particles in the shower, we integrate the lateral distribution function:

$$N = 2\pi \int_0^{\infty} \rho(r) r \mathrm{d}r \tag{2.19}$$

In figure 2.6, $s = 0$ corresponds to the point of first interaction X_1 and $s = 1$ corresponds to the point of depth of maximum development $X_{\max}$.

2.2. Shower detection methods

This section aims to describe the different shower detection methods and the dedicated apparatus. We have seen in previous sections that an EAS is characterized by its longitudinal and lateral profiles, but can be also described by its shower front. This shower front presents a curvature and has a high density near the shower axis which decreases with the distance to the shower axis, as described by the equation 2.18. The particles in the shower front induce principally an optical Cherenkov light, radio waves, and phenomena as air fluorescence as they propagate during the shower development. Nowadays, the air shower detection is done by recording the secondary particles reaching the ground, by optical Cherenkov detector arrays or by fluorescence detectors as Fly's Eye, HiReS, Auger or TA detectors. The radio emission and detection will be described in an upcoming dedicated chapter.

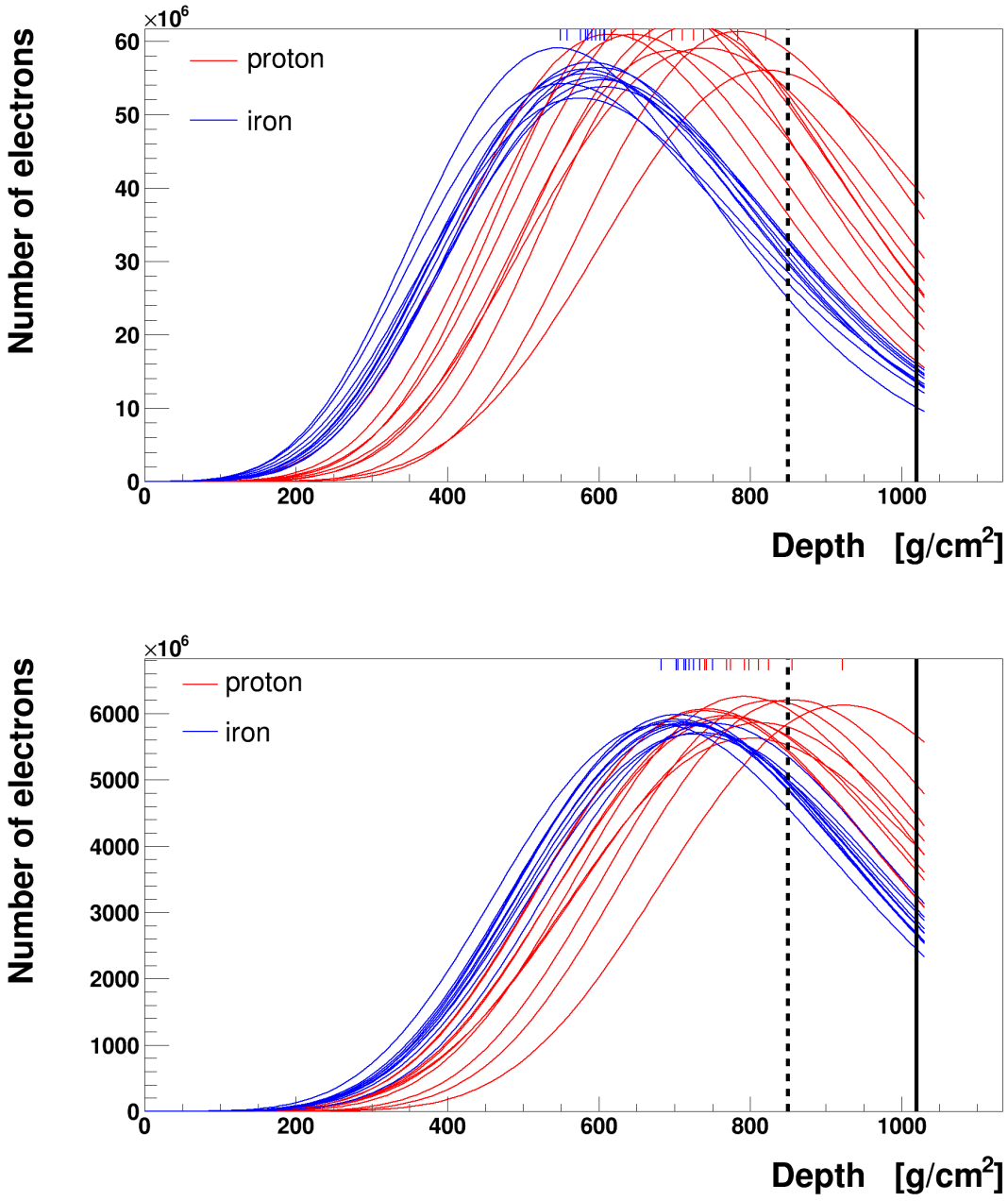


Figure 2.5 – *CONEX simulations of the longitudinal shower profile of 10 vertical protons (red) and 10 vertical irons (blue) induced EAS. Top:* 10^{17} eV. *Bottom:* 10^{19} eV. *The vertical plain line stands for the atmosphere depth corresponding to the altitude of CODALEMA, the vertical dashed line stands for the atmosphere depth corresponding to the altitude of Auger.*

2.2.1. Particle detector arrays

This kind of detectors was used since the beginning of the era of air shower studies, in the end of the 30's by Pierre Auger and collaborators. The bulk of this method is to detect in coincidence in several detectors packets of charged particles. For that, the detectors are deployed on ground. A partial two dimensional picture of the shower is obtained when the particle disk reaches the ground, corresponding to a precise stage of the shower development. One can note that the higher the density of the particle detectors, the better the sampling of the shower front, and the higher the precision on the primary cosmic ray characteristics. Today, two types of materials are used:

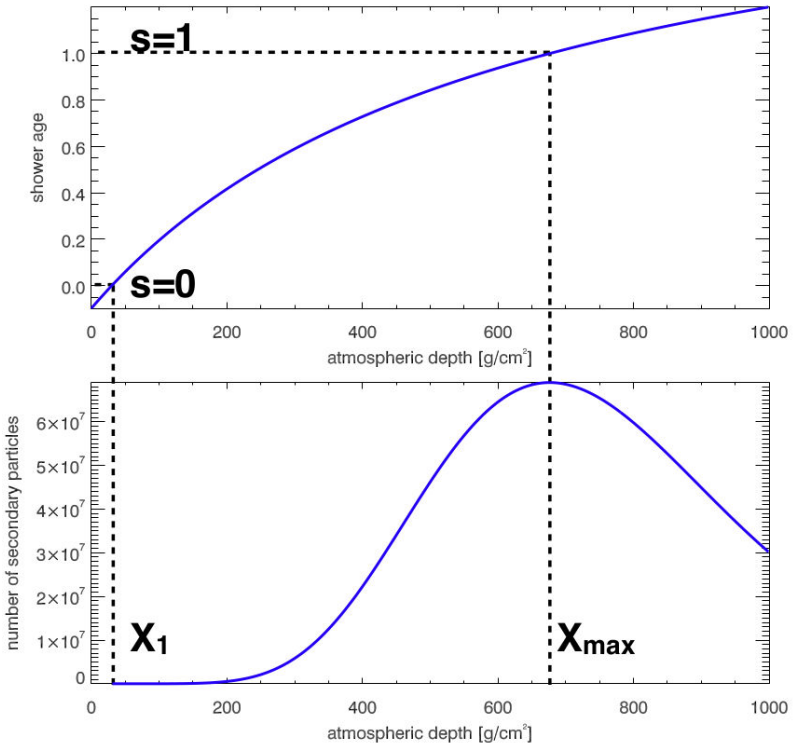


Figure 2.6 *– Top: shower age as a function of atmospheric depth. Bottom: number of secondary particles as a function of atmospheric depth. Adapted from [71].*

— The plastic scintillators: particles crossing the plastic induce light emission received by photomultiplier tubes. This phenomenon is due to the scintillation of organic molecules constituting the plastic. This type of scintillators is used in the CODALEMA and EXTASIS experiments (see Chapter 4, section 4.2.1), in the Telescope Array experiment and will be used on *AugerPrime*.
— The water tank, used for the Auger experiment (and at Haverah Park also), where the light emitted by the particles crossing water is due to the Cherenkov effect and is detected with photomultipliers.

The characteristics of the particle detectors have an influence on their detection efficiency. For example, a water tank is a volumetric sensor allowing a good detection efficiency for the inclined air shower, contrarily to plastic scintillators which are very thin. The geometry of the particle detector arrays is peculiar to each experiment, and relies on:

— The energy range to study.
— The required statistics of data (at constant density, a larger array will detect more showers than a smaller).
— The cost.

The upper limit of the studied energy range is fixed by the covered area, while the lower limit is fixed by the step of the grid of the particle detector arrays (i.e the distance between each detector). As explained previously, the flux is a quickly decreasing function with the energy, and the total number of particles in the shower increases with the energy. Thus, for high energy cosmic ray, a good compromise is to increase the step of the grid to cover a greater area. With this type of detector, the pieces of information obtained are:

— The arrival direction of the primary cosmic ray, triangulated with the different arrival times in the concerned detectors (different description is possible for the shower front curvature).
— The estimation of the number of secondary particles at ground level by the deposited signal in the different detectors.
— The lateral profile reconstruction using a NKG parameterization.

— The estimation of the particle shower core.
The duty cycle of this type of detector is ~ 100 %.

2.2.2. Air Cherenkov detector arrays

Air Cherenkov detector has a large aperture and a wide angle acceptance. It records the atmospheric Cherenkov radiation of air showers. In that case, as explained in section 2.1.2.6, photons are detected. The geometry of Cherenkov radiation is shown in figure 2.7. The position of the wavefront can be expressed by the equation 2.12. The air Cherenkov

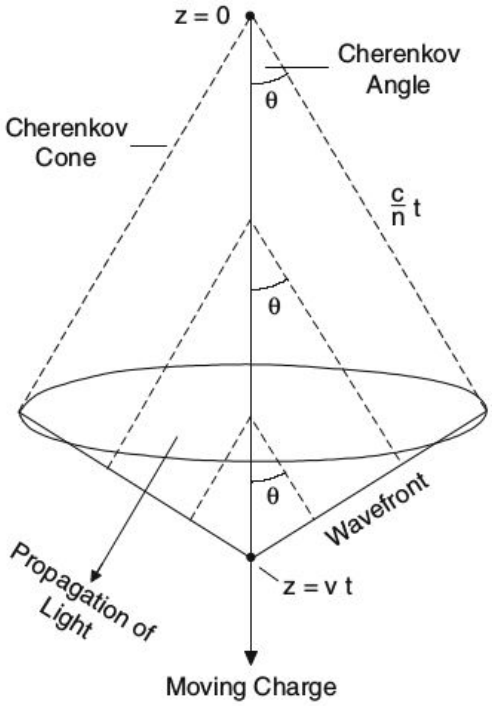


Figure 2.7 – *Geometry of Cherenkov radiation. Adapted from [58].*

detector covers a large fraction of the sky. The Cherenkov telescope used for gamma ray astronomy works differently and have a very narrow field of view, since in the case of gamma ray astronomy, we search for localized sources, while in the case of cosmic rays the source location is unknown and the cosmic rays come from all the sky directions. Nowadays, the air Cherenkov detector is used in the Tunka Valley for the Tunka Advanced Instrument for cosmic ray physics and Gamma Astronomy (TAIGA). The collected data are the number density, the arrival time of the optical photons emitted and the position of the concerned detector. This type of detectors works as a particle detectors, so the shower reconstruction works as the one of the particle detector, but data contain more information on the longitudinal profile because all steps of the shower development are seen (data from particle detectors contain more information on the lateral profile, but only the last step of the shower development can be seen at ground level).

2.2.3. Air fluorescence detectors

The fluorescence light emission induced by the charged particles during the shower development is isotropic. The components of the atmosphere, mainly the nitrogen, on the shower path are excited by the charged particles by absorption of photons (second positive band for the nitrogen molecules and first negative band for nitrogen ions). To return to their ground state, the nitrogen molecules emit photons in the UV range, so-called scintillation process, as shown in figure 2.8. In that case, the atmosphere can be

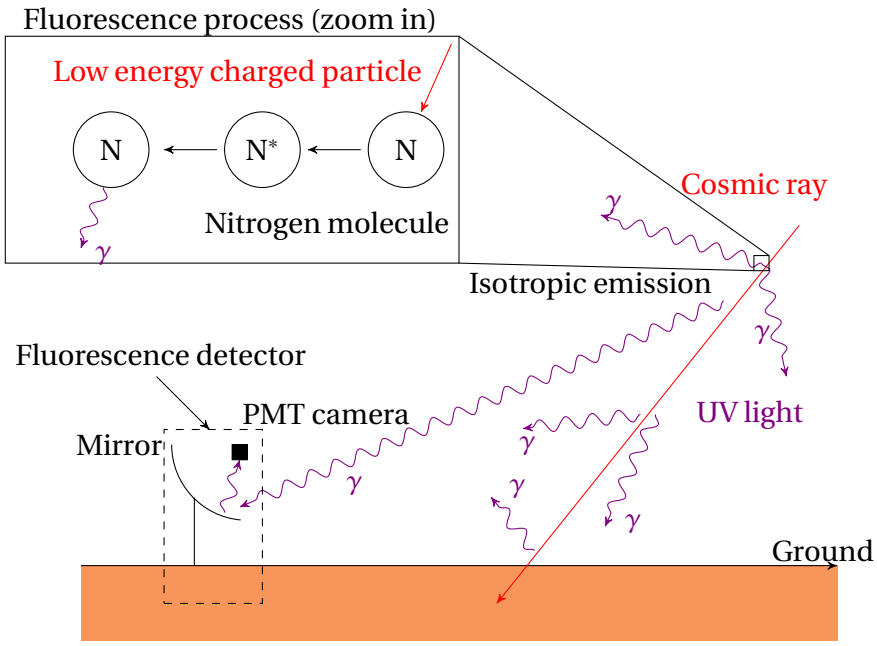


Figure 2.8 *– Cosmic ray fluorescence detection and process.*

seen as a scintillator. The isotropic emission of fluorescence light means that showers could be detected from all directions, meaning that the shower can be observed from its side. There are different geometries for this type of detectors, for example Auger and HiRes do not use the same detector, but the principle remains the same. Mirrors reflect the fluorescence light to a camera composed of photomultiplier tubes. Air showers can be observed up to ~ 20 km away. Each detector observes a defined part of the sky, and all the detectors allow an observation of approximately half the sky (2π sr) in the case of Telescope Array experiment and approximately 7400 km^2 · sr at 10^{19} eV for the Pierre Auger Observatory. The fluorescence detector is a very powerful tool to infer the composition on the primary cosmic ray, but has a duty cycle of only 14 % because this type of detectors operates in special conditions: a clear night, without moon or bad weather.

2.2.4. Other type of detection

2.2.4.1. RADAR

RADAR (Radio detecting and ranging) is a detection method which has been used in 1941 by Blackett and Lovell [72] to study the EAS when developing in the atmosphere by detecting their radar reflection. This method is usually used to track space objects or ballistic missiles, but also to detect aircraft, to measure distances and so on. In the case of EAS, Blackett and Lovell have been the pioneers in the study of this method, although their proposal did not receive a lot of attention, due to the complexity of the method, the background noise level and trigger considerations. In the 70's, a Tokyo research group carried out an experiment using a LORAN (Long Range Navigation) system to study the radio emission of EAS induced by primary cosmic ray above 10^{20} eV [73]. Nowadays, experimental efforts are carried out to study the cosmic ray air showers with this technique, such as at the Telescope Array experiment [74]. Recent works are investigating the radar technique for detecting neutrinos above 4 PeV in dense media like ice [75, 76, 77].

2.2.4.2. Acoustic detection

As said previously, depending on the characteristics of the EAS, the latter can reach the ground. When reaching the ground, several phenomena occur, such as the production of

a thermo-acoustic shock wave that should be detectable. This thermo-acoustic emission has been confirmed by accelerator experiments and previous studies, such as at Lake Baikal [78] where EAS shower front impacting on large volume of fresh water at high altitude has been explored, but also at AMADEUS [79] in Mediterranean Sea, ACoRNE in the Atlantic Ocean or at SPATS experiments [80] which is an extension of IceCube.

Showers induced by cosmic rays are similar to showers induced by neutrinos, and some of the detection techniques described in this chapter are also valid for neutrino induced showers. It is worth noticing that, for a better accuracy and a redundancy of the observables, hybrid detectors and couple experiments are highly recommended, such as it is done at the Pierre Auger Observatory, for instance.

***

In this chapter, we have seen the phenomenology of EAS, from the relevant electromagnetic processes for the shower development until their detection. As a reminder, the EAS permit to characterize the primary cosmic rays. For that, different methods of detection exist. We deliberately ignore the method that will interest us in the following, since it is largely detailed in the following chapter: the radio detection method of the EAS.

Part II

The Extensive Air Shower radio detection

"We'll walk this road together, through the storm. Whatever weather, cold or warm."

Marshall Bruce Mathers III

CHAPTER 3

Radio signal emission from Extensive Air Showers, simulation and detection

Outline

The ultimate goal of the study of cosmic ray is the reconstruction of the parameters of the primary which arrives in the atmosphere, especially its mass. Indeed, the nature of the primary is primordial to understand the mechanisms of acceleration of cosmic rays, their propagation and their sources. Nowadays, the measurements of the shower maximum via the fluorescence technique is mainly used to deduct the mass composition event by event. The measurements of the Pierre Auger Observatory are shown in figure 3.1. The mass composition is not well established, due to the lack of statistics at high energy, and also due to the poor duty cycle of 14 % of the fluorescence technique, as mentioned in section 2.2.3. A third detection method exists: the radio detection of EAS. First developed during the 1960's, then abandoned due to technological limitations but fully revisited at the digital era since 2002, this technique is based on the fact that, during

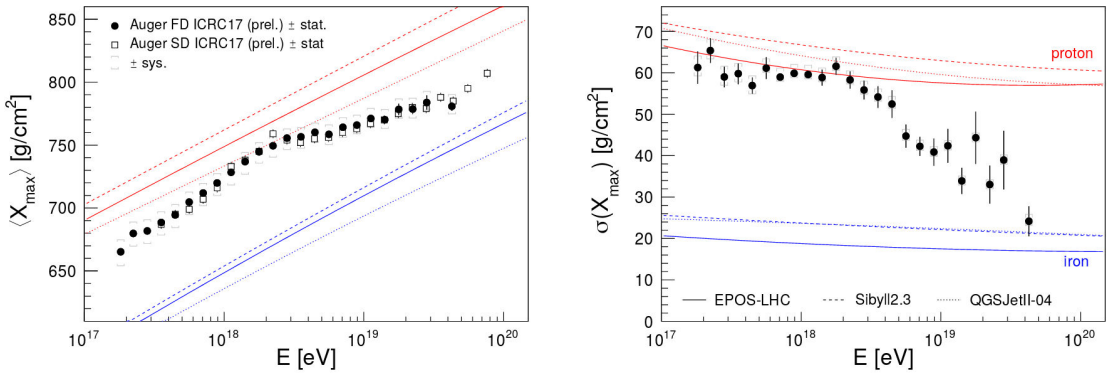


Figure 3.1 *– Left: measurement of $\langle X_{\max} \rangle$ as a function of energy by the Pierre Auger Observatory. Red and blue lines represent air-shower simulations of $\langle X_{\max} \rangle$ for proton and iron primaries. Right: its standard deviation. Taken from [18].*

the shower development in the atmosphere, the charged particles in movement generate a coherent transient electric field. The latter has two main origins: transverse current induced by the geomagnetic field and radial current induced by the charge excess mechanism [81, 82, 83, 84, 85]. In both cases, the currents vary during the shower development, leading to a radio emission [6]. The use of the radio method, sensitive to the whole EAS longitudinal profile and with a better duty cycle (~100 %), could improve the number of detection at high energy, and that is, with a better statistics, reduce the error bars on the measurements of the shower maximum.

In this chapter, we deal with the mechanisms responsible for the emission of an electric field during the shower development, which can be then detected at ground with dedicated radio antennas. The experiments and results leading to the radio detection are first treated, especially the key role of optical Cherenkov in the first investigation of the radio detection. Then, the main mechanisms that produce the radio emission induced by EAS are introduced. Lastly, the radio simulation and the radio detection of EAS are discussed.

3.1. On the way to radio transient detection

3.1.1. Pioneer works

Lucien Mallet(1885-1981) is one of the first french x-ray operator and radiation therapists. In 1926, he described an emission of a low wavelength light in water and organic substances exposed to gamma rays. The emitted spectrum associated to this emission is continuous. The Russians Sergey Vavilov and Pavel Cherenkov showed that this radiation is independent of the composition of the liquid, which was contradictory with an emission due to a fluorescence mechanism. Ilja Frank and Igor Tamm described then this radiation classically, and obtain in 1958 the Nobel Prize with Pavel Cherenkov. Due to the fact that the speed of light in a medium is lower than the speed of light in vacuum, a particle can move in this medium with a velocity higher than the spedd of light in this medium, inducing a radiation as described in section 2.1.2.6. In the mid 1950s, Jelley and Galbraith discovered an optical Cherenkov emission (see section 2.2.2 for the detection technique) in coincidence with the observation of an EAS [86]. Jelley proposed [87] to investigate the radio frequency band of the Cherenkov emission, described by the

Frank-Tamm relation:

$$\frac{\mathrm{d}P}{\mathrm{d}\nu} = \alpha \left(\frac{4\pi^2\hbar}{c}\right) z^2 \nu \left(1 - \frac{1}{\beta^2 n^2}\right) \ell \tag{3.1}$$

where ν is the frequency, α the fine-structure constant, $\hbar$ the reduced Planck constant and ℓ the the path length travelled by a particle of charge ze in a medium of refractive index n. One can see that when we pass from the visible domain to the radio domain, the spectrum decreases rapidly (by a factor of 10^7). At that time, the state of the art is as follows:

— The Cherenkov effect appears to be a good contributor to be a source of radio transient because of its success in optical band.
— An equal number of electrons and positrons are contained in the shower leading to destructive interferences and thus radio emission is impossible.
— Electronic components are not so fast (we know nowadays that the observed radio transients are extremely fast, around 20 ns) and the radio detection of EAS is difficult to carry out.

From that, the proposition of Jelley to study the radio domain of the Cherenkov radiation seemed to be complicated to be implemented. However, in 1962, Askaryan [82] stipulated that because of the annihilation of positrons with the electrons of the dielectric medium (the atmosphere in the case of EAS), an excess of electrons must exist in the shower front. Hence, the second item in the previous list disappears, and a radio emission becomes possible. The Askaryan emission mechanism, also known as the charge excess emission mechanism, will be described in section 3.2.2, page 52.

3.1.2. First observation of radio transients

The first observation of radio transients was carried out by Jelley [88]. The instrumental setup was as follows: Geiger counters were used as particle detectors, and radio pulses were searched in coincidence with Geiger counters signals by using 72 horizontal dipole antennas, working at 44 MHz. The energy of the primary cosmic ray responsible for the observed pulses was estimated to be larger than 5×10^{16} eV. Actually, the estimation was wrong, but a conclusion can still be drawn from these radio pulse observations: air showers emit detectable radio signal. However, unlike Jelley's proposition, these observations do not provide the affirmation that the radio emission is due to the phase-coherent Cherenkov process. Nevertheless, at that time, the relevant pieces of information were:

— There is a radio emission during the shower development, first observed by Jelley *et al.* at 44 MHz [88], and then from 2 MHz up to 500 MHz [89].
— Radio emission is dominated by the geomagnetic mechanism (described in section 3.2.1), and a correlation with the angle between air shower axis and the geomagnetic field starts to arise.
— There is a dependence between signal strength and lateral distance of antenna from the air shower axis.

All these assumptions were summarized in an equation by Allan [90]:

$$\epsilon_\nu = 20\,\mu\mathrm{V m^{-1} MHz^{-1}} \left(\frac{E_p}{10^{17}\mathrm{eV}}\right) \sin\alpha \cos\theta \exp\left(-\frac{R}{R_0(\nu,\theta)}\right) \tag{3.2}$$

where ϵ_ν is the radio pulse amplitude per unit bandwidth, α is the angle between the direction of arrival of the shower and the geomagnetic field vector, R is the distance to the shower axis and θ is the zenith angle. This equation is valid for $10^{17} < E_p < 10^{18}$ eV and for $R < 300$m. R_0 is a factor depending on θ and frequency (it increases when θ increases and ν decreases), and for example $R_0 = 110 \pm 10$ m for $\nu = 55$ MHz and $\theta < 35\,°$.

3.2. Physics of radio transient emission mechanisms

After a stop in the radio detection efforts in the mid-70s, the radio detection technique re-emerged in the 2000s, bringing renewed hope for an improved detection of cosmic ray induced air showers. As for the fluorescence, the radio detection method is a calorimetric measurement, because of the production of the radio signal during all the shower development and the transparency of the atmosphere for the studied frequency range. The method has a duty cycle of ~ 100 %, except during thunderstorms which enhance the amplitude of the radio signal which is no longer directly characteristic of the EAS. A radio antenna is cheaper than the other instruments used for the shower studies, permitting to cover very large areas. Nowadays, radio signal permits to reconstruct the shower core, the $X_{\max}$ (nature) and the energy of the primary cosmic ray. There has been a lot of progress on the detection, the understanding and the simulation of radio emission induced by EAS. One of the biggest evolution comes from digital electronics, allowing a very fast processing of the radio signal.
Let us now deal with the radio emission mechanisms occurring during shower development.

3.2.1. Geomagnetic emission mechanism

There are two main mechanisms involved in the production of the radio emission by cosmic ray showers [6] (see figure 3.2).

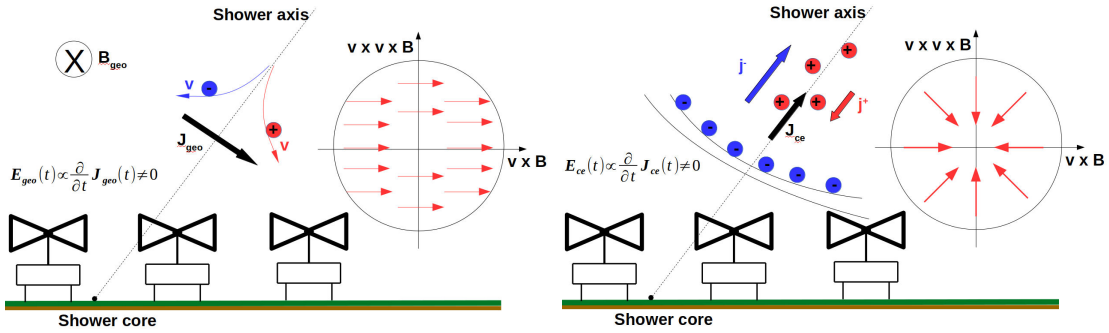


Figure 3.2 *– Left: Illustration of the geomagnetic mechanism. The arrows denote the direction of linear polarisation in the plane perpendicular to the air shower axis. Irrespective of the observer position, the emission is linearly polarised along the direction given by the Lorentz force, $\sim \vec{v} \times \vec{B}$ (east-west for vertical air showers). Right: Illustration of the charge excess (Askaryan) emission. The arrows illustrate the linear polarisation with electric field vectors oriented radially with respect to the shower axis.*

The main one is the emission associated with the geomagnetic field, well described by Kahn and Lerche [81]. Secondary electrons and positrons in the EAS are accelerated and deflected by the geomagnetic field, leading to a radio emission. This systematic deflection of the charged particles by the geomagnetic field can be explained by the Lorentz force: $\vec{F} = q\vec{v} \times \vec{B}$ where q is the particle charge, $\vec{v}$ its velocity and $\vec{B}$ the geomagnetic field vector. In the macroscopic description, the resulting currents will be, on average, perpendicular to the shower axis and are called *transverse currents*. They vary with the EAS development and the number of secondary particles first increases to reach a maximum, and then decreases as the shower dies out. This time variation will lead to an electromagnetic radiation. The resulting electric field is linearly polarized and aligned along $\vec{F}$ (thus along $\vec{n} \times \vec{B}$), where the direction of propagation of the secondary particles can be assimilated with the shower axis $\vec{n}$. At fixed energy, the intensity of the electric field depends on the direction of arrival of the primary cosmic ray, in particular

on the angle α between the direction of arrival of the shower and the geomagnetic field (also called the *geomagnetic angle*). In other words, for the CODALEMA experiment where the geomagnetic field points $\theta = 27\,°$ and $\phi = 270\,°$ and is oriented downward, the geomagnetic mechanism produces a larger electric field for showers coming from the North when all other parameters are fixed. The CODALEMA data show a good agreement with this assertion, as shown in figure 3.3, which presents a 10° gaussian smoothed sky map of 1,572 events detected by CODALEMA from the end of 2015 and the end of 2018.

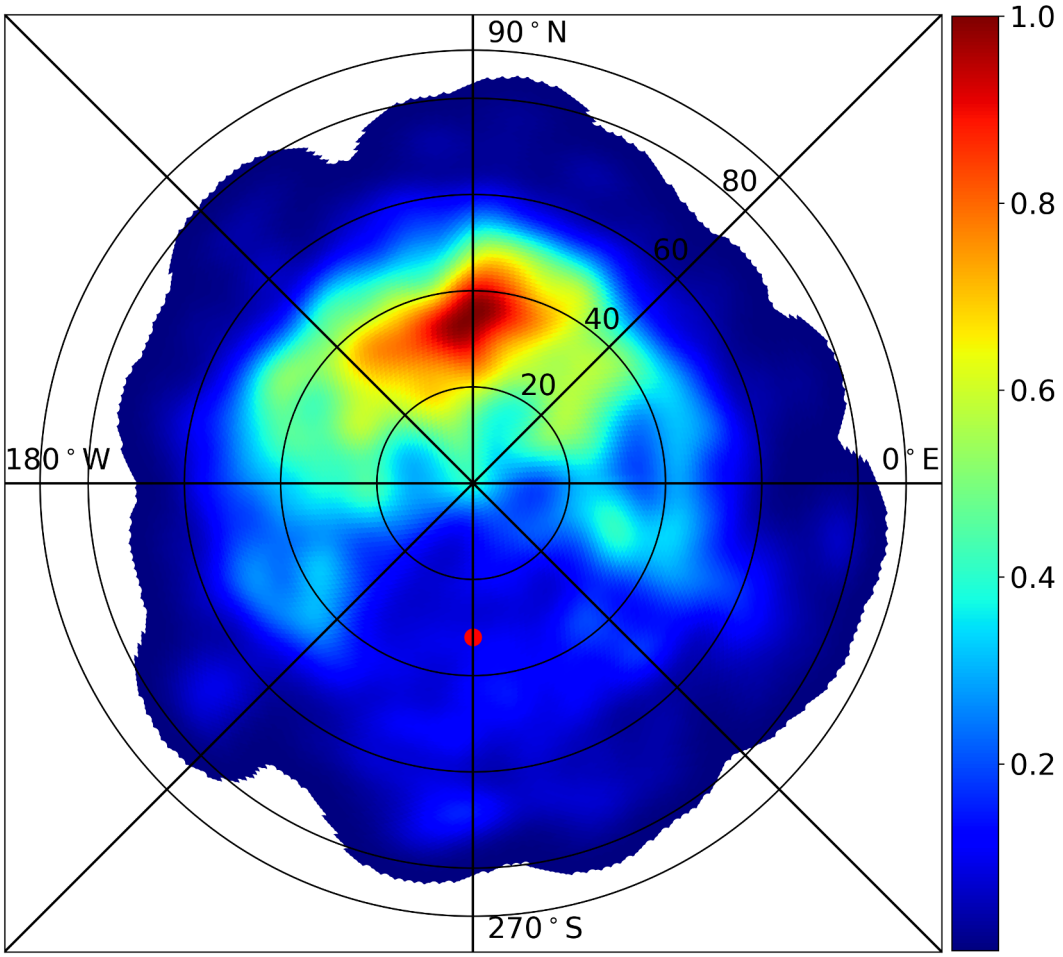


Figure 3.3 – 10° *gaussian smoothed sky map of events detected at CODALEMA. The azimuth angle is indicated on the outskirts of the circle, the zenith angle is indicated inside the circle.. The red point represents the geomagnetic field. The dataset contains 1572 events recorded by CODALEMA from the end of 2015 and the end of 2018, the construction of the figure is detailed in appendix B.*

As described in [83], the North–South asymmetry comes from the difference in the radio signal emission depending on the direction of arrival of showers, which is explained by the geomagnetic mechanism detailed just above, since the strength of the radio signal depends on the *geomagnetic angle*.

All the charged particles in the shower are concerned by this mechanism, however, the main contribution to the radio emission comes from the electrons and positrons because the emission scales as $1/m^4$ (with m the mass of the charged particles) and muons are ~ 200 times heavier than electrons and positrons (this is even worse for other charged particles, such as the pions). These particles propagate at approximately the speed of

light, resulting in a compression in the forward direction for the emission inducing short radio pulses, as shown in figure 3.10 in section 3.3.2.3 page 62. Figure 3.4 shows the polarization of the signals, obtained by plotting the temporal evolution of the radio pulse of the EW polarization as a function of the temporal evolution of the radio pulse of the NS polarization, for an event simulated with the SELFAS code (presented and explained in section 3.3.2 page 58). A primary proton of 10^{17} eV is simulated, with a direction of arrival optimizing the geomagnetic mechanism at CODALEMA site, namely, $\theta = 63\,°$ and $\phi = 90\,°$. The shower core is placed at the center of the array, in $(x = 0, y = 0)$. As explain

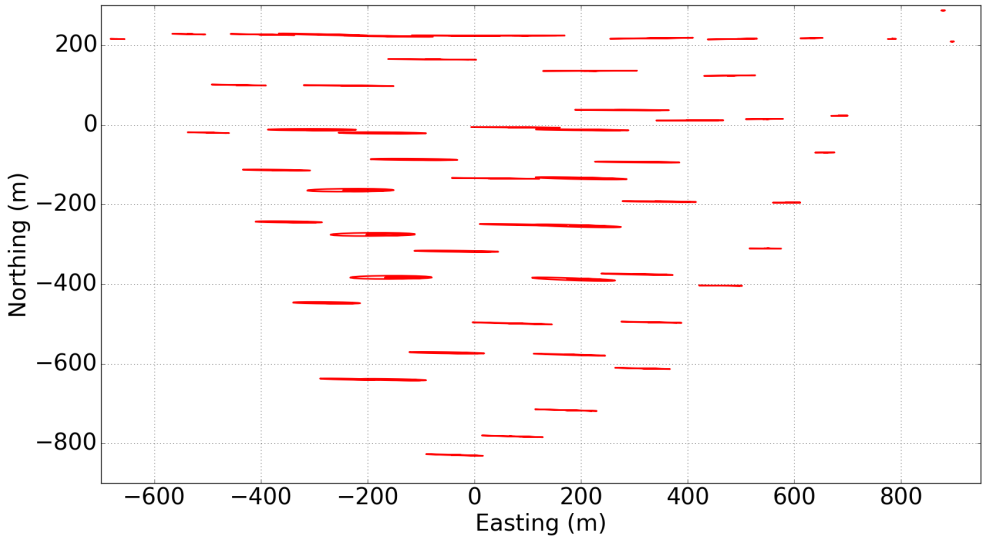


Figure 3.4 *– SELFAS simulation of a proton at* 10^{17} eV *coming from* $\theta = 63\,°$ *and* $\phi = 90\,°$ *optimizing geomagnetic effect and hitting the ground at* $(0,0)$. *Electric field polarizations are represented at the location of the CODALEMA antennas, and are aligned along* $\vec{F}$ *(thus along* $\vec{v} \times \vec{B}$*). The size of the pattern reflects the amplitude of the signal.*

previously, the electric field vector is linearly polarized and aligned along $\vec{F}$ (thus along $\vec{v} \times \vec{B}$) for the geomagnetic emission mechanism. In the case of CODALEMA, $B_x =\sim 0$, thus $\vec{F}_{NS}$ and $\vec{F}_{\text{Vertical}}$ are proportional. The Lorentz force can be expressed as:

$$\vec{F} = q \begin{pmatrix} -\cos\theta_B \sin\theta \cos\phi - \sin\theta_B \cos\theta \\ \cos\theta_B \sin\theta \cos\phi \\ \sin\theta_B \sin\theta \cos\phi \end{pmatrix} . \begin{pmatrix} \overrightarrow{e_{EW}} \\ \overrightarrow{e_{NS}} \\ \overrightarrow{e_{Vertical}} \end{pmatrix} \tag{3.3}$$

3.2.2. Charge excess emission mechanism

In addition to the geomagnetic contribution, there is the charge excess mechanism [82]. A negative charge excess of ≈ 10 - 20 % appears in EAS [84], due to the fact that the nitrogen molecules of the atmosphere are ionized by the air shower particles allowing the annihilation of the positrons with the atmospheric electrons. At high energy, the processes of pair production (section 2.1.2.2) and bremsstrahlung (section 2.1.2.1) dominate. As the shower develops, the average energy decreases and other processes appear, as Compton scattering (see section 2.1.2.5), delta-ray production, knock-on electrons (via Bhabha and Møller scattering), positron annihilation, resulting in a net negative charge excess. These processes are presented in figure 3.5, showing respectively the Feynman diagrams of Compton, Bhabha and Møller scattering.

The ionization electrons are contained in the curved shower front, while the heavier positive ions present a delay and travel upstream of the shower front. Obviously, the

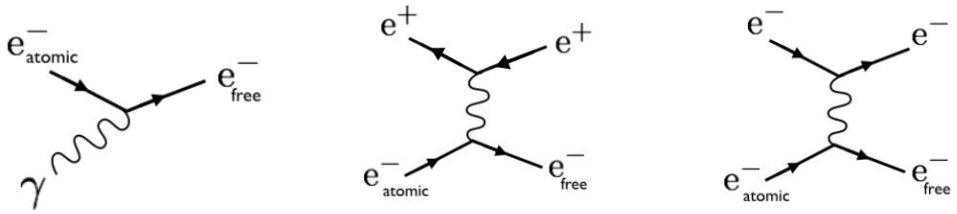


Figure 3.5 *– Feynman diagrams of Compton, Bhabha and Møller scattering.*

source of the radio emission being the same, an identical scheme as for the geomagnetic mechanism occurs: during the shower development, the net negative charge increases to reach a maximum and finally decreases when the shower dies out. The time variation of the evolution of the charge excess leads to a radio transient emission [6]. Contrarily to the geomagnetic mechanism, the electric field vector is radially polarized with respect to the shower axis, and its orientation depends on the location of an observer, as illustrated in figure 3.2.

Figure 3.6 shows the polarization for an event simulated with the SELFAS code. A primary proton of 10^{17} eV is simulated, with a direction of arrival optimizing the charge excess mechanism at CODALEMA site, namely, $\theta = 27\,°$ and $\phi = 270\,°$. The shower core is placed at the center of the array, in $x = 0$ and $y = 0$.

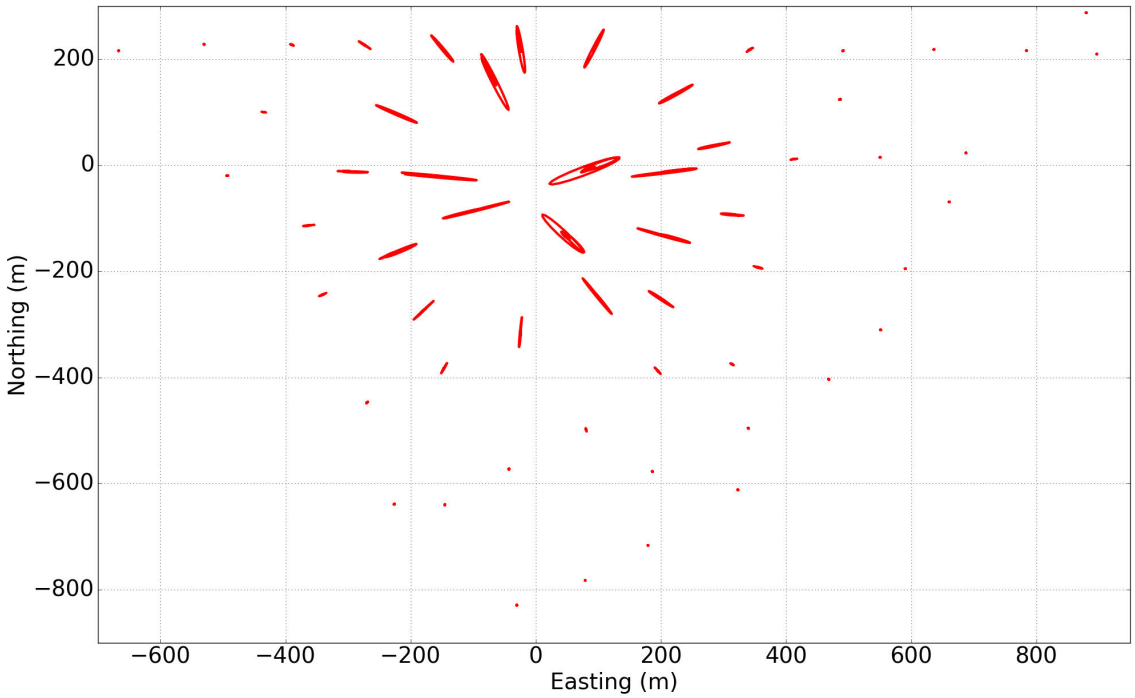


Figure 3.6 *– SELFAS simulation of a proton at* 10^{17} eV *coming from* $\theta = 27\,°$ *and* $\phi = 270\,°$ *and hitting the ground at* $(0, 0)$*. Electric field polarizations are represented at the location of the CODALEMA antennas, and are radially oriented with respect to the shower axis. The size of the pattern reflects the amplitude of the signal.*

The polarizations are thus an important information to distinguish the predominant mechanism for the radio emission event by event, since the geomagnetic and charge excess mechanisms have different polarization patterns. Moreover, we see for this simulations that the orientation of the polarizations seems to converge to a zone, due to

the radial orientation of them. In the present case, we know that the shower core is set in $x = 0$ and $y = 0$, and we observe that the polarizations converge to this zone. In the case where the Askaryan mechanism is dominant, the polarizations might indicate the shower core location. However, the polarizations do not point exactly to the shower core location. One hypothesis is that this is probably due to the lateral extension of the shower, which may leads to a small but sufficient contribution of the geomagnetic mechanism for perturbing the polarization pattern. The superposition of the two mechanisms is explained in the section 3.2.3 below.

3.2.3. Superposition and asymmetries

The total electric field is the summation of the two latter mechanisms. Depending on the observer location, the interferences can be destructive or constructive, as shown in figure 3.7, giving an asymmetric total electric field, and thus generating an asymmetry around the shower axis. For the example of a vertical shower used in figure 3.7, the electric field amplitude of the EW polarization is higher at the east side of the particle core, which leads to a shift of the radio core with respect to the particle core. This asymmetry has

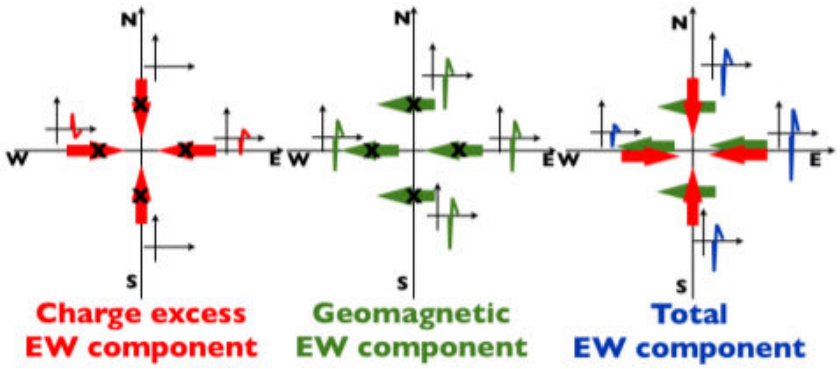


Figure 3.7 *– Interferences of the two main mechanism: depending on the observer location (represented by the black crosses), the interferences can be destructive or constructive, giving an asymmetric total electric field. The orientation of the measured electric field in the EW polarization at ground level is represented for four antennas at different positions but equivalent distance from the shower axis: left for the charge excess mechanism, center for the geomagnetic mechanism and right for the summation of the two mechanisms. Taken from [84].*

been experimentally observed by the CODALEMA experiment, as shown in figure 3.8.

3.2.4. Other mechanisms

We give in this subsection some other mechanisms which are able to produce a detectable radio transient.

3.2.4.1. Geoelectric charge separation mechanism

Some studies have been carried out on the geoelectric charge separation mechanism [91, 92]. Under normal atmospheric conditions, the ambient electric field reaches the value of about $100\ \mathrm{V \cdot m^{-1}}$. It increases significantly during thunderstorms. According to Wilson [91], the electrostatic field begins to be null after the interaction of the electron with the atmosphere. This leads to a fast detectable electric transient. Charman [92] completed and revisited this conclusion, and considered that a transverse charge separation occurred during the shower development, just like for the geomagnetic mechanism, due to the perpendicular component of the electric field, leading to a fast radio transient. For

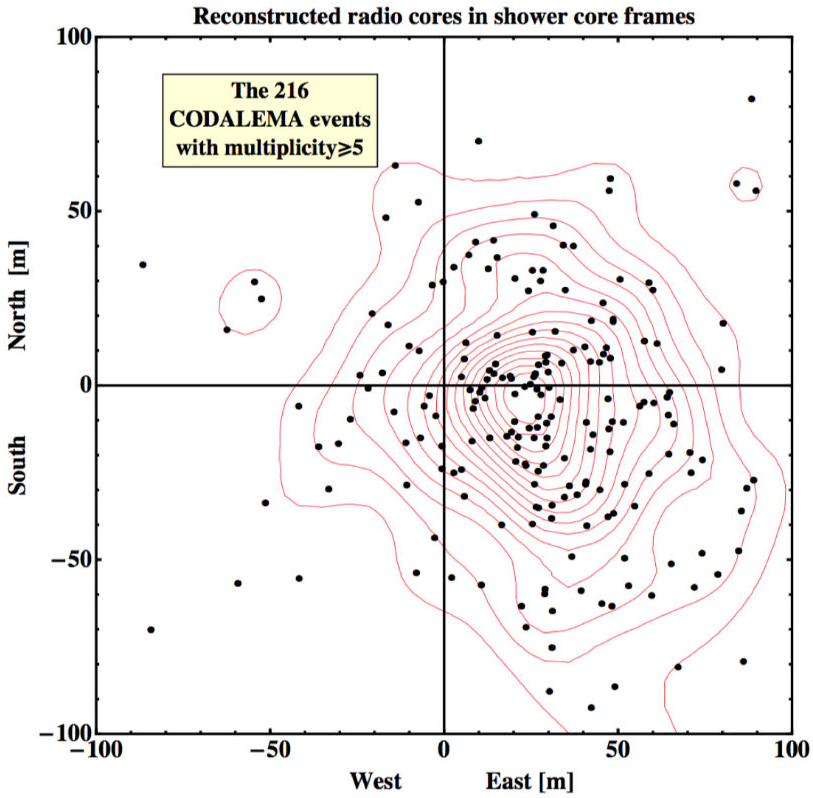


Figure 3.8– *Asymmetry observed in the CODALEMA data. Black dots represent the positions of the radio core relative to the shower core deduced from the particle detectors. Taken from [84].*

more details, please refer to [91, 92, 93]. The conclusion that can be drawn concerning this mechanism is that the electric field detected at ground is no longer directly proportional to the primary cosmic ray energy, since it has been enhanced by the atmospheric electric field. Using the radio detection method, the signal coming from this mechanism becomes non exploitable.

3.2.4.2. Transition radiation

The transition radiation was first predicted by Ginzburg and Franck in 1946. The transition radiation (TR), occurs when a charged-ultra-relativistic particle crosses the interface of two media with different dielectric constants. This transition can be described by classical electrodynamics [94]. In the sixties, the technology was insufficient to measure precisely this small signal. It is worth noticing that the radiation of a particle with a uniform movement when it crosses the interface of two media is similar to that of a particle with a non-uniform movement in vacuum. In both cases, the emission is correlated to the phase velocity of the electromagnetic wave and to the particle velocity. In the first case, this is the velocity of the wave which is changing, while in the second case this is the particle velocity. TR and the Bremsstrahlung have a common physical origin. The emitted field is characterized by the properties of the particle and the medium. Let's assume that, for simplicity, the two media are homogeneous. We note that the two fields in each of the respective media are different, because these media have different characteristics. The TR is due to the properties of polarization and depolarization of the crossed media by the charged particle [Ginzburg and Frank (1945 and 1946), Garibian (1958), Wartski (1976)]. Close to the trajectory, when the particle crosses a medium, the electrons of this medium are perturbed, forming atomic dipoles. They come back to their initial position by emitting a radiation. In a homogeneous medium, the positions of the

dipoles are symmetric around the particle, and the total radiation is null. If this symmetry is lost, as this is the case for an interface between two media or one medium and vacuum, a dipolar momentum appears, inducing the TR independently of the particle velocity. In other words, this is the rearrangement of the electrons in the two media which is at the origin of this radiation. The equations describing the TR can be found in the literature, as in *"Classical Electrodynamics"* by Jackson [94] or in *"Principles of Radiation Interaction in Matter and Detection"* by Leroy and Rancoita [57]. Radio emission from the TR is expected for a shower developing later in the atmosphere, i.e. for an event with a large number of electrons reaching the ground. The TR will be emitted when the shower front crosses the boundary between air and ground (only the negative charge excess will contribute). The TR will be emitted in forward and backward directions, the last one which can be detected by radio antennas. For recent studies on transition radiation: [95, 96].

3.2.4.3. Sudden death mechanism

The sudden coherent deceleration of the secondary electrons in the shower front when reaching the ground level emits a strong electric field which could be detectable. Let the reference time ($t = 0$) be the instant of impact at ground. The SDP arrives at the antenna at time $t_{\mathrm{SDP}} \backsimeq d_{\mathrm{core}}/c$, with d_{core} the distance between shower core and the antenna and c the speed of light in the medium. If at least 3 antennas are involved, the core position can be estimated by simple intersection of circles of radius d_{core}. If we can observe with the same antenna the electric field from the development in the air and the SDP, we could have an intrinsic time scale within the shower, which would allow a direct estimation of the distance between the X_{max} (the atmospheric depth of the maximum of the shower development) and the ground along the shower axis [97], providing an excellent way to obtain the nature of the primary cosmic ray. While the electric field amplitude on the ground created by the regular emission needs a two-dimensional description in general, the amplitude of the SDP decreases as $1/d_{\mathrm{core}}$. The medium frequency (MF) pulse amplitude strongly depends on the axis distance d_{axis} with a Gaussian decrease at first order, and simulation also predicts that the regular pulse can be detected at larger distances at low frequencies, as it will be shown in the following of this manuscript, as it will be shown in chapter 8. Moreover, this mechanism is also completely theoretically described in [98].

3.3. Simulation of the electric field emitted by air showers

In this section, we present an overview of the available simulation codes of the electric field emitted during the shower development. We detail the simulation code SELFAS (developed internally within the group), and we explain how we use it to simulate the physical process that we want to detect.

3.3.1. Existing simulation codes for air shower

All the existing models are inspired by the same electromagnetism bases, namely the Maxwell's equations. The differences between the models lie in the mathematical treatment of these equations when they are computerized. We distinguish two types of models: the macroscopic and the microscopic models. First ones treat globally the particles of the shower to compute an electric field, while second ones treat independently each particle of the shower and evaluate individual electric fields summed up to produce a global electric field. Both aim at giving the relation between the characteristics of an EAS and the characteristics of the emitted radio pulse.

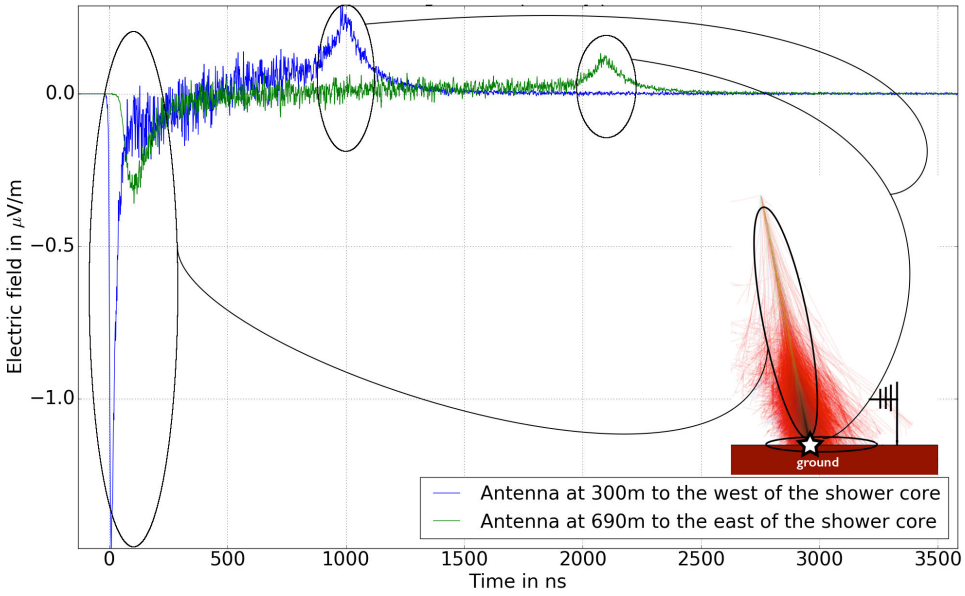


Figure 3.9 *– Sketch of the emission of the SDP. Vertical polarization of the electric field as a function of time obtained with a SELFAS3 simulation for an observer at* 300 m *(blue) and* 690 m *(green) of the shower core and for a vertical proton-induced shower at* 10^{18} eV. *The negative peaks at the beginning of the traces are due to the shower development in air. The second, positive peaks around have been interpreted as a new mechanism: the effect of the coherent deceleration of the shower front when hitting the ground, called the "sudden death pulse" (SDP). In the antenna located at 300 m from the core, the simulated SDP arrives* $\delta t = 300/c = 1$ µs *after the normal pulse, which is consistent with the propagation time from the shower core. The decreasing of the amplitude of the SDP as* $1/d_{\text{core}}$ *is also visible.*

3.3.1.1. Macroscopic models

The macroscopic models are mostly used to model the shower as a macroscopic current and charge density, and thus to model the macroscopic emission of the electric field. These are analytical models and can compute the electric field rapidly [6]. MGMR (Macroscopic Geo-Magnetic Radiation) [99, 100] is a macroscopic analytic model, in the time-domain, based on the deflection of charged particles of air shower due to the geomagnetic mechanism (this creates the transverse current). Secondary mechanisms such as charge excess mechanisms are also taken into account. It uses Monte-Carlo simulations, and predicts bipolar radio pulses. The EVA model [101] is also based on a macroscopic approach to calculate radio emission from EAS, particularly on the macroscopic current present in the shower front. It emphasises especially on the role played by the Cherenkov effects in the radio emission, and namely on the dependence of the radio pulse on the position to the shower maximum.

3.3.1.2. Microscopic models

With the microscopic models, the radiation is calculated for each single particle, and superposed to obtain the total radiation emitted by an EAS during its development.

REAS [102] is a Monte Carlo code based on the endpoint formalism. It consists in calculating the radiation emitted by the acceleration of a moving particle at the extremities of particle track [103]. It calculates and superposes the geosynchrotron radio emission from all the individual particles present in the EAS. It uses the histogrammed CORSIKA

showers [104] (software for simulation of EAS).

As the REAS code, CoREAS [105] uses the endpoint formalism which has been integrated directly into CORSIKA. Thus, it uses the electron and positron distributions obtained with CORSIKA.

ZHAireS [106] is a simulation code based on Monte Carlo simulations, using the AIRES (AIRshower Extended Simulations) shower code [107]. It permits the calculation of the electric field in time and frequency domains emitted during the shower development in the atmosphere or ice, by calculating the radiation emitted directly by the particle tracks (not only at the extremities of the track such as in the endpoint formalism).

The SELFAS code [108, 97] is the simulation code used in the following of this thesis. It is not based on a full Monte Carlo simulation, contrarily to the ZHAireS and CoREAS which use AIRES [109] and CORSIKA respectively, but uses the universality principle of shower for the distributions of the electron-positron in air showers [110]. In EAS, many distributions (in particular those related to electrons and positrons) depend practically only on the total number of particles in the shower and the age of the shower [70] which makes it possible to avoid a complete simulation with CORSIKA or AIRES. From that, no EAS simulation is needed to simulate the radio emission and the computation time is strongly reduced compared to CoREAS and ZHAireS.

3.3.1.3. Commentary on the different models

The main advantages of the macroscopic models is their speed of calculations (due to the fact that they are analytical models), only few minutes to obtain a simulated electric field. But the main problem is that this kind of approach simplifies a lot the description of a shower, and gives only a global view of the dominant mechanisms inducing the emission of the electric field, and consequently, a global view of the characteristics of a shower. This is confirmed by the comparisons between the macroscopic and microscopic models, which are not in agreement [105]. This result indicates that the physical processes are not fully described by the macroscopic models, furthermore, they do not fit the data as well as microscopic models.

For the microscopic models, the differences come from the Monte Carlo codes used and from the formalism used to calculate the radiation. Concerning the ZHAireS and CoREAS, the first one uses the ZHS formalism and the second one the endpoint formalism. These two formalisms were qualified as mathematically equivalent [6], but are not equivalent when dealing with computational calculations, leading to different time calculations and to the use of patches in particular cases, such as CoREAS which uses a patch for particles close to or on the Cherenkov cone (consisting in the use of the ZHS equations). However, a comparison of the electric-fields (amplitude, shape...) simulated by ZHAireS and CoREAS shows that they agree at a level of 20 % [111]. Moreover, the data are well fitted by these microscopic models. In the following, we will emphasize on the formalism used in SELFAS. In the following, we will limit ourselves to the use of one model, SELFAS, and occasionally, for comparisons, we will discuss the predictions from ZHAireS.

3.3.2. Formalism and prediction from SELFAS

3.3.2.1. Theoretical framework

Part of this section have been published in: **D. García-Fernández, et al.**, *Calculations of low-frequency radio emission by cosmic ray-induced particle showers*, Phys. Rev. D97 [98]. Let us begin by a recall of Maxwell equations:

$$\nabla \cdot \mathbf{D} = \rho \tag{3.4}$$

$$\nabla \wedge \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{3.5}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3.6}$$

$$\nabla \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{3.7}$$

where ρ is the charge density and $\mathbf{J}$ is the current density. We know that, in an homogeneous dielectric medium, $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ with ϵ the permittivity and μ the permeability of the considered medium. From the equation 3.6, we can write:

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{3.8}$$

where A is a vector potential. Thus, equation 3.7 can be rewritten as:

$$\begin{aligned} \nabla \wedge \mathbf{E} + \frac{\partial(\nabla \times \mathbf{A})}{\partial t} = \nabla \wedge (\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0 \\ \Rightarrow \mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla\Phi \\ \Rightarrow \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi \end{aligned} \tag{3.9}$$

We can rewrite Maxwell's equations in the form of two coupled equations by using the scalar and vector potentials. For this, we use the Lorentz gauge which can be expressed as $\nabla \cdot \mathbf{A} + \mu\epsilon\frac{\partial\Phi}{\partial t} = 0$, and we obtain:

$$\begin{aligned} \nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon\mathbf{E}) = \epsilon\nabla \cdot \mathbf{E} = \rho \\ = \epsilon_0\nabla(-\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi) = \rho \\ = -\nabla^2\Phi - \nabla(\frac{\partial \mathbf{A}}{\partial t}) = \frac{\rho}{\epsilon} \\ \Rightarrow \boxed{\nabla^2\Phi - \frac{1}{c_n^2}\frac{\partial^2\Phi}{\partial t^2} = -\frac{\rho}{\epsilon}} \end{aligned} \tag{3.10}$$

where c_n is the speed of light in the dielectric medium. The same calculation can be made for the vector potential giving:

$$\boxed{\nabla^2\mathbf{A} - \frac{1}{c_n^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\mu_0\mathbf{J}} \tag{3.11}$$

To solve equations 3.10 and 3.11, we use retarded solutions known as Green functions, giving:

$$\Phi(\mathbf{x}, t) = \frac{1}{4\pi\epsilon}\int \mathrm{d}^3x' \frac{1}{R}\left[\rho(\mathbf{x}', t')\right]_{\text{ret}} \tag{3.12}$$

$$\mathbf{A}(\mathbf{x}, t) = \frac{\mu_0}{4\pi}\int \mathrm{d}^3x' \frac{1}{R}[\mathbf{J}(\mathbf{x}', t')]_{\text{ret}} \tag{3.13}$$

where $R = |\mathbf{x} - \mathbf{x}'|$ and t' is the retarded time defined as $t' = t - \frac{1}{c_n}|\mathbf{x} - \mathbf{x}'(t')|$. Using equations 3.12, 3.13 by plugging them into equation 3.9, we can derived the electric field and obtained:

$$\mathbf{E}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon}\int \mathrm{d}^3x' \left[\frac{\hat{\mathbf{R}}}{R}[\rho(\mathbf{x}', t')]_{\text{ret}} + \frac{\hat{\mathbf{R}}}{c_nR}\left[\frac{\partial\rho(\mathbf{x}', t')}{\partial t'}\right]_{\text{ret}} - \frac{1}{c_n^2R}\left[\frac{\partial\mathbf{J}(\mathbf{x}', t')}{\partial t'}\right]_{\text{ret}}\right] \tag{3.14}$$

where $\hat{\mathbf{R}} = \frac{\mathbf{x}-\mathbf{x}'}{|\mathbf{x}-\mathbf{x}'|}$. Using the fact that $t' = t - \frac{1}{c_n}|\mathbf{x} - \mathbf{x}'(t')|$, we can write:

$$\left[\frac{\partial f(\mathbf{x}', t')}{\partial t'}\right]_{\text{ret}} = \frac{\partial}{\partial t}\left[f(\mathbf{x}', t')\right]_{\text{ret}} \frac{\partial t}{\partial t'} = \frac{\partial}{\partial t}\left[f(\mathbf{x}', t')\right]_{\text{ret}} \tag{3.15}$$

and noticing that R is not a function of t, we can rewrite equation 3.14 as:

$$\mathbf{E}(\mathbf{x}, t) = \frac{1}{4\pi\epsilon}\left[\int \mathrm{d}^3x' \frac{\hat{\mathbf{R}}}{R}[\rho(\mathbf{x}', t')]_{\text{ret}} + \frac{\partial}{\partial t}\int \mathrm{d}^3x' \frac{\hat{\mathbf{R}}}{c_n R}[\rho(\mathbf{x}', t')]_{\text{ret}} - \frac{\partial}{\partial t}\int \mathrm{d}^3x' \frac{1}{c_n^2 R}[\mathbf{J}(\mathbf{x}', t')]_{\text{ret}}\right] \tag{3.16}$$

Now that we have derived the calculation of the electric field from the Maxwell equations, we need to define the source whose the electric field is to be calculated. All the Monte Carlo codes with a microscopic approach described in the previous section use a particle track to calculate the electric field of a source. Let us define the instant t_1 before which the charge density in all space is zero. Let us consider a neutral atom. At t_1 and position $\mathbf{x}_1$, a point-like source separates from the atom with the charge $-q$ and travels in a straight line at a constant speed until its sudden stop at instant t_2 and position $\mathbf{x}_2$. The associated charge density, which certifies that the charge is conserved, can be written as:

$$\begin{aligned}\rho(\mathbf{x}', t') &= -q\delta^3(\mathbf{x}' - \mathbf{x}_1)\Theta(t' - t_1) \\ &\quad +q\delta^3(\mathbf{x}' - \mathbf{x}_1) - \mathbf{v}(t' - t_1))[\Theta(t' - t_1) - \Theta(t' - t_2)] \\ &\quad +q\delta^3(\mathbf{x}' - \mathbf{x}_2)\Theta(t' - t_2)\end{aligned} \tag{3.17}$$

where δ is the Dirac delta function and Θ the Heaviside step function. The associated current density can be written as:

$$\mathbf{J}(\mathbf{x}', t') = q\mathbf{v}\delta^3(\mathbf{x}' - \mathbf{x}_1) - \mathbf{v}(t' - t_1))[\Theta(t' - t_1) - \Theta(t' - t_2)] \tag{3.18}$$

In order to simplify the integration in equation 3.16, we rewrite the term containing the Dirac notation of the second line of equation 3.17 as:

$$\begin{aligned}\delta^3(\mathbf{g}(\mathbf{x}')) &= \delta^3\left(\mathbf{x}' - \mathbf{x}_1 - \mathbf{v}\left(t' - \frac{|\mathbf{x} - \mathbf{x}'|}{c_n} - t_1\right)\right) = \sum_i \frac{\delta^3(\mathbf{x}' - \mathbf{x}_{p,i}(\mathbf{x}, t))}{|\frac{\partial \mathbf{g}}{\partial \mathbf{x}'}|_{\mathbf{x}_{\mathrm{p,i}}}} \\ &= \sum_i \frac{\delta^3(\mathbf{x}' - \mathbf{x}_{p,i}(\mathbf{x}, t))}{|1 - \mathbf{v}\cdot\hat{\mathbf{R}}(\mathbf{x}_{p,i}(\mathbf{x}, t))/c_n|} = \sum_i \frac{\delta^3(\mathbf{x}' - \mathbf{x}_{p,i}(\mathbf{x}, t))}{\kappa_i}\end{aligned} \tag{3.19}$$

where i denotes the several possible retarded positions (we might have two retarded positions below and above the Cherenkov angle), $\mathbf{x}_{p,i}(\mathbf{x}, t)$ is the retarded position of the particle as a function of (**x**,t) and $\kappa_i = |1 - \mathbf{v}\cdot\hat{\mathbf{R}}(\mathbf{x}_{p,i}(\mathbf{x}, t))/c_n| = |1 - \mathbf{v}\cdot\hat{\mathbf{R}}_i/c_n|$. From the definition of the Dirac delta function, we see that $\mathbf{x}'$ will be evaluated only at the retarded position. Using equation 3.19 in equation 3.16 and considering the source described by equations 3.17 and 3.18, the electric field is:

$$\begin{aligned}\mathbf{E}(\mathbf{x}, t) &= \frac{q}{4\pi\epsilon}\left\{-\frac{\hat{\mathbf{R}}_1}{R_1^2}[\Theta(t' - t_1)]_{\text{ret}} - \frac{\hat{\mathbf{R}}_1}{c_n R_1}\frac{\partial}{\partial t}[\Theta(t' - t_1)]_{\text{ret}}\right. \\ &+ \left[\frac{\hat{\mathbf{R}}_1}{\kappa R^2}(\Pi(t', t_1, t_2))\right]_{\text{ret}} + \frac{1}{c_n}\frac{\partial}{\partial t}\left[\frac{\hat{\mathbf{R}}_1}{\kappa R}(\Pi(t', t_1, t_2))\right]_{\text{ret}} - \frac{\mathbf{v}}{c_n^2}\frac{\partial}{\partial t}\left[\frac{1}{\kappa R}(\Pi(t', t_1, t_2))\right]_{\text{ret}} \\ &+ \left.\frac{\hat{\mathbf{R}}_2}{R_2^2}[\Theta(t' - t_2)]_{\text{ret}} + \frac{\hat{\mathbf{R}}_2}{c_n R_2}\frac{\partial}{\partial t}[\Theta(t' - t_2)]_{\text{ret}}\right\}\end{aligned} \tag{3.20}$$

with $\Pi(t', t_1, t_2) = \Theta(t' - t_1) - \Theta(t' - t_2)$ and $R_{1,2}$, $\hat{\mathbf{R}}_{1,2}$ are respectively the distances and unit vectors from the starting (index 1) and ending (index 2) point of the particle track to the observer. The mathematical definition of the Heaviside function stipulates that it takes

values 0 or 1. In that case, the derivatives of the Heaviside functions of equation 3.20 can be expressed with the Dirac delta as:

$$\frac{\partial}{\partial t}[\Theta(t'-t_i)]_{\text{ret}} = \left[\frac{\partial}{\partial t}\Theta(t'-t_i)\right]_{\text{ret}} = [\delta(t'-t_i)]_{\text{ret}} \tag{3.21}$$

and the equation 3.20 can be rewritten as:

$$\begin{aligned}\mathbf{E}(\mathbf{x},t) &= \frac{q}{4\pi\epsilon}\left\{-\frac{\hat{\mathbf{R}}_1}{R_1^2}\Theta(t-t_1-R_1/c_n) - \frac{\hat{\mathbf{R}}_1}{c_n R_1}\delta(t-t_1-R_1/c_n)\right.\\ &+ \left[\frac{\hat{\mathbf{R}}_1}{\kappa R^2}(\Pi(t',t_1,t_2))\right]_{\text{ret}} + \frac{1}{c_n}\frac{\partial}{\partial t}\left[\frac{\hat{\mathbf{R}}_1}{\kappa R}(\Pi(t',t_1,t_2))\right]_{\text{ret}} - \frac{\mathbf{v}}{c_n^2}\frac{\partial}{\partial t}\left[\frac{1}{\kappa R}(\Pi(t',t_1,t_2))\right]_{\text{ret}}\\ &+ \left.\frac{\hat{\mathbf{R}}_2}{R_2^2}\Theta(t-t_2-R_2/c_n) + \frac{\hat{\mathbf{R}}_2}{c_n R_2}\delta(t-t_2-R_2/c_n)\right\}\end{aligned} \tag{3.22}$$

in which we have used $t' = t - R_i/c_n$ for $t' < t_1$ and $t' > t_2$. First (creation of the particle) and third (stop of the particle) lines of equation 3.22 contain a static Coulomb field which contributes at t_1 and t_2 respectively and an impulse radiation field with the form $\frac{\hat{\mathbf{R}}_i}{c_n R_i}\delta(t-t_i-R_i/c_n)$ which is due to the charge conservation and the use of a realistic charge density. The impulse radiation field is due to the changes in the charge density, which thus can be paired with the second line. The second line is equivalent to the Heaviside-Feynman expression used in [97] to calculate the electric field of a particle track, and is inseparable from the impulse radiation field. Equation 3.22 is able to describe the electric field created by a particle track [98], and has been implemented in the SELFAS Monte Carlo code.

3.3.2.2. Implementation in the SELFAS Monte Carlo code

As we said previously, SELFAS is not based on a full Monte Carlo simulation but it uses the universality principle and the "shower age" for the distributions of the electrons-positrons in air showers. CONEX [59] is used as Monte Carlo code for computing the longitudinal profile. Then SELFAS samples the particles obtained from CONEX, and propagate them through a defined thickness of material (15 $\text{g}\cdot\text{cm}^{-2}$ for the thickness, corresponding in our case to the atmosphere). Moreover, we see that equation 3.22 depends on the refractive index of the atmosphere ($c_n = c/n$). Thus, we need a precise atmosphere description. Initially, the atmosphere description was based on the US Standard model, which is a static model not including day/night nor seasonal effects. In the last version of SELFAS, we are using the data from Global Data Assimilation System (GDAS) [112] to have a more precise description of the atmosphere at the time of ours cosmic-ray detections [113] and for a location close to the experiment site (see appendix C).

For the implementation in SELFAS, we need first to treat the case when the particle reaches the ground. Indeed, equation 3.22 is valid for a single medium, thus if a particle track is straddling between medium 1 (air) and medium 2 (below the ground), this particle track must be modify to end just above the ground, and another one must begin at the same time just after the ground. However, in the case of the EXTASIS experiment (see chapter 4), operating from 1 to 6 MHz, we can not use the far-field hypothesis or the geometrical optics. Furthermore, we are in the case where the antenna is in the near-field zone of the ground (9 m for EXTASIS) and the particle tracks too. Thus, the contribution to the electric field is a mixture of a surface wave and a reflected wave. The calculation of this kind of contribution is detailed in [114]. The total electric field to be calculated can be separated in three contributions:

— A direct contribution coming from the track above the ground level, which can be calculated with the equation 3.22 at all frequencies.
— A contribution coming from the response of the interface (in our case, the ground) to the field emitted by the particle track above the ground level. For that contribution, a special treatment has to be used depending on the positions of the particle track and the antenna with respect to the ground level: if the track and the antenna are in the far-field zone (see section 3.4.1.3 page 68 for more details), geometrical optics and standard Fresnel coefficients can be used. In the case where the antenna is in the near-field zone and the track in the far-field zone, the reciprocity theorem can be used [115] permitting to calculate the antenna voltage with the far-field antenna pattern and the direct field of the track.
— A contribution coming from particle tracks below the ground, which are not taken into account since the attenuation losses of radio waves inside soil [116] are estimated to be larger than in the atmosphere inducing a contribution smaller than the other contributions.

3.3.2.3. Results obtained with SELFAS

Figure 3.10 shows the simulated radio pulses obtained in the EW polarization for a vertical proton of 10^{19} eV as observed at various distance from the shower axis. They present a bipolar shape as expected by the formalism used in SELFAS, have a duration of few tens of nanoseconds for an observer close to the shower axis (blue curve for example) and of few hundreds of nanoseconds for an observer far away from the shower (yellow curve for example). We notice that the duration of the pulse depends on the observer position from the shower axis. It is the same for the polarity of the signal, which corresponds to the sign of the maximum of the absolute amplitude of the signal. In the case of figure 3.10, pulses for an observer situated at less than ~ 215 m of the shower axis present a negative polarity, while pulses for an observer situated at more than ~ 215 m of the shower axis present a positive polarity.

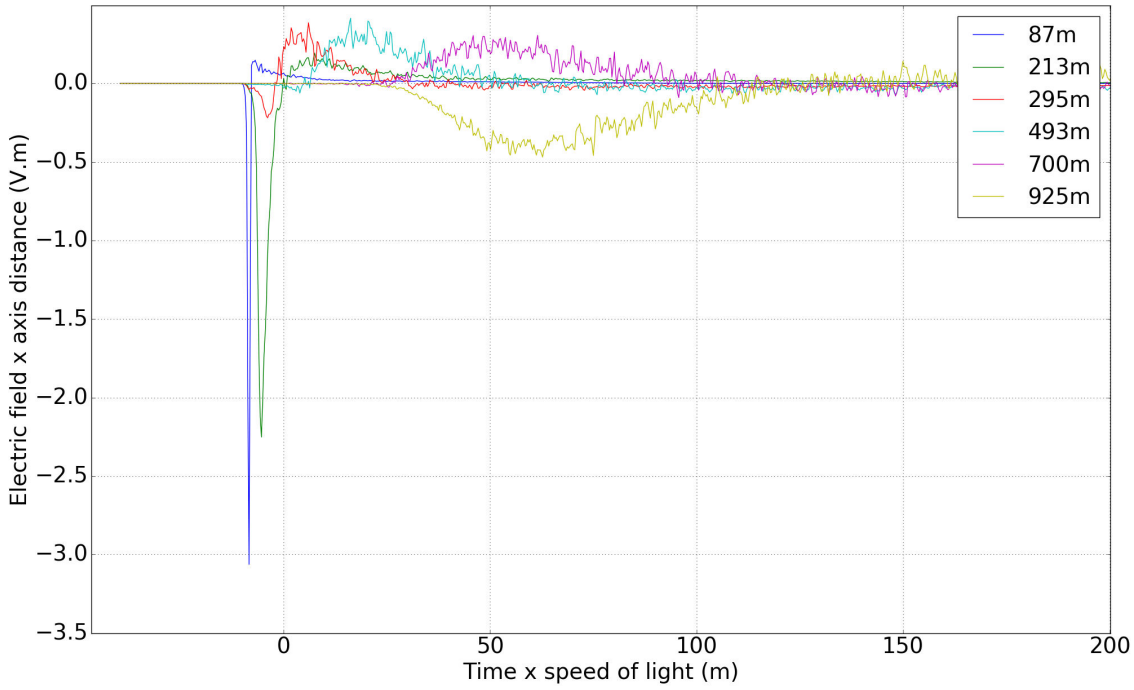


Figure 3.10 *– Radio pulses in the EW polarization for a vertical proton of* 10^{19} eV *as observed at various distances from the shower axis. Simulation done with SELFAS code.*

As for the time domain, where the received signal in the antennas has a characteristic

shape depending on the distance of the antenna from the shower axis, the frequency component of the emitted electric field depends also on the distance of the antenna from the shower axis, as shown in figure 3.11. Unsurprisingly, one can see that antennas

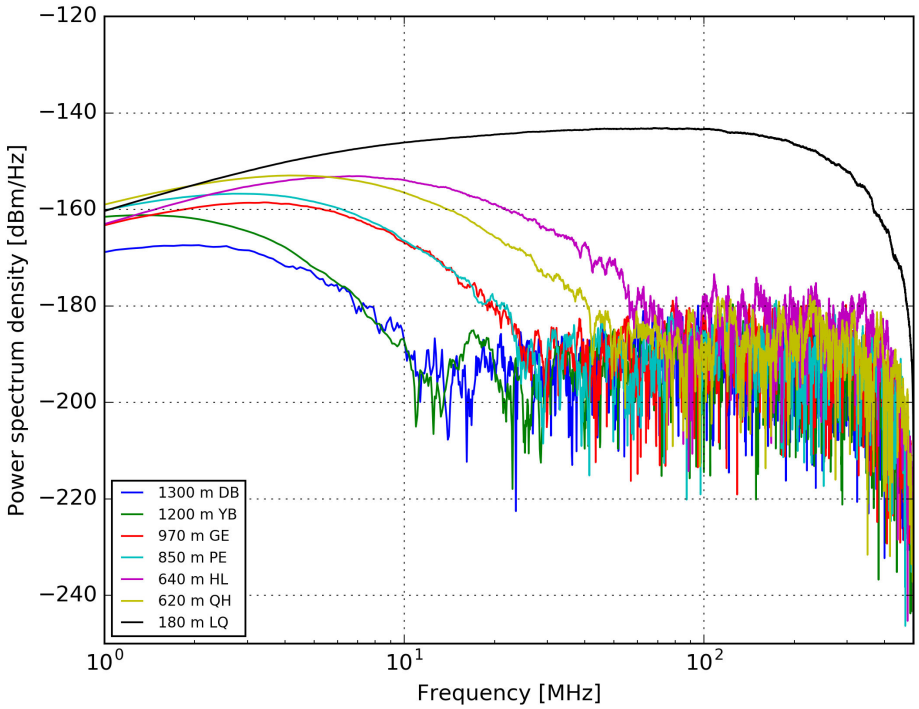


Figure 3.11 – *Simulated power spectrum density as a function of frequency for different shower axis distances, for a proton at* 10^{17} eV *arriving with* $\theta = 41\,°$ *and* $\phi = 145\,°$. *For the farthest antennas DB and YB, the coherence of the signal is lost at* 10 MHz, *followed by GE and PE for which the coherence is lost at* 30 MHz. *For the closest antenna, the signal is coherent over the full used band.*

close to the shower axis will detect a larger electric field than antennas far away from the shower axis. We see that the power spectrum densities (see appendix A for details) quickly drop in the high frequency band with the shower axis distance, while they decrease much more slowly in the low frequency band. In other words, the emission is still coherent at large shower axis distances for the lower frequencies. For antennas close to the shower axis, the source can be assimilated to a point-like source, and the radio emission will be coherent from 1 to 500 MHz while the source begins to be larger (the geometry induces a loss of coherence) for antennas far away from the shower axis distance, and the radio emission becomes non-coherent.

Figure 3.12 presents an event detected by CODALEMA in 2015, with a direction of arrival of $\theta = 32\,°$ and $\phi = 270\,°$, which is approximatively aligned with the geomagnetic vector, and thus which approximatively minimizes the geomagnetic mechanism. The simulation is done for this event, and the polarizations are presented in the figure on the left. The figure on the right corresponds to the real event detected by CODALEMA. Green lines close to colored circles represent the orientation of the polarization of each antennas. By applying the radio reconstruction method explained in chapter 7, we estimate the shower core at $x = (75 \pm 18)$ m and $y = (-229 \pm 17)$ m. We see that the polarizations converge to a zone in agreement with the shower core estimated with the method presented in chapter 7. As said previously, the direction of arrival of this event tends to maximize the Askaryan mechanism compared to the geomagnetic mechanism, and the polarizations do not

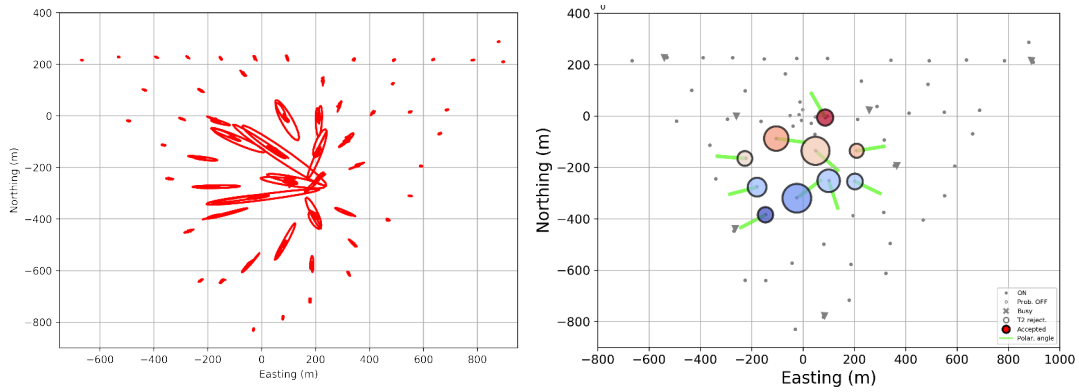


Figure 3.12 *– Left: electric field polarization for a simulated proton of* 10^{17} eV *for the CODALEMA experiment array with the corresponding direction of arrival. The polarizations are radially oriented with respect to the shower axis. Simulation done with SELFAS code. Right: the corresponding real event detected by CODALEMA. Green lines close to colored circles represent the orientation of the polarization of each antennas.*

point exactly to the estimated shower core location, probably because of the geomagnetic mechanism contribution which is still present although diminished.

As is it shown in figure 3.13, there is also a global good agreement between SELFAS3 and ZHAireS for the antenna with a shower axis distance below 200 m. The east-west component of the simulated electric field pulses is presented, corresponding to a vertical shower induced by a proton of 10^{17} eV for an altitude of 1400 m (corresponding to the altitude of the Pierre Auger Observatory). Amplitudes of the signals are multiplied by a factor increasing with the shower axis distance.

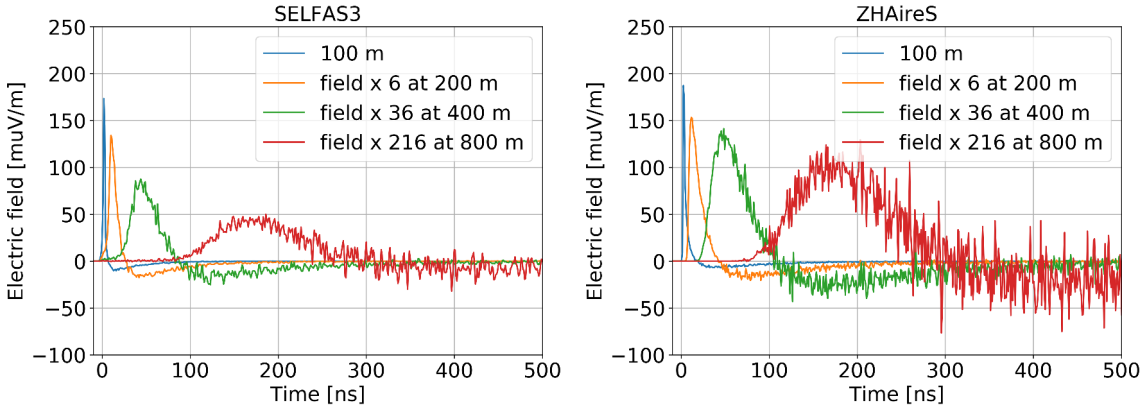


Figure 3.13 *– Comparaison between SELFAS3 and ZHAireS. Left: east-west component of the SELFAS3 simulated electric field pulses, corresponding to a vertical shower induced by a proton of* 10^{17} eV *for an altitude of* 1400 m. *Amplitudes of the signals are multiplied by a factor increasing with the shower axis distance. Right: same figure using ZHAireS. Both simulations have been made with an* $X_{\max}$ *set to* 712 g · cm^{-2}. *Amplitudes of the signals are multiplied by a factor increasing with the shower axis distance.*

However, by zooming in on the x-axis, clear differences appear, as shown in figure 3.14. The figure presents the comparison between SELFAS3 and ZHAireS for antennas close to the shower axis. Dashed lines are SELFAS3 east-west components, and plain lines are ZHAireS east-west components. We see that the pulse shape is different, especially at the

level of the undershoot which lasts much longer for ZHAireS simulations. Amplitudes of pulses are in good agreement, even if a time shift is present between the two codes. The principal difference consist in the down time of the pulse, which is really much longer in the case of ZHAireS simulations.

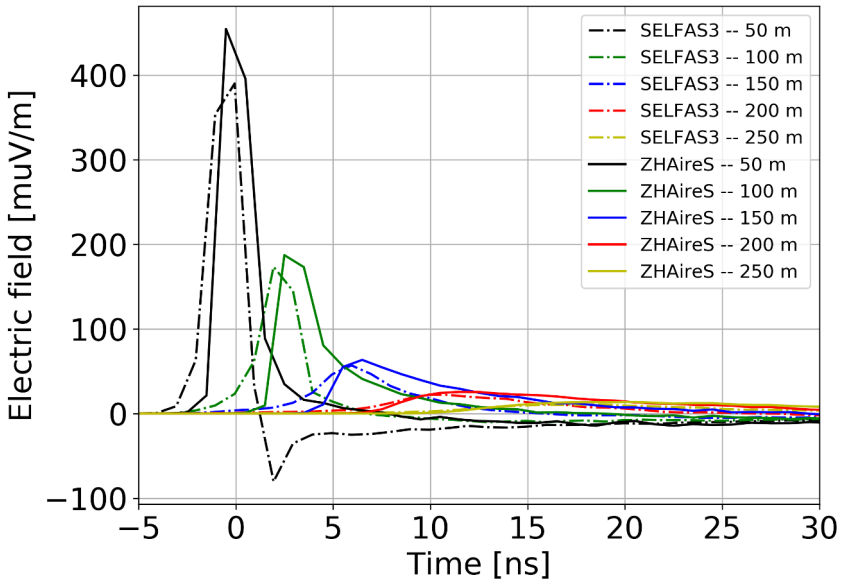


Figure 3.14 *– Comparaison between SELFAS3 and ZHAireS for antennas close to the shower axis. Dashed lines are SELFAS3 east-west components, and plain lines are ZHAireS east-west components.*

The differences in the pulse shape between the two codes reflect a different frequency content, which can be seen in figure 3.15, presenting the PSD comparaison between SELFAS3 and ZHAireS for the east-west component. The amplitude difference is clearly visible (values of the y-axis are directly linked to the amplitude of the signal). The differences in the shape of transients which reflect a different frequency content are also visible. It results a different form of spectra and a different frequency dependence. These results are also valid for the north-south and vertical polarization.

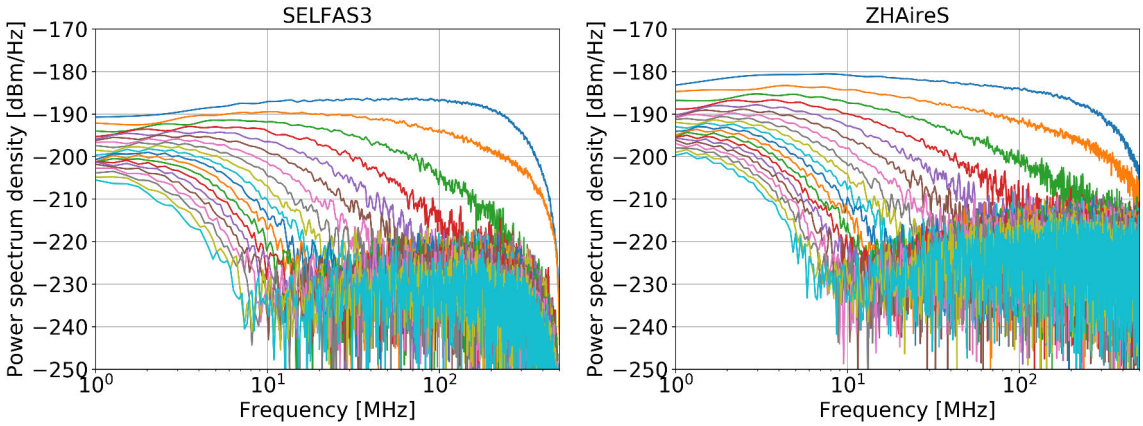


Figure 3.15 *– PSD comparaison between SELFAS3 and ZHAireS. Left: PSD of the east-west components of the SELFAS3 simulated electric field pulses. Right: PSD of the east-west components of the ZHAireS simulated electric field pulses.*

All the simulations and related results presented in this thesis have been made with

the SELFAS3 code. The quantitative and qualitative differences illustrated above (pulse shape, frequency spectra) have undoubtedly an impact on the results. For example, we know that the pulse shape and frequency spectra differences will induce a difference in the shape of the LDF for instance. This difference should impact the comparison between data and the SELFAS3 simulations, and thus impact the estimated parameters of the primary cosmic rays. However, we will show in the following of this manuscript that there is a really good agreement between the SELFAS3 simulations and data.

3.3.2.4. Simulation of the sudden death signal

Figure 3.16 shows the vertical polarization of the electric field obtained with SELFAS3 for an observer at 300 m of the shower core and a vertical proton as primary cosmic ray at 10^{18} eV for the location of the Nançay Radioastronomy Observatory [1].

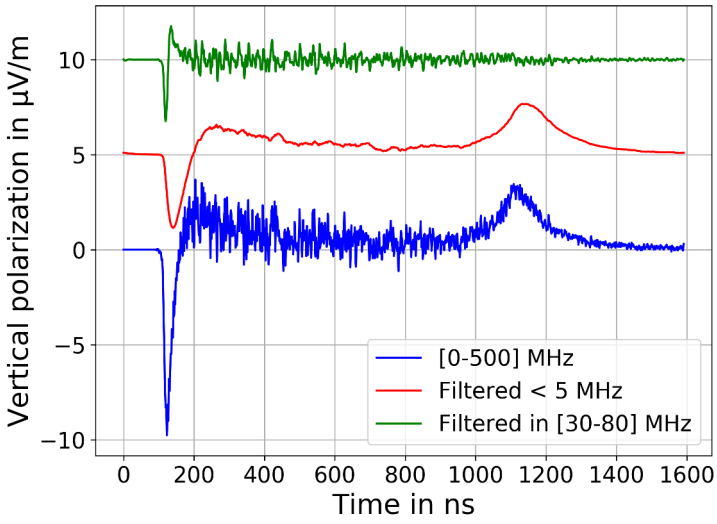


Figure 3.16 *– Vertical polarization of the electric field as a function of time obtained with a SELFAS3 simulation for an observer at* 300 m *of the shower core and for a vertical proton-induced shower at* 10^{18} eV *(blue) for the location of the Nançay Radioastronomy Observatory. Associated filtered responses in different bands (green for* [30 – 80] MHz *and red for* <5 MHz*) show that two pulses can be seen below* 5 MHz*. The red and green curves are positively shifted along* y *for clarity.*

On the full-band trace (blue line), the negative peak at 150 ns is due to the shower development in air. The second, positive peak around 1150 ns has been interpreted as a new mechanism: the effect of the coherent deceleration of the shower front when hitting the ground, called the “sudden death pulse” (SDP) [97, 98]. The SDP arrives 1 µs after the normal pulse, which is consistent with the propagation time from the shower core at ground to the antenna. After filtering (Butterworth 6th order) in different frequency bands, only the shower development peak survives in the MF band while both signals are still detectable for frequencies below < 5 MHz, confirming the interest of using LF antennas. Detecting and studying this phenomenon on one hand and understanding the low frequency counterpart of the radio emission of air showers on the other hand are the goals of the EXTASIS (EXTinction of Air-Shower Induced Signal) experiment, described in Chapter 4.

1. Altitude: 130 m asl, geomagnetic field amplitude of 24 μG with a unit vector oriented as: (B_x = 0.0030, B_y = 0.4548, B_z = −0.8906), x being the east-west direction (positive towards east), y the geographical north-south direction (positive towards north) and z the local vertical).

3.4. Radiodetection of air showers

As described in the previous section, SELFAS3 only predicts the electric field emitted by EAS, without needing the characteristics of the radio detector and its nearby environment. SELFAS computes the components of the electric field at various locations assuming a particular geomagnetic field and a specific atmosphere description. The description of the radio detector and its environment is made separately, the bulk of the work being to optimize the radio detector depending on what it has to detect. Thus, we first define what we want to detect, then we present different types of radio antennas and to finish we present how to detect the radio signal induced by EAS.

3.4.1. Detection in the modern era

As mentioned in section 3.1.1, the first observation of radio transients was carried out by Jelley in 1965. At that time, there was not fast digital oscilloscopes which appeared only during the nineties (actually, the analogical oscilloscopes were fast, but there was nothing to save data quickly). Thanks to the progress made on the acquisition electronic apparatus, the technique of detection has become more efficient, permitting to build comprehensive radio detectors. In the following, we will focus on the detection in the modern era, i.e on the detection carried out by the modern experiments.

3.4.1.1. Features for radio detecting EAS

The radio instrument is commonly composed of a radiator, a signal amplifier, filters, a digitizer (to digitize the signal by step of time). A sketch of a full equivalent diagram of a channel is shown in figure 3.17. The electromagnetic wave induces an electric potential

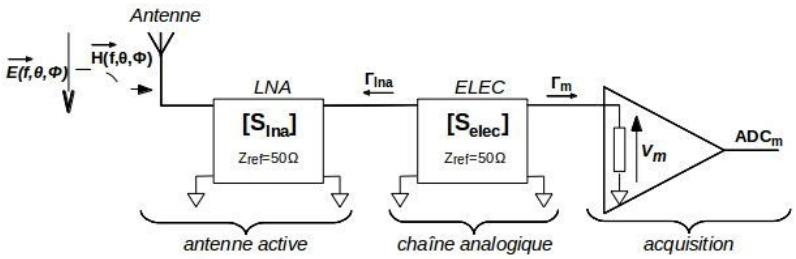


Figure 3.17 *– Typical detection chain in the digital era. From the left: the active antenna composed of the radiator and its low noise amplifier (LNA), the amplification and filtering analogical chain and then the acquisition system [117].*

difference on the radiator, giving an electric voltage of typically 100 µV (obviously, this value depends on the characteristics of the primary cosmic ray which induces the air shower). This voltage is amplified by the low noise amplifier (LNA), filtered or not and then sampled. If we want to retrieve the real observable of the event, namely the emitted electric field, one have to take into account the antenna response (see [118] for a complete calculation), as we will see in 3.4.1.4.

3.4.1.2. Radio antennas for EAS

Several types of radio antennas have been used by the radio experiments. Each type of radio detectors have its advantages and disadvantages (antenna response, sensitivity, cost). An ideal radio antenna would be an antenna with a perfect response over all direction, i.e an isotropic antenna because cosmic rays are coming from all the directions, wide band with a good sensitivity from 1 MHz to 1 GHz because the field from showers is coherent in this frequency band and which would not be sensitive to its environment

conditions. Moreover, to study the polarization pattern of the radio signal at ground, two perpendicular dipole antennas are needed.

Antennas used for the UHECR radio detection, namely on CODALEMA, can be seen as transducers which convert the incoming electric field to a voltage. The transducer is composed of an antenna connected to a Low Noise Amplifier (LNA). The incoming electric field being a fast radio transient, the antenna should be ultra wide band (UWB). The cosmic rays are arriving from all the sky direction, thus the transducer should be isotropic in 2π sr, implying the use of an antenna with a low gain in the whole bandwidth.

Concerning the LNA, and for a given antenna radiator geometry, it is designed in order to flatten the antenna-LNA response and to keep a minimum signal to noise ratio on the widest bandwidth : this is the so-called active antenna concept.

A complete work on the comparison of the different types of antennas has been done by the AERA collaboration [118].

3.4.1.3. Far-field and near-field zones

We commonly use the far-field hypothesis to simplify calculations and theories. The antenna field zone can be divided in two zones, the Fraunhofer zone also called far-field zone, and the Fresnel zone also called near-field zone. The Fresnel zone can again be divided in two zones, the reactive near-field zone and the radiating near-field zone.
The reactive near-field zone is defined as the antenna region where the electric field emitted has a longitudinal component, and thus depends on the distance from the antenna. The reactive near-field zone is limited by $R < 0.62 \times \sqrt{(D^3/\lambda)}$, where D is the dimension of the antenna and λ the wavelength. The boundary between this zone and the Fraunhofer zone is ambiguous and various criteria exist. One of them is that the frontier between these Fraunhofer and Fresnel zones can be expressed as the ratio between the dimension of the antenna and the wavelength. In the Fraunhofer region, the electric field components are only radially oriented and there is no dependence on the distance to the antenna (no longitudinal component, explaining why terms in "[ ]" are zero in equations 3.29, 3.30, 3.31). Thus, the amplitude pattern shape changes with the distance between source and antenna, as shown in figure 3.18.

