

Map making from transit interferometers observations for 21cm Intensity Mapping experiments: Application to Tianlai and PAON-4

Jiao Zhang

► To cite this version:

Jiao Zhang. Map making from transit interferometers observations for 21cm Intensity Mapping experiments: Application to Tianlai and PAON-4. Cosmology and Extra-Galactic Astrophysics [astro-ph.CO]. Université Paris-Saclay; University of Chinese academy of sciences, 2017. English. NNT: 2017SACLS158. tel-01587810

HAL Id: tel-01587810 https://theses.hal.science/tel-01587810

Submitted on 14 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.







NNT: 2017SACLS158

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SACLAY ET UNIVERSITY OF CHINESE ACADEMY OF SCIENCE PRÉPARÉE À L'UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE Nº 576 Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)

Spécialité de doctorat : Astroparticules et Cosmologie

par

Mme Jiao ZHANG

Map making from transit interferometers observations for 21cm Intensity Mapping experiments Application to Tianlai and PAON-4

Reconstruction de cartes à partir des observations d'interféromètres radio en mode transit pour les expériences de cartographie d'intensité à 21 cm Application à Tianlai et PAON-4

Thèse présentée et soutenue à Orsay le 26 juin 2017

devant le jury composé de

- M. Achille Stocchi
- M. Martin Bucher
- M. Tongjie Zhang
- M. Steve Torchinsky
- M. Réza Ansari
- M. Xuelei Chen

Président du jury

Directeur de thèse

Co-directeur de thèse

Rapporteur

Rapporteur

Examinateur

Abstract

The analysis of the statistical properties of the distribution of matter in the cosmos (LSS or Large Scale Structure) is one of the main cosmological probes that allow the study of the cosmological standard model, in particular the parameters characterizing dark matter and dark energy. Baryonic Acoustic Oscillations (BAO's) are one of the measurements that can be extracted from the study of matter distribution in large-scale structure (LSS).

The observation of the cosmic distribution of the matter from neutral atomic hydrogen (HI) 21 cm emission is a new method, complementary to the optical observation to map the distribution of matter in the cosmos.

In the last decade, the Intensity Mapping method has been proposed as an effective method for mapping the 21cm radio emission in three dimensions. In particular, it does not require the detection of individual objects (galaxies), and can therefore be performed with instruments smaller in size than those such as SKA or FAST, which are designed to detect 21 cm galaxies at cosmological distances. A radio interferometer using a set of fixed cylindrical or parabolic reflectors observing the sky in transit modes are suitable instruments for intensity mapping surveys.

The specific observational mode from this type of radio telescope by intensity mapping is studied in the context of this thesis. We show in particular that a specific sky maps reconstruction method from the visibilities can be applied to the observations of these interferometers operating in transit mode. This method corresponds to the m-modes decomposition of the spherical harmonics and is very efficient for the reconstruction of large sky areas observed in transit mode.

A reconstruction code based on this principle has been developed, as well as different criteria for the comparison of instrumental performances, such as the synthesized antenna lobe, the noise spectrum of the reconstructed maps and the overall instrument response in the the spherical harmonics (l,m) plane.

The method has then been applied to different configurations of interferometers composed of parabolic or cylindrical reflectors in the PAON-4 and Tianlai projects. In addition to optimizing the Tianlai and PAON-4 interferometer configurations, the work presented here includes a first application of the method to the PAON-4 data.

Résumé

L'analyse des propriétés statistiques de la distribution de la matière dans le cosmos (Grandes Structures, LSS or Large Scale Structure) est l'une des principales sondes cosmologiques qui permettent l'étude du modèle standard cosmologique, en particulier les paramètres caractérisant la matière noire et l'énergie noire. Les Oscillations Acoustiques Baryoniques (BAO's) sont l'une des mesures qui peuvent être extraites de l'étude de la distribution de matière à grande échelle (LSS).

L'observation de la distribution cosmique de la matière à partir de l'émission à 21 cm de l'hydrogène atomique neutre (HI) est une nouvelle méthode, complémentaire des relevés optiques pour cartographier la distribution de la matière dans le cosmos.

La méthode de cartographie d'intensité (Intensity Mapping) a été proposée depuis moins d'une dizaine d'années comme une méthode efficace pour cartographier en trois dimensions l'émission radio à 21 cm. Elle n'implique en particulier pas la détection des objets individuels (galaxies), et des relevés de ce type peuvent donc être effectués avec des instruments plus modestes en taille que ceux comme SKA ou FAST qui sont conçus pour détecter les galaxies à 21 cm à des distances cosmologiques.