3.4.1.4. Deconvolving method

Once the data are recorded in voltage, one of the objective is to retrieve the real observable of the event, i.e convert a voltage in V to an electric field in $\mathrm{V} \cdot \mathrm{m}^{-1}$. For that, we need to deconvolve the recorded voltage of the antenna response.
Remark: different terms exist as to the procedure for removing (or adding) the antenna response: to unfold, to deconvolve, to deconvolute... In the following, we choose to use to deconvolve and to convolve, associated with the name deconvolving and convolving.

Simulations of antenna response (gain and impedance) are performed with the version 4 of NEC engine (NEC4) [120] using far field conditions. The electrical field in space for a radio signal is calculated for a given frequency, and the antenna is broken into segments. Then, based on the diameter of the conductor and the wavelength of the signal, the induced voltage and currents at each segment are calculated. The model of the standalone antenna used in the following is presented in figure 3.19. It comprises the near environment of the antenna and a description of the ground.
The recorded voltage is the combination of the electric field $\mathbf{E}$ and the antenna vector effective length $H_{ant}(\nu, \theta, \phi)$, depending on the frequency ν and the direction of arrival of the shower θ and ϕ. For an antenna i in the frequency domain, the Fourier transform of

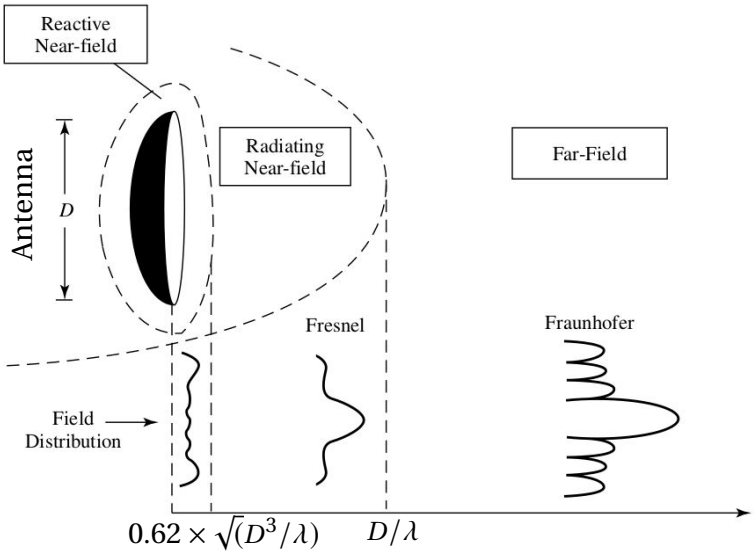


Figure 3.18 *– Amplitude pattern shape as a function of the distance between source and antenna. The antenna has a dimension equal to* D*. Source: [115, 119].*

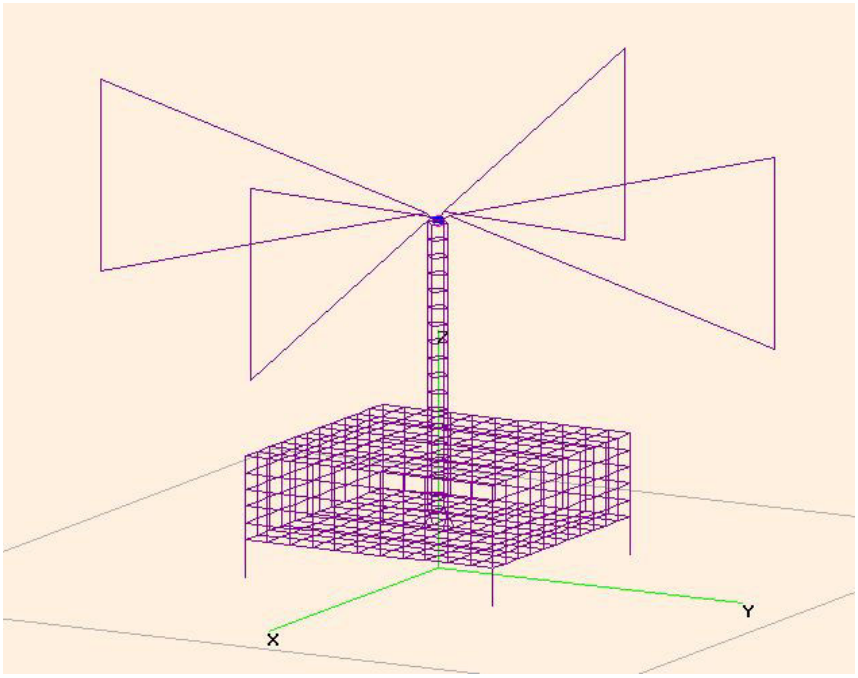


Figure 3.19 *– NEC model of the standalone antenna.*

the recorded voltage in ADC units can be written in spherical coordinates:

$$\mathscr{F}(ADC^i(t,\theta,\phi)) = \mathbf{H}^i_{tot}(\nu,\theta,\phi)\cdot\mathbf{E}(\nu,\theta,\phi) \tag{3.23}$$

where

$$\mathbf{H}^i_{tot}(\nu,\theta,\phi) = \mathbf{H}^i_{\theta,tot}(\nu,\theta,\phi)\overrightarrow{e_\theta} + \mathbf{H}^i_{\phi,tot}(\nu,\theta,\phi)\overrightarrow{e_\phi} \tag{3.24}$$

where

$$\overrightarrow{e_\theta} = \begin{pmatrix}\cos\theta\cos\phi\\ \cos\theta\sin\phi\\ -\sin\theta\end{pmatrix}, \overrightarrow{e_\phi} = \begin{pmatrix}-\sin\phi\\ \cos\phi\\ 0\end{pmatrix} \tag{3.25}$$

The spherical coordinates frame used for the deconvolving method is shown in figure 3.20. For the CODALEMA experiment, radio antennas are composed of Butterfly radiators,

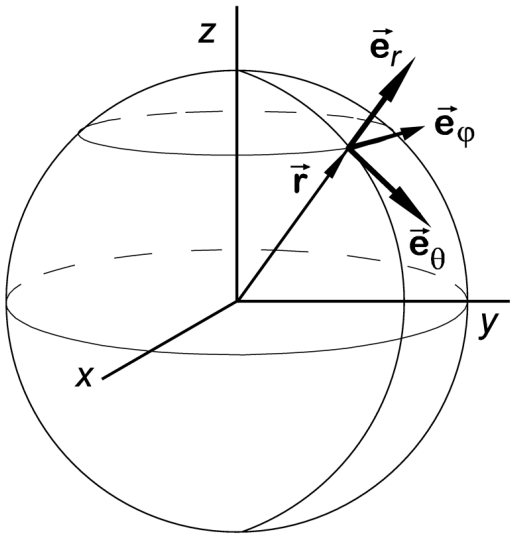


Figure 3.20 *– Spherical coordinates used for the deconvolving method.*

meaning composed of two horizontal perpendicular antennas, noted *i* and *j*. Combining equations 3.23 and 3.24 for the two antennas, we can write:

$$\begin{pmatrix} \mathscr{F}(ADC^i) \\ \mathscr{F}(ADC^j) \end{pmatrix} = \begin{pmatrix} \mathbf{H}^i_{\theta,tot} & \mathbf{H}^i_{\phi,tot} \\ \mathbf{H}^j_{\theta,tot} & \mathbf{H}^j_{\phi,tot} \end{pmatrix} \cdot \begin{pmatrix} E_\theta(\nu,\theta,\phi) \\ E_\phi(\nu,\theta,\phi) \end{pmatrix} \tag{3.26}$$

By inversion of the matrix, we obtain the electric field, and in the time domain, we have:

$$E_\theta(t,\theta,\phi) = \mathscr{F}^{-1}(E_\theta(\nu,\theta,\phi)) \tag{3.27}$$

$$E_\phi(t,\theta,\phi) = \mathscr{F}^{-1}(E_\phi(\nu,\theta,\phi)) \tag{3.28}$$

and finally we have:

$$E_i(t) = \cos\theta\cos\phi E_\theta(t) - \sin\phi, E_\phi(t)[+\sin\theta\cos\phi E_r(t)] \tag{3.29}$$

$$E_j(t) = \cos\theta\sin\phi E_\theta(t) + \cos\phi, E_\phi(t)[+\sin\theta\sin\phi E_r(t)] \tag{3.30}$$

$$E_{vertical}(t) = -\sin\theta E_\theta(t)[+\cos\theta E_r(t)] \tag{3.31}$$

We commonly use the far-field hypothesis, inducing that the terms in "[ ]" are zero since $E_r(t) = 0$. With these equations, the electric field in the time domain for the two polarizations of a butterfly antennas can be calculated. By reconstruction, the vertical component is also deduced. Thus, the electric field in three dimensions for one radio antenna is obtained. The transfer functions $\mathbf{H}^{i,j}_{\theta,tot}(\nu,\theta,\phi)$ and $\mathbf{H}^{i,j}_{\phi,tot}(\nu,\theta,\phi)$ are obtained by simulating the antenna response and by measurements of the different elements of the electronic chain, such as the LNA, the ADC and filters.

However, the deconvolving method has a major disadvantage. Indeed, the signal that has to be deconvolved is composed of the signal coming from the radio emission of the EAS (i.e from a direction (θ, ϕ)), and of the ambient noise (i.e a noise coming from all the directions). The deconvolving is made for a particular direction (θ, ϕ), and so the deconvolving of the signal composed of the ambient noise is incorrect, because in that case, the ambient noise is assumed to come from a particular direction (θ, ϕ). The result obtained after the deconvolving is hard to exploit, since having an artificial noise component added by the deconvolving procedure. An easier way is to work on the simulated electric field, then convolving it with the antenna response, as explained in the following section.

3.4.1.5. Convolving method

Inversely, when working with simulations, the observables are electric fields. In that case, if we want to compare the simulations with the actual event (as explained in chapter 7), we need to convolve the simulations by the response of the antenna. The simulated electric field is only composed of the signal coming from the EAS, and we do not have to deal with the problem of the treatment of the background noise. The sequence of operations is:

1. Simulations give $E_x(t)$, $E_y(t)$, $E_z(t)$. We make a change of reference frame, from Cartesian coordinates to spherical coordinates $E_\theta(t)$, $E_\phi(t)$, $E_r(t)$, by multiplying the electric field in Cartesian coordinates by the rotation matrix depending on the shower geometry. By using the far-field hypothesis, $E_r(t)$ is set to zero, and will not be treated in the following.
2. Calculation of the Fourier transforms, giving the Fourier coefficients a_k associated to each frequency f_k.
3. Interpolation of the transfer functions $\mathbf{H}^{i,j}_{\theta,tot}(\nu,\theta,\phi)$ and $\mathbf{H}^{i,j}_{\phi,tot}(\nu,\theta,\phi)$ for the considered direction of arrival (θ,ϕ) to the frequencies f_k. An angular step of 1 degree and frequency step of 1 MHz are chosen.
4. We obtain $\mathscr{F}(ADC^i)$ and $\mathscr{F}(ADC^j)$.
5. Finally, using the inverse Fourier transforms, we obtain ADC^i and ADC^j using equation 3.23, which represent the ADC channel values in V ·

The ADC channel values can directly be compared with the real event data. An example of the convolving method is shown on figure 7.4 of the chapter 7 page 163.

3.5. Modern radio experiments

The map of figure 3.21 represents the total geomagnetic intensity all over the world, where the location of the main radio experiments are represented by black squares: AERA (Auger Engineering Radio Array) [121], Tunka-Rex (Tunka radio extension) [122], TREND (Tianshan radio experiment for neutrino detection) [123], Yakutsk experiment [124], LOFAR (Low frequency array) [125, 126], LOPES (LOFAR prototype station) [127] and CODALEMA [128]. It is worth noticing that the local geomagnetic field, the altitude of the experiment site and the direction of arrival of the air-shower determine the properties of the emitted radio signal. Indeed, as we have seen in section 3.3.2, the amplitude of the electric signal depends on the strength of the geomagnetic field (actually, on the *geomagnetic angle*). Moreover, the orientation of the geomagnetic field plays also a role in the amplitude of the electric field and can generate an azimuthal asymmetry in the radio detection of EAS, as discussed in section 3.2.1. It is worth noting that Tunka-Rex and Yakutsk, in the middle and north of Siberia, are on favourable sites considering the strength of the magnetic field (the larger the intensity of the geomagnetic fields, the larger the Lorentz force, and the larger the amplitude of the signal), while AERA is the one of the seven experiments where the strength of the geomagnetic field is the smallest. The reader can also refer to the review of Frank G. Schröder [130] for more details.

3.5.1. AERA

AERA [121] is located within the Pierre Auger Observatory site, in Argentina. It aims at studying the radio emission of EAS in the so-called “infill” area of the Pierre Auger Observatory, set up for lowering the energy detection threshold thanks to a higher density of water Cherenkov tanks, and mostly to cross-check the mass sensitivity obtained with the radio method with the Auger detectors. It is composed of 153 radio detection stations spread over 17 km^2, as shown in figure 3.22. Figure 3.23 presents the two types of radio

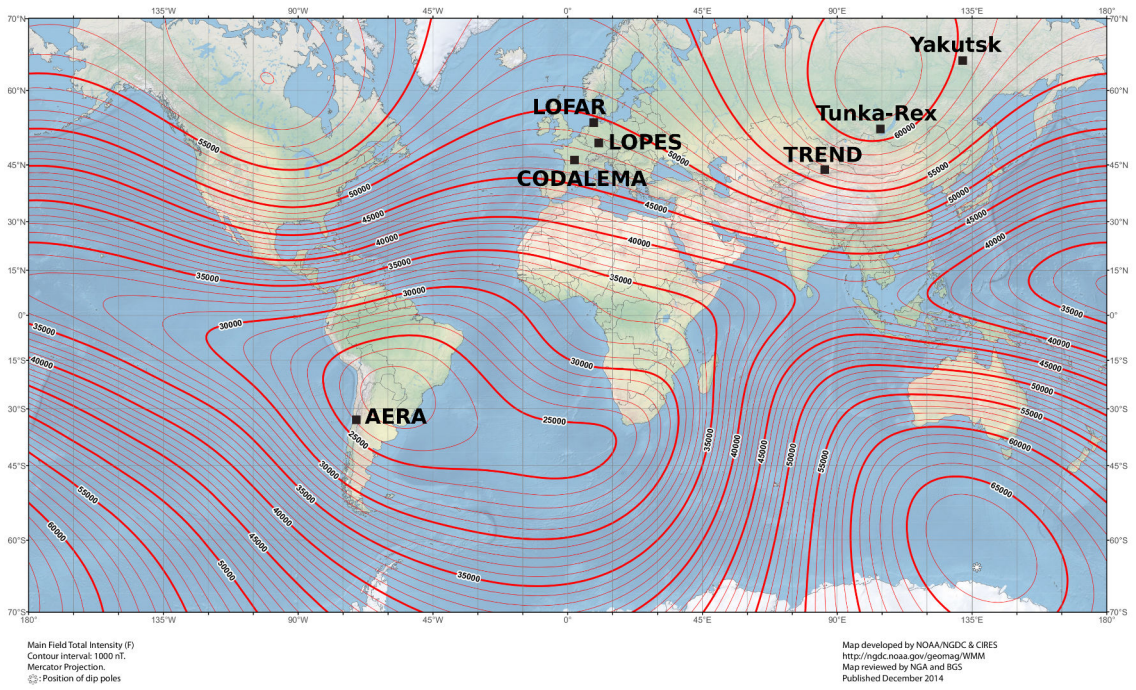


Figure 3.21 *– Total geomagnetic intensity map [129]. The black squares denote the location of radio experiments studying cosmic-ray air showers. See text for more details. The geomagnetic field intensity for each site is: AERA:* 24 µT*, Tunka-Rex:* 60.5 µT*, TREND:* 56.5 µT*, Yakutsk experiment:* 59.5 µT*, LOFAR:* 50 µT*, LOPES:* 49 µT *and CODALEMA:* 47 µT*.*

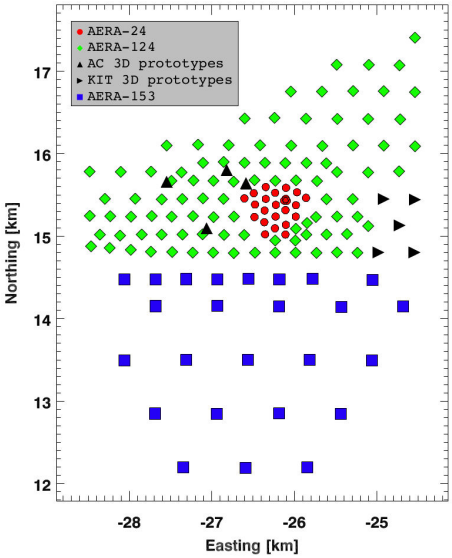


Figure 3.22 *– Map of the AERA experiment, with the different deployment stages [71]. The red circles represent the LPDA antennas. The green diamonds represent the Butterfly antennas externally and internally triggered, the blue squares represent the Butterfly antennas only internally triggered.*

antennas used at AERA, operating in the frequency band [30 – 80] MHz. The radio array works in parallel with the other instruments of the Auger site, such as the fluorescence telescopes and the water-Cherenkov tanks, allowing cross-calibration of the energy of the different observables: energy, X_{max}. Among other results, AERA was the first able to quantify the contribution of the charge excess mechanism to the radio emission [132].

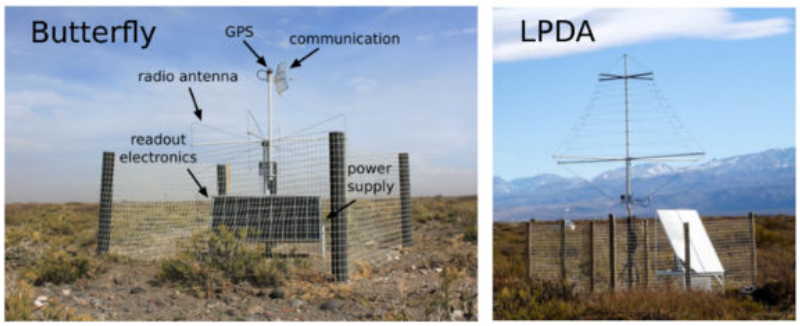


Figure 3.23 *– Photography of radio antennas at AERA [131].*

Concerning the estimation of the primary energy, AERA has a resolution of 17 % [133]. The energy deposited per unit area, also called the *energy fluence*, at the locations of the AERA radio detectors is determined via the reconstruction of the electric field vector. The energy fluence is then integrated to estimate the total deposited energy [134, 135].

3.5.2. Tunka Rex

Tunka-Rex [122] is the radio extension of the Tunka-133 and Tunka-Grande experiments. It is located near the Lake Baikal in Siberia. Similarly to AERA, the association of the radio array to other instruments, namely air-Cherenkov telescopes, permits cross-calibrations of the different features of the primary cosmic-ray. Currently, by the association of Tunka-Rex, Tunka-133 and Tunka-Grande, the energy threshold for radio detection is around 10^{17} eV. Tunka-Rex is composed of 38 dual-polarized radio antennas (Salla antennas) over 1 km^2, as shown in figure 3.24. A photography of one cluster center

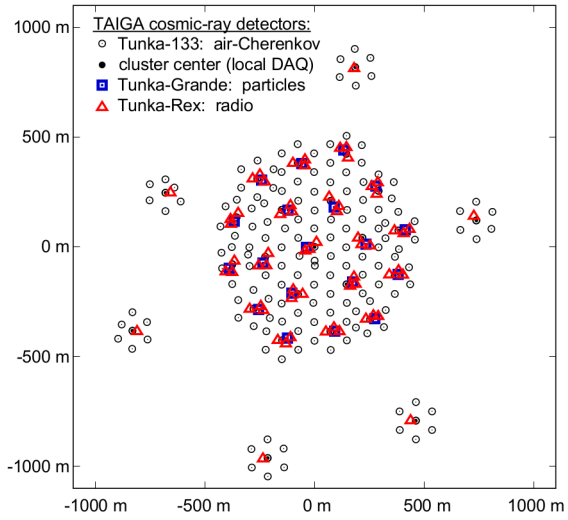


Figure 3.24 *– Map of Tunka experiment [136].*

with three antenna stations, a particle-detector station and underground scintillators, and a photomultiplier detector is shown in figure 3.25. From their set of 3-years data, they have shown that a reconstruction of the shower maximum ($X_{\max}$) with a precision of 35 g $\cdot$ cm^{-2} is possible [122, 138, 139].

3.5.3. TREND

TREND [123] is located in the XinJiang province in China and is considered as a pathfinder for GRAND (giant radio array for neutrino detection, which should be com-

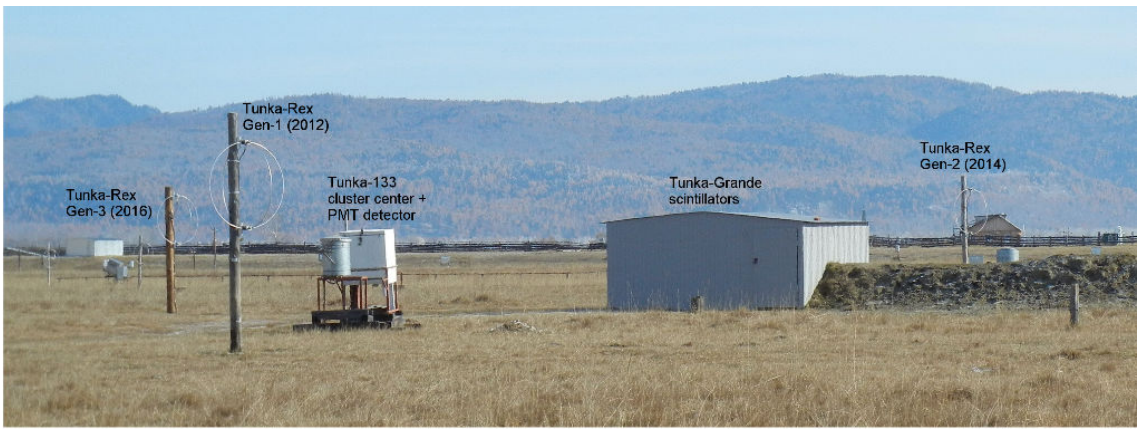


Figure 3.25 *– Photo of one cluster center at Tunka, with three antenna stations, a particle-detector station and underground scintillators, and a photomultiplier detector [137].*

posed of 100,000 radio antennas over 200,000 km^2 [140, 141]). It is currently composed of 50 self-triggered radio antennas spaced-apart from 50 m and operating in [50 − 100] MHz [142], as shown in figure 3.26.

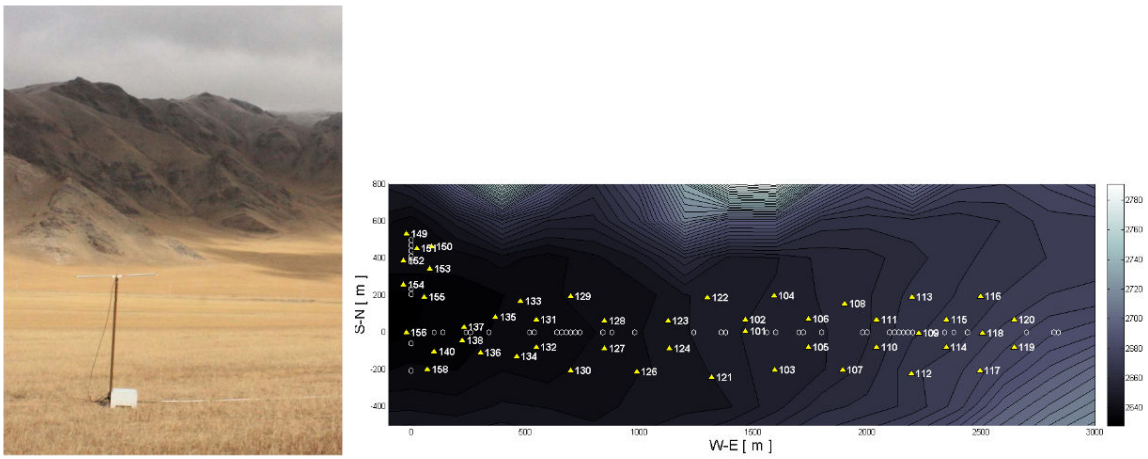


Figure 3.26 *– Left: photo of one radio antenna at TREND. Right: TREND layout [143]. The ground extension is* ∼ 2.7 km *for easting and* ∼ 0.8 km *for northing.*

3.5.4. Yakutsk

Yakutsk experiment [124] is running since 1972, and is located close to the Yakutsk city. The experiment is composed of particle detectors, air-Cherenkov telescopes and few antennas operating around 32 MHz, as shown in figure 3.27. It has hosted low-frequency radio instruments operating at 1.9 MHz at the beginning of the experiment.

3.5.5. LOFAR

LOFAR [125, 126] is a digital radio interferometer dedicated to radioastronomy. It is composed of several stations distributed among 48 sites, also spread over 5 European countries: 40 in Netherlands, 5 in Germany, 1 in France, England and Sweden. LOFAR has a key program on cosmic rays. One of the advantage of LOFAR is the hight antenna density, allowing a fine measurement of the variation of the electric field with respect to the shower axis distance, see figure 3.28 right for the antenna used. One station is composed of 96 Low Band Antenna (LBA) and 96 High Band Antenna (HBA). The LBA operates in

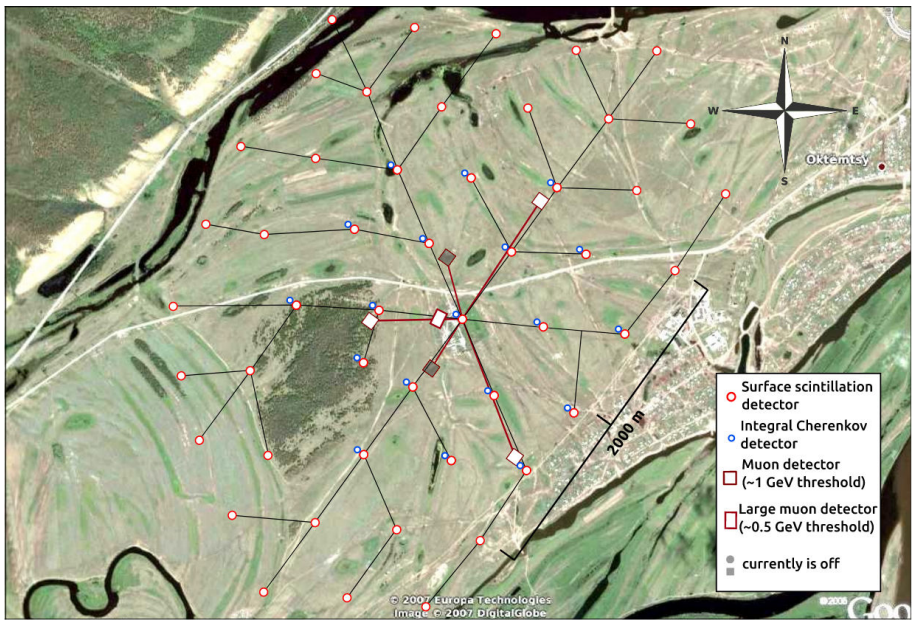


Figure 3.27 *– Photography of the layout of Yakutsk experiment. Adapted from [144].*

[10 – 80] MHz and the HBA operates in [120 – 240] MHz. The LBA are equipped with two simple dual polarization dipoles above a conducting ground plane, which could be related to the shape of the antenna of the Compact Array (see chapter 4). The HBA are composed of 16 bowtie shaped dual dipole antennas. They are assembled and arranged in a four-by-four grid. It is worth noticing that, contrary to CODALEMA, the radio signal of EAS is not measured over the entire frequency band [20 – 200] MHz with only one antenna, but it is measured separately by two types of antennas. We will see in chapter 7 that the measurement of the radio signal over the whole band with the same antenna is a specificity of CODALEMA and constitutes a major advantage in the reconstruction of the characteristics of the primary cosmic rays.

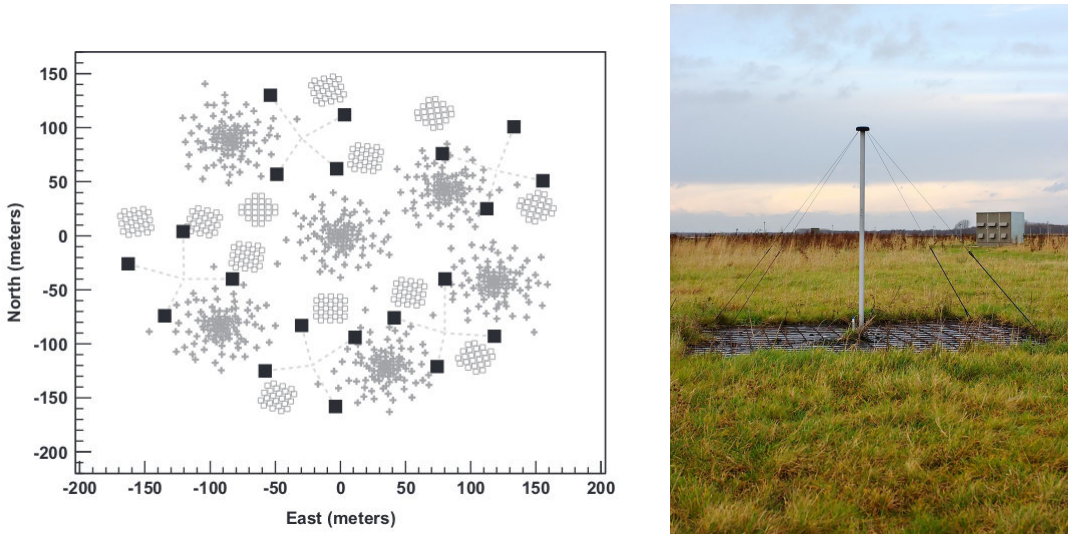


Figure 3.28 *– Left: the LOFAR core, near Exloo in Netherlands [145], containing the particle detectors (black squares) of the LOFAR Radboud air-shower array. The crosses and open squares represent the two different types of LOFAR radio antennas (LBA and HBA respectively). Right: radio antenna of LOFAR.*

It was the first experiment designed to measure radio emission from cosmic-ray induced air showers in parallel with interferometric observations. Thanks to the scintillators of the LOFAR Radboud air-shower array (LORA, see figure 3.28 left), the EAS are identified when a radio signal is in coincidence with the signals from the particle detectors.

3.5.6. CODALEMA & EXTASIS

The CODALEMA [128, 146] (COsmic ray Detection Array with Logarithmic Electro-Magnetic Antennas) experiment is hosted at the radio Observatory of Nançay, and is the support of the EXTASIS (EXTinction of Air Shower Induced Signal) experiment [9]. These two experiments will be discussed in detail in the chapter 4.

3.5.7. Radio experiments at the South Pole

The South Pole is one of the ideal locations for radio experiment, since there is few man-made emitters (mainly the galactic and atmospheric noise contribute to the background noise). The ARA (Askaryan Radio Array) aims to study high-energy neutrinos by detecting the radio pulses generated as they go through the ice. It is located at the South Pole, near the site of the IceCube experiment [147, 148]. The idea is to perform radio detection in a dense medium such as ice, which would allow, thanks to the 1 km-radio attenuation length [149], a sufficient sensitivity to study the ultra-high energy neutrinos. As ARIANNA (Antarctic Ross Ice Shelf antenna neutrino array) [150], it aims to prove that the radio-detection in ice can be performed for 10^{16}-10^{19} eV neutrinos.
The antennas of ARA are planned to work in the range [150 – 850] MHz, spaced from each other by one kilometer, and buried in the ice ($\sim$ 200 m underground). For ARIANNA, the frequency range is [50 – 1000] MHz for the in-ice 8×4 antennas.
Contrarily to ARA and ARIANNA, the ANITA (Antarctic Impulsive Transient Antenna) experiment uses an antenna array of 32 dual-polarized quad-ridged horn antennas operating in the [200 – 1200] MHz band, and equipped on a high altitude balloon ($\sim$ 36 km). ANITA aims to detect the Askaryan emission induced by the interaction of ultra-high energy neutrino in the Antarctic ice (up-going event), but also to detect the radio emission coming from the geomagnetic mechanism which is reflected off toward the balloon (down-going event) [151]. It probes the extremely high energy range $> 10^{19}$ eV. The latest results on two anomalous up-going event can be found in [152, 153].

3.5.8. GHz-radio experiments

The detection of a radio signal in the [1.5 – 6] GHz was reported in [154] in 2008. This signal has been interpreted as coming from the Molecular Bremsstrahlung Radiation (MBR) emission mechanism, which produces an unpolarized and isotropic radio signal. The MBR mechanism is supposed to occur during EAS development in the atmosphere. Several experiments have been developed to study this mechanism such as CROME (Cosmic-Ray Observation via Microwave Emission) [155] and EASIER (Extensive Air Shower Identification with ElectronRadiometer) [156]. This mechanism is an extension of the geomagnetic mechanism occurring in the EMF band [157], and the associated radio emission is strongly beamed along a Cherenkov cone. However, the associated wavelength is similar to the distance inter-particles leading to incoherent emission. This incoherence makes it difficult to detect the signal in this frequency range. Indeed, the MBR was not detected by the dedicated experiments such as CROME and EASIER [158].

3.5.9. Comparison

Figure 3.29 presents the ground extension of AERA, LOFAR, T-Rex and CODALEMA/EXTASIS experiments. The operating frequency range of each experiment is indicated in the legend, and is summarized in figure 3.30. One sees that CODALEMA measures EAS

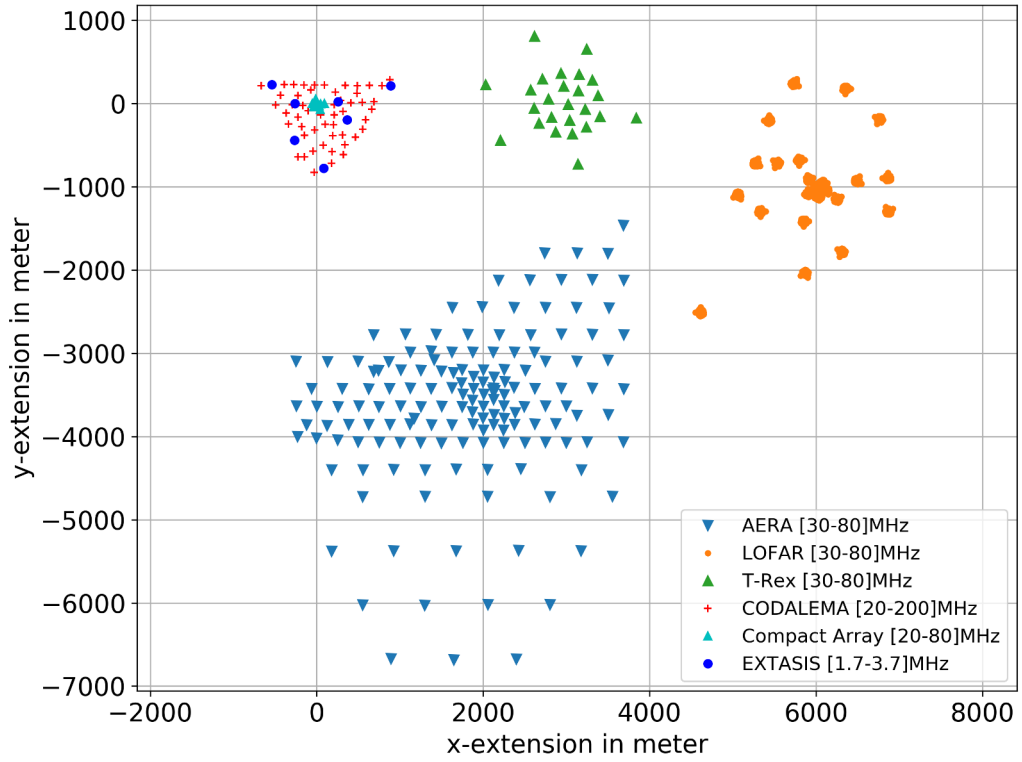


Figure 3.29 *– Ground extension of some modern radio experiments.*

from 1 to 200 MHz, and is the only one that can do radio-detection in a so wide frequency range with the same type of antennas. We also see that AERA has a large area of around 17 km^2 with sparse array, while LOFAR ground extension is smaller but with a high density array. As we have seen at the beginning of this chapter, the covered surface restricts the primary cosmic ray energy, corresponding to the upper limit of the energy range, and that the distance between detectors gives the lower limit of the energy range. Moreover, the smaller the distance between the detectors, the higher the precision on the electric field profile of individual showers.

Thus, each experiment is sized for the studies it wishes to conduct on the characterization of EAS parameters.
Figure 3.31 presents the typical dimensions of an EAS, and their associated frequency emission. Depending on the frequency band used, different parts of the shower will be probe.

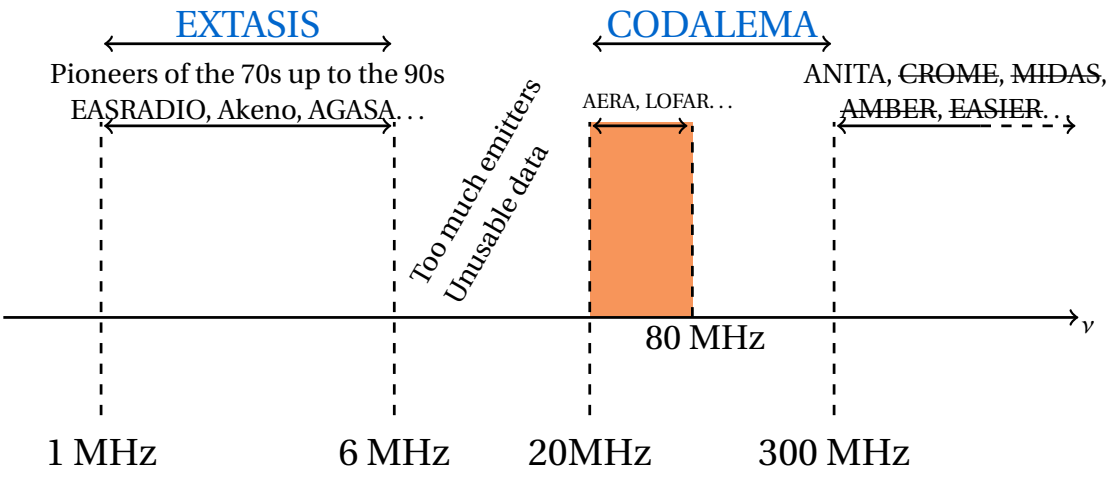


Figure 3.30 – *Frequency band of some modern radio experiments.*

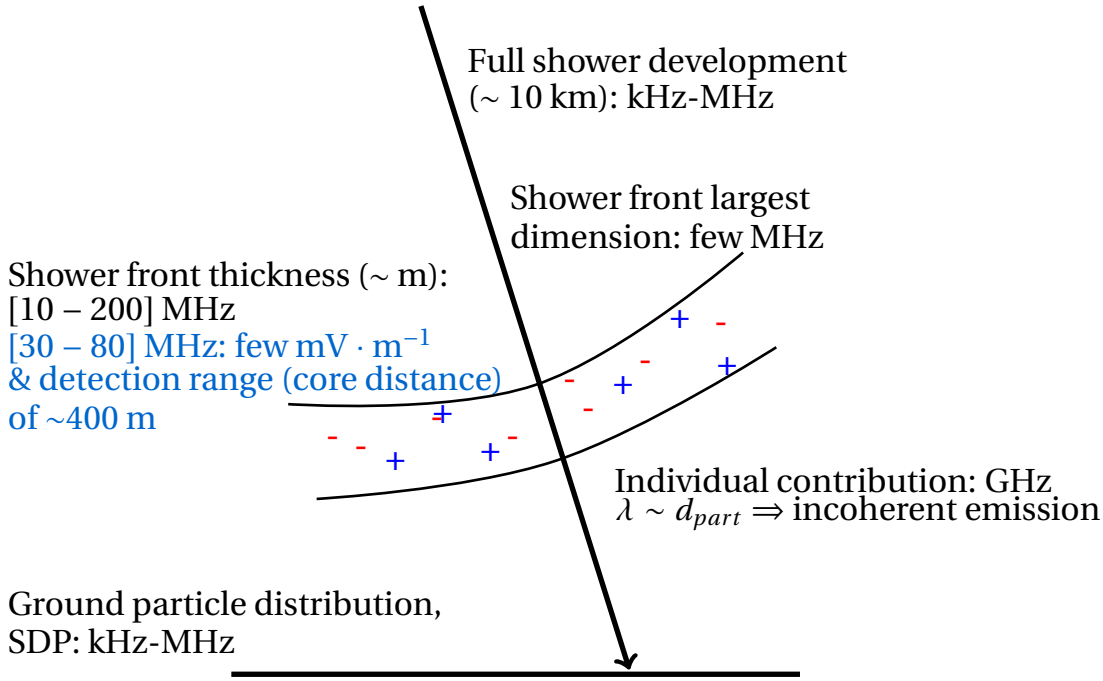


Figure 3.31 – *Typical shower dimensions and their associated frequency emission.*

"The true method of knowledge is experiment."

William Blake

CHAPTER 4

CODALEMA and EXTASIS: instrumental setup

Outline

The CODALEMA (*COsmic ray Detection Array with Logarithmic ElectroMagnetic Antennas*) experiment located within the Nançay Radio Observatory is one of the pioneering and promotor experiment of this revival of the radio detection of EAS, which is today adopted by a flurry of experiments in the world. Coupled with an increasingly sophisticated understanding of the processes involved allowing the use of high-performance simulation codes, EAS radio-detection now reaches a level of maturity sufficient to match the more traditional methods for the determination of the fundamental properties of UHECR. This chapter will focus on the description of the CODALEMA and EXTASIS (*EXTinction of Air Shower Induced Signal*) instrumental setups. CODALEMA and EXTASIS aim at exploring some possible solutions to improve the radio detection method:

— To use clever triggering algorithms to solve the autonomous detection problem such as the composite trigger developed by the group, which is a combination of several antennas in a clustered way. It could also permit to detect cosmic ray events without particle detectors, for example. This composite trigger will not be studied in this manuscript.

— Concerning the detection range, it has been demonstrated in [159, 9] that it is much larger at low-frequency (below 5 MHz) than in the classical band [30 – 80] MHz. This is one of the goals of the EXTASIS project. Let us note that the use of the composite trigger, namely phased combination of several antennas, will also increase the detection range in the classical band.

— As it is shown in appendix E, above 120 MHz the signal strength is clearly dependent on the distance to the shower axis. Moreover, the radio Cerenkov ring can

be well observed, as shown in figure E.4. This observation could help to better constrain the reconstruction of the shower core (see chapter 7), but also to solve the problem of the antenna density at short distance of the shower axis. The description of the electric field profile could be better defined over a large bandwidth where its fast variations are observed.

— Finally, a three-fold antenna has been installed [160] close to the center of the particle detectors of CODALEMA, in the middle of the covered field, in order to measure the complete 3D electric field produced by air showers. The idea is to check the validity of the far-field hypothesis commonly used by the radio community.

4.1. Nançay observatory

4.1.1. Facilities dedicated to cosmic ray observation

Since 2002, the Nançay Radio Observatory hosts the CODALEMA experiment, which is one of the pioneering experiments that have participated in the rebirth of radio detection of cosmic rays at the beginning of the 21st century. Over the years, the experiment has seen the development of a large collection of detectors, intended to study the properties of the radio emission associated with cosmic-ray induced air-showers in the energy range from 10^{16} to 10^{18} eV. The original array of CODALEMA made of logarithmic antennas rapidly replaced by short active dipole antennas, which in turns were replaced by the so-called Butterfly antennas. Figure 4.1 shows the experimental area at Nançay. In its current version, the instrumental setup is composed of:

— A square array (0.34×0.34 km^2) of 13 particle scintillator detectors.

— A set of 57 so-called “standalone” cross-polarized antennas (SA), operating in the [20 – 200] MHz band (noted EMF band), distributed over ~ 1 km^2.

— A so-called “Compact Array” of 10 cross-polarized antennas, arranged in a star shape of 150 m extension and whose signal acquisition (in [20 – 200] MHz band, noted MF band) is triggered by the particle detector.

CODALEMA is today the supporting experiment of the EXTASIS experiment [9], an array of 7 low-frequency (LF) antennas triggered by the particle detector, which takes advantage of its existing infrastructure. The LF antenna locations have been chosen to cover the overall Nançay area and also in a way that each LF antenna has a MF standalone antenna nearby. LF antennas are named DB, YB, GE, PE, HL, QH and LQ.

As show in figure 4.1 by green and purple circles, the FR606 LOFAR station and the NenuFAR radio telescope [163] are completely surrounded by CODALEMA. NenuFAR is an extension of LOFAR, comprising 96 islets of 19 antennas similar to those of the Compact Array. NenuFAR is part of the LOFAR Super Station in Nançay, since the Nançay Radio Observatory hosts the FR606 international LOFAR station. Thanks to their proximity, CODALEMA and NenuFAR will constitute an array of very high radio antenna density operating in the largest frequency bandwidth ([1 – 200] MHz). This association will lead to studies of several features of the electric field emitted by EAS, and could constitute in coming years a pathfinder for existing and future large radio experiment such as SKA-low [164, 165, 166].

Concerning the environment of SA, they are dispersed over ~ 1.1 km^2. This environment is variable with antennas exposed to the whole sky, others in the middle of a fairly dense forest, others near deciduous trees. Previous studies have concluded the weak influence of the environment [167], allowing to deploy the antennas in different environments without affecting the measurements of the electric field.

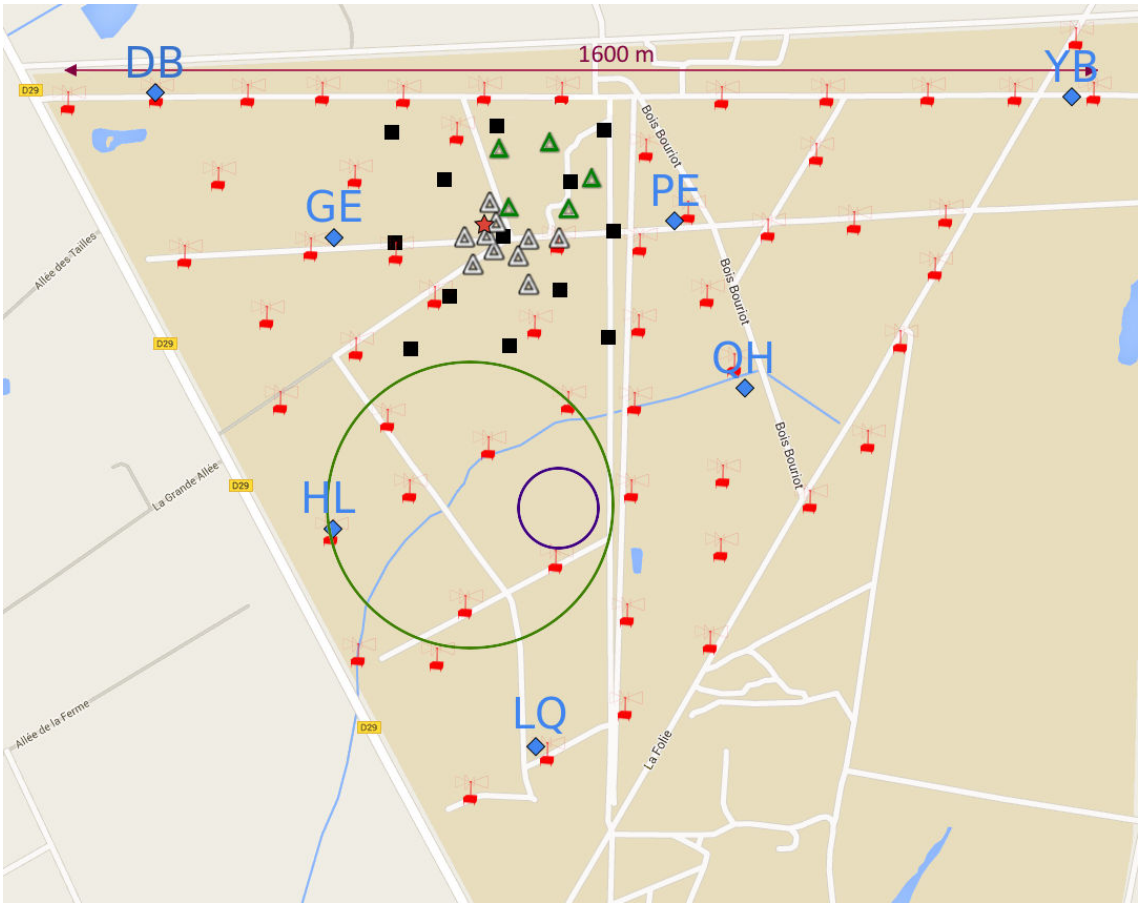


Figure 4.1 *– Experimental setup at Nançay [161, 162]. Red icons represent the 57 standalone antennas of CODALEMA, black squares the 13 scintillators, white triangles the 10 antennas of the Compact Array. The blue diamonds feature the low-frequency antennas and the red star the three-fold antenna. Green triangles are the 5 antennas of the "composite trigger array", which is not described in this chapter. The purple and green circles represent respectively the position of the LOFAR FR606 local station and NenuFAR.*

4.1.2. Sky and radio emitters

Figure 4.2 presents a typical day (top, red) and night (lower, blue) power spectrum density of the background noise of Nançay, over a wide frequency band. The y-axis represents the background noise level in dBm/Hz. It is important to note that the spectrum floor is driven by the ambient noise (galactic, atmospheric...) which is inherent to the environment of the experimental site, while the spectrum lines are due to anthropic emitters, for example AM and FM broadcastings.

For the case of CODALEMA which is operating in the EMF band, it appears obvious that the FM broadcasting will pollute our measurements. To overcome this problem, we will have to clean our frequency band by filtering the recorded signal, as explained in appendix A page 229. For the EXTASIS case, the cleanest band at low frequency lies in the range [2 – 4] MHz during day (referred to as Low Frequencies (LF) in the following), as it will be discussed in chapter 8. Moreover, if we now compare the day and night power spectrum density for the LF and EMF bands, one sees that the EMF spectrum floor is smaller by at least 10 dB compared to the spectrum floor in the LF band, indicating that a detection of transients will be easier in the EMF band than in the LF band. Let us now focus on the EMF band. One sees that the MF spectrum floor is the same during the day and during the night. Below 20 MHz, there are more emitters during day than during night, principally due to the AM broadcasting. The same kind of observations for

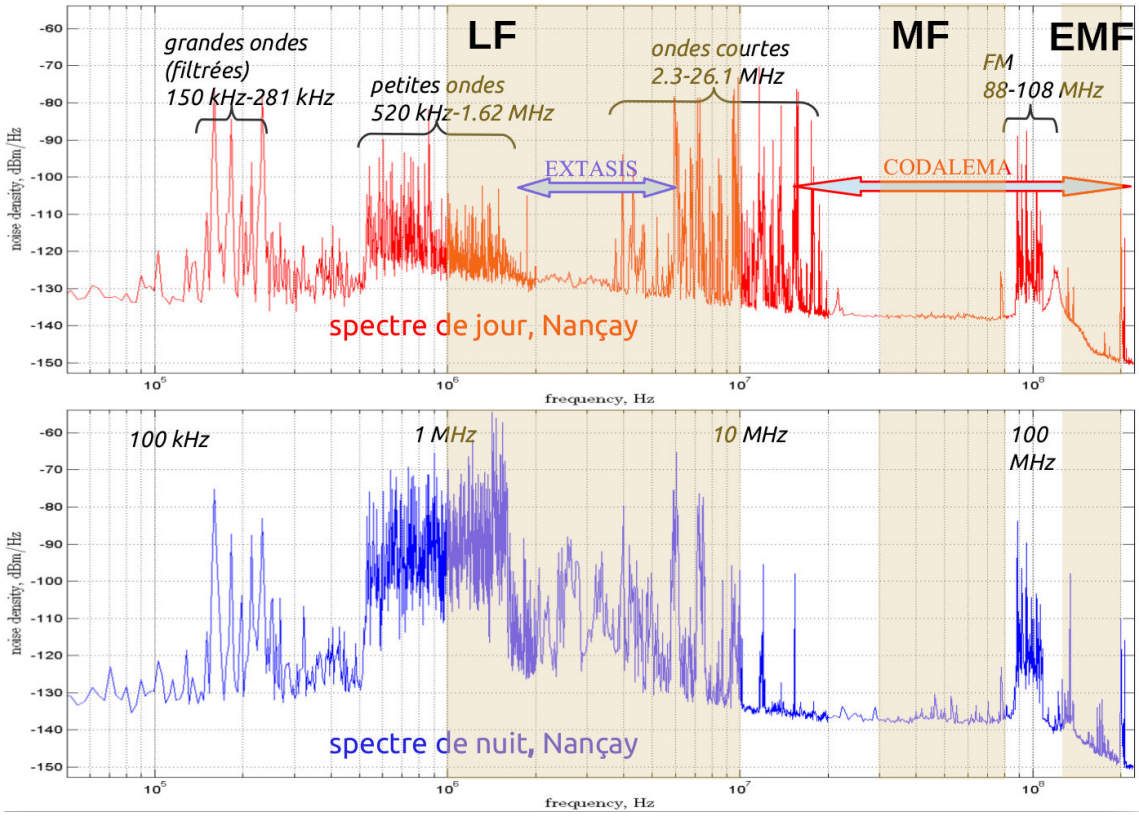


Figure 4.2 *– Day (top, red) and night (bottom, blue) power spectrum density of the background noise of Nançay. The y-axis represents the background noise level in* dBm/Hz. *A high value indicates a large background noise, and inversely. Figure adapted, extracted from a presentation of D. Charrier.*

the LF band will be largely discussed in chapter 8 dedicated to the EXTASIS results, and attributed to the behaviour of the atmosphere at these frequencies.

The next two sections will be dedicated to the description of the instrumental setup of CODALEMA and then EXTASIS.

4.2. CODALEMA

4.2.1. Particle detector array

This array is composed of 13 scintillators covering 340×340 m^2 with a variable spacing (approximately 120 m on the diagonal and 170 m on the NS and EW directions, see figure 4.1). It aims to provide the arrival direction, the size, the core location of the EAS and the energy of the primary cosmic ray [168]. The surface covered by the scintillator array restricts the upper limit of the studied energy range, and the distance between detectors gives the lower limit. Its characteristics lead to an energy threshold (lower limit) around the knee (10^{15} eV) and a maximum efficiency at 10^{16} eV (see figure 9.2 page 198).

One scintillator counter is composed of a NE102A plastic scintillator of 80×80 cm^2 and 4 cm thick associated with two XP3462 photomultipliers, as shown in figure 4.3. To maximise the light collection, the plastic scintillator and both photomultipliers are contained in an inverted pyramid stainless steel box coated with white diffusing paint, itself contained in an outer plastic box. A black tarpaulin covers the whole for weather and light protection. Each scintillator device is connected to a central shelter by a $\sim$ 300 m pair of cable. The two photomultipliers work at two different gains, high and low, for

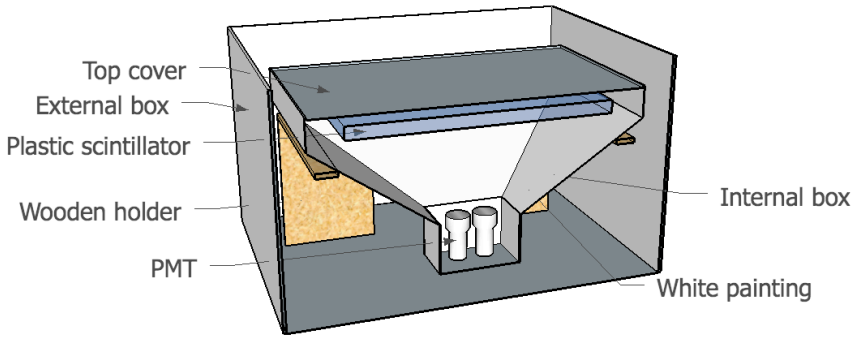


Figure 4.3 – *Schematic cut view of a scintillator.*

a matter of dynamics, i.e in order to have a dynamic range from 0.3 to 3,000 Vertical Equivalent Muon (VEM, assimilated to the charge deposited by a single vertical muon in the scintillator). Signals from the scintillator are connected to 8 4-channels, 12bits, 1 GS $\cdot$ s^{-1} digitizer boards (so-called MATACQ boards) [169].

The MATACQ boards are externally triggered by a 16-fold multiplicity module. The multiplicity module compares the high gain photomultiplier signal to a defined level and calculates the number of scintillators matching this criterion (called the multiplicity). A trigger is produced if the multiplicity is greater or equal to 5 scintillators among 13 within a 600 ns time window. The waveform is digitized on 2,560 points, corresponding to a 2.56 µs signal length. This setup gives an average event rate of 40 events per hour. The trigger signal is distributed to different instruments, currently the Compact Array, the LF antennas and a GPS datation station called MC250 dedicated to a precise GPS datation and providing a time reference for all the instruments. The reader can refer to chapter 6 for the description of the method used to reconstruct the arrival direction and to chapter 9 for the method used to estimate the core position and the energy of the primary particle.

Figure 4.4 presents the trigger rates of the scintillators since 2012. The trigger rate is stable over the years. The increase in the trigger rate near the beginning of 2014 coincides with a change in the multiplicity criterion for the trigger. Before 2014, the five central scintillators were required to build the trigger. Since the beginning of 2014, the trigger is built when at least any 5 out of the 13 scintillators are triggered within a given time window. Moreover, the periods where the trigger rate is zero correspond to periods of maintenance. The periods where the trigger rate is higher than the normal behaviour correspond to periods of test, where the required multiplicity to build the trigger has been lowered. For instance, during the period between November 2012 and November 2013, a lot of modifications concerning the acquisition system of the scintillators and their power supplies have been made.

Figure 4.5 presents the zenithal and azimuthal distributions observed for the internal events. The internal events are defined as events for which the largest measured signal is recorded in one of the five scintillators constituting the inner part of the array, allowing a reliable reconstruction of the core position and of the energy. The zenithal distribution is fitted with the empirical function given in [83]:

$$\frac{\mathrm{d}N}{\mathrm{d}\theta} = \frac{(a + b\theta)\cos(\theta)\sin(\theta)}{1 + \exp\left(\frac{\theta-\theta_0}{\theta_1}\right)} \tag{4.1}$$

where the $\cos(\theta)\sin(\theta)$ term describes the cosmic ray flux, $a = 2.5 \times 10^5$, $b = -4.3 \times 10^3$, $\theta_0 = 31.0\,°$ and $\theta_1 = 10.5\,°$. As expected, the azimuthal distribution is uniform. This confirms that the asymmetry seen in figure 3.3 page 51 can be attributed to the radio signals, i.e to the radio emission mechanisms.

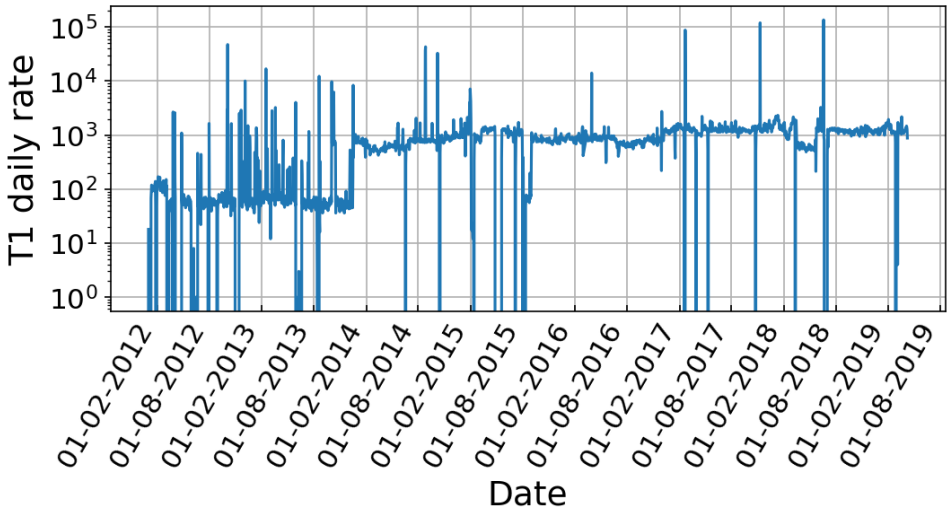


Figure 4.4 *– Trigger rates of the scintillators over 8 years of functioning.*

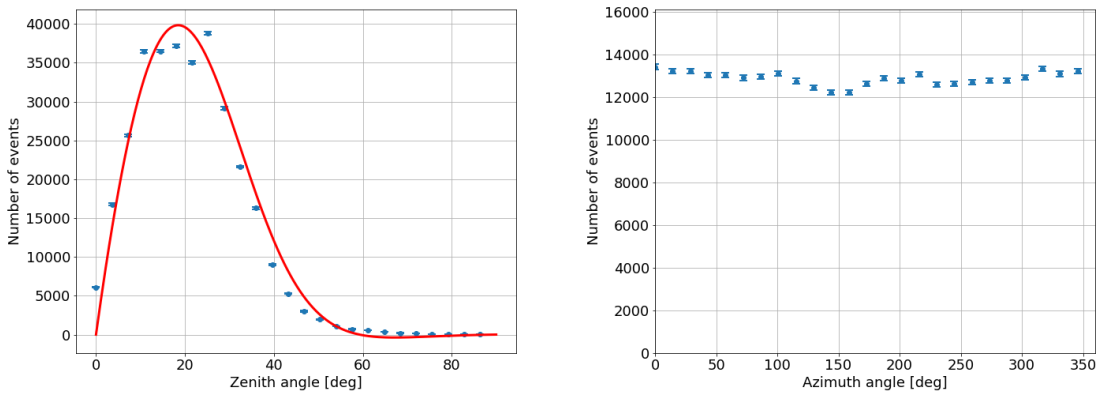


Figure 4.5 *– Zenith (left) and azimuth (right) angular distributions of internal events detected by the scintillators.*

4.2.2. Standalone antenna array

Historically, the radio part of CODALEMA used cabled antennas triggered by the particle detector array. It had the advantage to allow an unambiguous identification of radio transients as the counterpart of the shower particles and to establish the proof-of-principle of the radio detection method. To go further, the extension of the antenna array coverage was mandatory as already discussed. However, over large distances, signals are attenuated and dispersed in the cables, and it becomes impossible to cover very large area with this type of set up. The idea was to build an array of autonomous detection stations able to individually detect the radio signal from EAS. For their conception, some criteria have been defined:

— The criterion concerning the antenna have been defined in section 3.4.1.2 page 67;
— Each functionality (as triggering, coding, timing) has to be ensured by a dedicated electronics element.
— The station crate has to be electromagnetically shielded.
— Being installed in a protected radio environment, the data transfer must be done

through an optical fiber network.
— Commercial technologies and standard protocols have to be used.
— The station must have a low consumption.
— The station has to be remotely settable and accessible.

4.2.2.1. The Butterfly antenna

Contrarily to what the CODALEMA acronym might suggest, the spiral logarithmic antennas used for the first setup of CODALEMA were replaced by standalone antennas working in the [20 – 200] MHz range. They are 1.2 m fat crossed dipoles also called Butterfly antennas (see figure 4.6). They can be assimilated as transducers which converts the electric field component of an incoming electromagnetic wave to a voltage. They are composed of an antenna (radiator) connected to a low noise amplifier (LNA), situated in the central nut, shown in figure 4.6 on the right [170].

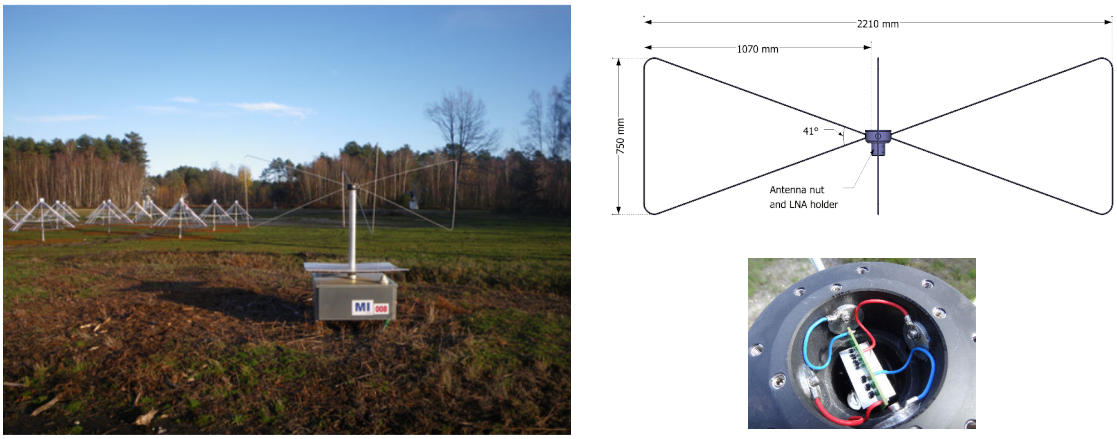


Figure 4.6– *Left: photography of a standalone antenna at Nançay (with an islet of NenuFAR in the background). Right: top, schematics view of the Butterfly crossed dipole antenna with its central holding nut. Bottom: zoom on the central nut, housing the LNA board connected to the dipoles.*

Since the electromagnetic input field induced by EAS is a transient, their design has been chosen to optimize the bandwidth (need to be ultra wide band), to be isotropic (since the cosmic rays arrive from all sky directions), to be sensitive (in order to detect the weakest electromagnetic transients). In other words, the antenna has been optimized to fill the criterion detailed in section 3.4.1.2. Simulations of the gain pattern made with NEC4 are presented in figure 4.7, for different zenith angles and frequencies. For a given azimuth angle ϕ (here $\phi = 0\,°$ to the East and $\phi = 90\,°$ to the North), secondary lobes appear with increasing frequency (left). Until 150 MHz, the secondary lobes are very small, and the gain of the antenna is quite the same for all the zenith angles. However, from 150 MHz and above, the secondary lobes are bigger, and the dips in the gain pattern might affect the measurements in these directions. Similarly, simulations of the gain pattern made with NEC4 are presented in figure 4.8, for different azimuth angles and frequencies. For a given zenith angle θ (here $\theta = 0\,°$ to the top and $\theta = 90\,°$ to the horizon), secondary lobes appear with increasing frequency. Let us now check this observation on an actual event. Figure 4.9 shows the power spectrum density (PSD) for the involved antennas for the EW polarization (left), and shows also the gain pattern of the antennas (right) for the corresponding arrival direction of the event ($\phi = 124\,°$ and $\theta = 51\,°$). On the PSD, we observe a dip around 140 MHz, probably due to the gain pattern of our antennas for this arrival direction, that is confirmed by the simulation of the gain (right figure), which presents a dip in the gain around $\theta = 50\,°$ (blue curve). The same result is observed on the NS polarization. It is therefore essential to know the response of this antenna to

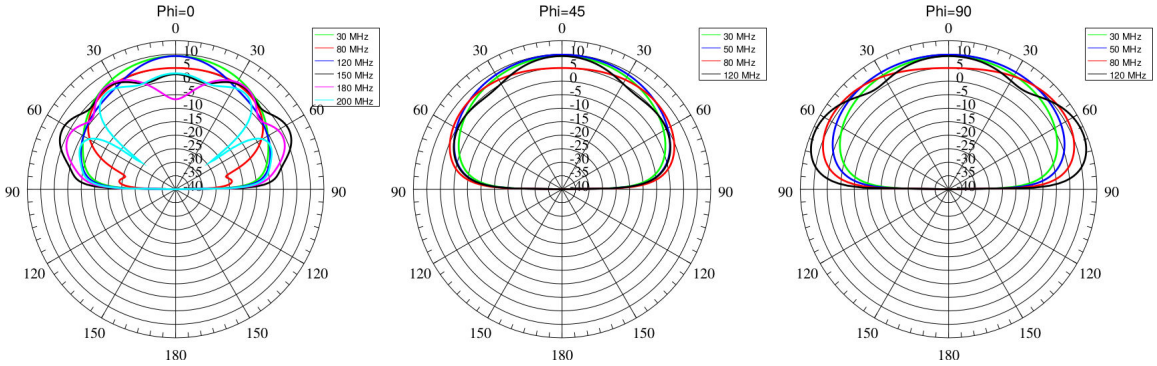


Figure 4.7 *– Simulation of gain pattern of CODALEMA antennas as a function of θ and frequency. Upper left: $\phi = 0\,°$. Upper right: $\phi = 45\,°$. Bottom: $\phi = 90\,°$. The radial axis is the gain of the antenna in* dB ·

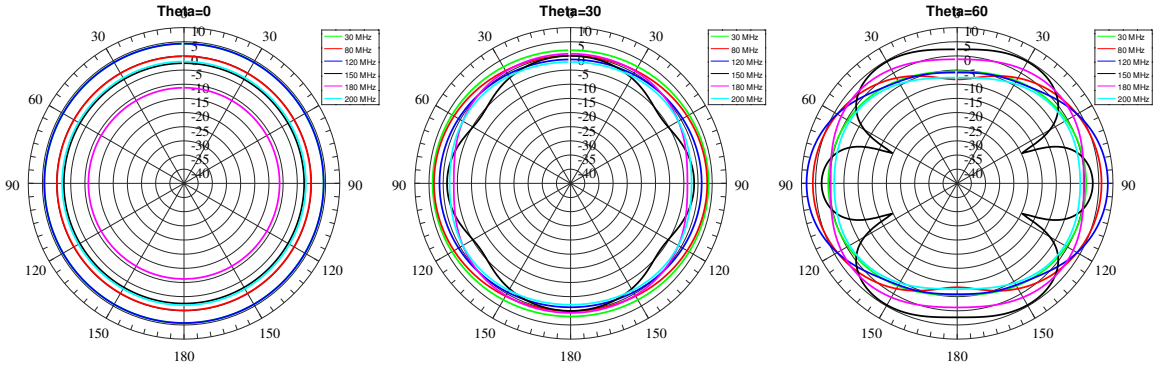


Figure 4.8 *– Simulation of gain pattern of CODALEMA antennas as a function of ϕ and frequency. Upper left: $\theta = 0\,°$. Upper right: $\theta = 30\,°$. Bottom: $\theta = 60\,°$. The radial axis is the gain of the antenna in* dB ·

correct and convolve / deconvolve the data correctly.

As explained in chapter 3, the radio signal emitted by EAS is created by the geomagnetic emission mechanism giving a signal polarized in the $\overrightarrow{v} \times \overrightarrow{B}$ direction and by the Askaryan mechanism giving a signal radially polarized with respect to the shower axis. For this reason, the expected signal has components in both East-West and North-South polarizations. However, AM and FM transmitters are contained in the operating frequency range, and the transient pulse is mixed with the noise, as shown in figure 4.10. It appears that by filtering in the MF range, which is also the band used by other radio experiment, the transient pulse clearly appears.

4.2.2.2. Electronic components of the standalone antennas

Figure 4.11 presents the electronic cards present in the station crate shown in figure 4.6 below the Butterfly antenna. The Butterfly antenna is placed on a pole at the center of the structure. The station crate is separated in two halves, one containing the electrical protections and a linear power transformer, the other one containing the electronic cards. Electrical, optical and analog connections between the two halves are made through a patch panel. Particular attention has been put on the electromagnetic insulation of the second half which contains the electronic cards. Most of the communications between the electronic cards are made through a backplane connection, and some of them are

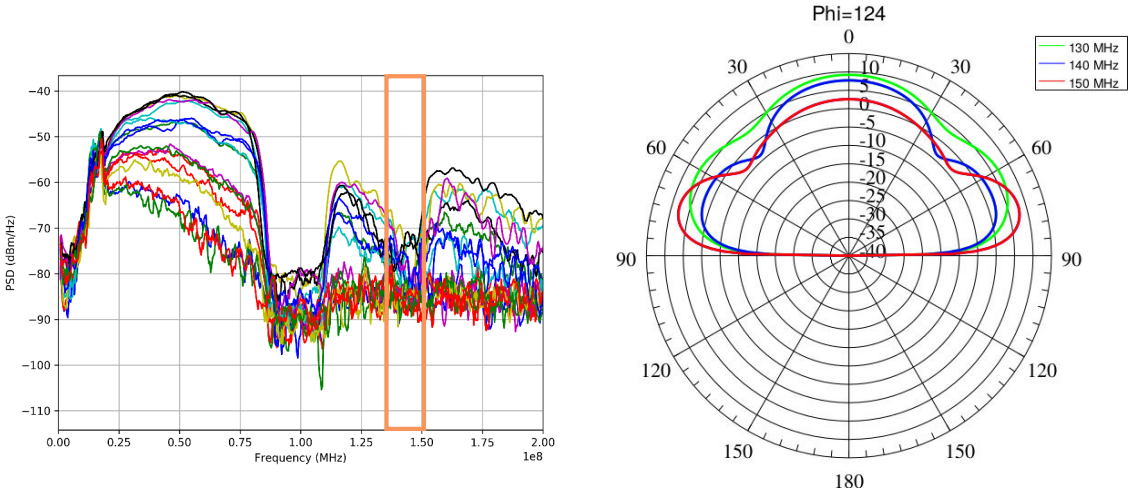


Figure 4.9 *– Simulation of the gain pattern of the CODALEMA antennas as a function of* θ *and frequency. Left: PSD of the event in the EW polarization. Right: simulation gain for* $\phi = 124\,°$*. A dip around* 140 MHz *is observed on the PSD (orange box, figure on the right), due to the gain pattern of our antennas for this arrival direction which also presents a dip in the gain around* $\theta = 50\,°$*. See text for details.*

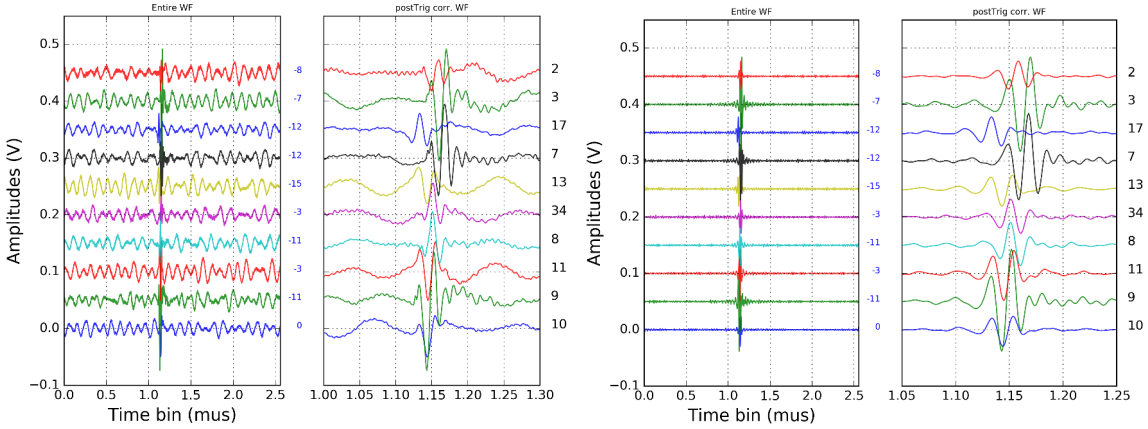


Figure 4.10 *– Left: Example of signals from 10 SA recorded in EW polarizations in the EMF band and corresponding to a cosmic ray event. Right: The same event filtered in* [30 – 80] MHz*. For each plot, the left canvas features the entire waveform and the right one shows a zoom around the transient pulse.*

made in front end by Ethernet or coaxial cables.

The electronic crate is composed as follows:

— The digitizer boards: the digitization is based on the MATACQ chip, the same used in the ADC of the particle detector acquisition, assimilated to a running circular analog memory read by an ADC. The idea is to continuously record the signal into the analog memory matrix. This memory matrix is frozen when a trigger signal is received, and the memory cells are digitized. The MATACQ board contains the MATACQ chip while the MATCOD board serves as controller for the MATACQ board and contains the ADC. More details can be found in [169]. It features two input signals (EW and NS polarizations) sampled at 1 GS $\cdot$ s^{-1} over 2,560 memory cells, corresponding to a 2.56 µs-record. This makes it possible to extend observations up to the current higher limit (200 MHz) of the Butterfly antenna. The signal

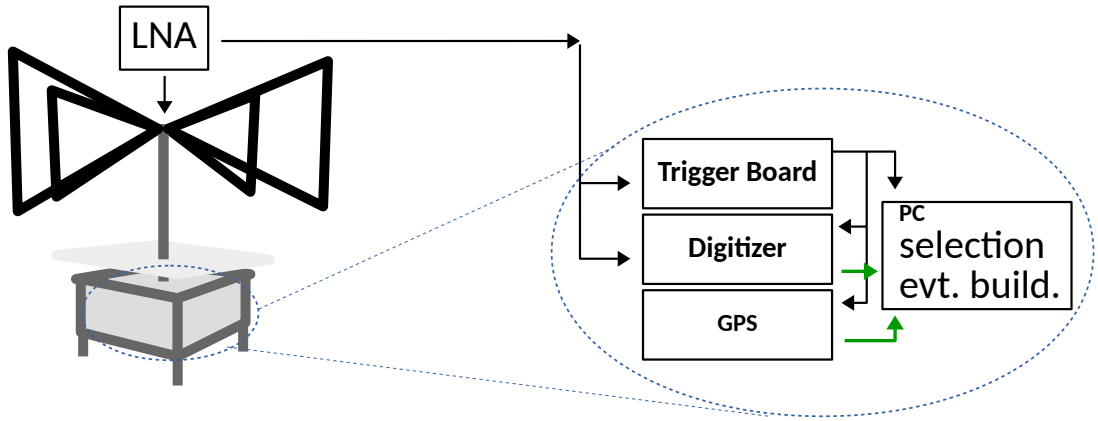


Figure 4.11 *– Sketch of the station crate of the Butterfly antenna.*

dynamic (±1 V) is coded on 14 bits leading to a least significant bit (LSB) value of 122 µV.

— The trigger board: It is composed of two antennas input channels and one external trigger input channel. An internal software trigger (called 1W trigger type) can also be activated for calibration purposes for instance (see chapter 5). Both antenna input channels, corresponding to the EW and NS polarizations, are split into two equal power parts. One part is filtered in two frequency bands [35 – 80] MHz and [120 – 200] MHz (by the so-called *ABCD* filter, whose the frequency response is shown in figure 4.12) and the voltage is compared to an amplitude threshold which is remotely adjustable depending on the noise condition of the previous events. If one or both signals exceed the defined threshold, the other part of the signals sent to the ADC board is digitized at a sampling rate of 1 GHz. In that case, this level 1 signal (noted T1) is also sent to the GPS board and to the onboard PC.
— The GPS board: it is equipped with a Motorola M12T especially designed by the *Laboratoire d'Astronomie de l'Observatoire de Besançon*. Its time resolution is about 1 ns, the differential precision between different GPS card is around 10 ns.
— The communication board: all the modules are connected to a dedicated board for communication/connection purposes.
— The power board provides voltage to the digitizer, trigger GPS and communication boards (5 ± 7 V) and to the PC (12 V). The communication board implements a Rabbit microcontroller.
— The PC: it is used to run the local acquisition, to perform the first level of data analysis and to manage outer communications. Communications with the other electronic boards are made exclusively via the Rabbit board.

4.2.2.3. Operating modes of the standalone antennas

There are two operating modes of the standalone antennas for which dedicated local acquisition programs have been developed. The first one is an autonomous mode based on different levels of triggers is implemented in the trigger board and in the onboard PC. The second one is a central data acquisition mode (based on the CDAQ software developed for AERA) has been implemented in a central computer communicating with all the stations. Practically, a third mode is used which combines trigger selections from the local acquisition and selections from the central acquisition. Figure 4.13 illustrated the triggering strategy and data saving scheme.

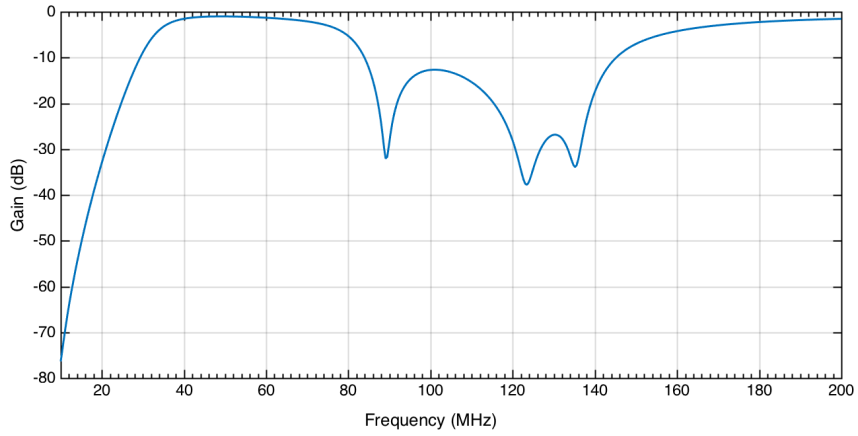


Figure 4.12 *– Frequency response of the ABCD filter.*

4.2.2.3.1. Autonomous mode

Assuming a low rate of cosmic-ray induced signals, no specific efforts had been put to design fast trigger and ADC boards. It takes almost 30 ms to select and record an event in a station. For the observation of a continuous event (meaning to trigger every 30 ms), it would make a marginal duty cycle of 0.0085 %. The actual observation time of the sky in these conditions of operation is 7.42 s (deadtime of 30 ms plus 2.56 µs of record) on one day. These conditions of operation correspond to a very noisy environment, causing problem of saturation of the local data acquisition. For the observation of cosmic rays, the recorded data in these conditions are unusable.

However, for the observation of relatively rare high energy cosmic ray event, 30 ms is a short duration and would lead to a duty cycle of 100 % for cosmic-ray induced transients. In the best conditions of operation, meaning in a radio quiet environment, the antennas will be mostly available to trigger on a potential signal linked to a cosmic ray.

Unfortunately, in Nançay, we are in intermediate conditions, and at the T1 level, the level of background noise can be very high (especially during the day) causing potential problem of saturation of the local data acquisition. Thus, a saturated station is always busy, and may miss cosmic ray events. A data analysis have been carried out since 2011, leading to a performing noise rejection algorithm described in chapter 6. Each SA operates autonomously, independently of the other ones. The recorded events of each SA are stored in a database, such as for the other instruments, and the coincidences between the different instruments are made off-line, as explained in section 6.5 page 151.

4.2.2.3.2. Central DAQ

CODALEMA uses the CDAQ software of AERA which was developed largely by the CODALEMA group involved in AERA and thought to be compatible with all the electronics used on both AERA and CODALEMA. Contrarily to the first mode presented above, the CDAQ software allows a sophisticated online data processing including combination of information from several stations, leading to a third level of trigger T3. The central DAQ of CODALEMA works as follows: each standalone antenna sends T2 times (at the nanosecond level) to the central DAQ process named `postmaster`, which forwards them to another process, the `t3maker`. The latter defines the third level of trigger. It builds time and space coincidences between the different standalone antennas. It is worth noticing that a T3 event does not systematically corresponds to a cosmic ray event, due to the detection of parasitic transients such as airplanes, lightning…Finally, once a T3 trigger is built, the `postmaster` requests the events, and sends them to the event builder `aevb`

which builds a radio coincidence and saves data to the disk (see figure 4.13).

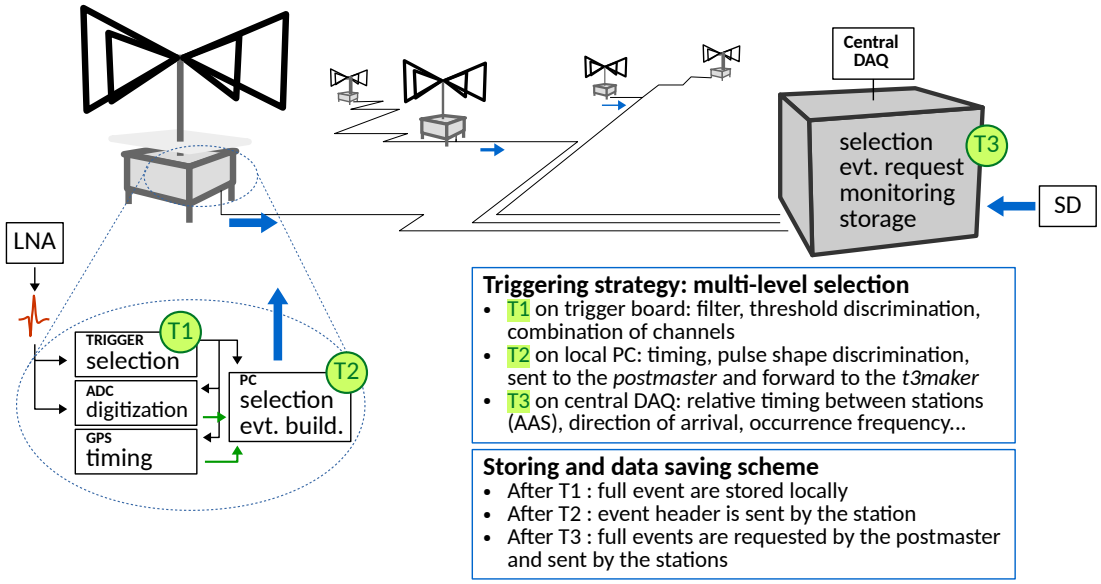


Figure 4.13 – *Schematics of the data taking at CODALEMA.*

4.2.3. The Compact Array

The Compact Array (CA) is a cluster of 10 antennas in dual linear horizontal polarizations, operating since 2013. It has a ground extension of 150×150 m^2, and is situated in the middle of the particle detector of CODALEMA. The CA antennas are active antennas, inverted V-shape, fat dipole (see figure 4.14). The CA presents two main advantages: a short distances between antennas allowing to study the variability of the detected electric field at short scale and a triggering selectivity and sensitivity.

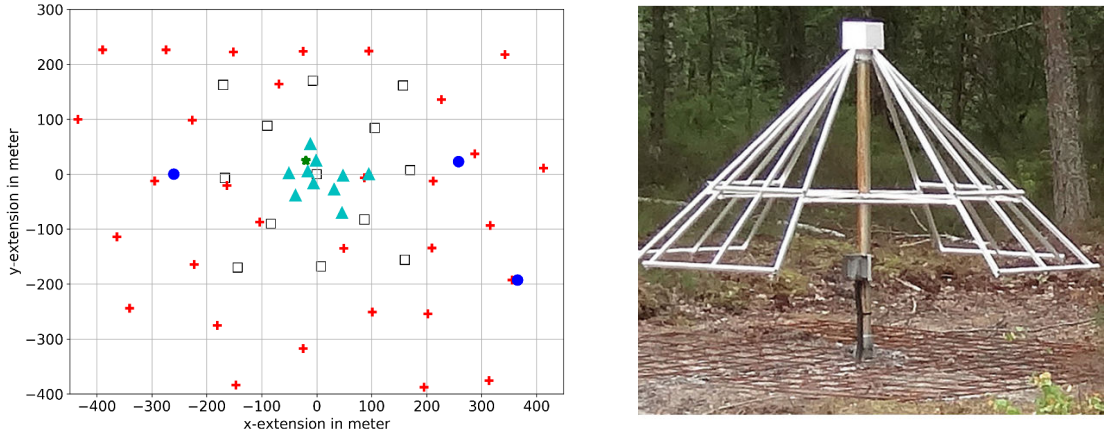


Figure 4.14 – *Left: CA ground extension represented by blue triangles, in the middle of the particle detectors of CODALEMA represented by the black empty squares, red crosses represent the CODALEMA SA. The green star stands for the position of the three-fold antenna (see section 4.2.4). Right: photography of one antenna of the CA.*

The CA is externally triggered by the scintillator array, working in the EMF band at

a sampling rate of 400 MS · s^{-1}. The recorded events are dated by the dedicated GPS station MC250. A threshold detection method is used to find transient in the recorded events, and the direction of arrival is reconstructed with the method explained in the next chapter in section 6.1. The event is selected as an actual shower if the reconstructed direction of arrival of the CA and the particle detector array agree within 20°. Being fully contained and externally triggered by the particle detector array, the CA is used as an indicator of the presence of the radio counterpart on some of the events, and its data enrich the events recorded with the standalone antenna array. The CA is also an offline test bench for the development of phased detection algorithms for online use.

The energy threshold for the radio detection of cosmic rays is clearly visible below 10^{17} eV as shown in figure 4.15. The energy distributions of the scintillators events and of the ARR events converge at the highest energy. This indicates that the relative efficiency of the radio detection technique, defined as the ratio of the number of events detected with the radio technique over the number of events detected by the scintillators, increases for the highest energy. Figure 4.15 right presents the efficiency of the radio detection technique as a function of the estimation of the energy carried out by the scintillators. It regularly increases above 10^{16} eV to reach approximately 50 % at 1.6×10^{17} eV. Beyond this energy, the statistics is low. Looking closely, above $\log(E) = 17.5$, we notice that some reconstructions are incorrect in the ARR, since the analysis of the data of this detector has not been optimized. Several methods are under investigation, such as the cross-correlation method (use of the notion of spatial coherence of a source of electromagnetic radiation) [171] or the phasing method. These techniques can improve the detection of shower signals from cosmic rays of future detectors such as GRAND [140, 141], or for the detection of weak signals produced by gamma rays [172], for example. With these types of improvement in the analysis, the efficiency should be probably greater than 80 % above 10^{18} eV.

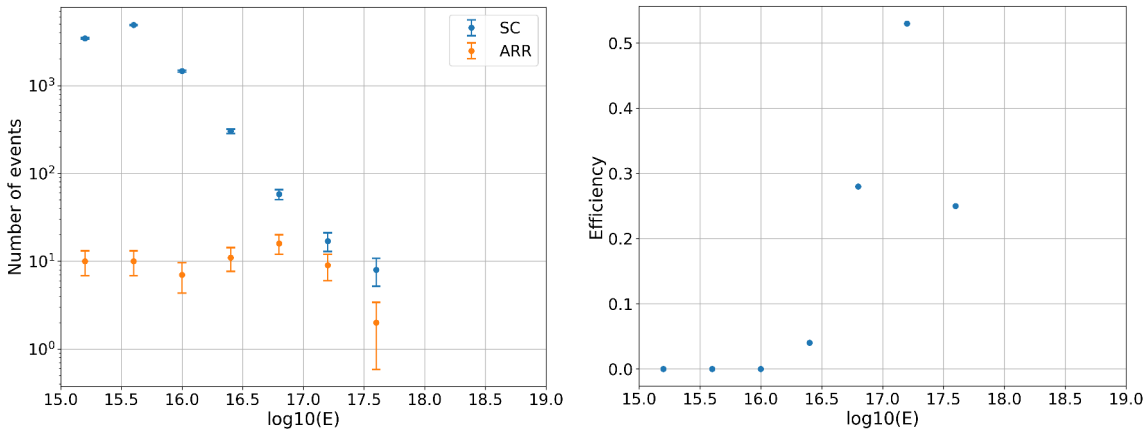


Figure 4.15 *– Left: energy distribution of internal events measured by the scintillators in coincidence with the ARR. Right: relative efficiency of the radio detection technique as a function of the estimation of the energy carried out by the scintillators.*

4.2.4. The three-fold antenna

A three-fold antenna has also been installed within the particle detector array to measure the incoming electric field in coincidence with the particle detector array [160]. This antenna allows to measure the three components of the electric field: East-West, North-South and Vertical polarizations. By detecting the three components of the electric

field, we aim at:

— obtaining the full 3D electric field components directly, without reconstructing the vertical polarization with the two horizontal ones ;
— consequently, testing and confirming the validity of the far-field hypothesis for the study of the electric field emitted by cosmic ray air showers.

The three-fold antenna is composed of three radiators, the same as for the Butterfly antenna of CODALEMA (see figure 4.16). The orientation of the three radiators is made by two rotations of 45° and 54.75° around respectively the X (EW) and Y (NS) axes. The rotation matrices from EW, NS, VE to the three-fold coordinate system can be written as:

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix} \tag{4.2}$$

$$R_y = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix} \tag{4.3}$$

and permits to recover the three-fold antenna polarizations in the same frame than the one for the CA.

The three-fold antenna is externally triggered by the particle detector. The signals of the three polarizations are recorded and GPS dated with MC250, with the same type of electronics as for the scintillators (2.56 µs record length, 1 GS · s^{-1}, 12-bit ADC). This ensures that the three-fold antenna signals are strongly correlated to particle EAS, though some accidental detection are possible, but extremely rare in the considered time window.

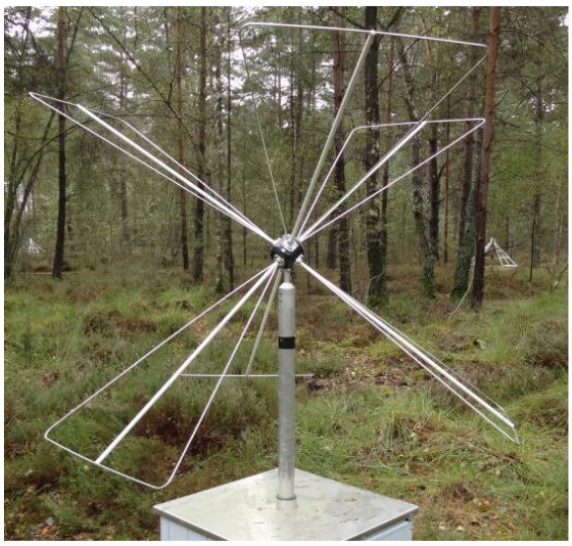


Figure 4.16 *– Photography of the three-fold antenna. Each Butterfly antenna is rotated twice to form a direct trihedron inclined by* 45°. *See [160] for more details.*

Once the three-fold antenna signals are rotated and recovered in the EW, NS et Vertical frame, the idea is to compare the polarization patterns between the three-fold antenna signals and the recovered electric field components obtained after deconvolving the CA antennas of the same event [160] (see section 3.4.1.4 page 68 for the deconvolving method). Figure 4.17 presents an illustrative event recorded by the three-fold antenna.

Figure 4.18 presents the polarization patterns of the three-fold antenna (top) and of the nearest CA antenna after deconvolving. The prediction of the polarization pattern from a pure geomagnetic event are represented by the black lines. From the direction of arrival of the event (ϕ = 120 ° and θ = 49 °) and the geomagnetic vector angle direction for the experiment, we expect a polarization pattern in agreement with the predicted

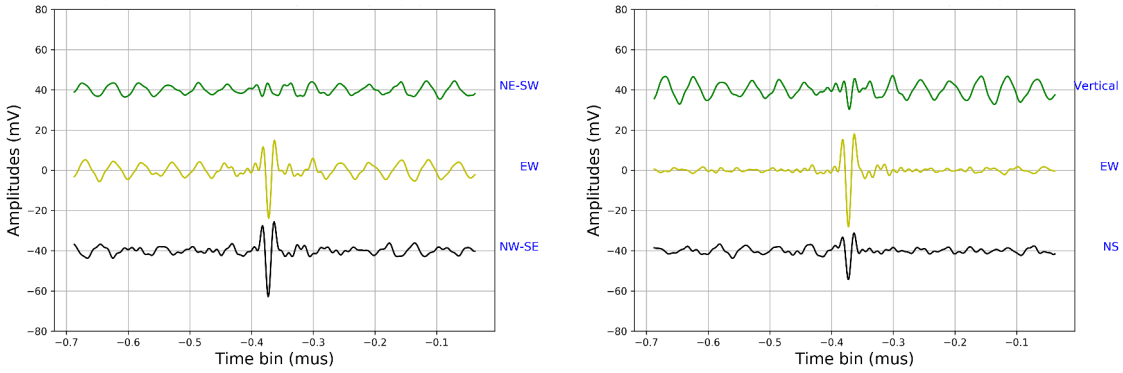


Figure 4.17 *– An illustrative event recorded by the three-fold antenna. Left: event in the three-fold antenna frame. Right: event in the CA frame after rotation. Signals are filtered in* [20 – 80] MHz.

polarization pattern, which is indeed the case for the three-fold antenna and the CA antenna. For this event, one note that the polarization patterns of the three-fold antenna and the CA antenna after deconvolving are similar, which tends to confirm the validity of the far-field hypothesis.

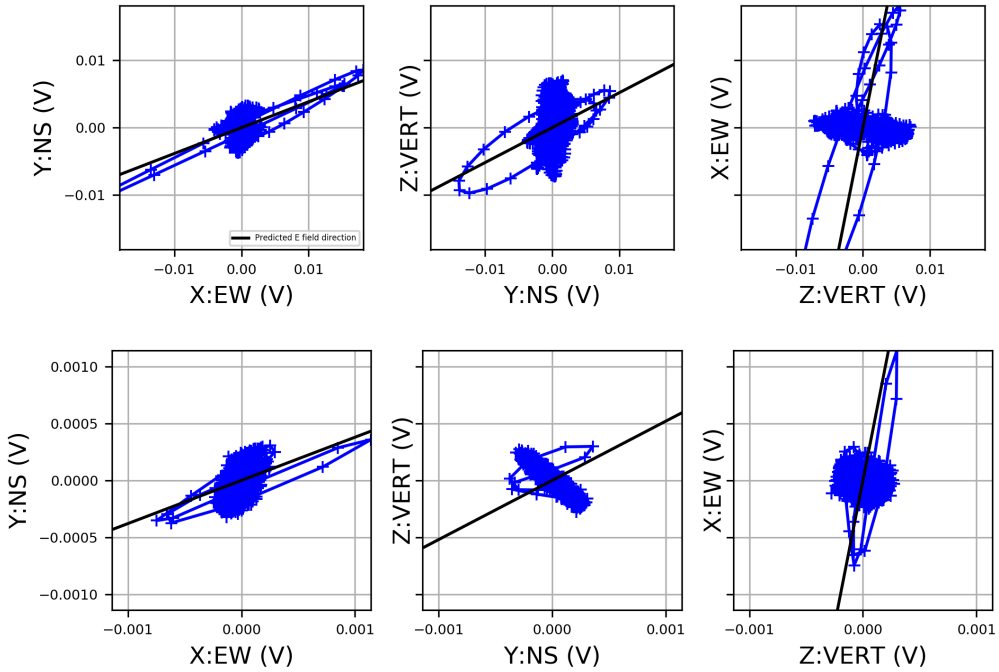


Figure 4.18 *– Top: polarization pattern for the three-fold antenna for the event of figure 4.17. Bottom: polarization pattern for the closest CA antenna after deconvolving, for the same event. The polarization ellipses are represented in blue and the predictions are represented by the black lines. Signals are filtered in* [20 – 80] MHz. *See section 3.2 page 50 for the construction of the polarization pattern.*

The same procedure has been applied on a large set of events recorded in 2017, selected as follows:

— Self-triggered event for the SA array with a minimal multiplicity of 4 within ±5 µs from a SD trigger.

— A minimal multiplicity of 7 for the CA array, with a detected signal in the CA antenna closest to the three-fold antenna.

— A value of atmospheric electric field at the moment of the event corresponding to normal conditions ($|E_{\text{atm}}| \leq 500\,\text{V}\cdot\text{m}^{-1}$).
— Finally, a detected signal in at least two polarizations of the three-fold antenna.

One has to note that for this set of events, there is no cut on the direction of arrivals. Thus, the pattern of some events may not match the geomagnetic prediction. The obtained results are shown in figure 4.19 for the comparison between the NS/EW polarization angles of the closest CA antenna and the three-fold antenna. A clear correlation is obtained between the closest CA antenna and the three-fold antenna.

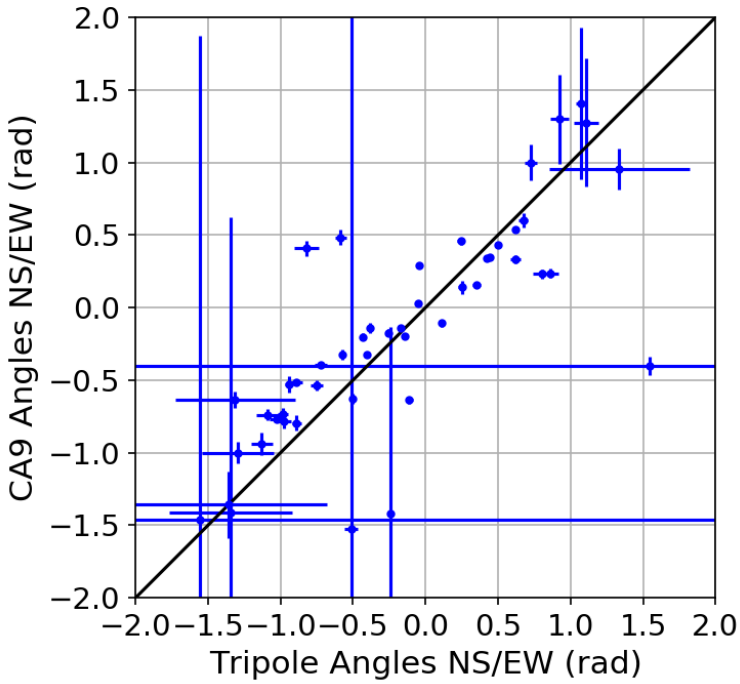


Figure 4.19 *– The NS/EW polarization angles: closest CA antenna angles vs the three-fold antenna angle. The black line stands for the correlation line.*

Figure 4.20 presents the comparison to the predicted geomagnetic polarisation angles in the 3 directions, where the blue points represent the angles of the closest CA antenna and the red points represent the angles of the three-fold antenna. A clear correlation is obtained with the predicted geomagnetic polarisation angles for the NS/EW polarization (figure 4.20 left). As for the NS/EW polarization, the correlation is clear for the EW/Vertical polarization angles (figure 4.20 right).

Concerning the Vertical/NS polarization, there is no correlation found. By construction, we expect a Vertical/NS polarization angle close to the geomagnetic zenith angle of the experiment site, i.e $\theta_{\text{B}} = \sim 27\,°$. This problem is under investigation.

Despite the result obtained for the Vertical/NS polarization, this short study tends to confirm the validity of the far-field hypothesis.

4.3. EXTASIS

Parts of this section has been published in: . . ., **A. Escudie et al.**, *Radio detection of cosmic rays below* 1.7 − 3.7 MHz*: the EXTASIS experiment*, Astroparticle Physics [9].

As we have seen in previous chapters, the coherent radio emission during the development of air shower has two main origins: transverse current induced by the geomagnetic field, and the radial current induced by the charge excess mechanism [81, 82, 83, 84, 85].

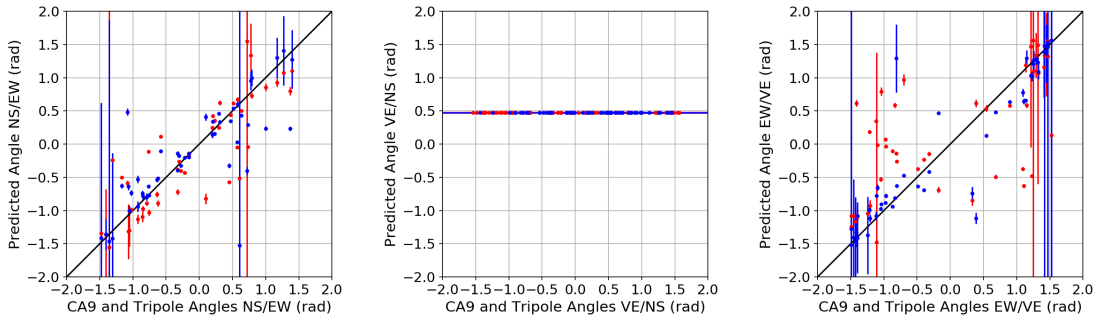


Figure 4.20 *– Comparison to the predicted geomagnetic polarisation angles in the 3 directions for the set of events. Blue points represent the angles of the closest CA antenna and red points represent the angles of the three-fold antenna. The black line stands for the correlation line.*

The resulting emission appears as a fast electric field transient lasting few tens of nanoseconds, which can be detected by large bandwidth antennas and fast acquisition systems. In the most common way, the observations are carried out in the MF band ([20 – 80] MHz) by experiments such as AERA [121], Tunka-Rex [122], TREND [123], Yakutsk experiment [124] or LOFAR [125]. The use of this band is mainly due to man-made broadcasting at low and medium frequencies (AM, FM bands) leading to the choice of relative low sampling rates ($\sim$ 200 MS $\cdot$ s^{-1}) of the digitizers used by experiments such as AERA and LOFAR. However, CODALEMA [173] works with a sampling rate of 1 GS $\cdot$ s^{-1}, making it possible to extend observations above the FM band where ARIANNA [174], ANITA [151] and CROME [175] were or are operating. The main limitation of the frequency band is then due to the bandwidth of the antenna used, which is optimized and well mastered in [20 – 200] MHz for CODALEMA, referred to as Extended Medium Frequencies (EMF) in the following.

Several detections at low frequencies (hereafter LF, below 20 MHz) have been carried out in the 70's and up to the 90's (a complete review can be found in appendix D page 237). A main conclusion can be drawn from these observations (partially summarized in table 4.1): the results are not well understood. Indeed, several experiments [176, 177, 178, 179] have reported that when the frequency decreases, a strong increase of the radio pulse amplitude is observed. For example, Prescott et. al [177] reported transients of about 300 µV $\cdot$ m^{-1} $\cdot$ MHz^{-1} at 3.6 MHz, averaged over 400 showers whose energy was not known. Nevertheless, the fact that most of those measurements of large electric field have not been reproduced casts doubt on the plausability of such large values [178]. Moreover, some estimations of shower energy seem to be incredibly low to permit a radio detection: for example, Stubbs [178] reported an energy of a detected primary cosmic ray of supposedly 2×10^{14} eV, and a transient amplitude of about 1 µV $\cdot$ m^{-1} $\cdot$ MHz^{-1} at 2 MHz averaged over 100 showers. However, most of the modern experiments have an energy threshold for detecting a radio contribution from a cosmic-ray-induced air showers around 10^{16} eV in the MF band. If one considers that the LF signal is, like the MF one, proportional to the number of secondary particles, this would mean that the LF signal has been produced by about at least a hundred times less particles than the smallest detectable MF signal. This would probably imply that either an unknown kind of enhancement of the LF signal acts to make it detectable above the radio noise at these frequencies (see section 8.1), or another emission process exists (which is discussed further in this article), or more likely, a wrong estimate of the shower energy was made at that time. Assuming erroneous shower energies, the rescaling made in [178] to get a normalization for a shower of an energy equal to 10^{17} eV gives huge and probably overestimated electric field amplitudes of about 500 µV $\cdot$ m^{-1} $\cdot$ MHz^{-1}. As shown in the current paper, such values would easily be detectable on a reasonable timescale regarding the expected

Ref	Frequency	Observations
[176]	3.6 MHz	Calgary (1049 m asl), signal at 3.6 MHz 3 times larger than at 22 MHz, no signal larger than noise at 10 MHz.
[177]	2 MHz	Haverah Park: signal at 2 MHz 200 times larger than at 32 MHz and 375 times larger than at 44 MHz
[178]	2 MHz	Buckland Park: 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$, estimated energy of 2×10^{14} eV, giving after rescaling 500 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ at 10^{17} eV. From that, signal at 2 MHz would be 250 times larger than at 44 MHz
[179]	3.6 MHz	Dominion Radio Astro Observatory: signal at 3.6 MHz 1 order of magnitude larger than in [20 – 60] MHz, but less by a factor of 3-4 than at 2 MHz
[180]	[0.026 – 0.3] MHz	Akeno, AGASA: unipolar and negative signal, with a width of 5 µs, field decreasing with decreasing distance
[181]	2.6 MHz	EASRADIO: increase in electric field strength when frequency decreases. Observation of low-frequency signal in coincidence with extensive air shower.

Table 4.1 *– Summary of some pioneer observations in the LF band: references, observations frequency (*MHz*), comments.*

number of showers at 10^{17} eV falling on any — even small — detection arrays. This is also in contradiction with the fact that no other LF experiment was able to detect them at that time. Despite these puzzling issues, Akeno experiment [180] found correlation of signals in the LF band with signals at higher frequencies. They reported signal amplitudes of about 100 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ in the range [26 – 300] kHz, a measurement that has been successfully repeated and which is consistent with other measurements carried out by EAS-RADIO [181].

From these observations, and in spite of the controversy, we have therefore chosen to retake the LF study, taking advantage of an improved version of SELFAS3 simulation code of the radio signal of cosmic-ray-induced air showers [108, 98], and by installing LF antennas in the CODALEMA experiment [173, 159] (see figure 4.21). A sketch of the instrumental setup of LF antenna is shown in figure 4.25.

4.3.1. The LF antenna

Currently, the EXTASIS experiment is made of dedicated LF antennas (figure 4.1 and figure 4.21) triggered by scintillators. They are based on the so-called “Butterfly” active antennas in use in CODALEMA (see the concept of these active antennas in section 4.2.2), with the same radiating element and same “LONAMOS” Low Noise Amplifier (LNA), except than the setting of the LNA is tuned for the frequency band below 10 MHz. The shape of the radiating element is a bow tie made of an aluminium rod with an overall length from end to end of 1.2 m. Apart the LNA, another difference with the CODALEMA antennas is that their dual, crossed polarizations are East-West and Vertical, since the SDP is expected to be mainly vertically polarized [98]. Regarding the LF band, this antenna is a short dipole, since its length is less than one tenth of the shortest wavelength. In

Figure 4.21 *– Photography of a LF antenna.*

that case, our antenna impedance is well described by a pure capacitance C_{ant} of 12 pF, value estimated using the complex impedance produced by NEC-4 simulations. As the antenna is active, the LNA is located at the antenna feedpoint, as illustrated on the left part of figure 4.22. Thanks to a low capacitive input impedance of the LNA (C_{lna} = 6.6 pF), the signal received by the antenna is transferred to the LNA input through a capacitive divider giving broadband characteristics. On the sketch of figure 4.22, the LNA noise density is modelized only by a voltage source v_{lna}^2, which is a good approximation as the input transistor of the LONAMOS is a CMOS one. Considering the atmospheric noise as a signal, the signal to noise ratio (SNR) of the active antenna can thus be written as:

$$\frac{v_{atm}^2}{v_{noise}^2} = \frac{T_{atm}\, R_{rad}}{T_0\, R_{loss} + \left(\frac{C_{ant}+C_{lna}}{C_{ant}}\right)^2 \frac{v_{lna}^2}{4k}} \tag{4.4}$$

T_{atm} is the minimum atmospheric day temperature of figure 8.1 of chapter 8, T_0 = 290 K the air temperature as previously defined, R_{rad} the antenna radiation resistance and R_{loss} the loss resistance due to the ground below the antenna. From equation 4.4, the LNA intrinsic noise increases independently of the frequency by a constant factor of 2.4 depending only on the capacitive divider ratio. That would be impossible to obtain with a

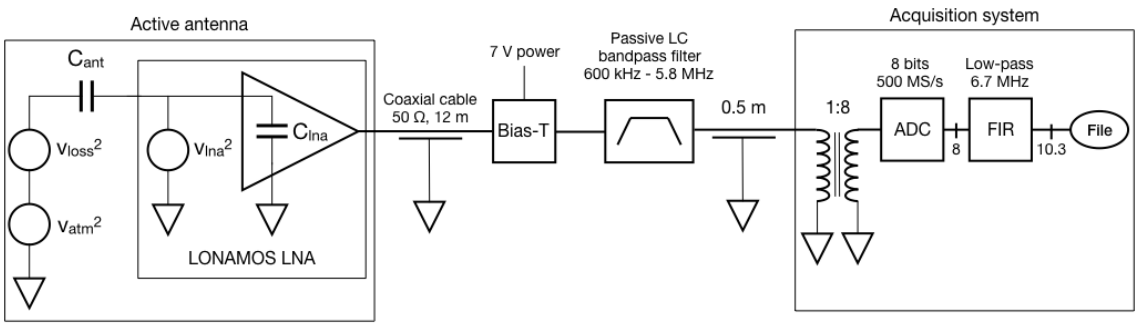


Figure 4.22 *– Simplified sketch of the active antenna ant its noise sources, the RF components, the ADC and signal processing.* $v^2_{atm} = 4k\,T_{atm}\,R_{rad}$ *is the equivalent noise source of the overall atmospheric noise seen in* 2π *sr by the antenna.* $v^2_{loss} = 4k\,T_0\,R_{loss}$ *is the equivalent noise source of the ground losses seen by the antenna.* v^2_{lna} *is the equivalent noise source of the noise of the LNA located at the feedpoint of the antenna.* v^2_{atm}, v^2_{loss} *and* v^2_{lna} *are noise densities in* $\mathrm{V}^2 \cdot \mathrm{Hz}^{-1}$ *unit. Triangles pointing down feature grounding. See text for details on FIR and acquisition.*

50 or 75 Ω input impedance LNA, unless using an antenna near its first resonance, which would imply a huge and heavy antenna with an end to end length around 60 m for the [1-6] MHz band. But a drawback of this short dipole is the low value of R_{rad}, around 1 Ω, as the antenna is used at a frequency 18 times lower than its 45 MHz resonance frequency. Hopefully, this low value of R_{rad} is compensated by the large value of T_{atm}, keeping the product $T_{\mathrm{atm}}\,R_{\mathrm{rad}}$ of equation 4.4 high enough against v^2_{noise}, making it possible to use such a short dipole from the SNR point of view.

As the longest wavelength is around 200 m, the lossy ground is in the near field of the antenna, implying losses than can not be neglected. The finite ground conductivity σ_{ground} and ground relative permittivity $\epsilon_{\mathrm{ground}}$ imply ground ohmic losses, represented by the loss resistance R_{loss} of equation 4.4 at ambient temperature T_0. In this paper, simulations of antenna gain and antenna impedance are performed with the NEC4 engine using the moment method with far field conditions. σ_{ground} is set to $5\ \mathrm{mS} \cdot \mathrm{m}^{-1}$ and $\epsilon_{\mathrm{ground}}$ is set to 13, which are typical values of "average" ground. From equation 4.4, even for a noiseless LNA with $v_{\mathrm{lna}} = 0$, the intrinsic SNR of the antenna is not infinite and depends on the ground losses. To guarantee a SNR of at least 10 dB, $T_0\,R_{\mathrm{loss}}$ should be kept ten times lower than $T_{\mathrm{atm}}\,R_{\mathrm{rad}}$. In order to lower ground losses, one could imagine to place the antenna at a 2 m height above a metallic mesh laid on the ground, but to be efficient this mesh would have to be huge, with a surface in the range of one wavelength squared ($\sim 9 \times 10^4\ \mathrm{m}^2$). Another solution consists in moving away the antenna from the lossy ground. Then, as the antenna height increases, R_{loss} decreases and the signal to noise ratio is increasing, as illustrated in figure 4.23. For this simulation where the LNA noise is set to zero, our criterium consists in choosing a minimum height of the antenna giving a minimum signal to noise ratio of 10 dB. It is fullfilled for a height of 9 m in the worst case of the antenna in the horizontal polarization and at the lowest (2 MHz) observing frequency. Thanks to the choice of short 1.2 m length dipole, antenna weight is minimized, easing the possibility to place it on a 9 m mast by minimizing the mechanical constraints.

Figure 4.24 presents the 9 m high antenna total gain patterns as a function of the zenith angle and azimuth angles for different frequencies and for both the horizontal and vertical antenna. For these polarizations, and at a fixed azimuth angle, the gain pattern is maximal for a zenith angle corresponding to a vertical direction of arrival for the horizontal polarization and for a zenith angle of 65° for the vertical polarization, and decreases as the zenith angle increases. The horizontal antenna is optimal to detect the LF counterpart of the radio signal emitted during the shower development. Inversely, the

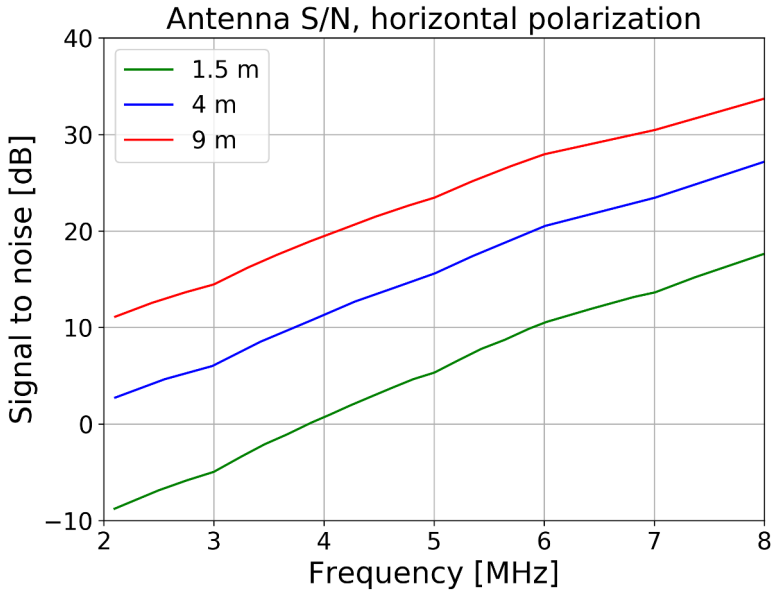


Figure 4.23 *– Signal to noise ratio of a horizontal Butterfly antenna with a noiseless LNA parameterized by its height above a lossy ground with* $\epsilon_{ground} = 13$ *and* $\sigma_{ground} = 5\,\mathrm{mS} \cdot \mathrm{m}^{-1}$.

vertical antenna is optimal to detect the sudden death signal coming from the ground level, thus the gain pattern is maximal for large zenith angles [1]. At a zenith angle of 41°, the gain pattern is homogeneous over the whole azimuth angle range, with a difference of about 3 dB between the gain of both antennas at 2.5 MHz. Figure 4.23 shows that lower heights give lower SNR values: therefore, if the antenna were placed at lower heights, the antenna gain pattern should be corrected by the same offset values. For instance, for an horizontal antenna, the zenith gain of 2.5 dB at 9 m would become -6.5 dB at 4 m, and -19 dB at 1.5 m.

4.3.2. From the antenna to the ADC

At frequencies below 1.7 MHz the power spectrum is highly dominated by a strong local AM 162 kHz transmitter and medium waves [526.5-1606.5] kHz AM transmitters. To allow the LNA to work in its linear zone, far enough from its compression point in daytime conditions, front end high pass and notch filters (not shown in figure 4.22) are added at the LNA inputs. More exactly than previously mentioned, the input impedance of the LNA is equivalent to a capacitance (C_{lna}) in parallel to a resistance, thus defining a first order high pass filtering. The setting of the LONAMOS are performed so as the resistance is 10 kΩ, giving a 900 kHz cut off frequency. A passive second order LC high pass filter is also placed in front of the LNA in addition to a LC 162 kHz notch filter. These three filters give a total attenuation of 53 dB at 162 kHz. As shown in figure 4.22, the output signal of the LNA is transmitted by a 12 m coaxial cable to a RF analog chain followed by an 8 bits commercial digital oscilloscope controlled by a dedicated acquisition software hosted by a local PC. The analog chain is composed of a bias-T, allowing to power the LNA via the signal cable, followed by a band pass filter with 600 kHz and 5.8 MHz cut off frequencies in order to attenuate the strong local 162 kHz transmitter and medium wave transmitters.

1. Here, the antenna gain simulation is performed by taking into account the ground and the far-field hypothesis. This implies that the NEC software considers that the source-to-antenna distance is always infinitely greater than the antenna-to-ground distance, and nothing is simulated beyond 90°. However, in the case of EXTASIS, the ground is in the near field of the antenna. A gain simulation taking into account the ground and the near field condition is therefore more accurate, but it requires much longer calculation times and it is difficult to implement.

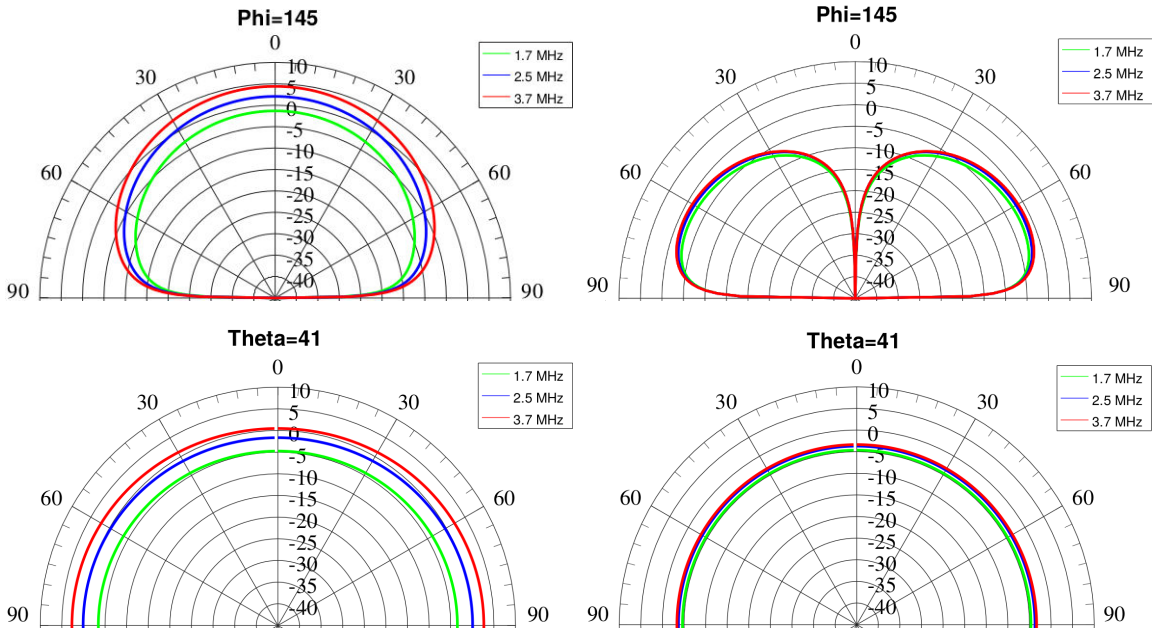


Figure 4.24 *– Simulation of gain pattern of EXTASIS antennas at 9 m height, as a function of* Θ *(top figures),* Φ *(bottom figures) for different frequencies. Top left:* $\Phi = 145\,°$ *for the horizontal antenna; bottom left:* $\Theta = 41\,°$ *for the horizontal antenna; top right:* $\Phi = 145\,°$ *for the vertical antenna; bottom right:* $\Theta = 41\,°$ *for the vertical antenna. Frequencies: green* 1.7 MHz*, blue* 2.5 MHz *and red* 3.7 MHz*.* $\Theta = 0\,°$ *corresponds to zenith,* $\Phi = 0\,°$ *to the East and* $\Phi = 90\,°$ *to the North.*

This chain is ended by an impedance transformer with a 1:8 impedance ratio stuck to the ADC input connector. It performs a 50 Ω matching (input reflexion coefficient lower than -18 dB) to the high impedance of the ADC input in a [230 kHz - 13 MHz] bandwidth, and adds a 9.7 dB voltage gain. Despite the RF signal is bandwidth limited to less than 6 MHz, the signal is oversampled [182] to 500 MS $\cdot$ s^{-1} in order to obtain 14 dB additional dynamic on the ADC thanks to a digital Finite Impulse Response (FIR) filtering applied by the acquisition software. Consequently, the 8 bits, 2 ns sampling digitizer is equivalent to a 10.3 bits digitizer but with a 10 MHz limited bandwidth. The FIR filter is a 201 coefficients gate weighted by a cardinal sine function, in order to obtain an extremely flat magnitude response up to 6 MHz and a minimum rejection of -40 dB from 10 MHz to beyond. During the FIR filtering, the DAQ software also performs a decimation by a factor of 25 resulting in a final record with a time resolution of 50 ns. Thus, lighter files are stored without any loss of information in view of the Nyquist condition. Oversampling factor, decimation factor and FIR filter coefficients can be changed in the DAQ software.

4.3.3. Trigger signal and acquisition

When at least any 5 out of the 13 scintillators of the CODALEMA particle detector are triggered within a given time window, called a “particle coincidence window”, a master trigger is built and sent to a nearby dedicated GPS station which dates the event at the ns precision, and to the EXTASIS LF antennas. Due to the extent of the array (several hundreds of meters), the trigger for the LF antennas is distributed over an Ethernet network, which takes on average 750 ± 250 µs to reach the antenna. The trigger signal received at each LF antenna is also dated allowing to correct for the network time jitter. From that, we know exactly the propagation time of the trigger signal, which is necessary to know where to seek the signal in the 2 ms record. As a reminder, the sought LF signal lasts few µs , which is approximately at least two order of magnitude less than the recorded

trace duration. The raw digitizer sampling time bin is 2 ns, the recorded trace contains 10^6 time bins, corresponding to 2 ms signal length, long enough to find the particle detector trigger time in the trace by subtracting the trigger emission time from the trigger reception time. Taking into account the extent of the LF antenna array around the particle detector and the propagation speed of the signal (approximated to the speed of light), if any LF radio transient has been recorded in coincidence with the air shower that has triggered the particle detector, it should be located within a 3 µs-wide time window around the reconstructed particle trigger time. The LF antennas have a reception crate, which contains as electronics boards:

- A Raspberry Pi (noted Rb in figure 4.25) computer which receives the distributed trigger and sends it to the conversion card.
- A conversion card (noted C) which converts the trigger signal from TTL to LVDS format, sent to the trigger card.
- A trigger card (noted T), which receives in input the trigger in LVDS format, and sends it again to the conversion card (noted C) which converts it from LVDS to TTL format.
- A digitizer (noted picoscope) which receives the trigger signal from the conversion card, and freezes the two polarizations.
- A local computer (noted PC) which registers the data and sends them to a file on a hard drive ;
- A GPS card (noted GPS) used to stamp the date on the registered event.

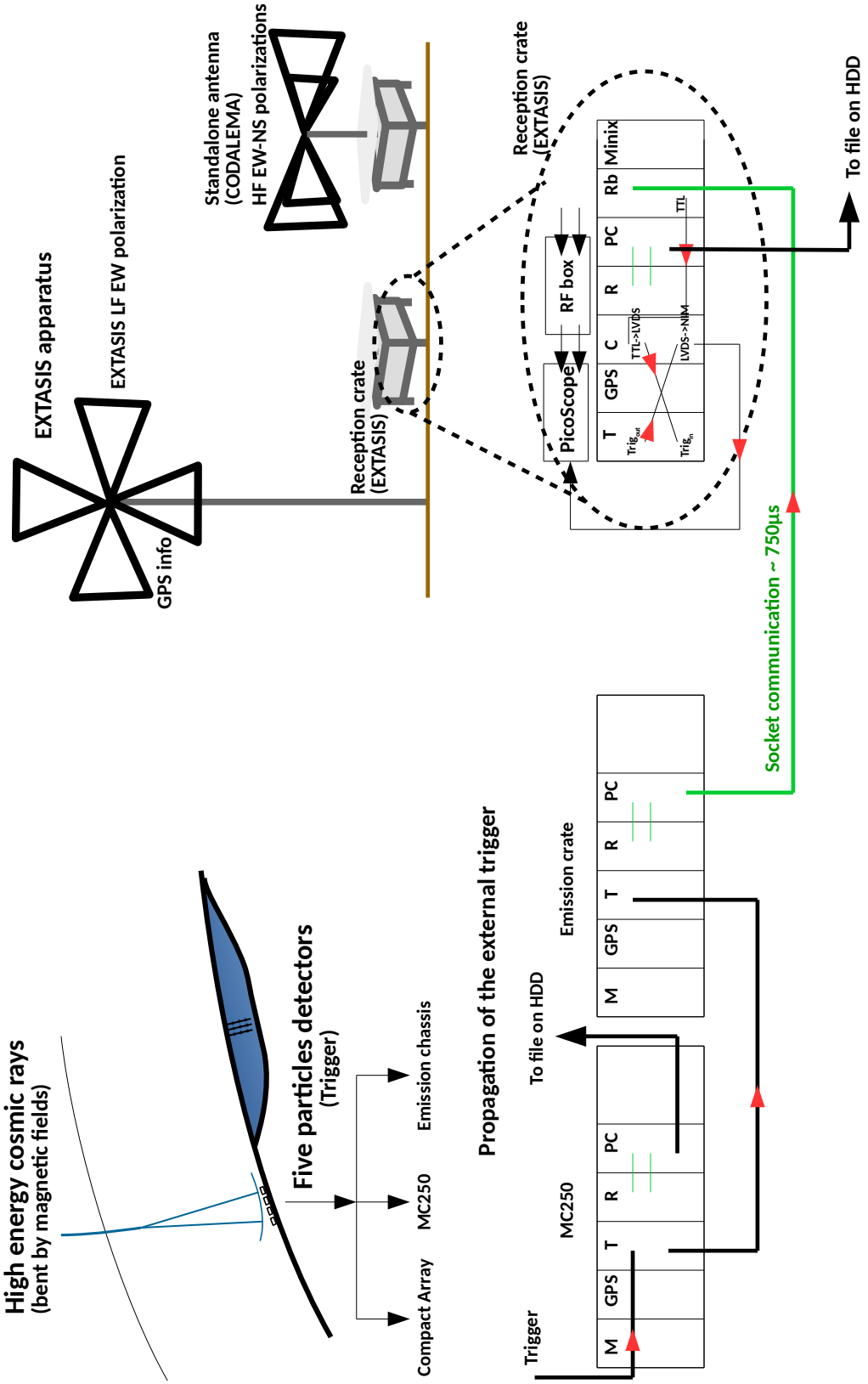


Figure 4.25 *– Sketch of the instrumental setup of EXTASIS. From upper left: when at least any 5 out of the 13 scintillators of the CODALEMA particle detector are triggered within a given time window, a master trigger is built and sent to a nearby dedicated GPS station which dates the event at the ns precision (called MC250), and to the EXTASIS LF antennas. When the trigger arrives in the emission crate, it is sent by a socket communication to the reception crate of each LF antennas. After some electronic conversions, the signal is sent to the digitizer to record the signals. See text for more details on each electronic module composing the reception crate.*

" If you can't explain it simply, you don't understand it well enough."

Albert Einstein

CHAPTER 5

Calibration of CODALEMA using the Galactic radio-emission

Outline

Parts of the results of this chapter have been presented in TeVPA2018 and ICRC2019 conferences, and published in [146].

One of the consequences of the study of ultra-high energy cosmic rays, namely to study them by indirect detection, is to have to go back to the characteristics of the primary cosmic rays, namely the energy and the nature. In the past, numerous disagreements have been noted on the amplitudes of the electric fields detected for a given energy [183], notably due to the difficulty of calibrating the radio detectors then used. In order to be as precise as possible on the measurement of the amplitudes, and thus on the estimation of the energy of the primary cosmic ray (see chapter 9), we propose in this chapter a method of calibration of our radio detectors. Indeed, we aim to cross calibrate our radio instruments (the SA and CA antennas, see 4.2.2 and 4.2.3 respectively) using the Galactic radio emission [184]. Three models of Galactic radio emission are studied and compared. Then, we present the results of the cross calibration of our radio antennas.

5.1. Method

In addition to the threshold trigger, the autonomous stations are equipped with a software trigger (called 1W trigger type) which is produced at regular intervals (typically every 5 minutes in regular functioning) over the full autonomous station array. 1W data contain mostly background noise, which is a combination of sky and system noise, radio emitters and natural phenomenon such as thunderstorms or magnetic solar disturbances (MSD). When the external disturbances are absent, it remains a background noise due to the Galactic radio emission, which can be used as a reference to cross calibrate our antennas (see figure 5.1).

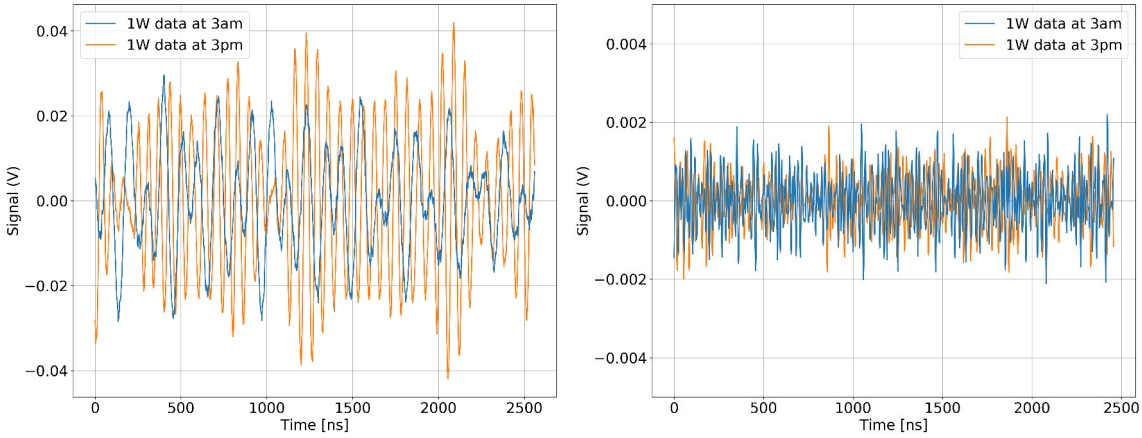


Figure 5.1 *– Left: Example of 1W data in fullband taken the 13rd April 2015. Right: Same as left filtered in* [30 – 80] MHz.

Figure 5.2 presents the 1W data in Local Sidereal Time (LST) [1] (blue points) for one SA. The RMS of one 1W data is the sigma value obtained from the Gaussian distribution [2] of the recorded signal (see figure 5.1 right). Then, for a given LST bin, the ⟨RMS⟩ is the mean value obtained from the Gaussian distribution of the RMS values of the 1W data of the given LST bin (orange points in figure 5.2).

Figure 5.3 presents the ⟨RMS⟩ (noted rms in lowercase in figures to lighten the notations) of the 1W data in LST for all the SA. The transit curves are stackable, meaning that the source of the emission is the same for all the antenna array. The behaviour of the ⟨RMS⟩ of the 1W data is similar for all the antennas over LST, and only an offset along the y-axis is present.

After the cross calibration, this offset along the y-axis is supposed to be reduced to zero, assuming that all the antennas have the same response. The obtained shape of the ⟨RMS⟩ of the 1W data is assumed to be due to the Galactic radio emission. To do that, we first need to estimate this emission. The galactic noise temperature received by the antenna is partly transferred to a low noise amplifier (LNA). This LNA is connected to an analog chain, composed of coaxial cables, amplifier and filters, and terminated by a 50 Ω terminal load, input impedance of an ADC. That is why the RMS voltage value $|V_{\text{acq}}|$ (acq stands for acquisition) of the total noise sampled by the ADC for a given bandwidth results both from the galactic noise temperature T_{ant} collected by the antenna and from additional system noise mainly dominated by the intrinsic noise of the LNA.

1. At a given location, a celestial object observed at one position in the sky and at a given LST will be observed at the same position in the sky on another day at the same LST.

2. The noise rms could also have been calculated from the mathematical standard deviation of the signal. However, in the case where a transient is present in the trace, the Gaussian fit allows a more accurate estimation of the noise rms.

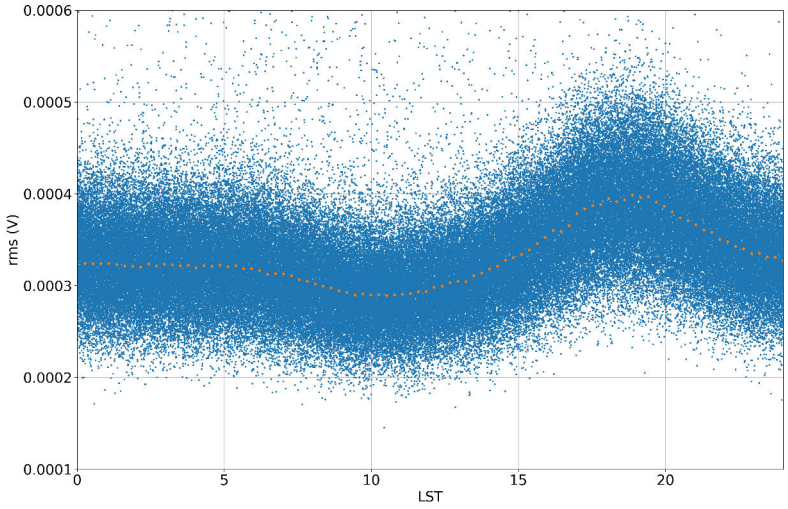


Figure 5.2 *– Example of 1W data in LST for one SA.*

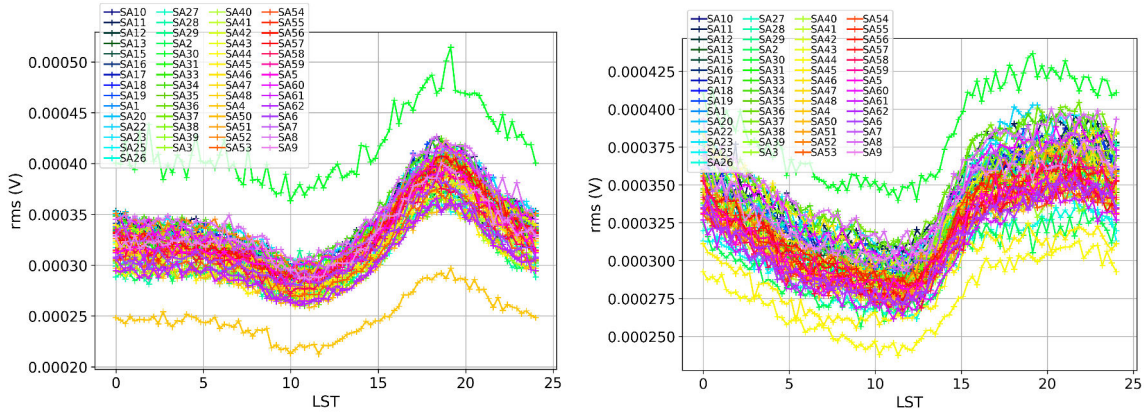


Figure 5.3 *– The* ⟨RMS⟩ *of the 1W data in Local Sidereal Time (LST) for the EW (left) and NS (right) polarizations at* 50 MHz *(sub-bands of* 10 MHz *width) and for an arbitrary month. For each LST bin, the* ⟨RMS⟩ *is calculated from the distribution of the RMS values of the background noise. Outliers in green and yellow are antennas with an electrical problem and come out of the range. However, we will see that this type of behavior is perfectly regulated by the proposed calibration.*

Remark: It is worth precising that the cross calibration of the radio detectors does not need the use of a Galactic radio emission model. We could use the average behavior of antennas as a reference, but the use of an external source (the Galactic radio emission in our case) is preferred. The use of the Galactic radio emission allows not only to cross calibrate the antennas, but also to check and validate the NEC4 modelization of our radio-detectors, which plays a major role in the convolving of simulations by the antenna response, and thus in the systematic comparisons between data and convolved simulations, as discussed in chapter 7.

The simulated RMS voltage value $|V_{\text{sim}}|$ due to the only galactic temperature is also computed. The available noise temperature in K received by an antenna can be expressed from [185] as:

$$T_{\text{ant}} = \frac{1}{\Omega}\int_0^{\pi}\int_0^{2\pi} T_{\text{gal}}(\theta,\phi)\cdot G_{\text{ant}}(\theta,\phi)\mathrm{d}\Omega \tag{5.1}$$

where $T_{gal}(\theta, \phi)$ is the brightness temperature of the Galaxy for each direction of the sky in Kelvin, $G_{ant}(\theta, \phi)$ is the gain of the antenna (dimensionless) and $\Omega = 2\pi$ sr since we perform the integration over an infinite plan ground ($\theta = 0\,°$ corresponds to the zenith, $\phi = 0\,°$ to the East and $\phi = 90\,°$ to the North).

Using the Boltzmann law, T_{ant} can be converted to a power for a given bandwidth B with:

$$P_{ant} = k_B \cdot T_{ant} \cdot B \tag{5.2}$$

with k_B the Bolstzmann's constant equal to 1.38×10^{-23} J $\cdot$ K^{-1}. This available power can be converted to a voltage $|V_{sim}|$ using a 50 Ω reference resistance R_{ref} by :

$$|V_{sim}| = \sqrt{P_{ant} \cdot R_{ref}} = \sqrt{k_B \cdot T_{ant} \cdot B \cdot R_{ref}} \tag{5.3}$$

Finally, for a narrow frequency range, a relative comparison of the $|V_{acq}|$ variations to the $|V_{sim}|$ variations can be performed.

An example of the sky map of the Galactic radio emission in galactic coordinates is shown in figure 5.4 top for a frequency of 50 MHz. The principal contribution comes from the center of the Galaxy and at second order from the galactic plane. Thus, the RMS of the received signal must be maximal when the lobe of the antenna points toward these zones. Inversely, the RMS of the received signal must be minimal when the lobe of the antenna points to the sky directions with a minimal emission, i.e to the Galactic poles (in black in figure 5.4 top).

From now, the objective is to obtain a model for the temperature of the Galactic radio emission. These models are presented in section 5.2. Once we have a model of Galactic radio emission, the next step is to calculate the contribution of each part of the sky by convolving the skymaps obtained for each frequency by the response of our radio instruments. This requires also a model of the latter that we describe in section 5.3.

5.2. Comparison of the Galactic radio emission models

Several experiments have worked on the understanding of the diffuse Galactic radio emission, and produced sky surveys at different frequencies. Historically, the motivations were mainly the study of the cosmic microwave background (CMB) [186, 187] (see section 1.2.4.1 page 22). We present in this section three models which gather these sky surveys: the Global Sky Model (GSM) [188, 189], the Low Frequency Sky Map model (LFmap) [190] and the Low Frequency Sky Model (LFSM) [191]. Their relative differences will be detailed.

5.2.1. Global Sky Model (GSM)

At the time of writing this thesis, two versions of GSM are available, one released in 2008 [188], the second one in 2016 [189]. The first version of GSM is based on measurements at different frequencies for a total of 23 maps, from 10 MHz up to 100 GHz, including 11 maps with a high quality. The 11 maps are combined and used to train spectral templates. These templates are then interpolated to produce a skymap at a desired frequency (Principal Component Analysis). The second version of GSM has been updated with 7 more maps, extending the model up to 5 THz. The training and the interpolation have also been improved. The second version of GSM has an angular resolution twice smaller than the original version over the whole frequency band. However, it seems that some sources such as *Cyg A* are missing in the updated model, excepted in the 45 MHz and the 1.4 GHz maps. Missing sources correspond to yellow/white spots in figure 5.4 bottom. It is worth noticing that even though, the final results, after taking into account

the antenna response, do not seem to be impacted. In other words, point sources are not important when treating with non-directional antennas. Nevertheless, for completeness, we decided to use the 2008 model, containing the sources.

From the documentation of GSM, the uncertainties on the predicted skymaps are estimated between 5 and 15 %.

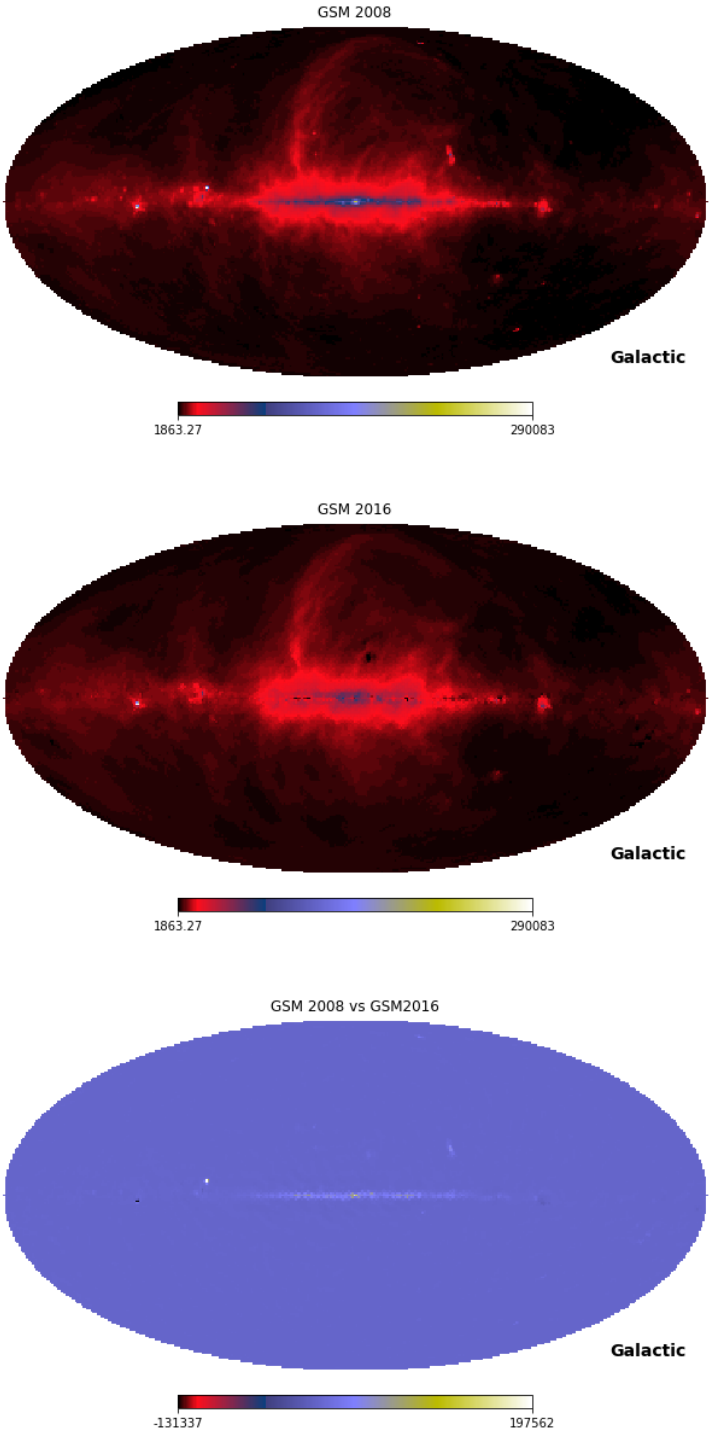


Figure 5.4 *– Top: skymap in galactic coordinates given by GSM 2008 at* 50 MHz *in Kelvin. Center: skymap in galactic coordinates given by GSM 2016 at* 50 MHz *in Kelvin. Bottom: skymap of differences in Kelvin in galactic coordinates between GSM 2008 and GSM 2016. The remaining sources are Cyg A (left hotspot) and Sgr A (center of the map).*

5.2.2. Low Frequency Sky Map (LFmap)

LFmap code was released in 2007 [190]. The goal was to obtain sky maps at any frequency in the Long Wavelength Array (LWA, radio telescope in the center of New Mexico) band ([35 – 80] MHz) [191]. The code is based on a temperature of the sky described by a power law. The brightness temperature at a given frequency ν and a given direction (α, δ) is:

$$T_{\text{sky}}(\nu, \alpha, \delta) = T_{\text{CMB}} + T_{\text{Iso}}(\nu) + T_{\text{Gal}}(\nu, \alpha, \delta) \tag{5.4}$$

where T_{CMB} is the temperature of the CMB (2.73 K), T_{Iso} is the isotropic component (mainly due to unresolved emission from extragalactic sources, estimated to be equal to $\sim$ 1025 K at 50 MHz) and T_{Gal} is the temperature of the Galaxy (order of magnitude: a few hundred thousand of K, dominating contribution in [30 – 200] MHz). The calculation is derived in [190]. Moreover, there is a break in the spectrum at around 200 MHz, and the power law description is not sufficient [190]. To deal with it, the idea is to start from the 408 MHz map given by Haslam in 1982 [192], to use the Platania analysis [193] to rescale the map at 200 MHz (frequency of the break in the spectrum) and then to use a lower frequency map to calculate the new indices. Caveat: the LF map used is incomplete, and some sources are also missing (*Cyg A, Cas A, Vir A* and *Tau A*). LFmap code handles the HII absorption, and corrects for missing discrete sources (only for *Cyg A* and *Cas A*). The differences between the original GSM and LFmap are shown in figure 5.5. LFmap has a higher temperature.

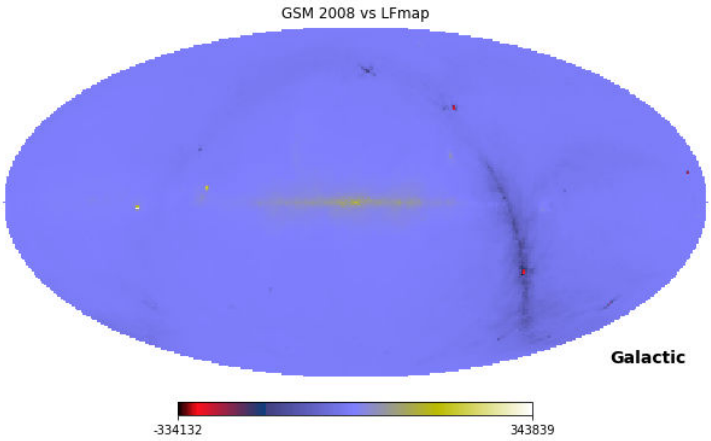


Figure 5.5 *– Skymap of differences in Kelvin in galactic coordinates between GSM 2008 and LFmap.*

5.2.3. Low Frequency Sky Model (LFSM)

LFSM was released in 2017 and has been updated in 2018. The code is based on a sky survey made by LWA1 [191] in [35 – 80] MHz plus other input surveys at higher frequency (400 MHz and 1800 MHz) completed by the Platania analysis. It is available on the LWA Data Archive. The modelling approach is quite similar to GSM. The differences between the original GSM and LFSM are shown in figure 5.6.

5.3. Response of SA and CA antennas

Both galactic (background emission) and antenna response are expected to be largely frequency dependent. The first step is to fix the granularity of the frequency scale. We choose to work in sub-bands of 10 MHz width each 5 MHz step for three reasons:

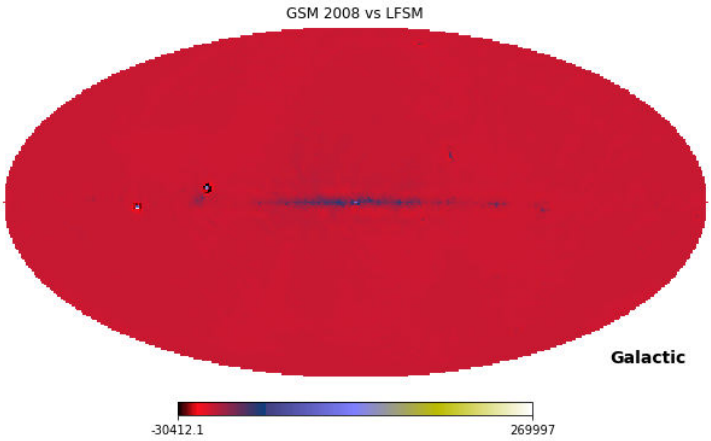


Figure 5.6 *– Skymap of differences in Kelvin in galactic coordinates between GSM 2008 and LFSM.*

— The GSM gives skymaps for a precise frequency, thus narrower the frequency band, the more the data will be consistent with the sky models. However, the frequency resolution associated to a signal of 2.56 µs is equal to ~ 390 kHz. The frequency band must therefore be as narrow as possible, nevertheless remaining wide enough to have good frequency resolution. A sub-band of 10 MHz is a good compromise.
— At high frequency, the contribution of the electronic noise is frequency-dependent, and the calibration need to be carried out in frequency sub-bands. Moreover, at high frequency, secondary lobes of the antenna appear at different frequencies and are also frequency-dependent.
— At last, other experiments such as LOFAR have already done this kind of calibration in sub-bands of 10 MHz [194], showing really good efficiency of this calibration method.

In the following, the SA will used as an example. We will base our argument on two test frequencies at 50 and 150 MHz, but it has been verified that the results obtained are consistent over the full band of interest ([20 – 200] MHz) and that the results are also valid for the CA antennas.

5.3.1. Modelling of the antenna

Firstly, a simple geometrical model was studied for the modelling of the SA. It consists in only taking into account the shape of the antenna in terms of geometry. The gain of the antenna is thus maximal for directions which are orthogonal to the considered polarization, and zero for parallel directions to the considered polarization. Then, the modelling of the SA was also made with the NEC4 engine (hereafter named the complete description of the SA). A comparison with NEC2 has been made and the obtained results are identical, as described in section 3.4.1.4 page 68. Figure 5.7 presents the result obtained for the geometrical description (noted mechanical in the figure) and the complete description of the SA. We immediately see that both approaches give different results. Although both approaches exhibit a variation peaking between 15 and 20 H LST, the predicted curves are significantly different. In the following, we show how it is important to have a complete description of the antenna, containing a description of the ground, the full electronic chain, the antenna gain and the antenna impedance to well fit data.

5.3.2. Combination of the antenna and the Galactic radio emission models

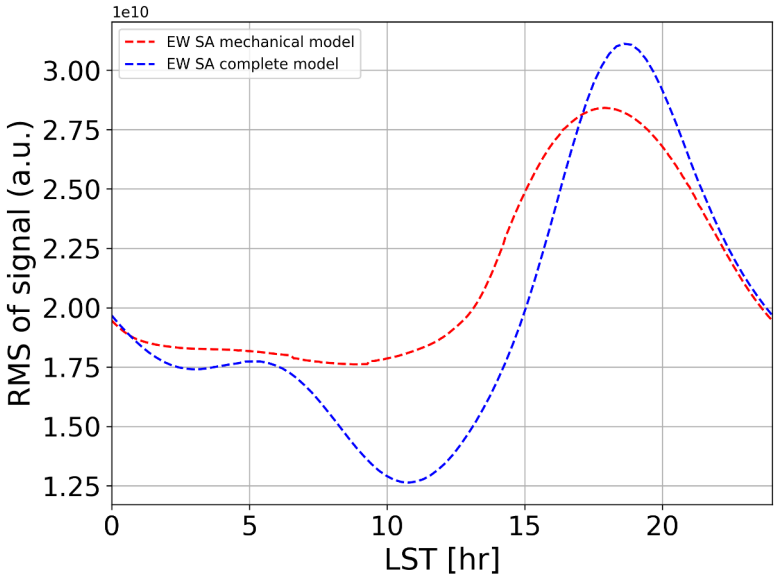


Figure 5.7 *– Prediction of transit curves at* 50 MHz *made with the complete (blue) and geometrical (red) models of the SA.*

5.3.2.1. Choice of the galactic model

To calculate the contribution of each part of the sky, we have to convolve the skymap by the lobe of the antenna as a function of the Local Sidereal Time (LST) (i.e to project the lobe of the antenna as a function of the Local Sidereal Time (LST) on the skymap). Then, we obtain the expected signal at a given LST by summing each pixels of the convoluted skymap (see equations 5.1 and 5.3). Data were binned in LST-bins of 16 minutes. The result obtained is called *transit curve*, that can be compared to the actual data. Figure 5.8 presents the comparison of the transit curves obtained with the different galactic models.

The three sky models presented here are consistent in terms of temperature between each others. Moreover, after the convolving of the skymaps with the antenna response, the three models agree at a 10 % level. We have chosen to work with **the first version of GSM**, because it is the most complete regarding the sources. It is also the reason why the NenuFAR collaboration chose it. As we were at some point working together to understand the galactic codes, it was also logical to use the same maps. The next step is to calculate the contribution of each parts of the sky by convolving the GSM skymaps obtained for each frequency by the response of our radio instruments.

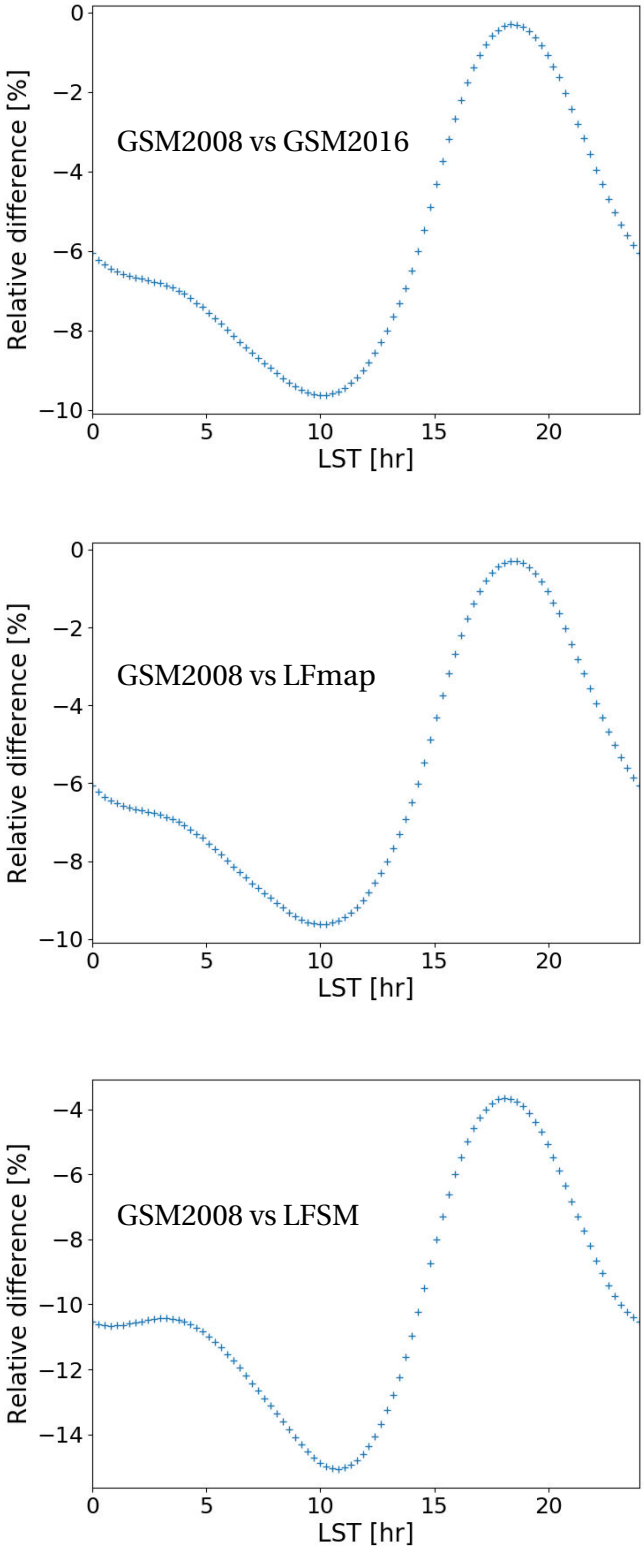


Figure 5.8 *– Comparison of transit curves obtained with the different galactic models at* 50 MHz. *Top: comparison of the predictions of transit curve for GSM2008 and GSM2016. Center: comparison of the predictions of transit curve for GSM2008 and LFmap. Bottom: comparison of the predictions of transit curve for GSM2008 and LFSM. The three models (GSM2008, LFmap and LFSM) agree at a* 10 % *level.*

5.3.2.2. Application to data

In figure 5.9, we present the predictions of the measured Galactic radio emission obtained with the NEC4 antenna model for 50 MHz and 150 MHz. The statistic error $\sigma_{\text{rms}}^{\text{stat}}$ corresponding to a LST-bin is calculated as follows:

$$\sigma_{\text{rms}}^{\text{stat}} = \frac{1}{\sqrt{N}} \sqrt{\frac{1}{N-1} \sum_{i}^{N} (\text{RMS}^i - \langle \text{RMS} \rangle)^2} \tag{5.5}$$

with N the number of RMS values for a given LST-bin. We noticed a good agreement between the prediction made with NEC4 (blue curve) and the data at 50 MHz (figure 5.9). The same conclusion can be drawn from the 150 MHz curves. The result shows here the importance of having a fine description of the antenna to fit properly the data, contrarily to the geometrical model presented above. As a matter of fact, the simple geometrical model is unable to predict the 3 maximums at 15, 18 and 21 H LST due to the particular lobe structures at this frequency.

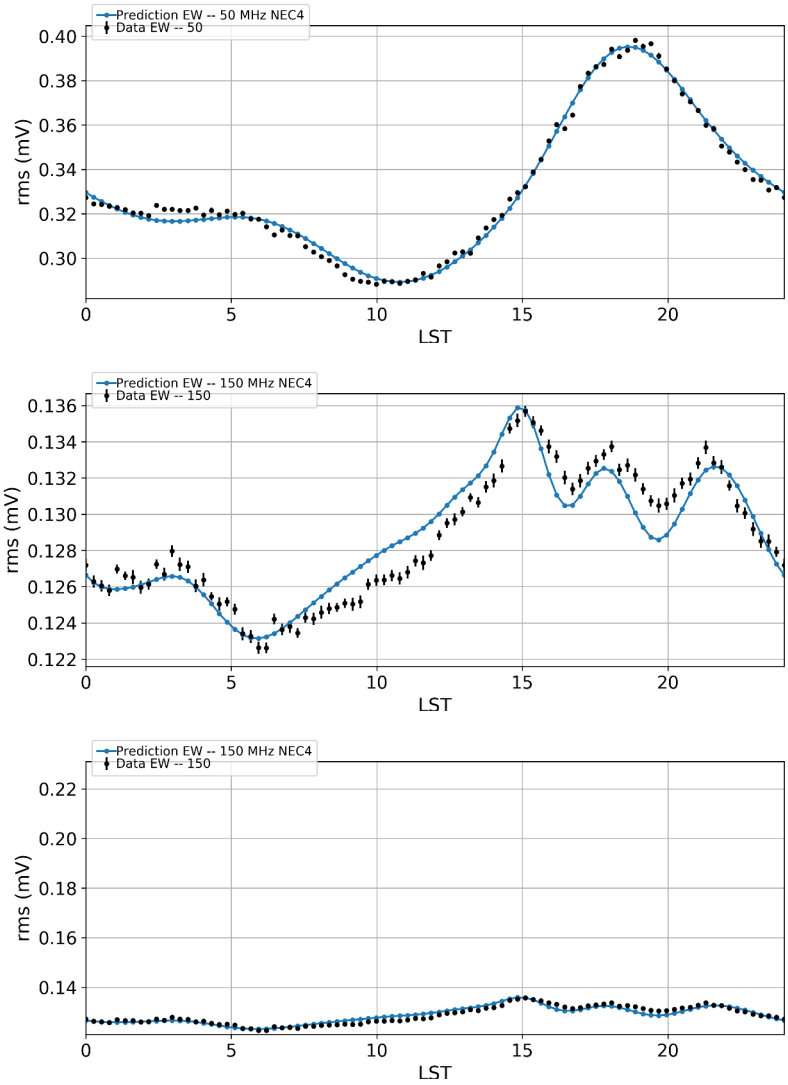


Figure 5.9 *– Top:* 50 MHz*-predictions of transit curve in the EW polarization of one SA fitted on data (black points). Center:* 150 MHz*-predictions of transit curve of the EW polarization fitted on data (black points). Bottom:* 150 MHz*-predictions of transit curve of one SA in the EW polarization fitted on data (black points) with the same y-axis extension that top figure.*

Remark: *figure 5.9 presents the comparison of the predictions of transit curves at* 150 MHz*, with an adjusted scale in y-axis (center), and with an y-axis scale similar to that of* 50 MHz *(bottom). The visible differences in the center figure are finally tiny when rescaling with the same y-axis extension than the transit curves at* 50 MHz.

5.3.3. Influence of polarization of the antenna

Figure 5.10 is the same as figure 5.9 for the NS polarization. As for the EW polarization, we noticed a good agreement between the prediction made with NEC4 (blue curve) and the data at 50 MHz. The same results is obtained for 150 MHz. Similarly, the visible differences in figure bottom are finally tiny when rescaling with the same y-axis extension than the transit curves at 50 MHz.

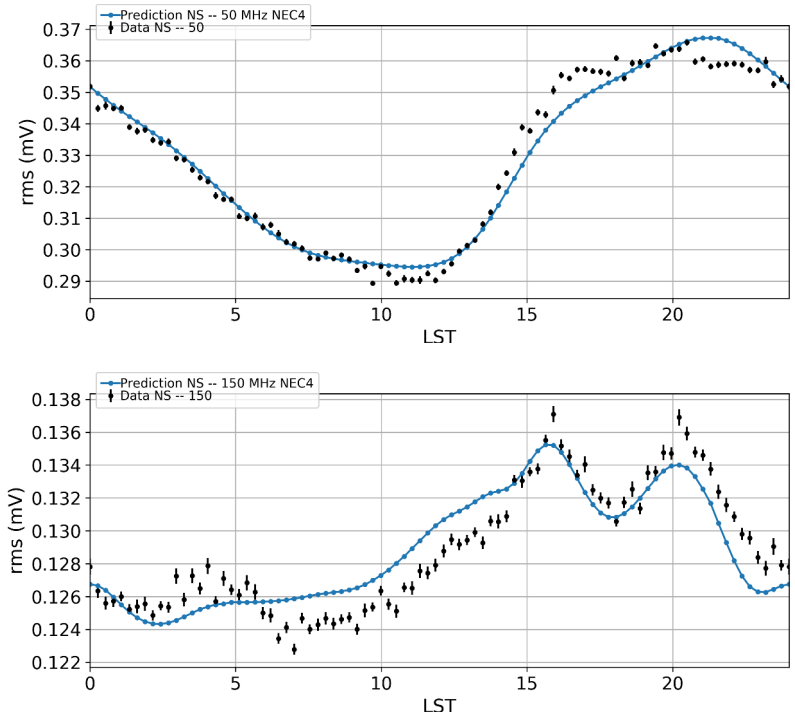


Figure 5.10 *– Left:* 50 MHz*-predictions of transit curve of one SA in the NS polarization fitted on data (black points). Right:* 150 MHz*-predictions of transit curve of one SA in the NS polarization fitted on data (black points).*

The behavior of the predictions fits the data in both polarizations and at both frequencies allowing us to conclude that our antenna model is rather good. The difference between the EW and NS curves (shape, number of peaks) can be explained by the different orientations of the lobes with respect to the Earth rotation axis.

5.3.4. Influence of the shape of the antenna

Figure 5.11 presents the predictions of the measured Galactic radio emission at 50 MHz (top) and 150 MHz (bottom) for the SA and the CA antenna, for the EW (left) and NS (right) polarizations, as a function of LST.

The antenna responses are different between the SA and the CA antenna, as expected since their shapes are different. We also see that at 50 MHz, the transit curves of both SA and CA antenna go through one minimum and one maximum, signing a principal large lobe. At 150 MHz, secondary lobes appear, the transit curves go through several maxima: three for the SA and two for the CA antenna for the EW polarization, and two for the SA and two for the CA antenna for the NS polarization. Concerning the sensitivity of

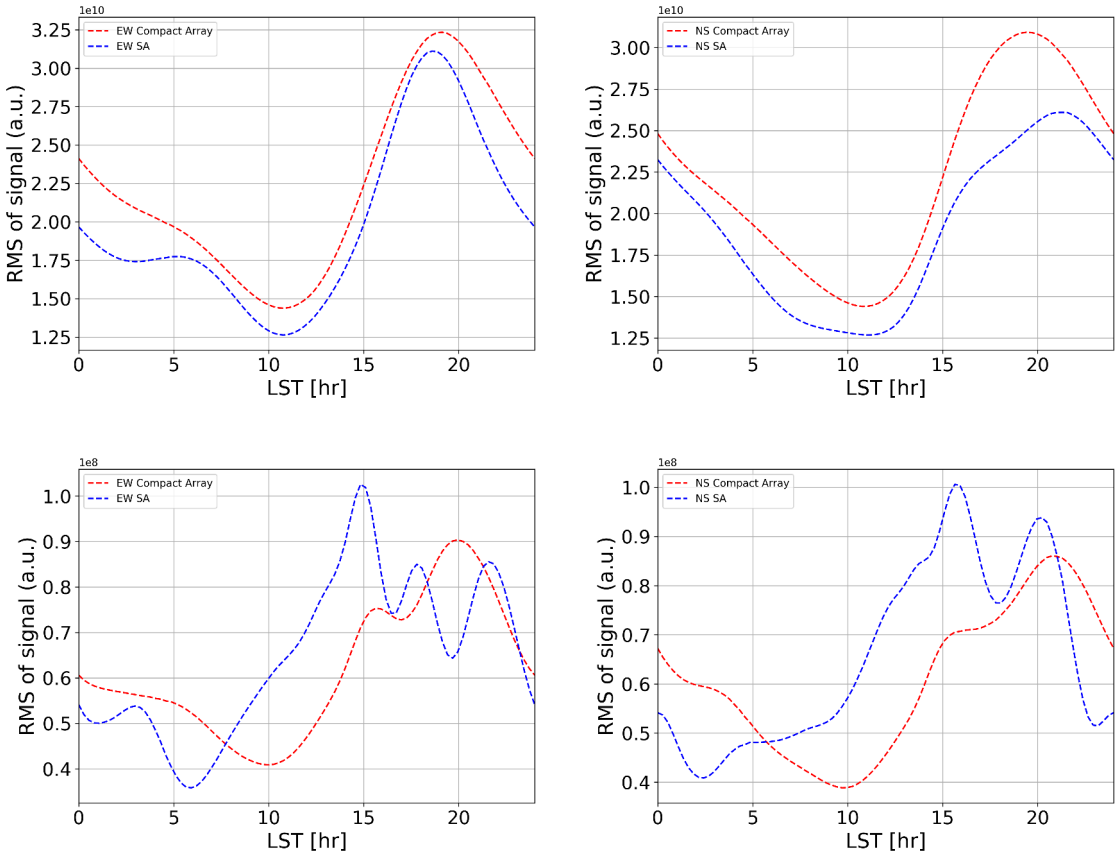


Figure 5.11 *– Simulated signal of the Galactic radio emission at* 50 MHz *(top) and* 150 MHz *(bottom) for the EW (left) and NS (right) polarization. The red lines are for the CA antenna, the blue lines are for the SA.*

the antennas, the greater the difference between the maximum and the minimum of the transit curve, the more the antenna is sensitive. The SA and CA antenna have the same sensitivity to the galactic variation.

5.3.5. Influence of the ground composition

The modelling of our radio detector response takes into account the ground characteristics (the ground conductivity σ_{ground} and the ground relative permittivity ϵ_{ground}). The influence of the composition of the ground is obviously important when looking the prediction curves without fit to the data, as shown in figures 5.12 which present the predictions of the transit curves in arbitrary units for several ground characteristics. The variability of the ground can be correlated with a seasonal variation, but it is difficult to know it at any moment, which also makes impossible an absolute calibration of the antennas. This is one of the reasons why we chose a relative method with a cross calibration of antennas permitting to get rid of the variations of the composition of the ground.

Figure 5.13 presents the predictions of transit curves fitted on data for the frequencies 50 MHz and 150 MHz and different ground compositions. Initially, an average ground composition was used (blue curve), corresponding to a finite ground conductivity $\sigma_{\text{ground}} = 5\,\text{mS}\cdot\text{m}^{-1}$ and a ground relative permittivity $\epsilon_{\text{ground}} = 13$. Other ground characteristics have been tested, only the two extreme ground compositions are shown in figure 5.13 (orange and green curves). A configuration with a ground plane, as the one equipped on the CA antennas, was also studied. We show that the use of an average

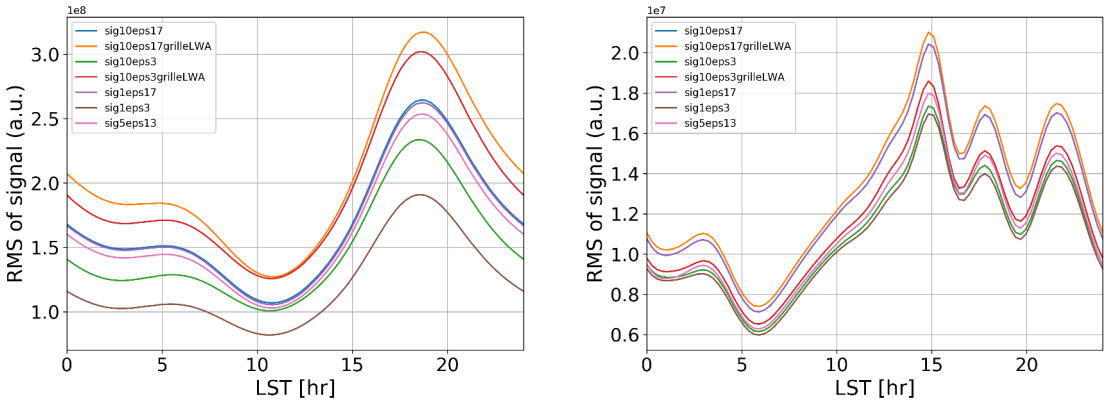


Figure 5.12 *– Top: predictions of transit curves of one SA at* 50 MHz *for the EW polarization and for different compositions of the ground. Bottom: same at* 150 MHz.

ground (in blue) permits to fit well the data (black points), and the differences between the different ground compositions are negligible when we perform a cross calibration. We have also found that the presence of the ground plane does not affect the results. This is even more true at 150 MHz, where the ground seems to have no influence.

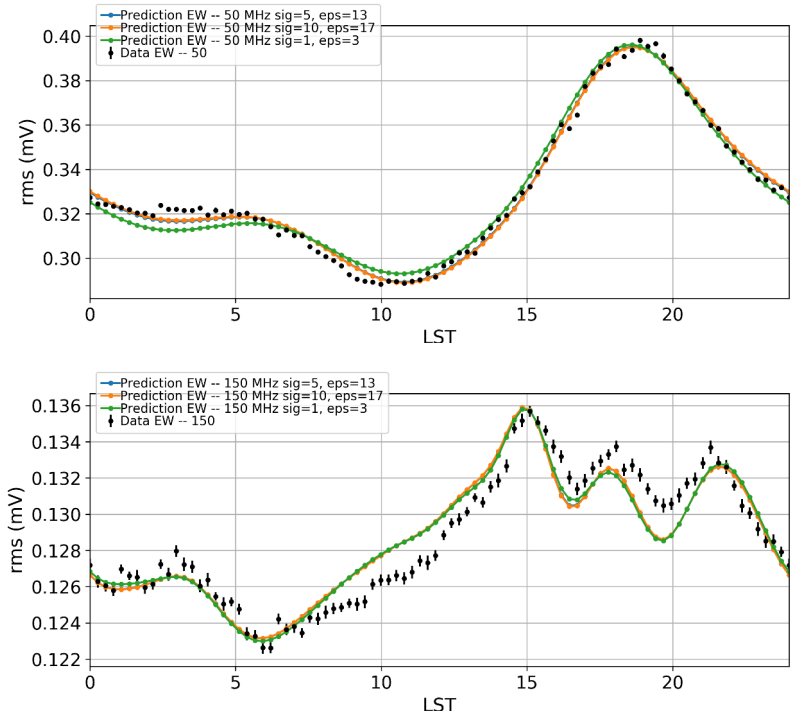


Figure 5.13 *– Top: predictions of transit curves of one SA fitted on data (black points) at* 50 MHz *for the EW polarization and for different compositions of the ground, i.e for different combinations of the finite ground conductivity* σ_{ground} *and ground relative permittivity* ϵ_{ground}*. Bottom: same at* 150 MHz.

We have now a complete description of our radio instruments and a prediction of the Galactic radio emission seen by our radio instruments for the [20 – 200] MHz frequency range. In the following, we propose a cross calibration of our radio instruments. The obtained calibration coefficients will contain all the inhomogeneities, mainly due to the

difference from the electronic chain and the environment of each antenna.

5.4. Cross calibration of the standalone antennas of CODALEMA

5.4.1. The data set

With the 1W data collected since 2015, a precise and continuous reference of background noise can be used to cross calibrate our radio instruments. Figure 5.14 presents the 1 W data as a function of time (UTC) for one week of February 2015 (left) and one week of August 2015(right). The daily galactic variation is visible: the maximum of the Galactic radio emission shifts in UTC of around twelve hours between February and August, which is consistent with the difference between the duration of one day in LST and UTC (four minutes offset each day, over 6 months). Moreover, an amplitude variation seems visible by comparing the mean behaviour of the RMS values of the week of February with the mean behaviour of RMS values of the week of August, which are shifted in amplitude along the y-axis.

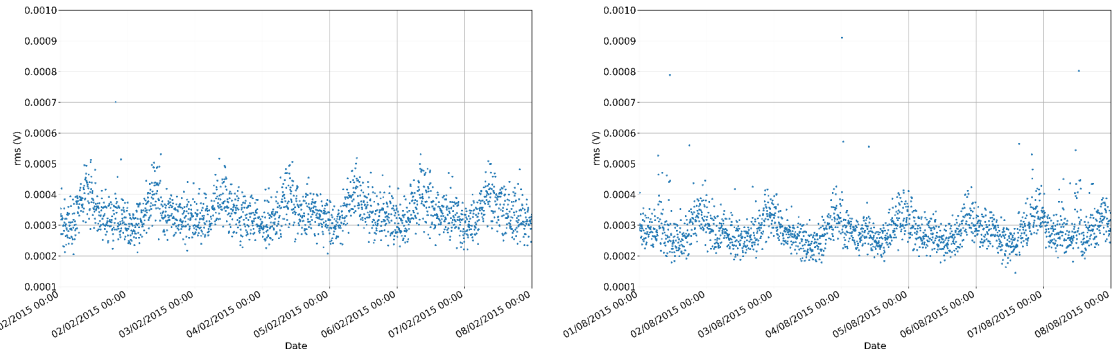


Figure 5.14 *– Illustration of the seasonal variation on two weeks of 2015. Left: February. Right: August. The seasonal variation is visible by comparing the mean behaviour of the RMS values of the week of February with the mean behaviour of RMS values of the week of August, which are shifted in amplitude along the y-axis.*

Indeed, as shown in figure 5.15 presenting the 1W data as a function of time (UTC) for the same SA that of figure 5.14 over 2.5 years, the fluctuation in the ambient noise level is clearly visible, and can be assimilated in first approximation to a seasonal variation. This variation is not common to all the antennas.

For instance, figure 5.16 presentes the 1W data as a function of time (UTC) for another SA over 2.5 years. The noise level detected by this SA is constant over approximately 2.5 years, on both polarizations, and the seasonal variation is not present.

Moreover a careful and systematic cleaning of our data set is made, by seeking for perturbation periods. We see different periods to exclude, namely between the 21st and 28th June 2015, corresponding to an abnormal increase of the noise level. This exclusion period is common to the whole antenna array (SA and CA), and is consistent with the magnetic solar disturbance (MSD) reported by [195]. The geomagnetic storms of solar cycle 24 which occurred in June 2015, in the active region (AR) 12371. It occurred over almost 1 week, between 18 - 25 June 2015. The 18th June at 00:33 UT, the AR12371 produced a coronal mass ejection (CME) associated to a flare of type M with a magnitude of 1.2 (noted M1.2) not directed to the Earth, as shown in figure 5.17 left. The same day, at 16:30 UT, a CME occurred, directed to the Earth, with an associated M3.0 flare that peaked at 17:36 UT. This flare was visible in our radio instruments, namely in the Compact Array

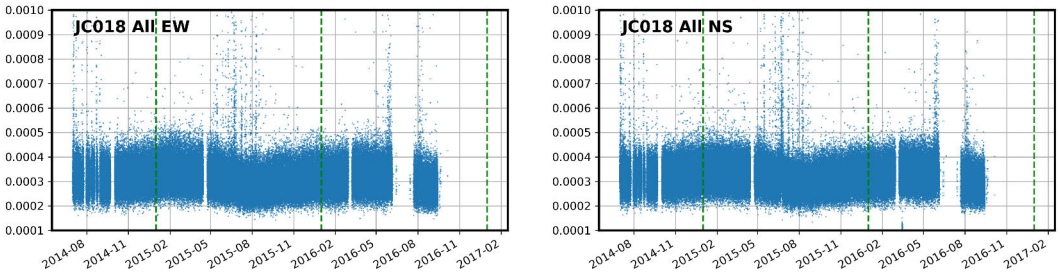


Figure 5.15 *– 1 W data for another station from 2014 to 2016. A fluctuation in the ambient noise level is present, which could be assimilated to a seasonal variation (hypothesis!). At the writing date of this manuscript, we have not found an explanation for these seasonal variations: this does not seem to be due to the environment of the antenna, since these seasonal variations are present on antennas with very different environments.*

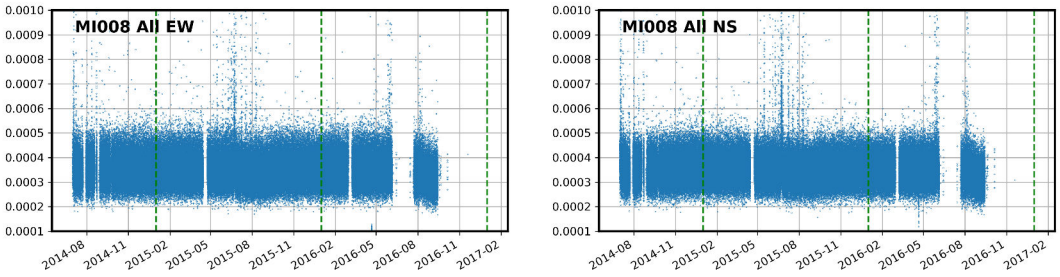


Figure 5.16 *– RMS value of one SA from 2014 to 2016. The noise level detected by the SA is constant over approximately two years. The vertical green lines stand for the beginning of each year.*

and in the SA. Several flares of different magnitudes follows as shown in figure 5.17 right, the most important arising on June 23rd.

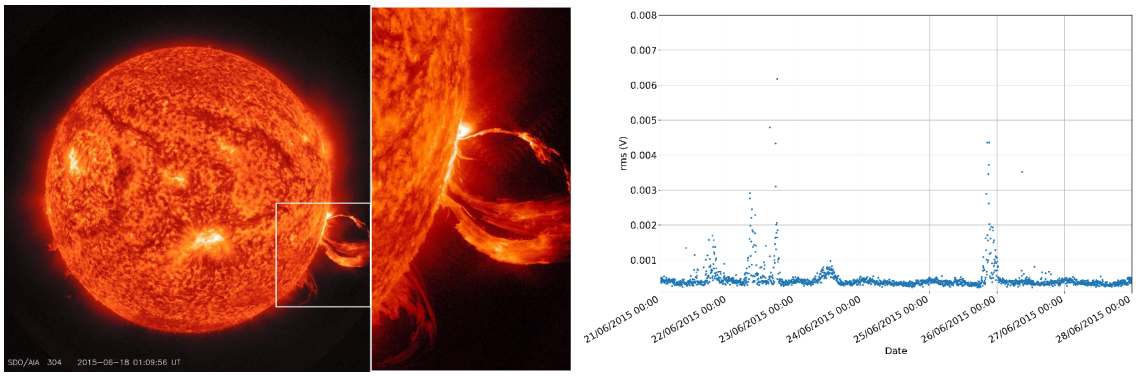


Figure 5.17 *– Left: Photography of the 18th June coronal mass ejection made by the Solar Dynamics Observatory with the telescope AIA (for Atmospheric Imaging Assembly) 304. From [196]. Right: Successive Solar flares observed by CODALEMA radio instruments.*

Remark: we have attempted to use this MSD to validate the calibration method described in this chapter. This was not conclusive. Indeed, as it can be seen in the figure 5.17 right, the observed flares are on the one hand decorrelated from the position of the Sun in the sky (flare observed at night), and on the other hand the mechanism of signal production

is poorly known: Arrival of solar charged particles into the atmosphere? Combination in a single wave of the Coronal Mass Ejections of solar flares previously described which go at different speeds [195] ? These solar flares are orders of magnitude lower than the one seen and studied in [197].

We choose data from 2015 as an absolute reference for the cross calibration. Figure 5.18 presents the 1 W data of one radio antenna for 2015 (red dots) and 2016 (blue dots) in LST. The predictions obtained as explained in section 5.3 are represented by the red (2015) and blue (2016) plain lines. We see that the behaviour of the ⟨RMS⟩ over two years is stable, permitting to choose one year as an absolute reference without biasing the results of the cross calibration. To alleviate the problem of the fluctuation in the ambient noise

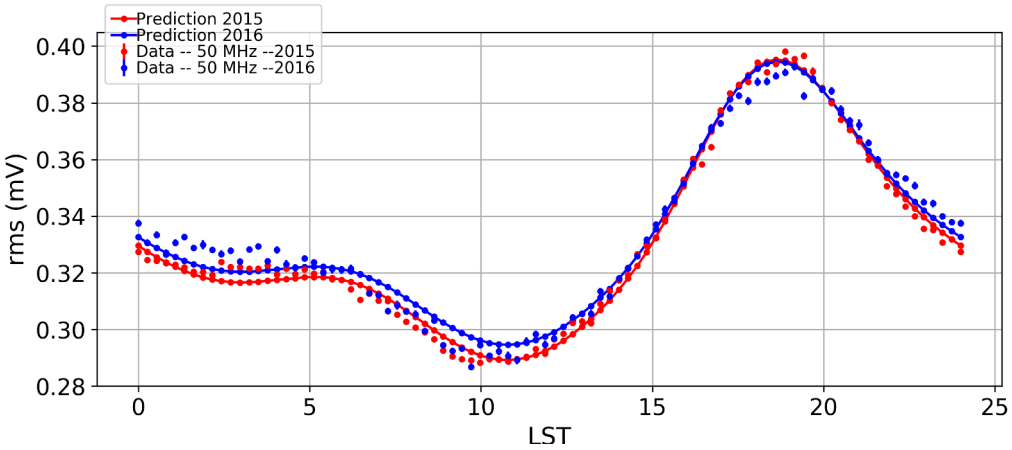


Figure 5.18 *– 1 W data for 2015 (red dots) and 2016 (blue dots) in LST. The predictions obtained as explained in section 5.3 are represented by the red (2015) and blue (2016) plain lines. Each dot is the average of all antennas over 16 min LST.*

level seen in different antennas, we have chosen to calculate the calibration coefficients month by month for each antenna. The calculation of the calibration coefficients, for a given frequency and a given antenna, is made as follows:

— The average transit curve (noted TC_{mean}) of the whole antenna array of 2015 is calculated: for each LST bin, the ⟨RMS⟩ is calculated as previously explained from the Gaussian distribution of the RMS of the 1W data.

— The galactic model is then adjusted to TC_{mean} with the equation $f(t) = k_{\text{mean}} \times TC_{\text{mean}}(t) + a_{\text{mean}}$ where t is the LST bin and k_{mean} and a_{mean} the two parameters to be adjusted, and will serve as the reference (noted Ref) to calculate the calibration coefficients. In general, the k parameter is related to the gain of the antenna, and the a parameter contain the information of the electronic chain of the antenna.

— The transit curve of the antenna is then adjusted to Ref with the same equation as above, giving the calibration coefficient couple (k, a) to apply to the transit curve of the antenna.

This procedure is made for all frequencies with a step of 1 MHz and all antennas. Figure 5.19 presents the k coefficient for the two antennas used in figures 5.16 and 5.15. A sine wave has been fitted to the coefficients, expecting that the seasonal variation follows this trend. The seasonal variation is visible on the right figure corresponding to the same antenna as for the figure 5.15, the fit of the sine wave gives a semi-period of ~ 5.7 months. As expected, the calculation of the calibration coefficients month by month permits to take into account this seasonal variation and to correct it.

Remark: the fluctuation in the ambient noise level seen in different antennas, interpreted as a seasonal variation, is most likely to be produced by a changing of the ground properties throughout the seasons.

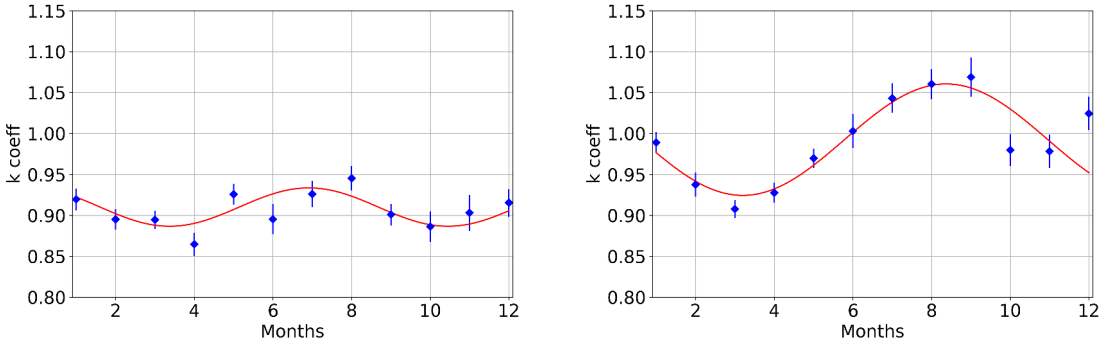


Figure 5.19 *– Illustration of the seasonal variation in the calibration coefficients. Left: calibration coefficient for the same antenna of figure 5.16. Right: calibration coefficient for the same antenna of figure 5.15. A sine wave has been fitted to the coefficients. The seasonal variation is visible on the right figure, with a semi-period of* ∼ 5.2 *months.*

5.4.2. Application of the procedure to the whole array and results

As explained above, for a given frequency and a given month, two calibration coefficients are calculated, noted in the following $k_{i,f,m}$ and $a_{i,f,m}$ with i the number id of the antenna, f the frequency and m the month. The calibration coefficients obtained are relative factors to the mean of the measured signal, antenna by antenna.

Then, the two calibration coefficients are applied to the measured signals as follows:

$$V^{cal}_{i,f,m} = k_{i,f,m} \times V_{i,f,m} + a_{i,f,m} \tag{5.6}$$

where $V_{i,f,m}$ is the measured signal in volt, $k_{i,f,m}$ and $a_{i,f,m}$ are the two calibration coefficients and $V_{i,f,m}$ is the calibrated measured signal. Figure 5.20 presents the results obtained after the cross calibration of the SA, at 50 MHz for an arbitrary month, for the EW (left) and NS (right) polarizations, to be compared with figure 5.3. After the cross calibration, the dispersion over the whole antenna array stands in around ∼ 6 %.

In other words, at 50 MHz and for the considered month, all the SA signals are now contained within a ∼ 20 µV wide ⟨RMS⟩ interval. For a typical shower event, the RMS, which is the sigma value obtained from the Gaussian distribution of the recorded signal, of one antenna is equal to ∼ 700 µV at 50 MHz. After the cross calibration, the obtained dispersion over the whole SA array is smaller than the RMS of one antenna by more than one order of magnitude.

5.5. Cross calibration of the Compact Array

The 1W trigger of the SA is not available for the CA antennas, which are externally triggered by the scintillators. However, CA waveforms are often composed only of background noise, allowing us to use the scintillator trigger to cross calibrate the CA array. We also know where the signal is expected to be in the waveform, allowing us to use, in case when a shower induced transient is present, the part of the waveform only composed of background noise. Figure 5.21 present ⟨RMS⟩ of the noise data in LST. As for the SA, the transit curves are stackable, meaning that the source of the emission is the same for all the antenna array, i.e the Galactic radio emission can be used to cross calibrate the CA array. The behaviour of the ⟨RMS⟩ of the noise data is similar for all the antennas over a day in LST, and only an offset along the y-axis is present. Unlike the SA, the CA antennas

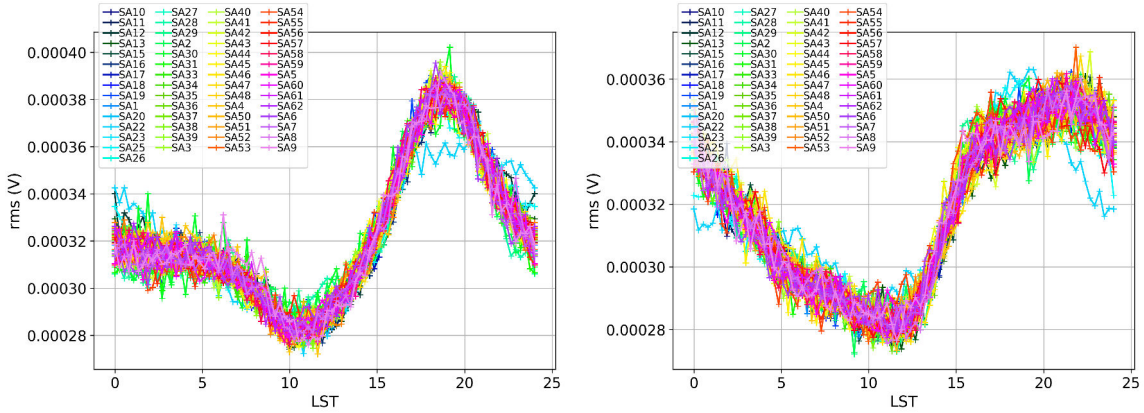


Figure 5.20 *– Results obtained after the cross calibration of the SA, at* 50 MHz *and for an arbitrary month, for the EW (left) and NS (right) polarizations. The blue curve that comes out of the global behavior of the antennas is understood, it is a reversal in the electronics of the polarizations for one antenna.*

does not present the seasonal variations. The transit curves are thus calculated over a complete year (against one month for the SA), explaining why the CA transit curves are smoother. The calibration coefficients are also calculated for a complete year.