Des interféromètres radio utilisant un ensemble de réflecteurs cylindriques ou paraboliques fixes, observant le ciel en mode transit sont adaptés à la cartographie d'intensité.

Le mode d'observation spécifique de ce type de radio télescope en cartographie d'intensité est étudié dans le cadre de ce travail de thèse. On montre en particulier qu'une méthode spécifique de reconstruction des cartes du ciel à partir des visibilités peut-être appliquée aux observations de ces interféromètres fonctionnant en mode transit. Cette méthode correspond à la décomposition en modes m des harmoniques sphériques et est très performante pour la reconstruction de grandes zones du ciel observées en mode transit.

Un code de reconstruction fondé sur ce principe a été développé, ainsi que différents critères de comparaison des performances instrumentales, comme le lobe d'antenne synthétisé, le spectre de bruit sur les cartes reconstruites et la réponse globale de l'instrument dans le plan (l,m) des harmoniques sphériques.

La méthode a été appliquée à différentes configurations des interféromètres composés de réflecteurs paraboliques ou cylindriques dans le cadre des projets PAON-4 et Tianlai. Outre l'optimisation des configurations des interféromètres Tianlai et PAON-4, le travail présenté inclut une première application de la méthode aux données PAON-4.

Présentation synthétique du manuscrit

Ce manuscrit de thèse rédigé en anglais comprend cinq chapitre, complété par trois annexes.

Le premier chapitre présente brièvement le modèle standard de la cosmologie, le métrique de Friedmann-Lemaître-Robertson-Walker, l'histoire thermique de l'univers et la formation des grandes structures. Certaines des principales sondes cosmologiques permettant l'étude de l'énergie noire sont ensuite discutées: les supernovae de type Ia, les amas de galaxies, les distorsions gravitationnelles et les oscillations acoustiques baryoniques (BAO's).

La cosmologie à 21 cm et les observations de l'émission de l'hydrogène atomique neutre (HI) sont abordé dans le second chapitre. Ce chapitre débute par une courte présentation de quelques uns des grands radiotélescopes et interféromètres radio. L'objet de l'étude présentée dans cette thèse étant la cartographie de la distribution du gaz HI, l'évolution de l'émission à 21 cm à travers les âges cosmiques est discutée, ainsi que les relevés radio ayant comme objectif la détection de l'émission à 21 cm extragalactique. Le second chapitre se termine par une présentation synthétique des instruments dédiées aux observations de l'époque de réionisation de l'univers, comme LOFAR ou MWA, des projets et instruments dédiés à la cartographie d'intensité à 21 cm pour des relevés cosmologique pour l'étude de l'énergie noire. Les instruments PAON-4 et Tianlai sont présentés dans ce cadre.

Les chapitres 3 et 4 constituent le coeur du travail effectué. Celui-ci porte en effet sur les problèmes spécifiques des interféromètres radio adaptés aux relevés HI pour la cosmologie. Ceuxci doivent avoir un grand champ de vue instantané et capable d'observer les émissions radio sur une grande largeur de bande en fréquence. C'est en effet la largeur de bande qui détermine l'intervalle en décalage vers le rouge ou redshift (z) observé ($\nu = \nu_0/(1+z)$). Les réseaux interféromériques denses, observant en mode transit sont les instruments envisagés pour les relevés cosmologiques à 21 cm utilisant la cartographie d'intensité. Le fonctionnement de ces instruments en mode transit, contrairement aux observations interféromériques classiques où une source est suivie dans le ciel, ainsi que les grandes surfaces du ciel observés imposent de repenser les méthodes de reconstruction du ciel à partir des visibilités.