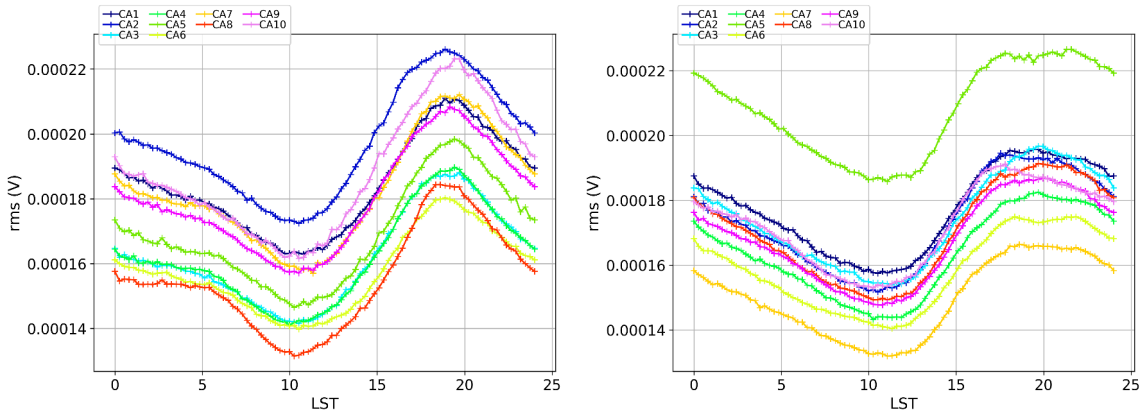


Figure 5.21 *–* ⟨RMS⟩ *of the CA data in LST for the EW (left) and NS (right) polarizations at* 50 MHz.

Figure 5.22 presents the predictions of transit curves for one CA antenna made with the NEC4 fitted on data (black points) at 50 MHz and 150 MHz. As for the SA, the predictions are in agreement with the data, validating the NEC4 model used to make the predictions. From that, the same procedure as the one used for the cross calibration of the SA was used for the cross calibration of the CA antennas of CODALEMA.

Figure 5.23 presents the results obtained after the cross calibration of the CA antennas, at 50 MHz for an arbitrary month, for the EW (left) and NS (right) polarizations, to be compared with figure 5.21. After the cross calibration, the dispersion over the whole CA is around ~ 3 % only.

As for the SA, the result obtained is quite good, since all the CA signals are contained within a 5 µV width ⟨RMS⟩ interval. For a typical event (same event as the one used for the SA), the RMS of the recorded signal of one antenna is equal to ~ 400 µV. After the cross calibration, the obtained dispersion over the whole CA array is smaller than the RMS of the signal recorded by one antenna by two orders of magnitude. The dispersion

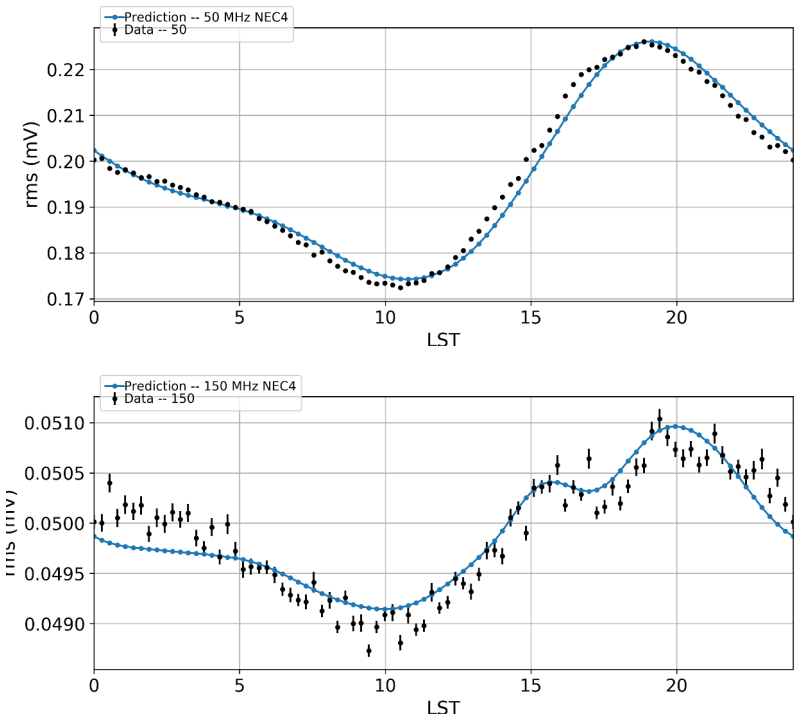


Figure 5.22 *– Top: predictions of transit curves fitted on data (black points) at* 50 MHz *for one CA antenna. Bottom: same at* 150 MHz *for one CA antenna.*

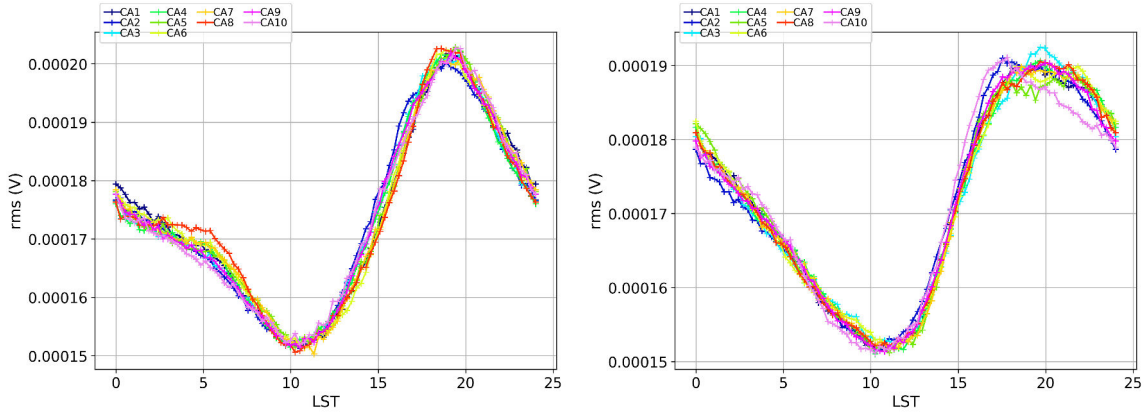


Figure 5.23 *– Results obtained after the cross calibration of the CA antenna, at* 50 MHz *and for an arbitrary month, for the EW (left) and NS (right) polarizations. All signals are contained within* 5 µV *interval.*

of the CA antennas is less important than the dispersion of the SA: due to different gain in the electronics chain, the RMS of the signal is weaker for the CA antennas than for the SA. Indeed, the dispersion of the CA antennas *before* the cross calibration was around 40 µV, which is approximately equal to the dispersion of the SA *after* the cross calibration. Before the calibration, the dispersion of the SA was around 70 µV excluding the two outliers. It is worth noticing that we also have 57 SA and only 10 CA antennas, the latter being spread over a much smaller area (the distance between two SA is ~ 150 m, all the CA antennas are contained in ~ 145 m for the two farthest CA antennas). One can infer that the local noise conditions are certainly more different for the SA than for the CA antennas, leading to a larger ⟨RMS⟩ noise figure.

5.6. Application of the calibration coefficients

Now that the calibration coefficients are obtained, we want to apply them to the recorded waveforms. Since the coefficients are calculated for each frequencies, we decide to apply them during the filtering procedure. We use a rectangle filter, on which we apply a Tukey window (also known as the tapered cosine window). It can be assimilated to a cosine lobe of width $\alpha(N+1)/2$ convolved with the rectangular window used of width $(1-\alpha/2)(N+1)$, where α is a parameter to adjust ($\alpha = 0$ corresponding to a rectangular window) and N the number of samples of the signal. The filtering procedure is:

— We apply the Fourier transform to the recorded signal ($N = 2{,}560$).
— We define a rectangle window, where we set to zero the Fourier coefficients corresponding to the frequencies that we want to filter and to one for the frequencies that we want to keep,.
— We apply a Tukey window on the defined rectangle window, with α set to 0.2.
— The matrix of the calibration coefficients is then applied on the final window containing the Fourier coefficients corresponding to the frequencies that we do not want to filter.
— We apply the inverse Fourier transform to retrieve the filtered signal in the time domain.

Figure 5.24 presents an example on an actual shower of the application of the calibration coefficients on signals filtered in [30 – 80] MHz for the EW polarizations. The shape of the signal is conserved after the application of the calibration, only the amplitude is changed. This is confirmed by the polarization of the signals (see figure 5.25), for which the orientation is globally not changed, only the amplitude of the ellipse of polarization is modified. Thus, the application of the calibration coefficients during the filtering procedure seems to be correct, since the phase of the signal is not modified.

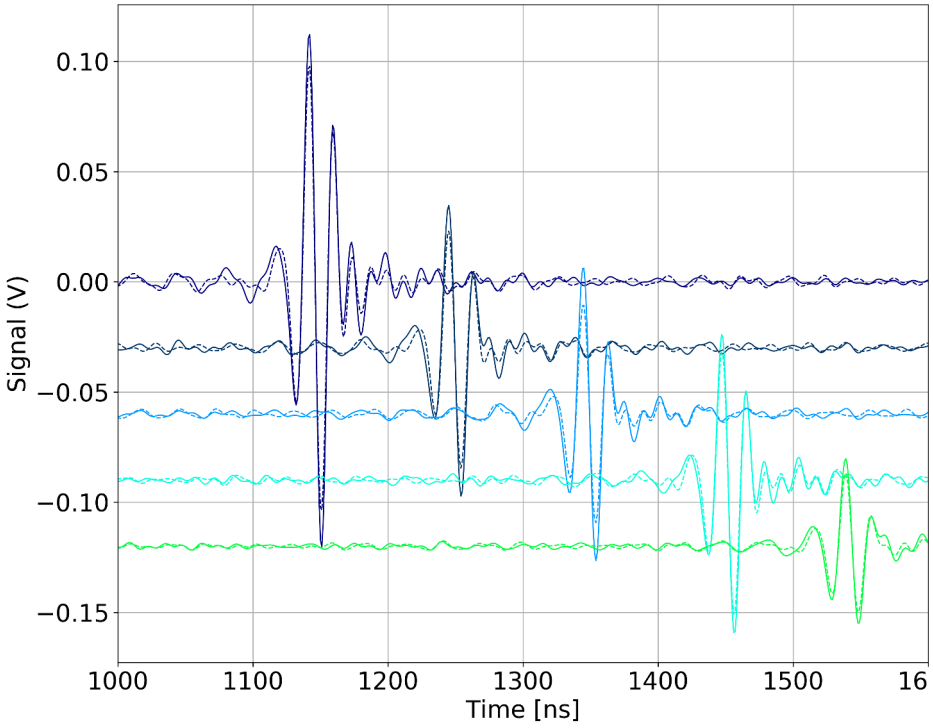


Figure 5.24 *– Example on an actual shower of the application of the calibration coefficients on signals filtered in* [30 – 80] MHz *for the EW polarizations. Plain lines are the filtered signals without calibration, dashed line filtered signal with calibration*

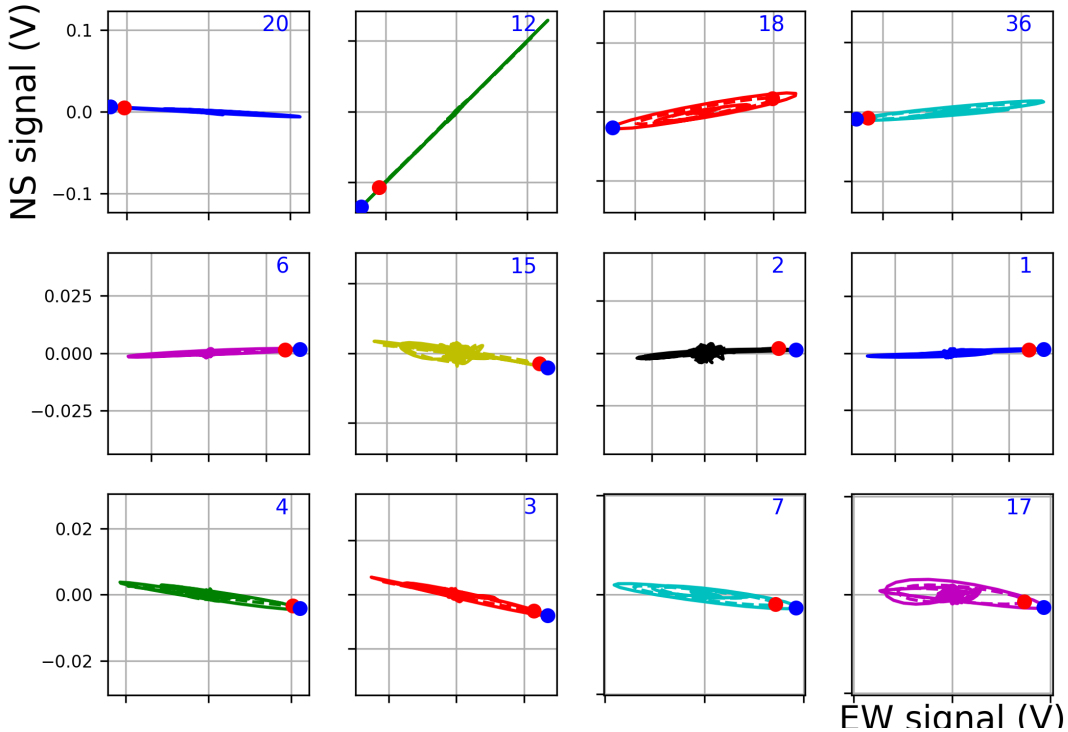


Figure 5.25 – *Example on an actual shower of the application of the calibration coefficients: polarization of signal with (dashed lines) and without (plain lines) calibration. The maximum of the ellipse of polarization is indicated by red (with calibration) and blue (without calibration) points.*

5.7. Validation and quantification of the effect of the calibration

We know for sure that it is necessary to correct the antenna/antenna disparities that exist on our radio arrays. Although it is obvious that this is the role of calibration, this section aims to validate first the calibration, and then to quantify its effect.
We want to draw the reader's attention to one point. We have shown that the calibration makes it possible to significantly reduce the disparities existing between the antennas of the same array, in particular by showing that the transit curves are all superimposed (the SA are contained within a 40 µV). But it could remain a significant offset (coefficient a), which could subsequently led to an overestimation of energy for example. This is the first check to be made: we can see in the figure 5.3 that all the curves are around the value 320 µV, which is always the case after the calibration (figure 5.20). This is a first indication that the calibration procedure leads to small correction factor and not to large extrapolation associated to large uncertainties. Moreover, the main validation of the global chain of analysis will be shown in chapter 7 with the improvement of the chi-square value with the generalized reconstruction method which includes the calibration (see section 7.4.3 page 172). We will focus here on the qualitative rather than the quantitative that will be treated in section 7.4.3.

5.7.1. Using the convolving method

We choose to illustrate the validation and quantification via the convolving method on an actual shower event recorded by CODALEMA (in coincidence in the SA and CA) in September 2016. We apply the reconstruction method detailed in chapter 7 (with the calibration), and we obtain the estimation of the shower core position, of the $X_{\max}$ and of the energy. With these parameters, we are able to compare the convolved simulation of

one antenna with the actual data recorded by the same antenna.

Figure 5.26 (top part) presents the results obtained for the same antenna that of figure 5.15 which featured a seasonal variation. After the reconstruction and the calibration, we see on figure top right, presenting the PSD of the recorded data and of the simulated antenna, a really good agreement over [20 – 200] MHz. This indicates that we well master the procedure of the convolving/deconvolving, and that the calibration is also correct. This result is obtained for the 22 SA involved in the event.

Figure 5.26 (bottom part) presents the results obtained for a CA antenna. The event seen in the CA being in coincidence with the event seen in the SA, we expect to also obtain a good agreement between the PSD of the recorded data and the simulated data. After the reconstruction and the calibration, we see on figure top right that the PSD are not in good agreement, namely at the amplitude level than the spectrum form. The wrong agreement could be explained by:

— The estimation of the energy which could be wrong, explaining why the amplitudes of the CA PSD are not in agreement: the amplitude of the PSD of the 22 SA are in agreement with the simulations, meaning that this point can be ruled out.
— An error in the procedure of the convolving/deconvolving: the procedure for the CA being the same as the one used on the SA, this point can be ruled out.
— An error or a bad sizing of the electronic chain of the CA: this point being different for the two arrays, this seems to be the better explanation. At the time of writing this thesis, the problem is being investigated.

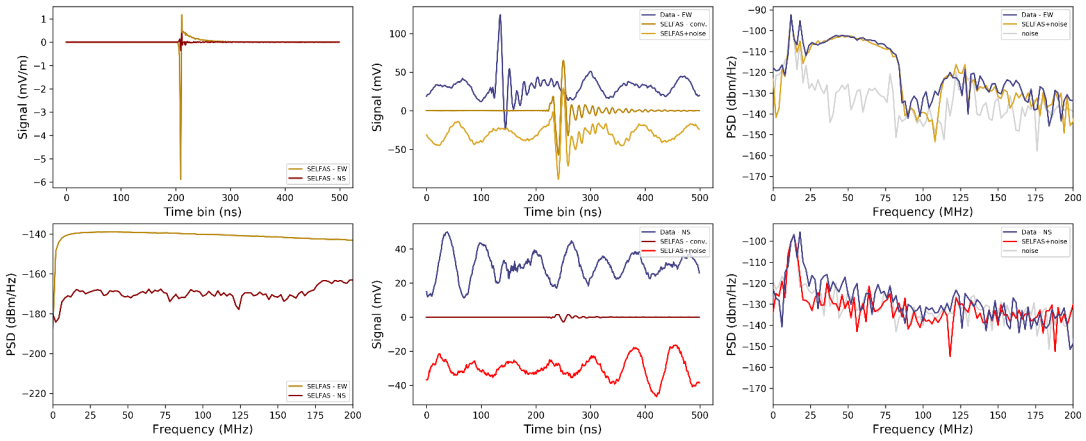
This poor mastering of the convolving of the CA is surely at the origin of the gaps that we observe in the following section.

5.7.2. Using the deconvolving method

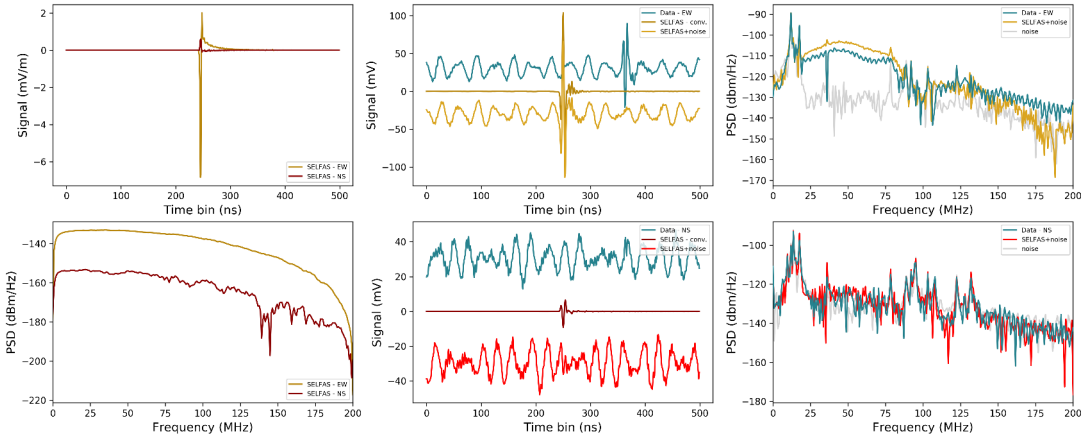
At present, the SAs are comparable to each other, and the CA antennas are comparable to each other. In order to validate our cross calibration method, we will verify that an SA is comparable to a CA antenna on two shower events. The first shower event used was detected in September 2016. It presents an interesting ground pattern. Indeed, the SA number 2 (SA2) and the CA antenna number 2 (ARR2), which are closed together ($\sim$ 40 m), have detected a signal. The calibration coefficients calculated above are thus applied on the waveforms of each antennas, and the deconvolving procedure is then applied. The hypothesis used is that a SA and CA antenna at the same position, i.e at the same position with respect to the shower axis, should detect the same electric field after calibration and deconvolving, if the calibration and the description of the response of the antennas are valid.

Figure 5.27 presents the comparison of SA2 and ARR2 deconvolved before (left column) and after (right column) the application of the calibration, in the MF band (top) and in [140 – 200] MHz (bottom). In the MF band, the SA2 and ARR2 were already in agreement before the calibration (PSD are contained within less 5 dB), and stay in agreement after the calibration (PSD are also contained within less 5 dB). The calibration in the MF band does not bring a major improvement. However, in [140 – 200] MHz, the disparities before the calibration are much visible ($\sim$ 10 dB), and are reduced after the calibration.

The same work has been made on another event recorded in May 2019. Figure 5.28 presents the comparison of SA2 and ARR2 deconvolved before (left column) and after (right column) the application of the calibration, in the MF band (top) and in [140 – 200] MHz (bottom). As for the first event, the SA2 and ARR2 were already in agreement before the calibration in the MF band, and stay in agreement after the calibration. In [140 – 200] MHz, and for this event, it is hard to draw a conclusion. The 165 MHz peak



(a) *SA*



(b) *CA*

Figure 5.26 *– Result figure obtained after the comparison of simulations and data. The canvas is divided in six figures. Top left: simulated signals in EW and NS polarizations. Bottom left: PSD of the simulated signal in the EW and NS polarizations. Top center: recorded signal, convolved simulated signal and convolved simulated signal with added noise in the EW polarization (the curves are positively shifted along x and y for clarity). Bottom center: same as top center in the NS polarization. Top right: PSD of the recorded signal and of the convolved simulated signal in the EW polarization. Bottom right: same as top right in the NS polarization. (a): results obtained for the SA. (b): results obtained for the CA.*

appears to be an artefact related to the deconvolving. During the deconvolving procedure, if the direction of arrival is changed of 2 degrees, the peak disappears. This problem is being investigated.

5.7.3. Validation and quantification on SELFAS simulations

In order to validate the calibration and to quantify its effect, we choose to work on one SELFAS simulation ($\theta = 22\,°$ and $\phi = 101\,°$), using the SA array for the simulated antenna array. This simulation serves as “fake” data for which we know the core position and the energy of the simulation. This is the simulation of a real shower event detected by CODALEMA, with a multiplicity of 8 antennas. We then apply the procedure explained in chapter 7 by fitting 50 simulations to the “fake” data. To test the utility of the calibration,

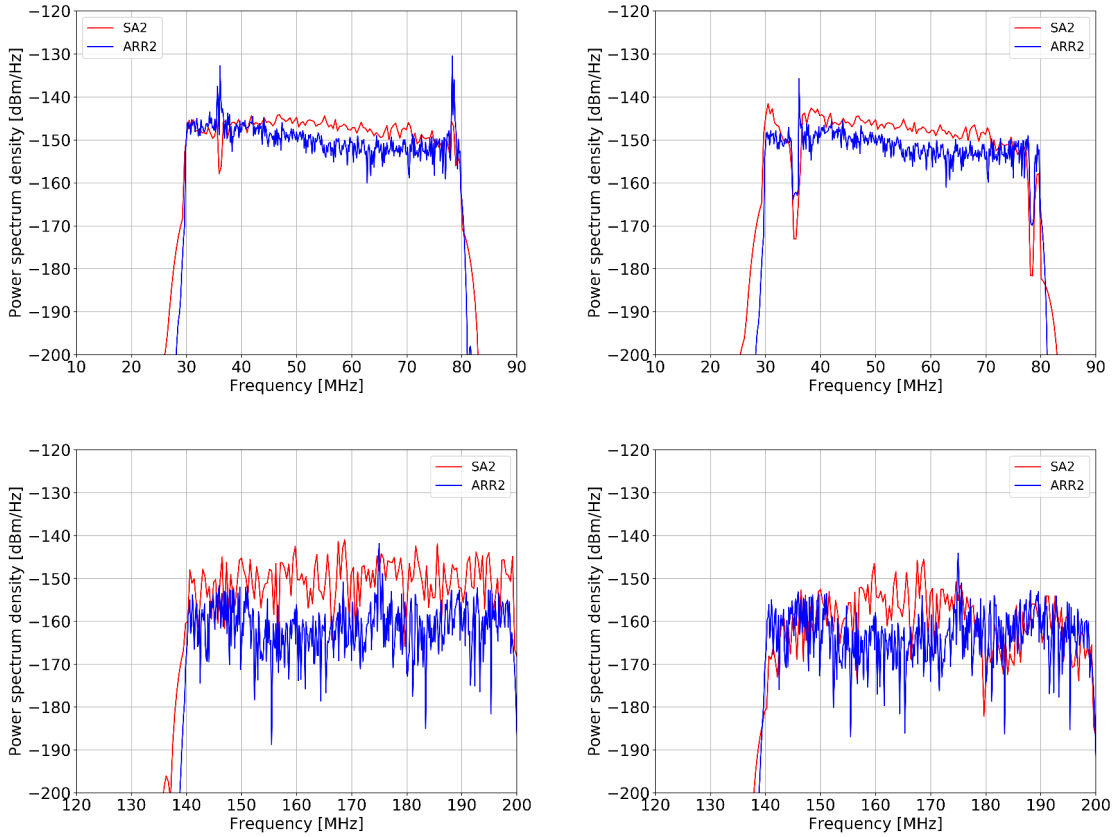


Figure 5.27 *– Top: Comparison of the PSD of the EW polarization of SA2 and ARR2 deconvolved before (left) and after (right) the application of the calibration in the MF band. Bottom: same as top but in* [140 – 200] MHz. *Transmitters at 37 and* 78 MHz *are systematically filtered. Bottom: same at EMF.*

the reconstruction is performed on the de-calibrated (inverse of the procedure explained above) "fake" data first and on the "fake" data then. Knowing the parameters to be reconstructed, we can evaluate the performances and the utility of the calibration. We expect a worst accuracy on the reconstructed parameters with the de-calibrated "fake" data. The results are shown in table 5.1. The deviation from the actual values of the parameters is indicated: X_c is the x-position of the shower core, Y_c is the y-position of the shower core and α is the scaling factor related to the energy (see chapter 7).

	ΔX_c [m]	ΔY_c [m]	$\Delta\alpha$ [%]
De-calibrated	-5 ± 1	24 ± 12	-26 ± 11
Calibrated	-5 ± 1	-7 ± 14	-18 ± 9

Table 5.1 *– Validation of the cross calibration on a SELFAS simulation ($\theta = 22°$ and $\phi = 101°$). The deviation from the actual value of the parameters is indicated. X_c is the x-position of the shower core, Y_c is the y-position of the shower core and α is the scaling factor related to the energy.*

Since the deviations from the actual values of the parameters are larger for the de-calibrated "fake" data, we conclude that the calibration is necessary. On this event, the calibration has a real impact on the determination of the Y_c position of the shower core and of the energy.

The same work has been made on another simulation of an actual shower event with

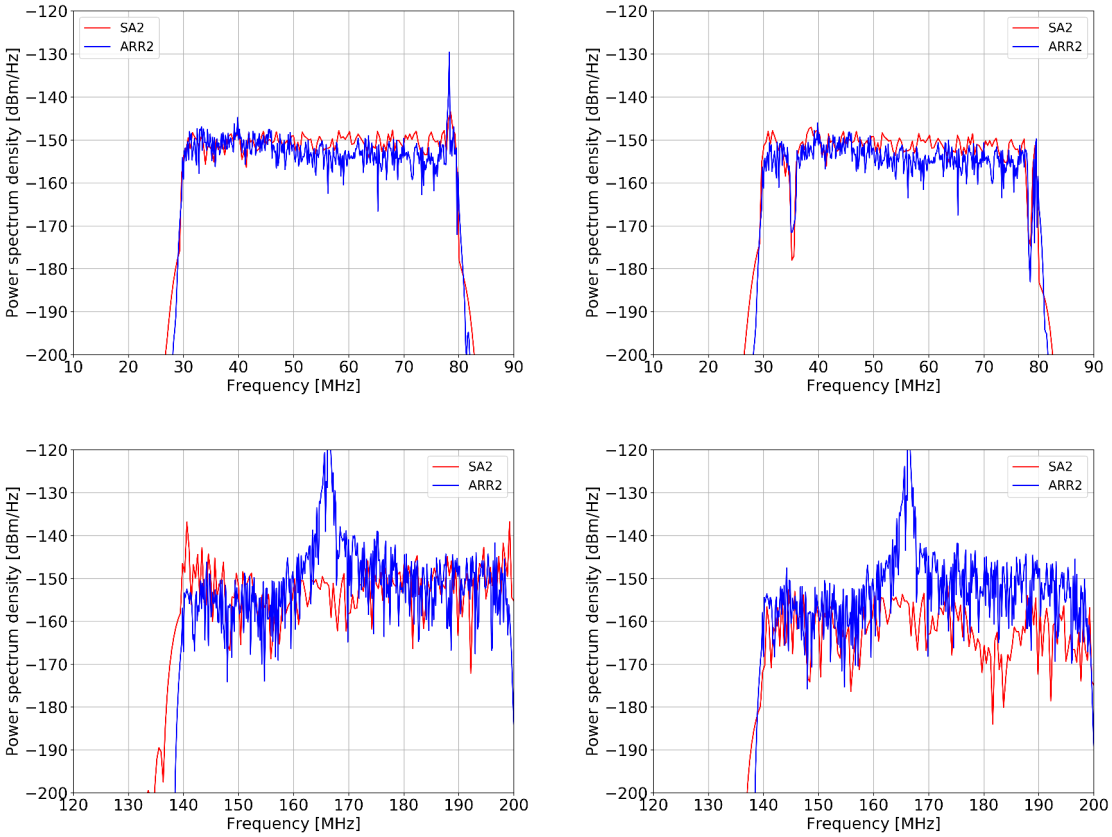


Figure 5.28 *– Top: Comparison of the PSD of the EW polarization of SA2 and ARR2 deconvolved before (left) and after (right) the application of the calibration in the MF band. Bottom: same as top but in* [140 – 200] MHz. *Transmitters at 37 and* 78 MHz *are systematically filtered. Bottom: same at EMF.*

a larger multiplicity (multiplicity of 47) and with a more inclined arrival direction ($\theta = 68\,°$ and $\phi = 204\,°$), and the results are summarized in table 5.2. As for the first simulation,

	ΔX_c [m]	ΔY_c [m]	$\Delta\alpha$ [%]
De-calibrated	-14 ± 14	-31 ± 26	3 ± 6
Calibrated	10 ± 17	-22 ± 25	-0.1 ± 6

Table 5.2 *– Validation of the cross calibration on an inclined SELFAS simulation ($\theta = 68\,°$ and $\phi = 204\,°$). The deviation from the actual value of the parameters is indicated. X_c is the x-position of the shower core, Y_c is the y-position of the shower core and α is the scaling factor related to the energy.*

the deviations from the actual values of the parameters are larger for the de-calibrated "fake" data. However, the impact of the calibration is lower on this simulation. This is explainable by the multiplicity, which is 47 for this event, against 8 for the first event. The large multiplicity of the event minimizes the effect of the correction of the calibration.

Based on the two example events described in this section, it seems that the cross-calibration of the antennas permits to better reconstruct the characteristics of the primary cosmic ray events. It is a proof of principle which, coupled with the validation made via two different instruments (CA antennas and SA) in the previous section, permits at this stage of the manuscript to validate the calibration method. We will see in chapter 9 that the calibration coupled with the reconstruction method described in chapter 7 gives

satisfactory final results, especially when comparing the energy reconstructed with the radio signal and the energy reconstructed with the particle detector array.

"We're not here just to take part.
We're here to take over!"
Conor McGregor

CHAPTER 6

Online and offline selection of radio events

Outline

In this chapter, we give some hints to improve the selectivity of the radio events for the CODALEMA experiment, mostly for the standalone antennas. Indeed, the detector is not triggered by a particle detector array but directly on the radio signal, and as we will see, the number of parasites becomes very important. It will not be discussed here the selection of events for the Compact Array and the three-fold antenna: since these two instruments are externally triggered by scintillators, a simple method of amplitude threshold or a beam forming technic combining the individual signals allow to highlight a potential transient related to a cosmic ray. On the other hand, although the low-frequency antennas are also triggered by the particle detectors, the signal and the noise environment in which it is sought is completely new, and therefore requires the development of analysis tools to be detected, which will be discussed in the second part of this chapter.

As it will be discussed, when an event is detected, the first information to reconstruct is the arrival direction before to select them. The method used to reconstruct the arrival direction is valid for all instruments of CODALEMA and EXTASIS, for both radio detectors and particle detectors.

6.1. Reconstruction of the arrival direction

The reconstruction of the arrival direction is known as the "Difference Of Arrival" problem. To estimate the arrival direction, we use the time of arrival of the signal in the stations and their positions. The chosen conventions are to set the North at $\phi = 90\,°$ and the East at $\phi = 0\,°$. $\theta = 0\,°$ corresponds to zenith. Assuming that the front shape is a plane (first approximation, simple to implement), we perform a plane fit just as follows to determine θ, ϕ and the time of arrival t_{det} at the position (0, 0):

$$\chi^2 = \frac{1}{N-3}\sum_{i=1}^{N}\left[c(t_{0i} - t_{\text{det}}) - (u.x_i + v.y_i)\right]^2 \tag{6.1}$$

where N is the number of involved antennas (also noted multiplicity of the event), t_{0i} the time corresponding to the detection of the signal in the i^{th} antenna, (x_i, y_i) the coordinates of the i^{th} antenna and u and v the direction cosines expressed as:

$$u = \sin\theta\cos\phi \tag{6.2}$$

$$v = \sin\theta\sin\phi \tag{6.3}$$

from which we determine the zenith angle:

$$\theta = \arcsin\left(\sqrt{u^2 + v^2}\right) \tag{6.4}$$

and the azimuth angle:

$$\phi = \arctan\left(\frac{v}{u}\right) \tag{6.5}$$

For this estimation, we suppose that all detectors are at the same altitude (z-coordinate), giving the simplified equation 6.1, from which we can mathematically define the matrix

$$A = \frac{1}{\sigma^2}\begin{pmatrix} \sum x_i{}^2 & \sum x_i y_i & -\sum x_i \\ \sum x_i y_i & \sum y_i{}^2 & -\sum y_i \\ \sum x_i & \sum y_i & -\sum 1 \end{pmatrix} \tag{6.6}$$

where σ is the uncertainty in time, set to 10 ns for the case of the standalone antennas, 30 ns for the scintillators, 10 ns for the Compact Array and 100 ns for the low-frequency antennas. The choice of these values will be discussed in section 6.5. By inverting the matrix A, we obtain the uncertainties and covariance:

$$\sigma_u^2 = A_{11}^{-1} \tag{6.7}$$

$$\sigma_v^2 = A_{22}^{-1} \tag{6.8}$$

$$\sigma_{t_0,s}^2 = A_{33}^{-1} \tag{6.9}$$

$$\text{cov}(u, v) = A_{12}^{-1} \tag{6.10}$$

We then can calculate the uncertainties on the arrival direction:

$$\sigma_\theta^2 = \sigma_u^2\left(\frac{\partial\theta}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial\theta}{\partial v}\right)^2 + 2\text{cov}(u, v)\left(\frac{\partial\theta}{\partial u}\right)\left(\frac{\partial\theta}{\partial v}\right) \tag{6.11}$$

$$\sigma_\phi^2 = \sigma_u^2\left(\frac{\partial\phi}{\partial u}\right)^2 + \sigma_v^2\left(\frac{\partial\phi}{\partial v}\right)^2 + 2\text{cov}(u, v)\left(\frac{\partial\phi}{\partial u}\right)\left(\frac{\partial\phi}{\partial v}\right) \tag{6.12}$$

Thus, for each event and each instrument, an uncertainty on the arrival direction is estimated. It is worth noticing that concerning the scintillator array, the signals of the PMTs are carried to the central shelter by 300-meter-long cables (~200-meter-long for the radio antennas of the Compact Array). The characteristic delay of each cable has

been measured by reflectometry. The delay of each cable is taken into account in the t_0 term of the equation 6.1 to retrieve the actual particle collection time at the scintillator ($t_{0i} = t_i - t_i(cable)$).

The angular resolution as a function of the multiplicity is discussed in the following section (see figure 6.3).

6.2. CODALEMA – Noise rejection algorithms

Several noise rejection algorithms can be used to improve the efficiency and the selectivity of the radio events recorded by the SA. For example, the events mostly detected by CODALEMA are "Anthropic and Almost Stationary" (AAS). They come from sources giving event accumulations more or less in their direction of arrivals such as public electrical transformers or the wind farms at the South of the Nançay Radio Observatory. We list two categories of sources of noise, noted RFI for Radio Frequency Interference. There are the stationary sources giving accumulations in the direction of arrivals and the moving sources, such as planes. Some rejection algorithms used online in the central DAQ are described in section 4.2.2.3.2 page 89. Concerning the AAS events, the rejection algorithm uses the fact that the sources are stationary. The initial version of the algorithm was using the arrival direction of the events. One found out that a simple angular cut to reject sources at the horizon was performing poorly especially when the arrival direction was artificially reconstructed above the horizon and following a pattern clearly signing a digital artifact (see figure 6.1 left). Time differences of pairs of stations revealed to be

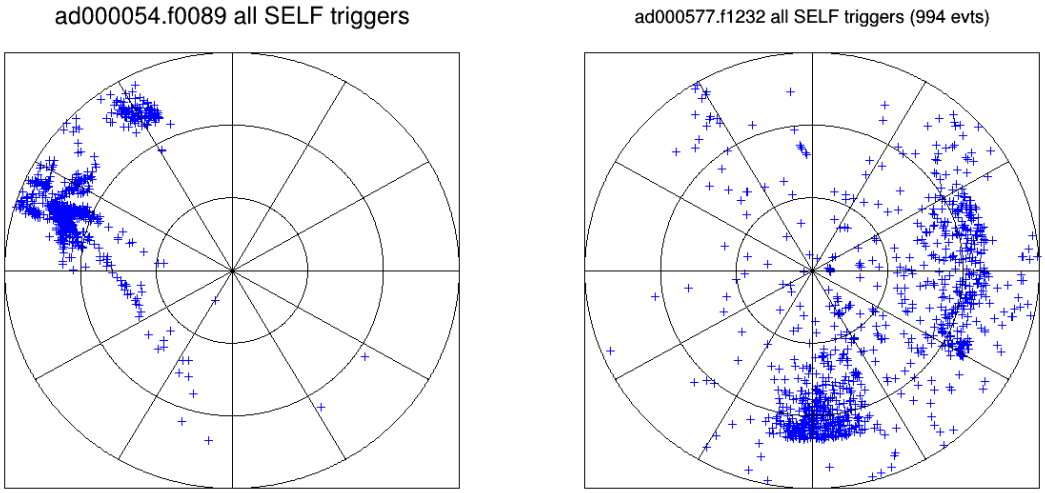


Figure 6.1 *– Two skymaps of some incoming parasitic events. The North is oriented with the geographical North (top). Left: Digital artifact in the direction of arrivals due to a poor reconstruction of the source at* ~ 160°. *Right: Effect of the angular rejection. A cut at* ~ 70° *is clearly visible at the bottom of the skymap, in a zone where the AAS rejection algorithm was performed.*

more robust and efficient criterion to reject dynamically noise sources. If all the time differences of all pairs of involved standalone antennas correspond to the time differences expected for an identified source, the event is rejected. For the event passing this selection, a second rejection level is applied. The quality of the angular reconstruction is evaluated by calculating the geometrical spread of the event defined as the chi-squared of the linear fit of the coordinates of the involved standalone antennas. Events are sorted

by their geometrical spread. The events with a small spread are rejected. As an example, for two weeks of data taking, 52.5 millions of events are built by the central DAQ and only 1 % are selected after the different stages of rejection (AAS and cone algorithms), corresponding roughly to 550,000 events. In this batch of events, only 34 are tagged as cosmic ray events (0.006 % of the ~550,000 recorded events). The capability of detection of cosmic rays is limited by our T1 dead time, which can not be decreased with the current electronics used on CODALEMA. The idea is thus to select the event at the T1 level, different methods being studied to increase the rejection rate. Figure 6.1 presents the skymap of the incoming events received by the `postmaster` of the central DAQ over two different time periods, illustrating the digital artifacts (left) and the effect of the angular rejection (right). Both skymaps comport roughly 1,000 events. Figure 6.1 left corresponds to a 35-minutes period during nighttime (average rate of parasitic transients of ~ 0.5 Hz), while figure 6.1 right corresponds to a 10-minutes period during daytime (average rate of parasitic transients of ~ 1.7 Hz). This illustrates what was shown in figure 4.2, where emitters are much more present during daytime than during nighttime in the [10 – 20] MHz range.

In this section, we propose to illustrate the different offline noise rejection algorithms applied after the T2 and T3 levels, which can be used to improve the radio event selectivity. They are based on specific features of the shower transients or of the parasitic signals. These are based on the rise time of the transient (noted RTA), on the compacity of the event (noted CtA) and on the polarizations of the event (noted PA). Their results are presented hereafter.

Let us consider one typical day randomly chosen. For that day, among the ~ 3.7 millions of T3 considered by the CDAQ, 50,096 events have been selected and passed the AAS, spread and cone criteria. In this batch of events, only 1 is tagged as an actual shower, meaning that 50,095 are fortuitous. These fortuitous events should not be recorded. The skymap of the reconstructed directions of arrival of the 50,096 events is shown in figure 6.2 left where two plane trajectories are visible, corresponding to the curved trajectories formed by the succession of the reconstructed direction of arrivals. The sources of noise are numerous, reconstructed in a diffuse way (explaining the total sky's occupation) and distributed all over the horizon confirming the contents of the snapshots of the CDAQ illustrated by the figure 6.1.

The uncertainties on the zenith and azimuth angles as a function of the multiplicity for these events are presented in figure 6.3. The uncertainty is calculated with the equation 6.11 derived in section 6.1. As expected, the larger the multiplicity is, the smaller the uncertainty on the angles is. The uncertainties on the zenith angle can explain the total sky's occupation and the spreading patterns of the sources on the skymap of the figure 6.1 right.

6.2.1. Rise Time Algorithm

This algorithm is based on the rise time of the transient recorded by one radio antenna, defined as the time between two chosen values of the normalized cumulative function of the signal:

$$C(i) = \frac{\sum_{k=b_{start}}^{b_{start}+i} s(k)^2}{\sum_{k=b_{start}}^{b_{end}} s(k)^2} \tag{6.13}$$

where $s(k)$ is the filtered signal of one polarization. Usually b_{start} and b_{end} are set to 1,000 and 1,500 in order to consider a narrow window around the trigger position (set to 1,100) in the signal set in the local acquisition. The algorithm is optimized for the classical frequency band, and the rise time is set between 70 % and 10 % of the cumulative function:

$$R_t = C(70\,\%) - C(10\,\%) \tag{6.14}$$

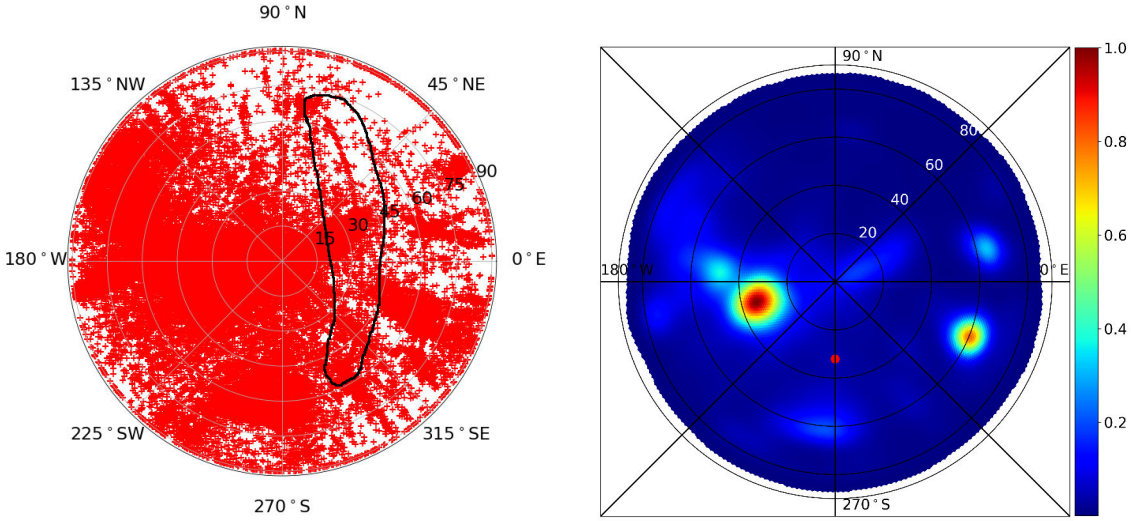


Figure 6.2 *– Left: Skymap of events accepted by the central DAQ. The reconstructed direction of arrivals are represented by the red crosses. The azimuth angle is indicated on the outskirts of the circle, the zenith angle is indicated inside the circle. The North is oriented with the geographical North (top). Two plane trajectories are visible, corresponding to the curved trajectories formed by the succession of the reconstructed direction of arrivals. One of them is surrounded in black. Right: same as left,* 10° *gaussian smoothed sky map.*

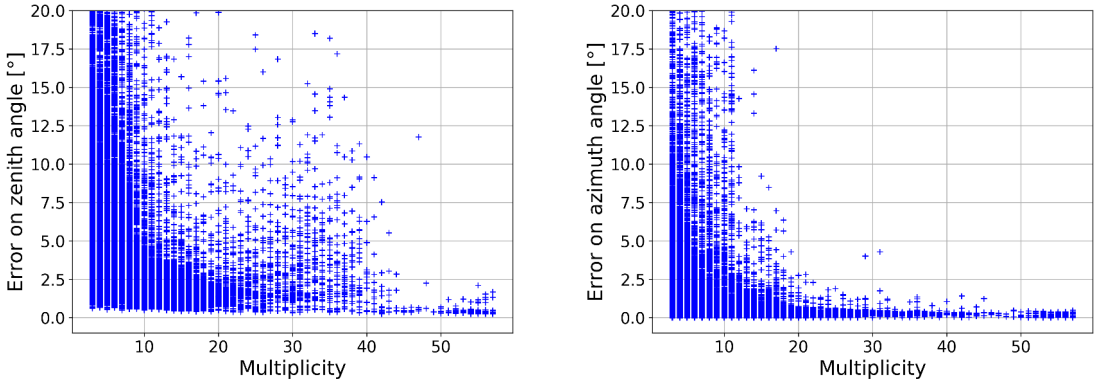


Figure 6.3 *– Left: uncertainty on the zenith angle as a function of the multiplicity. Right: uncertainty on the azimuth angle as a function of the multiplicity. The uncertainty is calculated with the equation 6.11.*

where R_t is the time needed to go from 10 to 70 percent of the normalized cumulative function. A strong correlation between the two defined points of the cumulative has been demonstrated in [198, 62], and the rejection criterion is based on a χ^2-like test, as described in [62].

The criterion on the cumulative function previously described has been updated. As illustrated by the figure 6.5, signals featuring several distinct pulses are not ideally considered by this narrow window since only one (and generally the first) peak is taken into account. The new criterion is based on the entire time window of the recorded waveform, i.e the [2.56] µs, permitting to take into account all the transients present in the waveform (see figure 6.4). The two defined points of the cumulative which have been

used in [62] have also been updated after testing on a batch of events, and the rise time is now calculated between 25 % and 65 % of the cumulative function, and events with a rise time greater than 20 ns are rejected. This value is conservative, and would require further studies to be finely tuned, but the high rejection efficiency of the present RTA is satisfactory: with this updated RTA, 99.77 % of the fortuitous events are rejected, and no cosmic ray event is eliminated.

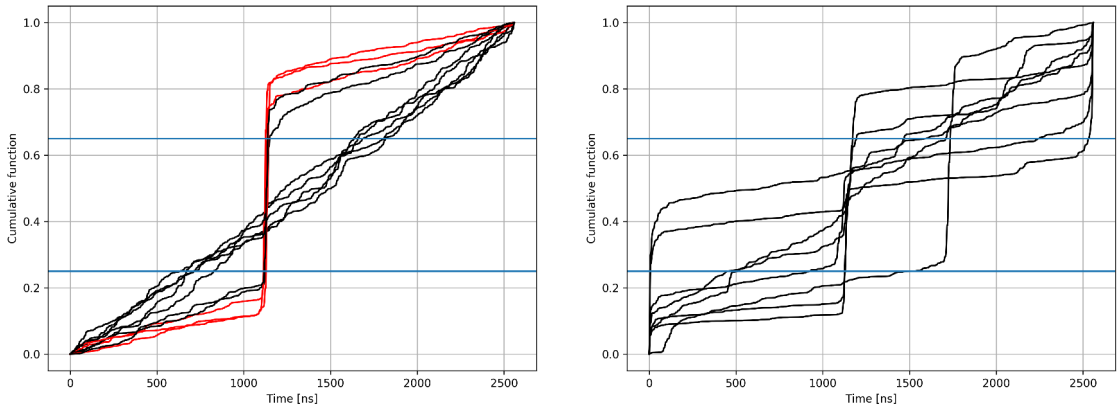


Figure 6.4 *– Left: rise times for a shower event, coming from East (ϕ = 203 °). Right: rise times for a fortuitous event. The horizontal lines stand for* 25 % *and* 65 % *of the cumulative function. Red curves correspond to rise times lower than* 20 ns *and black curves correspond to rise times greater than* 20 ns. *For the shower event, we see that several antennas have a rise time lower than* 20 ns, *while for the fortuitous event, no antenna passes the selection. Moreover, the fortuitous event presents several distinct pulses in the traces, corresponding to the different steps in the cumulative function. For the shower event (left), the cumulative curves with a fairly linear trend correspond to the rise times of the EW polarizations which have no signal since the event is coming from the East.*

Applying the updated RTA, only 114 events over 50,096 remain, including the event tagged as an actual shower. Figure 6.5 left presents the directions of arrival of the remaining events after the application of the RTA. The two plane trajectories are still visible. The RTA is efficient to reject most of the events, namely related to stationary sources. This can be explained by the waveforms associated to this kind of sources, which present multiple transients, thus leading to a large rise time, as shown in figure 6.5 right. By way of comparison, figure 6.22.(c) presents transients related to a cosmic ray event.

6.2.2. Compacity Algorithm

For this algorithm, the idea is to have a criterion on the number of standalone antenna involved in the event with respect to the number which should have been involved. A shower event presents a regular pattern at ground, meaning that all the standalone antennas around the shower core should be involved. In other words, there should be no missing standalone antennas in the pattern at ground close to the shower axis. This is not the case for a fortuitous event, where the pattern at ground can be irregular, with missing standalone antennas: the high repetition rate of most noise events lead to high, but not necessarily synchronous occupation of the SA. To obtain a criterion on that behaviour, we calculate the convex hull of the event [199]. It consists in finding the smallest convex hull which contains all the involved standalone antennas. We then calculate the number of involved standalone antennas present in the convex hull and compare it with the number of antennas that actually exists in the convex hull. Figure 6.6 presents the convex hull of

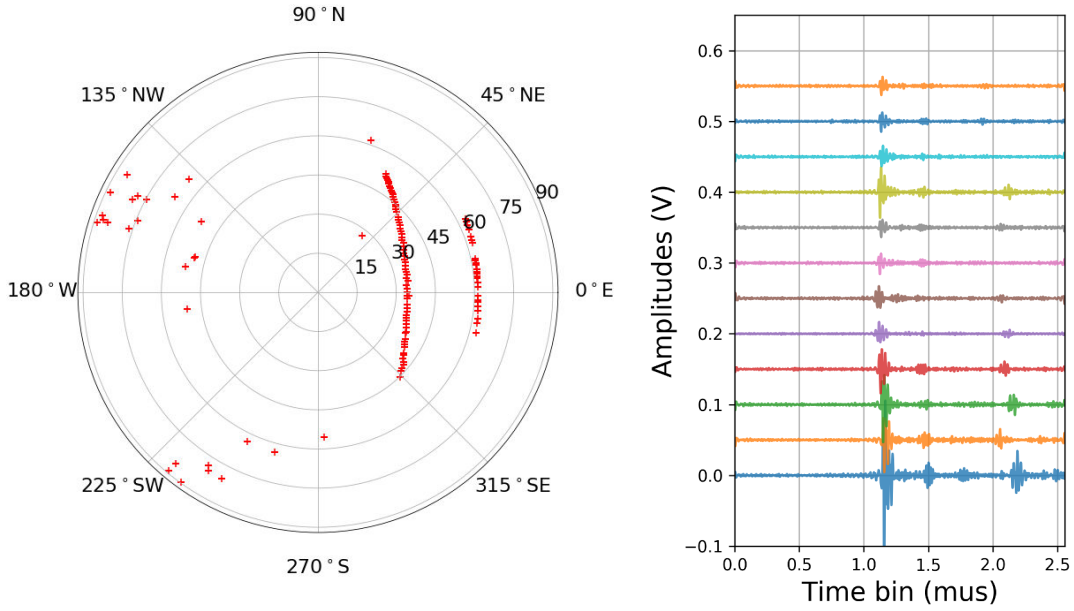


Figure 6.5 *– Left: skymap of the remaining events after application of the Rise Time Algorithm. Two plane trajectories are visible. Right: example of waveforms associated to a stationary source, presenting multiple transients and thus leading to a large rise time.*

two events: one is a shower event (left), the other one a fortuitous event (right). Most of the time, the behaviors on a shower event or a noise event are different.

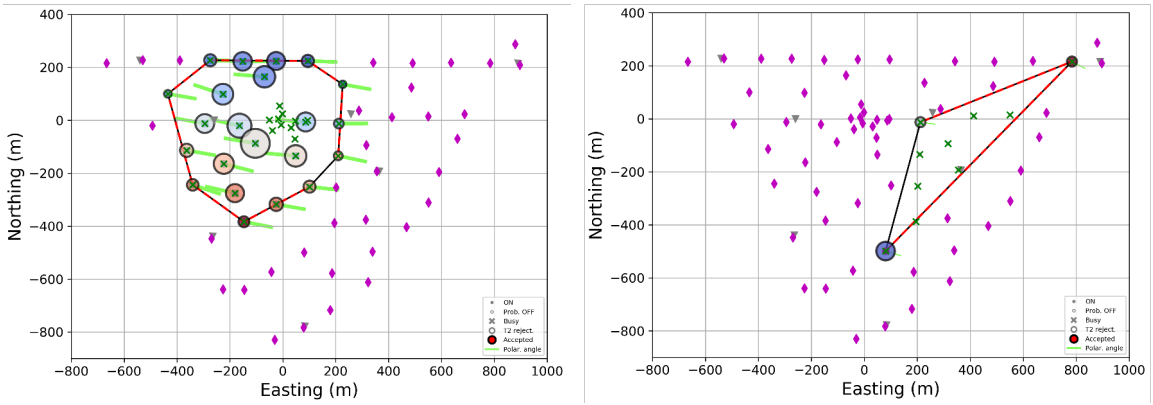


Figure 6.6 *– Left: convex hull for a shower event. Right: convex hull for a fortuitous event. The violet diamonds represent the standalone antennas, the involved standalone antennas are represented by coloured circles, whose colour indicates the order in which the signal has been seen by the antennas (from blue, earliest, to red, latest) and area of circles reflects the relative amplitude of the signal. The convex hull is delimited by the red lines. The standalone antennas inside the convex hull are represented by green crosses. All the standalone antennas inside the convex hull are present on the event on the left figure (shower event), while some of them are missing on the event on the right figure (fortuitous event).*

On this basis, the criterion used to select or reject an event is expressed as follows:

$$C = \frac{\mathrm{SA_{trig}}}{\mathrm{SA_{in}}} \tag{6.15}$$

where $\mathrm{SA_{trig}}$ is the number of triggered standalone antennas, $\mathrm{SA_{in}}$ is the number of all standalone antennas inside the convex hull and A_{CH} is the area of the convex hull. Then,

we tag the event as a good event if C is greater or equal to 70 % of the convex hull area (i.e. 70 % of the antennas in the convex hull are present in the event). The coefficient 70 % has been chosen to be conservative, and further studies are needed to optimize the algorithm. Applying CtA on the full data set of 50,096 events, 28,814 events remain, meaning that 42.4 % of events have been rejected. Figure 6.7 presents the direction of arrivals of the remaining events after the application of the CtA. The idea is simply to illustrate the performance of the algorithm when applied alone, a further study on the percentage of antennas to be present in the convex hull should be carried out to obtain a better rejection efficiency. Currently, this criterion is less efficient that the RTA for instance, but it makes it possible to consolidate the batch of exploitable data: over the 114 events remaining after the application of the RTA, 21 more events are removed. After the RTA plus the CtA, only 93 events remain.

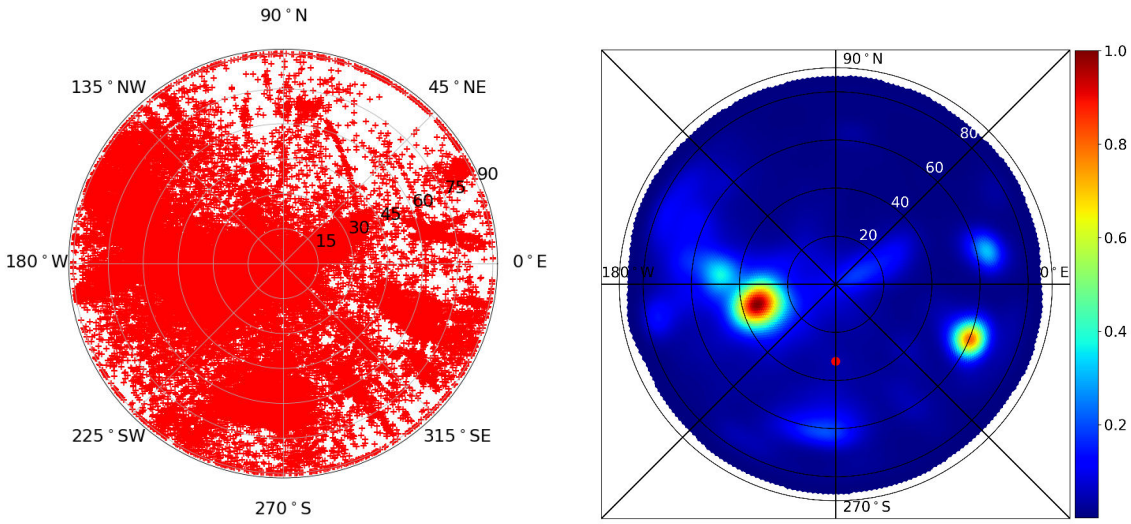


Figure 6.7 *– Left: Skymap of the remaining events after applying the Compacity Algorithm on the initial batch of 50,096 events. 28,814 events remain, rejecting* 42.4 % *of the initial batch. Right: same as left,* 10° *gaussian smoothed sky map.*

6.2.3. Polarization Algorithm

The total electric field is the summation of the electric fields produced by the geomagnetic and charge excess mechanisms. Since the main contribution to the electric field comes from the geomagnetic mechanism detailed in section 3.2.1, the resulting electric field is expected to be linearly polarized and aligned along $\vec{n} \times \vec{B}$ for most of the detected events ($\vec{n}$ is the shower axis vector). The Polarization Algorithm (PA) aims to exploit this expectation, namely to compare the predicted polarization angle and the detected polarization angle. For each event, we calculate the predicted polarization angle as follows:

$$\mathrm{pol}_{\mathrm{th}} = \arctan2 \frac{\left(\vec{n} \times \vec{B}\right)_{\mathrm{NS}}}{\left(\vec{n} \times \vec{B}\right)_{\mathrm{EW}}} \tag{6.16}$$

where arctan2 is the arc-tangent choosing the quadrant correctly so that $\arctan2(x_1, x_2)$ is the angle signed between the x-axis and the radius ending at the origin and passing through the point (x_2, x_1), $\vec{n}$ is the shower axis vector determined from the reconstructed direction of arrival and $\vec{B}$ is the geomagnetic field vector for the site of the experiment.

Then, for each antenna i, we calculate the polarization angle as follows:

$$\text{pol}_i = \arctan2 \frac{E^i_{\text{NS}}}{E^i_{\text{EW}}} \tag{6.17}$$

where E_{NS} and E_{EW} are amplitude of the electric field of the standalone antenna i in the NS and EW polarizations respectively taken when the maximum of the signal (either of the EW or NS polarization) is reached. Example of such values is shown in figure 6.23.(b). We calculate an average behavior of the polarization angles for the antennas of the event by taking the median value of the polarization angles (permit to take into account antennas with a problem on a polarization that would distort the measurement of the polarization angle) and compare it to the predicted polarization angle: the events are selected if the deviation from the predicted polarization angle is less than 15° (set after testing on a batch of events). After the PA, 19,467 events remain from the original data set (61.2 % of events have been rejected).

Key point: The PA aims to exploit the polarization pattern expected for the geomagnetic mechanism, namely to compare the predicted polarization angle and the detected polarization angle. Thus, this algorithm is reliable for events maximizing the geomagnetic mechanism, and can not be applied on events having a significant contribution coming from the charge excess mechanism. It is thus important to define an area of the sky on which we will not apply this criterion. This zone of the sky corresponds to the directions of arrival close to the direction geomagnetic field at CODALEMA, which minimizes the geomagnetic effect and maximizes the charge excess mechanism. Based on figure 6.8, the zone where we will not apply the PA is defined as: $\theta \in [20, 40]$ ° and $\phi \in [225, 315]$ °, delimiting a fairly conservative zone around the geomagnetic field vector direction.

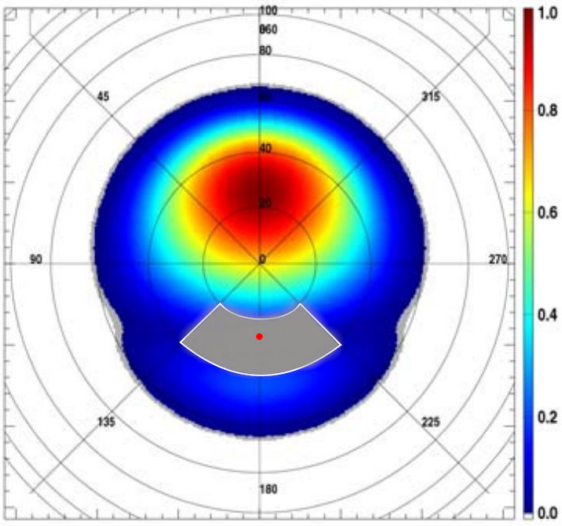


Figure 6.8 *– Sky map calculated by considering the EW component of the Lorentz force multiplied by the trigger coverage map. The color scale is normalized to 1 in the direction of the maximum. The exclusion zone for the application of the RtA is delimited by the white area and colored in grey. Adapted from [83].*

The method based on the arctan2 has a main limitation, which is to use only the maximum of waveforms of the electric field for the calculation. Another method can be used: the orthogonal distance regression (ODR). This method is supposed to be more accurate than the calculation with the arctan2, because the ODR is applied on the entire ellipse of polarization, i.e on the full waveforms of the electric field. Let us introduce

the mathematical framework of the ODR. We assume that there is an explicit linear relationship between the EW and NS polarizations, that we note f, and which can be expressed as follows:

$$\text{NS} \backsimeq f(\text{EW}; \beta) \tag{6.18}$$

where EW and NS are the electric field data in EW and NS polarizations respectively, β is the list of parameters to be determined. In our case, both electric fields in the EW and NS polarizations are detected with some errors, permitting to determine β by ODR [200]. Assuming (x_i, y_i) our observed set of data, with i from 0 to n =2,559 the number of samples in the waveforms, observed with error δ_i and ϵ_i respectively, we can write:

$$y_i = f_i(x_i + \delta_i \beta) - \epsilon_i \tag{6.19}$$

The ODR aims to find the parameters β by minimizing the sum of the squares of the n orthogonal distances from the linear curve $f(x; \beta)$ to the n data points. This can be written:

$$\min_{\beta,\delta,\epsilon} \sum_{i}^{n} (\epsilon_i^2 + \delta_i^2) \tag{6.20}$$

with the constraints:

$$y_i = f_i(x_i + \delta_i \beta) - \epsilon_i \tag{6.21}$$

The constraints on ϵ_i are linear, we can remove them from equations 6.20 and 6.21, leading to:

$$\min_{\beta,\delta} \sum_{i}^{n} \left(\omega_{\epsilon_i} [f_i(x_i + \delta_i \beta) - \epsilon_i]^2 + \omega_{\delta_i} \delta_i^2 \right) \tag{6.22}$$

where ω_{ϵ_i} and ω_{δ_i} are the weights on ϵ_i and δ_i respectively. They are used to give more impact to the large electric field values in the waveforms. After applying the PA with the ODR, 21,294 events remain (57.5 % of events have been rejected), whose the direction of arrivals are largely the same as for the first method, as shown in figure 6.9. Both methods give approximately the same percentage of rejected events. It is important to note that the event tagged as a cosmic ray event is not rejected by both methods.

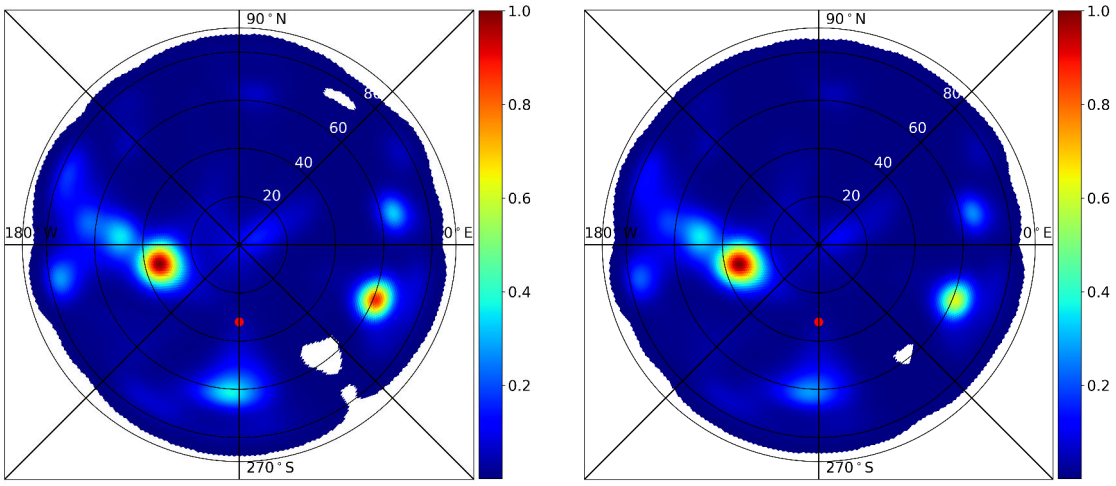


Figure 6.9 – 10° *gaussian smoothed sky maps of the remaining events after the Polarization Algorithm processed with the arctan method (left) and the ODR method (right).*

6.2.4. Order of application of the algorithms

The combination RTA plus CtA plus PA in that order gives the best result for the noise rejection, since only 28 events remain, whose the event tagged as an actual shower, corresponding to 0.05 % of the initial batch of events (see figure 6.10). The RTA must be applied first, since it is efficient to reject most of the events related to stationary sources. Then, for the remaining event, we would like first to test if the pattern at ground is characteristic of a cosmic ray event, in order to consolidate the set of remaining event after the application of the RTA. The PA is then applied last.

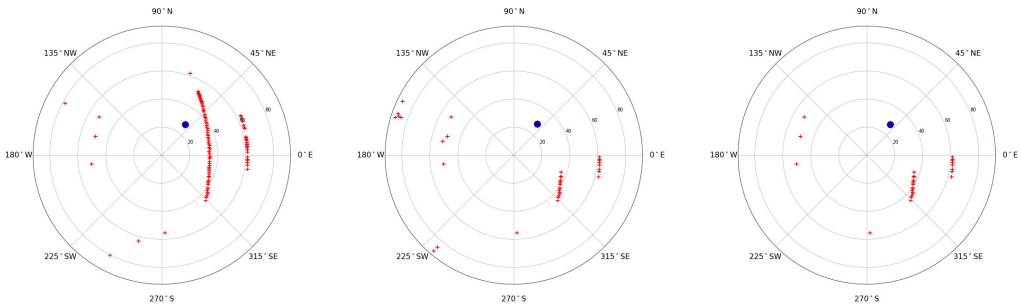


Figure 6.10 *– Direction of arrivals of the remaining events after applying the noise rejection algorithms in that order: RTA, CtA and PA. The actual cosmic ray event is represented by the blue point.*

6.2.5. Noise rejection algorithms performances

After testing the three noise rejection algorithms on a small set of data (one day), we illustrate in this section their performance on a large set of data (one month), corresponding to 1,704,838 events which have been selected by the AAS and cone algorithm at the central DAQ level. On this period, 58 events have been identified as cosmic ray events, tagged with the particle detectors. Figure 6.11 presents the remaining events after the application of the noise rejection algorithms. Only 701 events remain, corresponding to 0.04 % of the initial batch of events. Additional algorithms are considered in order to further reduce the number of parasitic transients from the overall set. For instance, we see in figure 6.11 that most of the remaining events seems to be aircraft events, which could be removed with an algorithm looking for the temporal and angular environment of two consecutive events. We have checked that all the detected cosmic ray events are present after the application of the noise rejection algorithms, corresponding roughly to 1.7 events per day. In the energy range of CODALEMA ($[10^{17} - 10^{18}]$ eV), $\sim$ 5 events per day are expected, which leads to an estimation of the efficiency for this batch of events equal to about 33 %. For one day, around $\sim$ 19 events are from an unknown origin (parasitic transients, cosmic ray events without counterpart in the scintillators). Farther works are needed, namely to look carefully at these $\sim$ 19 events per day to estimate if they are parasitic transients or cosmic ray events without counterpart in the scintillators, and in the latter case, the detection efficiency would be improved.

6.3. Conclusion on the selection method of the radio events

Most of the events detected by CODALEMA are anthropic, static and almost stationary (coming from sources giving accumulations in their directions of arrival). Several rejection algorithms are used online in the CODALEMA central DAQ to significantly suppress

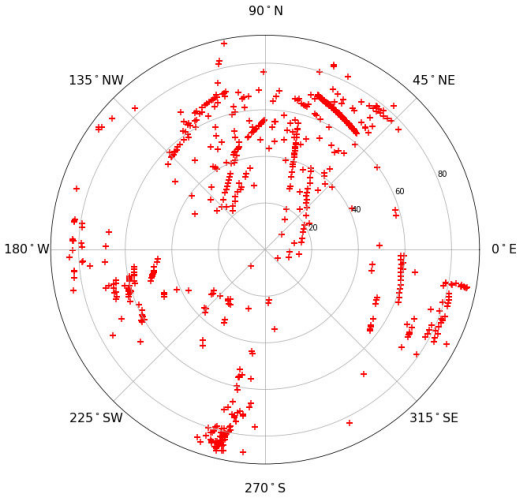


Figure 6.11 *– Skymap of the remaining events after applying noise rejection algorithms on one month data.*

the parasitic events. However, a lot of parasitic noises remain in the recorded events since the rejection algorithm rate was set to reach sustainable number of events written on disk. Different methods have been developped and tested offline to improve the selection efficiency. Figure 6.12 presents the general structure of the online and offline rejection algorithms. The number of events, corresponding to one month of data taking, is indicated in red for each step.

The offline algorithms were not implemented online since the current CODALEMA duty cycle is mostly constrained by the T1 analog level efficiency and over simplicity. Indeed, as previously seen, the bottleneck for the efficiency of the radio method at CODALEMA is the electronics used, inducing a dead time of 30 ms, corresponding to an acquisition rate of $\sim$ 30 Hz. For the two periods of acquisition of figure 6.1 for which the average rate of parasitic transients is estimated to $\sim$ 0.5 Hz and $\sim$ 1.7 Hz, an acquisition rate of $\sim$ 30 Hz should be sufficient for cosmic ray observation. In quiet conditions (no anthropic noises), no cosmic ray events should be missing. Real duty cycle is difficult to determine, since largely dependent on external noises.

Even if noise rejection algorithms have been explored, only the improvement of the selectivity of the radio event is possible. A serious change in trigger electronics would be necessary to implement all or part of the offline rejection algorithms at the T1 level, which would be necessary to increase the duty cycle. Still offline currently, a much more thorough study of these algorithms is necessary to ensure the independence of the radio method, namely to be able to get rid of another type of detectors such as a particle detector to ensure the origin of the event (cosmic ray event, parasitic event). At that time, due to lack of time, further studies could not be conducted, and like other experiments such as LOFAR and AERA, the use of the information of the particle detectors is mandatory to ensure that the recorded radio events are from cosmic ray origin.

Moreover, we have already mentioned in chapter 3 in section 3.2.4.1 that the atmospheric electric field makes the radio signal non exploitable for the reconstruction of the primary cosmic ray characteristics. We will see in chapter 8 that the atmospheric field plays a major role in the detection of LF events, for instance. Thus, a rejection algorithm based on the atmospheric electric field has also been explored, and can be associated with the polarization algorithm discussed above. Indeed, one of the characteristics of an event detected during period of abnormal atmospheric electric field conditions is

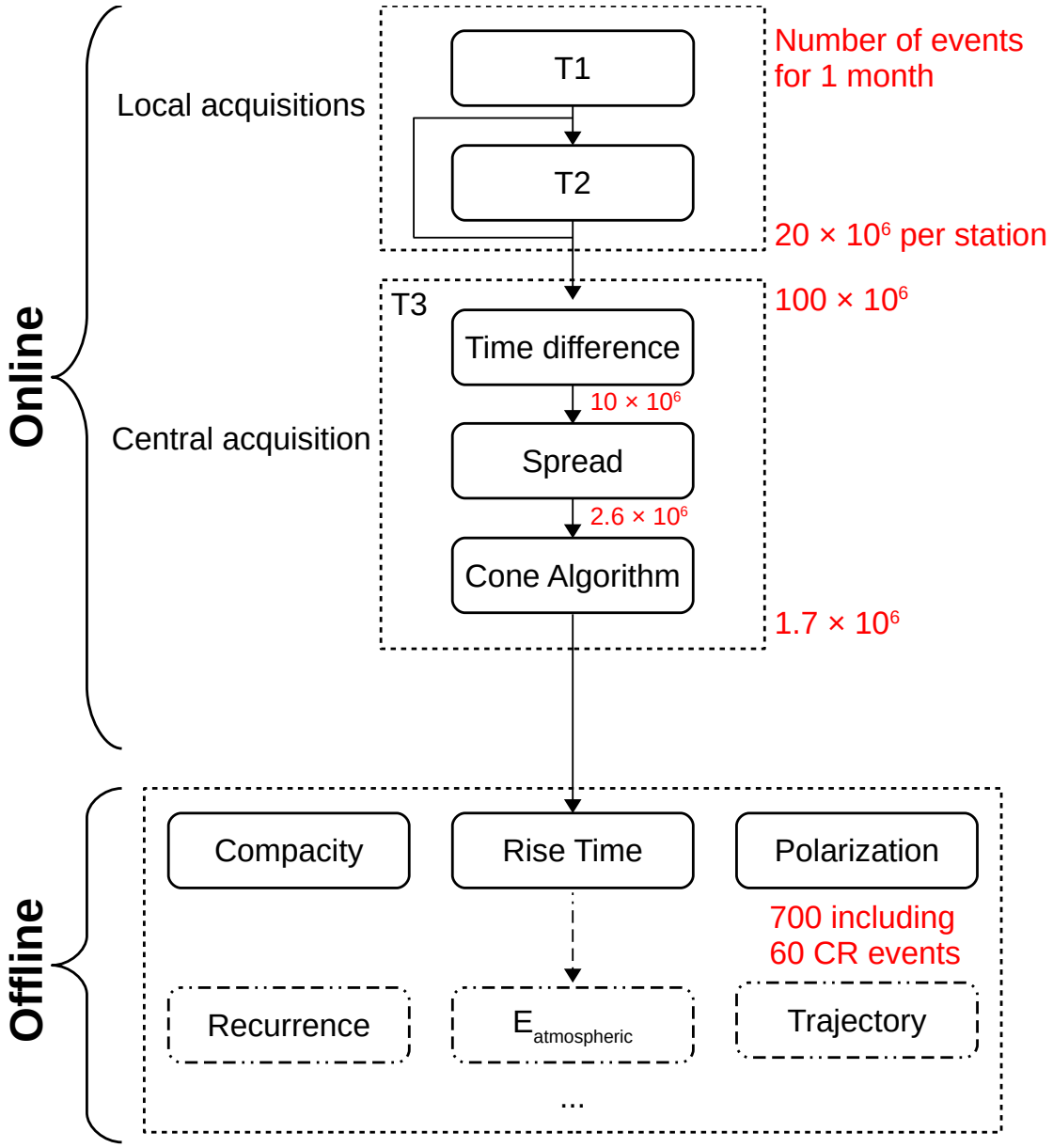


Figure 6.12 *– General structure of the online and offline rejection algorithms. The number of events, corresponding to one month of data taking, is indicated in red for each step.*

the orientation of the polarization, which tends to move away from the orientation predicted by the geomagnetic mechanism, as shown in figure 6.13. It presents the skymap of the arrival directions of the CODALEMA events coupled with the information of the polarization orientation of the event (the arrows denote the polarization angle of the event). We expect for the event produced mainly by the geomagnetic mechanisms an arrow orthogonal to the direction given by the direction of the event (θ, ϕ) and the center of the plot (the zenith), which is the case for most of the black arrows.

Red arrows are for events with an abnormal atmospheric electric field: $|E_{\text{atm}}| \geq 0.5\,\text{kV}\cdot\text{m}^{-1}$ (figure 6.13 left) and $|E_{\text{atm}}| \geq 10\,\text{kV}\cdot\text{m}^{-1}$ (figure 6.13 right). This limit has been empirically determined on the basis of the batch of events presented here. We observed for abnormal atmospheric electric field values that the polarization angle move away from the orientation predicted by the geomagnetic mechanism. This is the reason why a cut on the value of the atmospheric electric field value is done in chapter 4 for the

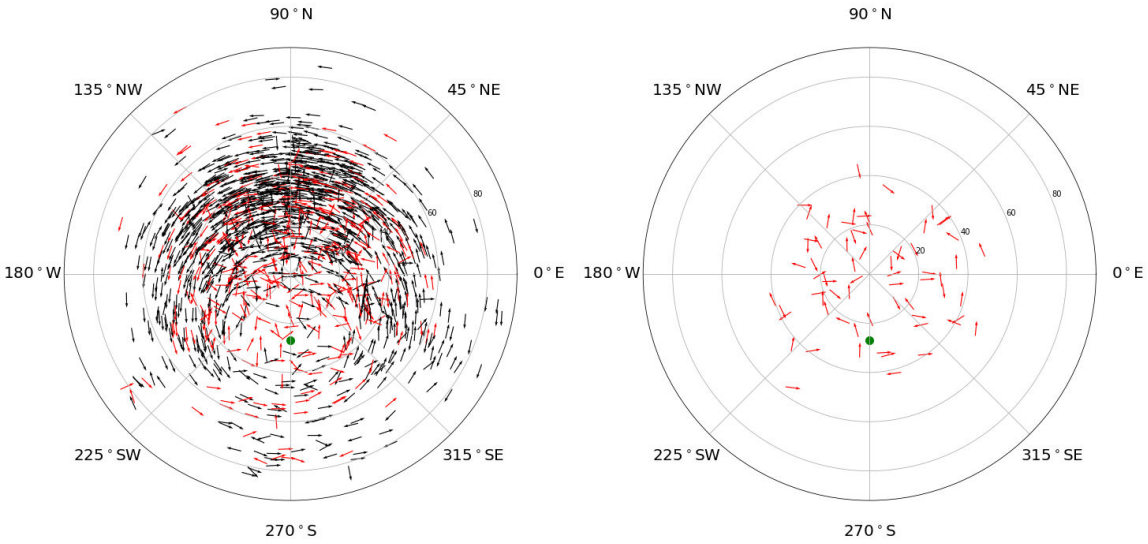


Figure 6.13 *– Skymap of the arrival directions of the CODALEMA events coupled with the information of the polarization orientation of the events. For the figure on the left, red arrows are for events with an abnormal atmospheric electric field* ($|E_{\text{atm}} \geq 0.5\,\text{kV}\cdot\text{m}^{-1}|$)*, while for the figure on the right, red arrows are for events with* $|E_{\text{atm}} \geq 10\,\text{kV}\cdot\text{m}^{-1}|$*. See text for comments.*

selection of the batch of events for the study with the three-fold antenna. Moreover, in left figure, we see that some red arrows, tagged as events with an abnormal atmospheric electric field values, present a polarization angle consistent with the prediction of the geomagnetic mechanism. However, even if the the polarization angle is consistent with the prediction, the abnormal atmospheric electric field value will enhance the radio signal, leading to an overestimated energy of the primary cosmic ray. For the batch of events used to make the figure 6.13 (1,573 events), 436 events have an $|E_{\text{atm}}| \geq 0.5\,\text{kV}\cdot\text{m}^{-1}$, of whom 182 have been rejected with the polarization rejection algorithm (PA).

At this point, some questions remain:

— Should we include in the rejection algorithms the value of the atmospheric electric field? If so, how to define what is an abnormal electric field? We will try to answer this question in the chapter 8.

— How can we clean up even more the skymap of the figure 6.11? One possibility would be to look at the temporal environment of the remaining events, for example to check the trigger rate in the minutes preceding the event. If this rate is high, then the event in question is probably the same type of events from the previous minute. It would therefore be a fortuitous event.

— Once the skymap is cleansed of the remaining spurious events, how do the remaining events look like? Do they correspond to the criteria of a cosmic event? If so, is it a cosmic event without scintillator counterpart? In this case, it should be verified that the particle detector was in use. If the particle detectors were in operation, we should estimate the position of the shower core of the radio event: if the shower core of the radio event is far away from the particle detector array, it should be an actual cosmic ray event scintillator counterpart ; otherwise the event should have been detected by the scintillators, and in that case it is necessary to investigate more to make sure of the radio event in question.

After the application of the algorithms presented above to the batch of events recorded

by CODALEMA between July 2016 and September 2018, the remaining dataset appears unbiased. Indeed, figure 6.14 presents the zenithal and azimuthal distributions of these events. As for the scintillators, the zenithal distribution is fitted with the empirical function given in [83]: $\mathrm{d}N/\mathrm{d}\theta = (a + b\theta)\cos(\theta)\sin(\theta)/(1 + \exp((\theta - \theta_0)/\theta_1))$. We obtain $a = 33.6$, $b = 4.77$, $\theta_0 = 55.1\,°$ and $\theta_1 = 5.20\,°$. As expected, the azimuthal distribution presents a North/South asymmetry (see figure 3.3 and [83]) directly attributed to the radio emission mechanisms.

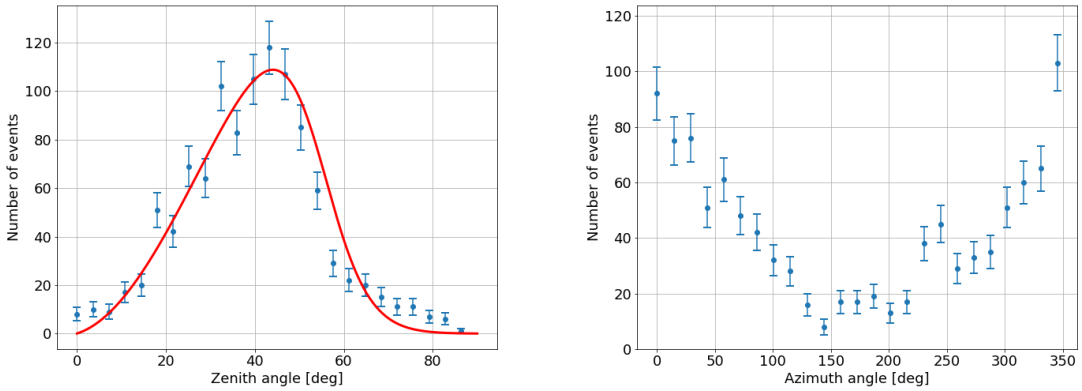
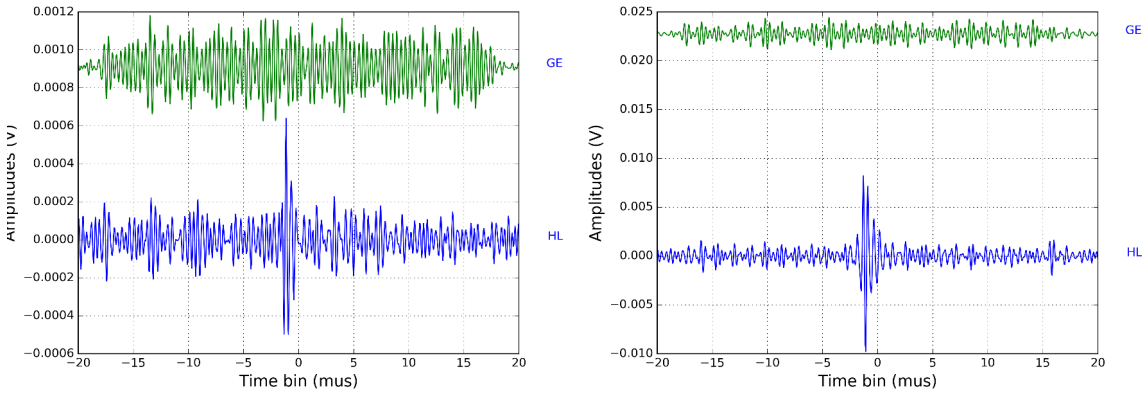


Figure 6.14 *– Zenith (left) and azimuth (right) angular distributions for the radio events. For the azimuth distribution, the North corresponds to* 0° *and* 360° *and the West to* 90°*. The distribution is in agreement with the asymmetry shown in figure 3.3.*

6.4. EXTASIS – Finding a low-frequency pulse

Parts of this section has been published in: . . ., **A. Escudie et al.**, *Radio detection of cosmic rays below* 1.7 – 3.7 MHz*: the EXTASIS experiment*, Astroparticle Physics [9].

The low-frequency antennas are externally triggered by the particle detectors. The signal and the noise environment in which it is sought is completely new leading to the need of development of analysis tools to detect it. In the frequency range below 10 MHz, the ambient noise is important, and the shower transients are expected to be often buried in the noise. Signal cleaning requires a high-performance signal processing method. Three detection methods have been developed and tested: simple threshold (minimum bias method), linear prediction coding (LPC) [201] and a combination of wavelet analysis [202] and neural networks. They have been compared in order to select the most efficient one. In figure 6.15 is shown a low-frequency event in two antennas (GE and HL), one having a transient in the two polarizations horizontal and vertical (see figure 4.1 for their locations) and the other one not having a transient. We used this event to illustrate the three methods.



(a) *EW polarization.* ***(b)*** *Vertical polarization.*

Figure 6.15 *– Event in GE (green) and HL (blue) antennas in the EW (left) and vertical (rgiht) polarizations. The x-axis represent time bin in* µs a*nd the y-axis in the recorded voltage. The green curve is positively shifted along* y *for clarity.*

6.4.1. Method

6.4.2. The amplitude threshold

The first and basic method is to use a simple threshold set on each filtered signal, above which a radio transient could be detected. Moreover, we know that the expected transient has to be in the window $[-5; 5]$ µs within the LF radio trace, because of the geometry of our antenna array and because the trigger position is set to at 0 µs. We search the maximum of the signal in this window and find its position. Once again, we define a window $[-1.5; 1.5]$ µs around this maximum, and calculate the standard deviation of the signal in this window, that we compare to the standard deviation of the signal performed on a window define as $[-40; -20]$ µs before the trigger, in which only noise is expected. Then, we apply a threshold as 2.7 times the standard deviation of the noise (this threshold was defined after different tests on the global batch of events, where it maximizes the efficiency of the method). With this condition, the transient in the HL antenna is detected on both polarizations. This condition permit to not detect noise on the GE antenna which does not have transient.

6.4.3. The LPC method: Linear Prediction Coding

The LPC method is a linearly predictive method, and is principally used in speech coding, speech synthesis, speech recognition, but also in neurophysics [201], in geophysics [203]. The LPC was first proposed in 1984 to encode human speech. The LPC tries to predict future values of the input signal using past values. When the LPC tries to predict future signal values, a prediction error is calculated. This error is obviously weak when the LPC correctly predicts the future signal, and is obviously large when the LPC is unable to predict the future signal. In science for example, it is used to detect and suppress unexpected peaks in recorded signals, peaks that can not be predicted by the LPC. In our case, we used the LPC method the other way, i.e to detect a transient in our experimental signal. A review of the complete mathematical description can be found in [204]. In these methods, the sample n can be defined as a linear combination of the $n-1$ past samples:

$$s_p(n) = \sum_{k=1}^{n} a_k s(n-k) \tag{6.23}$$

where a_k are the predictor coefficients. We call prediction error the difference between the signal $s(n)$ and the predicted signal $s_p(n)$. In the present case of a search for a transient signal in a noise composed of the contribution of several transmitters and the atmospheric noise, the prediction error represents well the expected transient signal. The predictor coefficients are determined by minimizing the sum of squared differences between the true samples and the linearly predicted samples. The predictor coefficients a_k can be calculated by solving the n equations with n unknowns (giving a hermitian Toeplitz matrix [1]). For more information on the determination of the predictor coefficients, please refer to [204]. These resolutions lead to $n^3/3 + O(n^2)$ operations and n^2 storage locations. The resolution can be made with a Levinson-Durbin algorithm [205, 206, 207].

In figure 6.16, is shown the prediction error of the LPC method on the event of the figure 6.15 for the two antennas. The red line corresponds to the definition of the threshold set as $\max(\text{se}) \geqslant \mu_{\text{se}} + K\sigma_{\text{se}}$, where μ_{se} and σ_{se} are respectively the mean of the squared prediction error (se) and the standard deviation of the squared prediction error. K is a factor empirically defined as 14 [208]. Parameter K is determined according to the percentage of false detection, corresponding to few percents for $K = 14$. We can easily see that the transient in HL antenna is mispredicted. Consequently, the maximum of the squared prediction error is higher than the threshold, and the transient is thus detected. We also see that the procedure does not find any transient in the GE antenna, since the GE antenna does not present a signal. We note that there is a 5 orders of magnitude difference between the squared prediction error on GE which has no transient and the squared prediction error on HL which has one.

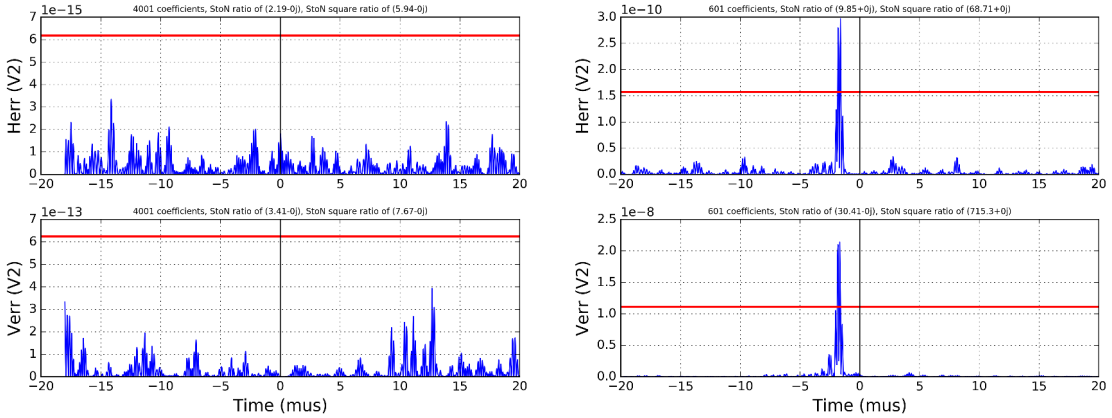
6.4.4. The wavelet and neural network combination

6.4.4.1. The wavelet analysis

The wavelet transform can be used to process time series containing nonstationary power at different frequencies. However, one must have a time series, *x(t)*, with an equal time spacing and a mother wavelet, $\Psi(\eta)$, which depends on a nondimensional time parameter η. The method can be categorized as discrete (DWT) or continuous (CWT) wavelet transform. The CWT performed a continuous translation, and is used mainly for pattern matching for discontinuity and transient detection, and the best example to give is the observation of gravitational wave the 12 February 2016 by the LIGO Scientific Collaboration and Virgo Collaboration [209]. The mother wavelet must:

— have zero mean, and so be oscillating

1. The Toeplitz matrix is a matrix in which each diagonal value is constant, and in that case, symmetric and positive.



(a) *LPC on GE antenna.* ***(b)*** *LPC on HL antenna.*

Figure 6.16 *– LPC method on the event of the figure 6.15. Here is shown the prediction error of the LPC method for the two antennas. The red line corresponds to the definition of the threshold. See text for more details.*

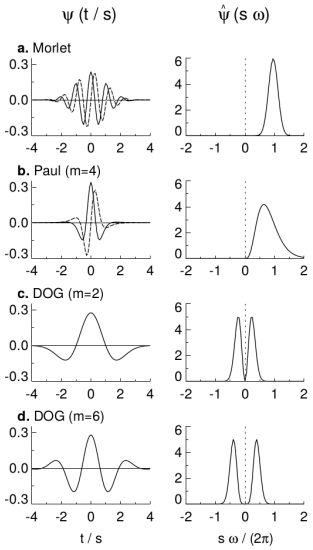
— be localized in both time and frequency domain
— have zero integral
— be square integrable
— may be complex valued

Let $x(t)$ be a real function of a real variable . The wavelet transform of x is:

$$g(a,b) = \frac{1}{\sqrt{a}} \int_{t=-\infty}^{t=\infty} x(t)\overline{\psi}_{a,b}(t)\,\mathrm{d}t \tag{6.24}$$

where the function $\psi_{a,b}(t)$ is obtained by translating the mother wavelet $\Psi(\frac{t-b}{a})$ where b determines the position and a gives the scale. It either dilates or compresses the signal. The factor $\sqrt{a}$ is for energy normalization across the different scales. In our case, a is the frequency and b the time. One can interpret the equation 6.24 as an inner product measuring the similarity between the signal $x(t)$ and the deformed mother wavelet. Figure 6.17 presents different wavelet bases, that can be used as mother wavelets. We decided to use for our study the Mexican Hat wavelet as mother wavelet (i.e the wavelet of figure 6.17 left, letter *c*), which presents basic features of a peak, a symmetry and one important positive peak [202, 210, 211].

Initially, the wavelet method was combined with a neural network to pick out wavelet pictures by pattern recognition. However, the procedure was complicated and the training of the neural network was not so easy. Currently, after testing, we performed the CWT at 11 scales on the filtered signal. As for the threshold method, a threshold is defined on these coefficients and is set to 3.5 times the standard deviation of the coefficients in a window where no signal is expected. In figure 6.18, are shown the results of the wavelet method applied to the event of figure 6.15 for the two antennas, i.e the 2D plot of the CWT of the time series. The x-axis represents the bin numbers. The y-axis represents the width of the window used for the wavelet transform (also corresponding to the scale which is frequency-related) from small (top of figure) to large (bottom of figure). The colored plot represents the CWT coefficient image. Green color represents low amplitude, red high. The plot in the upper right corner represents the best CWT coefficient scale. The transient is detected for HL antenna for the event of the figure 6.15, and no signal is detected for the GE antenna.



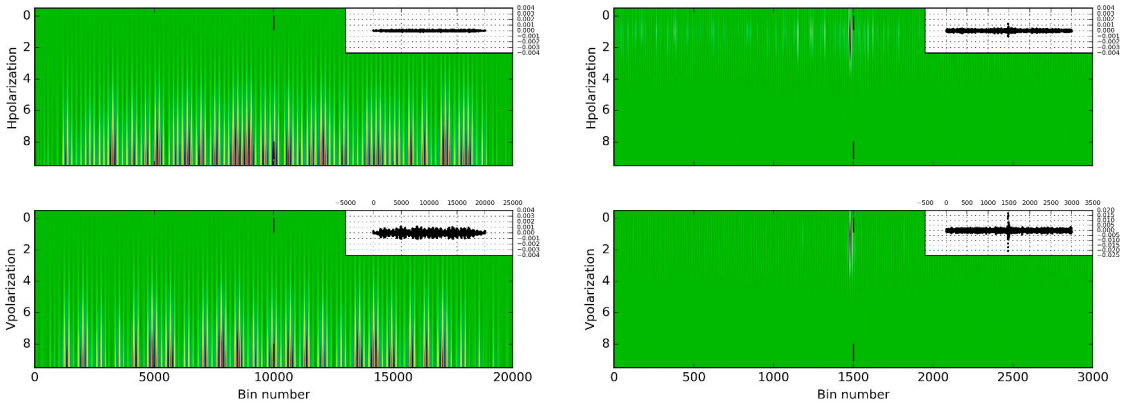
(a)

Name	$\psi_0(\eta)$	$\hat{\psi}_0(s\omega)$	e-folding time τ_s	Fourier wavelength λ
Morlet (ω_0 = frequency)	$\pi^{-1/4} e^{i\omega_0\eta} e^{-\eta^2/2}$	$\pi^{-1/4} H(\omega) e^{-(s\omega-\omega_0)^2/2}$	$\sqrt{2}s$	$\frac{4\pi s}{\omega_0+\sqrt{2+\omega_0^2}}$
Paul (m = order)	$\frac{2^m i^m m!}{\sqrt{\pi(2m)!}}(1-i\eta)^{-(m+1)}$	$\frac{2^m}{\sqrt{m(2m-1)!}} H(\omega)(s\omega)^m e^{-s\omega}$	$s/\sqrt{2}$	$\frac{4\pi s}{2m+1}$
DOG (m = derivative)	$\frac{(-1)^{m+1}}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} \frac{d^m}{d\eta^m}\left(e^{-\eta^2/2}\right)$	$\frac{-i^m}{\sqrt{\Gamma\left(m+\frac{1}{2}\right)}} (s\omega)^m e^{-(s\omega)^2/2}$	$\sqrt{2}s$	$\frac{2\pi s}{\sqrt{m+\frac{1}{2}}}$

H(ω) = Heaviside step function, H(ω) = 1 if $\omega > 0$, H(ω) = 0 otherwise.
DOG = derivative of a Gaussian; $m = 2$ is the Marr or Mexican hat wavelet.

(b)

Figure 6.17 *– Four different wavelet bases (a), their basis functions and their properties (b) [210]. (a) On the left are the real part (solid) and imaginary part (dashed) for the wavelets in the time domain. On the right are the corresponding wavelets in the frequency domain. For plotting purposes, the scale was chosen to be $s = 10\,\delta t$. (b) Three wavelet basis functions and their properties. DOG is for derivative of Gaussian and H the Heaviside step function.*



(a) *Wavelet on GE antenna.* ***(b)*** *Wavelet on HL antenna.*

Figure 6.18 *– Wavelet method on the event of the figure 6.15. The x-axis represents the bin numbers, and the y-axis the width of the window used for the wavelet transform. The colored plot represents the CWT coefficient image. Green (red) color represents low (high) amplitude. The plot in upper right corner represents the best CWT coefficient scale. Two small vertical black lines are manually added on the plot to indicate where the trigger instant is. We see that a detection is done by the wavelet methods when a thin red pattern appears in the upper part of the plot.*

6.4.5. Comparison of the methods

Our batch of events contains two days of data recorded during winter, corresponding to 2,535 events triggered by the particle detector. To compare the three methods cited above, a fake transient with a known shape and a known position but with a random amplitude has been randomly added to our raw data (probably containing no transients

related to cosmic events). The signals are filtered in the range [1.7 – 3.7] MHz and then, the three methods have been applied to each event. This test is only intended to select the most efficient method, that is why we do the test on a batch of events recorded under the worst conditions of background noise.

We want to analyse the calculation time, the time accuracy of the detected transient and the percentage of detection of each methods. In figure 6.19 is represented the calculation time of each methods (the calculation time presented for the wavelet takes into account only the wavelet analysis, not the neural network processing). One can immediately see that wavelet method is the fastest, with about 0.05 s per calculation. Then come the LPC method with about 0.3 s and then the threshold method with a large distribution centered on about 5 s.

Remark: the threshold method has a larger calculation time compared to the two other methods, while it is the simplest method to implement. This is because the threshold method is coupled with a Time over Threshold method to check the number of transient repetitions exceeding the defined threshold in a given time delay. If the number of repetition is important, it is surely a parasitic signal that we do not want to select.

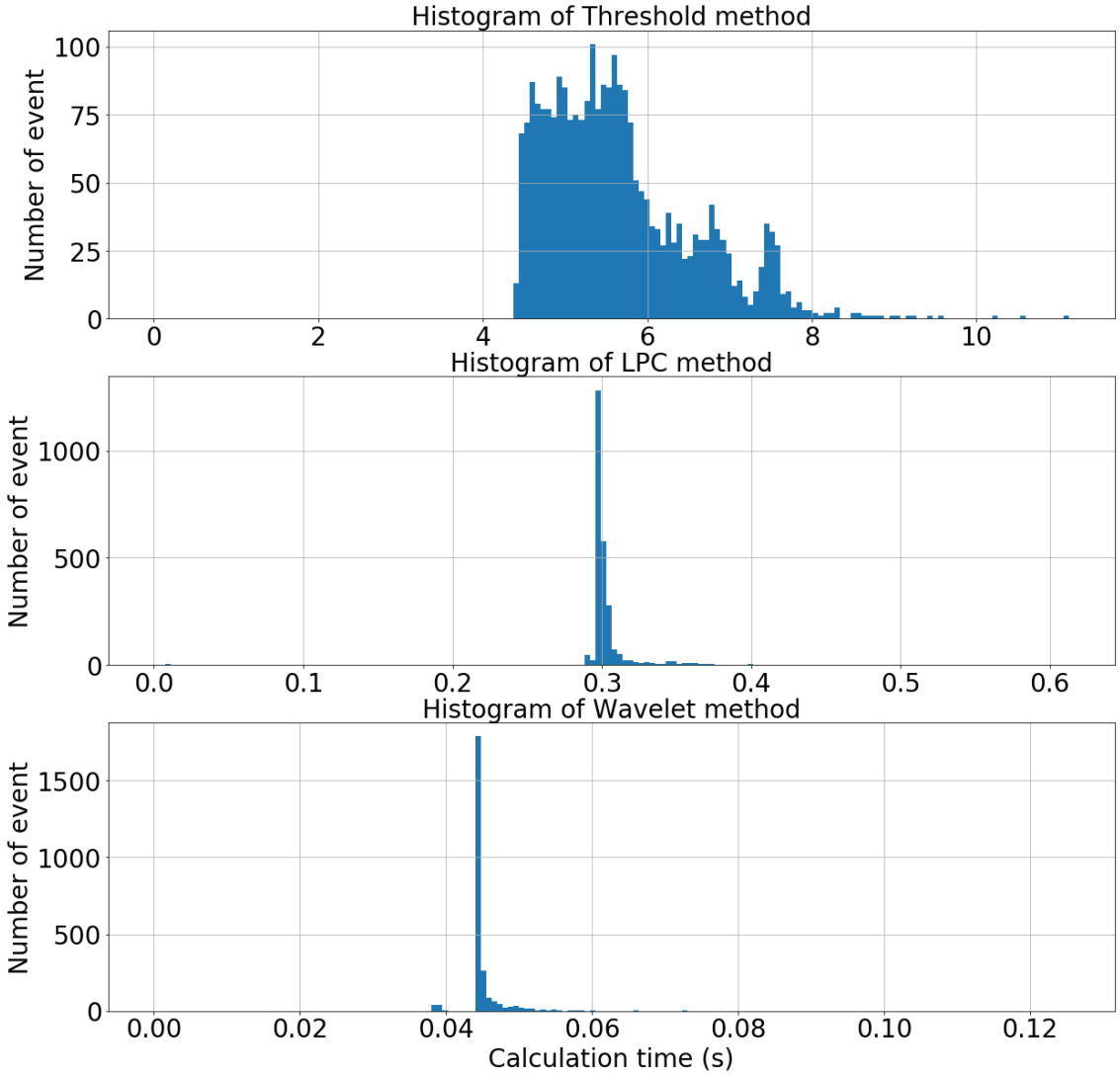


Figure 6.19 *– Calculation time of each methods.*

In figure 6.20 is represented the percentage of detection of the two antennas, for each

method, as a function of the amplitude of the simulated pulse relative to the observed noise peak to peak value. The LPC and wavelet method have the best efficiency at low pulse amplitude compared to the threshold method. The gap decreases for high amplitude pulse, but the LPC and wavelet are still more efficient.

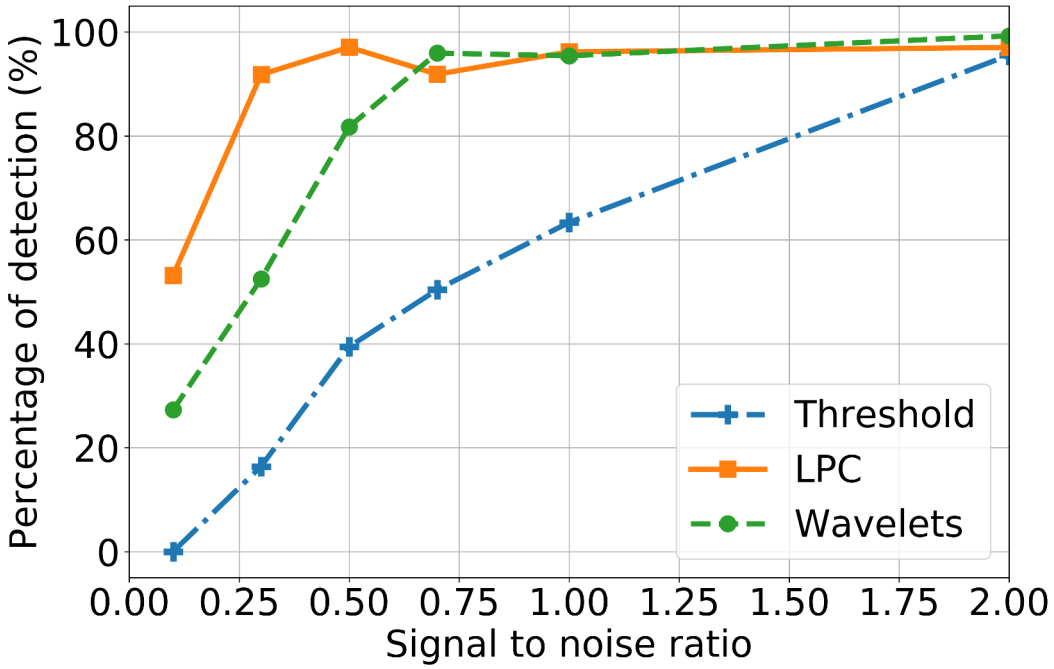


Figure 6.20 *– Detection efficiency of the three tested methods, as a function of the Signal to Noise ratio. The signal is the amplitude of the simulated pulse while the noise is the peak to peak value of the observed noise. Plain line is for the LPC method, dashed line for the threshold method and dotted-dashed line for the wavelet method.*

We have compared the percentage of detection of each methods, but let us now compare the percentage of false detection. The percentage of false detection (when a method detect a transient while there was no addition of transient in the traces) of the threshold method is around 1 %, around 4 % for the LPC method and 10 % for the wavelet method. We could have set the thresholds of the different methods to obtain the same false detection rate and compare the efficiency of real detection (or the opposite). However, we are not seeking here to achieve significant purity of the registered batch. The main objective is to detect the maximum number of events, limiting the number of false detections. The events recorded in LF are then treated in the general analysis pipeline, which will apply cuts using the MF signal, and thus eliminate false detections.

Let us now compare the time accuracy of the detected transient, shown in figure 6.21. The values of abscissae are the difference between the real position of the transient in the signal and the position found by the methods. The threshold and wavelet methods give the best results. The LPC method shows a systematic shift corresponding to the semi-period of the simulated filtered transient, the prediction beginning to mismatch at this time. However, our goal is to detect offline a transient in our data, allowing us to correct the systematic shift in our analysis.

At this point, we decide to eliminate the threshold method, because of its calculation time and also its poor efficiency at low pulse amplitude compared to the two other methods. Consequently, our preference goes to the LPC method, because of its high percentage of detection at low and high amplitude transient, its low percentage of false

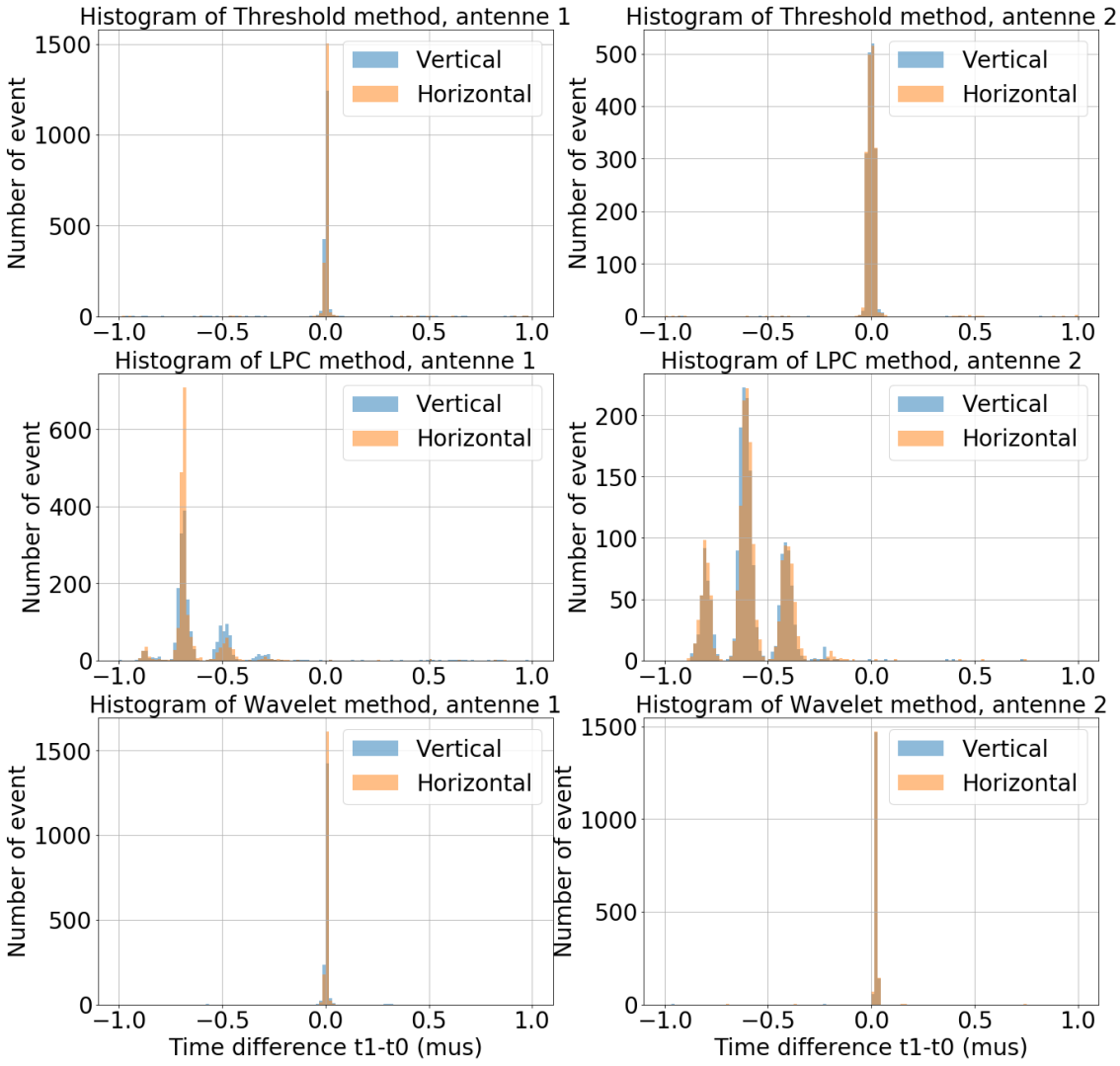


Figure 6.21 *– Time accuracy on the detected transient of each methods. The values of abscissae are the difference between the real position of the transient in the signal and the position found by the methods.*

detection and its fast calculation time.

6.5. Building a cosmic ray event

This section aims to describe the selection procedure of the events seen simultaneously by multiple instruments of CODALEMA and EXTASIS, and the elaboration of a so-called "cosmic ray event" associating several of these instruments. As it has been discussed in section 4.3.3, a trigger is generated by the particle detector and distributed to the Compact Array and to the LF antennas. For the 57 standalone antennas, no particle trigger is sent. Transients detected in coincidence on several of these standalone antennas build a "radio coincidence", characterized by an average radio event time that will be compared to the particle detector event. The criterion is that at least three standalone antennas are triggered within a time interval compatible with the propagation of a plane wave at the speed of light (typically ~ 5 µs for a horizontal shower and antennas covering the maximal extension of CODALEMA, i.e at a distance of 1.6 km from each other). The radio event is promoted as an actual shower if its timing is compatible with the timing of the scintillators within ±3 µs and if the reconstructed arrival directions agree within 20° [173].

The information of the Compact Array and the three-fold antenna are added and this constitutes the complete cosmic ray event. To register and archive the cosmic ray events (around 1 to 3 per day) and to be able to exploit them for further analysis, an ID-card is build, with the data of the various instruments: waveforms for the antennas, power spectrum densities, polarization figures, footprint for particle detector and radio detectors, information on the compacity and the rise time...

An illustration of a cosmic ray event is shown hereafter in figures 6.22, 6.23 and 6.24. For the 10 CA antennas and the 20 involved SA, both EW and NS polarizations have a signal. A transient is also seen in the three-fold antenna and the LF antennas. The directions of arrival reconstructed by the SA, CA and scintillators are respectively ($\theta_{SA} = 60 \pm 1\,°$, $\phi_{SA} = 154 \pm 1\,°$), ($\theta_{CA} = 58 \pm 5\,°$, $\phi_{CA} = 155 \pm 5\,°$) and ($\theta_{SC} = 61 \pm 16\,°$, $\phi_{SC} = 154 \pm 5\,°$).

Figure 6.22 presents:
- (a): The signals of the 20 SA recorded in EW polarizations in fullband. For each waveform plots, the left canvas features the entire waveform and the right one a zoom around the transient pulse.
- (b): Same as (a) for the NS polarizations.
- (c): The signals of the 20 SA in EW polarizations and filtered in [30 – 80] MHz. We note a shift of the position of the transient of the SA4, assimilated to a problem on the digitization card (MATACQ).
- (d): Same as (c) for the NS polarizations.
- (e): The signals of the 10 CA recorded in EW polarizations and filtered in [20 – 80] MHz.
- (f): Same as (e) for the NS polarizations and filtered in [20 – 80] MHz.
- (g): PSD of the SA in the EW polarization.
- (h): PSD of the SA in the NS polarization.

Concerning the signals seen in the CA, their shape looks like an inverted Ricker pulse (see figure A.3). The uncertainty on the determination of the maximum amplitude of the signal, and the corresponding timing used for the arrival direction reconstruction is estimated at 10 ns, corresponding to four sampling bins. For the SA, the shape of the pulse is clearly visible in figure 6.22.(c) right panel. The uncertainty on the determination of the maximum amplitude of the signal, and the corresponding timing is estimated at 10 ns, corresponding to a semi-period of the filtered transient. Figure 6.23 presents:
- (a): Filtered CA polarizations. Each square canvas represents the ellipse of polarization of one CA antenna. The straight black line accounts for the prediction from the geomagnetic mechanism. The red point indicates the maximum amplitude of the ellipse.
- (b): Same as (a) for the SA. The ellipse of polarization of the SA4 is chaotic, due to

the shift of the position of the transient seen in figure 6.22.(c).

— (c): The signals of the three-fold antenna in fullband. The green curve is the NW-SE polarization, the yellow curve is the EW polarization and the black curve is the NE-SW polarization (see section 4.2.4).
— (d): Same as (c) filtered in [30 – 80] MHz.
— (e): Figure obtained from the monitoring of the scintillators. The upper part of the figure presents the superimposed waveforms, a map of the triggered scintillators and the information on the multiplicity and the arrival direction. The lower part represents the waveforms of each scintillator.

Concerning the scintillators. We see different structures in the recorded signals, which lead to the choice of 30 ns for the uncertainty used for the reconstruction of the arrival direction. Figure 6.24 presents:

— (a): The signals of the 7 LF antennas recorded in the horizontal polarization, filtered in [1.7 – 3.7] MHz.
— (b): Same as (a) for the vertical polarization.
— (c): Map of the event. Grey dots represent the radio antennas. The colored triangles indicate the positions of the LF antennas that detected the shower. The involved standalone antennas are represented by coloured circles, whose colour indicates the timing order in which the signal has been seen by the antennas (from blue, earliest, to red, latest) and area of circles reflects the relative amplitude of the signal (linear scale). The green lines close to the circles indicate the orientation of the measured polarization of each MF antenna. For this event, the orientation of the measured polarization is nearly orthogonal to the direction of arrival of the event as expected from the dominant geomagnetic mechanism.
— (d): Same as (c) for the CA antennas.

6.6. Conclusion

In this chapter, we have presented three noise rejection algorithms to improve the efficiency of the radio array of CODALEMA, leading to a rejection rate of 99.96 % of the events recorded over one month of data taking. We have also presented different methods for the detection of signal in the LF band. We have chosen the LPC method, which presents a better efficiency at low pulse amplitude and a weaker percentage of false detection compared to the two other methods. Lastly, we have described the method to build a cosmic ray event and gave an illustration of the ID-card which is build for each event.

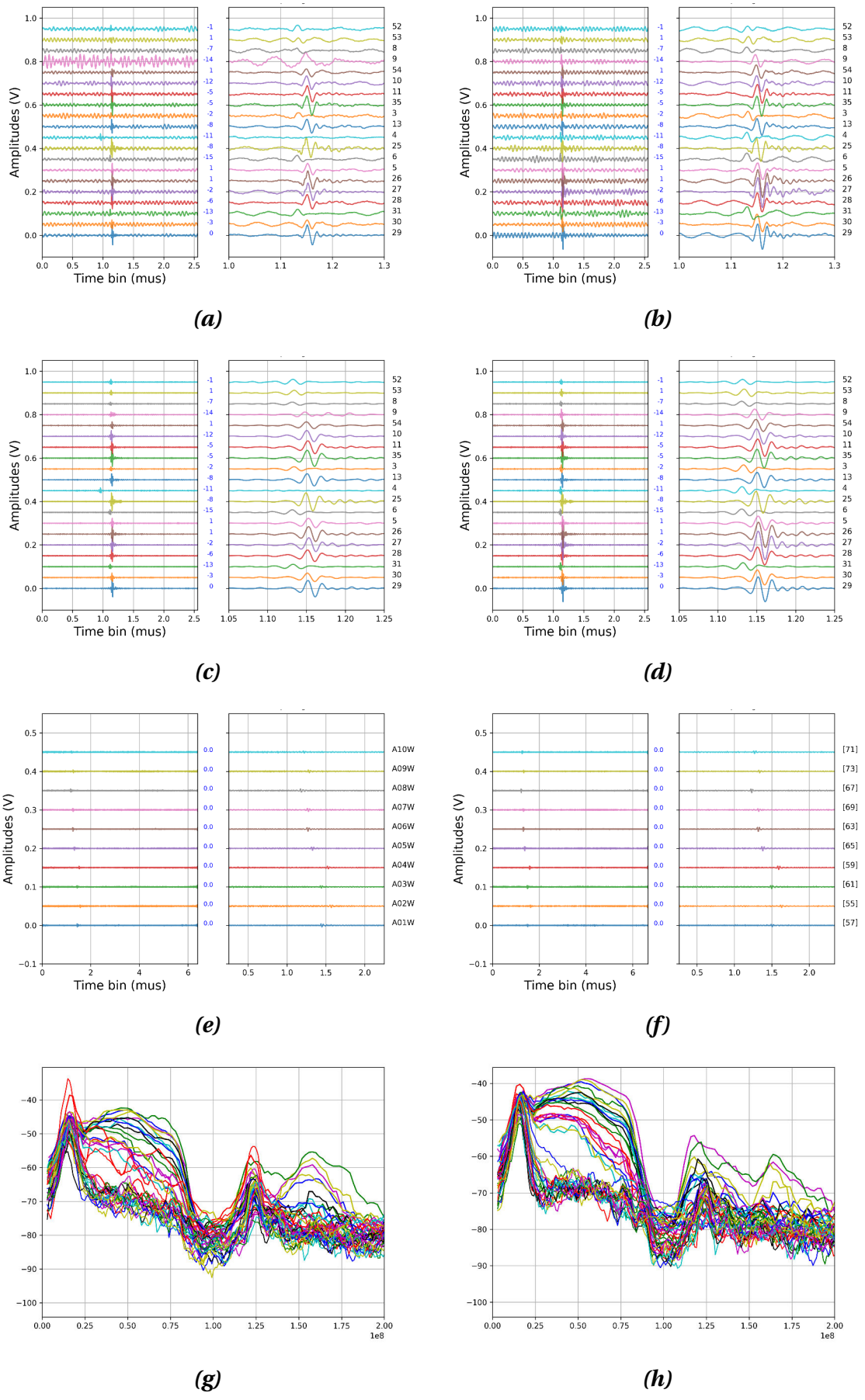


Figure 6.22 – *Example of waveforms and PSD of SA and CA antennas.*

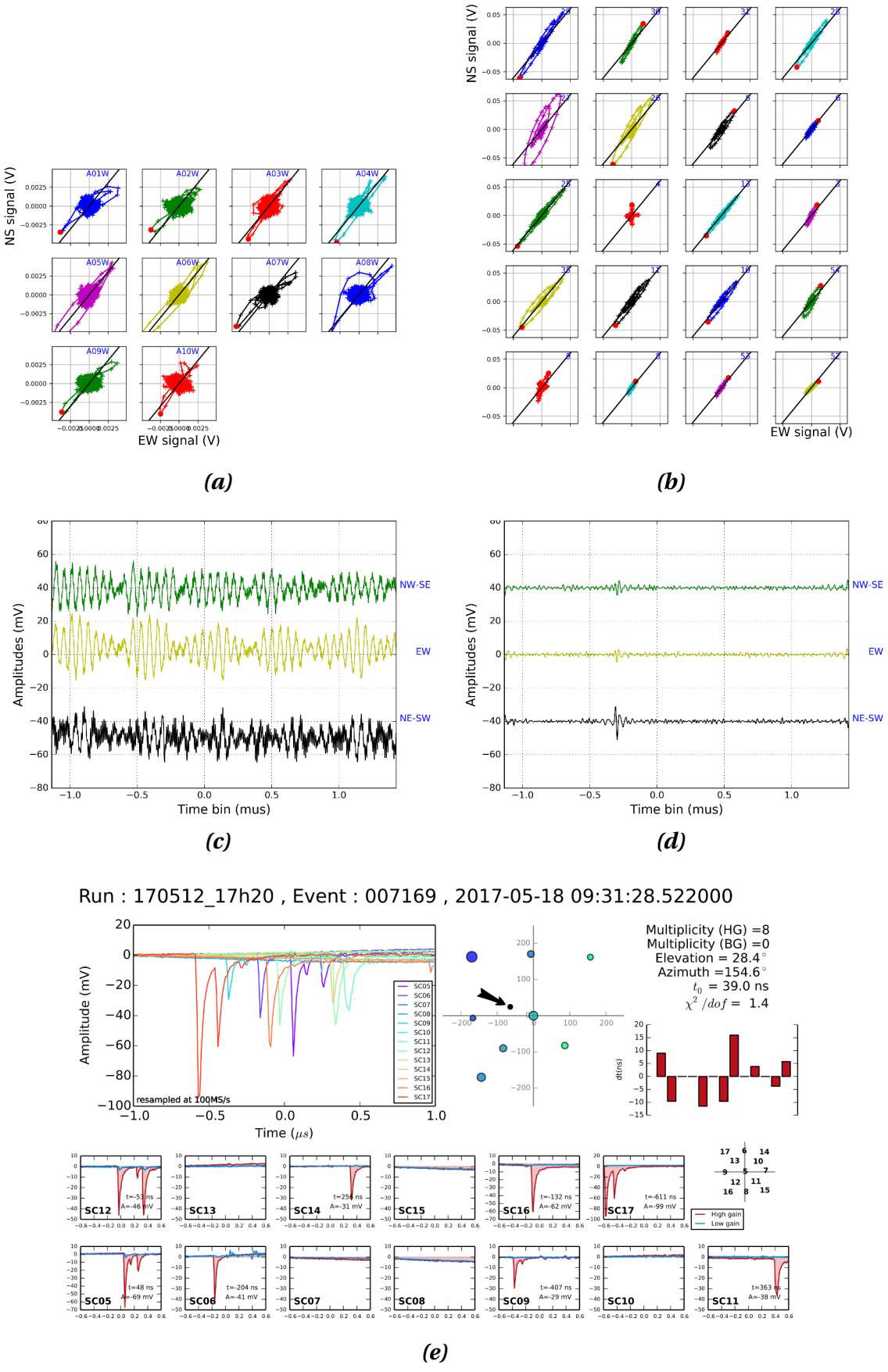
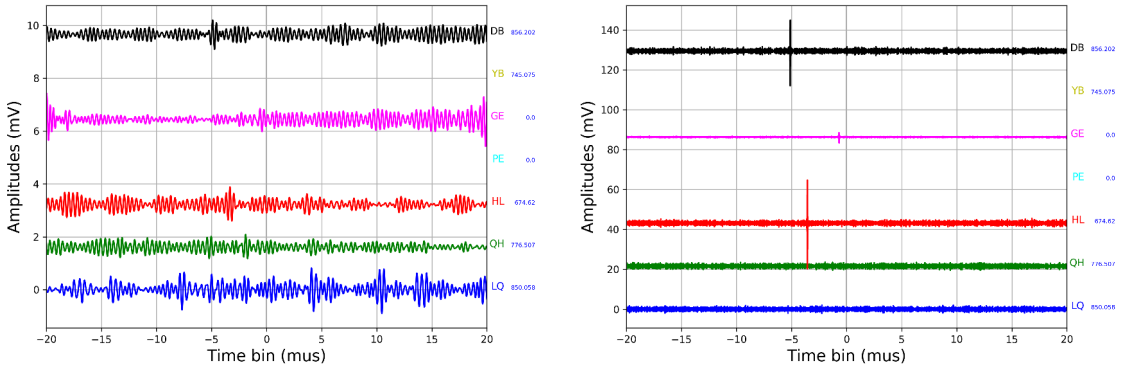
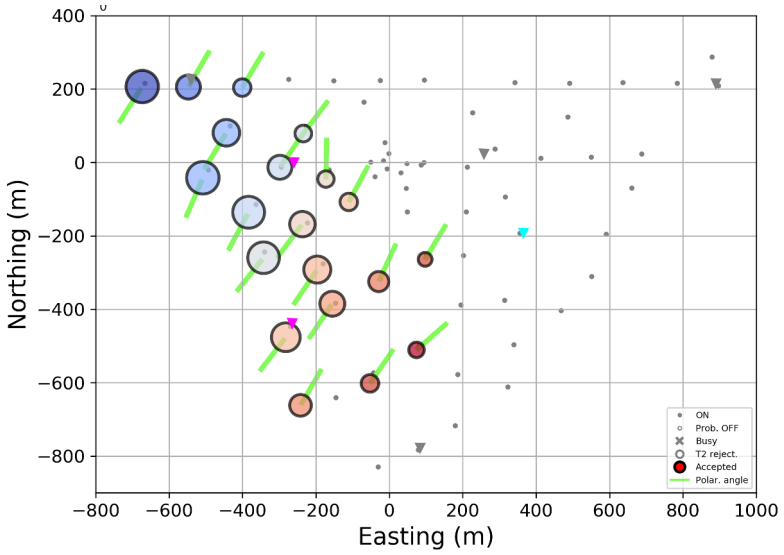


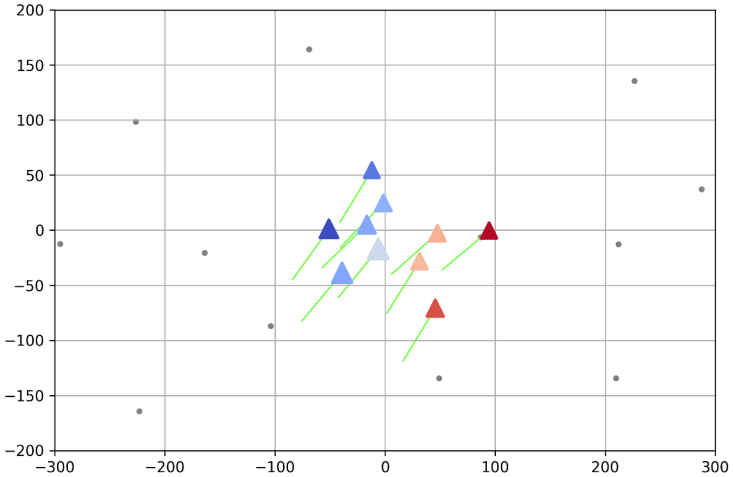
Figure 6.23 – *Example of SA and CA polarizations and three-fold antenna event.*



(a) *(b)*



(c)



(d)

Figure 6.24 *– Example of LF antenna waveforms and of SA and CA map of event.*

Part III

Analysis & Results from the CODALEMA/EXTASIS instruments

"You can do anything you set your mind to."
Benjamin Franklin

CHAPTER 7

Automatic analysis and radio reconstruction of Extensive Air Showers

Outline

Parts of the results of this chapter have been presented in ICRC2017, TeVPA2018, UHECR2018, URSI-France 2019 and ICRC2019 conferences, and published in [53, 146, 212, 213].

7.1. Introduction and incentive

We have seen in chapter 3 (for example in figure 3.31) that the radio signal is particularly rich in information on the shower. Indeed, all the information about the primary cosmic ray is contained in the electric field profile (or the lateral distribution function, noted *ldf*). We know that the intensity of the electric field is directly proportional to the energy of the primary [90, 6, 130] ; the slope of the *ldf* and the frequency spectra are linked to the mass of the primary (see figure 7.1) and to the geometry of the event ; the timing of the transients gives via triangulation the arrival direction of the primary ; and the position of the radio core is roughly the same as the position of the particle core.
From that point, the challenge is to use reliable radio features to reconstruct the information on the primary cosmic ray. Different method for the estimation of the shower parameters exist [71, 134, 135, 214, 139, 215]. In this thesis, we decided to develop the method proposed in [213]. The main feature of the method is the comparison of the data to several simulations. Indeed, no direct observables allow to go back to the X_{max} and it is thus necessary to go through a comparison with simulations to estimate it. Ditto for the position of the shower core and the primary energy. Considering the large amount of

collected data, it is also mandatory to build an automatic analysis and reconstruction of each interesting event.

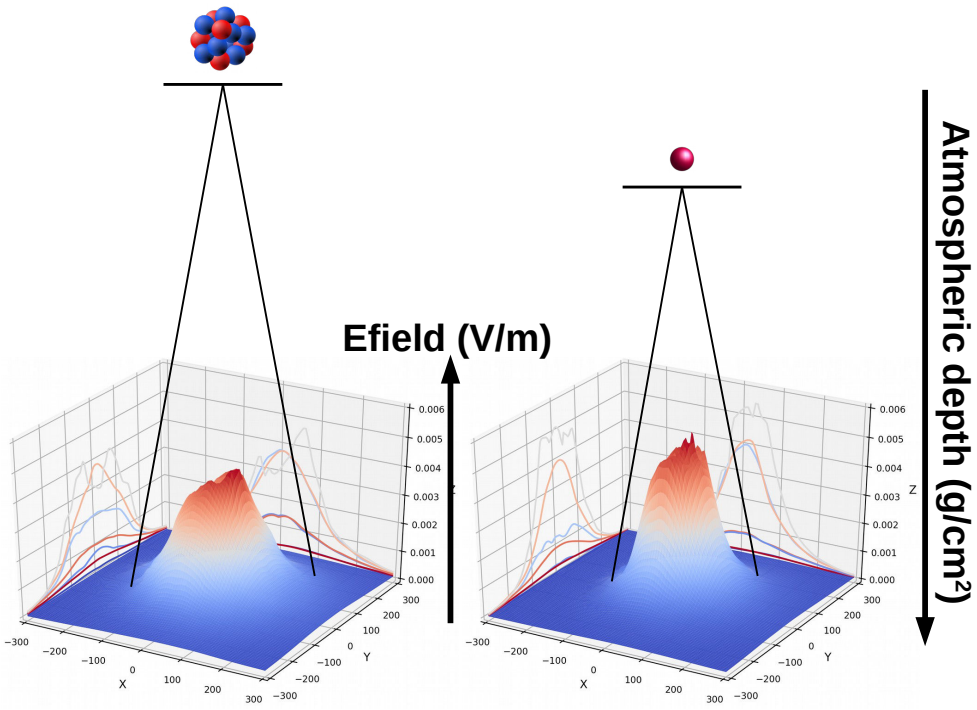


Figure 7.1 *– Difference of the simulated LDF depending on the mass of the primary, left for an iron nucleus and right for a proton. The vertical axis represents the atmospheric depth in* $\mathrm{g \cdot cm^{-2}}$ *and the horizontal lines stands for the* $X_{\max}$ *values for proton and iron nucleus induced showers (not at scale). For heavy particles such as iron, the* $X_{\max}$ *is reached high in altitude on average leading to large radio footprint on ground, while for light particles such as proton, the* $X_{\max}$ *is reached closer to the ground leading to smaller radio footprint. The size and shape of the radio footprint and the amplitude of the LDF can be measured to determine the* $X_{\max}$ *value.*

7.2. Basic procedure to analyse cosmic ray events and reconstruct their parameters

Figure 7.2 presents a flow diagram of the method. Each step will be discussed in details in the next subsections. It is an iterative method where the energy is adjusted for different tested core positions.

7.2.1. Preparation of the set of simulations

The first step is to build a cosmic ray event as explained in section 6.5. After this step, we obtain some pieces of information, such as the date and time of the event, the raw waveforms of the involved antennas and the arrival direction angles θ and ϕ obtained via triangulation made with radio detectors and particle detectors (see section 6.1 page 130). The results of the cross calibration described in chapter 5 are applied on the event. Based on previous studies [71, 213], we make 50 simulations of that event with SELFAS (see section 3.3 page 56). The number of simulations is statistically sufficient considering the concept of shower universality [110]. SELFAS needs several inputs, such as:

— The mass of the primary: since the most probable composition beyond 10^{17} eV lies between protons and iron nuclei (see section 1.2.4.1), we make 40 simulations

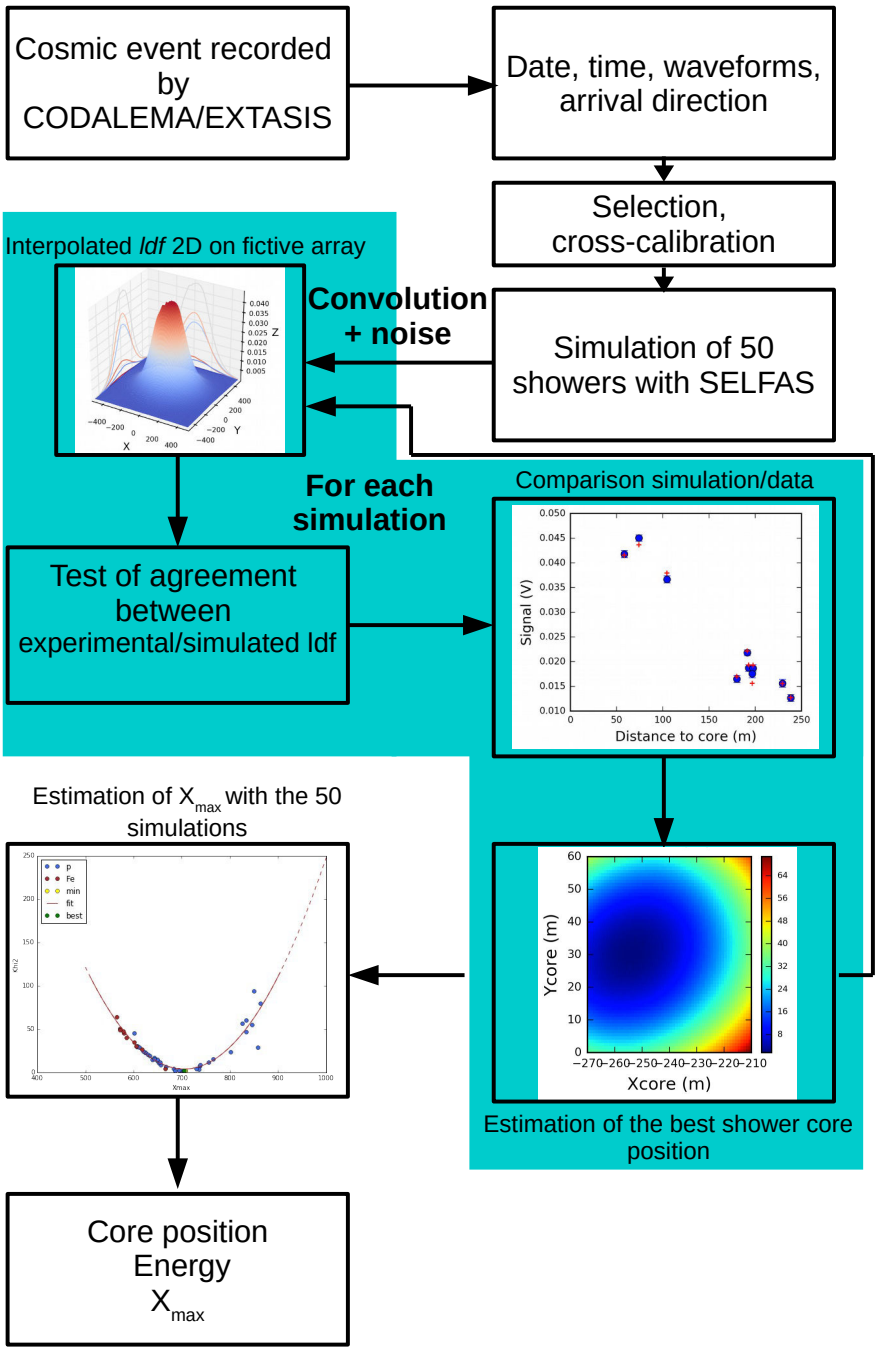


Figure 7.2 *– Flow diagram of the method. The blue area delimits the procedure made for each simulation. In the standard way, it is thus repeated 50 times*

with a proton as primary and 10 with an iron as primary (this proportion is to take into account the larger fluctuations of $X_{\max}$ depth for protons as discussed in chapter 2).

— The geomagnetic field vector, depending on the location of the experiment.

— The arrival direction of the event we want to reconstruct.

— The energy of the primary cosmic ray, which is arbitrarily set to 10^{17} eV, because we know that the amplitude of the electric field is directly proportional to the

energy of the primary, and we treat as a free parameter (which is a scaling factor) the amplitude of the computed *ldf*. With this, we avoid a lot of simulations at different energies.

— The antenna and core positions in a frame chose by the user. In our case, the core position is set to (0, 0) corresponding to the center of the particle detector array, and at the origin of a dense virtual array composed of 16 arms, each composed of 20 antennas and disposed in star (see figure 7.3) This virtual array is symmetric in the shower frame.

— The relative position of the simulated antennas.

— The ground altitude, depending on the location of the experiment.

— The number of secondary particles to be used in the simulation (not reflecting the "real" number in the actual shower).

— The first interaction depth, provided by CONEX which needs a model of hadron interactions at high energy: we choose to work with EPOS LHC (latest version of EPOS).

— A model of the atmosphere typical of the day and the location where the event was recorded. In the latest version of SELFAS, we work with the GDAS (see section 3.3.2.2 and appendix C).

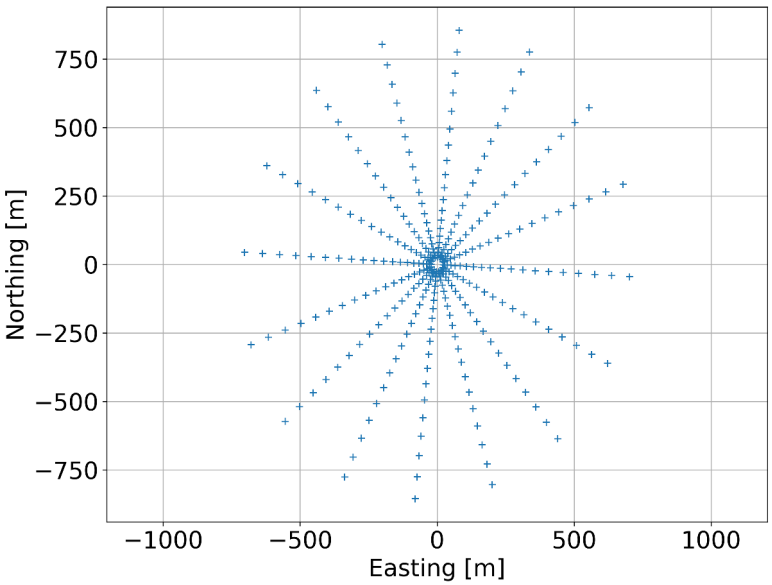


Figure 7.3 *– Example of fictive antenna array used in SELFAS simulations.*

Based on preparatory studies, setting the number of secondary particles to 20,000,000 permits to avoid the numerical noise in the simulations. Finally, the last consideration is the choice of the model for the high energy interaction model. The main models are QGSJET-II.04 [216], Sibyll 2.1 [217] and EPOS LHC [69]. EPOS takes into account the energy-momentum correlations between re-scatterings, and predicts larger number of muons than QGSJET. EPOS is the model which is the most in agreement with experimental data of the Pierre Auger Observatory, as explained in [218]. Even if the EPOS LHC and QGSJET-II.04 were updated with the last results from the LHC, the models can not match correctly the data. One can just notice that EPOS LHC model presents the lowest

uncertainty and is the closest to data values. This is the reason why, in the following, all the simulations are done with EPOS as high energy interaction model.

7.2.2. Convolving of SELFAS simulations

SELFAS computes the electric field, while CODALEMA and EXTASIS record a voltage. Thus, we have to choose which one to convert in order to compare them. Contrarily to the method presented in section 3.4.1.4, which consists in deconvolving the experimental data to retrieve an electric field, we choose to convolve the simulations (see section 3.4.1.5). This step transforms the simulated waveforms from $V \cdot m^{-1}$ to V, allowing us to compare them "directly" to the experimental data (with a scaling factor coming from the arbitrary choice of the primary cosmic ray energy to simulate). The convolving consists in taking into account the acquisition chain and antenna response, and converting the 3 components of the electric field given by SELFAS in voltage, then in ADC counts. For that, and as explained in section 3.4.1.4, the NEC software is used to simulate the antenna response, associated with measurements to consider the full detection chain. Figure 7.4 represents the convolving of a SELFAS simulation for the EW and NS polarizations and the associated power spectrum densities. There is a really good agreement between

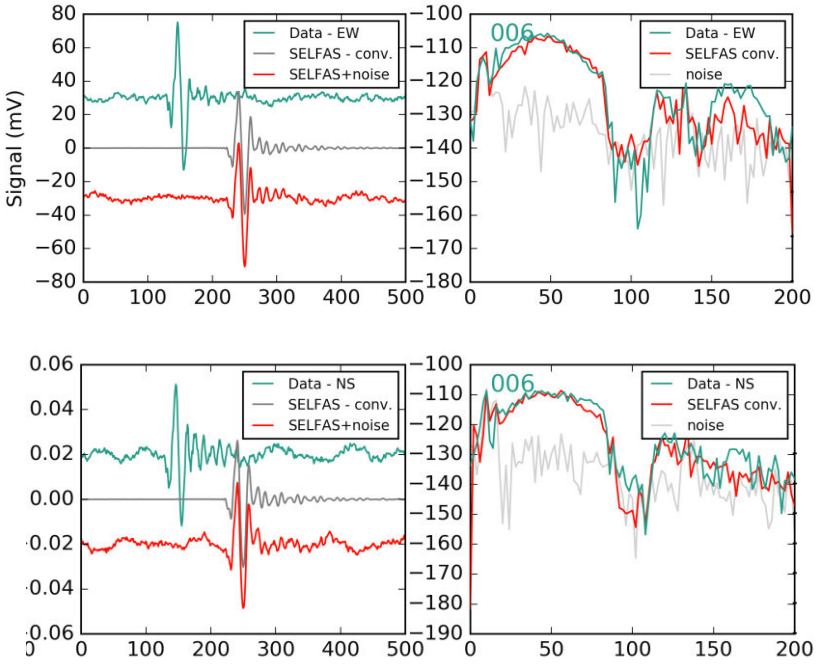


Figure 7.4 *– Convolution of SELFAS simulation. Up: the EW polarization. Down: the NS polarization. The figure on the left represents the waveforms of data (green), convolved SELFAS (grey) and convolved noise added SELFAS (red). The signals are in full band. The traces are shifted along x and y for clarity. The figure on the right represents the comparison of the power spectrum densities of convolved noise added SELFAS simulation (red) and the experimental data (green).*

the power spectrum densities, that is, between the convolved and noise added SELFAS simulation and the experimental data. This example of agreement is obtained after the reconstruction described in section 7.2.3 of this chapter. Wrong core location and/or wrong energy does not lead to such good matching between the experimental signal of a given antenna and the simulated signal from the nearest virtual antenna.

It is worth noticing that, by introducing the noise and adding it to the convolved

SELFAS simulation, the simulated transient could be buried in the noise, as shown in figure 7.5, where the signal is still visible but hard to see by eye compared to the signal of figure 7.4. This can also be seen on the power spectrum density of convolved and noise added SELFAS simulation, which is at the same level of power than the ones of the data and the noise. To add noise to the simulated signal, we proceed as follows:

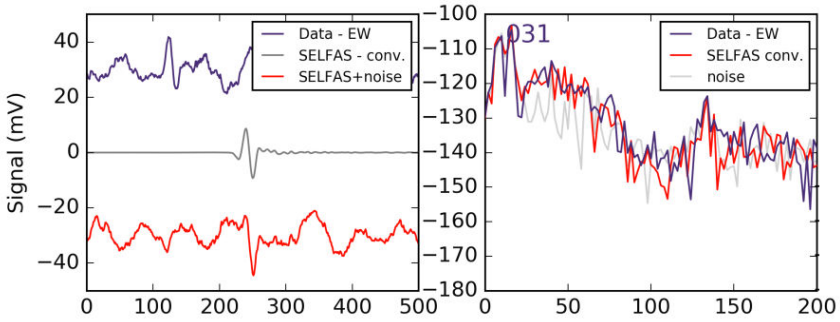


Figure 7.5 *– Convolution of SELFAS simulation for the farthest antenna with a detectable signal.*

— once the estimation of the shower core is obtained (see section 7.2.3), we position the center of our fictive array at the position of the estimated shower core, and we associate each simulated antenna to the closest real antenna,
— we use the recorded waveform of the closest real antenna, and define a window where only noise is expected (the end of the waveform),
— we add to the simulated signal the noise of the closest real antenna taken in the defined window.

We have verified that the window defined for the noise had no impact on the final result (we have tested a window taken at the beginning of the waveform in particular). For the simulated antennas far from the estimated shower axis, the addition of noise makes the transient not only invisible but also undetectable, which is really consistent with our observations. Indeed, it is well known that the detection range is directly dependent on the shower axis distances.

7.2.3. Reconstruction of core position, energy and $X_{\max}$

Remark: The initial method does not used the calibration coefficients calculated in chapter 5. They will only be implemented and used in the generalized method presented in section 7.4.

In order to have the voltage at each position of the simulated array (not only at the simulated antenna positions), we need to compute the simulated *ldf* and to linearly interpolate it. Then, we test the agreement between the interpolated simulated *ldf* and the measured voltage at different shower core positions (*i,j*). A chi-square minimization is made:

$$\chi^2(x_i, y_j) = \frac{1}{n_{ant} - 3} \sum_{k=1}^{n_{ant}} \left(\frac{A_k^{det} - a_{ij} A_{ijk}^{sim}}{\sigma_k^{det}} \right)^2 \times f(a_{ij}) \tag{7.1}$$

where:

— n_{ant} is the multiplicity of the event, i.e the number of involved antennas,
— A_k^{det} is the amplitude in volts of the filtered signal measured by the antenna number k, equal to the quadratic sum of the EW and NS polarizations,
— σ_k^{det} is the corresponding uncertainty, calculated from the RMS of the noise and the formulae of the propagation of the error on the quadratic sum,

— A_{ijk}^{sim} is the amplitude of the simulated signal convolved and filtered for the antenna k with respect to the tested core position (i, j) (the convolving is made in [10 – 200] MHz and then the convolved signal is filtered by a Tukey window in [30 – 80] MHz),

— a_{ij} is the scaling factor (related to the energy) at the tested core position i, j (the scaling factor is the same for all the antennas at a given tested shower core),

— $f(a_{ij})$ is an empiric function introduced to avoid small chi-square values (which indicate a good agreement) related to small simulated voltages (occurring when the tested core position is very far from the true core position ,see [71]), and is equal to $1 + 10^{10} \exp\left[-\frac{a_{ij}-\langle A^{det}\rangle}{0.2\langle A^{det}\rangle}\right]$.

Let us now consider as examples two experimental events detected by CODALEMA in the last two years (see figure 7.6). One is internal ($\theta = 26\,°$, $\phi = 161\,°$), and the other one is external ($\theta = 32\,°$, $\phi = 353\,°$) to our antenna array.

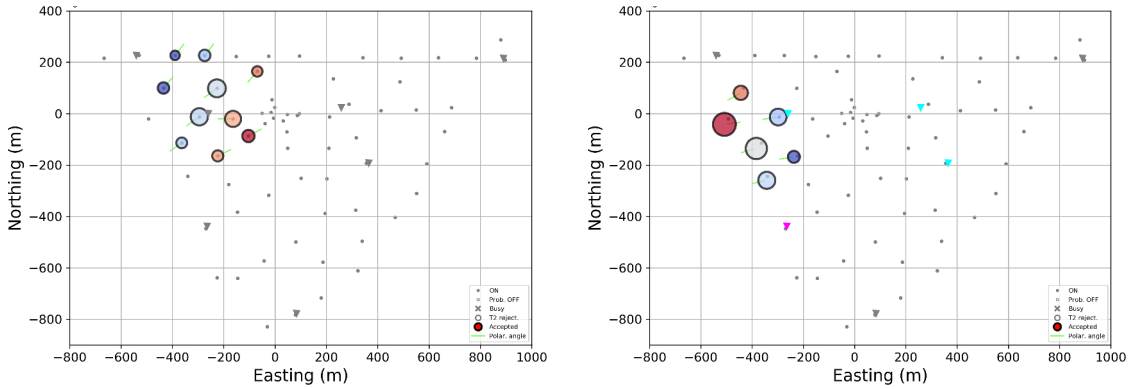


Figure 7.6 *– Footprint of two example events. Grey dots represent the standalone antennas. The involved standalone antennas are represented by coloured circles, whose colour indicates the timing order in which the signal has been seen by the antennas (from blue, earliest, to red, latest) and area of circles reflects the relative amplitude of the signal (linear scale).*

The interpolated simulated *ldf* are compared to the data for different shower core positions, as explained previously. The results of the comparison are shown in figure 7.7. These two plots are the results of the comparisons at the best core positions, i.e for the lowest chi-square value. One can see that SELFAS simulation is fairly accurate to reproduce the data. Generally, it is easier to reconstruct the parameters of an event contained in our antenna array (internal) than to reconstruct the parameters of an event not contained in our antenna array (external). Indeed, in the latter case, the experimental *ldf* is not well defined (lack of antennas) and the comparison with the simulated *ldf* might not be optimum. However, the method allows to find an estimation of the core position even outside the array.
This step in our analysis is done for all the tested core positions, giving for each of them a reduced chi-square value. These values can be represented on a density map as shown in figure 7.8. A clear minimization of the reduced chi-square values can be seen, indicating the best estimated shower core position.

This procedure is repeated for all the set of simulations (50 times). So, for each comparison between data and one simulation, we obtain a best estimated shower core corresponding to the lowest reduced chi-square value.

The set of simulations has been produced with randomly chosen $X_{\max}$, covering a large

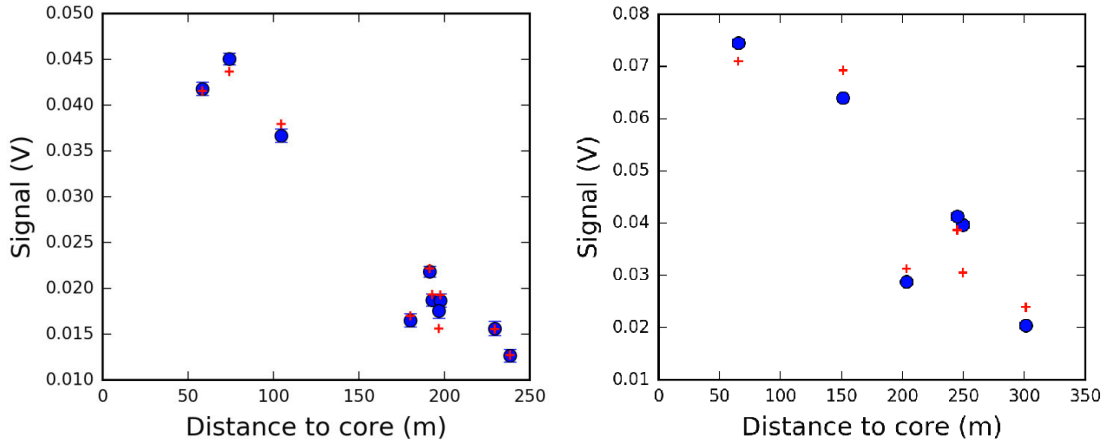


Figure 7.7 *– Comparison of the amplitude as a function of the shower core distance between simulations (red crosses) and data (blue points). Left: the event is internal. Right: the event is external. These two plots are the results of the comparisons at the best core positions (so, for the lowest reduced chi-square). The error bars on the experimental data are the RMS of the signal.*

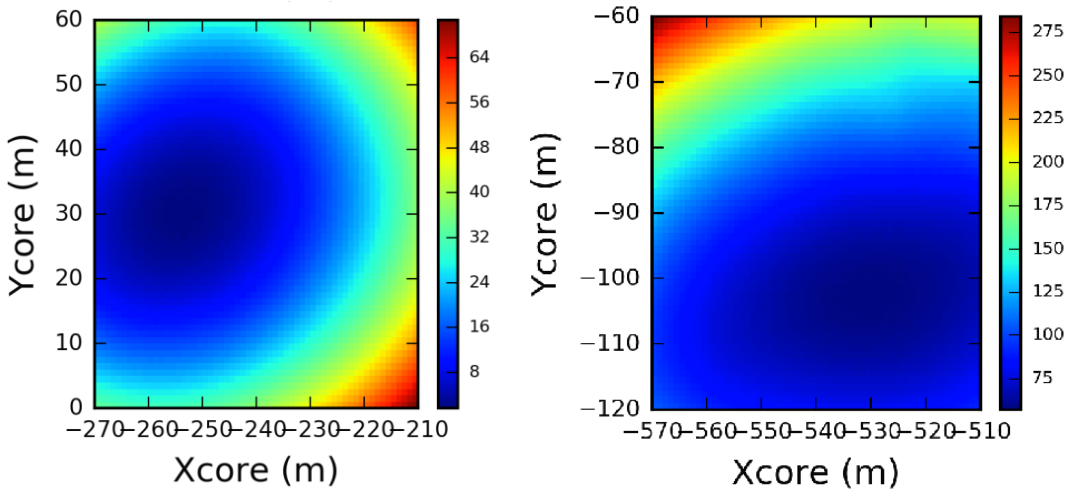


Figure 7.8 *– Density map of the reduced chi-square values for a given simulation. Left: the event is internal. Right: the event is external. The colorbar indicates the reduced chi-square values.*

window of possible values. We can thus test the agreement between data and simulation as a function of the $X_{\max}$ values. The best reduced chi-square value (obtained for the best estimated shower core position) for each comparison is saved, and plotted as a function of the corresponding $X_{\max}$, as shown in figure 7.9. By fitting the values, we obtain the $X_{\max}$ corresponding to the best agreement. For our two examples, the estimated parameters are shown in table 7.1. Let us now try to estimate the errors on the reconstructed parameters.

7.3. Uncertainty estimation on the reconstructed parameters

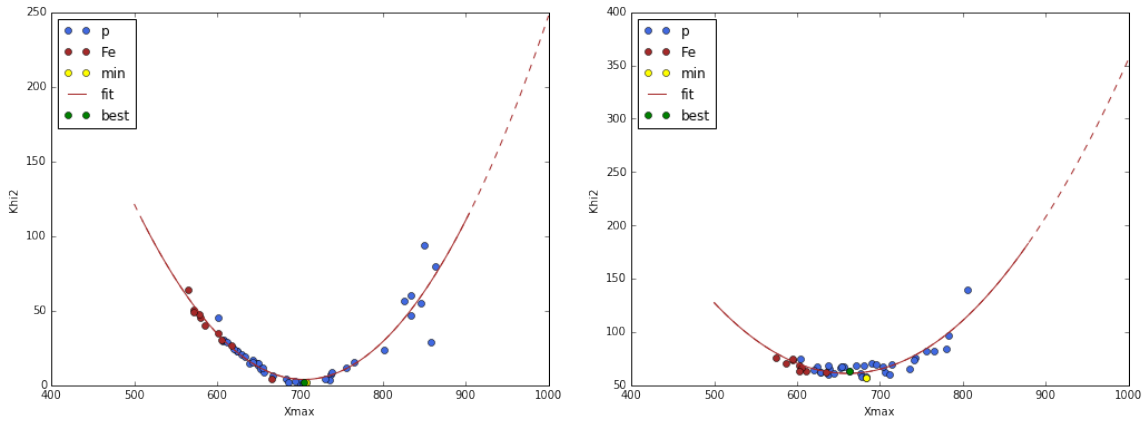


Figure 7.9 *– Agreement between data and simulations as a function of $X_{\max}$. Left: example of the internal event. Right: example of the external event.*

	Internal event	External event
Xcore	−257 m	−539 m
Ycore	29 m	−107 m
Xmax	705 g · cm^{-2}	663 g · cm^{-2}
Energy	0.87 EeV	2.18 EeV

Table 7.1 *– Estimated parameters of the two considered examples.*

7.3.1. Framework

In the case of a poor data statistics (small numbers of data), it is although possible to deliver an estimator D of any parameter ζ characterizing all or part of the data. The reliability of the estimation can be measured thanks to an interval of confidence, which reflects the statistical precision of that estimator. This interval of confidence is a sort of criterion on the quality of our estimator, and based on the difference between the estimator and the parameter, also called the error estimation. In our case, the distribution of our experimental data is Gaussian. In the following, we choose 1σ as interval of confidence. So, the probability P_r that our estimator falls in our interval of confidence can be written as:

$$P_r = \frac{1}{\sqrt{2\pi\sigma}} \int_{\mu-\sigma}^{\mu+\sigma} \exp\left(-\frac{x-\mu}{2\sigma^2}\right) \mathrm{d}x = 68\,\% \text{ by definition} \tag{7.2}$$

where x is the estimation of the mean, μ is the mean and parameter to determine.
The idea is to calculate randomly the errors within a Gaussian distribution centred on the amplitude of the signal and with standard deviations calculated from the errors on the signal amplitude. We have tested three methods of error propagation. The first one is the method described in [71] where the error propagation is made on the experimental data. The second one is fairly the same method, but the propagation is made on the simulated data. Finally, the third method is a method used by LOFAR and presented in [219]. For the sake of clarity, only the classical method (the first one) will be discussed in the following, it is the chosen one.

7.3.2. Uncertainties on EAS parameters using the uncertainties on data

For this method, the Gaussian distribution is centred on the experimental signal amplitude and its standard deviation ($\sqrt{(\langle(|x-\langle x\rangle|^2)\rangle)}$) is calculated from the experimental errors on the measured signal. A new value for the experimental amplitude is randomly picked within the distribution, and this procedure is made for each involved antenna. We

then apply the procedure explained in section 7.2.3, and we obtain a new set of estimated parameters (core position, energy). We repeat this step sixty times (statistically sufficient). The uncertainties can be determined by the obtained distribution. Indeed, we determine the estimated parameter as the centroid and the interval of confidence is the standard deviation of the obtained distribution. For the shower core position, we obtain the distribution shown in figure 7.10. For the internal event, we obtain $Xc = (-253 \pm 8)$ m and

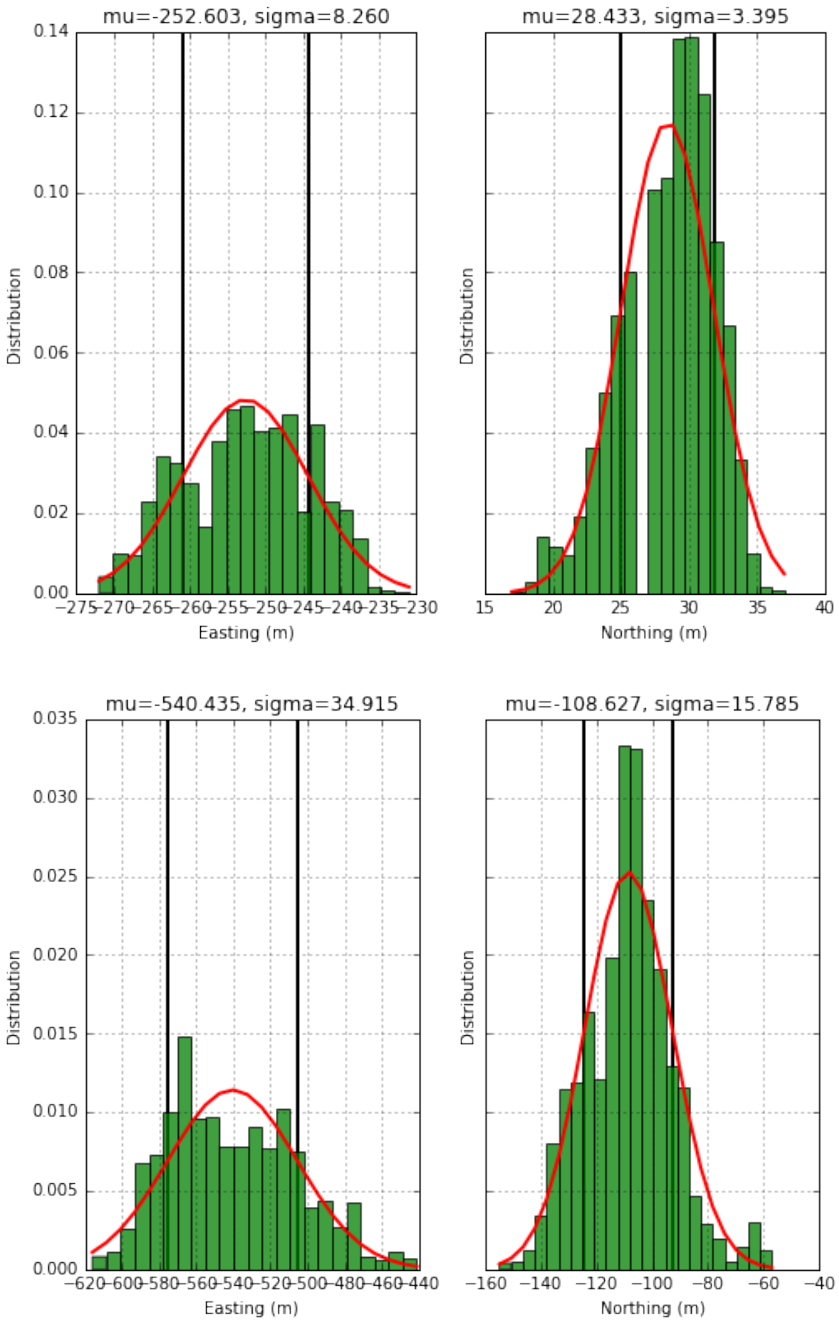


Figure 7.10 *– Probability distribution of the estimated shower core position in easting (left) and northing (right) units. The vertical lines represent the standard deviation. Top: example of the internal event. Bottom: example of the external event.*

$Yc = (28 \pm 3)$ m. For the external event, we obtain $Xc = (-540 \pm 35)$ m and $Yc = (-108 \pm 16)$ m.

The distribution of the energy is shown in figure 7.11. For the internal event, we obtain $E = (0.9 \pm 0.1)$ EeV. For the external event, we obtain $E = (2.3 \pm 0.4)$ EeV.

Finally, the distribution of the $X_{\max}$ is shown in figure 7.12. For the internal event, we

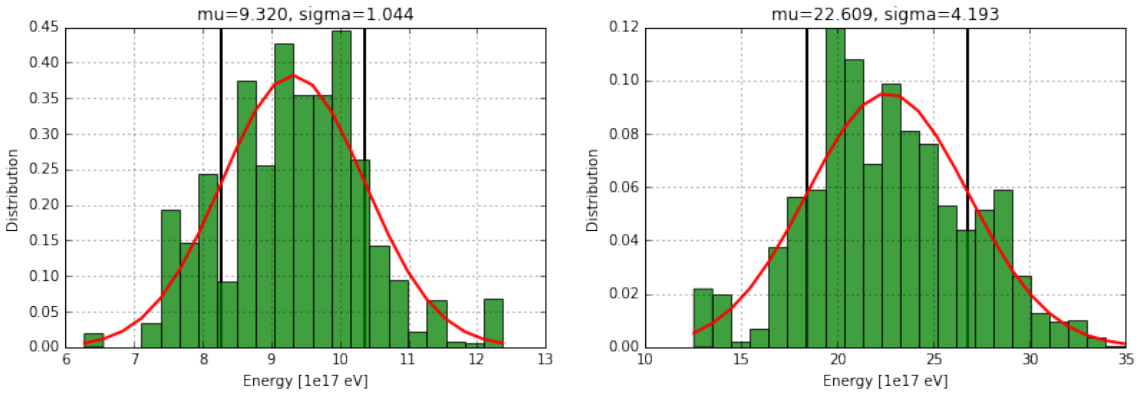


Figure 7.11 *– Probability distribution of the estimated energy. The vertical lines represent the standard deviation. Left: example of the internal event. Right: example of the external event.*

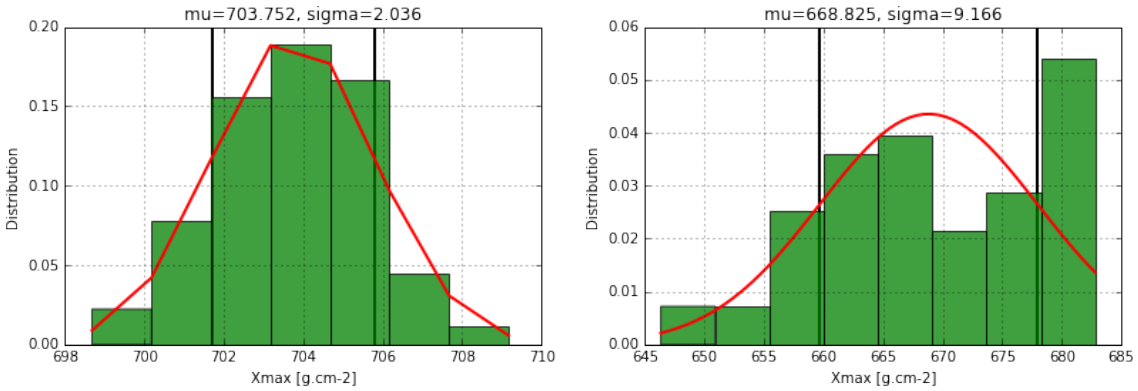


Figure 7.12 *– Probability distribution of the estimated $X_{\max}$. The vertical lines represent the standard deviation. Left: example of the internal event. Right: example of the external event.*

obtain $X_{\max} = (703 \pm 2)\,\mathrm{g \cdot cm^{-2}}$. For the external event, we obtain $X_{\max} = (668 \pm 9)\,\mathrm{g \cdot cm^{-2}}$. The estimated parameters and their uncertainties are shown in table 7.2.

The energy, core position and $X_{\max}$ values are compatible with the ones found by the first procedure without the error estimation, see table 7.1.

This method has already been used in previous work [71] and the self consistency of this latter has been verified. This allow us to affirm that the results obtained in this section are reliable, and in agreement with the results of the previous section.

7.4. Generalization of the method

Remark: As previously mentioned, the calibration coefficients calculated in chapter 5 are implemented and used in the generalized method presented in following section.

	Internal event	External event
Xcore	-253 ± 8 m	-540 ± 35 m
Ycore	28 ± 3 m	-108 ± 16 m
Xmax	703 ± 2 g $\cdot$ cm^{-2}	668 ± 9 g $\cdot$ cm^{-2}
Energy	0.9 ± 0.1 EeV	2.3 ± 0.4 EeV

Table 7.2 *– Estimated parameters and their uncertainties for our two examples.*

7.4.1. High frequency signal exploitation

Initially, the reconstruction of the shower parameters was made by using filtered signals in the range [30 – 80] MHz, and by quadratically summing both polarizations EW and NS [71, 220, 213]. The initial restriction to the [30 – 80] MHz band is mainly due to man-made broadcasting at low and medium frequencies. However, as said previously, CODALEMA works with a sampling rate of 1 GS $\cdot$ s^{-1} and the bandwidth of the antenna is optimized and well mastered up to 200 MHz. To fully exploit the capabilities of the CODALEMA autonomous stations, we have decided to also use the high frequency data for the event reconstruction, leading to a reconstruction in the [30 – 200] MHz band. As it is discussed below, the use of the continuity in the spectra of CODALEMA data is very useful for inclined showers. Indeed, inclined showers develop much higher in altitude. Thus, the source of the radio-emission is farther than for a vertical shower, which leads to a larger ground pattern, but also to a flatter lateral distribution [221]. The idea is to use the [120 – 200] MHz pieces of information to better constrain the comparisons, an thus to better reconstruct this type of events.

In figures 7.13, two events seen by CODALEMA in the classical band are shown. The size of the circle surface is proportional to the sum of the square of the two polarization amplitudes. The amplitude gradient is weak, and no particular pattern appears, which is why the comparisons do not converge efficiently.

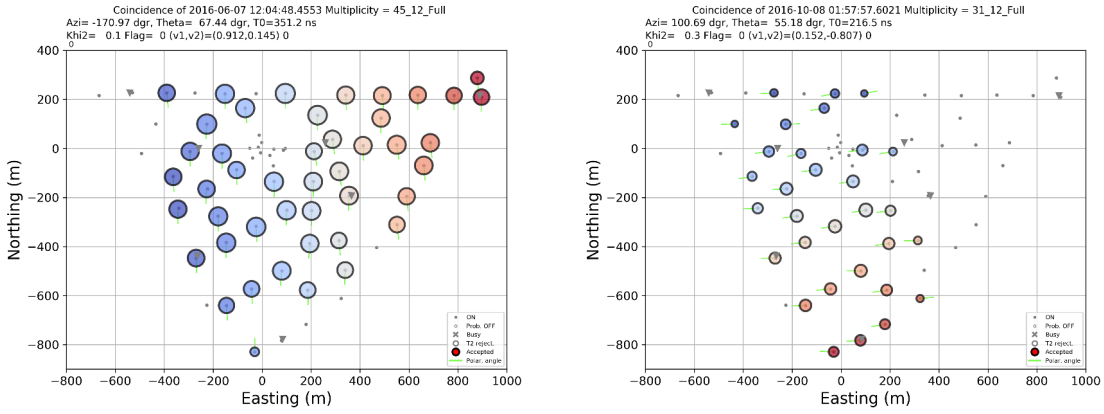


Figure 7.13 *– Left: event seen by CODALEMA in the classical band* [30 – 80] MHz. *Right: another event seen by CODALEMA in the classical band. See text.*

If we now study these events in the frequency range [120 – 200] MHz, we obtain the figures 7.14. One can easily see that amplitude gradients become visible. Moreover, one can see that a particular pattern appears: a part of a Cherenkov ring is visible on these two events. This particularity can be exploited for further improvements and is discussed in appendix E.

It is worth noticing the importance of the use of a wide-band antenna, such as the Butterfly used for the CODALEMA experiment, working in the [20 – 200] MHz range as

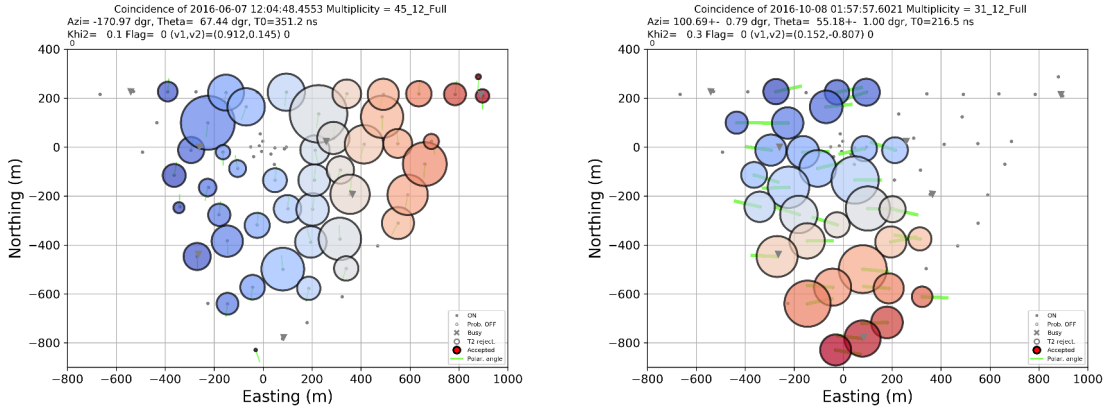


Figure 7.14 *– The two events seen by CODALEMA in* [120 – 200] MHz. *The amplitude of the signals (size of the circles) have been scaled for visibility.*

shown in section 4.2.2. As mentioned above, the idea is to take advantage of the high frequency part and especially the non-homogeneous pattern at ground to better reconstruct the shower parameters. Using the generalized radio method, the estimation of the shower core is estimated at (x_c =−46 ± 9 m,y_c =−373 ± 20 m) for the event presented in figure 7.14 on the left. The energy is estimated at $(3.9 \pm 0.3) \times 10^{17}$ eV. The chi-square of the generalized method is divided by three compared to the chi-square of the initial method. Thus, including the [120 – 200] MHz band in the reconstruction of the events permits to better constrain the comparisons.

7.4.2. Decoupling both polarizations

The quadratic summation of both polarizations EW and NS makes us lose information about each polarization and about the polarity of the signal. In order to fully exploit the capabilities of the CODALEMA autonomous stations, we decide to decouple the polarizations in the procedure, which means to treat them independently. To take benefit of the two polarizations and the full frequency band, equation 7.1 becomes:

$$\begin{aligned}\chi^2(x_i,y_j) = \frac{1}{4\times n_{ant}-3}\sum_{k=1}^{n_{ant}}\Bigg[&\left(\frac{AEW_k^{det}-a_{ij}AEW_{ijk}^{sim}}{\sigma EW_k^{det}}\right)^2_{\mathrm{MF}}+\left(\frac{ANS_k^{det}-a_{ij}ANS_{ijk}^{sim}}{\sigma NS_k^{det}}\right)^2_{\mathrm{MF}}\\ &+\left(\frac{AEW_k^{det}-a_{ij}AEW_{ijk}^{sim}}{\sigma EW_k^{det}}\right)^2_{\mathrm{HF}}+\left(\frac{ANS_k^{det}-a_{ij}ANS_{ijk}^{sim}}{\sigma NS_k^{det}}\right)^2_{\mathrm{HF}}\Bigg]\times f(a_{ij})\end{aligned} \tag{7.3}$$

where:

— n_{ant} is the multiplicity of the event, i.e the number of involved antennas,

— AEW_k^{det} (ANS_k^{det}) is the amplitude of the filtered signal (in MF for [30 – 80] MHz and in HF for [120 – 200] MHz) measured by the antenna number k in the East-West (EW) polarization (North-South polarization, NS),

— σEW_k^{det} (σNS_k^{det}) is the corresponding uncertainty for the EW (NS) polarization,

— AEW_{ijk}^{sim} (ANS_{ijk}^{sim}) is the amplitude of the filtered simulated EW (NS) polarization (in [30 – 80] MHz and [120 – 200] MHz) at the relative position k with respect to the tested core position i, j,

— a_{ij} is the scaling factors at the tested core position i, j,

— $f(a_{ij})$ is an empiric function introduced for calculation purposes (see [71]), and is equal to $1 + 10^{10} \exp\left[-\frac{a_{ij}-\langle A^{det}\rangle}{0.2\langle A^{det}\rangle}\right]$.

The use of polarization removes some ambiguities [213]. For example, events with abnormal polarization due to a strong atmospheric electric field could not be reconstructed correctly, since these events would not be simulated correctly.

7.4.3. Discussion on the generalization of the method

We illustrate the performance of the generalized method (decoupling of the polarizations, using of the high frequency band and including the calibration parameters) compared to the initial one on a reduced set of 361 events. The set was reduced mainly due to the time calculation. The results are shown in figure 7.15. We see that the generalized method gives better reduced chi-square values, meaning a better confidence in the estimated shower parameters.

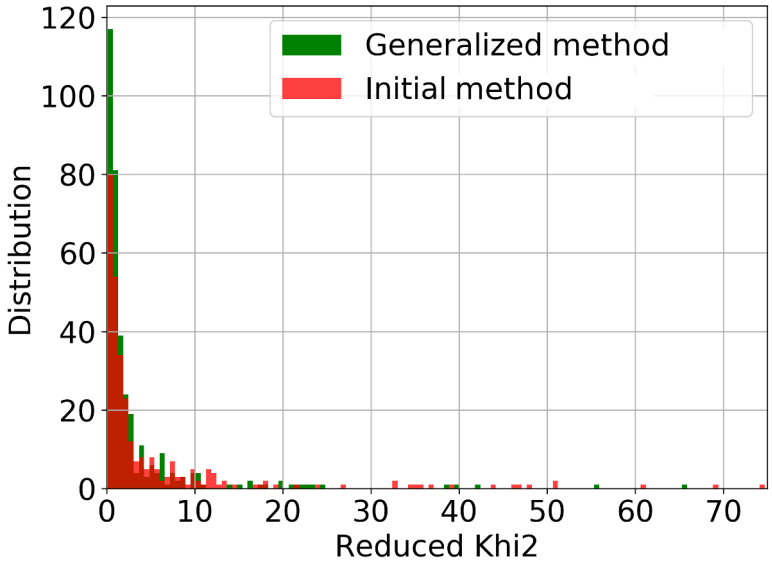


Figure 7.15 – *Comparison of reduced chi-square values of the initial and generalized method on a set of 361 events.*

In order to compare in detail the performances of both methods, we choose to work on three events and their associated simulations. For each event, and their whole set of simulations, we pick up one simulation and transform it into "fake" event. In addition to the convolving and the addition of the noise, the modifications consist in setting the shower core to a fake position and in using a fake energy scaling factor. We then apply the procedure explained in section 7.2.3 by fitting the remaining 49 simulations to the fake event, with the initial and generalized methods. Knowing the parameters to be reconstructed, we can evaluate the performances of both methods. The results are shown in table 7.3. The deviations from the real value of the parameters are indicated. Based on the three events used with known parameters, it seems that the generalized method is more accurate to estimate the shower parameters.

#	ΔX_{max} [g/cm^2]		ΔX_c [m]		ΔY_c [m]		$\Delta\alpha$ [%]	
1	-13 ± 84	5 ± 89	-6 ± 7	-5 ± 7	-20 ± 12	-16 ± 18	2.2 ± 12	-1.1 ± 10
2	-14 ± 102	-12 ± 84	-10 ± 21	-7 ± 22	12 ± 40	6 ± 35	-0.9 ± 8	-1.6 ± 8
3	16 ± 115	16 ± 109	0 ± 2	0 ± 4	1 ± 13	1 ± 11	1.3 ± 13	0.9 ± 13

Table 7.3 *– Comparison of the reconstruction methods. The deviation from the actual value of the parameters is indicated. For each parameter: left columns are for the initial method and right columns are for the generalized method.*

7.5. Automatic reconstruction of shower parameters based only on the radio signal

From now, the objective is to apply the previous analysis on all events detected by the CODALEMA and EXTASIS experiments. For that, an automatic procedure was build, whose analysis pipeline is represented in figure 7.16. Every hour, a script is executed on a local machine to check if a new event is existing. If it is the case, the event is processed (the construction of a cosmic ray event is explained in section 6.5 of chapter 6). The data are cleaned of transmitters (narrowband-transmitters, AM, FM), and antennas showing abnormal signs of functioning are flagged. The directions of arrival are reconstructed as explained in section 6.1 page 130. The necessary inputs for SELFAS simulations are stored in a global pickle file (used for serializing and de-serializing a Python object structure).

Then the SELFAS simulations are launch at the Computing Center (CC) of Lyon via a script hosted by a local machine, which is also executed every hour. The simulated array of 320 antennas is produced from the direction of arrival information. Forty protons and ten irons as primary cosmic rays are simulated (this proportion is required to take into account the dispersion on X_{max} values, as shown in figure 2.5 page 39). The GPS time of the event is also needed, to use the description of the atmosphere given by the GDAS (Global Data Assimilation System) data, as explained in [113] and in appendix C. The GDAS files are provided every weeks. Thus, the events are put in a queue waiting for the GDAS file to be provided. A dedicated script hosted by a local machine uploads each week the GDAS file from the GDAS Archive Information server. At the end of each of the 50 simulations, a file containing the simulated electric field is produced. The next step is to convolve the simulated electric field file with the response of the antenna with the method described in section 3.4.1.5 page 71. As an example, for an event with a zenith angle of 42° and an azimuth angle of 61°, one SELFAS3 simulation plus the convolving process take 15h48min on the worker *ccwsge0641* of the CC of Lyon, with as inputs a GDAS file for the description of the atmosphere, a simulated antenna array of 320 antennas and 20×10^6 particles.

Once the simulations are done, they are stored at CC Lyon. Another script hosted by a local machine is executed every two hours to check if new simulations are existing on the remote disk at CC Lyon, and to automatically launch the procedure previously described in section 7.2, with the estimation of the uncertainties thanks to the procedure explained in section 7.3.2 and with all the improvements described in section 7.4. The results of the comparison are stored at the same place than the simulations of the event in pickle files. As an example, for the same event mentioned above, the procedure described in section 7.2 takes seven minutes on the worker *ccwsge0641* of the CC of Lyon. For one simulation, sixty one files containing the results of the comparison are produced, meaning that for a complete event with fifty simulations, 3,050 pickle files are produced.

A third script is executed every day from a local machine to synchronize the remote path containing the simulations and the results of the comparisons and a local path, and

to remove files older than seven days from the remote path to free space. The average storage space for one event is around 8 GB. Finally, a last script analyses the outputs of the comparisons between simulations and event and perform another chi-square minimization to produce plots such as figure 7.9, but also to obtain an estimation of the $X_{\max}$ with its uncertainties, an estimation of the core position with its uncertainties and an estimation of the coefficient factor (giving the energy) with its uncertainties. The estimation of the $X_{\max}$ value and the energy will permit us to estimate the mass composition of cosmic rays detected by CODALEMA/EXTASIS, as we will see in chapter 9, to be compared with the estimation of the mass composition given by Auger, as shown in figure 3.1.

Remark: Programming efforts have been made to minimize some calculation times. However, at no time was there any question of a thorough optimization involving in particular the programming language, the computing location, etc.

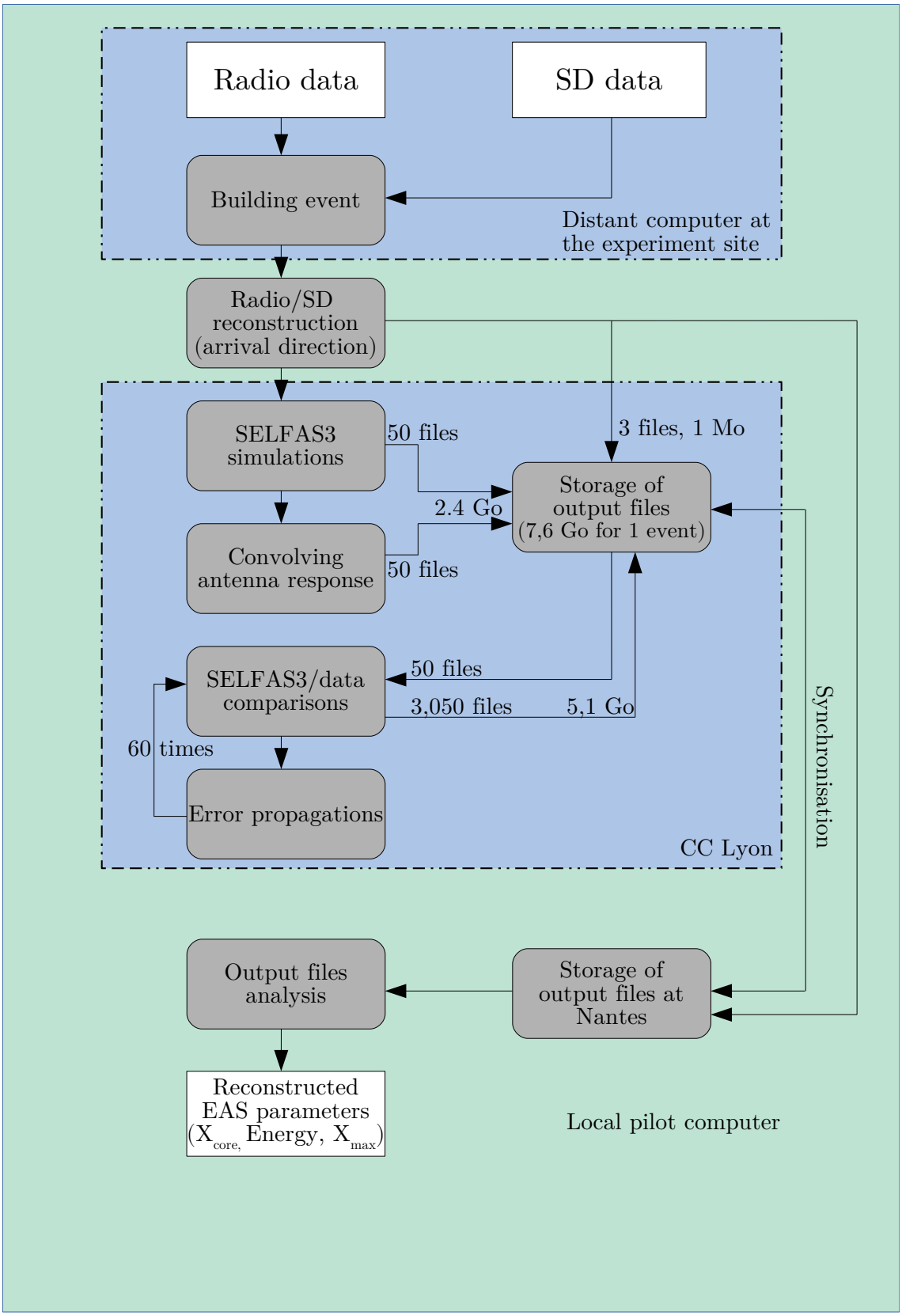


Figure 7.16 *– Analysis pipeline of the automating of the reconstruction of the shower properties: X_{max}, shower core and energy.*

"They didn't know it was impossible so they did it."
Mark Twain (1835-1910)

CHAPTER 8

Main features of low-frequency radio emission of Extensive Air Showers

Outline

Parts of the results of this chapter have been presented in ICRC2017, UHECR2018, TeVPA2018, URSI-France 2019 and ICRC2019 conferences, and published in [53, 146, 159, 212].

Part of this chapter has been published in: . . ., **A. Escudie et al.**, *Radio detection of cosmic rays in [1.7-3.7] MHz: the EXTASIS experiment,* Astroparticle Physics [9]. Since its submission, others LF events have been detected. We accordingly update the number of detected events and the induced results.

8.1. General characteristics of the sky at LF

Being extensively used, the MF and EMF bands are nowadays well known in terms of background noise. This is not the case of the LF band and the studies made in the 70's probably need to be reevaluated considering the strong evolution of the radiocommunications over the last 50 years. Consequently, the first work to carry out is to explore the LF environment, in particular the LF sky of our experiment site. The frequency range studied is dominated by the high brightness temperature of the atmospheric noise. In this section, we remind some facts on the ionosphere layers, and then focus on the background noise at LF.

8.1.1. Absorption of radio emission lines at low frequencies in the ionosphere

The overall background radio noise at low frequencies is less important during the day than during the night. This can be explained by the absorption of radio emission lines at these frequencies by the ionosphere. Indeed, as explained in [222], the ionosphere is composed of layers of ionized plasma constrained by the geomagnetic field. The properties of the layers depend on the free electron density, the altitude and the season. Notably, during the night, some layers disappear or shrink, and some others combine. These changes lead to the absorption, refraction, attenuation, depolarization and dispersion of radio emission lines. For example, the D layer of the ionosphere ([70 – 90] km of altitude) absorbs the radio emission lines from 0.3 MHz to 4 MHz, which tends to decrease the noise level in this frequency band. However, the D layer is only present during the day and vanishes or diminishes at night, no longer absorbing the radio emission lines. The behaviour of the D layer is only partly responsible for the daily variation of the noise. The F layer, which is composed of two sub-layers combining at night (from 150 to ~ 470 km of altitude), will tend to increase the sky-wave propagation during the night by reflecting and refracting them. Consequently the noise level after dark increases, due to both the reflection/refraction by the F layer and the disappearance of the D layer. This daily variation is expected to be visible in LF antenna data.

8.1.2. Overall background noise

Between few hundred of kHz and 10 MHz, the background noise is the superposition of three dominant contributions:

— the minimal thermal noise of the atmosphere (in other words, its brightness related to its temperature),
— the galactic radio emission,
— the so-called atmospheric noise, made of the contribution of the electric discharges from thunderstorms at planetary scale — ~ 100 per second —, and of any kind or man-made noise (radio transmitters, anthropic transients).

Thunderstorm discharges and man-made noise are not absorbed by the D atmospheric layer at night, and are reflected and refracted between the ionosphere layers and the ground. This causes an increase of the atmospheric noise during night. Contrarily to the galactic noise, the atmospheric background noise is not predictable and varies as function of the location on Earth, the season and the time of the day (see also section 8.1.3).

Figure 8.1 presents the background noise brightness temperature as a function of the frequency. Data are extracted from the International Telecommunication Union (ITU) [223] and corrected for the site of Nançay. For the atmospheric noise, only the contribution of the electric discharges from thunderstorms is taken into account in this reference, though an estimate of the made-made noise is given for several types of sites.