Dans un interféromètre radio, les ondes électromagnétiques captées par les antennes sont converties en signal électrique, représentant les variations en amplitude et phase de l'onde. Les visibilités V_{ij} correspondent à la corrélation des signaux entre deux antennes i, j, pour un signal quasi monochromatique $V_{ij} = \langle s_i^* s_j \rangle$. Il est en général plus facile de représenter le lien entre les visibilités et les émissions du ciel dans l'espace de Fourier des fréquences angulaires, notées (u, v). Compte tenu des spécificités des interféromètres fonctionnant en mode transit et observant une grande surface du ciel, il est plus pertinent de travailler en coordonnées sphériques (θ, φ) et de se placer dans l'espace des harmoniques sphériques $Y_{\ell m}(\theta, \varphi)$. Un formalisme adapté à l'analyse des visibilités et une méthode optimale de reconstruction des cartes utilisant les harmoniques sphériques ont été développés et sont présentés dans le chapitre 3. On montre en particulier qu'on peut tirer parti des observations en mode transit, couvrant la totalité du domaine angulaire selon la direction de l'angle horaire (α) pour décomposer le problème global de reconstruction de cartes, en un ensemble de problèmes de dimension réduite en terme de taille de système linéaire (nombre d'inconnus), en projetant sur les composantes en m des harmoniques sphériques.

Un logiciel de simulation et de reconstruction de cartes a été développé dans une approche orientée objet, dans le langage C++. Le code a été optimisé et peut tirer parti des architectures multi-coeurs pour calculer en parallèle différents modes m. Un certain nombre de critères permettant de quantifier la qualité de reconstruction et les performances d'un réseau d'antennes ont été définis dont le calcul a été incorporé dans le code. Mentionnons à titre d'exemple le calcul de la fonction de réponse de l'instrument (les matrices \mathbf{R}_m) et la fonction de transfert $T(\ell)$, ainsi que le spectre de bruit $C^{\text{noise}}(\ell)$. Rappelons que ce spectre de bruit est une quantité fondamentale pour quantifier correctement les observations en cartographie d'intensité.

Le code et les outils développés dont les fondements sont exposés dans le chapitre 3 ont été utilisés pour l'étude et l'optimisation de la configuration des instruments PAON-4 et Tianlai. La présentation des résultats obtenus fait l'objet du chapitre 4.

PAON-4 est un interféromètre de test comprenant 4 antennes de 5 mètres de diamètre, utilisant la chaîne électronique et d'acquisition BAORadio et déployé à l'observatoire radioastronomique de Nan?ay. Il a été conçu et construit par le l'observatoire de Paris, LAL (CNRS/IN2P3) et l'Irfu (CEA), avec comme objectif de tester les méthodes de calibration et de reconstruction de cartes pour la cartographie d'intensité avec les interféromètres en mode transit.

Le projet Tianlai dans sa phase actuelle comporte deux instruments, un réseau de 16 antennes paraboliques de 6 mètres de diamètre chacune, ainsi qu'un ensemble de trois réflecteurs cylindriques de 40 mètres de long et 15 mètres de largeur. L'ensemble comporte actuellement 96 récepteurs à double polarisation, réparties sur les 3 cylindres.

L'optimisation de la configuration du réseau des 16 antennes a fait l'objet de travaux présentés dans le chapitre 4. On peut noter en particulier l'intérêt d'un relevé autour du pôle céleste, où le réseau des 16 antennes de Tianlai serait suffisamment sensible pour détecter les structures HI proches ($z \leq 0.2$).

Une étude a également été menée pour optimiser le placement et la répartition des récepteurs le long de l'axe des cylindres. On a montré en effet que la répartition irrégulière des récepteurs permettait d'atténuer fortement les lobes secondaires correspondant aux ordres supérieurs de diffraction des réseaux réguliers. La répartition et le placement des récepteurs sur les cylindres Tianlai est basée sur l'étude menée et présentée dans le chapitre 4.

Le chapitre 5 est dédié à une courte présentation de quelques observations effectuées avec l'instrument PAON-4. Le travail de nettoyage des données et la calibration en gain et en phase de l'instrument a été mené par Qizhi Huang qui prépare également sa thèse de doctorat dans le cadre de Tianlai et PAON-4. Une première évaluation partielle des performances de l'instrument est présentée, ainsi qu'une comparaison des signaux de visibilités observées avec celles prédites à partir des cartes connues du ciel (relevé LAB à 21 cm, cartes synchrotron Haslam à 408 MHz, catalogues de sources radio ...). Le très bon accord entre les observations et les simulations permet de valider la méthode et le code de simulation et de reconstruction de cartes. Des cartes reconstruites à partir des observations PAON-4 de novembre 2016 pour deux fréquences sont également présentés, ainsi qu'un méthode plus simple et rapide de reconstruction de cartes (QuickMap).