If there were absolutely no other sources of noise, the absolute limit for a detectable signal would be set by the minimal thermal noise density N_{ref}, defined as

$$N_{\mathrm{ref}} = 10 \log \left(\frac{k T_0}{1 \text{ mW}} \right) = -174 \text{ dBm} \cdot \text{Hz}^{-1} \tag{8.1}$$

where T_0 =290 K is the reference air temperature for a reference power of 1 mW and k is the Boltzmann constant. In all cases this limit is surpassed by the galactic background, until about 150 MHz. Let F_{am} ("am" stands for atmospheric and median) be the difference between the noise coming from thunderstorms and the minimal thermal noise N_{ref}. F_{am} is then expressed as

$$F_{\mathrm{am}} = 10 \log \left(\frac{T_{\mathrm{n}}}{T_0} \right) \quad \text{dB} \tag{8.2}$$

where T_{n} is the temperature of the sky. F_{am} depends on this temperature, on the location of the observing site, on the day-night cycle and on the seasons. At 1 MHz, the daily

variation of F_{am} is 40 dB for winter and 35 dB for summer: this high variability of the atmospheric noise is easily observable with LF antennas. At frequencies above 20 MHz (MF), the observation limit is set by the galactic background noise, whose variations are visible by the CODALEMA antennas for instance: the atmospheric noise is no more dominant whatever the hour of the day. At 4 MHz, the brightness temperature of the atmospheric noise is 20 dB below the galactic noise level in the best daytime case while it is 17 dB above during the night. These are the real detection limits at LF which are anyway surpassed by the man-made noise level even for a quiet receiving site. To sum up, we show in figure 8.1 that for a quiet observation site, the noise limit is set by man-made activities during the day (the galactic noise being barely competitive), while during the night it is dominated by the atmospheric noise, and in all cases the LF noise level is well above the noise at MF. We can assess that the detection of air-shower transients will be complicated

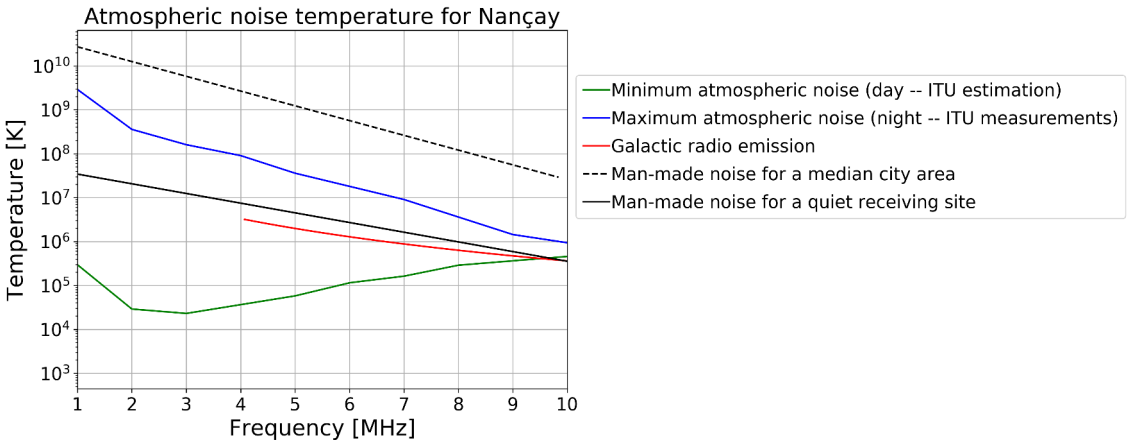


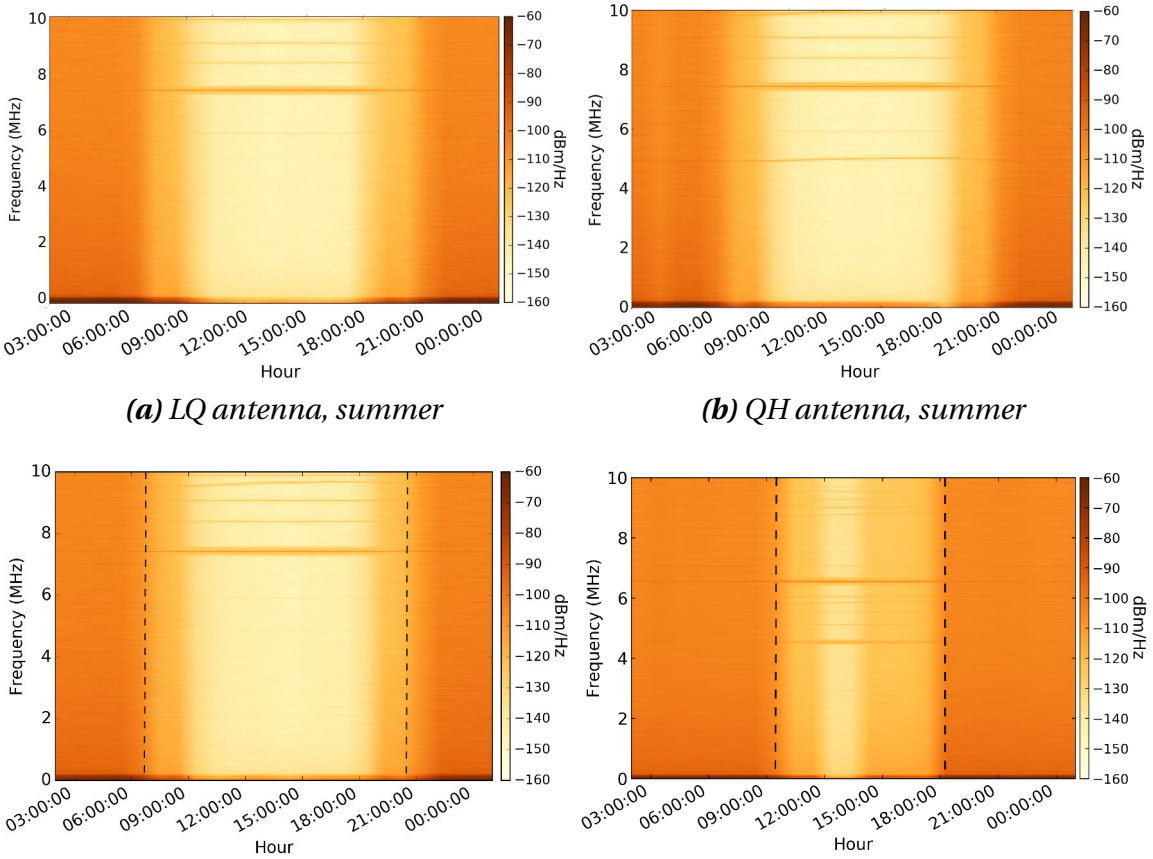
Figure 8.1 *– Atmospheric noise and galactic noise temperature as a function of the frequency, computed from the raw ITU data and corrected for the site of Nançay. Man-made noise temperature estimates have been added for a city and a quiet site. The galactic emission (considered apart from the atmosphere) is computed with the model established by Cane [224], on the basis of ground measurements above* 4 MHz *and satellite measurements below this frequency, where the atmosphere becomes mostly opaque to the incoming radiation. The galactic noise contribution thus experiences a cut-off below* 4 MHz *for realistic observation conditions, and is no longer dominant regarding to the atmospheric noise and man-made noise for a quiet receiving site.*

at LF, especially if their strength does not increase when the frequency drecreases, as it will be shown in the following.

8.1.3. Power spectrum density of the low frequency sky at Nançay

Figure 8.2 presents time-frequency diagrams for different LF antennas of EXTASIS. The time-frequency diagram consists in a daytime power spectrum density (PSD, in $\mathrm{dBm \cdot Hz^{-1}}$). The darker the color, the noisier the environment. The diagram gives a view of the environment of each antenna, which varies with its position (pictures (a) to (c)) and during one day, regardless of the position (day/night dependence is well observed). Moreover, a seasonal variation can also be seen, as shown in pictures (c) and (d), which presents time/frequency diagrams for one LF antenna taken during summer (c) and winter (d). The vertical black-dashed lines represent the sunrise and sunset times, delimiting night time and daytime. Of course, the two time periods do not have the same duration depending on the season, the duty cycle being maximal during summer and minimal during winter.

From figure 8.1, one could wonder whether the Nançay site is classified as a quiet (rural) or urban site, which would drastically hamper any observation at LF. Figure 8.3

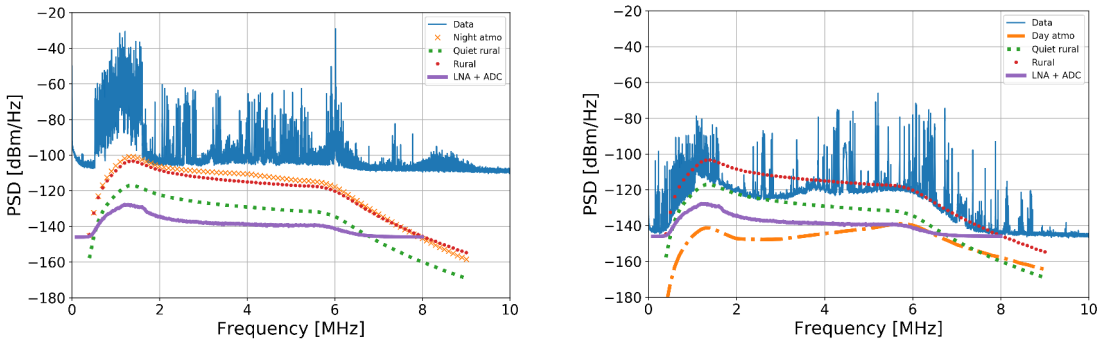


(a) *LQ antenna, summer* ***(b)*** *QH antenna, summer*

(c) *HL antenna, summer* ***(d)*** *HL antenna, winter*

Figure 8.2 *– Time-frequency diagram for different LF antennas. Fig. (a), (b) and (c) report environmental variations due to the location of the antennas, while Fig. (c) and (d) report a seasonal variation for the same antenna. The vertical black-dashed lines represent the sunrise and sunset. See text for details.*

presents the PSD of one LF antenna as recorded by our analog and digitization chain, averaged over 200 events during nighttime (a) and daytime (b), for the same day and same antenna than in figure 8.2.(a). The atmospheric and estimated man-made noises of figure 8.1 (calculated at the ADC input from the antenna simulations and the measured transfer function of the analog chain) are plotted along with the noise from the electronic chain. A lot of man-made broadcasting radio emissions (RFI) are present in the LF band, day and night. However, during daytime, the quiet rural noise level expected from ITU specifications is reached in most of the [1.7 – 3.7] MHz band, which contains less RFI and presents a noise floor 20 dBm $\cdot$ Hz^{-1} below the noise floor during nighttime. This is the reason why we have chosen to restrict our observations to this band. From these daytime PSDs, one can conclude that the Nançay site can be considered as a quiet rural site regarding the specifications of ITU. We also observe that the LNA noise is at least 15 dB lower than the minimum noise, which is not the case for the horizontal polarization (not shown here) for which the limitation is given by the LNA noise from 1.5 to 2.1 MHz. Therefore for the vertical polarization, the minimum signal to be detected is not limited by the sensitivity of our detector. It is worth noticing that, in [1.7 – 3.7] MHz, the root mean square (rms) of the noise at night is $\sim$ 100 times higher than the noise during the day. It means that, to be detected during the night, a pulse should be 100 times higher than a pulse observed during the day. In fact, as it will be shown in section 8.3.1, no valuable observations have been made during night time.



(a) *Nighttime PSD.* ***(b)*** *Daytime PSD.*

Figure 8.3 *– Night (left) and day (right) PSD of one LF antenna in vertical polarization, averaged over 200 events, for the same day than figure 8.2.(a). The atmospheric noise of figure 8.1 is shown together with the noise from the electronic chain. The figure on the right shows that the Nançay site can be considered as a quiet rural site up to 3.2 MHz.*

8.2. Results

In order to reach the EXTASIS objective to detect and exploit LF signals, data from the CODALEMA instruments and EXTASIS were combined. In the following, we first present how a complete cosmic-ray event is detected and analysed, and then we illustrate a LF signal detection.

8.2.1. Event reconstruction

This section aims to describe the selection procedure of the events seen simultaneously by multiple instruments of CODALEMA and EXTASIS, and the elaboration of a so-called “cosmic ray event” associating several of these instruments. As it has been discussed in section 4.3.3, the trigger is generated by the particle detector and distributed to the Compact Array and to the LF antennas. For the 57 standalone antennas, no particle trigger is sent. Transients detected in coincidence on several of these standalone antennas build a “radio coincidence”, characterized by an average radio event time that will be compared to the particle detector event. The criterion is that at least three standalone antennas are triggered within a time interval compatible with the propagation of a plane wave at the speed of light. The radio event is promoted as an actual shower if its timing is compatible with the timing of the scintillators and if the reconstructed arrival directions agree within 20° [173].

Since the installation of the complete instrumental setup of EXTASIS (March 2017) and until the date of writing of this manuscript, 1,660 cosmic events have been detected by the CODALEMA standalone antennas, with a potentially detectable LF signal. Among them, 633 present a pulse detected by the LPC method on at least one LF antenna. We checked that LF events with a multiplicity of 1 and 2 are mostly fortuitous, their detection time being incompatible with the expected time deduced from the MF signal reconstruction. Moreover, LF events with a multiplicity lower than 3 are not interesting for us in this study, where we aim to know whether the LF counterpart can improve the shower reconstruction or not. We thus decided to ignore LF events with a multiplicity lower than 3, knowing also that a large majority of them are probably accidentals, taking into account the transient noise rate and the wide time window of search. For LF events with a multiplicity of 3, the coincidence rate falls dramatically, since only 25 of these 1660 events present a LF counterpart according to the LPC method (see section 8.3). We

apply a conservative angular criterion on the selection of the LF events, because the uncertainties on the directions of arrival are estimated around ~ 10° (compared to ~ 1° for the standalone antennas), mostly due to the transient duration in the LF band. It is nevertheless important to note that there is a good agreement between the arrival directions estimated by the LF and by the MF antennas.

These events are summarized in table 8.1. They have a timing in agreement with the usual radio emission when the shower develops in the air and not with what we expect from the shower sudden death (corresponding to a good agreement between the arrival directions). Unfortunately, the next reconstruction steps (core location, X_{max} and energy estimation from the MF signals) can not be performed for most of these events for various reasons:

— a core location clearly well outside of the CODALEMA array requires an hazardous extrapolation of the electric field predicted by SELFAS3,
— a low multiplicity leads to a very poor accuracy on the shower parameters,
— an abnormal and undetermined atmospheric electric field kills the correlation between the electric field measured in the antennas and the shower development.

As a consequence, the fact that the energy can not be estimated in the EMF band for these 25 events does not allow to deduce properties on the amplitude of the LF signal. The unique event (number 4) for which the reconstruction appears reliable is detailed in the following section.

#	N_{SA}	θ_{SA} (°)	ϕ_{SA} (°)	$\Delta\alpha$ (°)	E_{stat}(kV · m^{-1})	Probability
1	6	31.8	353.1	2.1	7.5	1.91×10^{-3}
2	20	60.0	153.8	6.0	0.3	2.85×10^{-1}
3	11	28.8	68.9	3.2	1.8	6.35×10^{-3}
4	11	40.6	145.2	11.3	-0.1	1.14×10^{-1}
5	13	40.9	32.6	20.4	-0.1	1.14×10^{-1}
6	8	56.1	58.9	3.4	0	7.42×10^{-1}
7	3	34.2	252.9	3.8	0.3	2.85×10^{-1}
8	4	53.4	95.9	13.3	0.1	4.97×10^{-1}
9	4	44.4	76.6	22.9	0.1	4.97×10^{-1}
10	7	16.2	210.8	6.3	-12.3	1.19×10^{-3}
11	25	38.4	206.5	3.4	-15.7	5.09×10^{-4}
12	3	77.7	14.5	9.1	0.1	4.97×10^{-1}
13	5	24.5	235.3	3.7	-0.5	3.23×10^{-2}
14	22	23.0	92.3	6.0	-11.1	1.51×10^{-3}
15	21	46.5	109.8	5.0	-3.7	7.31×10^{-3}
16	20	25.9	74.3	14.1	-7.6	3.07×10^{-3}
17	24	33.5	23.4	10.3	2.1	5.25×10^{-3}
18	8	7.4	203.8	1.5	6.5	2.20×10^{-3}
19	5	51.7	346.4	3.4	0.1	4.97×10^{-1}
20	7	44.7	45.2	3.3	0.2	4.40×10^{-1}
21	45	61.0	34.2	7.0	-0.3	4.23×10^{-2}
22	9	51.2	150.0	15.6	-0.1	1.15×10^{-1}
23	6	7.7	265.3	2.2	-7.6	2.87×10^{-3}
24	15	41.1	237.0	6.7	-6.4	3.72×10^{-3}
25	28	31.0	165.1	12.1	-1.3	1.89×10^{-2}

Table 8.1 *– List of the 25 LF events selected. The first column is the event identification number. The second one gives the multiplicity of the standalone antennas of the event. The third and fourth columns give respectively the zenith and azimuth angles of the event reconstructed by the standalone antennas. The fifth column gives the 3D-angular difference of the arrival direction of the event reconstructed by the MF and LF antennas. The sixth column indicates the value of the atmospheric electric field recorded during the detection of the event, and the corresponding probability (see section 8.3.2).*

8.2.2. Example of low-frequency signal detections

A ground map of the event 4 of table 8.1 is shown in figure 8.4. Eleven standalone antennas (circles) at the south part of the MF array have recorded a signal. LF counterparts were registered in four LF antennas. The small green lines close to the circles indicate the orientation of the polarization of each MF antenna, expected to be nearly orthogonal to the direction of arrival of the event.

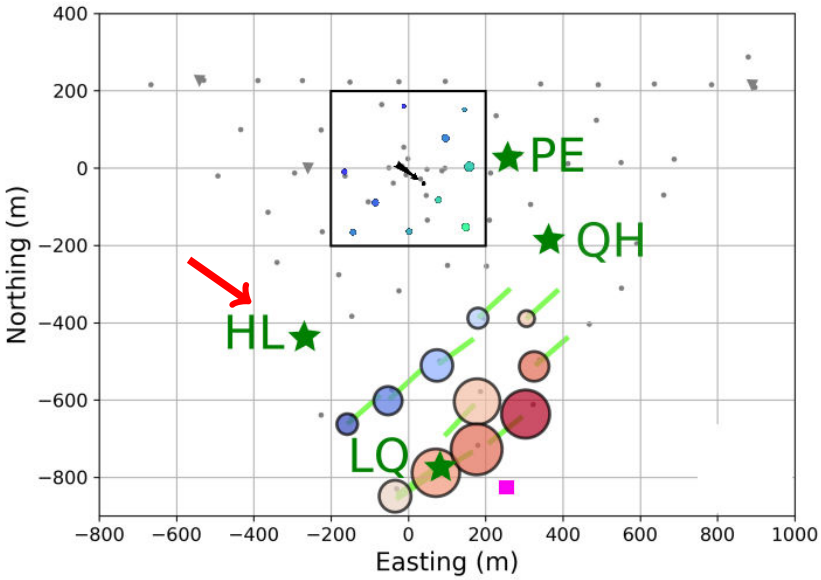
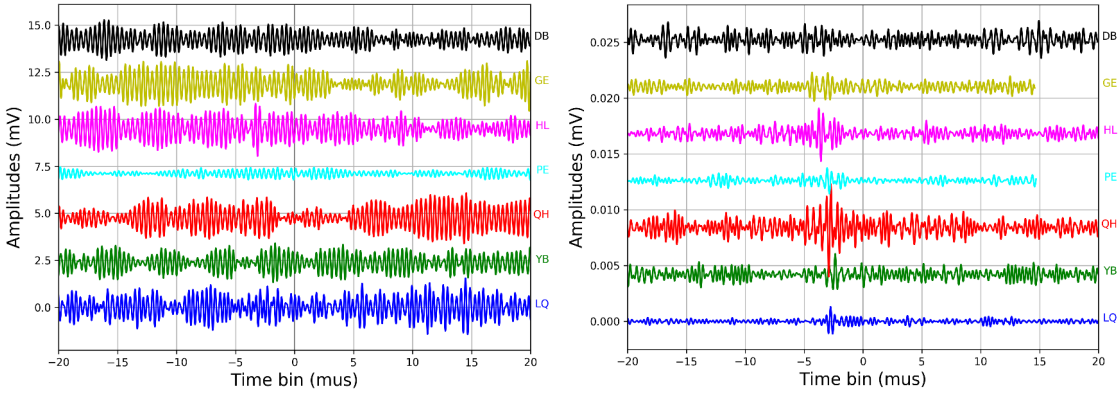


Figure 8.4 *– Footprint of the event 4 seen by the particle detector and some LF and MF antennas. The arrival direction is represented by the red arrow, and its energy* $(3.7 \pm 0.6) \times 10^{18}$ eV*, from the CODALEMA standalone antenna reconstruction. Grey dots represent the standalone antennas, the square area represents the scintillators array region. The green stars indicate the positions of the LF antennas that detected the shower. The involved standalone antennas are represented by coloured circles, whose colour indicates the timing order in which the signal has been seen by the antennas (from blue, earliest, to red, latest) and area of circles reflects the relative amplitude of the signal (linear scale). The small green lines close to the circles indicate the orientation of the measured polarization of each MF antenna, nearly orthogonal to the direction of arrival of the event as expected from the dominant geomagnetic mechanism. The estimated shower core location is represented by the magenta square (*$x = 259 \pm 35$ m *and* $y = -809 \pm 30$ m*). See text for more details on the cosmic ray and shower properties of this event.*

The recorded traces are shown in figure 8.5-(a), ordered by time of arrival in the LF antennas. The pulses located around −2.5 µs on GE and YB antennas (2^{nd} and 6^{th} traces from top) are fortuitous transients, rejected by both the LPC procedure and by the Direction Of Arrival (DOA) reconstruction. We see that the transients are not visible on the filtered traces (figure 8.5-(a)), but they appear after the LPC processing (figure 8.5-(b)).

Extensive, iterative and systematic comparisons of the EMF signals with SELFAS3 simulations were performed to deduce the cosmic ray and shower properties. It includes antenna by antenna spectrum comparisons (on both polarizations) over a substantial set of simulations spanning a large range in core position, shower X_{max} and cosmic-ray energy (see [71, 213] for more explanations on the method). Moreover, the amplitude and spectral index variations observed in the eleven MF antennas are well reproduced by the simulation selected at the end of this iterative process. It gives us a strong confidence on the event reconstruction reliability. The best core position, represented by a magenta square in figure 8.4, is $x = 259 \pm 35$ m and $y = -809 \pm 30$ m (our reference position is located at the center of the particle detector array). The method also gives an estimate



(a) *Filtered* [1.7 – 3.7] MHz *LF signals.* ***(b)*** *Error of prediction on LF signals.*

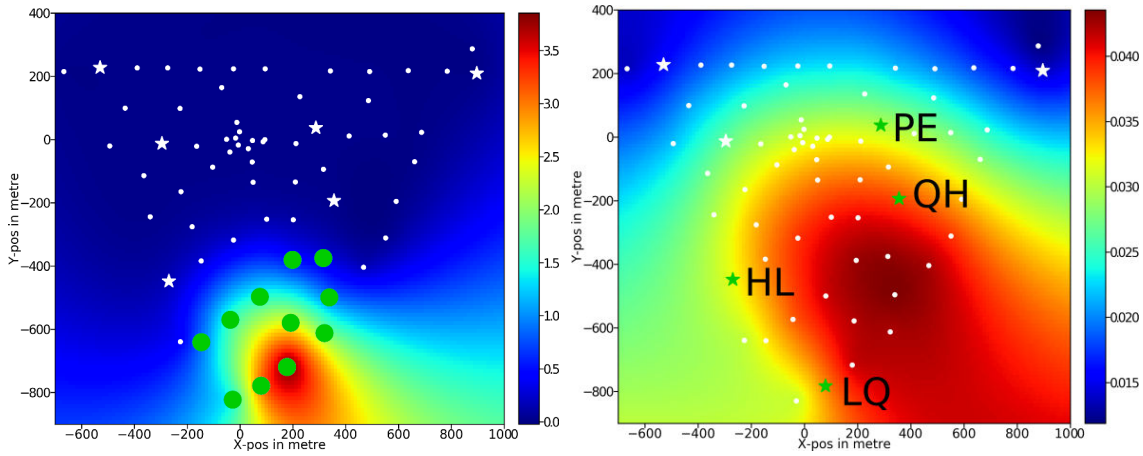
Figure 8.5 *– LF events seen in the horizontal polarizations. Traces are positively shifted along y for clarity. Left: LF signals, as a function of time, filtered in* [1.7 – 3.7] MHz*, ordered by time of arrival of the cosmic ray signal in antennas. Right: error of prediction of LF signals. Actual cosmic ray transients are detected on traces 3, 4, 5 and 7 from top, in the time window* [−5; 0] µs *after applying a simple threshold method on the LPC prediction error. DB (trace 1) did not detect any transient. Although transients are visible in traces 2 and 6, located at a time not compatible with the shower geometry.*

of $X_{\max}$ of 715 ± 19 g $\cdot$ cm^{-2} and an energy of $(3.7 \pm 0.6) \times 10^{18}$ eV. The latter is in good agreement with the energy estimated by the particle detector of $(2.75 \pm 1.05) \times 10^{18}$ eV. This event core location being external to the particle detector area, the shower core location determined by the radio method has been used for the particle detector energy reconstruction. This explains the large uncertainty on the energy estimated with the scintillator data.

Figure 8.6 presents the interpolated simulated electric field of the horizontal polarization in [30 – 80] MHz and [1.7 – 3.7] MHz. At low frequency (figure 8.6, right), the electric field distribution appears much wider and flatter than at medium frequency (left), with a considerably increased detection range. Indeed, the LF antenna PE located around ($x = 300$ m ; $y = 20$ m), see figure 4.21, has detected the shower at 850 m from the reconstructed shower core location, while the most distant MF antenna is only at 400 m from the latter. There is no MF counterpart in the standalone antenna associated with PE. This hints an electric field detection threshold of about 23 ± 4 µV $\cdot$ m^{-1} at low frequency in the horizontal polarization, the value detected on the PE antenna after correction for the antenna equivalent length and acquisition chain gains: the GE antenna, located at ($x = -250$ m ; $y = 0$ m) more or less at the same distance of the shower core, has not detected the simulated electric field of 23 µV $\cdot$ m^{-1}.

Moreover, as it can be seen in figure 8.6 where the color scale is expressed in mV $\cdot$ m^{-1}, the electric field in the LF band is actually smaller than in the EMF band. This result disagrees with the pioneer observations, which reported that when frequency decreases, a clear evidence of a strong increase of the radio pulse amplitude was seen.

Figure 8.7-left shows the simulated PSD as a function of frequency and for different LF antenna locations, in color for the involved LF antennas and in black for the others. The PSD quickly drops in the EMF band with the shower axis distance, while it decreases much more slowly in the LF band. Figure 8.7-right presents the PSD of the signal of the shower development over the whole frequency band. LF data are represented by the green line, EMF data by the blue line, and in red and black dashed lines the convoluted simulated power spectrum density in LF and EMF band respectively, in which we have added the noise of the corresponding band. The noise-added, convoluted simulations



(a) [30 – 80] MHz. ***(b)*** [1.7 – 3.7] MHz.

Figure 8.6 – *Lateral distribution of the electric field depending on the frequency range predicted by SELFAS3. Left:* [30 – 80] MHz. *Green circles represent the involved MF antennas in the event. Right:* [1.7 – 3.7] MHz. *Stars represent the LF antennas, green ones correspond to the involved LF antennas in the event. The LF antenna (PE) located around* (x = 300 m ; y = 20 m), 850 m *from the shower core location at ground, gives the extent of the detection zone at low frequency. The color scale, expressed in* $\mathrm{mV \cdot m^{-1}}$, *is not the same for the two plots: the detected electric field in the LF band is smaller.*

are in good agreement with the data, showing a good understanding of our LF and MF instruments, but also a good radio reconstruction of the characteristics of the primary cosmic ray. This result indicates again that the detection range should be larger at the LF

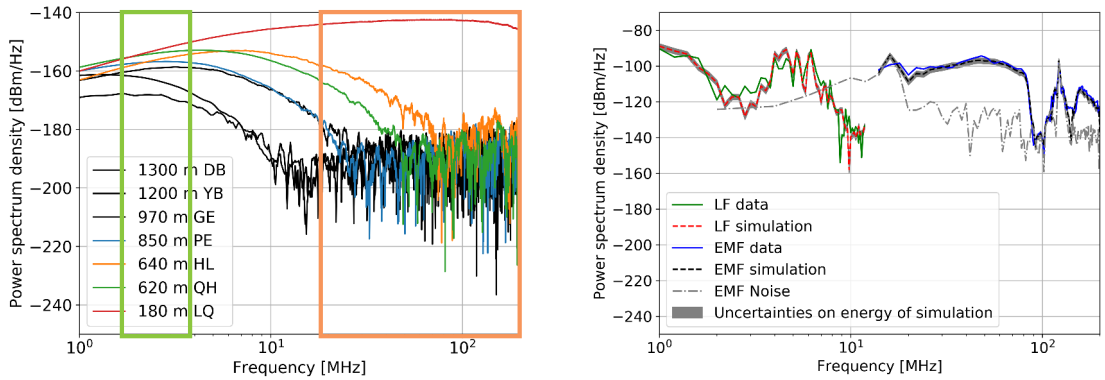


Figure 8.7 – *Left: simulated power spectrum density as a function of frequency calculated at the LF antenna locations, in colour for the involved LF antennas and in black for the others. Distance to shower axis is also indicated. The green band indicates the range of the LF band, and the peach band indicates the range of EMF band. Right: convoluted power spectrum density as a function of frequency for the southernmost LF and MF antennas, labelled LQ on the left figure.*

band than in the EMF band. Our data confirm this expectation. One way to quantify the detection range is to consider the axis distance. For a given event, if we know the core position, we can compute the axis distances between each detector and the shower axis. The maximum value of these axis distances is the maximum axis distance for this event and is an indication of the detection range. We don't have a proper core reconstruction

for our 25 events but we can define by eye a confidence zone where it should be, as suggested by the ground pattern of the triggered MF detectors. We used such circular zones for each event, with a confidence radius varying from 200 m (for internal events) up to 600 m (for external events). This is a conservative way to get an estimate of the core position and, consecutively, of the axis distances. Then, we compute the average and rms of the maximum axis distance for each event using a large number of sample core positions, taken inside the pre-defined confidence zone. These values are shown in figure 8.8, in black and red for the MF detectors and LF detectors, respectively. We see that the maximum axis distance is larger for LF detectors than for MF detectors, for almost all events. This was expected since a long time through the various simulations reported in the literature, but this is the first time that it is confirmed by an actual detection.

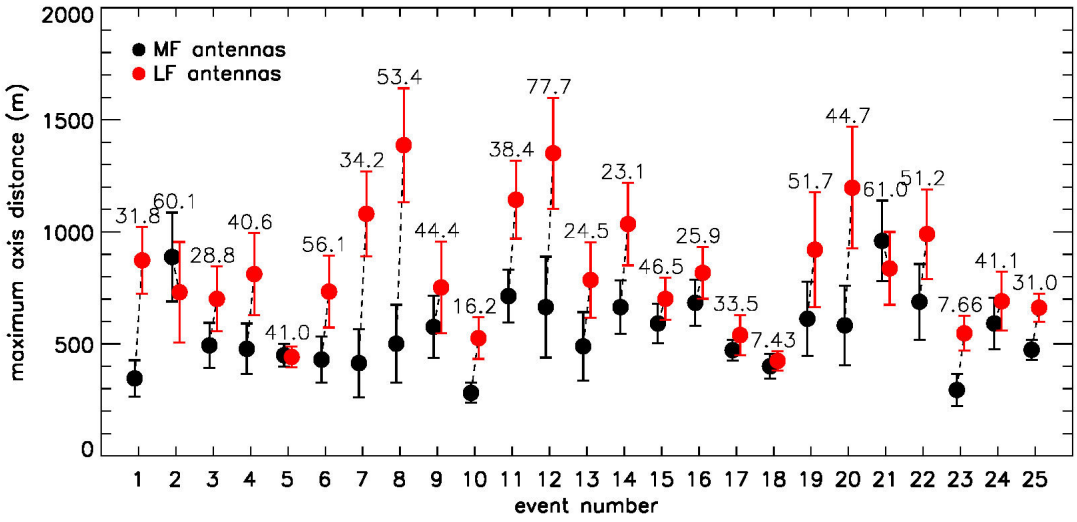


Figure 8.8– *Maximum axis distances for the LF and MF detectors of the 25 events of table 8.1, illustrating that the detection range is higher for LF detectors that the MF detectors. The zenith angles of the events are indicated at the top of the error bars. Large error bars correspond to external events with a large uncertainty on the core position. See text for details.*

At last, for the LF antenna in the MF zone (LQ), and thus at a given shower axis distance, there are 10 dBm $\cdot$ Hz^{-1} between the maximum in the EMF band and the maximum in the LF band, showing that the signal is larger in the classical band than in the LF band. For the southernmost LF antenna and its MF companion (LQ), the simulated power spectrum density has been convoluted with the antenna and acquisition chain responses to obtain the equivalent in ADC counts, in order to compare them with the raw data.

8.3. Discussions

8.3.1. How to explain the low detection rate in the low-frequency band?

As already mentioned in section 8.1.2, the LF sky is dominated by the atmospheric noise and the noise level at night is $\sim$ 100 times higher than during the day. Consequently, over one year of observation the duty cycle is reduced by a factor of 2. This seasonal variation shown in figures 8.2.(c) and (d), which considerably reduces the available daily time, makes a LF detection during winter highly unlikely.
Due to the noise, the signal to noise ratio is expected to be much smaller at LF than at MF. As an illustration, let us study the event shown in figure 8.4. For this event, accurate simulations have been carried out as explained in section 8.2.2 and the response of the LF

antennas has been taken into account, by convolving the simulations. Figure 8.9 depicts the detected signals for the southernmost LF antenna (blue curve) and its simulation (orange curve). This antenna has the highest detected signal of the event. The expected signal has been superimposed at the time bin where the actual signal has been detected.

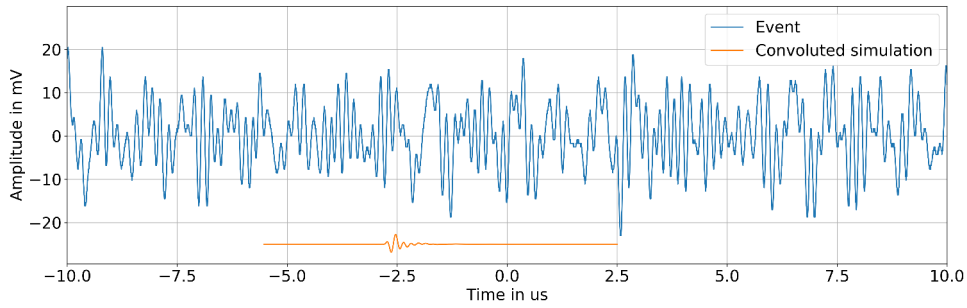


Figure 8.9 *– Blue line: full band signal of the LF antenna closest to the shower core of event of figure 8.4. Orange line: simulation of the signal at the same location and in the same frequency band, convoluted with the antenna response and placed at the time bin where the actual signal has been detected (shifted downwards for visibility).*

The rms of the background noise is 10 times larger than the amplitude of the convoluted signal, explaining why it is not visible by eye. However, it has been detected using the LPC method. With the LPC method (orange line of figure 6.20 in section 6.4.5 page 149), the detection efficiency is around 50 % for an amplitude one order of magnitude below the noise rms. For the considered event, 4 antennas over 7 present a signal after the LPC processing. For the closest of the three antennas without detection, the transient amplitude is estimated from the simulations to be equal to 1 % of the rms of the background noise (figure 8.10). For this amplitude level, the detection efficiency is much smaller than 50 %, explaining why the transient can not be seen even with the LPC method. These observations permit to determine the minimum amplitude of a detectable signal compared to the background noise. In that case, the signal in the farthest LF antenna that has detected the shower has an amplitude of 20 % of the rms of the background noise.

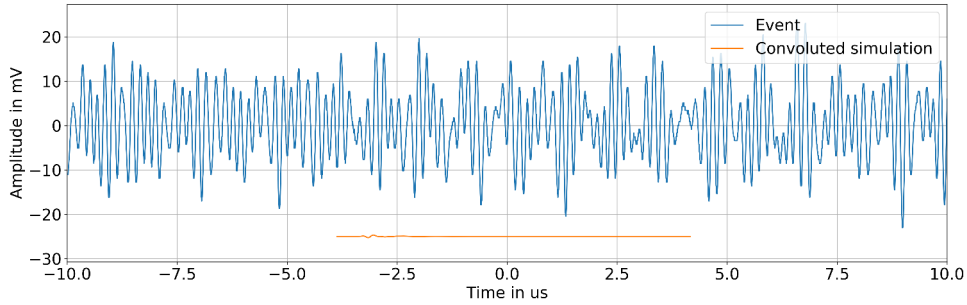


Figure 8.10 *– Same as figure 8.9 for the closest LF antenna without detection. Barely visible, the transient amplitude (around* −3 µs*) is equal to 1 % of the background noise RMS.*

The low detection rate of LF signals can be thus explained by the atmospheric background noise level which, in the best case, is 10 times higher than the expected minimum detectable signal and makes the detection of the LF counterpart of the shower development unlikely, at least for the shower properties expected at the CODALEMA site. Moreover, as it will be explained in the next sub-section, it is likely that even fewer showers would have been detected if particular atmospheric conditions did not probably amplify the radio signal.

8.3.2. Correlation with the atmospheric electric field

A static electric field sensor is installed at the CODALEMA site, giving every 3 seconds the value of the static vertical component of the atmospheric electric field ξ. In normal conditions (fair weather), the value of the atmospheric electric field is around $140\,\mathrm{V}\cdot\mathrm{m}^{-1}$. In thunderstorm conditions, the absolute value can reach $10^5\,\mathrm{V}\cdot\mathrm{m}^{-1}$ at ground level. The probability density function of the atmospheric electric field values is presented in figure 8.11. For each of the 25 LF events listed in table 8.1, we know the atmospheric

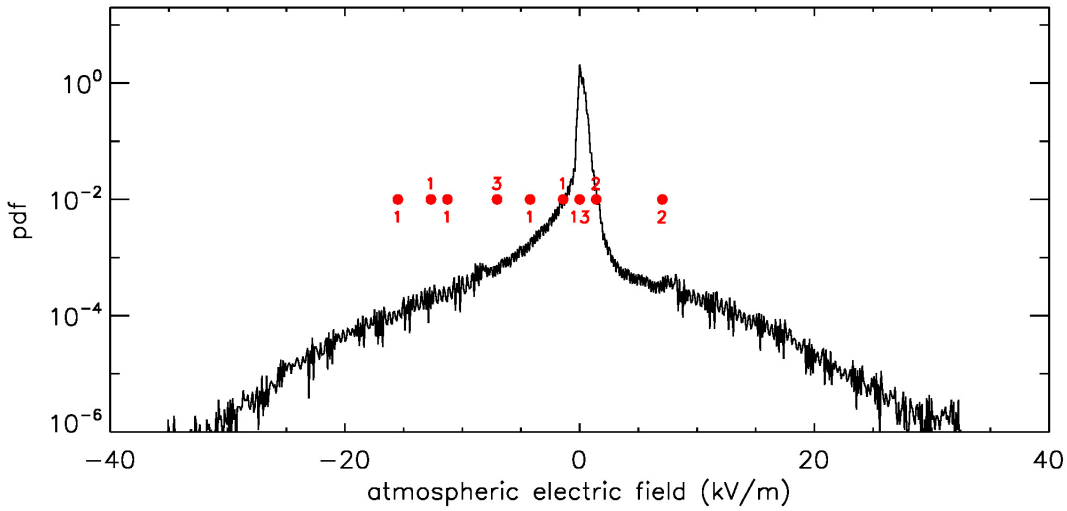


Figure 8.11 *– In black: probability density function of the atmospheric electric field measurements carried out since 2014/09/01, in Nançay (bin size is* $100\,\mathrm{V}\cdot\mathrm{m}^{-1}$*). The atmospheric electric field values at the time of the 25 LF events are represented by the red dots, with the number of corresponding events (bin size is* $1.4\,\mathrm{kV}\cdot\mathrm{m}^{-1}$*).*

electric field value within less than 1.5 s of the time of the event; these values are displayed in red in figure 8.11, together with the number of LF events in the corresponding bin (bin size is $1.4\,\mathrm{kV}\cdot\mathrm{m}^{-1}$). Ten of them correspond to thunderstorm conditions, i.e. outside of the main peak. Using the probability distribution function, we can compute the following probabilities, at the time of the LF event detection: $P(|\xi| > |\xi_{\mathrm{atm}}(t_{\mathrm{LFevent}})|)$. These probabilities are shown in the 6$^{\mathrm{th}}$ column of table 8.1. In order to check whether the atmospheric electric field values at the time of detection of the LF events are compatible or not with the global probability density function, we perform the Fischer combined probability test [225, 226]. We find that the χ^2_{data} associated to the 25 individual probabilities is 179.8. According to Fischer, this value is extracted from a χ^2 law with 36 degrees of freedom. The p-value of χ^2_{data} is 1.1×10^{-16}. The conclusion is that the high values of the atmospheric electric field at the time of detection of the LF events are not compatible with a random coincidence: the LF detection of cosmic rays is strongly favored by thunderstorm conditions.

As an example, let us consider one of the 25 LF events detected during storm conditions. During the day of the event, the atmospheric electric field had a chaotic behavior from 09:00 to 18:00, exhibiting large electric field values. Around the time of the event (11:28), the atmospheric electric field was equal to $-12.3\,\mathrm{kV}\cdot\mathrm{m}^{-1}$, about 20 σ from the average value during normal conditions ($140\,\mathrm{V}\cdot\mathrm{m}^{-1}$). It is worth noticing that the event presented in figure 8.4 was detected under normal atmospheric electric field conditions. Furthermore, any abnormal atmospheric electric field would strongly complicate the analysis comparing the observed MF signals to the simulated ones since the latter would require to perform shower simulations assuming a minimum knowledge of the atmo-

spheric electric field profile as a function of altitude.

As already observed in the past [227], it is likely that the radio signal experiences an amplification due to the local atmospheric electric field, making it possible to be detected even for low energy showers. This amplification of the LF signal due to thunderstorm conditions could be at the origin of the large electric field values recorded at the time of the pioneer experiments though, apart in [181], atmospheric electric field conditions are not mentioned, making it impossible to confirm this hypothesis.

8.3.3. How to explain the non-observation of the SDP at Nançay?

A strong contribution due to the sudden death of the shower could be obtained only if a lot of particles reach the ground, as shown in figure 8.12-bottom. This figure was obtained by simulating seven proton showers per bin of energy and zenith angle, assuming an altitude of 130 m corresponding to the Nançay site. At fixed energy, the number of particles reaching the ground decreases with increasing zenith angle. For example, for a primary energy around 3×10^{18} eV, and for vertical shower ($\theta \leqslant 10°$), the number of particles reaching the ground is of the order of 10^9. We can infer that the sudden death signal for an event such as the one shown in the previous section ($E = 4 \times 10^{18}$ eV, $\theta = 41°$) should not have been expected, because of a too small number of particles reaching the ground, estimated to be less than 6×10^8. This is confirmed by Figure 8.12-top, featuring the total expected amplitude of the SDP for an antenna at 200 m north of the shower core as a function of the primary energy and the shower zenith angle for the altitude of Nançay. Both figures are similar, corroborating the fact that the amplitude of the signal is directly related to the number of particles reaching the ground. This analysis shows that the SDP could be detectable for a number of particles at ground level larger than 10^9. If we consider the detection threshold at low frequency as previously estimated, i.e. 23 $\mu V \cdot m^{-1}$, showers giving a detectable SDP at the altitude of Nançay should have an energy larger than 4×10^{18} eV and a zenith angle smaller than the value indicated by the red dashed-line. We expect of the order of 0.3 shower per year having these characteristics at Nançay (assuming a duty cycle of 50 % due to the day/night effect). This considerably hampers the possibility of detection and thus the confirmation of the existence of the sudden death phenomenon. However, as shown in figure 8.13, the observation of the SDP could be significantly easier with LF antennas installed at places of higher altitudes such as the Pierre Auger Observatory (1400 m, Figure 8.13-top) or even better the IceTop site (2800 m, Figure 8.13-bottom). In Figure 8.13-bottom we display the 23 $\mu V \cdot m^{-1}$ contour for IceTop but we also added the same contours for the altitudes of Auger (1400 m) and Nançay (130 m). Going to higher altitudes implies a much larger number of particles at ground for showers of a few EeV, which considerably increases the chances of observation of this phenomenon. For fixed area and observation time, we give in table 8.2 the ratio of the number of detectable showers as a function of zenith angle.

Zenith angle	0°	30°	50°
Ratio IceTop/Auger	1.6	2.7	30
Ratio IceTop/EXTASIS	4	16	342
Ratio Auger/EXTASIS	2.5	6	11.4

Table 8.2 *– Ratio of the number of detectable events for the sites of IceTop, Auger and EXTASIS at fixed area and observation time as a function of zenith angle, assuming the same detection threshold of* 23 $\mu V \cdot m^{-1}$.

These numbers are due to the evolution of the total number of secondary particles reaching the ground level as a function of the zenith angle and the observation site altitude. For instance, Figure 8.14 presents this number as a function of the zenith angle at the Auger and IceTop sites for different energies. If we take into account the zenithal

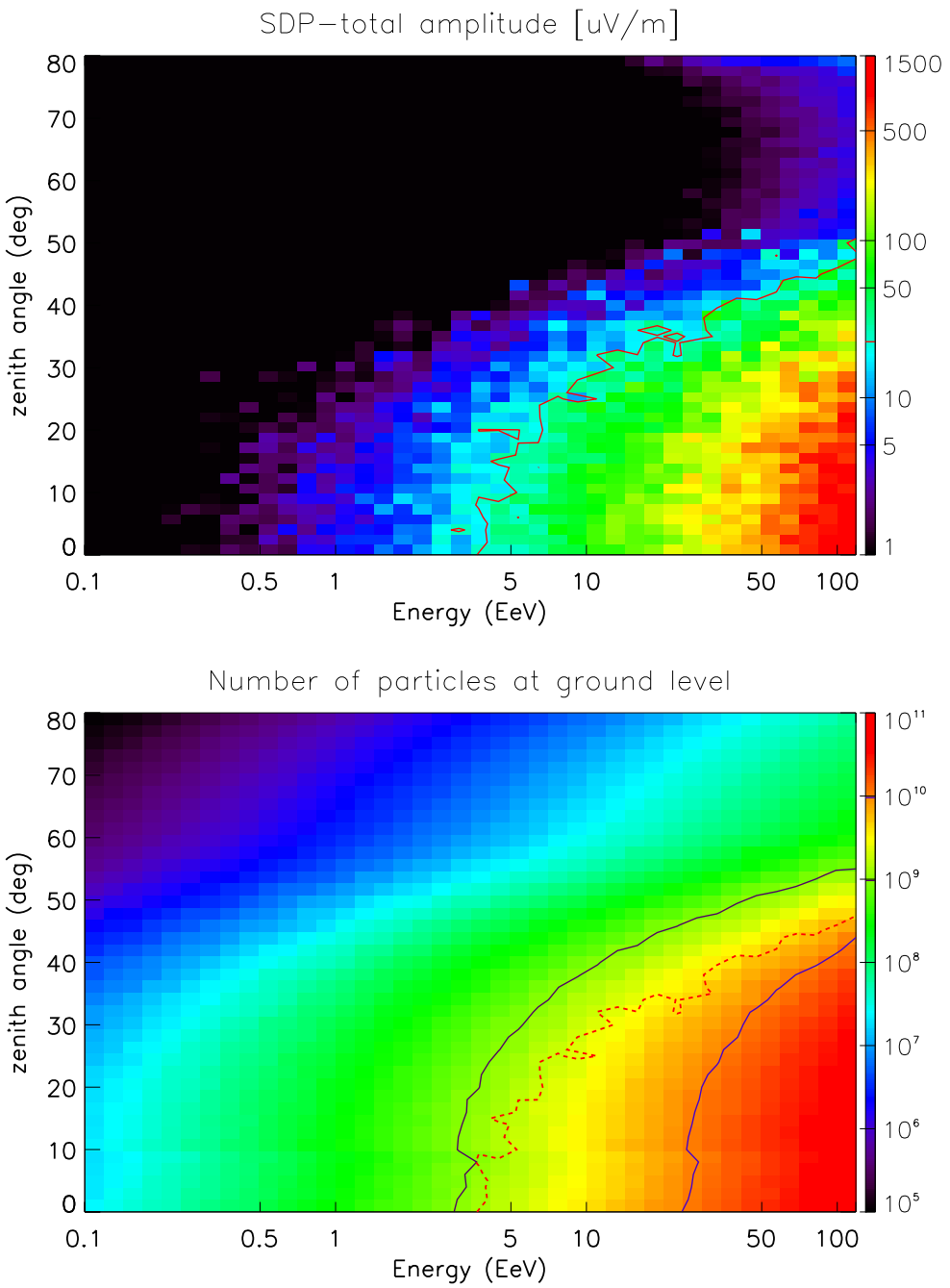


Figure 8.12 *– Top: SDP total amplitude as a function of the primary proton energy and shower zenith angle for the altitude of Nançay and for an antenna at* 200 m *north of the shower core. Each bin contains the amplitude averaged over 5 showers. The frequency band is* [1 − 10] MHz. *The red contour corresponds to the detection threshold (*23 µV · m^{-1}*). Bottom: number of particles reaching the ground at CODALEMA. Each bin contains the number averaged over* 7 *showers. The two solid contours correspond to* 10^9 *and* 10^{10} *particles. The dashed contour corresponds to the detection threshold of* 23 µV · m^{-1}.

acceptances of Figure 8.13-bottom, an effective detection area of 1 km^2, a duty cycle of 50 % and an integration time of 1 year, then the number of showers that could be observed with the sudden death signal is: 0.33 at the altitude of EXTASIS, 0.89 at the altitude of Auger and 1.75 at the altitude of IceTop (see appendix F for the detail of the calculation). We considered the spectral indices of the cosmic ray flux measured by the Telescope Array experiment given in [228]. The size of IceTop and EXTASIS are roughly the same,

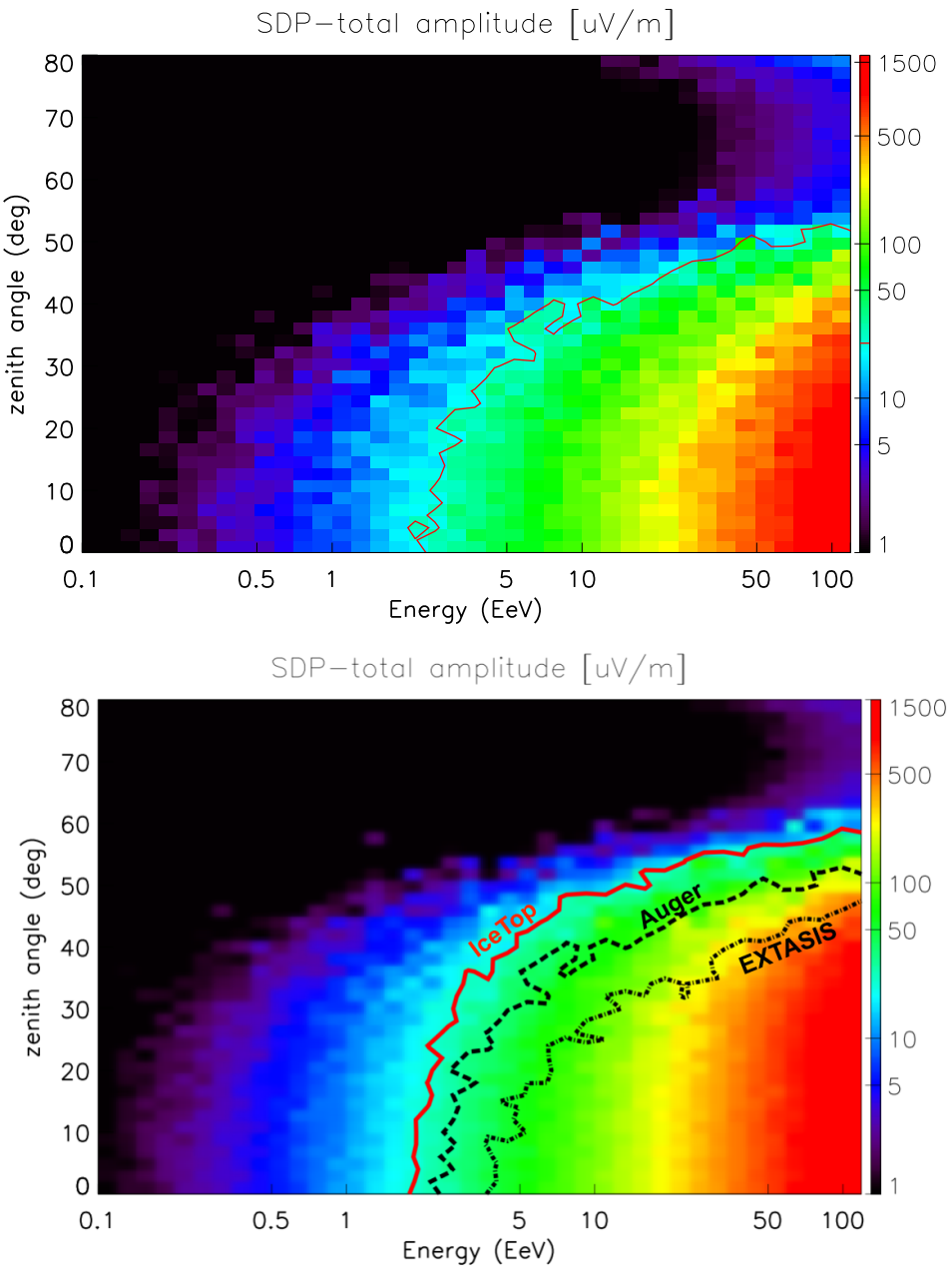


Figure 8.13 *– Top: same as Figure 8.12-top but with a ground altitude of* 1400 m *corresponding to the altitude of the Pierre Auger Observatory site; the contour corresponds to the detection threshold of* 23 $\mu V \cdot m^{-1}$. *Bottom: same figure for a ground altitude of* 2800 m *corresponding to the IceTop site. For an easier comparison, the contours obtained for the altitudes of EXTASIS (dot-dash line) and Auger (dash line) have been superimposed.*

1 km^2, so that the search for the sudden death signal seems to be compromised. The Auger site is much more interesting as we can expect of the order of 2,600 events per year (using 3,000 km^2). A site at 3000 m of altitude and covering an area of 200 km^2 (such as GRAND [140]), would provide around 350 events per year.

8.4. Conclusion and outlooks

The EXTASIS experiment has detected several LF signals in correlation with cosmic-ray events. They have been seen in coincidence with CODALEMA, allowing for some of them to reconstruct the characteristics of the primary cosmic ray by combining MF

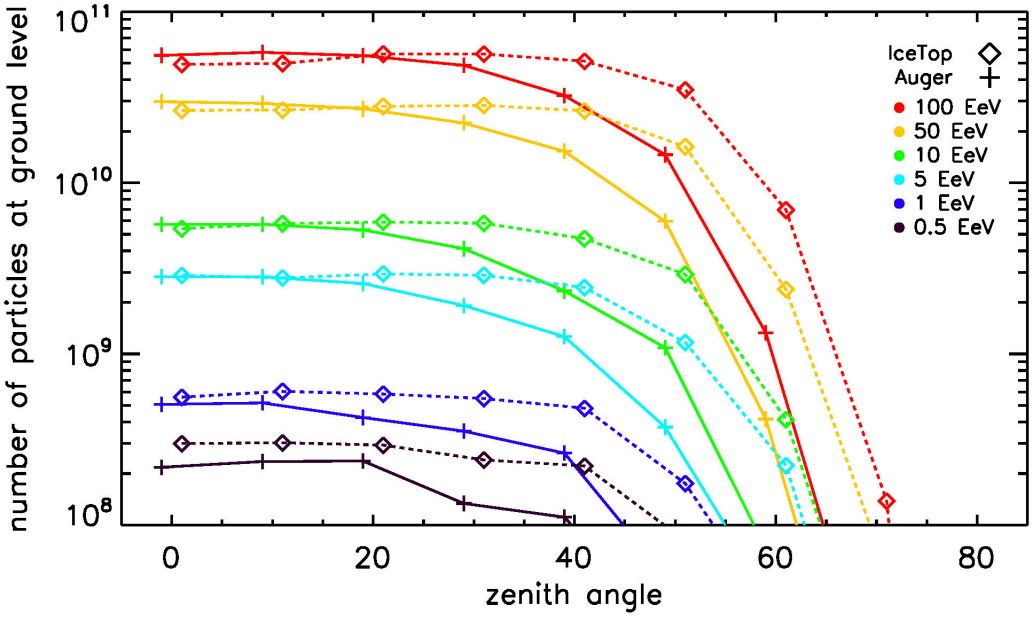


Figure 8.14 *– Number of particles at ground level as a function of the zenith angle, at various energies, for the Pierre Auger Observatory site (altitude* 1400 m*) and the IceTop site (altitude* 2800 m*).*

signals with SELFAS3 simulations. Using these results, we have performed a precise simulation and compared it to the LF and MF data. While we expected an agreement for the comparison with the MF data, due to the fact that it is the standard band used for the radio reconstruction method, we have seen that the comparison with the LF data is also satisfactory. This is a very nice validation of the SELFAS3 code. This result shows, for the first time, the frequency spectrum of air showers measured over a large frequency range from 1 to 200 MHz, despite the strength of the AM and FM bands in the Nançay vicinity. We have also confirmed that the detection range at low frequency is larger than in the classical band, as depicted in figure 8.8.

A search for LF events based only on the presence of low-frequency transients among the events recorded by EXTASIS, however, did not yield any evidence of events without a MF counterpart, confirming the conclusion that, when it exists, the low-frequency contribution of the radio signal of the atmospheric particle showers is much smaller and more difficult to detect than the contribution in the EMF band. Unfortunately, harsh atmospheric noise conditions hamper the detection at low frequency, for which the efficiency is already very poor. We have shown that the low rate of detection in the low-frequency band is mainly due to the atmospheric noise, which, in the best case, remains 10 times higher than the amplitude of the signal that we want to detect at the low altitude of Nançay; the duty cycle is around 50 %, the noise being too high during night time. We also found a correlation with the atmospheric electric field, that probably amplifies the transient signal and lowers again the real detection efficiency in normal conditions. This amplification could explain at least part of the large shower electric field values recorded in the past, though this can not be definitely demonstrated a posteriori.

Concerning at last the observation of the expected SDP, our simulations show that the EXTASIS antennas, though sensitive enough regarding the LF noise conditions, are not installed on a favorable site. Higher altitude sites coupled with a large detection area are much more favorable: 2800 m (IceTop) would be better than 1400 m (Auger) but the size of the IceTop site provides only around 1 shower per year detectable with the SDP. Auger is well suited with its area of 3,000 km^2 and could observe 2,600 showers producing

a detectable SDP per year. It would be very interesting to perform MHz measurements in the Auger-Horizon project [229], initially dedicated to the precise measurement of inclined showers in the usual range 30-80 MHz. Also the GRAND site, if confirmed at 3000 m of altitude, could be very well suited to the SDP search (around 350 events per year assuming an area of 200 km^2).

Remark: The initial aim of the EXTASIS programme was to re-investigate the low-frequency band and in particular the results reported by pioneering experiments. The programme, financed by the Région Pays de la Loire, and supported by the Programme National des Hautes Energies of CNRS/INSU, has made it possible to strengthen manpower in order to conduct sensitivity studies of the experiment, and in particular to improve the MC simulation code SELFAS. It should also be noted that only the EXTASIS experiment was able to estimate the background noise of the low-frequency band, which after study is the main limitation to the detection of radio transients in this frequency range. At the beginning of EXTASIS, although based on the results obtained for the detection of radio transients in the classical frequency band, knowledge of the low-frequency transient sky was almost nil. The same efforts and work as in the early 2000s for the classical band were to be carried out.

"Why do you think I came all this way?"

Bran Stark, Game of Thrones, Season 8: The Iron Throne

CHAPTER 9

Cosmic Ray mass composition using radio signal with CODALEMA

Outline

As seen in section 6.3, the use of the particle detectors is a precious information to ensure that the recorded radio events are from cosmic origin. In the introduction of this chapter, we are quickly discussing the analysis method to reconstruct the shower parameters with the particle detectors. The parameters obtained with the particle detectors, and the information on the quality of the reconstruction will be useful to build a batch of events seen in coincidence by the scintillators and the radio antennas. The analysis pipeline developed during this thesis allows us to have a sufficient number of reconstructed events, and thus a sufficient statistic for the parameters obtained (X_{max}, energy) from the radio signal. It becomes thus natural to use this batch of events to propose a cosmic ray mass composition using the radio information provided by the antennas of CODALEMA.

9.1. Analysis of the scintillator array

This work was developed internally within the group. The reconstruction method used in this section is a commonly applied method [230, 231]. We will focus on the reconstruction of the shower core position and on the reconstruction of the energy of the primary particle because these two parameters are of first interest for the analysis made in the second part of this chapter. Thus, we decide to not detail the part concerning the event selection and the data reduction, as well as the part concerning the monitoring and the calibration.

9.1.1. Core position reconstruction

The core position reconstruction can be carried out using a NKG function to describe the scintillator data (see 2.1.3 page 33). We define it as follows:

$$\rho(r) = N_e c(s) \left(\frac{r}{r_M}\right)^{s-2} \left(1 + \frac{r}{r_M}\right)^{s-4.5} \tag{9.1}$$

where N_e is the number of charges, r is the distance to the shower axis, s is the shower age taken equal to 1.2, r_M is the Moliere radius taken equal to 80 m and $c(s)$ is a normalization function depending on the shower age and defined as follows:

$$c(s) = \frac{\Gamma(4.5 - s)}{2\pi r_M^2 \Gamma(4.5 - 2s)} \tag{9.2}$$

where Γ is the mathematical function. The shower age and the Moliere radius have been fixed from simulations. A likelihood function is used and defined as follows:

$$L = \Pi_i P_G(\rho_i, \rho_i^{\text{fit}}) \times \Pi_j P_{\text{zero}}(\rho_j^{\text{fit}}) \tag{9.3}$$

where P_G is the Gaussian probability for the detection above a threshold set to $\rho_i > 0.3$ VEM. The density ρ_i is an estimator of the number of particles per surface unit and can be seen as the energy deposit per surface unit expressed in units of vertical-equivalent muon, noted VEM, and equal to:

$$P_G(\rho_i, \rho_i^{fit}) = \frac{1}{\sqrt{(2\pi)}\sigma_i} \mathrm{e}^{-\frac{(\rho_i - \rho_i^{fit})^2}{2\sigma_i^2}} \tag{9.4}$$

where σ_i is the estimated uncertainty of the measurement and ρ_i^{fit} is the density predicted by the NKG fit. In the case where the measured density is not above the defined threshold, a Poissonian no-hit probability is used:

$$P_{\text{zero}}(\rho_j^{\text{fit}}) = (\rho_j^{\text{fit}})^0 \frac{\mathrm{e}^{-\rho_j^{\text{fit}}}}{0!} = \mathrm{e}^{-\rho_j^{\text{fit}}} \tag{9.5}$$

We then minimize the negative of the logarithm ($-2 \log L$) of the equation 9.3, and removing the constant term, we obtain the alternative likelihood function:

$$F = -2 \log L = \sum_i \frac{(\rho_i - \rho_i^{\text{fit}})^2}{\sigma_i^2} + 2 \sum_j \rho_j^{\text{fit}} \tag{9.6}$$

where F is the function to minimize and is depending on the NKG function ρ_i^{fit} containing three free parameters, N_e and the coordinates of the core position (x_c, y_c) of the shower. The NKG function can be expressed as:

$$\rho_i^{\text{fit}}(N_e, x_c, y_c) = \frac{N_e \Gamma(3.3)}{2\pi r_M^2 \Gamma(2.1)} \left(\frac{\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}}{r_M}\right)^{-0.8} \left(1 + \frac{\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}}{r_M}\right)^{-3.3} \tag{9.7}$$

where (x_c, y_c) are the coordinates of the core position and (x_i, y_i) are the coordinates of each scintillator on the shower plane. An example of the NKG fit is shown in figure 9.1, representing the charge of the scintillators as a function of their distance to the shower axis.

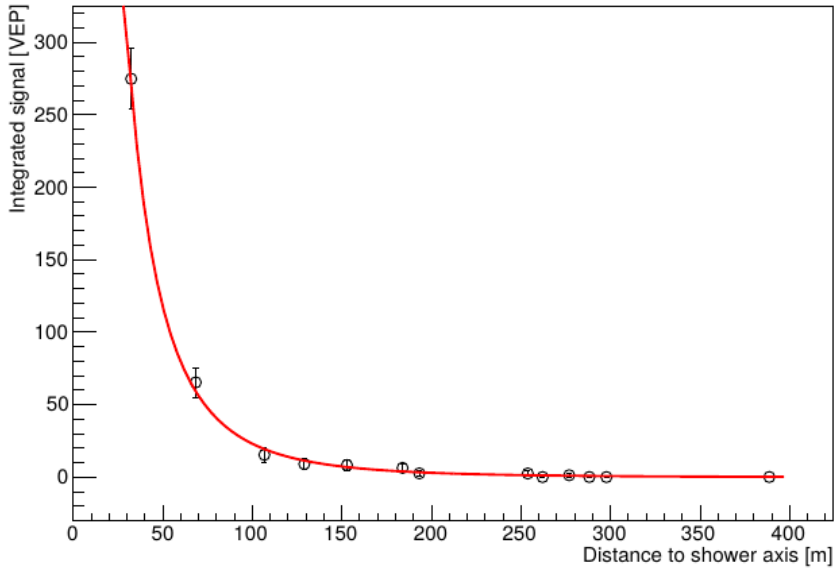


Figure 9.1 *– Example of integrated signal in VEP (Vertical Equivalent charged Particle, same as VEM) as a function of the distance to the shower axis. The uncertainties for each point are shown as error bars. The red line represents the result of the NKG fit.*

9.1.2. Energy reconstruction

To estimate the energy, we fit the following function:

$$\log_{10}(E/\text{eV}) = p_0 + p_1 \log_{10} N_e + p_2 (\log_{10} N_e)^2 \tag{9.8}$$

where the parameters p_0, p_1 and p_2 have been determined on a large set of AIRES Monte Carlo simulations [107] performed to estimate the precision on the arrival direction and on the core position but also to find the relationship between the energy of the primary particle and the shower size parameter N_e. Figure 9.2 presents the energy distribution of the internal events measured by the scintillators from November 2014 to the end of 2018, corresponding to 1211 effective days of stable data acquisition as shown in figure 4.4. This criterion called *centrality*, defines events for which one of the 5 central scintillators (constituting the inner part of the scintillator array, see figure 4.14 for instance) has measured the largest signal. The cosmic-ray energy threshold of our trigger setting (which is currently 5 out of the 13 scintillators) is evaluated between 5×10^{15} and 10^{16} eV, which is clearly visible in figure 9.2. The red line stands for the fit with an exponential function, whose the parameters indicate a slope of −3.3, which is overall in agreement with the power law given in section 1.2.2. The overall behavior of our spectrum seems consistent with the literature, since at 5×10^{17} eV the slope is $\sim$ 3.3. Due to a lack of a careful and constant monitoring of our instrument, we do not master precisely the acceptance of our detector, and thus we can not derive an absolute flux of cosmic rays observed by CODALEMA.

9.1.3. Reconstruction accuracy

In 2009, the scintillator array was composed of 17 scintillators. The uncertainties on the reconstructed parameters were estimated to $\sim$ 10 m on the position (x and y) of the shower core for the central events, and 30 % on the energy at 10^{17} eV [168]. Since then, 4 of the most central scintillators have been removed. The modification of the geometry of the array does not necessarily change a lot the resolution because the 4 scintillators removed were very internal to the array. Removing the 4 inner scintillators has also increased the

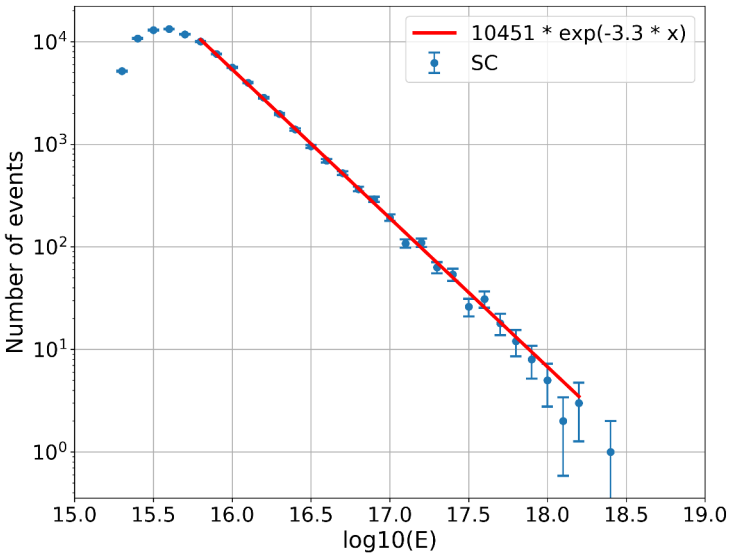


Figure 9.2 *– Energy distribution of internal events measured by the scintillators. The red line stands for the fit with an exponential function.*

energy threshold. On the other hand, the inhomogeneity of scintillators (over time or between them) in recent years (ageing of the plastic scintillator for instance) has certainly degraded the quality of the reconstructions and therefore the resolutions in position and energy. The uncertainties reported above for an array of 17 scintillators therefore seem to be greater for the case with 13 scintillators.

9.2. Correlation between radio and particle measurements

Contrarily to what was made in the past on other experiments [71, 220], the events detected by CODALEMA between mid of 2017 and the end of 2018 have been reconstructed with the radio method, with no selection made with another type of detectors such as fluorescence detectors. The applied cuts remove from the data set:

— the events with a multiplicity lower than 3 antennas ;
— the events with a bad rise time, as explained in chapter 6 section 6.2.1 ;
— the events with a bad compacity, as explained in chapter 6 section 6.2.2 ;
— the events with a bad polarization, as explained in chapter 6 section 6.2.3.

The number of events after the radio cut is 423 events. The quality of the estimated shower parameters reconstructed with the radio method can be obtained by comparing them with the reconstructions made with the particle detectors.

9.2.1. Particle detector internal events

First of all, an additional cut-off should be applied to the batch of 423 events, on the quality of the reconstruction of particle detectors (chi-square value). Of course, this cut has been applied to produce the figure 9.2 above. Applied on the batch of radio events, the number of events goes down from 423 to 256.

In this section, we will focus on the events which correspond to events tagged as central by the analysis of the particle detectors since they have the most reliable energy estimation. In that case, the number of radio reconstructed events goes down from 256 to 37. To explain this drop and the difference between “central” and all events, we try to estimate the surface of detection of each array. The area of the convex hull (see section 6.2.2) of the inner part of the scintillators is equal to 40000 m^2, and the area of the the convex hull which contains all the CODALEMA radio array is equal to 870000 m^2. The

ratio of the two areas gives ~ 0.05. The ratio between the events tagged as central and all the detected events gives ~ 0.09. The drop in the number of events when central events are selected can be very roughly explained by the actual detection area of the arrays.

From the set of internal events, we can observed that the energy, x and y core positions deduced from the particle detectors and the radio signals compare relatively well as shown in figure 9.3.

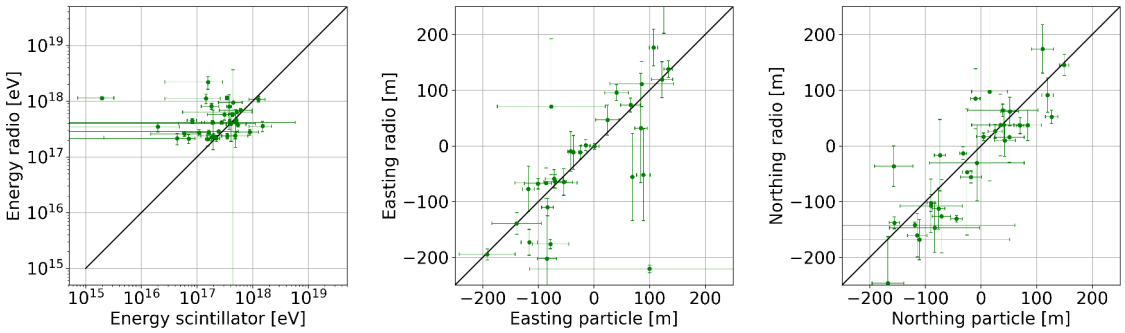


Figure 9.3 *– Left: correlation between the estimated primary cosmic ray energy with the radio method and with the particle detector method for the central events. Center: correlation between the estimated easting core position with the radio method and with the particle detectors method for the central events. Right: same as center for northing core position. The straight line accounts for a one-to-one correlation.*

9.2.2. Whole batch of events

Figure 9.4 left presents the comparison between the primary cosmic ray energy estimated with the radio method and with the particle detectors method for the full batch of 256 events. For the external events (of the scintillator array), the estimation of the shower core from the scintillator reconstruction is difficult to constrain, and thus the estimation on the energy is less accurate. The estimated energy from the particle detectors are mostly lower than the estimated energy from the radio method. This is due to the fact that the surface covered by the autonomous station array is much larger than the one of the particle detector array. Thereby, most of the events falls outside the particle detector array, meaning that the estimated shower core positions are not reliable (the shower core position is pulled inside the particle detector array limits) and that the energies are underestimated.

Thereby, for the external events, we are relying on the shower core determined from the radio method in the scintillator reconstruction to estimate the energy as shown in figure 9.4 right. In that case, the new energy from the particle detectors tends to increase and the global correlation is clearly improved. This result indicates that the energy estimated with the radio method is reliable. In a further step, the cut made on the quality of the reconstruction of the scintillators will be removed, allowing a pure radio study of the mass dependence of the cosmic rays with energy, increasing the batch of events to ~ 400 events (see section 9.3.2).

9.3. X_{max} reconstruction

9.3.1. Using cuts on the scintillator reconstruction

The X_{max} values are also reconstructed from the systematic comparisons of the radio signals with the simulated events, and a composition estimation versus binned energy

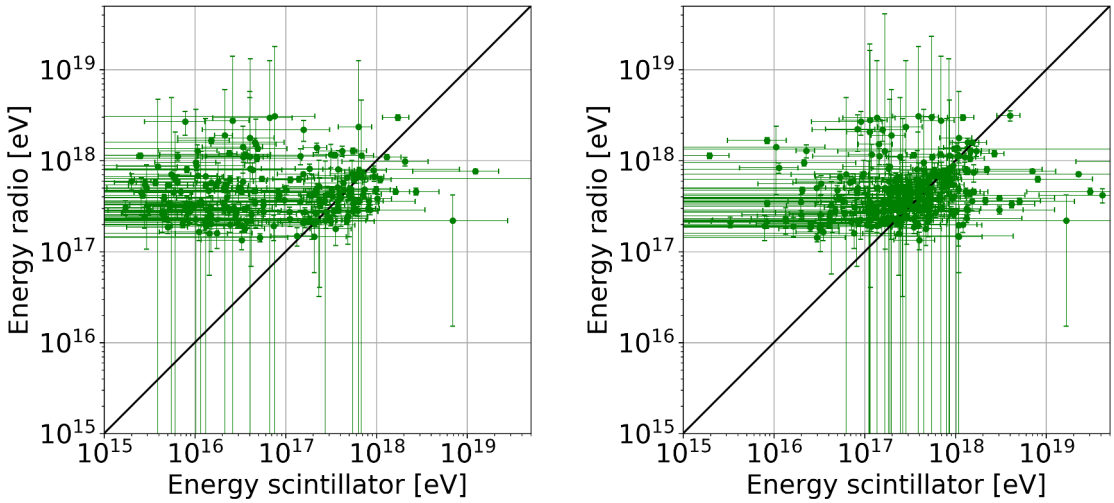


Figure 9.4 *– Left: correlation between the estimated primary cosmic ray energy with the radio method and with the particle detectors method. Right: the same but the core location used for the particle detector is taken from the radio signal analysis. The straight line accounts for a one-to-one correlation.*

can thus be tentatively derived. The obtained results are compared to the $X_{\max}$ estimated at the Pierre Auger Observatory (fluorescence technique), LOFAR (radio technique), Hires/MIA (combination of fluorescence and particle technique), Yakutsk (Cherenkov technique), TALE (fluorescence technique) and Tunka (Cherenkov technique), and presented in figure 9.5 left. Values from the other experiments are extracted from these references [232, 233, 234].

The error $\delta X_{\max}$ corresponding to a bin in energy, i.e corresponding to the error on the $\langle X_{\max} \rangle$, is calculated as follows:

$$\delta X_{\max} = \frac{1}{\sqrt{N}} \sqrt{\frac{1}{N-1} \sum_{i}^{N} (X_{\max}^{i} - \text{mean}\,(X_{\max}))^2} \tag{9.9}$$

with N the number of $X_{\max}$ values for a given bin in energy.

Remark: It is worth noticing that the acceptance of the radio array of CODALEMA has not been established. It is therefore impossible to correctly estimate statistical uncertainties. The same applies to the systematic uncertainties.

The data seem to follow a linear increase with the energy with the exception of the first and the last points. The reconstructed values are in very good agreement with the measurements of the other experiments, and are compatible with a light composition at energies above $10^{17.6}$ eV.

9.3.2. Full data set

We have shown previously that the energy reconstructed with the radio method was reliable when compared to the energy reconstructed by the scintillators. In this section, we perform a blind selection of radio events recorded between mid 2017 and mid 2019 (565 events), i.e without the quality cuts on the scintillator reconstruction applied before. After the application of the rejection algorithms described in the chapter 6 (and recalled in section 9.2), and after application of a cut on the quality of the radio reconstruction, the batch contains 405 events. The $\langle X_{\max} \rangle$ dependence with the energy is presented in

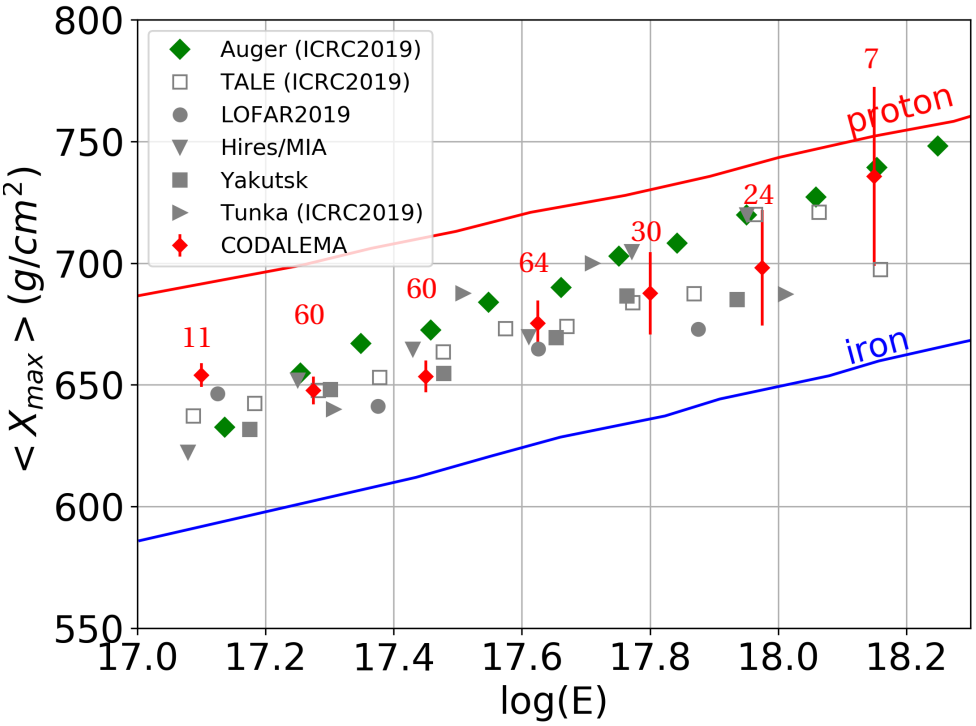


Figure 9.5 *– $\langle X_{\max} \rangle$ as a function of the energy of the primary by energy bins, with cuts on scintillator reconstruction (256 events). The red and blue lines are the simulated $\langle X_{\max} \rangle$ values obtained from the hadronic interaction code EPOS-LHC. The green diamonds are the $\langle X_{\max} \rangle$ measurements at the Pierre Auger Observatory, grey symbols are the mean values from LOFAR, Hires/MIA, Yakutsk, Tunka, TALE and the red diamonds are the results obtained with this work. For readability reasons, the error bars of the other experiments are not indicated.*

figure 9.6 left. The errors on the $\langle X_{\max} \rangle$ values are calculated with the equation 9.9. The reconstructed values are still in very good agreement with the other experiments. We note nevertheless that we are systematically below the values proposed by Auger, probable a sign of systematic effects related to the radio detection not yet well known. Figure 9.6 right presents the energy distribution of the corresponding events, reconstructed with the radio method. The red line stands for the fit of an exponential function, whose the parameters indicate a slope of −3.3, which is overall in agreement with the power law given in chapter 1 section 1.2.2 page 15, but also which is consistent with the power low found for the energy distribution of the scintillators. This indicates that our batch of events used for the composition is representative of the range of energy we are working on, and that this batch is therefore not strongly biased by the selections made.

As for Auger, the composition tends to become lighter with increasing energy. The main conclusion is that the results from CODALEMA does not contradict the results obtained with other methods such as Cherenkov light or air fluorescence. The fine proportion of the mass component is discussed in the next section. To check the consistency of the batch of events, and in order to verify if it is not biased by the applied selections, different tests have been applied to work on sub-batches of events to study the evolution of the $\langle X_{\max} \rangle$ values: dependence on the zenith angle, on the multiplicity etc. All of the previous results remain consistent. For instance, figure 9.7 right presents the results obtained on a batch of 304 events, selected a cut on the uncertainty of the reconstructed energy (30 %). The seven energy bins are respectively composed of 11, 69, 77, 75, 34, 30 and 8 events. The errors on the $\langle X_{\max} \rangle$ are calculated with the equation 9.9.

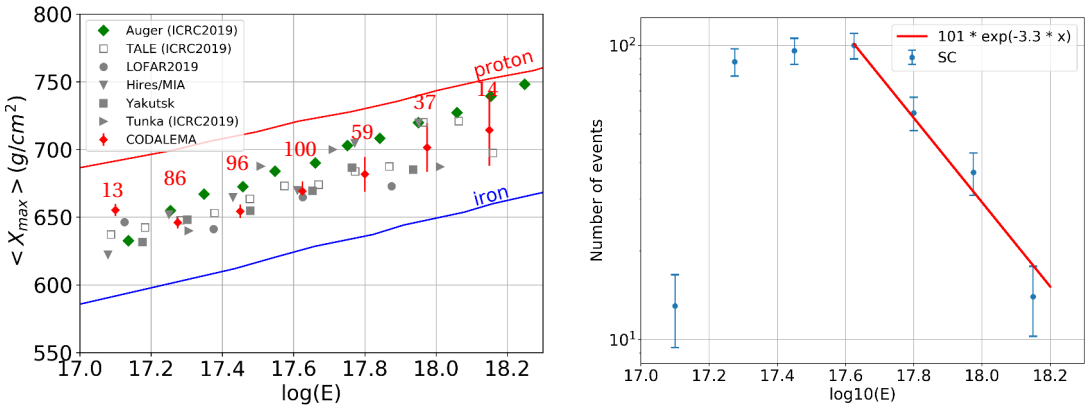


Figure 9.6 *– Left: $\langle X_{\max} \rangle$ values as a function of the energy of the primary by energy bins. Right: the associated energy distribution.*

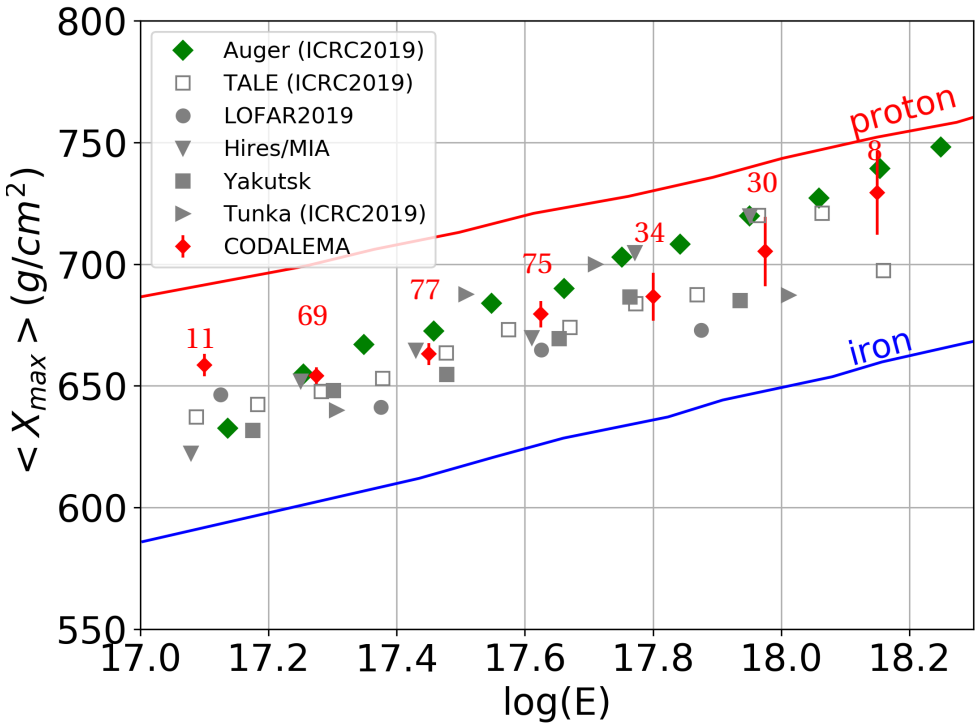


Figure 9.7 *– $\langle X_{\max} \rangle$ values as a function of the energy of the primary by energy bins, with cut on the uncertainty of the reconstructed energy (304 events).*

9.4. Attempt to estimate the mass composition of cosmic rays

The analysis developed in this section is based on an analysis carried by LOFAR in [232]. For each shower of our batch, we calculate a a-parameter defined as follows:

$$a = \frac{\langle X_{\max}^{\text{proton}} \rangle - X_{\max}}{\langle X_{\max}^{\text{proton}} \rangle - \langle X_{\max}^{\text{iron}} \rangle} \tag{9.10}$$

where $\langle X_{\max}^{\text{proton}} \rangle$ and $\langle X_{\max}^{\text{iron}} \rangle$ are respectively the mean value of $\langle X_{\max} \rangle$ for proton and iron obtained from the hadronic interaction code EPOS-LHC, and $X_{\max}$ is the estimated value for the considered event. We note that a pure proton composition will present a distribution centered on 0 while a pure iron composition will present a distribution

centered on 1. Due to the larger fluctuations of X_{max} values, the proton distribution is larger than the iron distribution. Then, the cumulative function of the probability density function (also called cumulative distribution function, noted CDF) of the a-parameters is calculated. The same procedure is made for four primary cosmic ray masses: proton, helium, nitrogen and iron as shown in figure 9.8 left. We then fit a four-component model to the CDF of our data, letting as free parameter the mixing ratio of the four primary cosmic ray masses. The best fit (brown curve) is found for: 1 % protons, 50 % Helium, 33 % Nitrogen and 16 % Iron. In the same way that the propagation of errors made in chapter 7, we estimate the errors for the a-parameter, and calculate the mean CDF and the region containing 99 % of the realizations (see the blue curve for the mean CDF and the shaded region for the level of confidence in figure 9.8 left). Figure 9.8 right presents the distribution of the chi-square values obtained for the four-component model when the proton and helium fractions are fixed, and when only the mixing ratio between nitrogen and iron is tested. The best fit is obtained for 1 % protons, 50 % Helium.

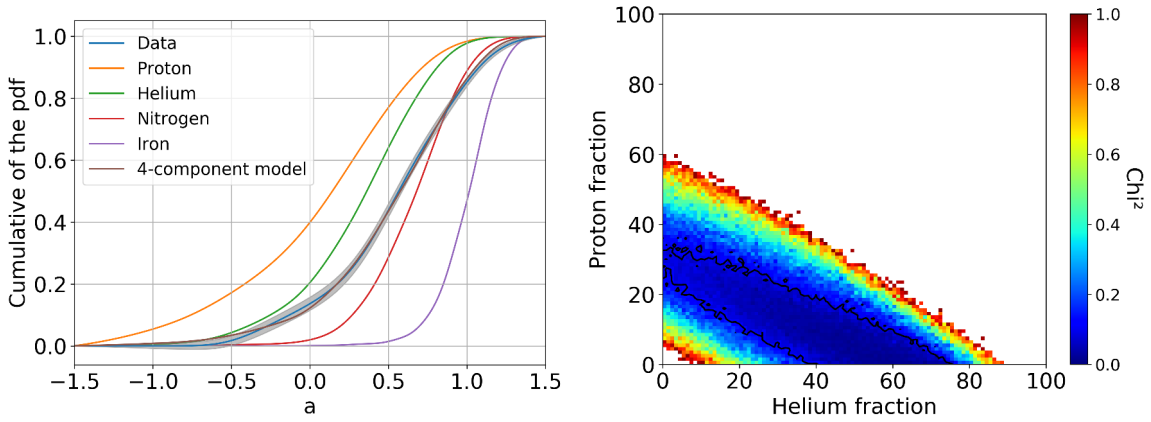


Figure 9.8 *– Left: Cumulative Distribution Function of the parameters a for the data and the pure composition protons, Helium, Nitrogen and Iron obtained with EPOS-LHC. The shaded region corresponds to the region containing* 99 % *of the realizations of the CDF of data (see text for details). Right: Distribution of the chi-square values obtained for the four-component model when the proton and helium fractions are fixed, and when only the mixing ration between nitrogen and iron is tested. The contour indicates all combinations of proton-Helium for which the fit is weakly deteriorated when replacing Helium by protons.*

9.5. Conclusions

To draw some conclusions on the results presented in the chapter:

— We have exploited the information from the 13 particle detectors, which permitted to show that a one-to-one correlation exists between E_{radio} and $E_{particles}$. Indeed, the energy obtained with the radio method compares well with the energy obtained from the particle detectors within the scintillator array acceptance and resolution. The energy obtained with the radio coupled with MC simulation codes seems reliable.

— The results obtained are preliminary, but the composition derived from the CODALEMA data is in agreement with the other experiments in $\log_{10} E \in [17.0 - 18.2]$, in particular with the composition proposed by the Pierre Auger Observatory [52].

— The results tend to confirm that the composition becomes lighter with increasing energy.

Thereby, above 10^{17} eV, the method (described in chapter 7) using only the radio signal is able to reconstruct the shower parameters and to estimate the cosmic ray composition. However, results presented in this chapter are preliminary, and some studies and checks have been carried out and some others are in progress:

— In order to confirm the results presented in figure 9.6 left with 405 events, some cuts have been applied to check if the results are not biased by the batch of events used: cuts on the quality of the radio reconstruction and on the atmospheric electric field for instance. The result on the composition remains stable after different cuts, as shown for example in figure 9.6 right.

— Events of 2019 are being processed to increase the statistics, and thus to permit to applied post-reconstruction cuts in order to obtain a very high accuracy on the parameters of events used in the analysis.

— As a reminder, we have shown in section 3.3.2.3 that significant discrepancies exist between SELFAS and ZHAireS simulation codes: differences in the timing, the amplitude and the shape of the pulses. A new version of SELFAS has been produced in July 2019 (new random choice of the shower particles, new calculation of the charge excess, change to multi-thread), and some discrepancies are still present compared to ZHAireS code (see figure 9.9): differences in timing and also in amplitude of the pulses. It would be interesting to quickly restart all the simulations of the events used to check if shower parameters and the proposed composition remains the same. In addition, it would be interesting to do the same work using another simulation code such as ZHAireS to compare. As a quick check, only one event has been processed with the new version of SELFAS, leading to a difference of ~ 23 % in the estimation of the energy, and a difference of ~ 20 g $\cdot$ cm^{-2} in the estimation of the $X_{\max}$ value. In this case for instance, the event would have contributed to the next energy bin, with a higher $X_{\max}$ value, passing the mean $X_{\max}$ value of the bin from 729 ± 17 g $\cdot$ cm^{-2} to 719 ± 18 g $\cdot$ cm^{-2}. The use of the new version of SELFAS adds a difference in the energy and $X_{\max}$ reconstructed values, but they remain compatible.

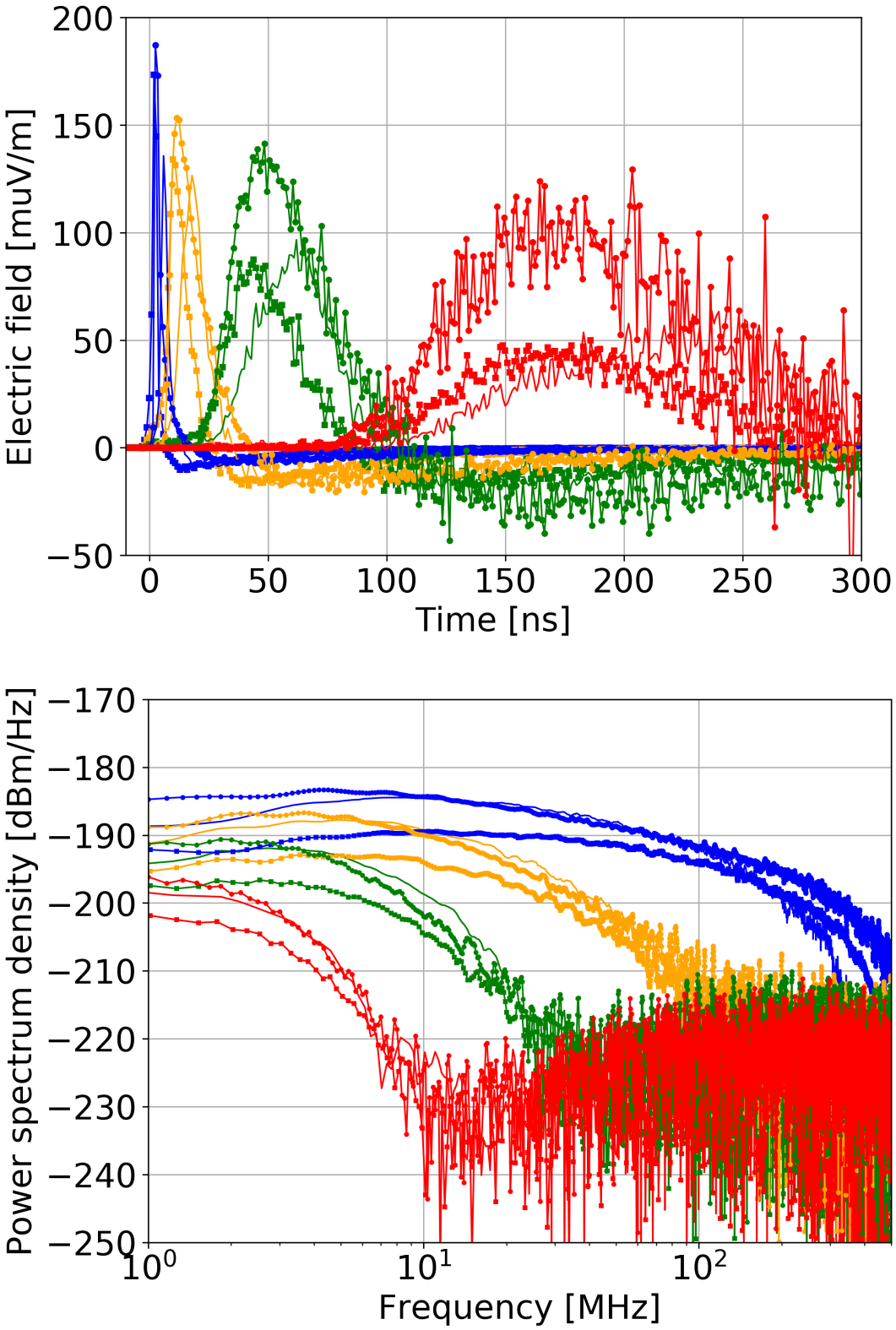


Figure 9.9 *– Comparaison between updated SELFAS3 (plain lines), SELFAS3 (squares) and ZHAireS (points) for different shower axis distances: blue for* 100 m*, orange for* 200 m*, green for* 400 m *and red for* 800 m*. Top: East-west component of the simulated electric field pulses, corresponding to a vertical shower induced by a proton of* 10^{17} eV *for an altitude of* 1400 m*. Amplitudes of the signals are multiplied by a factor increasing with the shower axis distance (1 for the blue curves, 6 for the orange curves, 36 for the green curves and 216 for the red curves). Bottom: PSD of the east-west components of the simulated electric field pulses.*

Part IV

Epilogue

"We're done when I say we're done."

Walter White, Breaking Bad

CHAPTER 10

Discussion and outlook

Closing remarks on this work

Over the years, significant efforts have been devoted to the understanding of the radio emission from extensive air showers. From the feasibility of the method demonstrated by LOPES and CODALEMA to the very recent results of LOFAR and CODALEMA, on the understanding of the main features of the radio signal, it seems today that the radio method is able to reconstruct the main characteristics of primary cosmic rays. It is often reported that this method is enough mature to compete with other methods such as fluorescence, with the advantage that its duty cycle is close to 100 % (compared to about 14 % for the fluorescence), making it possible to significantly increase the number of detected events, and thus to better understand the acceleration mechanisms of cosmic rays and to better constrain the properties of their sources, flux and nature. New antenna array projects for shower detection are planned (SKA, AugerPrime, neutrino telescopes).

However, to get there, a fine understanding of the phenomenon is necessary, from the propagation of secondary particles in the atmosphere up to the mechanisms of production of the radio emission, through the understanding of the atmosphere itself and the use of precise models such as the GDAS. All these ingredients are needed to build the most realistic simulation codes of the radio signal of cosmic-ray-induced air showers, which, by systematic comparisons with the data, allow us to reconstruct the primary cosmic ray characteristics.

Low-frequency band exploration

The band available below 20 MHz has remained unused for almost 20 years. It has been reported by some pioneering experiments that extensive air showers (EAS) emit a strong electric field and that there is evidence of a large increase of the radio pulse amplitude with decreasing frequencies. We have re-investigated the low-frequency (LF, namely the [1.7 – 3.7] MHz band) band with the EXTASIS experiment. We have observed 25 events in coincidence in CODALEMA and EXTASIS clearly identified as cosmic-ray events. We have estimated a detection threshold of $23 \pm 4\ \mu V \cdot m^{-1}$ from comparisons with detailed SELFAS3 simulations, and we have confirmed that the detection range at LF is larger than in the classical [30 – 80] MHz band. However, we have demonstrated that the harsh atmospheric noise conditions hamper the detection at LF. The low rate of detection in this band is mainly due to the atmospheric noise, which remains in the best case 10 times higher than the amplitude of the signal that we want to detect. We have also highlighted a strong correlation of the LF signal observation with the atmospheric electric field. Concerning the so-called "sudden death" contribution (the expected electric field radiated by the shower front when hitting ground level), no signature of such signal has been observed at Nançay. Its probability of detection is favored for large instruments

installed at high altitude such as Auger or GRAND combining large coverage and high number of particles at ground level. However, even if the expected amplitude of the SDP is larger than the one of the regular pulse at LF, such a detection should face the same problem as the one here encountered, i.e. surpassing the atmospheric noise level. Moreover, if identified, using this signal to improve the intrinsic time scale within the shower should be an additional difficult challenge.