L'annexe A présente une étude qui a été menée au début du travail de thèse et qui analyse les performances d'un réseau d'antenne et la reconstruction de cartes dans un cadre simplifié à une

dimension. L'annexe B présente brièvement la structure du code de reconstruction, et l'annexe C regroupe la première page des deux articles qui ont été publiés sur la base des travaux des chapitres 3 et 4.

Acknowledgements

I would like to express my sincere gratitude to my two advisors: Prof. Reza Ansari and Prof. Xuelei Chen, for their continuous support of my Ph.D study and related research. Prof. Chen introduced me to the amazing field of cosmology, and provided me many opportunities to do cutting edge research and to receive international education, and give me many encouragements and suggestions. Prof. Ansari invited me to visit France, and he carefully and patiently guided me through my study, research and the writing of this thesis. Both of my advisors are very generous and dedicated to science, I learned greatly from them.

Besides the two advisors, I benefited from interacting with many faculty and staff members in both France and China. In France, Dr. Jean-Eric Campagne and Dr. Christophe Magneville helped me a lot during my research. And Dr. Ana Sofia Torrento also helped me, both in practical aspects of life in France, as well as in my scientific work. I also wish to thank Dr. Jean-Michel Martin and Dr. Pierre Colom who have guided me through radioastronomy subtleties during my visits to Nançay. I extend my thanks to Dr. Marc Moniez and to the engineering team in charge of PAON-4, Daniel Charlet and Monique Taurigna among others, with special thanks to Claude Pailler , who has been in charge of data acquisition and observations with PAON-4.

In China, talking with Prof. Youjun Lu always delights me and helps me to understand many problems. Dr. Yougang Wang and Dr. Fengquan Wu helped me in many aspects of my research, and explained to me various aspects of the Tianlai experiment. I also wish to thank the secretaires, Ms. Tiantian Li and Ms. Jingyu Zhang, who are my good friends and helped me in my international study.

During my study I learned much from my fellow students. I am particularly in debt to Qizhi Huang, who analyzed the PAON-4 data with the code described in the thesis. Shifan Zuo collaborated with me on the research presented in this thesis. I also thank Zhigang Li, Yichao Li, Jixia Li.

My visit to Universite de Paris-Sud was supported by the China Scholarship Council. I would also like to thank LAL and its director Prof. Achille Stocchi, for providing me this wonderful opportunity to conduct research here, and to give continued support.

I would like also to thank all my jury members: M. Achille Stocchi, M. Steve Torchinsky and especially the two referees, M. Martin Bucher and M. Tongjie Zhang, who reviewed my manuscript.

Finally, my deepest gratitude is for my husband Zhongxiao Xu and my parents, who had always standed behind me and supported me with their love.

Contents

Al	ostrac	t	3
Ré	ésumé		5
Ac	know	ledgements	9
In	trodu	ction	23
1	Cos	mology	25
	1.1	Friedmann equations	25
	1.2	Thermal history and Structure growth	26
		1.2.1 Thermal history	27
		1.2.2 Structure growth	28
	1.3	Cosmic probes of Dark Energy	30
		1.3.1 Type Ia Supernovae	31
		1.3.2 Galaxy Clusters	31
		1.3.3 Weak Lensing	32
		1.3.4 Baryon Acoustic Oscillations	32
2	HI (Observations	35
	2.1	Radio telescope	35
		2.1.1 Single Dish	35
		2.1.2 Interferometer Array	36
	2.2	21 cm line and HI evolution	37
		2.2.1 Hydrogen 21cm line	37
		2.2.2 21cm line evolution	38
	2.3	HI galaxy surveys	39
	2.4	Projects for Epoch of Reionization	42
	2.5	Intensity mapping and Dark Energy Observation	43
	2.6	PAON-4 and Tianlai projects	45
		2.6.1 The PAON-4 array	45
		2.6.2 The Tianlai project	47

3	Map) makin	g for transit interferometers	51
	3.1	Visibil	ities	51
3.2		Classic	cal radio interferometry	52
	3.3	Non-tra	acking transit interferometers in the planar geometry	53
		3.3.1	Convolution theorem for reconstruction	54
		3.3.2	East-west transit observations	54
		3.3.3	Instrument pointing in the north-south plane	55
	3.4	Non-tra	acking transit interferometers in the spherical geometry	57
	3.5	Solving	g the system	61
	3.6	Variou	s indicators	64
		3.6.1	Reconstructed maps and PSF	64
		3.6.2	Instrument response and transfer function	66
		3.6.3	Error covariance matrix and noise power spectrum	67
		3.6.4	Filtering in (ℓ, m) space and angular masking $\ldots \ldots \ldots \ldots \ldots$	69
	3.7	Extens	ion to polarisation	69
4	Арр	lication	to different dense configurations	73
	4.1	Applic	ation to PAON-4	73
		4.1.1	Beams and response matrix features	73
		4.1.2	PAON4 beam and (ℓ, m) plane response	77
		4.1.3	PAON4 noise power spectrum and transfer function	81
	4.2	Applic	ation to the Tianlai Dish Array	85
		4.2.1	The Tianlai dish array mid-latitude survey	87
		4.2.2	The Tianlai dish array polar cap survey	93
		4.2.3	Tianlai 16-dish array sensitivity	98
	4.3	Applic	ation to the Tianlai Cylinder array	100
	4.4	The reg	gular array configuration	101
	4.5	The irr	egular array configuration	107
5	PAC	N-4 des	scription and first maps	113
	5.1	PAON	-4 detailed description	113
		5.1.1	PAON-2	113
		5.1.2	PAON-4	113
	5.2	PAON	-4 observations and data	119
	5.3 Data Processing and Map making		rocessing and Map making	123
		5.3.1	Data cleaning and calibration procedure	123
5.3.2 Quick Map Makin		5.3.2	Quick Map Making	127
	5.4	Instrun	nent and Map making performance	132
		5.4.1	Comparing data and simulations for the auto and cross correlation signals	133
		5.4.2	Reconstructed maps from PAON4 data using m-mode decomposition	136
		5.4.3	Effect of eigenvalue threshold on reconstructed maps	139

CC	CONTENTS	
Su	mmary and Outlooks	141
Bi	bliography	143
Ap	opendices	157
A	One-dimension reconstruction A.1 East-West direction A.1.1 Pseudo-inverse of matrix A A.1.2 Some results of response matrix $B * A$ (EW) A.2 North-South direction	159 159 160 161 162
B	A.2.1 Some results of response matrix $B * A$ (NS)	165 169
С	The two published papers	173

CONTENTS

List of Figures

1.2.1 Th	nermal history of the universe.	27
1.2.2 The ani (20	ne power spectrum of the cosmic microwave background radiation temperature isotropy in terms of the angular scale. The data shown comes from the WMAP 006), Acbar (2004) Boomerang (2005), CBI (2004), and VSA (2004) instru-	• •
me	ents. The theoretical model (solid line) is also shown.	29
1.3.1 Laı [32	arge-scale redshift-space correlation function of the SDSS LRG sample from 2]. The error bars are from the diagonal elements of the mock-catalog covari-	21
and		54
2.1.1 Let	eft: The 100 meter Green Bank Telescope; Right: FAST	36
2.1.2 Ver	ery Large Array	37
2.2.1 The is 2 ran by pro- eve sur the esc sig	the 21cm signal evolve with redshift, from [40]. In this figure, the highest redshift 25, so the first stage is not shown. The different linetype present different pa- meter models, so the 21cm signal evolution varies. The fiducial model is shown the solid curve. The long-dashed curve take the model with stronger escape obability of ionizing photon, and the universe is obviously ionized much earlier, en before heating, so there is no emission signals. The short-dashed curve as- me increasing the heating rate, further decreasing the strength and duration of e absorption epoch. The dot-dashed curve shows the model with both greater cape probability of ionizing photon and greater the heating rate, so the emission gnals appear earlier with short duration.	40
2.3.1 Plo as v stri	ot of the normalised redshift distribution of the HIPASS and ALFALFA surveys well as the distribution of the ALFALFA sources in the three, spatially separated ips.	41
2.3.2 Up	opper left: The ALFALFA HIMF; Bottom left: histogram of the number of HI	
sou	urces; Right hand: histogram of the contribution to Ω_{HI} of galaxies binned by	
HI	I mass.	41
2.5.1 BII	NGO	44
2.5.2 CH	HIME	44
2.5.3 HI	IRAX	44
2.6.1 PA	AON-4 interferometer configuration	46