Reconstruction of the shower parameters

The understanding of the behaviour of the radio signal with respect to the polarization, the timing or the frequency requires a complete understanding of the instruments. In this thesis, we have shown that a thoroughly modelling of the CODALEMA radio antennas, containing a description of the ground, the near environment of the antenna, the antenna gain and the antenna impedance is mandatory to satisfactory fit the data, but also in order to convolve the simulations to compare them to the data for the reconstruction of the primary cosmic ray characteristics. One of the main contributions of this thesis was the development of an automatic analysis framework including the noise rejection algorithms previously mentioned and a method of cross calibration of the CODALEMA radio antennas using the galactic radio emission in order to correct the differences between antennas related to the antennas themselves and to their electronic noise. The cross calibration using the galactic radio signal permitted to strongly reduce the disparities between the radio antennas, leading to a homogeneous behaviour of the SA. This analysis framework makes it possible to reconstruct all the properties of the EAS with the radio signal, such as the direction of arrival, the energy of the primary cosmic ray but also the nature of the primary cosmic ray. It is based on systematic comparisons between simulations and data, developed at Subatech for the purpose of AERA and CODALEMA data analysis. Exploiting the capabilities of the CODALEMA instruments, the reconstruction has been generalized by taking into account the high frequency part of the signals leading to a reconstruction in the EMF band (very useful for inclined showers) and by decoupling the polarizations in the procedure, which means to treat them independently. We have demonstrated that the generalization of the method allows to obtain better results on the reconstructed shower parameters, mainly for the very inclined showers, the uncertainties on the estimated parameters obtained with the new method becoming comparable to those obtained by the fluorescence technique for example. To validate the reconstructed parameters with the radio method, we were able to use the reconstructions made with the particle detector array, to compare the reconstructions of the primary cosmic ray energy and to reach a satisfactory agreement.

We have shown that, by using the core position estimated with the radio method, the energy estimated from the particle detectors can be re-evaluated, and the correlation between the reconstructed energy with the radio and the particle methods is clearly improved. It is worth noticing that the particle detector of CODALEMA is composed of only 13 heterogeneous and ageing scintillators. Despite this heterogeneity, we were able to assess the relevance of the shower parameters derived from the radio and we have shown that the preliminary results on the composition of cosmic rays proposed by CODALEMA is in agreement with the other experiments. Several works and checks are in progress, especially to increase statistics.

Noise rejection

A comprehensive study of the transient background noise seems nowadays mandatory. Chosen for its radio quietness when it was selected, the Nançay Radio Observatory

has revealed itself to be strongly affected by transient noise sources. CODALEMA has witnessed the (bad) evolution of its nearby environment over almost twenty years. Indeed, in order to build the purest batch of radio events for the reconstruction of the primary cosmic ray characteristics, the analysis framework requires solid noise rejection algorithms. Although, the galactic background noise allows us to calibrate our antennas since it is predictable, the man-made background noise is one of the limitations for the standalone detectors, due to the triggering on noise transients and to the dead time induced by our slow acquisition system. Failing to be able to improve our first level detection efficiency, different noise rejection algorithms have been explored to improve the selectivity of the radio events. These developments require a thorough understanding of the behavior of signals from EAS (understanding of the two fundamental radio emission mechanisms, polarizations, timing...). Even if a much more thorough study of these algorithms is mandatory to ensure the independence of the radio method, namely to be able to do without the additional help from another type of detectors such as a particle detector to confirm the origin of the event, we have seen that it should be possible to obtain a satisfactory purity of the radio events. To date, we are able to combine several pieces of information together to tag unambiguously an event as a cosmic ray event with the only radio signal, meaning that the latter is now well understood to use it to recognise a shower event from a parasitic event.

Ongoing or tested improvements

CODALEMA and EXTASIS have explored some possible solutions to improve the radio detection method, such as:

- To use clever triggering algorithms to solve the autonomous detection problem. It is about the composite trigger developed within the group, which is a combination of several antennas in a clustered way. It could also permit to detect cosmic ray events without particle detectors, for example. Time is lacking to detail this promising way.
- Concerning the detection range, it has been demonstrated in [159, 9] that it is much larger at low-frequency (below 5 MHz) than in the classical band [30 – 80] MHz, confirming the pioneering observations and the theoretical expectations. This was one of the goals of the EXTASIS project.
- As it is shown in appendix E, above 120 MHz the signal strength is clearly dependent on the distance to the shower axis. Moreover, the radio Cerenkov ring can be well observed, as shown in figure E.4. This observation could help to better constrain the reconstruction of the shower core (see chapter 7), but also to solve the problem of the antenna density at short distance of the shower axis. The description of the electric field profile could be better defined over a large bandwidth where its fast variations are observed.
- Finally, a three-fold antenna has been installed [160] close to the center of the particle detectors of CODALEMA, in the middle of the Compact Array, in order to measure the complete 3D electric field produced by air showers. The far-field hypothesis commonly used by the radio community is in the process of being verified and validated.

Pros & Cons of the radio detection method

In view of the conclusions drawn above, i.e on several facets of CODALEMA and the radio detection, a summary of the pros and the cons can be tentatively made:

Pros:

— The radio detection method has a potential duty cycle greater than the one of the fluorescence technique, making it one of the most effective methods on paper.
— The method is sensitive to the totality of the development of the shower, i.e it informs us on all the stages of development and life of the shower (against only a final view of the development of the shower for the technique of detection of particles). However, it should be noted that to date, no one has studied the correlation between the development of the shower and the temporal part of the signal. At a second level, i.e. not directly related to the study of cosmic rays, the radio method gives also the possibility to probe the atmosphere conditions trough the information on the polarization [235, 236].
— The atmosphere being transparent for the radio emission (no absorption or scattering), the radio detection method can be used for an absolute energy calibration of other shower detection methods.
— The method is nowadays almost autonomous meaning that with an effort on the rejection algorithms, this method could do without other methods, and because of its low cost, could allow a very large scale deployment [237].
— Last but not least, it has been demonstrated that the radio detection permits to reconstruct the shower parameters and to propose a composition of the cosmic rays with a very good accuracy.

Cons:
— The duty cycle and the detection efficiency of the radio method are hampered by the parasitic background noise which is really the main limitation for the standalone detectors. The number of parasitic transients is several orders of magnitude bigger than the cosmic ray expected rate.
— The overall radio efficiency, including the detection and the capability to reconstruct an event up to the shower parameters, is actually too low to hope to remarkably increase the statistics at high energies. For instance, LOFAR reported on few hundreds of events [232, 238] collected over several years. Similarly, CODALEMA reported (in this work and here [146]) few hundreds of events corresponding to roughly 3 years of data taking. AERA, the largest existing radio array operating since several years, did not publish so far results based on a large collection of radio events.
— From this, we immediately see that a long time is needed to collect sufficient statistics, and the hope for a major increase of the statistics at high energy seems hampered. Further works on optimizing the array size and the antenna pitch with respect to the target energy window are probably needed. As a matter of fact, the CODALEMA radio array topology was mostly influenced by the deployment constraints and by its initial goals to demonstrate the feasibility of the EAS radio detection.

Own commentaries on the radio detection method

The following section contents more personal comments, and therefore not necessarily shared by the cosmic ray community. As you will have probably understood, via the elements given above, the promise to increase the statistics at very high energy by using the method of the radio detection is not realistic so far. Today, and with the results obtained in the past years, the capabilities and the limitations of the radio method are clear. Several questions remain opened.

Radio detection, what for? We have shown that the radio detection is able to reconstruct shower events. But we have also shown that, to be able to reconstruct accurately

the shower parameters, several selections have to be made, decreasing the initial statistics. One has to wonder what the radio detection bring to the community. For instance the Pierre Auger Observatory results on the composition of the cosmic rays have been obtained with the fluorescence detectors principally. The results on the composition from the radio array of Auger, AREA, concerned the $[10^{17} - 10^{18}]$ eV, range already well described by the fluorescence and surface detectors. The goal of increasing the statistics at 10^{19} eV and above has not been reached, and in this sense, it seems logical to say that the radio did not bring a major contribution to a better understanding of the field of cosmic rays at the Pierre Auger Observatory. It is reported in [237] that the Pierre Auger Observatory will be upgraded, *AugerPrime* project [66], especially with radio antennas operating in [30 – 80] MHz which will equip each of the 1600 water-Cherenkov tank, in order to observe inclined EeV air showers [239]. Inclined showers produce large but rather uniform radio footprint in the usual frequency band which clearly disfavours an accurate estimate of the shower parameters Useful variations of the radio profile can only be observed by looking at higher frequency. Thus, the methods using comparisons between data and simulations can not be applied efficiently in [30 – 80] MHz. Moreover, the antenna used is optimized for zenith angle greater than 50°, since the objective is to observe inclined shower. It should be noted that most of the parasitic signals seen at cosmic ray observation sites comes from the horizon (anthropic noise), which may significantly pollute the measurements (the average rate of parasitic transients is estimated to ~ 15 kHz [240]). It thus will be complicated to use a self-triggered array, and this is probably the reason why each antenna will be triggered by the water-Cherenkov tank on which it will be installed. Thus, in this case, the radio will only supplement the area of the sky not visible by the surface detectors. When a shower is inclined, its electromagnetic component is almost completely absorbed into the atmosphere, and only the muons reach on the ground. The idea of this upgrade is to measure by the radio method the electromagnetic component of the shower, and measure with the ground particle detectors the muon component. The $X_{\max}$ would be then derived from the e/μ ratio. In other words, contrarily to what is claimed in [237], it seems that the antennas of this upgrade can not be used alone, because it will then be very complicated to extract shower parameters with signals seen in [30 – 80] MHz by antennas spaced 1600 m apart.

What remains to be done? Are the efforts to be made worth it? The current situation in France seems to show that radio detection for cosmic rays is no longer the fashionable method, even if new "hybrid" projects are emerging. For instance, the future Square Kilometer Array (SKA), especially the low-frequency core (SKA-low), will be composed of ~70,000 crossed dipoles radio antennas over 0.5 km^2 operating in [50 – 350] MHz [166, 241] (see figure 10.1).

From simulations (the procedure is the same as used in chapter 7 in section 7.4.3 using “fake” data and comparisons with simulations), the resolution obtained on the reconstructed $X_{\max}$ is expected to ~ 10 g · cm^{-2}, which is twice as better than the resolution reached with the fluorescence method for example. The energy range is between ~ 10^{16} eV up to ~ 10^{18} eV, and the number of shower events detected per year above 10^{17} eV is surprisingly estimated around 10,000. These estimations are made from simulations, without considering background noise neither the atmospheric electric field properties. The number of events potentially detectable over a year will then probably be much smaller than 10,000. However, we can expect that the statistics will remain sufficient to deliver a precise cosmic ray composition, if all the detected events can be reconstructed. I would like to mention that the composition in $10^{16} - 10^{17}$ eV is already provided by several experiments using different methods, such as Kascade [242] and Kascade-Grande (electrons and muons measurements at ground) [243], Tunka-133 (Cherenkov light) [244], Tale (air fluorescence, Cherenkov light) [245] or IceTop [246]. It is worth noting that SKA-low is not primarily designed for cosmic ray physics. If the instrument is fully available and

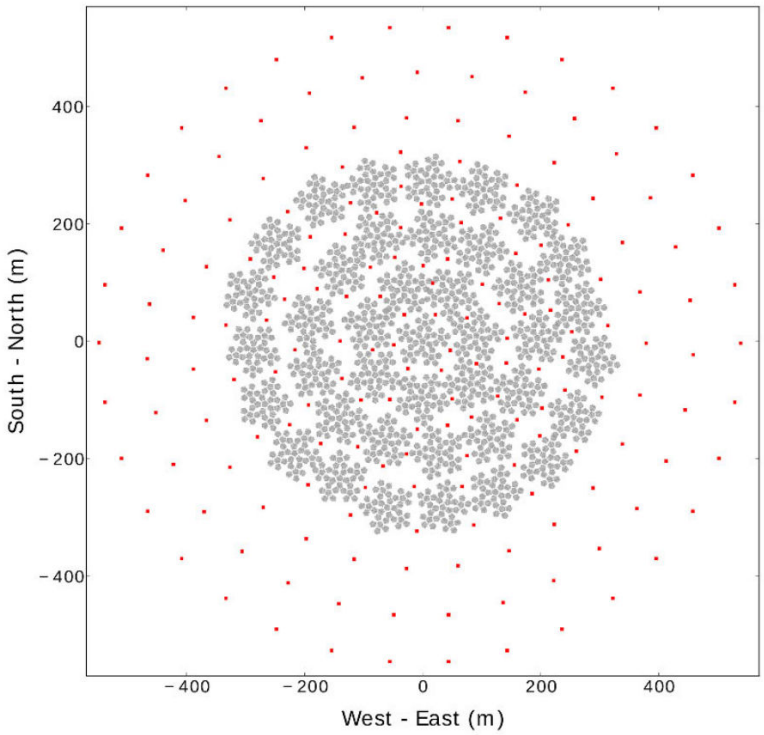


Figure 10.1 *– The SKA core, composed of ~70,000 antennas (grey dots) and of ~180 particle detectors (red squares). From [166].*

capable of conducting this type of observations without the need of significant hardware modifications and assuming that the cost of the associated scintillator array is marginal, it would be probably worth doing it. But again, my point of view is that radio detection of showers with SKA-low will bring nothing more on cosmic ray composition than what is already done by other experiments with other types of detection methods, even more so if we have to take into account the time it will take to build the instrument. However, the proposed analysis on particle and shower physics are very interesting, as well as the use of air-shower studies to probe the physics of thunderstorms and lightning [166].

Despite some "hybrid" projects as SKA-low or *AugerPrime*, the current trend is rather the use of the radio detection method to detect neutrinos, as on ARA [147, 148] or GRAND [141, 247] experiments. ARA, as detailed in section 3.5, aims to study high-energy neutrinos by detecting the radio pulses generated as they go through the ice. The experiments using the radio detection in-ice were actually proposed to study *cosmogenic neutrinos*, the neutrinos produced by the interaction of cosmic rays with the CMB (see equations 1.17 and 1.18). Unlike ultra-high energy (UHE) cosmic rays [1], UHE neutrinos point back to their sources since they are neutral and have a very small interaction cross-section. Their study could provide information on UHE cosmic ray sources but also on acceleration mechanisms. To date, ARA has estimated a projected neutrino flux limits over 3 years of 3×10^{10} cm$^{-2} \cdot$ s$^{-1} \cdot$ sr$^{-1} \cdot$ GeV^{-1} at 10^{18} eV for their ARA-37 array [147, 148]. Nowadays, IceCube measurements provide a sufficient statistics up to $\sim 10^{15}$ eV, and ARA and ARIANNA are sensitive at $\sim 10^{18}$ eV. The energy gap between $\sim 10^{15}$ eV and $\sim 10^{18}$ eV are currently under attention. RADAR detection method, first proposed and used in 1941 by Blackett and Lovell, is investigated to radio detect UHE neutrinos in dense media like

1. As a reminder, protons ($Z = 1$) with an energy above 5×10^{19} eV might point to their sources. Cosmic ray protons with an energy smaller than 5×10^{19} eV do not point back to their sources.

ice [75, 76, 77], permitting to determine the cosmic neutrino flux from several 10^{15} eV up to few 10^{18} eV. The principle of detection is based on the fact that an UHE-neutrino induced air-shower will produce a ionization plasma left behind while propagating. This induced ionization plasma, depending on its properties (lifetime, free charge collision rate), can be detected via radar reflection technique [76]. From [77], it has been reported with a 2.36σ level of confidence a signal detection consistent with a radio reflection on ionization plasma produced by a run1-testbeam at the SLAC National Accelerator Laboratory of 10^{10} eV electrons on a high-density polyethylene target. This results is very promising, hinting the experimental feasibility of detecting showers via radar method. To date, the run2-testbeam is under analysis, and a proposal for an in-nature test is in progress. The astrophysical and cosmogenic neutrinos studies will play a major role for a better understanding of the astrophysical sources of cosmic rays, and may allow to reach the highest part of the spectrum. The advantage of the radio detection in ice is that the attenuation length is about 1 km, permitting to equip large volumes of dense medium with sparse antennas operating in [10 – 1000] MHz. Several experimental approaches, based on existing instrumental setups such as ARA or ARIANNA, are under investigation (deep/shallow/surface antennas, localization...) to propose a future neutrino observatory: the Radio Neutrino Observatory (RNO). The RNO aims to complete the IceCube flux from several PeV (radio detection threshold in ice $\sim$ 1 PeV) and to reach the highest energies to study the GZK neutrinos. Concerning IceCube and its upgrade IceCube-Gen2 [248], namely the radio part [249], the idea is substantially the same as the Auger upgrade. The combination of radio antennas and scintillators coupled with IceCube will permit to discriminate the electromagnetic part from the muonic part of the events. Radio could help to solve the problem of muon deficit in shower models, for instance. It will also give access to $X_{\max}$ measurements, and provide a veto for cosmic ray events and anthropic backgrounds noise to IceCube and future in-ice radio experiment of UHE neutrinos.

If the size of SKA-low is rather similar to the ones of current radio experiments, the proposed GRAND experiment [141, 247] aspires to cover an unprecedented huge surface. At the moment, in its final version GRAND will be composed of 20 sub-arrays of ~10,000 radio antennas, spread over $\sim$ 200000 km^2 in mountains. It will be an observatory for UHE gamma, cosmic rays and neutrinos, with the objective to bring answers to the origin of the UHE cosmic rays. It aims principally to detect Earth-skimming tau lepton induced air showers, and secondly cosmic ray induced air showers (see figure 10.2). The tau lepton is produced by a tau neutrino interacting under Earth via the charged-current channel interaction. Due to its short lifetime, the tau lepton could decay to produce an air shower when emerging from the Earth. Before to comment the instrumental setup, here are the information given in [141]: the antenna has been optimized for horizontal air showers, with a detection rate of $\sim$ 100 UHE cosmic rays per day above 10^{19} eV. The predicted angular resolution is $< 0.5°$, the energy and $X_{\max}$ resolutions are 15 % and $20 - 40$ g $\cdot$ cm^{-2} respectively. It is worth noting that the resolution on the energy and $X_{\max}$ are approximately the same than for AERA for instance. From the information given in [141, 247], several questions arise: the antenna being optimized for horizontal events, how to manage the important noise coming from the horizon? ; how to trigger them? ; how to discriminate the so-called cosmic events from parasitic events? ; the sought events being mostly horizontal, how to deal with such event horizontally propagating (direct, reflected and transmitted radiation fields, attenuation, far-field approximation...)? ; in a practical way, how to deploy so many antennas (200,000 for the last stage of GRAND), in areas probably far from ground infrastructures (one will have to balance between radio quietness and access and deployment easiness), and those on a surface equivalent to twice the area of Iceland? ; how to monitor them ? ; and finally, how to manage data? how to send data to the central DAQ (huge flow of data transfer)? how to treat them: by sub-array first, globally then?. So many questions that remain unanswered for the

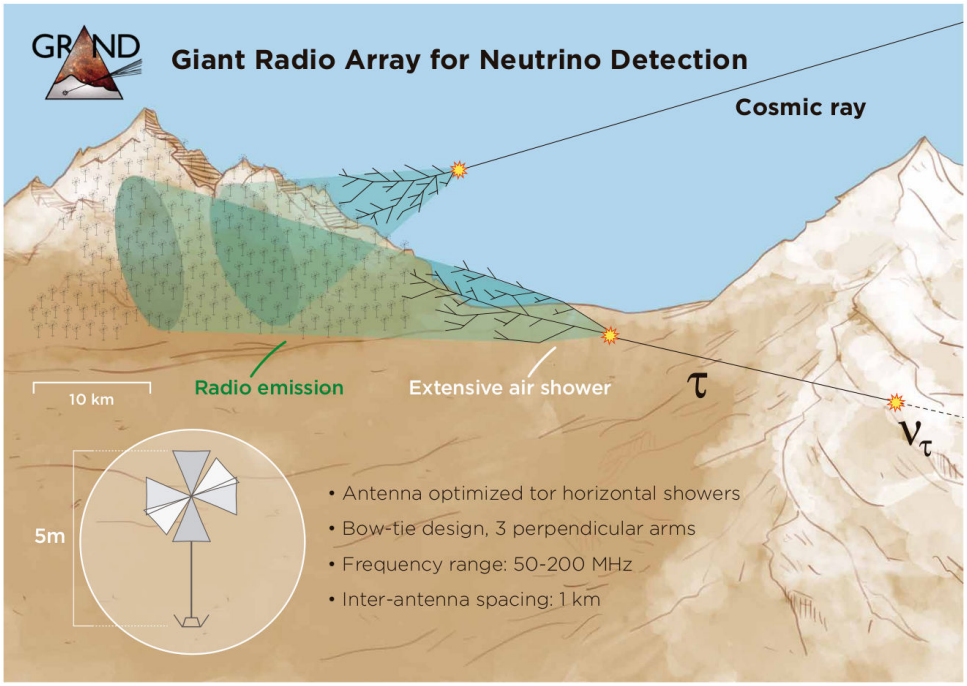


Figure 10.2 *– GRAND detection schematic diagram, for one sub-array of ~10,000 radio antennas. The lower left insert presents a sketch of the radio antenna designed for GRAND. From [141].*

moment. If the proposal seems interesting, many technical challenges will have to be met. One should insist on the fact that the expected tau neutrino observation rate is orders of magnitude below the expected cosmic ray rate which itself will be probably orders of magnitude smaller than the noise transient rate.

As we have seen through this chapter, the radio-detection technique, which is at a standstill for cosmic rays, remains valid for the study of neutrinos above the PeV. I would now like to mention a new project developed within my group of researches, which consists in radio detecting the showers initiated by gamma rays [250]. As for the radio detection of cosmic rays, the gamma astronomy has reached a mature stage, thanks to the instrumental efforts carried out and the results obtained by the H.E.S.S. [251], MAGIC [252], VERITAS [253] ground telescopes and the Fermi satellite [254]. A deep study of the gamma sky observed at energies from a few tens of GeV to more than 100 TeV will be possible with the CTA [255] project. However, one of the disadvantages of the technique is that the current telescopes observing the Cerenkov radiation of the showers created by the photons in the atmosphere have a rather low operating cycle, due to severe constraints similar to the ones of the fluorescence technique. Although CTA should be able to reach the energy limit currently held by HAWC [256] ($\simeq$ 100 TeV), the standard average sensitivity of other telescopes is restricted to a few tens of TeV for the most energetic photons.

For the UHECR detection, it is impossible to know in which direction to “look for" since the cosmic rays arrive from all the directions of the sky. Therefore the detection systems must have a maximum angular acceptance and availability in order to cover all possible directions of arrival. Concerning the radio method, assuming an energy threshold of detection found for the UHECR on observations with individual antennas (of the order of few 10^{16} eV), the detection of air showers generated by hundreds of TeV gamma rays, though they are similar in terms of charge contents to that of UHECR (see section 2.1.3), could be complicated, and the method seems not suitable in that case.

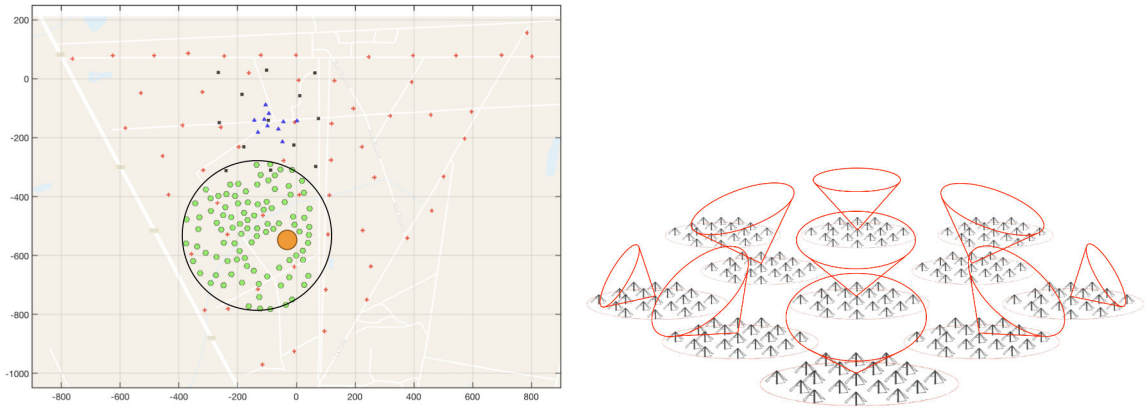


Figure 10.3 *– Left: Map of the Nançay observatory (North on top), featuring some of the setups currently involved in the radio detection of EAS. Red crosses feature the 57 CODALEMA autonomous radio detection stations, black squares the 13 scintillators, blue triangles the 10 antennas of the compact array. The location of the LOFAR FR606 LBA station is figured by the orange circle, while the NenuFAR radiotelescope mini-arrays are the green hexagons which fit into the black circle [250]. The scale is indicated on the axis. The area covered by CODALEMA is* $\simeq 1.1\ km^2$*. Right: scheme of principle, each mini array is analogically phased in different directions, covering the whole sky [257].*

However, the signals of several antennas may be combined in a given direction, in order to gain in sensitivity of detection, which also reduces the angular field of view: it is the principle of interferometers, widely used on past and current generations of so-called “digital" radio telescopes. The idea is therefore to combine the principle of detection of atmospheric shower electric field transients, used in CODALEMA, with the ability of antennas to be phased towards a known source direction. Contrarily to the case of UHECR, the direction where the signal from the high energy gamma ray comes from is known, allowing to improve the gain in detection sensitivity. Moreover, by combining several antennas to form one single beam, the detector becomes less sensitive to transient noises generally located close to the horizon. As already discussed, the duty cycle for this type of detection is close to 100 %, since the day/night alternation and the weather have no influence on the detection itself. Based on our experience of ultra-fast radio transient detection and on the context of the Nançay radio astronomy station which hosts the NenuFAR radio telescope [163] (see figure 10.3 left), we propose to explore the possibilities offered by this idea and to try for the first time to detect the radio signal produced by ultra-high energy gamma of a few hundred TeV to a few tens PeV from an identified astrophysical source - provided they can produce gammas at such energies.

Also known as the LOFAR Super Station (LSS) in Nançay, NenuFAR is a digital radio telescope consisting - in the long term - of 1824 crossed-dipoles antennas arranged in 96 mini-arrays (hereafter MA_{19}) of 19 antennas [2]. It has recently been recognized as a "pathfinder" of SKA. The NenuFAR antennas use the active amplifier developed for CODALEMA and are identical to the ones of its Compact Array. The 96 MA_{19} are distributed over an approximate circle of 500 m radius (0.2 km^2). In classical operation, each of the MA_{19} is analogically phased in any direction of the sky (see figure 10.3 right) and in the whole or any sub-band of the [10-85] MHz frequency window, leading to an instantaneous sensitivity $\sqrt{19}$ times higher than that of a single antenna. The whole NenuFAR thus has an instantaneous sensitivity much higher than that reached by a single LOFAR station of 96 single antennas [125], like the FR606 unit present at Nançay.
As already mentioned, the specificity of the gamma-ray sources is that we know their

2. In 2018, 54 MA_{19} of those 96 are installed and operational

position in the sky, thus the arrival direction of the high-energy photon, contrarily to the UHECR case. It is therefore possible to use one or several mini-arrays phased together in the direction of the source and use them as a trigger on transient events. We can take benefit of the large memory depth of the remaining MA_{19}, which would be triggered by the trigger arrays, to roll-back in time and find the transient in their memories. Each MA_{19} would then be a sampling point of the shower footprint at ground, exactly as it is done in UHECR radio detection arrays with single antennas. This would allow knowing the electric field distribution and use it to recover the properties of the primary photon, notably its energy thanks to simulations. We have shown in [250, 172] that it would be possible to detect gamma with energy above 100 TeV, meaning coming from galactic sources, such as the Crab Nebula which is now known as a source emitting gamma up to 450 TeV [258]. It has also been shown that gamma of 1 PeV should be detectable and distinguishable from cosmic ray events.

Ultimately, the radio detection of extensive air showers is a solid approach, able of providing all the pieces of information concerning the primary cosmic ray that has initiated the shower. Although not reaching the expectations in terms of statistics at the highest energies for the cosmic ray domain, and although being at a standstill for this physics, there are several opportunities and new projects around the use of this method for gamma and neutrino physics. Despite various development levels reached so far and various expected difficulties, lets hope that they will benefit from the efforts made during the last 20 years and pursue fruitfully the development of the radio detection of extensive air showers.

"Never jump the fence if you're not willing to face what's on the other side."

Dexter Morgan, Dexter, Season 7: Surprise, Motherfucker!

CHAPTER 11

Summary

Discovered more than a hundred years ago, cosmic rays are produced by galactic cataclysmic phenomena, but also from the confines of the Universe. What are their origins? How are they created? How are they accelerated? So many questions remain unanswered today, although recent studies attempt to address these challenges.

Cosmic rays, messengers of the cosmos

Despite the discovery of cosmic rays by Victor Hess in 1912, more than a hundred years ago, many questions remain unanswered today: what are cosmic rays, how are they created and where do they come from? To date, we know that the Earth is constantly bombarded by these cosmic rays, which are charged particles of extraterrestrial origin. Cosmic rays are produced and accelerated by astrophysical sources. These sources must have particular characteristics to be able to produce and accelerate these particles to such energies, meeting the Hillas criterion which involves the size and magnetic field of the source. The main candidates for the creation/acceleration of cosmic rays are type II supernovae (the last stage of a star's life associated with a gigantic final explosion), pulsars, which are fast rotating and highly magnetized neutron stars that periodically emit an electromagnetic signal, gamma-ray bursts (short bursts of gamma-ray of a few MeV , which can be produced during the coalescence of binary systems of neutron stars) and active galaxy nuclei (black hole of a few 10^6 to 10^{10} solar masses ejecting a beam of matter at 99 % of the speed of light).

The flux of cosmic rays decreases considerably with their energy and extends over ~32 orders of magnitude, for 12 orders of magnitude in energy. This remarkably regular behaviour, unique in physics, could sign a mechanism for the universal production of these cosmic rays. This is where their study is important, because the enigma to be solved lies in this spectrum, and in its remarkable regularity. Below 10^{14} eV, it is still possible to detect them directly thanks to dedicated satellites such as the AMS experiment installed on the International Space Station or balloon experiments such as those conducted by Victor Hess in particular. Beyond an energy of 10^{14} eV, it is almost impossible to detect primary cosmic rays directly, due to the low flux arriving on Earth: 1 cosmic ray per m^2 per year at $\sim 3 \times 10^{16}$ eV, falling to 1 cosmic ray per km^2 per century beyond 10^{20} eV.

The study of cosmic rays with energies greater than 10^{14} eV is done indirectly. Indeed, when a cosmic ray arrives on Earth, i.e. at the top of the atmosphere, it interacts with the atoms of the atmosphere. These interactions can take place at energies much higher than those reached in the laboratory. It is common to compare the energy of the most energetic cosmic rays to the 13 TeV in the mass centre in proton-proton collisions produced at the Large Hadron Collider (LHC) at CERN: the energy reached by cosmic rays is 800 times higher in this context. We would prefer here to compare the energy of the LHC proton beam, which is 6.5 TeV, with the incident energy of the most energetic cosmic ray ever detected, namely 3.2×10^{20} eV, which is therefore orders of magnitude higher! These 50 J correspond to a tennis ball going at 160 km $\cdot$ h^{-1}. The circumference of the LHC is

about 27 km. In comparison, with the same magnetic field, we would then have to build a particle accelerator with the circumference of Mercury's orbit, or 10^7 km, to reach these energies. In any case, these energies are not reproducible on Earth.
When the cosmic ray enters the atmosphere, the first interaction takes place at an atmospheric depth noted X_1 (which is the *point of first interaction*), generating a cascade of secondary charged particles commonly referred to as *shower of particles*. The number of particles in the shower will increase to the atmospheric depth $X_{\max}$, then decrease until the shower extinction. This shower is detectable on the ground, and provides us with information on the primary cosmic ray: it is indirect detection. The objective is to go back to the characteristics of the primary cosmic ray that generated the shower of particles, thus determining its arrival direction, its nature and its energy. The experimental challenge lies in the ability to determine these three quantities with a sufficient level of accuracy. This detection can be done during the development of the shower in the atmosphere, via the fluorescence technique which consists in collecting the fluorescence light emitted by the deexcitation of the nitrogen during the passage of the charged particles of the shower, or directly on the ground, with dedicated detectors counting the particles. However, the arrival of a shower on the ground depends on several parameters, such as the energy of the primary cosmic ray or its angle of arrival. At 10^{18} eV for example, the cosmic ray flux is estimated at only 0.2 particles per km^2 per day. Also, the solution is to build detectors covering huge areas, such as the Pierre Auger Observatory in Argentina, which covers $\sim$ 3000 km^2 (surface area equivalent to the urban surface of Nantes and its countryside).

It is customary to call cosmic rays of ultra-high energy (UHECR) those with an energy greater than 10^{17} eV. The objectives of the study of the RCUHE are to be able to fully characterize these messengers of the cosmos:

— What are the sources capable of creating these primary particles (active galaxy nuclei, gamma-ray bursts)?
— What are the acceleration mechanisms (Fermi acceleration by relativistic shocks)?
— What are they (proton, iron nuclei, intermediate elements)?

Like other particles such as photons resulting from LHC collisions that probe the matter [1], the anti-neutrinos that can probe nuclear reactors [2, 3] or gamma photons from radionuclides for medical imaging [4], cosmic rays are used as probes of particle physics and of the universe. One of the recent applications is the use of cosmic muons to probe the Great Pyramid of Khufu [5].

Despite the advent of major cosmic ray observatories, such as the Pierre Auger Observatory, which today make it possible, through the study of large atmospheric showers, to reconstruct the characteristics of the most energetic cosmic rays in the Universe, the acceleration mechanisms, the astrophysical sources and the high-energy part of the spectrum remain poorly understood. Being charged particles, they are deflected by galactic and extra-galactic magnetic fields, making it difficult to correlate arrival directions with point sources in the sky.

In order to better constrain the characteristics of cosmic rays and provide answers to the questions still open today, a large statistic, i. e. a large number of detected events, is necessary. It is therefore important to use detection methods with an optimal duty cycle. Radio detection of atmospheric air showers can solve this problem, with a duty cycle of $\sim$ 100 % (against only $\sim$ 14 % for the fluorescence method). First developed in the 1960s, abandoned due to technological constraints, and then completely revised in the digital era since 2002, this method is based on the fact that, when the shower develops in the atmosphere, the charged particles in movement generate the emission of a very short electric field pulse, detectable on the ground with dedicated radio antennas, over a large frequency band.

Radio emission of extensive air showers

The radioelectric emission produced during the development of the shower is coherent and has two main origins: the negative charge excess in the shower and the variation of transverse current induced by the Earth's magnetic field. This last mechanism, associated with the geomagnetic field, is called *geomagnetic mechanism*. The electrons and secondary positrons present in the showers are accelerated in the Earth's magnetic field, which leads to the emission of an electric field. This effect is related to the Lorentz's force: $\overrightarrow{F} = q\overrightarrow{v} \times \overrightarrow{B}$ where q is the charge particle, $\overrightarrow{v}$ its velocity and $\overrightarrow{B}$ the Earth's magnetic field vector. For particles moving along the shower axis, the resulting currents will be on average perpendicular to the axis of the shower and are called "transverse currents". These currents vary according to the evolution of the shower: the number of secondary particles increases to reach a maximum at atmospheric depth $X_{\max}$, then decreases until the shower extinction. This temporal variation in the number of particles in the shower leads to a temporal variation in the transverse currents inducing the emission of an electric field. Its polarization is linear, aligned along $\overrightarrow{F}$. The intensity of the electric field depends on the arrival direction of the primary cosmic ray, in particular the angle α between the arrival direction of the shower and the direction of the geomagnetic field.

At the second order, the mechanism of *negative charge excess* complements the geomagnetic contribution. At high energy, the processes of pair and bremsstrahlung production dominate. As the shower develops, the average energy decreases and other processes appear, such as Compton scattering, delta ray production and positron annihilation with preexisting free atmospheric electrons. This results in an excess of negative charges of 10 to 20 %. The ionization electrons are contained in the front of the shower, while the positive ions, much heavier, remain behind the front. As the shower develops, the absolute negative charge in the shower grows, reaches a maximum and finally decreases when the shower dies out. There is again a temporal variation of this charge excess, which results in the emission of an electric field. Unlike the geomagnetic mechanism, the electric field vector is oriented radially with respect to the axis of the shower and its orientation depends on the position of the observer.

The work carried out during this thesis aims to continue the studies of this radio signal resulting from the superposition of these two mechanisms and which takes the form of a fast electric field transient, lasting a few tens of nanoseconds, that can be detected by high bandwidth antennas and fast acquisition systems, such as those used by the CODALEMA and EXTASIS experiments. The originality of this work lies essentially in the range of the frequency range studied, from 1 to 200 MHz, for the first time, on the same detected events.

Radio antennas to observe the sky

Since 2002, the CODALEMA experiment, located within the Nançay Radio Observatory, has been one of the pioneering and promoting experiments in the renewal of radio-detection of extensive air showers, which is now being adopted by a multitude of experiments around the world. Coupled with an increasingly sophisticated understanding of the processes involved in radio emission, allowing the use of high-performance simulation codes, the radio detection of showers is now reaching a level of maturity sufficient to be competitive with more traditional detection methods. Radio detection is usually performed in a frequency range restricted to [20-80] MHz (noted MF in the following for medium frequencies) by experiments such as AERA, Tunka-Rex, TREND, Yakutsk or LOFAR. This restriction is mainly due to low and medium frequency (AM, FM bands) artificial broadcasting, which has led to relatively low sampling rates ($\sim 200\ \mathrm{MS} \cdot \mathrm{s}^{-1}$) but

sufficient to properly sample signals on 60 MHz. However, CODALEMA works with a sampling rate of 1 GS $\cdot$ s^{-1}, which allows observations to be extended to the [110-200] MHz band. The main limitation of the frequency band is then due to the antenna bandwidth used,[20-200] MHz for CODALEMA, hereinafter referred to as "Extended Medium Frequencies (EMF)". To date, the CODALEMA experiment is composed of:

— A square array (0.34×0.34 km^2) of 13 particle scintillator detectors.
— A set of 57 so-called “standalone” cross-polarized antennas, operating in the EMF band, distributed over $\sim$ 1 km^2.
— A so-called “Compact Array” of 10 cross-polarized antennas, arranged in a star shape of 150 m extension and whose signal acquisition (in EMF band) is triggered by the particle detector.

The array of autonomous stations is self-triggered on the radio signals, which means that each station is independent. Transients from showers or any other sources (noise, aircraft. . .) are either stored on a remote disk for offline analysis or sent directly to a central acquisition (DAQ) capable of building the event online based on multiple station signals, meeting multiple selection criteria. Cross-checking can be performed offline with events detected by particle detectors or one of the instruments externally triggered by particle detectors (compact array or low frequency antennas).

CODALEMA is today the host experiment of EXTASIS, an array of 7 low-frequency antennas triggered by particle detectors, which takes advantage of existing infrastructures. EXTASIS aims to re-examine the [1-10] MHz frequency band, and in particular to study the so-called "Sudden Death" contribution, the electric field expected and emitted by the shower front when it reaches the ground and when the particles suddenly disappear.

Galactic radio emission to make everyone agree

In the past, many disagreements have been reported about the amplitudes of the electric field detected for a given energy, in particular because of the difficulty in calibrating the radio detectors then in use. Part of the work of this thesis was to propose a method for calibration of the CODALEMA radio antennas (autonomous stations and compact array antennas). The idea was to use the Galactic radio emission as an external source to calibrate the antennas. In order to carry out this calibration, two things are necessary: a description of the antenna and a description of the galactic radio emission.

Concerning the second point, we compared three models of galactic radio emission. We showed that the models were in agreement with each other, allowing us to choose one without risk of biasing our procedure. We chose the Global Sky Model (GSM) from Oliveira-Costa et. al. We then coupled a simulated antenna model with NEC software with the galactic emission model. We calculated the contribution of each part of the sky, by convolving the sky map obtained with GSM by the antenna lobe as a function of local sidereal time (LST). A simulation of the signal measured at a given LST time is obtained by summing each pixel of the convoluted sky map with the antenna response. The result obtained is called *transit curve*, which can be compared to actual data. Calibration is done by sub-band of frequencies of width 10 MHz and by step of 5 MHz.

We have shown that, in the context of a cross calibration between antennas (as opposed to an absolute calibration), the galactic emission model or the environmental conditions near the antenna (such as ground composition) did not have an impact on the final results. We have highlighted a variation present on some autonomous stations, which we have assimilated as a first approximation to a seasonal variation. In order to take into account its effect, we decided to calculate the calibration coefficients month by month, and demonstrated that this made it possible to take this variation into account

and to correct its effects. The simulation of the signal obtained is in accordance with the data detected by the antennas. This observation is of paramount importance, because in addition to allowing us to calibrate our antennas, the use of galactic radio transmission allows us to validate the antenna model used. This antenna model is at the centre of most of the analyses of this thesis, in particular the one consisting in convolving the simulations of the electric field of the shower by the antenna response and comparing them with the real data to estimate the parameters of the shower and thus of the primary cosmic ray.

Before calibration, a significant disparity in signal level existed between the antennas, in the order of 70 µV for autonomous stations and 40 µV for antennas in the compact array. These disparities are greatly reduced after calibration, and are in the order of 20 µV for autonomous stations and 5 µV for antennas in the compact array. These residual disparities are much lower than the RMS of the noise recorded by the antennas during an event, which allows a first estimate of the quality of the calibration.

Cosmic events

The ground measurement of the electric field emitted during the development of the shower is made difficult by the background noise. In [20 – 200] MHz, it is composed of a predictable galactic background noise (which has allowed us to calibrate our antennas) and a parasitic background noise due to artificial emissions, interferences or electric discharges from storms on a global scale. This background noise is one of the limitations of autonomous detectors: triggering the acquisition on noise results in a significant dead time, preventing detection on a so-called "cosmic ray" event. It is therefore essential to have a good understanding of the background environment in order to develop an analytical framework capable of generating the purest possible batch of radio events that can be analysed at a later stage. During this thesis work, different noise rejection algorithms were explored in order to develop and define a complete analytical framework to reconstruct the characteristics of primary cosmic rays. They are based on the signal rise time, the compacity of the event and the polarization behaviour of the event. We have shown that, although exploratory, these three coupled methods allow an interesting selection/rejection rate (99.96 % of the recorded events are rightly rejected, in other words 99.96 % of the recorded events are fortuitous events not resulting from a shower initiated by a cosmic ray).

In [1.7 – 3.7] MHz (the low frequency band of the EXTASIS experiment), background noise is the superposition of three dominant contributions: the minimum thermal noise of the atmosphere (in other words, its luminosity related to its temperature) and so-called atmospheric noise, consisting of the contribution of electrical discharges from storms on a global scale, and all types of man-made noise (radio transmitters, anthropic transients). Atmospheric noise conditions severely hamper low-frequency detection. One of the main contributions of this thesis was the development of an analytical framework dedicated to the low frequency band, the associated signal being often drowned in noise.

Information from messengers of the cosmos

We have shown that an advanced knowledge of the instruments is necessary. Indeed, to return to the characteristics of primary cosmic rays, namely energy and composition, systematic comparisons between data and simulations must be made. The calibration of our radio detectors as well as the convolving of simulations by antenna response to compare them to the data requires a detailed description and understanding of the detectors used. One of the important contributions of this thesis was the development of an

automatic analysis framework coupled with the antenna calibration and noise rejection algorithms mentioned above in order to be able to reconstruct all the properties of the showers with the radio signal, such as the direction of arrival, the energy but also the nature of the primary cosmic ray. We started from a simulation/data comparison method already used within the group, and generalized it in order to exploit all the capabilities of CODALEMA instruments: decoupling of the EW and NS polarizations of the antennas (allows to keep the polarization and polarity information, which was lost with the old method using the quadratic sum of the two polarizations) and use of the entire frequency range (allows to better reconstruct very inclined shower events, i.e. with a large zenithal angle). We have demonstrated that the generalization of the method allows better results to be obtained on the reconstructed sheaf parameters, the uncertainties on the parameters obtained with the new method becoming comparable to those obtained by the fluorescence technique for example.

Concerning EXTASIS, we observed 25 low-frequency events detected in coincidence with CODALEMA instruments and estimated a detection threshold of 23 ± 4 $\mu V \cdot m^{-1}$ based on comparisons with simulations. We also found a strong correlation between the observation of low-frequency signals and the value of the atmospheric electric field at the time of detection. We showed that the detection of the sudden death signal was very unlikely for low-altitude experiments, such as EXTASIS, and that in order to increase the chances of detecting this contribution, higher-altitude sites such as the Pierre Auger Observatory were to be preferred.

Concerning CODALEMA, after selection and reconstruction of events, we were interested in the mass composition of cosmic rays. The analytical framework put in place allowed us to obtain sufficient statistics for the number of events reconstructed. We have shown, in a preliminary way, that the results obtained concerning the estimation of energy and the parameter X_{max} for CODALEMA events were in agreement with the results obtained by other experiments such as LOFAR or Auger. Various tests were carried out to establish the results obtained, in particular by applying different cuts/selections to the batch of events to be processed. The final result obtained is in agreement with those proposed by the community, thus validating all the steps of our analysis framework, from the calibration to generalization of the simulation/data comparison method and the improvement of event selection.

Outlook

Although we have, throughout this work, shown that radio detection technique is capable of providing us with the expected information about primary cosmic rays, namely their energy and nature, the picture is not all rosy. Indeed, when we look at the efficiency of the radio detection, as well as the percentage of reconstructable events, the conclusion is that radio detection technique cannot meet all expectations in the short term, particularly with regard to increasing statistics. Although the radio detection technique is now mature enough to compete with other techniques in terms of reconstruction quality, the observation is as follows: radio detection seems to have made little contribution to the cosmic ray domain. Indeed, other techniques have already in the past, and continue to provide the necessary and expected information: the composition of cosmic rays of the Pierre Auger Observatory, or Telescope Array, is provided by fluorescence light telescopes.

While the technique of radio detection of cosmic rays seems to be reaching its limits, new projects are emerging, such as *AugerPrime* which aims to instrument the 1600 Auger tanks with radio antennas to study inclined showers, or the SKA "cosmic rays" key

program. At the same time, the radio detection of ultra-high energy neutrinos is widely investigated, by the ARA or ARIANNA experiments in Antarctica, but also by new projects such as RNO and IceCube-Gen2.

Similarly, a project within the group is emerging for the radio detection of atmospheric showers initiated by gamma rays. Unlike cosmic rays, whose sources are not identified and which therefore require instruments covering the entire sky, gamma ray sources are known. It then becomes possible, thanks to a radio telescope, to point directly towards the sources to increase the detection sensitivity, which is necessary since gamma rays do not reach the energies of cosmic rays, and the produced shower develops very high in the atmosphere, reaches only slightly the ground and the associated radio signal is weaker. This is the objective of the GammuFAR project, which involves the use of the NenuFAR radio telescope located within the Nancay Radio Observatory.

This manuscript deals with a complex subject, both in terms of the complexity of the physical processes involved and their plurality. It has been constructed in such a way that the reader has a progressive vision and understanding of the elements studied during these three years. The studies carried out during this thesis and developed throughout this manuscript are at the boundary between different disciplines, such as particle physics, cosmic ray physics and astrophysics. The manuscript is divided into four parts. The first part introduces the notion of cosmic rays, gives the minimal theoretical framework to the understanding of the domain and concludes with the study of large atmospheric showers. The second part of the manuscript focuses on the radio signal emitted by the shower, detected by the CODALEMA and EXTASIS experiments, whose instrumental setup is discussed. We then detail the calibration method used, as well as the methods for selecting the events detected. The third part presents the results obtained during this thesis work, in particular the results obtained with EXTASIS on the low frequency band [1.7 – 3.7] MHz, but also the results obtained with CODALEMA on the composition of cosmic rays and the analyses developed to achieve this. Finally, the last part contains a conclusion on the work carried out, and opens with a critical discussion on the technique of radio detection of cosmic rays and other astroparticles.

Part V

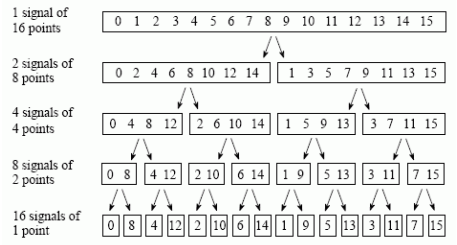
Appendices

APPENDIX A

Signal processing tools

A.1. The Fast Fourier Transform

In signal processing, the use of Fourier transform is essential. But when it comes to digitally calculate it, this is the discrete Fourier Transform (DFT) that must be used, and there are different possibilities to compute the DFT. In this thesis, we chose to use the Fast Fourier Transform (FFT) algorithm. This is the basis of all our procedures (used in filtering, wavelets, PSD $\cdots$) and this is why we decide to detail how the FFT works. The `fft` function in Python computes the one-dimensional n-point discrete Fourier Transform (DFT) with the FFT algorithm. Figure A.1 shows the process of the FFT. The N time domain signal samples are decomposed by the FFT algorithm into N-single time domain samples (figure A.1.(a)), usually carried out by the inversion of bit number (figure A.1.(b)). Then, the N-single frequency values corresponding to the N time domain samples are calculated, giving the frequency spectrum corresponding to the input signal. In order to match up sample numbers with their good frequency, the `fftshift` function is used. Figure A.2 shows an illustration of this DFT on an actual recorded signal.



(a) *The FFT decomposition.*

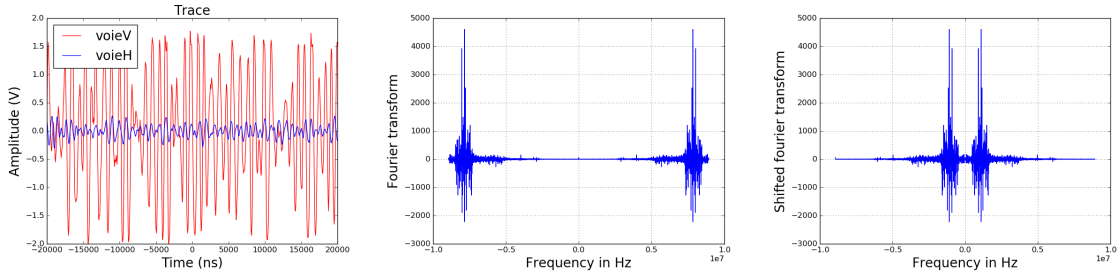
Sample numbers in normal order		Sample numbers after bit reversal	
Decimal	*Binary*	*Decimal*	*Binary*
0	0000	0	0000
1	0001	8	1000
2	0010	4	0100
3	0011	12	1100
4	0100	2	0010
5	0101	10	1010
6	0110	6	0100
7	0111	14	1110
8	1000	1	0001
9	1001	9	1001
10	1010	5	0101
11	1011	13	1101
12	1100	3	0011
13	1101	11	1011
14	1110	7	0111
15	1111	15	1111

(b) *The FFT bit reversal sorting.*

Figure A.1 *– The FFT procedure [259].*

A.2. Frequency filtering

To filter in frequency the recorded signals, we need to set to zero the Fourier coefficient corresponding to the frequencies that we want to filter, and apply the inverse FFT operation to come back in time domain. Figure A.3 shows different signals and their filtered response. The upper plot shows the simulated expected pulse of a sudden death signal and its [1.7 – 3.7] MHz filtered response. The middle plot shows a simulation of the standard geomagnetic pulse and its [20 – 80] MHz filtered counterpart, as it may be observed in standalone stations. The lower plot shows a Ricker signal and its filtered response in [1.7 – 3.7] MHz, used as *mother wavelet* for the selection procedure using



(a) *Signals.* ***(b)*** *Fourier transform of signal.* ***(c)*** *Shifted Fourier transform.*

Figure A.2 *– Basic processing of signals. Left: the vertical (red) and horizontal (blue) signals, in time domain. Center: the Fourier transform for the vertical polarization, obtained with the FFT algorithm of Python. Right: after applying the shifted Fourier transform function.*

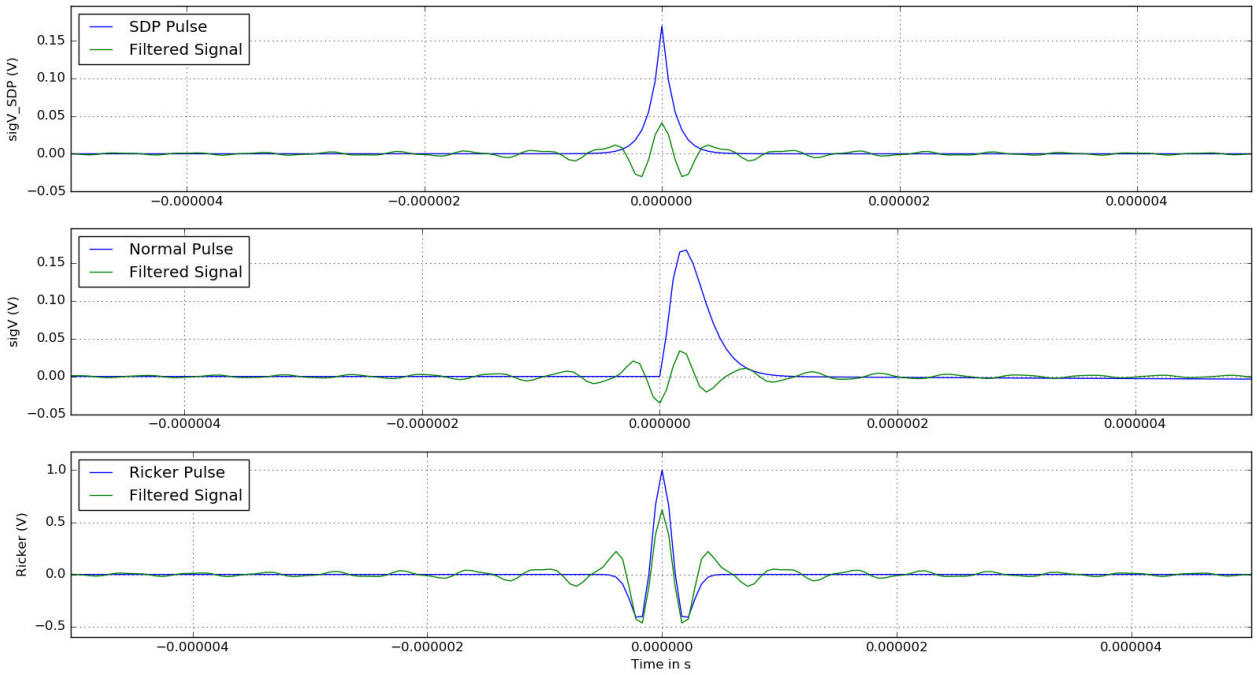


Figure A.3 *– Filtered simulated pulses (see text for explanations).*

the wavelet transforms as discussed in section 6.4.4 page 145.

Most of the filtering procedures in this manuscript are made with a rectangle filter, on which we apply a Tukey window (also known as the tapered cosine window). It can be assimilated to a cosine lobe of width $\alpha(N+1)/2$ convolved with the rectangular window used of width $(1-\alpha/2)(N+1)$, where α is a parameter to adjust ($\alpha = 0$ corresponding to a rectangular window) and N the number of samples of the signal. We chose $\alpha = 0.2$. Some of the procedures have been made with a Butterworth filter 6th order. Both Tukey and Butterworth filters are designed to have a flat response inside their passband.

A.3. Power spectrum density

The antenna converts electric field $\mathrm{V}\cdot\mathrm{m}^{-1}$ received from every space directions to a voltage density (V) applied to the LNA input. The conversion factor is called the antenna effective length which depends on the frequency and also the space direction as the

antenna receiving pattern is not isotropic. In our case of an active antenna we include the LNA response (voltage gain) in the antenna effective length. Thus the antenna "output" is the LNA output whose reference impedance Zref is 50 Ω. In this manuscript, we use the power spectrum density (PSD). For this, we have to perform the FFT algorithm of a signal in V generated at the LNA output and convert it to a voltage density ($V/Hz^{1/2}$) depending on the frequency. We also convert voltage density to power density ($dB \cdot m \cdot Hz^{-1}$) using the 50 Ω reference value this way:

$$\begin{aligned} dB \cdot m \cdot Hz^{-1} &= 10 \log \left[(V/Hz^{1/2})^2 / Zref / 10^{-3} \right] \\ &= 10 \log \left[(V/Hz^{1/2})^2 / Zref \right] - 30 dB \cdot m \end{aligned} \tag{A.1}$$

We can thus apply the conversion shown above to get the PSD in the correct units, and build the PSD for any individual signal or the map of the average PSD for one day, as shown in figure 8.2.

APPENDIX B

HEALPix and map projections

B.1. Data analysis and visualization on the sphere

HEALPix is an acronym for Hierarchical Equal Area isoLatitude Pixelisation, comporting a set of functions for astrophysics and cosmology data representation. It provides algorithms allowing to subdivide the surface of a sphere in pixels of same area, and moreover to arrange them into lines of equal latitude (simplification of spherical harmonic analysis), as shown in figure B.1. Then, a Mollweide projection can be performed. The three essential properties are:

1. the original sphere is divided into mosaic (curvilinear quadrilaterals). The resolution is determined as $12 \times NSIDE^2$. In our case, the $NSIDE$ is set to 64, corresponding to a resolution of 49,152 pixels ;
2. the areas of these pixels are identical ;
3. and finally, these pixels are arranged into lines of equal latitude.

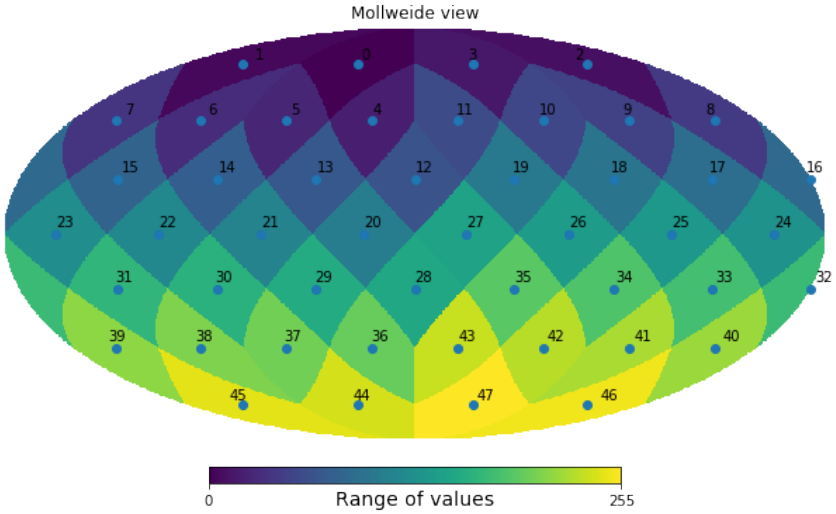


Figure B.1 *– Hierarchical Equal Area isoLatitude Pixelisation: subdivision of the surface of a sphere in pixels of same area of the original spherical surface.*

This is what is done for the figure 3.3. The angles of arrival direction reconstructed by CODALEMA is used, and a Mollweide projection is made, figure B.2.(a), and is smoothed by a native HEALPix function (b).

B.2. Lambert projection

To came back in the local frame, we proceed to a Lambert projection, which is a conic map projection. The latitude and longitude corresponding to our origin, in that

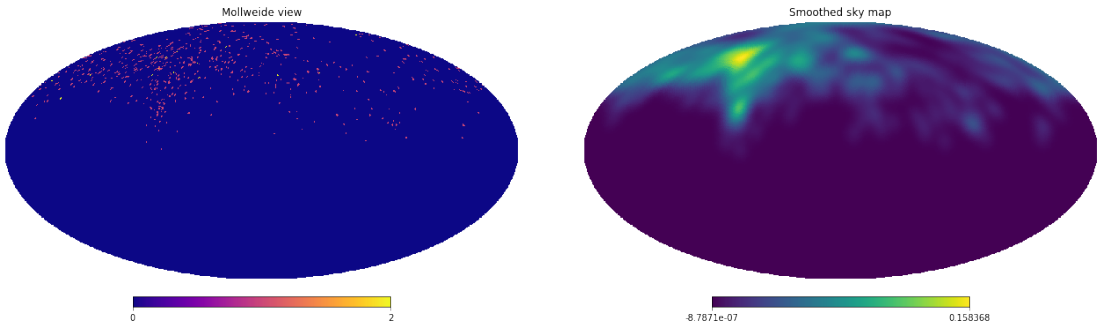


Figure B.2 *– Left: Mollweide projection of the arrival direction reconstructed by CODALEMA. Right: same as left, smoothed by a native HEALPix function.*

case the experiment site, and the angles given by the Mollweide map are needed. The transformation can be expressed as:

$$\begin{aligned} cp &= \cos(\theta), \\ sp &= \sin(\theta), \\ \mathrm{d}\phi &= \phi - \lambda_0, \\ s\mathrm{d}\phi &= \sin(\mathrm{d}\phi), \\ c\mathrm{d}\phi &= \cos(\mathrm{d}\phi), \\ k &= \sqrt{(2./(1 + sp_1 \times sp + cp_1 \times cp \times c\mathrm{d}\phi))}, \\ x &= -k \times cp \times s\mathrm{d}\phi, \\ y &= k \times (cp_1 \times sp - sp_1 \times cp \times c\mathrm{d}\phi). \end{aligned} \tag{B.1}$$

where θ, ϕ are the angle given by the Mollweide map, λ_0 is the longitude of the experiment site, sp_1 and cp_1 are respectively the sinus and cosine of the latitude of the experiment site. This projection gives the figure 3.3 page 51.

APPENDIX C

Global Data Assimilation System

As explained in chapter 3, some characteristics of the atmosphere are needed to simulate the electric field emitted by EAS, namely the air refractive index and the density (see the derivation in section 3.3.2) which depends on the temperature, the total pressure and the relative humidity of the atmosphere. These characteristics can be derived from the GDAS, by using the data coming from NCEP's GDAS which gather measurements of the physicochemical characteristics of the atmosphere.

The data contained in the uploaded file from [112] are on a latitude-longitude grid, with a spacing of one degree, corresponding exactly to 119.1984 km at the reference point and contains a week of data. Thus, we firstly need to extract the data corresponding to the location of the CODALEMA experiment, which is 33.58 in latitude and −16.97 in longitude. This location corresponds to the grid point (344.0124.6), however the used script only allows to extract full degrees location. Thereby, the location used in all our simulations corresponds to the grid point (344.0125.0), meaning that an approximation of 44.5 km is done on the location of the experiment.

Moreover, the GDAS is run 4 times a day and contains measurements for every 3 hours and for 24 geopotential heights (vertical levels) including a measurement at the surface. The description of the vertical levels can be found in table 3 of [112], with the corresponding absolute pressure. A complete description on how the GDAS data are converted for a use in SELFAS3 can be found in [113].

An example of GDAS data after extracting the values corresponding to the location of CODALEMA is shown in figure C.1. The set of data corresponding to an interval of three hours is composed of 24 lines. The first one contains the Meteorological Fields at ground level which are described in table 2 of [112]. The following 23 lines contains the Meteorological Fields for the 23 vertical levels. The first three columns indicate the year, the month and the day. The fourth column indicates the time slot (3 hours by 3 hours), followed by vertical level number (from 0 to 23). Columns 6 to 11 contain the U parameters which can also be found in table 2 of [112], such as: geopotential height, temperature and relative humidity.

```
19     3    29     0     0   1009.31   1029.40      0.00      0.07      0.07    -48.52
     0.00     91.89     -2.75     -2.11    279.20     13.00    150.14      0.00
    -0.03    283.76      5.70    152.69    277.89      0.00      0.00      0.00
     0.00      0.00      0.00     -0.52      0.00      0.00     13.00
19     3    29     0     1    239.36    281.34     -6.76     -4.21     -0.00     79.03
19     3    29     0     2    449.32    282.89    -11.51     -3.21     -0.00     67.03
19     3    29     0     3    665.07    282.44    -10.27     -0.71     -0.00     65.03
19     3    29     0     4    884.78    280.61     -8.77      0.79     -0.00     68.03
19     3    29     0     5   1110.69    278.69     -7.27      1.80     -0.00     74.02
19     3    29     0     6   1573.83    274.96     -6.26      1.80      0.00     72.03
19     3    29     0     7   2062.50    274.39     -6.27     -0.70      0.00     23.00
19     3    29     0     8   2578.92    272.74     -5.51     -0.20      0.00     18.00
19     3    29     0     9   3125.86    268.51     -4.27     -0.71     -0.00     33.00
19     3    29     0    10   3703.97    264.01     -4.07     -0.90     -0.00     57.00
19     3    29     0    11   4318.98    259.93     -4.69     -1.33     -0.00     75.00
19     3    29     0    12   4976.26    255.73     -5.46     -1.84     -0.00     78.00
19     3    29     0    13   5683.49    250.90     -5.47     -2.01     -0.00     75.00
19     3    29     0    14   6448.19    244.91     -5.76     -2.24     -0.00     73.00
19     3    29     0    15   7280.83    238.23     -5.12     -2.05     -0.00     69.00
19     3    29     0    16   8197.60    230.51     -3.90     -1.77     -0.00     85.00
19     3    29     0    17   9218.32    222.11     -3.18     -2.93     -0.00     85.70
19     3    29     0    18  10385.61    216.51     -6.41     -3.78      0.00     36.10
19     3    29     0    19  11803.10    218.85     -4.49     -4.15      0.00      6.10
19     3    29     0    20  13651.83    219.12     -1.92     -5.34     -0.00      1.60
19     3    29     0    21  16231.76    215.91     -0.18     -5.69      0.00      1.30
19     3    29     0    22  20606.15    216.40     -1.32     -1.70      0.00      0.00
19     3    29     0    23  26476.78    222.90     -7.60    -10.09      0.00      0.00


19     3    29     3     0   1009.04   1028.60      0.00      0.04      0.02    -40.90
     0.00     93.99     -2.44     -0.93    278.17      0.00    150.14      0.00
    -0.01    284.51      6.40     44.79    277.06      0.00      0.00      0.00
     0.00      0.00      0.00     -8.55      0.00      0.00      0.00
19     3    29     3     1    232.28    280.03     -5.24     -1.06     -0.00     85.04
19     3    29     3     2    442.56    281.78     -8.50      0.94     -0.00     72.04
19     3    29     3     3    656.64    282.38     -8.00      1.94     -0.00     61.00
19     3    29     3     4    876.68    280.70     -6.99      1.44     -0.00     63.01
19     3    29     3     5   1102.90    278.72     -6.49      0.95     -0.00     69.02
19     3    29     3     6   1566.77    275.23     -6.99     -1.05      0.00     65.98
19     3    29     3     7   2054.15    274.73     -8.75     -1.05      0.00     24.00
19     3    29     3     8   2573.34    272.75     -6.99      0.44     -0.00     17.00
19     3    29     3     9   3119.08    268.58     -5.49      0.44     -0.00     24.00
19     3    29     3    10   3697.93    263.90     -5.43      0.51     -0.00     44.00
19     3    29     3    11   4311.14    259.44     -5.97      0.71     -0.00     70.00
19     3    29     3    12   4967.22    255.28     -5.68      0.76      0.00     79.00
19     3    29     3    13   5674.00    250.49     -4.80      0.15      0.00     78.00
19     3    29     3    14   6436.46    244.40     -4.44     -0.65      0.00     79.00
19     3    29     3    15   7267.74    237.81     -3.19     -1.22     -0.00     74.00
19     3    29     3    16   8183.37    230.32     -1.61     -1.88      0.00     85.00
19     3    29     3    17   9202.99    222.07     -2.68     -2.69      0.00     69.32
19     3    29     3    18  10372.61    217.25     -5.21     -2.62      0.00     28.60
19     3    29     3    19  11793.01    218.95     -3.57     -3.93     -0.00      6.20
19     3    29     3    20  13641.05    219.01     -1.64     -4.86      0.00      1.60
19     3    29     3    21  16221.63    215.78     -0.22     -4.62      0.00      1.30
19     3    29     3    22  20593.97    216.30      0.54     -2.02      0.00      0.00
19     3    29     3    23  26465.78    222.79     -7.41     -9.74      0.00      0.00
```

Figure C.1 *– An example of GDAS data after extracting the values corresponding to the location of CODALEMA. The set of data corresponding to an interval of three hours is composed of 24 lines. See text for details.*

APPENDIX D

Review of the pioneer works on low-frequency range

D.1. Introduction

This appendix aims to review the pioneer works on the low-frequency range carried out in the 70's and up to the 90's. Indeed, one of the most important works for the EXTASIS project is the study and the understanding of the previous results obtained in the past. For that, each paper dealing with the low-frequency detection of a signal was summarized: the experiment site, the name of the experiment, the geomagnetic field (obtained via https://www.ngdc.noaa.gov/geomag-web/#igrfwmm between October, 18th and 19th of 2017). Then, a description of apparatus of the experiment is done. Finally, the results of the papers are given, with sometimes some own commentaries. Each subsection corresponds to a particular paper, chronologically ordered, and which is written as a keynote. Several detections correlated with air showers at low frequencies have been realized in the 70's and up to the 90's, and their observations are summarized in the table D.1, and a conclusion can be drawn from these observations: the results are not well understood.

At that time, several experiments [176, 177, 178, 179] have shown that when frequency decreases, there is a clear evidence of a strong increase of the radio pulse amplitude. Let us try to shortly review all these pioneer works. We first make a short review of the pioneer works, and then make a complete review of each paper.

D.2. Short review of the pioneer works

Table D.1 summarizes this section. The table is composed of seven columns: date of paper, the reference of the paper, the experiment site, the instrumental setup, the frequency and the last column gathers the main results of the paper. The papers are sorted in chronological order.

Generally speaking, one can see that the results seem to be not well understood. The fact that some measurements of large electric field have not been reproduced raises doubt on the obtaining of very large electric field values. Moreover, some estimations of shower energy seem to be incredibly low to permit a radio detection. Owing to these probably erroneous shower energies, the normalization to a shower of energy equal to 10^{17} eV gives huge but probably overestimated electric field amplitudes.

Table D.1 *– Summary of low-frequency bibliographic review.*

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1969	Prescott, Hough, Pidcock, 11th ICRC 3, 717 [176]	Dominion Radio Astrophysical Observatory at Penticton (384 m asl), Canada. NS-component (+N/-S) : 17,099.3 nT, EW-component (+E/-W) : 4,774.7 nT, VE-component (+D/-U) : 51,836.1 nT	See text for a complete description of the instrumental setup	3.6 MHz, 10 MHz, 22 MHz	Calgary (1049 m asl), signal at 3.6 MHz 3 times larger than at 22 MHz, no signal larger than noise at 10 MHz. Results consistent with previous studies made by Spencer, Allan et. al.
1970	Allan, Clay, Nature 225, 253 [177]	Haverah Park (220 m asl). NS-component (+N/-S) : 18,209.4 nT, EW-component (+E/-W) : 18,204.8 nT, VE-component (+D/-U) : 45,929.5 nT	See text for a complete description of the instrumental setup	2 MHz, 408 MHz	Signal at 2 MHz 200 times larger than at 32 MHz and 375 times larger than at 44 MHz.
1971	Stubbs, Nature 230, 172 [178]	Buckland Park (13 m asl). NS-component (+N/-S) : 23,207.0 nT, EW-component (+E/-W) : 3,206.1 nT, VE-component (+D/-U) : −54,266.4 nT	See text for a complete description of the instrumental setup	2 MHz	Signal at 2 MHz 250 times larger than at 44 MHz.

Continued on next page

Table D.1 – continued from previous page

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1971	Hough, Prescott, Clay, Nature 232, 14 [179]	Dominion Radio Astrophysical Observatory at Penticton (384 m asl), Canada. NS-component (+N/-S) : 17,099.3 nT, EW-component (+E/-W) : 4,774.7 nT, VE-component (+D/-U) : 51,836.1 nT	See text for a complete description of the instrumental setup	3.6 MHz	Signal at 3.6 MHz 1 order of magnitude larger than in [20 – 60] MHz, but less by a factor of 3-4 than at 2 MHz.
1972	Allan, Nature 237, 384 [260]	None	None	5 MHz	Upper limit of µV/m/MHz , upper limit proportional to ν, becoming smaller with decreasing frequency !?
1972	Felgate, Stubbs, Nature 239, 151 [261]	Buckland Park (13 m asl). NS-component (+N/-S) : 23,207.0 nT, EW-component (+E/-W) : 3,206.1 nT, VE-component (+D/-U) : −54,266.4 nT	See text for a complete description of the instrumental setup	6 MHz	Signal observed due to a geomagnetic component polarized E-W, and a non-geomagnetic component (34% of the total field).

Continued on next page

Table D.1 – continued from previous page

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1973	Gregory, Clay, Prescott, Nature 245, 86 [262]	Buckland Park (13 m asl). NS-component (+N/-S) : 23,207.0 nT, EW-component (+E/-W) : 3,206.1 nT, VE-component (+D/-U) : −54,266.4 nT	See text for a complete description of the instrumental setup	100 kHz	The first one showing a trace of radio pulse at low-frequency in coincidence with extensive air shower, $\epsilon(\nu)$ proportional to $\nu^{-1.5}$, nature of low-frequency signal different from that at frequency above 20 MHz.
1973	G.B. Khristiansen, 13th ICRC 4, 2399A	Buckland Park (13 m asl). Yakutsk, URSS, altitude 117m asl NS-component (+N/-S) : 13,283.3 nT, EW-component (+E/-W) : −3,721.9 nT, VE-component (+D/-U) : 58,047.9 nT	See text for a complete description of the instrumental setup	100 kHz	The first one showing a trace of radio pulse at low-frequency in coincidence with extensive air shower, $\epsilon(\nu)$ proportional to $\nu^{-1.5}$, nature of low-frequency signal different from that at frequency above 20 MHz.
1975	Allan, Sun, ICRC 8, 3082	Haverah Park (220 m asl). NS-component (+N/-S) : 18,209.4 nT, EW-component (+E/-W) : 18,204.8 nT, VE-component (+D/-U) : 45,929.5 nT	See text for a complete description of the instrumental setup	1.8 MHz	Haverah Park (220 m asl), period of 8 weeks, no low-frequency signal observed in coincidence with extensive air shower, and the biggest observed shower had an energy of 1.6×10^{18} eV observed at 60 MHz.

Continued on next page

Table D.1 – continued from previous page

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1978	Sivaprasad, Aust. J. Phys. 31, 439-45 [263]	None	None	10 kHz to 10 MHz	Gather all data from previous studies =>two groups, one for which observations are in agreement with the theory, the other one with observations greater that theory. Geoelectric mechanism does not produce fields larger than the geomagnetic one produces.
1985	Suga, Kakimoto, Nishi, ICRC 7, 268 [264]	Akeno, AGASA, 900 m asl. NS-component (+N/-S) : 29,991.4 nT, EW-component (+E/-W) : −4,039.5 nT, VE-component (+D/-U) : 35,958.0 nT	See text for a complete description of the instrumental setup	50 kHz, 170 kHz, 1.647 MHz	Akeno (900 m asl), field strengh »geomagnetic and geoelectric mechanisms.
1985	Nishimura, ICRC 7, 308	None	None	10 kHz, 100 kHz and 1 MHz	Askaryan effect preponderant for low-frequency and transverse current of geomagnetic for high frequency. First mention of "sudden death" mechanism (transition radiation).

Continued on next page

Table D.1 – continued from previous page

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1987	Suga, Nishi, ICRC 6, 125 [180]	Akeno, AGASA, 900 m asl. NS-component (+N/-S) : 29,991.4 nT, EW-component (+E/-W) : −4,039.5 nT, VE-component (+D/-U) : 35,958.0 nT	See text for a complete description of the instrumental setup	26 to 300 kHz	Unipolar and negative signal, with a width of 5µs, field decreasing with decreasing distance.
1987	Datta, Pathak, ICRC 6, 129	None	None	1 to 20 MHz	High field strength observed experimentally can not be explained by theoretical ldf.
1991	Castagnoli, Ghia, Gomez, Trivero, ICRC 4, 343 [181]	EASRADIO, altitude 2000 m asl NS-component (+N/-S) : 24,185.8 nT, EW-component (+E/-W) : 1,366.8 nT, VE-component (+D/-U) : 39,880.0 nT	See text for a complete description of the instrumental setup	470 kHz, 2.6 MHz	Increase in electric field strength when frequency decreases. Observation of low-frequency signal in coincidence with extensive air shower.
1992	Baishya, Datta, Kalita, Mazumdar, Pathak, NCimC16, 17 [265]	Cosmic Ray Reseach Laboratory, Gauhati, 51.8 m asl		2 MHz, 9 MHz	Contrary to Nishimura 1985, other mechanisms (not only transition radiation), transition radiation contribution depending on field strength.

Continued on next page

Table D.1 – continued from previous page

Year	Reference	Experience site	Instrumental setup	Frequency	Main results
1993	Kadota, Suzuki, Nishi, Kakimoto, ICRC 4, 262 [266]	Akeno, AGASA, 900 m asl. NS-component (+N/-S) : 29,991.4 nT, EW-component (+E/-W) : −4,039.5 nT, VE-component (+D/-U) : 35,958.0 nT	See text for a complete description of the instrumental setup	??	~300 events with a signal at 5sigma, mainly negative unipolar signal, positive uni/bipolar signals.

D.3. Complete and extensive review of the pioneer works

Properties of radio emission of extensive air showers [176]

Experiment site

Dominion Radio Astrophysical Observatory at Penticton (384 m asl), Canada.
Declination : 15.6014°, Inclination : 71.0941°
Horizontal intensity : 17, 753.4 nT
NS-component (+N/-S) : 17, 099.3 nT, EW-component (+E/-W) : 4, 774.7 nT, VE-component (+D/-U) : 51, 836.1 nT
Total : 54, 792.1 nT

Instrumental setup (see page 2 of [176])

The main apparatus is well described. It is composed of 8 lines of 4 wide band dipole antennas, EW polarization, $\lambda/8$ above the ground, et forming the main radiotelescope (22.25 MHz) of Penticton. Symmetrically, there are 5 scintillators. A little farther to the East, there are 16 crossed-dipole antennas. The idea is to detect the signal coming from the charge excess mechanism with the latter apparatus.

For the 10 MHz contribution, the apparatus is similar: 5 lines of 4 half-wave folded dipoles, NS polarization, $\lambda/8$ above the ground. This system is triggered by Geiger counters.

For the 3.6 MHz contribution, 1 half-wave folded dipoles in EW polarization, installed on the roof of the university. Also triggered by Geiger counters.

Comments

For the signal, they obviously know where to seek the signal in the trace. So, they measure the noise variations at different positions in the trace and compare them. They are aware of the possible effect of their electronics on the data. For 10 MHz, they have 305 events, and no significant signal. However, for 3.6 MHz, and 527 events, the result is positive in the window where the signal is expected. They also know the daily variation of the ionosphere layers, and record the data during the day.

At 3.6 MHz, the signal is equal to 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ (see page 6 of [176]), and for 22 MHz the signal is equal to 0.3 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ (see page 6 of [176]). These amplitudes are determined via the calibration of the amplifier. The size of the showers is estimated and approximatively equal to 10^6 particles. They deal with large and undefined uncertainties. They compare their results with the measures carried out at 42 MHz and 32 MHz, and think that the results are consistent.

***

Frequency spectrum of air shower radio pulses [177]

Experiment site

Haverah Park (220 m asl).
Declination : -1° 17’ 40", Inclination : 68° 22’ 24"
Horizontal intensity : 17, 753.4 nT
NS-component (+N/-S) : 18, 209.4 nT, EW-component (+E/-W) : 18, 204.8 nT, VE-component (+D/-U) : 45, 929.5 nT
Total : 49, 407.5 nT

Instrumental setup (see page 1 of [177])

For 2 MHz, the antenna used is the same that the one use for the compact array, 1 m height, 4 wires of 1.5 m. Using the BBC emitter, they check that the expected voltage is equal to $E_\nu \delta_\nu / 2$. Then, they filter the BBC, and the maximum gain of the antenna is at to 1.8 MHz. The antenna is triggered by the particle detectors of the experiment when $E \geq 10^{17}$ eV. The traces are recorded on 10 µs.

Comments

The main shape of their signal is a sinusoid of 2 MHz, with an amplitude equal to 1 mV with an irregular modulation of about 20 %. They said that for 1 event, they do not see significant signal. So, they use the same method as [176]. They study three regions, and obtain 434 and 437 units (au) for aleatory regions, and 458 units for the region where the signal is expected (see page 1 of [177] colonne de droite). The error is ± 4 au. They check for non-shower event that the increase is not present. So, they obtain a significant signal, 5 sigmas above the noise. From that point, they do a calculation and assess that the signal has an amplitude of 300 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ at 2 MHz.

Now, their idea is to compare their result with the studies of the other experiments. They say that the main mechanism is the geomagnetic one, and need to take into account the angle between the arrival direction and the one of the geomagnetic field. This consideration gives a horizontal component of 750 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ at 300 m from the shower axis, to be compared with the other results. But, be careful, this is not the detected amplitude, but the calculated one. They will all do this kind of extrapolation in the following, with hypothesis not well-understood in their opinion.

***

Evidence of radio pulses at 2 MHz from cosmic ray air showers [178]

Experiment site

Buckland Park (13 m asl).
Declination : 7.8658°, Inclination : −66.6496°
Horizontal intensity : 23, 427.4 nT
NS-component (+N/-S) : 23, 207.0 nT, EW-component (+E/-W) : 3, 206.1 nT, VE-component (+D/-U) : −54, 266.4 nT
Total : 59, 107.4 nT

Instrumental setup (see page 2 of [178])

At 1.98 MHz : 178 individual half-wave dipoles over 1 km^2. The system is triggered by an Cherenkov optical system (described in the paper). The DAQ for the low-frequency antennas is also described in the paper. It is not really clear, because the trigger system can be used only during clear night, without clouds, but the low-frequency measurements have to be carried out during the day...

Comments

First, the author says that Prescott [176] has done other measurements, but without results. It was impossible to obtain the results of [177] where the amplitudes were large (see page 1 of [178]). Then, via the count rate, he shows that a signal is seen in the low-frequency band (5 sigmas above the noise), see the figure page 2 of [178] on the right column. With the optical system, he estimates the energy at 2×10^{14} eV and assess that

the transients have an average amplitude of 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$. From that, he says that the amplitude of the electric field is proportional to the energy of the primary, and by using the same consideration than the previous works, the amplitude of the signal is around 500 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ for an energy of 10^{17} eV. This result has to be compared with the previous measures, and he conclude that it is consistent with the fact that the signal is higher at low-frequency than at high-frequency.

***

Observation of radio pulses from extensive air showers at 3.6 MHz [179]

Experiment site

As for the previous paper, Dominion Radio Astrophysical Observatory at Penticton (384 m asl), Canada.
Declination : 15.6014°, Inclination : 71.0941°
Horizontal intensity : 17, 753.4 nT
NS-component (+N/-S) : 17, 099.3 nT, EW-component (+E/-W) : 4, 774.7 nT, VE-component (+D/-U) : 51, 836.1 nT
Total : 54, 792.1 nT

Instrumental setup (see page 1 of [179])

Half-wave folded dipole, centered on 3.6 MHz and installed on the roof of the university. The trigger system is a triangle of 30 m aside, composed of Geiger counters. The trace is recorded on 20 µs.

Comments

They have about 5 showers per hour of about 5×10^5 particles. They know where the signal is supposed to be in the trace, and take care of the spurious phenomenon. They record 832 showers and study 300 of them. For this type of showers, they obtain a signal of $2.5 \pm 0.6\ \mu V \cdot m^{-1} \cdot MHz^{-1}$.

Remark: We now need to pay attention to their considerations: they choose to rescale their measure to a shower of 10^{17} eV*, with an arrival direction aligned with the geomagnetic field.*

From that, they say that:
— the 750 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ of [177] becomes 900 $\mu V \cdot m^{-1} \cdot MHz^{-1}$
— the 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ of [178] becomes 1300 $\mu V \cdot m^{-1} \cdot MHz^{-1}$
— and their$2.5 \pm 0.6\ \mu V \cdot m^{-1} \cdot MHz^{-1}$ becomes 300 $\mu V \cdot m^{-1} \cdot MHz^{-1}$.

So, the large values reported in the papers are not the measured values, but the normed values. We will see in the next papers that the hypothesis used to rescale their measures are not well-mastered.

Different questions to be asked:
— they say that the normalization is not well-mastered, so, are not the reported values of electric field wrong ?
— to do the normalization, one have to take into account $\sin\theta_g$, and well-mastered the energy of the showers. I think that this is not the case for the latter point. We are talking about 10^{14} eV in [178]. Is not the radio limit well above ?

***

Low frequency radio emission from extensive air showers [260]

Experiment site

None.

Instrumental setup

None.

Comments

This is a theoretical paper aiming to define an upper limit for the electric field potentially detectable at low-frequency. Allan talk about an upper and generous limit. For this study, he uses two models: one without charge excess mechanism, the other one with 10 % of charge excess. For the first model, at 5 MHz and for a shower of 10^{17} eV, he gives an upper limit of 100 $\mu V \cdot m^{-1} \cdot MHz^{-1}$, which is a really generous limit in his point a view. Via the experiment, one will detect less signal: *"This value will not be approached in any real situation"* (see page 1 of [260]) right column, and *"The contradiction is complete"*. Comforting quotation. For the second model, he gives an upper limit of 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$. He conclude that in those two extreme cases, the value could be 100 $\mu V \cdot m^{-1} \cdot MHz^{-1}$, and that it will tend to 10 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ experimentally. He adds that these limits, calculated via his formula, depend on the frequency. From that, if the frequency decreases, the upper limit should also decrease. This result seems to be in contradiction with the fact that the amplitude of the signal is higher at low-frequency than at high frequency.

***

Polarization of extensive air shower emission at 6 MHz [261]

Experiment site

Buckland Park (13 m asl).
Declination : 7.8658°, Inclination : −66.6496°
Horizontal intensity : 23, 427.4 nT
NS-component (+N/-S) : 23, 207.0 nT, EW-component (+E/-W) : 3, 206.1 nT, VE-component (+D/-U) : −54, 266.4 nT
Total : 59, 107.4 nT

Instrumental setup (see page 1 of [261])

The instrumental setup is well described on page 1 of [261], right column. It is about 89 dipoles arranged on a square of 91 m aside, provided for 2 MHz, but allowing to work at 6 MHz. This is the same trigger as in the previous paper [178].

Comments

They have 210 events, and a repeat of [178]. For eV, they have $0.8 \pm 0.2\ \mu V \cdot m^{-1} \cdot MHz^{-1}$ for the EW polarization and $0.5 \pm 0.2\ \mu V \cdot m^{-1} \cdot MHz^{-1}$ for the NS polarization, see page 2 of [261] left column, second paragraph. Rescaling their data, they obtain 140 $\mu V \cdot m^{-1} \cdot MHz^{-1}$. But there is a large incertitude on the normalization: *"We emphasize that there is still considerable doubt as to the energy dependence of the radio field strength and in our normalization we have assumed that the dependence is essentially linear."*.

***

Low frequency radio emission from extensive air showers [262]

Experiment site

Buckland Park (13 m asl).
Declination : 7.8658°, Inclination : −66.6496°
Horizontal intensity : 23, 427.4 nT
NS-component (+N/-S) : 23, 207.0 nT, EW-component (+E/-W) : 3, 206.1 nT, VE-component (+D/-U) : −54, 266.4 nT
Total : 59, 107.4 nT

Instrumental setup (see page 1 of [262])

Antenna centered on 100 kHz, copper wire of 98 m at 3 m above the ground. Sampling of 1 µs over 256 µs. Trigger via 5 scintillators, disposed on a squared of 31 m aside, with one scintillator in the centre. The trigger is send when a coincidence is build in a window of 150 ns between the central scintillator and at least 3 others (see page 1 of [262], left column).

Comments

This is really interesting, because they write that they have around 1 shower every 500 s, corresponding to showers of 10^{15} eV, 10^{16} eV. This is how they estimate the energy. They record 10000 events. They are aware about the spurious, and notice a correlation between signal amplitude and the energy of the primary. With the antenna described above, the calculations are complicated, so they install another antenna, which is a 9 m long vertical whip antenna. With that, they record the rms of the signal, which is equal to 500 $\mu V \cdot m^{-1} \cdot MHz^{-1}$. Then, they explain that the observed pulses (shown in the paper on page 1, right column) have not always the expected shape, and that they have observed different shapes of transients. They conclude that they have seen signals with a good S/N, but that the mechanism must be different than at high frequency.

***

The investigation of radioemission from EAS with the frequencies 32 MHz, 58 MHz and 2 MHz

Experiment site

Yakutsk, Russie (enfin, URSS) où il peut faire froid, vraiment froid (record enregistré : −64.4 °C), altitude 117m asl
Declination : −15.6527°, Inclination : 76.6318°
Horizontal intensity : 13, 794.9 nT
NS-component (+N/-S) : 13, 283.3 nT, EW-component (+E/-W) : −3, 721.9 nT, VE-component (+D/-U) : 58, 047.9 nT
Total : 59, 664.5 nT

Instrumental setup (see page 2 of ...)

They present the entire array, which is well described. For the low-frequency apparatus at 2 MHz, we have to move to the page 6 of the paper. It is about two low-frequency antennas, situated at 400 m of the center of the array, in EW polarization.

Comments

After 100 hours acquisition, they have recorded 60 events. With the normalization ($\sin\theta_g = 1$), they have around 1 $\mu V \cdot m^{-1} \cdot MHz^{-1}$ for showers from 6×10^6 to 1.7×10^8 particles and at 600 m and 1300 m.

***

Experimental work on radio emission from EAS at Haverah Park

Experiment site

Haverah Park (220 m asl).
Declination : -1° 17' 40", Inclination : 68° 22' 24"
Horizontal intensity : 17, 753.4 nT
NS-component (+N/-S) : 18, 209.4 nT, EW-component (+E/-W) : 18, 204.8 nT, VE-component (+D/-U) : 45, 929.5 nT
Total : 49, 407.5 nT

Instrumental setup

The classical setup is well described in the paper of the Denver conference in 1973. In addition, they use a mutlifolded half wave dipole antenna, $\lambda/10$ above the ground, centered on 1.8 MHz.

Comments

Low-frequency data are recorded in the same time of those at 60 MHz, during the day and during 8 weeks. The biggest shower seen by the classical array has an energy estimated equal to 1.6×10^{18} eV, falls down at 200 m of the antenna, but nothing was seen in the low-frequency antenna. They said that, the biggest shower observed, after the normalization and at this frequency, gives a upper limit equal to 0.6 $\mu V \cdot m^{-1} \cdot MHz^{-1}$, which is consistent with the spectrum, but absolutely not consistent with the results of the previous experiments.

Remark: Let us notice that, until now, the indicated values (without the normalization) are in fact extrapolated values, via the control of their antenna etc. The values seems to be weak to envision a radio detection.

***

Extensive air showers and radio frequency electromagnetic fields [263]

Experiment site

None.

Instrumental setup

None.

Comments

Sivaprasad do a summary of the previous experimental data, and splits them in two categories: data which are in agreement with the theoretical upper limits, and the others which are not in agreement. He says that the experiments which have given large values in the past, are completely unable to reproduce such large values. He also says that if the measure are correct, they are not from showers, but from parasites, and namely from storms (geoelectric effect).

***

Radio signals from very large showers [264]

Experiment site

Akeno, AGASA, 900 m asl.
Declination : -7° 40' 15", Inclination : 49° 54' 58"
Horizontal intensity : 30, 262.2 nT
NS-component (+N/-S) : 29, 991.4 nT, EW-component (+E/-W) : −4, 039.5 nT, VE-component (+D/-U) : 35, 958.0 nT
Total : 46, 997.6 nT

Instrumental setup (see page 1 of [264])

In addition of the AGASA installations, the antennas are installed at 2 km of the center of the array, and the DAQ is installed in a primary school, not far away. They are two types of antennas working simultaneously. This is what we call the ball antennas, at 10 m above the ground. The frequency response of the apparatus is flat in the range [1 − 10000000] Hz. Then, they are also the vertical wire antennas, also at 10 m above the ground. The frequency response is the same that the other ones, but with a larger gain. The DAQ is well described in the paper, pages 1 and 2.

Comments

They show a signal for a shower of 10^{20} eV, of around 10^7 to 2×10^9 particles, and up to 2 km the center of the array. To explain that, they propose a mechanism: the radiation should be produced by the electrons produced by the ionisation of the shower particles in the atmosphere, ionisation due to the acceleration by the geoelectric field (see page 3 of [264]).

***

Radio wave emitted by an extensive air showers in 10 kHz to 1 MHz region

Experiment site

None.

Instrumental setup

None.

Comments

Their goal is to estimate the electric field at 10 kHz, 100 kHz and 1 MHz, but also the electric field produced at the impact of the shower on the ground: they deal with the transition radiation. They show that at low-frequency, this is the charge excess which is the main contribution of the radiation. For a shower of 10^{20} eV and at 1 km, they estimate an electric field of 100 $\mu V \cdot m^{-1} \cdot MHz^{-1}$. The transition radiation would be the main contribution below 1 MHz and above 1 km. They present two figures in support to their results (see page 4 of the paper). They deal with a shower of 10^{20} eV, but we can see on the figure that E_0 =10^{26} eV. What does E_0 stand for ?

***

VLF-LF signals from large air showers [180]

Experiment site

Akeno, AGASA, 900 m asl.
Declination : -7° 40' 15", Inclination : 49° 54' 58"
Horizontal intensity : 30, 262.2 nT
NS-component (+N/-S) : 29, 991.4 nT, EW-component (+E/-W) : −4, 039.5 nT, VE-component (+D/-U) : 35, 958.0 nT
Total : 46, 997.6 nT

Instrumental setup

This is the same as for [264]. For the low-frequency band, a vertical antenna of 10 m+LNA+FFTanalyzer at 26 kHz and 300 kHz is installed. They filter online. They also monitor the geoelectrical field. There is a scheme of the experimental line (see page 2 of [180]), left above. They do the same thing at 8 kHz and 90 kHz and 155 kHz and 186 kHz.

Comments

For the first frequency couple, they detect unipolar and negative transients. For the last frequency couple, rather, this is wave packets. They say that the signal is inversely proportional to the shower axis distance. The signal is detected at 2 km for a shower of 10^9 particles and until 2.5 km for showers of $\leq 10^9$ particles if the background is low. The larger of the signal for the first frequency couple is around 5 µs, and the dependency with the core position, R and the arrival direction is not clear for them.

Are they dealing with the sudden death signal ?

***

On the low-frequency radio-emission from extensive air showers

Experiment site

None.

Instrumental setup

None.

Comments

This is a theoretical paper. The geomagnetic effect can not explain on its own the amplitudes reported by the previous experiments. For different ldf (Capdevielle, Hillas et NKG), they study the contribution of the charge excess. For a shower of 10^{16} eV, the ldf shapes are similar, increasing when the frequency decreases. This result is in contradiction with the experimental measures, as shown on the page 2 of the paper. The observations are not explainable by the charge excess on its own.

***

On the production mechanism of MF-HF radiopulses associated with large extensive air showers

Experiment site

Cosmic Ray Reseach Laboratory, Gauhati, 51.8 m asl.
All data are on the page 2 of the paper, at the beginning of the second section.

Instrumental setup

The DAQ is sophisticated, and is sketched up on page 3. They have 5 scintillators, one folded dipole radio antenna optimized for 9 MHz. They explain how they tune the antenna to wort at 2 MHz.

Comments

The paper deals with the transition radiation. The authors show the frequency spectrum of the transition radiation and the geomagnetic mechanism (see page 5 of the paper). They compare the spectra to the data, and there is a food agreement with the transition radiation spectrum. But if we look carefully, they put only two points on the figure, one of them is on the two curves. Moreover, the points are normalized to the theoretical data obtained at 9 MHz. With the other figures, they claim that the transition radiation is probably not the only mechanism which has to be taken into account.

***

LF-MF radio signals from large air showers

Experiment site

Akeno, AGASA, 900 m asl.
Declination : -7° 40’ 15", Inclination : 49° 54’ 58"
Horizontal intensity : 30, 262.2 nT
NS-component (+N/-S) : 29, 991.4 nT, EW-component (+E/-W) : −4, 039.5 nT, VE-component (+D/-U) : 35, 958.0 nT
Total : 46, 997.6 nT

Instrumental setup

The setup is well described on page 1. The apparatus is triggered when $E \geq 10^{17}$ eV.

Comments

They have 279 events. Transients observed with the low-frequency antennas have width of 1 µs to 5 µs. They are unipolar and negative, but also positive and bipolar. By referring to a shower of 10^{18} eV, via the same hypothesis than previously, they have:

— at 250 m, the field is equal to 8 mV $\cdot$ m^{-1} $\cdot$ MHz^{-1} ;

— at 2 km, the field is equal to 100 µV $\cdot$ m^{-1} $\cdot$ MHz^{-1}.