2.6.2 The regular or square array (Left) and the circular array (right) configurations for	10
2.6.3 The Tianlai Dish and Cylinder Array.	40 50
3.3.1 The reconstructed map in the planar geometry	57
3.4.1 Schematic representation of m-mode decomposition in spherical harmonics space of the skymap reconstruction from visibilities.	61
3.6.1 Input maps for our simulation. Top: the sky brightness temperature at 1420.4 MHz (21-cm) from the LAB survey; Bottom: the sky brightness temperature at 1250MHz (radio continuum) computed using the GSM model.	65
3.6.2 Weight function $W(\ell)$ applied to computed $\widehat{\mathcal{I}(\ell,m)}$. The black curve suit for PAON4 with autocorrelation, while the red curve suit for PAON4 without autocorrelation. The blue curve is for Tianlai dish array without autocorrelation.	70
4.1.1 The beam patterns in spherical harmonics $\mathcal{L}_{\ell,m}$ with dish size $D = 4.5m$. Top left: auto-correlation located at equator; top right: cross-correlation beam from a pair of dish in North-South direction with $d_{sep} = 7m$ located at equator; bottom left: cross-correlation beam from a pair of dish in East-West direction with $d_{sep} = 7m$ located at equator; bottom right: cross-correlation beam from a pair of dish in East-West direction with $d_{sep} = 7m$ located at declination $\delta = 60^{\circ}$.	74
4.1.2 Examples of \mathbf{R}_m matrix. Here we assume that the band-like region between $35^{\circ}23'$ and $59^{\circ}23'$ is surveyed. Top: the \mathbf{R}_m matrix for an auto-correlation of a single dish with $D_{\text{eff}} = 4.5$ m. Top Left: $m = 0$; Top Center: $m = 50$; Top Right: $m = 100$. Bottom: the \mathbf{R}_m matrix for an auto-correlation of a pair of dishes in EW baseline and $d_{\text{exp}} = 7$ m. Bottom Left: $m = 50$: Bottom Center:	
m = 100; Bottom Right: $m = 250$.	75
4.1.3 The R matrix for 4 dishes with 15m baselines and pointing at $35^{\circ}23'$ and $59^{\circ}23'$. 4.1.4 The R (ℓ m) for 4 dishes with 15m baselines observing at 1420 MHz (top) and	77
1250 MHz (bottom) for the survey of the region with $35^{\circ}23'^{\circ} < \delta < 59^{\circ}23'^{\circ}$.	78
4.1.5 Comparison of the PAON-4 synthetic beam(left panel) with that of the compact 2×2 array (center panel) and the D=15.5 m single dish (right panel).	78
4.1.6 Comparison of R matrix for PAON-4 (left), 2×2 (center) and D=15.5 m single dish (D _{eff} = 14 m) configuration (right)	79
4.1.7 The matrix R with (top) and without (bottom) auto-correlations for the PAON-4 case with 1420 MHz (Left) 1250 MHz (Middle) and 1200 MHz (Right) frequencies	79
4.1.8 Comparison of the error variance matrix for PAON-4 (left), 2×2 (center) and	17
D=15.5 m single dish ($D_{eff} = 14 \text{ m}$) configuration (right)	81
4.1.9 Comparison of the transfer function $T(\ell)$ (left panel) and the noise power spectrum $C^{\text{noise}}(\ell)$ (right panel) for PAON-4 array in red, compact 2×2 array in blue and the single dish in black. Solid lines: map-making with auto-correlations, Dashed	
lines: map-making without auto-correlations.	82