This is not in agreement with the model, and amplitudes are 10 times larger! This raise a question: are their hypothesis to normalize their data correct ? Is the estimation of the energy correct ? Moreover, the intensity of the shower coming from East or West is larger, so they propose a new model, and it is in agreement with the experimental data. But not totally, the width of the transient is not consistent with the theory, as well as the shape of the transients.

***

Measurements of LF-MF radio pulses from EAS at Gran Sasso

Experiment site

EASRADIO, altitude 2000 m asl
Declination : 3° 14’ 4", Inclination : 58° 43’ 27"
Horizontal intensity : 24, 224.4 nT
NS-component (+N/-S) : 24, 185.8 nT, EW-component (+E/-W) : 1, 366.8 nT, VE-component (+D/-U) : 39, 880.0 nT
Total : 46, 660.9 nT

Instrumental setup

EASTOP + two low-frequency vertical antennas of 15 m height, and at 200 m and 400 m on each side of the array. They are triggered by EASTOP, but an auto-trigger mode is possible. The setup is well descrived (see page 1 and 2 of the paper)

Comments

They present a transient in the band [2.3 – 2.9] MHz, with a width of 2.5 µs. Then, for a shower of 10^8 particles, the maximum of the amplitude is 400 µV $\cdot$ m^{-1} $\cdot$ MHz^{-1} at 2.6 MHz. The signal is even more important at lower frequency. They highlight a dependency between the intensity of the signal and the distance to the shower axis.

APPENDIX E

High frequency signal exploitation & Cherenkov ring in the radio data of CODALEMA

E.1. Cherenkov effect

The Cherenkov effect (also known as Vavilov-Cherenkov effect) was discovered by Sergueï Vavilov et Pavel Cherenkov. This phenomenon is produced when an object is moving with a velocity greater than the speed of light of the dielectric medium in which it moves. Mathematically, the phenomenon is described by this equation:

$$\cos\theta = \frac{1}{\beta n} \tag{E.1}$$

with n the refractive index of the medium, β the ratio of v to the speed of light c in vacuum and θ is the angle of the emission cone. An analogy can be done with the shock-wave emitted by a plane moving with a velocity greater than the speed of the sound in the atmosphere. In our case, the plan is replaced by a particle and the sound by the light. By generalizing this analogy, we can say that a shock-wave is emitted when a particle is moving faster than the speed of the wave in a medium.
It was proposed that this phenomenon occurs during the shower developments in the atmosphere [267, 268]. In that case, this is the shower front (mainly composed of electrons and positrons) which moves with almost the speed of light, and so with a velocity greater than the velocity of the electromagnetic waves in the atmosphere, which is equal to c/n. The emission is amplified on the cone of aperture equal to θ (see equation E.1), the axis of revolution of the cone being merged with the shower axis, as shown in figure E.1.

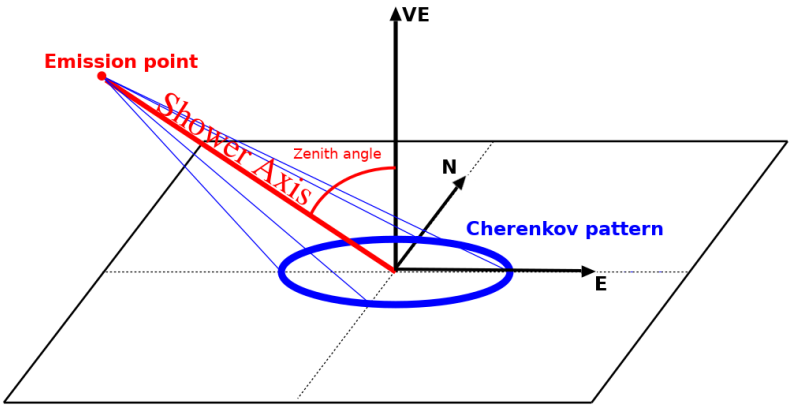


Figure E.1 – *Shower geometry for Cherenkov emission consideration. Adapted from [268].*

Thus, the signal pattern on the ground should present a ring of amplified emission, corresponding to the projection of the cone on the ground, with a diameter around 100 m

depending on the arrival direction of the shower [267]. This pattern, for frequencies above 120 MHz has been seen in the CODALEMA data, as we will see in the next following.

E.2. Observation of Cherenkov ring with CODALEMA

In this section, we present 2 examples of CODALEMA events processed in the frequency range [120 – 200] MHz. These 2 examples present a Cherenkov ring in the emission pattern. Moreover, the radio-reconstruction of these events in the classical frequency range [30 – 80] MHz did not succeed because the signal amplitude gradient is extremely weak, therefore not allowing a good and precise reconstruction of the shower core, the energy and the $X_m athr mmax$ as described in the chapter 7. The idea is to use the [120 – 200] MHz pieces of information to better constrain the comparisons, an thus to better reconstruct this type of events.
In figures E.2 are shown two events seen by CODALEMA in the classical band. The size of the circle surface is proportional to the sum of the square of the two polarization amplitudes. Colors indicate the order in which the signal is seen by the antennas (the earliests in blue, the latests in red). The amplitude gradient is homogeneous, and no particular pattern appears, which is why the comparisons go wrong.

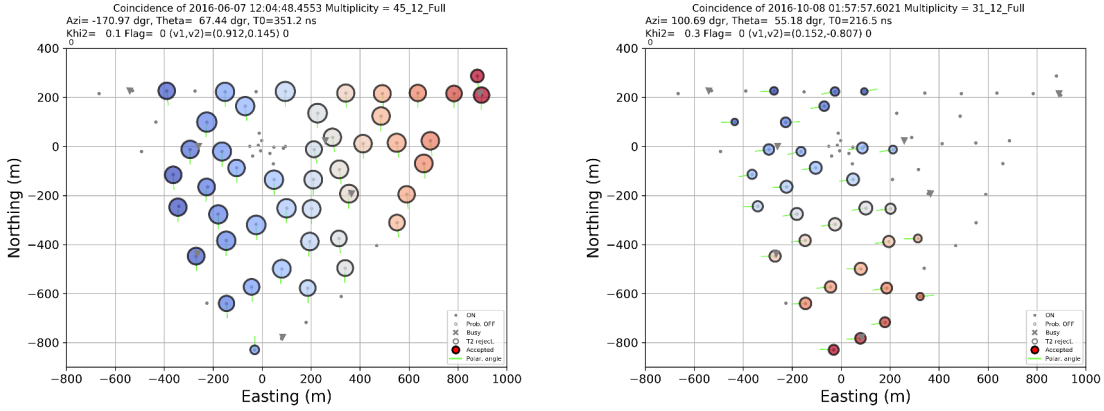


Figure E.2 *– Left: event seen by CODALEMA in the classical band* [30 – 80] MHz. *Right: another event seen by CODALEMA in the classical band.*

If we now study these events in the frequency range [120 – 200] MHz, we obtain the figures E.3. One can easily see that the amplitude gradient is completely different and non-homogeneous. Moreover, one can see that a particular pattern appears: a part of the Cherenkov ring is visible on these two events.

One has to note that the regularity of the distribution of the CODALEMA antennas allows to identify the Cherenkov ring by eye in our data. It is worth noticing the importance of the use of a wide-band antenna, such as the Butterfly used for the CODALEMA experiment, working in the [20 – 200] MHz range as shown in section 4.2.2.
As mentioned above, the idea is to take advantage of the high frequency part and especially the non-homogeneous pattern at ground to better reconstruct the shower parameters. The Cherenkov ring could, in principle, help us to better constrain the search of the core position, which is in the case of very inclined event very difficult due to the homogeneity of the pattern at ground. Let us now try to find the parameters of the Cherenkov ring, by fitting an ellipse on the ground pattern.

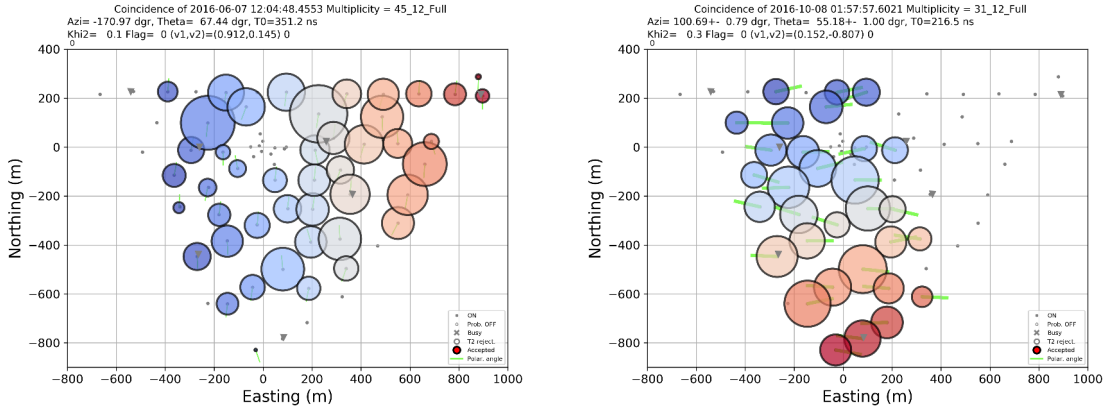


Figure E.3 *– Left: event seen by CODALEMA in* [120 – 200] MHz. *Right: another event seen by CODALEMA in* [120 – 200] MHz.

E.3. Fitting an ellipse to the radio emission pattern

E.3.1. Mathematical framework

The mathematical framework described in the following is based on [269]. Let us first define what is a conic. In affine space, a general conic a set of points verifying:

$$F(\mathbf{a},\mathbf{x}) = \mathbf{a}\cdot\mathbf{x} = ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{E.2}$$

where $\mathbf{a} = \begin{bmatrix} a & b & c & d & e & f \end{bmatrix}^T$ and $\mathbf{x} = \begin{bmatrix} x^2 & xy & y^2 & x & y & 1 \end{bmatrix}^T$. To fit this conic, one has to minimize the sum of $F(\mathbf{a},\mathbf{x})^2$:

$$\mathcal{D}(\mathbf{a}) = \sum_{i=1}^{N} F(\mathbf{x}_i)^2 \tag{E.3}$$

where $\mathcal{D}$ represent the algebraic distance to of a point to the conic and N is the number data points. For mathematical consideration, the parameter vector $\mathbf{a}$ has to be constrain, as we will see in the following. The equation E.3 is for a general conic. So, if we want to fit ellipses, we need to constrain the parameter vector $\mathbf{a}$. In fact, to be an ellipse, the constrains on $\mathbf{a} = \begin{bmatrix} a & b & c & d & e & f \end{bmatrix}^T$ are:

— a and c are non-zero and same sign ;
— the discriminant $b^2 - 4ac$ is negative.

From that, we impose the equality constraint $b^2 - 4ac = -1$ as it is done in [269], which can be rewritten as:

$$\mathbf{a}^T \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{a} = 1 \tag{E.4}$$

where the matrix will be noted $\mathbf{C}$ in the following. Now, by generalization, we note $\mathbf{D} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 & \mathbf{x}_5 & \mathbf{x}_6 \end{bmatrix}^T$ which is a $n\times 6$ matrix (also called design matrix). The equation E.3 can be rewritten as:

$$\mathcal{D}(\mathbf{a}) = \mathbf{a}^T\mathbf{S}\mathbf{a} \tag{E.5}$$

where $\mathbf{S}$ is the scatter matrix $\mathbf{D}^T\mathbf{D}$. From that, we have introduce the Lagrange multiplier λ and obtain the system of equations:

$$\begin{cases} \mathbf{Sa} = \lambda \mathbf{Ca} \\ \mathbf{a}^T\mathbf{Ca} = 1 \end{cases} \tag{E.6}$$

To obtain this system of equations, we have introduce the Lagrangian $\mathcal{L}(\mathbf{a}) = \mathcal{D}(\mathbf{a}) - \lambda(\mathbf{a}^T\mathbf{Ca} - \Phi)$ where Φ is a positive number, and we have minimized this Lagrangian ($\partial_{\mathbf{a}}\mathcal{L}(\mathbf{a})$ results in the first equation of the system E.6). Now, to solve the system, we can rewrite the first equation of the system as a general eigenvalues problem, and solve the problem to find $\mathbf{a}$ which is the eigenvector corresponding to the largest eigenvalue $1/\lambda$. For that, let us note $(\lambda_i, \mathbf{u}_i)$ a eigenvalue-eigenvector pair. If this pair solves the first equation of the system, the pair $(\lambda_i, \mu\mathbf{u}_i)$ also solve the first equation of the system for any μ. Using the second equation of the system, we find:

$$\mu_i = \sqrt{\frac{1}{\mathbf{u}_i^T\mathbf{Cu}_i}} = \sqrt{\frac{\lambda_i}{\mathbf{u}_i^T\mathbf{Su}_i}} \tag{E.7}$$

There can be up to six solutions, but the chosen solution will be the one which gives the lowest residual $\hat{\mathbf{a}}_i^T S\hat{\mathbf{a}}_i = \lambda_i$. In [269], it is shown that the constrain due to the sign of the discriminant yields to one solution for the minimization. Obviously, the solution corresponds to an ellipse, due to the constraint on the discriminant of the conic.

Now that we have made the minimization and find the corresponding $\mathbf{a}$, we can compute the different pieces of information of the fitted ellipse: center, angle of rotation, axes. The relations between $\mathbf{a}$ and these characteristics are:

— for the center of the ellipse:

$$x_0 = \frac{cd - bf}{b^2 - 4ac} \tag{E.8}$$

$$y_0 = \frac{af - bd}{b^2 - 4ac} \tag{E.9}$$

— for the semi-axes:

$$a' = \sqrt{\frac{2(af^2 + cd^2 + gb^2 - 2bdf - acg)}{(b^2 - 4ac)\left[\sqrt{(a-c)^2 + 4b^2} - (a+c)\right]}} \tag{E.10}$$

$$b' = \sqrt{\frac{2(af^2 + cd^2 + gb^2 - 2bdf - acg)}{(b^2 - 4ac)\left[-\sqrt{(a-c)^2 + 4b^2} - (a+c)\right]}} \tag{E.11}$$

— and for the angle of rotation (counterclockwise angle from the x-axis to the semi-major axis of the ellipse):

$$\Phi = \begin{cases} 0 & \text{for } b = 0 \text{ and } a < c \\ \frac{\pi}{2} & \text{for } b = 0 \text{ and } a > c \\ \frac{1}{2}\arctan^{-1}\left(\frac{a-c}{2b}\right) & \text{for } b \neq 0 \text{ and } a < c \\ \frac{\pi}{2} + \frac{1}{2}\arctan^{-1}\left(\frac{a-c}{2b}\right) & \text{for } b \neq 0 \text{and } a > c \end{cases} \tag{E.12}$$

Moreover, we know that the eccentricity of an ellipse can be expressed as $\frac{\sqrt{a^2-b^2}}{a}$ and we know the relation between the eccentricity of the ellipse, the angle of aperture of the cone and the angle between the plan of the ellipse and the axis of the cone. From that relation, we have:

$$\theta = \arcsin\left(\frac{\sqrt{a^2 - b^2}}{a} \times \cos\alpha\right) \tag{E.13}$$

with θ the angle between the axis of the cone and the plan of the ellipse, α the angle of aperture of the cone and a and b the semi-axes of the ellipse. In our case, the angle θ will correspond to the zenith angle of the arrival direction of the atmospheric shower. The fit of ellipse on our data will give us an estimation of the shower core (ellipse center), an estimation of the zenith angle θ and an estimation of the azimuthal angle Φ.

E.3.2. Experimental results

Let us deal with the second example. Fitting an ellipse on the involved antennas gives the result shown in figure E.4, where the ellipse is represented in red, and where the red cross represents the center of the fitted ellipse. By using the characteristics of the fitted

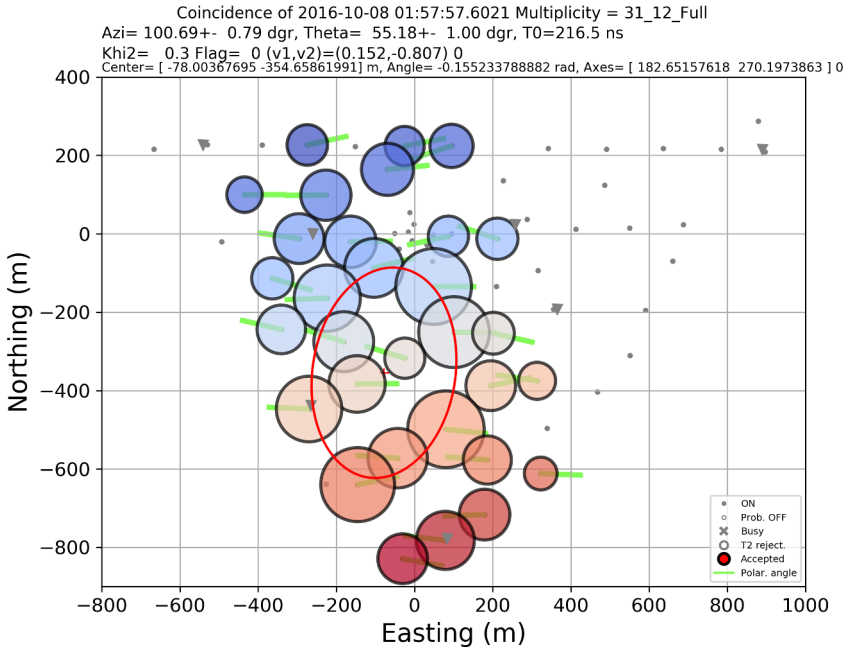


Figure E.4 *– Fit of an ellipse to the radio emission pattern. The red cross represents the center of the fitted ellipse.*

ellipse, the shower core is estimated at (x_c =−78 m,y_c =−355 m). We can also estimate the azimuth angle Φ =84° and the zenith angle θ =47°. The estimation of the arrival direction given by the standalone antennas is Φ =101° and θ =55°. Using the improved radio method, the estimation of the shower core is estimated at (x_c =−46 ± 9 m,y_c =−373 ± 20 m). The energy is estimated at $3.9 \pm 0.3 \times 10^{17}$ eV. The reduced chi-square of the reconstruction of the event using the improved method is divided by three compared to the reduced chi-square of the reconstruction of the event using the initial method. Thus, including the [120 – 200] MHz band in the reconstruction of the events permits to better constrain the comparison, since the amplitude gradient at ground becomes non-homogeneous at higher frequencies.

E.4. Discussion

The observation of a Cherenkov ring in the CODALEMA data is not systematic, and has been observed mainly for inclined shower. The development of fitting an ellipse on the ground pattern for such events was intended to better constrain the position of the shower core for the reconstruction with the radio method. Indeed, the more the

shower is inclined, the more difficult it is to constrain the position of the shower core, and thus to well reconstruct the parameters of the event. The use of the Cherenkov ring parameters should help to better reconstruct this kind of events. For lack of time, this development could not be conducted on a large batch of events, and is presented here only for informational purposes. It could be interesting to invest this development at greater length later.

APPENDIX F

Calculation of the number of expected air showers

In this appendix, we give the necessary elements for calculating the cosmic ray flux for a given experiment. The spectral indices of the cosmic ray flux measured by the Telescope Array experiment given in [228] are used.
In our case, we will focus on the high energy part of the spectrum, i.e above the knee at 3×10^{15} eV. The differential flux can be expressed as:

$$\frac{\mathrm{d}^4 N}{\mathrm{d}S\mathrm{d}t\mathrm{d}\Omega\mathrm{d}E} \backsimeq 10^{25.748} \left(\frac{E}{1\ \mathrm{eV}}\right)^{-\alpha} \mathrm{m}^{-2} \cdot \mathrm{s}^{-1} \cdot \mathrm{sr}^{-1} \cdot \mathrm{EeV}^{-1} \tag{F.1}$$

where $\mathrm{d}N$ is the number of cosmic rays differential element, $\mathrm{d}S$ is the surface differential element, $\mathrm{d}t$ is the time differential element, $\mathrm{d}\Omega$ is the solid angle differential element, $\mathrm{d}E$ is the energy differential element and $\alpha = 3.1$. Passing the energy in EeV, and moving to adapted units, the flux can now be expressed as:

$$\frac{\mathrm{d}^4 N}{\mathrm{d}S\mathrm{d}t\mathrm{d}\Omega\mathrm{d}E} \backsimeq 28.1 \left(\frac{E}{1\ \mathrm{EeV}}\right)^{-\alpha} /\mathrm{km}^2/\mathrm{yr}/\mathrm{sr}/\mathrm{EeV} \tag{F.2}$$

leading to a number of expected events given by:

$$\mathrm{d}^4 N \backsimeq 2\pi \times 28.1 \int_S \int_t \int_{E=1\ \mathrm{EeV}}^{100} \left(\frac{E}{1\ \mathrm{EeV}}\right)^{-\alpha} \int_{\theta=0}^{\theta_{\max}} \sin\theta \mathrm{d}\theta \mathrm{d}S\mathrm{d}t\mathrm{d}E \tag{F.3}$$

In case of CODALEMA/EXTASIS, we will consider an effective area (depending on θ) of $1 \cos\theta$ km^2. We will consider a data taking time of around 3 years, since the installation of the latest version of EXTASIS. In the case of events detected with CODALEMA, the duty cycle can be assimilated to $\sim$ 100 %, and in the case of events detected with EXTASIS, the day/night atmospheric noise penalty lead to a duty cycle of 50 %.

For CODALEMA, the number of expected events is given by:

$$\mathrm{d}^2 N \backsimeq 2\pi \times 28.1 \times 3 \int_S \int_t \int_{E=1\ \mathrm{EeV}}^{100} \left(\frac{E}{1\ \mathrm{EeV}}\right)^{-\alpha} \int_{\theta=0}^{\theta_{\max}} \sin\theta \mathrm{d}\theta \mathrm{d}E \tag{F.4}$$

which gives a number of expected events equal to $\sim$ 112.

For EXTASIS, the number of expected events is given by:

$$\mathrm{d}^2 N \backsimeq 2\pi \times 28.1 \times 1.5 \int_S \int_t \int_{E=1\ \mathrm{EeV}}^{100} \left(\frac{E}{1\ \mathrm{EeV}}\right)^{-\alpha} \int_{\theta=0}^{\theta_{\max}} \sin\theta \mathrm{d}\theta \mathrm{d}E \tag{F.5}$$

which gives a number of expected events equal to $\sim$ 56. Moreover, as discussed in chapter 8, a detection threshold has been derived and is equal to 23 $\mu\mathrm{V} \cdot \mathrm{m}^{-1}$. To calculate the expected detectable events in the case of the SDP, we need to take it into account, namely to take into account the dependence between the zenith angle and the energy which

corresponds to the detection threshold, and add a Heaviside function $\Theta(E)$. Thus, the number of expected events for the detection of the SDP is now given by:

$$\mathrm{d}^2N \backsimeq 2\pi \times 28.1 \times 3 \int_S \int_t \int_{E=1\ \mathrm{EeV}}^{100} \left(\frac{E}{1\ \mathrm{EeV}}\right)^{-\alpha} \int_{\theta=0}^{\theta_{\max}(E)} \Theta(E) \sin\theta \mathrm{d}\theta \mathrm{d}E \tag{F.6}$$

which gives, after a numerical integration using the linear interpolation of the contour presented in figure 8.13, a number of expected events equal to ~ 0.33.

APPENDIX G

Abbreviations and Acronyms

Table G.1 lists the abbreviations and the acronyms used in this manuscript.

Table G.1 *– List of abbreviations and acronyms.*

AERA	Auger Engineering Radio Array
AGASA	Akeno Giant Air Shower Array, in Japan
AGN	Active Galactic Nuclei
AIRES	AIR shower Extended Simulation
ANITA	Antarctic Impulsive Transient Antenna experiment
ARA	Askaryan Radio Array
ARIANNA	Antarctic Ross Ice Shelf antenna neutrino array
CA	Compact Array of CODALEMA
CDAQ	Central Data Acquisition
CMB	Cosmic Microwave Background
CODALEMA	COsmic ray Detection Array with Logarithmic ElectroMagnetic Antennas
CORSIKA	COsmic Ray SImulation for KAscade
EAS	Extensive Air Shower
EMF	Extended Medium Frequency, [20 – 200] MHz
EPOS	Energy Parton Off-shell Splitting, MC simulation code
EXTASIS	EXTinction of Air Shower Induced Signal
EW	East-West polarization
FD	Fluorescence Detector
GeV	Giga electronvolt (equal to 10^9 eV)
GRAND	Giant Radio Array for Neutrino Detection
GRB	Gamma Ray Burst
GSM	Global Sky Model
GZK	Greisen-Zatsepin-Kuzmin cutoff
LDF, ldf	Lateral Distribution Function
LF	Low Frequency
LFmap	Low Frequency Sky Map model
LFSM	Low Frequency Sky Model
LHC	Large Hadron Collider, at CERN, Geneva in Switzerland
LNA	Low Noise Amplifier
LOFAR	Low frequency array
LST	Local Sidereal Time
MF	Medium Frequency, [20 – 80] MHz
NEC	Numerical Electromagnetics Code
NKG	Nishimura-Kamata-Gresein function

Table G.2 *– List of abbreviations and acronyms (following).*

NS	North South polarization
pc	parsec, equal to 3.086×10^{16} m
PSD	Power Spectrum Density
RMS	Root Mean Square
RNO	Radio Neutrino Observatory
SA	Standalone Antennas of CODALEMA
SD	Surface Detector
SDP	Sudden Death Pulse
TA	Telescope Array, in Utah, USA
TREND	Tianshan radio experiment for neutrino detection
UHECR	Ultra High Energy Cosmic Ray
VE	Vertical polarization

References

[1] Shreyasi Acharya et al. *Measurement of the inclusive isolated photon production cross section in pp collisions at $\sqrt{s}$ = 7 TeV. Submitted to: Eur. Phys. J.*, 2019. [Quoted in pages xii and 220.]

[2] Sebastian Böser, Christian Buck, Carlo Giunti *et al. Status of Light Sterile Neutrino Searches*. 2019. [Quoted in pages xii and 220.]

[3] H. de Kerret et al. *First Double Chooz θ_{13} Measurement via Total Neutron Capture Detection*. 2019. [Quoted in pages xii and 220.]

[4] Mateusz Sitarz, Etienne Nigron, Arnaud Guertin *et al. New Cross-Sections for natMo(,x) Reactions and Medical 97Ru Production Estimations with Radionuclide Yield Calculator. Instruments*, 3(1), 2019. [Quoted in pages xii and 220.]

[5] Kunihiro Morishima et al. *Discovery of a big void in Khufu's Pyramid by observation of cosmic-ray muons. Nature*, 552(7685):386–390, 2017. [Quoted in pages xii and 220.]

[6] T. Huege. *Radio detection of cosmic ray air showers in the digital era. Physics Reports*, 620:1–52, March 2016. [Quoted in pages 1, 33, 48, 50, 53, 57, 58, and 159.]

[7] A. S. Torrentó Coello. *The AMS experiment*, pages 389–397. Springer Netherlands, Dordrecht, 2005. [Quoted in page 1.]

[8] D. J. Bird, S. C. Corbató, H. Y. Dai *et al. Results from the fly's eye experiment. AIP Conference Proceedings*, 338(1):839–854, 1995. [Quoted in page 2.]

[9] D. Charrier, R. Dallier, Antony Escudie *et al. Radio detection of cosmic rays in [1.7-3.7] MHz: the EXTASIS experiment. Astroparticle Physics*, 113:6 – 21, 2019. [Quoted in pages 3, 76, 79, 80, 94, 144, 177, and 211.]

[10] *The history of electromagnetism.* http://www.podcastscience.fm/dossiers/2013/05/16/lhistoire-de-lelectromagnetisme/. Accessed: 2016-10-15. [Quoted in pages 7, 8, 9, and 11.]

[11] *Ampère et l'histoire de l'électricité.* http://www.ampere.cnrs.fr/parcourspedagogique/zoom/coulomb/fortification/. Accessed: 2016-10-15. [Quoted in pages 7 and 11.]

[12] James Clerk Maxwell. *A Dynamical Theory of the Electromagnetic Field. Dover Publications*, 1865. [Quoted in page 13.]

[13] James Clerk Maxwell. *A Treatise on Electricity and Magnetism. Dover Publications*, 1873. Dover Publications, New York. 1873. ISBN 0-486-60636-8 (Vol. 1). [Quoted in page 13.]

[14] James Clerk Maxwell. *A Treatise on Electricity and Magnetism. Dover Publications*, 1873. Dover Publications, New York. 1873. ISBN 0-486-60637-6 (Vol. 2). [Quoted in page 13.]

[15] *1912: Victor Hess's balloon flight during total eclipse to measure cosmic rays. American Physical Society*, 19, 2010. doi:10.1007/978-3-642-28371-01. [Quoted in page 15.]

[16] John P. Wefel. *Cosmic Rays up to TeV energy: A Brief Overview*. 2016. [Quoted in page 15.]

[17] T. Stanev. *High Energy Cosmic Rays*. Springer Praxis Books. Springer Berlin Heidelberg, 2004. [Quoted in pages 15, 16, 18, 23, and 30.]

[18] Alexander Aab et al. *Highlights from the Pierre Auger Observatory. PoS*, ICRC2017:1102, 2018. [35,1102(2017)]. [Quoted in pages 17 and 48.]

[19] Kumiko Kotera and Angela V. Olinto. *The Astrophysics of Ultrahigh-Energy Cosmic Rays. Annual Review of Astronomy and Astrophysics*, 49(1):119–153, Sep 2011. [Quoted in page 17.]

[20] Matthews, James H. and Bell, Anthony R. and Araudo, Anabella T. and Blundell, Katherine M. *Cosmic ray acceleration to ultrahigh energy in radio galaxies.* In *Ultra High Energy Cosmic Rays (UHECR 2018) Paris, France, October 8-12, 2018*, 2019. [Quoted in page 18.]

[21] T. Abu-Zayyad, D. Ivanov, C. C. H. Jui *et al. The Knee and the Second Knee of the Cosmic-Ray Energy Spectrum.* 2018. [Quoted in page 18.]

[22] B. P. Abbott, R. Abbott, T. D. Abbott *et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. Physical Review Letters*, 119(16):161101, October 2017. [Quoted in page 18.]

[23] Kseniya V. Ptitsyna and Sergei V. Troitsky. *Physical conditions in potential sources of ultra-high-energy cosmic rays. I. Updated Hillas plot and radiation-loss constraints. Physics-Uspekhi*, 53:691–701, 2010. [Quoted in page 19.]

[24] Enrico Fermi. *On the Origin of the Cosmic Radiation. Physical Review*, 75:1169–1174, Apr 1949. [Quoted in page 19.]

[25] E. Parizot. *Accélération des particules : les mécanismes de Fermi.* Physique et Astrophysique du Rayonnement Cosmique, Ecole CNRS de Goutelas XXVI (2003). [Quoted in page 20.]

[26] M. Takeda, N. Hayashida, K. Honda *et al. Extension of the Cosmic-Ray Energy Spectrum beyond the Predicted Greisen-Zatsepin-Kuz'min Cutoff. Physical Review Letters*, 81:1163–1166, August 1998. [Quoted in page 22.]

[27] Dmitri Ivanov for the Pierre Auger Collaboration and the Telescope Array Collaboration. *Report of the Telescope Array - pierre Auger Observatory Working Group on Energy Spectrum.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in pages 22 and 27.]

[28] M. Tanabashi, K. Hagiwara, K. Hikasa *et al. Review of Particle Physics. Physical Review D*, 98:030001, Aug 2018. [Quoted in page 22.]

[29] G. T. Zatsepin and V. A. Kuz'min. *Upper Limit of the Spectrum of Cosmic Rays. Soviet Journal of Experimental and Theoretical Physics Letters*, 4:78, August 1966. [Quoted in page 22.]

[30] Kenneth Greisen. *End to the Cosmic-Ray Spectrum? Physical Review Letter*, 16:748–750, Apr 1966. [Quoted in page 22.]

[31] Daniele Gaggero and Mauro Valli. *Impact of cosmic-ray physics on dark matter indirect searches. Advances in High Energy Physics*, 2018:3010514, 2018. [Quoted in page 23.]

[32] Etienne Parizot. *Cosmic rays and cosmos radiation.* Habilitation à diriger des recherches, Université Paris-Diderot - Paris VII, December 2005. [Quoted in page 23.]

[33] E. Roulet. *Large-scale anisotropies above 0.03 EeV measured by the Pierre Auger Observatory.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 23.]

[34] A. di Matteo. *Full-sky searches for anisotropies in UHECR arrival directions with the Pierre Auger Observatory and the Telescope Array.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 23.]

[35] L. Caccianiga. *Anisotropies of the Highest Energy Cosmic-ray Events Recorded by the Pierre Auger Observatory in 15 years of Operation.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 23.]

[36] Alexander Aab et al. *Observation of a Large-scale Anisotropy in the Arrival Directions of Cosmic Rays above* 8×10^{18} *eV. Science*, 357(6537):1266–1270, 2017. [Quoted in pages 23, 24, and 25.]

[37] Alexander Aab et al. *The Pierre Auger Cosmic Ray Observatory. Nuclear Instruments and Methods*, A798:172–213, 2015. [Quoted in page 24.]

[38] P. Erdogdu, J. P. Huchra, O. Lahav *et al. The dipole anisotropy of the 2 Micron All-Sky Redshift Survey. Monthly Notices of the Royal Astronomical Society*, 368:1515–1526, June 2006. [Quoted in page 24.]

[39] Alexander Aab et al. *An Indication of anisotropy in arrival directions of ultra-high-energy cosmic rays through comparison to the flux pattern of extragalactic gamma-ray sources. Astrophysical Journal*, 853:L29, 2018. [Quoted in pages 24 and 25.]

[40] W. B. Atwood, A. A. Abdo, M. Ackermann *et al. The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission. Astrophysical Journal*, 697:1071–1102, June 2009. [Quoted in page 24.]

[41] H. Kawai, S. Yoshida, H. Yoshii *et al. Telescope Array Experiment. Nuclear Physics B - Proceedings Supplements*, 175-176:221 – 226, 2008. Proceedings of the XIV International Symposium on Very High Energy Cosmic Ray Interactions. [Quoted in page 25.]

[42] R. U. Abbasi et al. *Evidence of Intermediate-Scale Energy Spectrum Anisotropy of Cosmic Rays* $E \geq 10^{19.2}$ *eV with the Telescope Array Surface Detector. Astrophysical Journal*, 862(2):91, 2018. [Quoted in page 25.]

[43] Kawata, K., di Matteo, A., Fujii, T. *et al. TA Anisotropy Summary. EPJ Web Conferences*, 210:01004, 2019. [Quoted in page 26.]

[44] Alexander Aab et al. *Inferences on mass composition and tests of hadronic interactions from 0.3 to 100 EeV using the water-Cherenkov detectors of the Pierre Auger Observatory. Physical Review D*, D96(12):122003, 2017. [Quoted in pages 26 and 27.]

[45] Alexander Aab et al. *The Pierre Auger Cosmic Ray Observatory. Nuclear Instruments and Methods*, A798:172–213, 2015. [Quoted in page 27.]

[46] Francesco Fenu for the Pierre Auger Collaboration. *The cosmic ray energy spectrum measured using the Pierre Auger Observatory*. In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 27.]

[47] Hanlon, William. *Measurements of UHECR Mass Composition by Telescope Array. EPJ Web of Conferences*, 210:01008, 2019. [Quoted in page 28.]

[48] AbuZayyad, Tareq, Deligny, Olivier, Ikeda, Daisuke *et al. Auger-TA energy spectrum working group report. EPJ Web Conf.*, 210:01002, 2019. [Quoted in page 27.]

[49] O. Deligny. *The energy spectrum of ultra-high energy cosmic rays measured at the Pierre Auger Observatory and at the Telescope Array*. In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 27.]

[50] Vitor de Souza for the Pierre Auger Collaboration and the Telescope Array Collaboration. *Testing the agreement between the Xmax distributions measured by the Pierre Auger and Telescope Array Observatories*. In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 27.]

[51] Jon Paul Lundquist Thomas Stroman William Hanlon, Daisuke Ikeda and Yana Zhezher. *Telescope Array Composition Summary*. In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 27.]

[52] W. Halon. *Interpreting Auger and Telescope Array Composition Observations*. In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in pages 27 and 203.]

[53] A. Escudie, D. Charrier, R. Dallier *et al. Radio detection of atmospheric air showers of particles.* 2019. [Quoted in pages 29, 159, and 177.]

[54] A. Aab et. al. *Evidence for a mixed mass composition at the 'ankle' in the cosmic-ray spectrum. Physics Letters B*, 762:288 – 295, 2016. [Quoted in page 29.]

[55] Bruno Rossi. *Über den Ursprung der durchdringenden Korpuskularstrahlung der Atmosphäre. Zeitschrift für Physik*, 68(1):64–84, Jan 1931. [Quoted in page 30.]

[56] P. Auger, P. Ehrenfest, R. Maze *et al. Extensive Cosmic-Ray Showers. Reviews of Modern Physics*, 11:288–291, July 1939. [Quoted in page 30.]

[57] Claude Leroy and Pier-Giorgio Rancoita. *Principles of radiation interaction in matter and detection; 2nd ed.* World Scientific, Singapore, 2009. [Quoted in pages 30 and 56.]

[58] Peter K.F Grieder. *Extensive Air Showers: High Energy Phenomena and Astrophysical Aspects - A Tutorial, Reference Manual and Data Book.* Springer, Berlin, Heidelberg, 2010. [Quoted in pages 30, 34, 36, 37, and 41.]

[59] T. Bergmann, R. Engel, D. Heck *et al. One-dimensional hybrid approach to extensive air shower simulation. Astroparticle Physics*, 26(6):420 – 432, 2007. [Quoted in pages 35, 38, and 61.]

[60] W. Heitler. *The quantum theory of radiation / by W. Heitler.* Clarendon Press Oxford, 3rd ed. edition, 1954. [Quoted in page 35.]

[61] J. Matthews. *A Heitler model of extensive air showers. Astroparticle Physics*, 22(5–6):387 – 397, 2005. [Quoted in pages 35 and 37.]

[62] Diego Torres Machado. *Radio detection of ultra high energy cosmic rays : commissioning and data analysis of an array of autonomous stations.* Phd thesis, University of Nantes, October 2013. [Quoted in pages 35, 133, and 134.]

[63] Harm Schoorlemmer. *Tuning in on cosmic rays. Polarization of radio signals from air showers as a probe of emission mechanisms.* PhD thesis, Nijmegen U., 2012. [Quoted in page 37.]

[64] Alexander Aab et al. *The Pierre Auger Observatory Upgrade - Preliminary Design Report.* 2016. [Quoted in page 37.]

[65] Daniele Martello. *The Pierre Auger Observatory Upgrade. PoS*, ICRC2017:383, 2018. [,147(2017)]. [Quoted in page 37.]

[66] Antonella Castellina. *AugerPrime: the Pierre Auger Observatory Upgrade. EPJ Web Conf.*, 210:06002, 2019. [Quoted in pages 37 and 213.]

[67] T. K. Gaisser and A. M. Hillas. *Reliability of the method of constant intensity cuts for reconstructing the average development of vertical showers. International Cosmic Ray Conference*, 8:353–357, 1977. [Quoted in page 37.]

[68] M. Gabriele O. Catalano, G. D' Alí Staiti and L. La Fata. *The longitudinal EAS profile at* $E > 10^{19}$ *eV: A comparison between gil analytical formula and the predictions of detailed Monte Carlo simulations. International Cosmic Ray Conference*, pages 498–501, 2001. [Quoted in page 37.]

[69] T. Pierog and K. Werner. *EPOS Model and Ultra High Energy Cosmic Rays. Nuclear Physics B - Proceedings Supplements*, 196:102–105, 2009. [Quoted in pages 38 and 162.]

[70] Koichi Kamata and Jun Nishimura. *The Lateral and the Angular Structure Functions of Electron Showers. Progress of Theoretical Physics Supplement*, 6:93–155, 02 1958. [Quoted in pages 38 and 58.]

[71] F. Gaté. *Estimation of the composition of cosmic rays using the radio signal.* Phd thesis, Ecole des Mines de Nantes, 2016. [Quoted in pages 40, 72, 159, 160, 165, 167, 169, 170, 172, 183, and 198.]

[72] Patrick Maynard Stuart Blackett and Alfred Charles Bernard Lovell. *Radio echoes and cosmic ray showers. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 177(969):183–186, 1941. [Quoted in page 42.]

[73] K. Suga, H. Sakuyama, S. Kawaguchi *et al. Evidence for a Primary Cosmic-Ray Particle with Energy* 4×10^{21} *eV. Physical Review Letter*, 27:1604–1607, Dec 1971. [Quoted in page 42.]

[74] S. Kunwar, R. Abbasi, C. Allen *et al. Design, construction and operation of a low-power, autonomous radio-frequency data-acquisition station for the TARA experiment. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 797:110 – 120, 2015. [Quoted in page 42.]

[75] Krijn D. de Vries, Kael Hanson, and Thomas Meures. *On the feasibility of RADAR detection of high-energy neutrino-induced showers in ice. Astroparticle Physics*, 60:25–31, 2015. [Quoted in pages 42 and 215.]

[76] Krijn D. de Vries, Paul Coppin, Aongus O'Murchadha *et al. On the Radar detection of high-energy neutrino-induced cascades in ice; From Radar scattering cross-section to sensitivity*. 2018. [Quoted in pages 42 and 215.]

[77] S. Prohira et al. *Suggestion of Coherent Radio Reflections from an Electron-Beam Induced Particle Cascade*. 2018. [Quoted in pages 42 and 215.]

[78] V. Aynutdinov, A. Avrorin, I. Belolaptikov *et al. Towards high energy neutrino acoustic detector in Lake Baikal: Current status and perspectives. AIP Conference Proceedings*, 1535(1):176–179, 2013. [Quoted in page 43.]

[79] J.A. Aguilar, I. Al Samarai, A. Albert *et al. AMADEUS—The acoustic neutrino detection test system of the ANTARES deep-sea neutrino telescope. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 626-627:128 – 143, 2011. [Quoted in page 43.]

[80] Lee F. Thompson. *Acoustic detection of ultra-high-energy neutrinos. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 588(1):155 – 161, 2008. Proceedings of the First International Conference on Astroparticle Physics. [Quoted in page 43.]

[81] F. D. Kahn and I. Lerche. *Radiation from Cosmic Ray Air Showers. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 289(1417):206–213, 1966. [Quoted in pages 48, 50, and 94.]

[82] GA Askaryan. J. Exp. Theor. Phys. 21 (1962) 658. [Quoted in pages 48, 49, 52, and 94.]

[83] D. Ardouin, A. Belletoile, C. Berat *et al. Geomagnetic origin of the radio emission from cosmic ray induced air showers observed by CODALEMA. Astroparticle Physics*, 31:192–200, April 2009. [Quoted in pages 48, 51, 83, 94, 137, and 143.]

[84] A. Bellétoile, R. Dallier, A. Lecacheux *et al. Evidence for the charge-excess contribution in air shower radio emission observed by the CODALEMA experiment. Astroparticle Physics*, 69:50–60, September 2015. [Quoted in pages 48, 52, 54, 55, and 94.]

[85] The Pierre Auger Collaboration. *Probing the radio emission from cosmic-ray-induced air showers by polarization measurements. Physical Review D*, 89(052002), March 2014. [Quoted in pages 48 and 94.]

[86] W. Galbraith and J.V. Jelley. *Simulation of Radiation Energy Release in Air Showers. Nature*, 171(349), 1953. [Quoted in page 48.]

[87] Boris M Bolotovskiĭ and E M Leĭkin. *J.V. JELLEY, "CERENKOV RADIATION AND ITS APPLICATIONS". Soviet Physics Uspekhi*, 2(6):987–997, jun 1960. [Quoted in page 48.]

[88] J.V. Jelley, J.H. Fruin, N.A. Porter *et al.* *Radio pulses from extensive air showers.* *Nature*, 205:327–328, 1965. [Quoted in page 49.]

[89] R. E. Spencer. *Radio Pulses from Cosmic Ray Air Showers at 44, 105, 239 and 408 MHz.* *Nature*, 222:460–461, May 1969. [Quoted in page 49.]

[90] H.R. Allan. *Radio Emission from Extensive Air Showers.* *Progress in Elementary Particles and Cosmic Ray Physics*, 10:171–302, 1971. [Quoted in pages 49 and 159.]

[91] Robert R. Wilson. *Atmospheric Signals Caused by Cosmic-Ray Showers.* *Physical Review*, 108:155–156, Oct 1957. [Quoted in pages 54 and 55.]

[92] W. N. Charman. *Journal of Atmospheric and Solar-Terrestrial Physics*, 30(195), 1968. [Quoted in pages 54 and 55.]

[93] S. Buitink, T. Huege, H. Falcke *et al.* *Monte Carlo simulations of air showers in atmospheric electric fields.* *Astroparticle Physics*, 33:1–12, February 2010. [Quoted in page 55.]

[94] John David Jackson. *Classical electrodynamics.* Wiley, New York, NY, 3rd ed. edition, 1999. [Quoted in pages 55 and 56.]

[95] Krijn D. de Vries, Stijn Buitink, Nick van Eijndhoven *et al.* *The cosmic-ray air-shower signal in Askaryan radio detectors.* *Astroparticle Physics*, 74:96–104, 2016. [Quoted in page 56.]

[96] Kalpana Roy Sinha and Pranayee Datta. *Transition Radiation as a Tool for Identification of Primary Cosmic Rays.* *globe*, 3(4), 1965. [Quoted in page 56.]

[97] B. Revenu and V. Marin. *Radio emission from the air shower sudden death.* *ArXiv e-prints*, November 2012. [Quoted in pages 56, 58, 61, and 66.]

[98] Daniel García-Fernández, Benoît Revenu, Didier Charrier *et al.* *Calculations of low-frequency radio emission by cosmic ray-induced particle showers.* *Physical Review D*, D97(10):103010, 2018. [Quoted in pages 56, 58, 61, 66, and 96.]

[99] O. Scholten, K. Werner, and F. Rusydi. *A macroscopic description of coherent geomagnetic radiation from cosmic-ray air showers.* *Astroparticle Physics*, 29(2):94 – 103, 2008. [Quoted in page 57.]

[100] Krijn Dominique de Vries. *Macroscopic modelling of radio emission from ultra-high-energy-cosmic-ray-induced-air showers*, 2012. [Quoted in page 57.]

[101] Krijn D. de Vries, Olaf Scholten, and Klaus Werner. *The EVA code; Macroscopic modeling of radio emission from air showers based on full MC simulations including a realistic index of refraction.* *AIP Conference Proceedings*, 1535:133–137, 05 2013. [Quoted in page 57.]

[102] T. Huege, R. Ulrich, and R. Engel. *Monte Carlo simulations of geosynchrotron radio emission from CORSIKA-simulated air showers.* *Astroparticle Physics*, 27(5):392 – 405, 2007. [Quoted in page 57.]

[103] Clancy W. James, Heino Falcke, Tim Huege *et al.* *General description of electromagnetic radiation processes based on instantaneous charge acceleration in "endpoints".* *Physical Review E*, 84:056602, Nov 2011. [Quoted in page 57.]

[104] D Heck, Johannes Knapp, J-N Capdevielle *et al.* *CORSIKA: A Monte Carlo code to simulate extensive air showers.* 01 1998. [Quoted in page 58.]

[105] T. Huege, M. Ludwig, and C. W. James. *Simulating radio emission from air showers with CoREAS.* *AIP Conference Proceedings*, 1535:128, 2013. [Quoted in page 58.]

[106] Jaime Alvarez-Muñiz, Washington R. Carvalho, Enrique Zas *et al.* *Radio pulses from electromagnetic, hadronic and neutrino-induced showers up to EeV energies.* *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 662(Supplement 1):S187 – S190, 2012. 4th International workshop on Acoustic and Radio EeV Neutrino detection Activities. [Quoted in page 58.]

[107] Sergio J. Sciutto. *AIRES home page*. http://www2.fisica.unlp.edu.ar/auger/aires/, 2017. Accessed: 2017-11-17. [Quoted in pages 58 and 197.]

[108] Vincent Marin and Benoît Revenu. *Simulation of radio emission from cosmic ray air shower with SELFAS2. Astroparticle Physics*, 35(11):733 – 741, 2012. [Quoted in pages 58 and 96.]

[109] E. Zas, F. Halzen, and T. Stanev. *Electromagnetic pulses from high-energy showers: Implications for neutrino detection. Physical Review D*, 45:362–376, Jan 1992. [Quoted in page 58.]

[110] S. Lafebre, R. Engel, H. Falcke *et al. Universality of electron-positron distributions in extensive air showers. Astroparticle Physics*, 31:243–254, April 2009. [Quoted in pages 58 and 160.]

[111] T. Huege. *Theory and simulations of air shower radio emission. AIP Conference Proceedings*, 1535:121, 2013. [Quoted in page 58.]

[112] GDAS Archive Information. *Technical Report, NOAA*. https://ready.arl.noaa.gov/gdas1.php, Accessed on April 2019. [Quoted in pages 61 and 235.]

[113] F. Gaté, B. Revenu, D. García-Fernández *et al. Computing the electric field from extensive air showers using a realistic description of the atmosphere. Astroparticle Physics*, 98:38 – 51, 2018. [Quoted in pages 61, 173, and 235.]

[114] Daniel García-Fernández, Benoît Revenu, Antony Escudie *et al. Influence of a planar boundary on the electric field emitted by a particle shower. Physical Review D*, 99(6):063009, Mar 2019. [Quoted in page 61.]

[115] C.A Balanis. *Antenna Theory, Analysis and Design*. John Wiley & Sons, Inc, 3rd edition, 2005. [Quoted in pages 62 and 69.]

[116] Ben K. Sternberg and Tsylya M. Levitskaya. *Electrical parameters of soils in the frequency range from 1 kHz to 1 GHz, using lumped-circuit methods. Radio Science*, 36(4):709–719, 2001. [Quoted in page 62.]

[117] Benoit Revenu and Didier Charrier. *Deconvolution of the antenna response*, 2016. Intern article. [Quoted in page 67.]

[118] P. Abreu, M. Aglietta, M. Ahlers *et al. Antennas for the detection of radio emission pulses from cosmic-ray induced air showers at the Pierre Auger Observatory. Journal of Instrumentation*, 7:P10011, October 2012. [Quoted in pages 67 and 68.]

[119] Y. Rahmat-Samii, L. I. Williams, and R. G. Yaccarino. *The UCLA bi-polar planar-near-field antenna-measurement and diagnostics range. IEEE Antennas and Propagation Magazine*, 37(6):16–35, Dec 1995. [Quoted in page 69.]

[120] C. Burke. *Numerical Electromagnetic Codes - User's Manual*. http://physics.princeton.edu/~mcdonald/examples/NEC_Manuals/NEC4UsersMan.pdf. [Quoted in page 68.]

[121] Alexander Aab et al. *The Pierre Auger Cosmic Ray Observatory. Nuclear Instruments and Methods*, A798:172–213, 2015. [Quoted in pages 71 and 95.]

[122] P.A. Bezyazeekov, N.M. Budnev, O.A. Gress *et al. Measurement of cosmic-ray air showers with the Tunka Radio Extension (Tunka-Rex). Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 802:89 – 96, 2015. [Quoted in pages 71, 73, and 95.]

[123] Olivier Martineau-Huynh for the TREND collaboration. *Status of the TREND project. ArXiv e-prints*, April 2012. [Quoted in pages 71, 73, and 95.]

[124] S. P. Knurenko and I. S. Petrov. *Air Shower Radio Emission with Energy $E_0 \geq 10^{19}$ eV by Yakutsk Array Data. Journal of Experimental and Theoretical Physics*, 104:5, 2016. [Quoted in pages 71, 74, and 95.]

[125] M. P. van Haarlem, M. W. Wise, A. W. Gunst *et al. LOFAR: The LOw-Frequency ARray. Astronomy & Astrophysics*, 556:A2, August 2013. [Quoted in pages 71, 74, 95, and 217.]

[126] A. Nelles, S. Buitink, A. Corstanje *et al. Detecting radio emission from air showers with LOFAR.* In R. Lahmann, T. Eberl, K. Graf *et al.*, editors, *American Institute of Physics Conference Series*, volume 1535 of *American Institute of Physics Conference Series*, pages 105–110, May 2013. [Quoted in pages 71 and 74.]

[127] H. Falcke, W. D. Apel, A. F. Badea *et al. Detection and imaging of atmospheric radio flashes from cosmic ray air showers. Nature*, 435:313–316, May 2005. [Quoted in page 71.]

[128] *The CODALEMA/EXTASIS experiment: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, 2017. [Quoted in pages 71 and 76.]

[129] *National Center for Environmental Information.* https://ngdc.noaa.gov/geomag/WMM/DoDWMM.shtml. Accessed: 2017-08-28. [Quoted in page 72.]

[130] Frank G. Schröder. *Radio detection of cosmic-ray air showers and high-energy neutrinos. Progress in Particle and Nuclear Physics*, 93:1 – 68, 2017. [Quoted in pages 71 and 159.]

[131] Ewa M. Holt for the Pierre Auger Collaboration. *Recent Results of the Auger Engineering Radio Array (AERA).* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 73.]

[132] Alexander Aab et al. *Probing the radio emission from air showers with polarization measurements. Physical Review D*, D89(5):052002, 2014. [Quoted in page 72.]

[133] A. Aab, P. Abreu, M. Aglietta *et al. Energy estimation of cosmic rays with the Engineering Radio Array of the Pierre Auger Observatory. Physical Review D*, 93(12), 6 2016. [Quoted in page 73.]

[134] Qader Dorosti Hasankiadeh. *Advanced Reconstruction Strategies for the Auger Engineering Radio Array.* 2017. [Quoted in pages 73 and 159.]

[135] Tim Huege. *Radio detection of cosmic rays with the Auger Engineering Radio Array. arXiv e-prints*, page arXiv:1905.04986, May 2019. [Quoted in pages 73 and 159.]

[136] Frank G. Schröder. *Overview on the Tunka-Rex antenna array for cosmic-ray air showers (ICRC 2017). PoS*, ICRC2017:459, 2017. [Quoted in page 73.]

[137] T.R. Collaboration F. Schröder. *Overview on the Tunka-Rex antenna array for cosmic-ray air showers.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 74.]

[138] P. A. Bezyazeekov et al. *Reconstruction of cosmic ray air showers with Tunka-Rex data using template fitting of radio pulses. Phys. Rev.*, D97(12):122004, 2018. [Quoted in page 73.]

[139] D. Kostunin, P. A. Bezyazeekov, N. M. Budnev *et al. Improved measurements of the energy and shower maximum of cosmic rays with Tunka-Rex. arXiv e-prints*, page arXiv:1711.11067, Nov 2017. [Quoted in pages 73 and 159.]

[140] Tueros, M. *GRAND, a Giant Radio Array for Neutrino Detection: Objectives, design and current status. 8th workshop on Acoustic and Radio EeV Neutrino Detection Activities (ARENA 2018) Catania, Italy, June 12-15, 2018*, ARENA2018, 2018. [Quoted in pages 74, 91, and 191.]

[141] Jaime Alvarez-Muñiz et al. *The Giant Radio Array for Neutrino Detection (GRAND): Science and Design.* 2018. [Quoted in pages 74, 91, 214, 215, and 216.]

[142] D. Charrier, K.D. de Vries, Q. Gou *et al. Autonomous radio detection of air showers with the TREND50 antenna array. Astroparticle Physics*, 2019. [Quoted in page 74.]

[143] Sandra Le Coz, Didier Charrier, Krijn D. de Vries *et al. Detection of Extensive Air Showers with the self-triggered TREND radio array.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 74.]

[144] A. Sabourov, A. Glushkov, M. Pravdin *et al. Mass composition of cosmic rays with energy above* 10^17 *eV according to the data of surface detectors of Yakutsk EAS array.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 75.]

[145] S. Thoudam, S. Buitink, A. Corstanje *et al. LORA: A scintillator array for LOFAR to measure extensive air showers. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 767:339 – 346, 2014. [Quoted in page 75.]

[146] A. Escudie, D. Charrier, R. Dallier *et al. From the Observation of UHECR Radio Signal in [1-200] MHz to the Composition: CODALEMA and EXTASIS Status Report.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in pages 76, 103, 159, 177, and 212.]

[147] P. Allison, J. Auffenberg, R. Bard *et al. Design and initial performance of the Askaryan Radio Array prototype EeV neutrino detector at the South Pole. Astroparticle Physics*, 35(7):457 – 477, 2012. [Quoted in pages 76 and 214.]

[148] P. Allison, J. Auffenberg, R. Bard *et al. First constraints on the ultra-high energy neutrino flux from a prototype station of the Askaryan Radio Array. Astroparticle Physics*, 70:62 – 80, 2015. [Quoted in pages 76 and 214.]

[149] S. Barwick, D. Besson, P. Gorham *et al. South Polar in situ radio-frequency ice attenuation. Journal of Glaciology*, 51(173):231–238, 2005. [Quoted in page 76.]

[150] S.W. Barwick, E.C. Berg, D.Z. Besson *et al. A first search for cosmogenic neutrinos with the ARIANNA Hexagonal Radio Array. Astroparticle Physics*, 70:12 – 26, 2015. [Quoted in page 76.]

[151] H. Schoorlemmer, K. Belov, A. Romero-Wolf *et al. Energy and flux measurements of ultra-high energy cosmic rays observed during the first ANITA flight. Astroparticle Physics*, 77:32 – 43, 2016. [Quoted in pages 76 and 95.]

[152] Derek B. Fox, Steinn Sigurdsson, Sarah Shandera *et al. The ANITA Anomalous Events as Signatures of a Beyond Standard Model Particle, and Supporting Observations from IceCube. Submitted to: Phys. Rev. D*, 2018. [Quoted in page 76.]

[153] Ian M. Shoemaker, Alexander Kusenko, Peter Kuipers Munneke *et al. Reflections On the Anomalous ANITA Events: The Antarctic Subsurface as a Possible Explanation.* 2019. [Quoted in page 76.]

[154] P. W. Gorham, N. G. Lehtinen, G. S. Varner *et al. Observations of microwave continuum emission from air shower plasmas. Phys. Rev. D*, 78:032007, Aug 2008. [Quoted in page 76.]

[155] R. Šmída et al. *First results of the CROME experiment.* In *Proceedings, 32nd International Cosmic Ray Conference (ICRC 2011): Beijing, China, August 11-18, 2011*, volume 3, pages 17–20, 2011. [Quoted in page 76.]

[156] Alexander Aab et al. *The Pierre Auger Observatory: Contributions to the 33rd International Cosmic Ray Conference (ICRC 2013).* In *Proceedings, 33rd International Cosmic Ray Conference (ICRC2013): Rio de Janeiro, Brazil, July 2-9, 2013*, 2013. [Quoted in page 76.]

[157] R. Smida, F. Werner, R. Engel *et al. Observation of Polarised Microwave Emission from Cosmic Ray Air Showers. arXiv e-prints*, June 2013. [Quoted in page 76.]

[158] Tim Huege and Dave Besson. *Radio-wave detection of ultra-high-energy neutrinos and cosmic rays. Progress of Theoretical and Experimental Physics*, 2017(12):12A106, Dec 2017. [Quoted in page 76.]

[159] A. Escudie, D. Charrier, R. Dallier *et al. Low frequency observation of cosmic ray air shower radio emission by CODALEMA/EXTASIS.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in pages 79, 96, 177, and 211.]

[160] R. Dallier, H. Carduner, D. Charrier *et al. Direct measurement of the vertical component of the electric field from EAS.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in pages 80, 91, 92, and 211.]

[161] R. Dallier, J.L. Béney, H. Carduner *et al. Multi-scale and multi-frequency studies of cosmic ray air shower radio signals at the CODALEMA site.* In *34th ICRC, The Hague, Netherlands*, The Hague, Netherlands, July 2015. [Quoted in page 81.]

[162] B. Revenu, D. Charrier, R. Dallier *et al. The CODALEMA/EXTASIS experiment: a multi-scale and multi-wavelength instrument for radio-detection of extensive air-showers.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 81.]

[163] P. Zarka, J. N. Girard, M. Tagger *et al. LSS/NenuFAR: The LOFAR Super Station project in Nançay.* In S. Boissier, P. de Laverny, N. Nardetto *et al.*, editors, *SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pages 687–694, December 2012. [Quoted in pages 80 and 217.]

[164] T. Huege, J. Bray, S. Buitink *et al. Precision measurements of cosmic ray air showers with the SKA. Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 148, April 2015. [Quoted in page 80.]

[165] T. Huege, J. Bray, S. Buitink *et al. High-precision measurements of extensive air showers with the SKA.* In *34th International Cosmic Ray Conference (ICRC2015)*, volume 34 of *International Cosmic Ray Conference*, page 309, July 2015. [Quoted in page 80.]

[166] T. Huege, J. D. Bray, S. Buitink *et al. Ultimate precision in cosmic-ray radio detection — the SKA.* In *European Physical Journal Web of Conferences*, volume 135 of *European Physical Journal Web of Conferences*, page 02003, March 2017. [Quoted in pages 80, 213, and 214.]

[167] Thibault Garcon. *Towards autonomous radio detection of ultra high energy cosmic rays.* Theses, Université de Nantes, December 2010. [Quoted in page 80.]

[168] Colas Rivière. *From radio signals to cosmic rays.* Theses, Université Joseph-Fourier - Grenoble I, December 2009. [Quoted in pages 82 and 197.]

[169] D. Breton, E. Delagnes, and M. Houry. *Very high dynamic range and high sampling rate VME digitizing boards for physics experiments. IEEE Transactions on Nuclear Science*, 52(6):2853–2860, Dec 2005. [Quoted in pages 83 and 87.]

[170] D. Charrier. *Antenna development for astroparticle and radioastronomy experiments.* In *4th International workshop on Acoustic and Radio EeV Neutrino detection Activities (ARENA 2010)*, volume 662, Supplement 1 of *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, pages S142–S145, Nantes, France, January 2012. [Quoted in page 85.]

[171] electronic & computer engineering School of electrical. *Digital Signal Processing EEE305 & EEE801("Part A")*, 2018. [Quoted in page 91.]

[172] D. Charrier A. Coffre S. Corbel L. Denis C. Dumez-Viou A. Escudie D. Garcia-Fernandez J. Girard J.M. Grießmeier L. Koopmans A. Loh L. Martin M. Pandey-Pommier B. Revenu P. Zarka R. Dallier, L. Bondonneau. *Exploration of the Potential of the Radio-Detection Technique for the Detection of High-Energy Gamma Rays.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in pages 91 and 218.]

[173] Didier Charrier, Richard Dallier, Antony Escudie *et al.* *The CODALEMA-3 experiment for radio-detection of Ultra High Energy Cosmic Rays*, 2019. In preparation. [Quoted in pages 95, 96, 151, and 181.]

[174] S.W. Barwick, D.Z. Besson, A. Burgman *et al.* *Radio detection of air showers with the ARIANNA experiment on the Ross Ice Shelf.* *Astroparticle Physics*, 90:50 – 68, 2017. [Quoted in page 95.]

[175] R. Šmída, F. Werner, R. Engel *et al.* *First Experimental Characterization of Microwave Emission from Cosmic Ray Air Showers.* *Physical Review Letters*, 113(22):221101, November 2014. [Quoted in page 95.]

[176] J. R. Prescott, J. H. Hough, and J. K. Pidcock. *Properties of radio emission of extensive air showers.* In *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, page 717, 1970. [Quoted in pages 95, 96, 237, 238, 244, and 245.]

[177] H. R. Allan and R. W. Clay. *Frequency Spectrum of Air Shower Radio Pulses.* *Nature*, 225:253–254, January 1970. [Quoted in pages 95, 96, 237, 238, 244, 245, and 246.]

[178] T. J. Stubbs. *Evidence of Radio Pulses at 2 MHz from Cosmic Ray Air Showers.* *Nature*, 230:172–173, April 1971. [Quoted in pages 95, 96, 237, 238, 245, 246, and 247.]

[179] J. H. Hough, J. R. Prescott, and R. W. Clay. *Observation of Radio Pulses from Extensive Air Showers at 3.6 MHz.* *Nature Phys Sci*, 232:14–15, July 1971. [Quoted in pages 95, 96, 237, 239, and 246.]

[180] K. Nishi and K. Suga. *Vlf-Lf Signals from Large Air Showers.* In *International Cosmic Ray Conference*, volume 6 of *International Cosmic Ray Conference*, page 125, 1987. [Quoted in pages 96, 242, and 251.]

[181] C. Castagnoli, P.L. Ghia, F. Gomez *et al.* *Measurements of L.F. - M.F. Radio Pulses from E.A.S. at Gran Sasso (2000 m a.s.l.).* In The Institute for Advanced Studies, editor, *Proc. of the 22nd ICRC, Dublin, Ireland*, volume 4, page 363, 11-23 August 1991. [Quoted in pages 96, 189, and 242.]

[182] Silicon Laboratories. *Improving ADC resolution by oversampling and averaging.* https://www.silabs.com/documents/public/application-notes/an118.pdf, 2013. [Quoted in page 100.]

[183] H. R. Allan et al V. B. Atrashkevich, O. V. Vedeneev. *Soviet Journal of Nuclear Physics*, 28:366, 1978. [Quoted in page 103.]

[184] J. Lamblin and the CODALEMA Collaboration. *Radiodetection of astronomical phenomena in the cosmic ray dedicated CODALEMA experiment.* In *30th International Cosmic Ray Conference (ICRC), Merida, Mexico*, Merida, Mexico, July 2007. [Quoted in page 103.]

[185] J.D Kraus. *Antennas.* McGraw-Hill Book Company, 2sd edition, 1988. [Quoted in page 105.]

[186] Max Tegmark and George Efstathiou. *A method for subtracting foregrounds from multi-frequency CMB sky maps.* *Monthly Notices of the Royal Astronomical Society*, 281, 08 1995. [Quoted in page 106.]

[187] Rajib Saha, Simon Prunet, Pankaj Jain *et al.* *CMB anisotropy power spectrum using linear combinations of WMAP maps.* *Physical Review D*, 78:023003, Jul 2008. [Quoted in page 106.]

[188] Angelica de Oliveira-Costa, Max Tegmark, B. M. Gaensler *et al.* *A model of diffuse Galactic Radio Emission from 10 MHz to 100 GHz.* *Monthly Notices of the Royal Astronomical Society*, 388:247, 2008. [Quoted in page 106.]

[189] H. Zheng, M. Tegmark, J. S. Dillon *et al.* *An improved model of diffuse galactic radio emission from 10 MHz to 5 THz.* *Monthly Notices of the Royal Astronomical Society*, 464(3):3486–3497, 2017. [Quoted in page 106.]

[190] Emil Polisensky. *LFmap: A Low Frequency Sky Map Generating Program*, 2007. [Quoted in pages 106 and 108.]

[191] Jayce Dowell, Gregory B. Taylor, Frank K. Schinzel *et al. The LWA1 Low Frequency Sky Survey. Monthly Notices of the Royal Astronomical Society*, 469(4):4537–4550, 2017. [Quoted in pages 106 and 108.]

[192] C. G. T. Haslam, C. J. Salter, H. Stoffel *et al. A 408 MHz all-sky continuum survey. II - The atlas of contour maps. Astronomy and Astrophysics Supplement Series*, 47:1, January 1982. [Quoted in page 108.]

[193] Paola Platania, E. Caserini, D. Maino *et al. Full sky study of diffuse galactic emission at long wavelengths. Astron. Astrophys.*, 410:847–864, 2003. [Quoted in page 108.]

[194] K. Mulrey, A. Bonardi, S. Buitink *et al. Calibration of the LOFAR low-band antennas using the Galaxy and a model of the signal chain. Astroparticle Physics*, 111:1 – 11, 2019. [Quoted in page 109.]

[195] Y. D. Liu, H. Hu, R. Wang *et al. Plasma and Magnetic Field Characteristics of Solar Coronal Mass Ejections in Relation to Geomagnetic Storm Intensity and Variability. The Astrophysical Journal Letters*, 809:L34, August 2015. [Quoted in pages 116 and 118.]

[196] *Solar Dynamics Observatory*. https://sdo.gsfc.nasa.gov/data/rules.php. Accessed: 2019-05-29. [Quoted in page 117.]

[197] Arnaud Bellétoile. *Development and data analysis of a radiodetection of ultra high energy cosmic rays experiment*. Theses, Université de Nantes, October 2007. [Quoted in page 118.]

[198] Jennifer Maller. *Characterization of radio transient signals at the Pierre Auger Observatory*. Theses, Ecole des Mines de Nantes, October 2014. [Quoted in page 133.]

[199] Joseph O'Rourke. *Computational Geometry in C*. Cambridge University Press, New York, NY, USA, 2nd edition, 1998. [Quoted in page 134.]

[200] Paul T. Boggs and Janet E. Rogers. *Orthogonal Distance Regression*. NISTIR 89–4197. [Quoted in page 138.]

[201] T. Bohlin. *Comparison of Two Methods of Modeling Stationary EEG Signals. IBM Journal of Research and Development*, 17(3):194–205, May 1973. [Quoted in pages 144 and 145.]

[202] I. Daubechies. *The Wavelet Transform, Time-frequency Localization and Signal Analysis. IEEE Transactions on Information Theory*, 36(5):961–1005, September 2006. [Quoted in pages 144 and 146.]

[203] Enders A. Robinson. *PREDICTIVE DECOMPOSITION OF TIME SERIES WITH APPLICATION TO SEISMIC EXPLORATION. GEOPHYSICS*, 32(3):418–484, 1967. [Quoted in page 145.]

[204] J. Makhoul. *Linear prediction: A tutorial review. Proceedings of the IEEE*, 63(4):561–580, April 1975. [Quoted in page 145.]

[205] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press, 1964. [Quoted in page 145.]

[206] J Durbin. *The Fitting of Time Series Models. Review of the International Statistical Institute*, 28:233–244, 01 1960. [Quoted in page 145.]

[207] Paolo Castiglioni. *Levinson–Durbin Algorithm*, 07 2005. [Quoted in page 145.]

[208] S. Valcares. *Cosmic ray properties from the electric fields measured by the CODALEMA experiment*. Theses, Université de Nantes, September 2008. [Quoted in page 145.]

[209] B. P. Abbott. *Observation of Gravitational Waves from a Binary Black Hole Merger. Physical Review Letter*, 116:061102, Feb 2016. [Quoted in page 145.]

[210] Christopher Torrence and Gilbert P. Compo. *A Practical Guide to Wavelet Analysis. Bulletin of the American Meteorological Society*, 79:61–78, 1998. [Quoted in pages 146 and 147.]

[211] Pan Du, Warren A. Kibbe, and Simon M. Lin. *Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics*, 22(17):2059, 2006. [Quoted in page 146.]

[212] Escudie, Antony, Charrier, Didier, Dallier, Richard *et al. Multi-wavelength observation of cosmic-ray air-showers with CODALEMA/EXTASIS. EPJ Web Conf.*, 210:05003, 2019. [Quoted in pages 159 and 177.]

[213] L. Martin, R. Dallier, A. Escudie *et al. Main features of cosmic ray induced air showers measured by the CODALEMA experiment.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in pages 159, 160, 170, 172, and 183.]

[214] Washington R. Carvalho and Jaime Alvarez-Muñiz. *Determination of cosmic-ray primary mass on an event-by-event basis using radio detection. Astroparticle Physics*, 2019. [Quoted in page 159.]

[215] P. Bezyazeekov, N. Budnev, O. Fedorov *et al. Advanced Signal Reconstruction in Tunka-Rex with Matched Filtering and Deep Learning. arXiv e-prints*, page arXiv:1906.10947, Jun 2019. [Quoted in page 159.]

[216] S. Ostapchenko. *QGSJET-II: Towards reliable description of very high energy hadronic interactions. Nuclear Physics B - Proceedings Supplements*, 151:143–146, 2006. [Quoted in page 162.]

[217] Eun-Joo Ahn, Ralph Engel, Thomas K. Gaisser *et al. Cosmic ray interaction event generator SIBYLL 2.1. Physical Review D*, D80:094003, 2009. [Quoted in page 162.]

[218] A. Aab, P. Abreu, M. Aglietta *et al. Muons in air showers at the Pierre Auger Observatory: Mean number in highly inclined events. Physical Review D*, 91:032003, Feb 2015. [Quoted in page 162.]

[219] S. Buitink et al. *Method for high precision reconstruction of air shower X_{max} using two-dimensional radio intensity profiles. Physical Review D*, D90(8):082003, 2014. [Quoted in page 167.]

[220] Buitink *et al. Shower Xmax determination based on LOFAR radio measurements.* 33^{rd} ICRC, Rio de Janeiro, Brazil (2013). [Quoted in pages 170 and 198.]

[221] Olga Kambeitz. *Measurement of horizontal air showers with the Auger Engineering Radio Array. EPJ Web Conf.*, 135:01015, 2017. [Quoted in page 170.]

[222] John S. Seybold. *Introduction to RF Propagation.* John Wiley & Sons, Inc, 2005. [Quoted in page 178.]

[223] International Telecommunication Union. *Recommendation ITU-R P.372-13, Radio noise.* https://www.itu.int/rec/R-REC-P.372-13-201609-I/en, 2016. [Quoted in page 178.]

[224] H. V. Cane. *Spectra of the non-thermal radio radiation from the galactic polar regions. Monthly Notices of the Royal Astronomical Society*, 189:465–478, November 1979. [Quoted in page 179.]

[225] Frederick Mosteller and R. A. Fisher. *Questions and Answers. The American Statistician*, 2(5):30–31, 1948. [Quoted in page 188.]

[226] R.A. Fisher. *Stastitical methods for research workers.* Oliver and Boyd, 1925. [Quoted in page 188.]

[227] W.D. Apel, J.C. Arteaga, L. Bähren *et al. Thunderstorm observations by air-shower radio antenna arrays. Advances in Space Research*, 48(7):1295 – 1303, 2011. [Quoted in page 189.]

[228] V. Verzi et al. *Measurement of Energy Spectrum of Ultra-High Energy Cosmic Rays. Prog. Theor. Exp. Phys.*, 12(12A103), 2017. [Quoted in pages 190 and 261.]

[229] J. Hörandel. *A large radio array at the Pierre Auger Observatory. 8th workshop on Acoustic and Radio EeV Neutrino Detection Activities (ARENA 2018) Catania, Italy, June 12-15, 2018*, ARENA2018, 2018. [Quoted in page 193.]

[230] D. Barnhill et al. *Measurement of the lateral distribution function of UHECR air showers with the Pierre Auger Observatory.* In *Proceedings, 29th International Cosmic Ray Conference (ICRC 2005) - by Forschungszentrum Karlsruhe, Institute for Nuclear Physics, and University Karlsruhe, Institute for Experimental Nuclear Physics: Pune, India, August 3-11, 2005*, volume 7, pages 291–294, 2005. [Quoted in page 195.]

[231] Alexander Aab et al. *Reconstruction of inclined air showers detected with the Pierre Auger Observatory. JCAP*, 1408(08):019, 2014. [Quoted in page 195.]

[232] S. Buitink, A. Corstanje, H. Falcke *et al. A large light-mass component of cosmic rays at 10^{17}-$10^{17.5}$ electronvolts from radio observations. Nature*, 531:70–73, March 2016. [Quoted in pages 200, 202, and 212.]

[233] T. AbuZayyad. *TALE FD Cosmic Rays Composition Measurement.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 200.]

[234] S. Buitink. *Towards an improved mass composition analysis with LOFAR.* In *International Cosmic Ray Conference*, Madison, WI, U.S.A., July 2019. [Quoted in page 200.]

[235] O. Scholten, T. N. G. Trinh, A. Bonardi *et al. Measurement of the circular polarization in radio emission from extensive air showers confirms emission mechanisms. Physical Review D*, 94(10):103010, Nov 2016. [Quoted in page 212.]

[236] T. N. G. Trinh, O. Scholten, A. Bonardi *et al. Thunderstorm electric fields probed by extensive air showers through their polarized radio emission. Physical Review D*, 95(8):083004, Apr 2017. [Quoted in page 212.]

[237] Jörg R. Hörandel and Pierre Auger Collaboration. *Precision measurements of cosmic rays up to the highest energies with a large radio array at the Pierre Auger Observatory.* In *European Physical Journal Web of Conferences*, volume 210 of *European Physical Journal Web of Conferences*, page 06005, Oct 2019. [Quoted in pages 212 and 213.]

[238] Arthur Corstanje for the Cosmic Rays KSP. *Cosmic ray composition measurements with LOFAR.* LOFAR Science Week, May 22, 2019. [Quoted in page 212.]

[239] A. Aab, P. Abreu, M. Aglietta *et al. Observation of inclined EeV air showers with the radio detector of the Pierre Auger Observatory. Journal of Cosmology and Astro-Particle Physics*, 2018(10):026, Oct 2018. [Quoted in page 213.]

[240] Tim Huege and Christoph B. Welling. *Reconstruction of air-shower measurements with AERA in the presence of pulsed radio-frequency interference. arXiv e-prints*, page arXiv:1906.05148, Jun 2019. [Quoted in page 213.]

[241] T. Huege A. Zilles, S. Buitink. *Extreme-precision measurements of cosmic rays via radio detection with the SKA.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 213.]

[242] T. Antoni, W.D. Apel, A.F. Badea *et al. KASCADE measurements of energy spectra for elemental groups of cosmic rays: Results and open problems. Astroparticle Physics*, 24(1):1 – 25, 2005. [Quoted in page 213.]

[243] C.J. Arteaga-Velázquez, D. Rivera-Rangel, W.D. Apel *et al. Measurements of the muon content of EAS in KASCADE-Grande compared with SIBYLL 2.3 predictions.* In *International Cosmic Ray Conference*, Bexco, Busan, South Korea, July 2017. [Quoted in page 213.]

[244] V. V. Prosin, S. F. Berezhnev, N. M. Budnev *et al.* *Results from Tunka-133 (5 years observation) and from the Tunka-HiSCORE prototype.* In *European Physical Journal Web of Conferences*, volume 121, page 03004, Jul 2016. [Quoted in page 213.]

[245] R. U. Abbasi, M. Abe, T. Abu-Zayyad *et al.* *The Cosmic Ray Energy Spectrum between 2 PeV and 2 EeV Observed with the TALE Detector in Monocular Mode.* *The Astrophysical Journal*, 865(1):74, sep 2018. [Quoted in page 213.]

[246] IceCube Collaboration. *Cosmic Ray Spectrum and Composition from PeV to EeV Using 3 Years of Data From IceTop and IceCube.* *arXiv e-prints*, page arXiv:1906.04317, Jun 2019. [Quoted in page 213.]

[247] Olivier Martineau-Huynh. *The GRAND project and GRANDProto300 experiment.* In *European Physical Journal Web of Conferences*, volume 210 of *European Physical Journal Web of Conferences*, page 06007, Oct 2019. [Quoted in pages 214 and 215.]

[248] IceCube-Gen2 Collaboration, :, M. G. Aartsen *et al.* *IceCube-Gen2: A Vision for the Future of Neutrino Astronomy in Antarctica.* *arXiv e-prints*, page arXiv:1412.5106, Dec 2014. [Quoted in page 215.]

[249] Frank G. Schröder. *Physics Potential of a Radio Surface Array at the South Pole (ARENA 2018).* *arXiv e-prints*, page arXiv:1811.00599, Nov 2018. [Quoted in page 215.]

[250] Antony Escudie, Richard Dallier, Didier Charrier *et al.* *Radio detection of atmospheric air showers of particles.* *arXiv e-prints*, page arXiv:1903.02889, Mar 2019. [Quoted in pages 216, 217, and 218.]

[251] H.E.S.S. collaboration. *H.E.S.S. Publications and Conference Contributions.* https://www.mpi-hd.mpg.de/hfm/HESS/pages/publications/. [Quoted in page 216.]

[252] MAGIC collaboration. *MAGIC Publications and Conference Contributions.* https://magic.mpp.mpg.de/public/publications/. [Quoted in page 216.]

[253] VERITAS collaboration. *VERITAS Publications and Conference Contributions.* https://veritas.sao.arizona.edu/veritas-science/veritas-results-mainmenu-72. [Quoted in page 216.]

[254] Fermi collaboration. *Fermi Publications and Conference Contributions.* https://fermi.gsfc.nasa.gov/ssc/library/pubs/. [Quoted in page 216.]

[255] CTA collaboration. *CTA Publications and Conference Contributions.* https://www.cta-observatory.org/science/library/. [Quoted in page 216.]

[256] HAWC collaboration. *HAWC Publications and Conference Contributions.* https://www.hawc-observatory.org/publications/. [Quoted in page 216.]

[257] R.Dallier, C.Charrier, L.Denis *et al.* *Recent results from CODALEMA and the Nançay radio facilities related to cosmic-ray measurements.* *EPJ Web Conf.*, 135:01002, 2017. [Quoted in page 217.]

[258] M. Amenomori, Y. W. Bao, X. J. Bi *et al.* *First Detection of Photons with Energy Beyond 100 TeV from an Astrophysical Source.* *arXiv e-prints*, page arXiv:1906.05521, Jun 2019. [Quoted in page 218.]

[259] Steven.W Smith. *The Scientist and Engineer's Guide to Digital Signal Processing.* California Technical Pub, 1st edition, 1997. [Quoted in page 229.]

[260] H. R. Allan. *Low Frequency Radio Emission from Extensive Air Showers.* *Nature*, 237:384–385, June 1972. [Quoted in pages 239 and 247.]

[261] D. G. Felgate and T. J. Stubbs. *Polarization of Extensive Air Shower Emission at 6 MHz.* *Nature*, 239:151–152, September 1972. [Quoted in pages 239 and 247.]

[262] A. G. Gregory, R. W. Clay, and J. R. Prescott. *Low Frequency Radio Emission from Extensive Air Showers.* *Nature*, 245:86–87, September 1973. [Quoted in pages 240 and 248.]

[263] K. Sivaprasad. *Extensive air showers and radio frequency electromagnetic fields. Australian Journal of Physics*, 31:439–445, October 1978. [Quoted in pages 241 and 249.]

[264] K. Suga, F. Kakimoto, and K. Nishi. *Radio signals from very large showers.* In F. C. Jones, editor, *International Cosmic Ray Conference*, volume 7 of *International Cosmic Ray Conference*, pages 268–271, August 1985. [Quoted in pages 241, 250, and 251.]

[265] R. Baishya, P. Datta, P. M. Kalita *et al. On the production mechanism of MF-HF radiopulses associated with large extensive air showers. Nuovo Cimento C Geophysics Space Physics C*, 16:17–22, February 1993. [Quoted in page 242.]

[266] K. Kadota, Y. Suzuki, K. Nishi *et al. LF-MF Radio Signals from Large Air Showers.* In *International Cosmic Ray Conference*, volume 4 of *International Cosmic Ray Conference*, page 262, 1993. [Quoted in page 243.]

[267] Krijn Vries, A M. van den Berg, Olaf Scholten *et al. Coherent Cherenkov Radiation from Cosmic-Ray-Induced Air Showers. Physical review letters*, 107:061101, 08 2011. [Quoted in pages 255 and 256.]

[268] J. Alvarez-Muñiz, W. R. Carvalho, Jr., A. Romero-Wolf *et al. Coherent radiation from extensive air showers in the ultrahigh frequency band. Physical Review D*, 86(12):123007, December 2012. [Quoted in page 255.]

[269] Andrew Fitzgibbon, Maurizio Pilu, and Robert B. Fisher. *Direct Least Squares Fitting of Ellipses*, 1996. [Quoted in pages 257 and 258.]

List of contributions

The publications can be found in my dropbox: publications. The associated presentations can also be found in my dropbox: presentations.

Published articles with significant personal contribution

D. Charrier, R. Dallier, A. Escudie*, D. García-Fernández, A. Lecacheux, L. Martin, B. Revenu (2018)
Radio detection of cosmic rays in [1.7-3.7] MHz: the EXTASIS experiment
Astroparticle Physics 113:6 - 21, 2019
* *corresponding author*

Published articles as a member of the CODALEMA/EXTASIS Collaboration

D. Charrier, R. Dallier, A. Escudie, D. García-Fernández, L. Martin, B. Revenu (2018)
The CODALEMA-3 experiment for radio-detection of Ultra High Energy Cosmic Rays
To be submitted to NIMA

D. García-Fernández, B. Revenu, A. Escudie, L. Martin (2019)
Influence of a planar boundary on the electric field emitted by a particle shower
Physical Review D, 99, 063009

D. García-Fernández, B. Revenu, D. Charrier, R. Dallier, A. Escudie, L. Martin (2018)
Calculations of low-frequency radio emission by cosmic ray-induced particle showers
Physical Review D, 97, 103010

F. Gaté and B. Revenu and D. García-Fernández and V. Marin and R. Dallier and A. Escudie and L. Martin (2018)
Computing the electric field from extensive air showers using a realistic description of the atmosphere
Astroparticle Physics, 98(38): 0927-6505

Conference Proceedings

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2019)
From the observation of UHECR radio signal in [1-200] MHz to the composition: CODALEMA and EXTASIS status report

36th International Cosmic Ray Conference, July 24th-August 1st 2019, Madison, Wisconsin, USA

R. Dallier, D. Charrier, A. Coffre, L. Denis, C. Dumez-Viou, A. Escudie, D. Garcia-Fernandez, J.B. Jarnoux, L. Martin, B. Revenu, P. Zarka (2019)
Exploration of the potential of the radio-detection technique for the detection of high-energy gamma rays
36th International Cosmic Ray Conference, July 24th-August 1st 2019, Madison, Wisconsin, USA

Daniel García-Fernández, Richard Dallier, Antony Escudie, Lilian Martin, Benoît Revenu (2019)
Electric field emitted by a particle track in two semi-infinite media
36th International Cosmic Ray Conference, July 24th-August 1st 2019, Madison, Wisconsin, USA

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2019)
Radio detection of atmospheric air showers of particles
URSI-France, Approaching both infinities through electromagnetic waves, 26-27 March 2019, Observatoire de Versailles Saint-Quentin-en-Yvelines

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2018)
Multi-wavelength observation of cosmic-ray air-showers with CODALEMA/EXTASIS
5th edition of the international symposium on Ultra High Energy Cosmic Rays 2018, 8-12 October 2018, Ecole Supérieure de Chimie, Paris

B. Revenu, D. Charrier, R. Dallier, A. Escudie, D.García-Fernández, A.Lecacheux, and L. Martin (2018)
CODALEMA/EXTASIS status report: radio detection in 1-200 MHz of cosmic rays (knee/ankle)
26th Extended European Cosmic Ray Symposium and 35th Russian Cosmic Ray Conference, 6-10 July 2018, Barnaul-Belokurikha-Altai Mountains, Russia

D. García-Fernández, D. Charrier, R. Dallier, A. Escudie,A. Lecacheux, L. Martin, B. Revenu (2018)
CODALEMA/EXTASIS: Radio detection of cosmic rays at low frequencies
20th International Symposium on Very High Energy Cosmic Ray Interactions, 21-25 May 2018, Nagoya, Japan

B. Revenu, D. Charrier, R. Dallier, A. Escudie, D.García-Fernández, A.Lecacheux, and L. Martin (2017)
The CODALEMA/EXTASIS experiment: a multi-scale and multi-wavelength instrument for radio-detection of extensive air-showers
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

L. Martin, R. Dallier, A. Escudie, D. García-Fernández, F. Gaté, A.Lecacheux and B. Revenu (2017)
Main features of cosmic ray induced air showers measured by the CODALEMA experiment

35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

A. Lecacheux, D. Charrier, R. Dallier, A. Escudie, D. García-Fernández, L. Martin and B. Revenu (2017)
On timing accuracy in observing radio impulses associated with Extensive Air Showers
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

R. Dallier, H. Carduner, D. Charrier, L. Denis, A. Escudie,D. García-Fernàndez, A. Lecacheux, L. Martin, B. Revenu (2017)
Direct measurement of the vertical component of the electric field from EAS
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

D. García-Fernández, D. Charrier, R. Dallier, A. Escudie,A. Lecacheux, L. Martin, B. Revenu, M. Tueros (2017)
Near-field radio emission induced by extensive air showers
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2017)
Low frequency observations of cosmic ray air shower emission by CODALEMA/EXTASIS
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

B. Revenu, F. Gaté, V. Marin, R. Dallier, A. Escudie, D. García-Fernández, L. Martin (2017)
Computing the electric field from Extensive Air Showers using a realisticdescription of the atmosphere
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

R. Dallier, D. Charrier, L. Denis, A. Escudie, F. Gaté, A. Lecacheux, L. Martin and B. Revenu (2017)
Recent results from CODALEMA and the Nançay radio facilities related to cosmic-ray measurements
7th International Conference on Acoustic and Radio EeV Neutrino Detection Activities (ARENA 2016)

Presentations

Talks

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2019)
From the observation of UHECR radio signal in [1-200] MHz to the composition: CODALEMA and EXTASIS status report
36th International Cosmic Ray Conference, July 24th-August 1st 2019, Madison, Wisconsin, USA

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2019)
Radio detection of atmospheric air showers of particles
URSI-France, Approaching both infinities through electromagnetic waves, 26-27 March

2019, Observatoire de Versailles Saint-Quentin-en-Yvelines

A. Escudie, R. Dallier, L. Martin (2018)
Oral Presentation (popular science) on Cosmic Rays
Open Days of the Nançay Radio Observatory, 16 September 2018, Nançay Radio Observatory

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, L. Martin and B. Revenu (2018)
Multi-wavelength observation of cosmic-ray air-showers withCODALEMA/EXTASIS
2018 TeV Particle Astrophysics conference, 27-31 August 2018, Berlin, Germany

A. Escudie (2018)
Real-time observation of cosmic-ray air-showers with CODALEMA/EXTASIS
Doctoral School Day, 5 June 2018, Le Mans, France

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, L. Martin and B. Revenu (2018)
Multi-wavelength observation of cosmic-ray air-showers withCODALEMA/EXTASIS
CFRCos: Workshop of the French cosmic ray research community, 26th-28th March 2018, Paris, France

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2017)
Low frequency observations of cosmic ray air shower emission by CODALEMA/EXTASIS
35th International Cosmic Ray Conference, 12-20 July 2017, Busan, South Korea

Posters

A. Escudie, D. Charrier, R. Dallier, D. García-Fernández, A. Lecacheux, L. Martin and B. Revenu (2018)
Multi-wavelength observation of cosmic-ray air-showers with CODALEMA/EXTASIS
5th edition of the international symposium on Ultra High Energy Cosmic Rays 2018, 8-12 October 2018, Ecole Supérieure de Chimie, Paris

A. Escudie (2017)
Low frequency observations of cosmic ray air shower radio emission by CODALEMA/EXTASIS
Doctoral School Day, 30 June 2017, Angers, France

A. Escudie (2016)
Study of ultra-high energy cosmic rays at low radio frequency
ISCRA: 20th Course: "Particle, Gamma-ray and Neutrino Astrophysics in the 21st Century", 1-7 August 2016, Erice, Sicily, Italy

UNIVERSITE / MATIERE
BRETAGNE \ MOLECULES
LOIRE / ET MATERIAUX

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Titre : De l'observation du signal radio des RCUHE dans [1-200] MHz à la composition avec les expériences CODALEMA et EXTASIS

Mots clés : rayons cosmiques, gerbes atmosphériques, radio détection, mort subite, calibration, composition

Résumé :

Malgré la découverte des rayons cosmiques il y a plus de cent ans, de nombreuses questions restent aujourd'hui sans réponse : que sont les rayons cosmiques, comment sont-ils créés et d'où viennent-ils ? Depuis 2002, l'instrument CODALEMA, basé sur le site de l'Observatoire de radio-astronomie de Nançay, étudie les rayons cosmiques d'ultra haute énergie (RCUHE, au-delà de 10^{17} eV) qui arrivent dans l'atmosphère terrestre. Leur faible flux rend impossible une détection directe à ces énergies. Ces rayons cosmiques vont cependant interagir avec les atomes de l'atmosphère, engendrant une cascade de particules secondaires chargées communément appelée gerbe de particules, détectable depuis le sol, et dont on va extraire des informations sur le rayon cosmique primaire. L'objectif est de remonter aux caractéristiques du primaire ayant engendré la gerbe de particules, donc de déterminer sa direction d'arrivée, sa nature et son énergie. Lors du développement de la gerbe, les particules chargées en mouvement engendrent notamment l'émission d'une impulsion de champ électrique très brève, que CODALEMA détecte au sol avec des antennes radio dédiées, sur une large bande de fréquences (entre 1 et 200 MHz). L'avantage majeur de la radio-détection est sa sensibilité au profil complet de la gerbe et son cycle utile proche des 100 %, qui pourrait permettre d'augmenter le nombre d'événements détectés à très haute énergie, et donc de mieux contraindre les propriétés des RCUHE. Au fil des ans, des efforts importants ont été consacrés à la compréhension de l'émission radio-électrique des grandes gerbes de particules dans la gamme [20-80] MHz mais, malgré certaines études menées jusqu'aux années 90, la bande [1-10] MHz est restée inutilisée pendant près de 30 ans. L'une des contributions de cette thèse porte sur l'expérience EXTASIS, adossée à CODALEMA, qui vise à ré-investiguer cette bande et à étudier la contribution dite de "mort subite", impulsion de champ électrique créé par les particules de la gerbe lors de leur arrivée et de leur disparition au sol. Nous présentons la configuration instrumentale d'EXTASIS, composée de 7 antennes basses fréquences exploitées dans [1.7-3.7] MHz, couvrant environ 1 km2. Nous rapportons l'observation, sur 2 ans, de 25 événements détectés en coïncidence par CODALEMA et EXTASIS et estimons un seuil de détection de 23$\pm$4 µV/m à partir de comparaisons avec des simulations. Nous rapportons également une forte corrélation entre l'observation du signal basse fréquence et le champ électrique atmosphérique. L'autre contribution majeure de cette thèse porte sur l'étude du champ électrique émis par les gerbes et l'amélioration des performances du détecteur dans la bande [20-200] MHz. Nous proposons dans un premier temps une méthode de calibration des antennes de CODALEMA en utilisant l'émission radio de la Galaxie. Nous investiguons aussi plusieurs algorithmes de réjection de bruit afin d'améliorer la sélectivité des événements enregistrés. Nous présentons ensuite une méthode de reconstruction des paramètres du rayon cosmique primaire, mettant en œuvre des comparaisons combinant des informations de polarisation et fréquentielles entre les données enregistrées et des simulations, nous menant enfin à une proposition de composition en masse des rayons cosmiques détectés.

Title: From the observation of UHECR radio signal in [1-200] MHz to the composition with the CODALEMA and EXTASIS experiments

Keywords: cosmic rays, air shower, radio detection, low frequency, calibration, mass composition

Abstract:

Despite the discovery of cosmic rays there are more than one hundred years ago, many questions remain unanswered today: what are cosmic rays, how are they created and where do they come from? Since 2002, the CODALEMA instrument, located within the Nançay Radio Observatory, studies the ultra-high energy cosmic rays (UHECR, above 10^{17} eV) arriving in the Earth atmosphere. Their low flux makes it impossible to detect them directly at these energies. These cosmic rays, however, will interact with the atoms of the atmosphere, generating a cascade of secondary charged particles, commonly known as extensive air shower (EAS), detectable at ground level, and from which we will extract information on the primary cosmic ray. The objective is to go back to the characteristics of the primary that generated the EAS, thus to determine its direction of arrival, its nature and its energy. During the development of the shower, these charged particles in movement generate a fast electric field transient, detected at ground by CODALEMA with dedicated radio antennas over a wide frequency band (between 1 and 200 MHz). The major advantage of radio-detection is its sensibility to the whole profile of the shower and its duty cycle close to 100 %, which could increase the number of events detected at very high energy, and thus to better constrain the properties of the RCUHE. Over the years, significant efforts have been devoted to the understanding of the radio emission of extensive air shower (EAS) in the range [20-80] MHz but, despite some studies led until the nineties, the [1-10] MHz band has remained unused for nearly 30 years. One of the contributions of this thesis concerns the EXTASIS experiment, supported by the CODALEMA instrument, which aims to reinvestigate the [1-10] MHz band and to study the so-called "sudden death" contribution, which is the expected impulsive electric field created by the particles at their arrival and their disappearance on the ground. We present the instrumental setup of EXTASIS, composed of 7 low frequency antennas exploited in [1.7-3.7] MHz, covering approximately 1 km^2. We report the observation, over 2 years, of 25 low-frequency events detected in coincidence by CODALEMA and EXTASIS and estimate a detection limit of 23$\pm$4 µV/m from comparisons with simulations. We also report a strong correlation between the observation of the low frequency signal and the atmospheric electric field. The other major contribution of this thesis concerns the study of the electric field emitted by the EAS and the improvement of the detector's performances in the [20-200] MHz band. First, we propose a calibration method for CODALEMA antennas using the radio emission of the Galaxy. We are also investigating several noise rejection algorithms to improve the selectivity of recorded events. We then present a method for reconstructing the parameters of the primary cosmic ray, implementing systematic comparisons combing polarization and frequency information between the recorded data and simulations, leading finally to a proposal for a mass composition of cosmic rays detected.