LIST OF FIGURES

4.1.10Comparison of the transfer function $T(\ell)$ after filtering and masking (left panel) and noise power spectrum $C^{\text{noise}}(\ell)$ after filtering and masking (right panel) for PAON-4 in red, compact 2×2 in blue and single dish in black. Solid lines: map- making with auto-correlations, Dashed lines: map-making without auto-correlations.	
4.1.1 Histograms of reconstructed noise map pixel values for PAON-4. Top panel: pixel-	82
lation with HEALPix $n_{side} = 256$, black curve: the raw map; red curve: filtered map with $\ell_{max} \sim 420$. Bottom panel: blue curve: filtered map with $n_{side} = 128$	
and $\ell_{\text{max}} \sim 375$; magenta curve: $n_{\text{side}} = 64$ with $\ell_{\text{max}} \sim 190$	83
figuration (right) as a function of separation scale d_{sep} and antenna pointing angle with respect to the zenith.	86
4.2.2 Comparison of Tianlai 16-dish configuration (circular) with a regular 4×4 array (center). The beam for the Tianlai 16-dish circular configuration, without the autoparalation signal is shown on the right papel. $4.4 \times 4.4 \text{ dog}^2$ high resolution	
area extracted from the reconstructed maps, centered on a point source position.	86
4.2.3 Comparison of the response matrix $\mathbf{R}(\ell, m)$ for the 4 × 4 regular 16 dish config- uration (top) and the Tianlai circular 16 dish configuration (bottom) at 1420MHz	0.0
 (left) and 1250MHz (right). 4.2.4 Reconstructed map for the Tianlai-16 dish configuration (bottom) at 21 cm, and the original 21 cm map (top) from the LAB survey. Modes with l ≤ 100 have 	88
been filtered out of the LAB map. Only a band in declination $20^{\circ} < \delta < 70^{\circ}$ is shown, the planned survey area is $30^{\circ} < \delta < 60^{\circ}$ as discussed in the text.	89
4.2.5 Comparison of the error variance matrix for Tianlai-16 dish circular configuration (right) with a regular 16 dish 4×4 array (left) at 1420MHz.	90
 4.2.6 The transfer function for the 16-dish regular (left) and circular (right) configurations. The curves are purple: 1420MHz, red: 1250MHz, green: 1200MHz. 4.2.7 Comparison of the noise power spectrum C^{noise}(ℓ). The green and blue curves 	90
correspond to the 16-dish circular layout, while red and orange curves correspond to the regular 4×4 array. Blue and orange noise power spectrum have been	
obtained after filtering in (ℓ, m) space and masking	92
Black curve: raw map; red curve: filtered map with $\ell_{max} \sim 1100$. Bottom panel: $n_{side} = 256$; blue curve: filtered map with $\ell_{max} \sim 780$; magenta curve: filtered	
map with $\ell_{\text{max}} \sim 420$	92 93
4.2.10Tianlai 16-dish polar cap survey. Top: the response matrix $\mathbf{R}(\ell, m)$; Bottom: The	04
 4.2.1 ILeft: the input polar cap map after a high pass filtering of ℓ > 100. Right: The reconstructed map from mock visibilities corresponding to the Tianlai 16-dish polar cap survey. The polar cap region with a radius of 20° is represented, while the largest dedication survey of the magnetic (150) 	74
lowest declination scan of the survey is at 15° .	95

4.2.12The transfer function $T(\ell)$ (left panel) and noise power spectrum $C^{\text{noise}}(\ell)$ (right	
panel) for the Tianlai 16-dish polar cap survey.	96
4.2.13Histograms of reconstructed map pixel values for the polar cap survey. Top panel:	
$n_{side} = 512$. Bottom panel: $n_{side} = 256$. Black curve: raw map; red curve:	
filtered map with $\ell_{max} \sim 1100$. Bottom panel: $n_{side} = 256$. Blue curve: filtered	
map with $\ell_{\rm max} \sim 780$; magenta curve: filtered map with $\ell_{\rm max} \sim 420$	97
4.2.14Noise power spectrum for the Tianlai mid-latitude survey and PAON-4, compared	
to Galactic 21 cm signal power spectrum. The power spectrum of the LAB map si	
shown in black, and the red curve corresponds to its extension to higher resolution	
$(\ell > 750)$ using GALFA. The blue curve is the noise power spectrum for the	
PAON-4 case, and the magenta curve represent the Tianlai 16-dish array mid-	
latitude surevy. The noise power spectrum have been computed without the auto-	00
	98
4.2.15The expected angular noise power spectra $C^{\text{noise}}(\ell)$ at 1050 MHz for a 2 year	
survey of a 20° band in declination for square arrays with 10 (black), 64 (brown), 144 (vallow) 256 (blue) dishes with $T_{\rm eff} = 100 {\rm K}$. The group surve correspond	
to the 256 dish array with $T_{sys} = 50$ K and the red curve shows the expected	
cosmological signal at 1050 MHz filtered by the instrument response $C^{21}(\ell) \times$	
$T(\ell)$.	99
4 3 1 Regular configuration of the cylinder array	100
4.3.2 GSM (Global Sky Model) map at 750 MHz, used as the true sky for the recon-	100
struction with the Tianlai cylinder configuration.	102
4.4.1 The primary beam (left) and synthetic beams for the regular 1 (center) and regular	
2 (right) configurations.	103
4.4.2 The beam patterns in spherical harmonics $\mathcal{L}_{\ell,m}$ with size $L_x = 13.5$ m, $L_y = 13.$	
0.3m and centered at latitude 44.15°. Top Left: auto-correlation of a feed; Top	
Right: cross-correlation for a EW baseline with $d_{sep} = 15m$; Bottom Left: cross-	
correlation beam for a NS baseline with $d_{\rm sep}=12{ m m}$; Bottom Right: cross-	
correlation for a SE-NW baseline with $(\Delta x, \Delta y) = (15m, 12m)$.	104
4.4.3 Comparison of the R matrix for regular 1(left) and regular2 (right) configurations.	105
4.4.4 Comparison of the error variance matrix for regular 1(left) and irregular2 (right)	
configurations.	105
4.4.5 Reconstructed sky map for the Tianlai cylinder configuration at 750 MHz. Top:	
Regular 2 configuration; Bottom: Regular 1 configuration. The input map is the	100
GSM map at 750 MHz.	106
4.5.1 The synthetic beam for Tianlai cylinder irregular configuration.	108
4.5.2 Comparison of the R matrix for the Irregular 1 (left) and Irregular2 (right) config-	100
urations	108
4.5.5 Comparison of the error variance matrix for irregular configurations.	109
4.5.4 Reconstructed sky map for the irregular 1 (top) and irregular 2 (bottom) configu-	110
	110

LIST OF FIGURES

4.5.5 Comparison of the transfer function $T(\ell)$ (left panel) and the noise power spectrum $C^{\text{noise}}(\ell)$ (right panel) for the Irregular 1 and Irregular 2 configurations.	1 . 110
5.1.1 PAON2.	. 114
5.1.2 Cross-correlations between aligned polarizations during the observation of CygA	L
transit	. 114
5.1.3 Cross-correlations between aligned polarizations during the observation of CygA transit.	. 115
5.1.4 PAON-4 interferometer configuration	. 116
5.1.5 The PAON-4 accessible sky area	. 116
5.1.6 DIGFFT board (4 analog channels sampled at 500 MSample/s with 8 bits dynamic range) capable of digital filtering (FFT) developed for the BAORadio project.	: . 117
5.1.7 The current PAON-4 cabling scheme (figure provided by Claude Pailler).	. 118
5.2.1 The current arrangement of visibilities over the matrix rows (<i>Table provided by Claude Pailler</i>).	, . 119
5.2.2 The example of visibility matrix/FITS file header.	. 120
5.2.3 The summaries of the observations towards Cygnus A and Cassiopeia A (CasA performed with PAON-4 in the last quarter of 2016 (<i>Tables provided by Claude Pailler</i>).) ? . 121
5.2.4 The auto-correlation of the raw TOD at 1400MHz (top panel) and 1420MHz (bot tom panel).	- . 122
5.2.5 The cross-correlation of the raw TOD at 1400MHz (left panel) and 1420MHz (right panel).	<mark>.</mark> . 122
5.3.1 A: The $g(\nu)$ by auto-correlation for the third antenna panel of PAON-4. B: Time dependent gain $G(t)$.	e . 124
5.3.2 2D time-frequency map for cross-correlation $(6m, 0)$.	. 125
5.3.3 Data Masking for identify and remove the RFI.	. 125
5.3.4 Corrected TOD using Gain	. 126
5.3.5 Left: The auto-correlation with amplitude calibration. Right: sigma noise and system temperature for each channel.	<mark>1</mark> . 127
5.3.6 Phase correction term.	. 129
5.3.7 Quick map making at 1420Mhz; Top: all observed region. Left: extracted from top figures from RA=55° to RA=95°; Right: extracted from top figures from RA=280° to RA=325°.) . 130
5.4.1 The auto-correlation signals comparison of PAON-4 data (red curve) and simu	-
lated visibilites (black curve) for the eleven scans at frequency 1420.4 MHz.	. 134
5.4.2 The cross-correlation signal comparison between simulated signals (black curve from LAB+CygA map and the PAON-4 observed visibilities (red curve) at 1420.4 MHz. The red curves show the cross-correlation from PAON-4 November data) - ,
and the black curves is simulated from LAB data.	. 134

5.4.3 The corrected cross-correlating signals for the pointing of 0° at 1420MHz for base-	
line with Δz . The red curves show the cross-correlation from PAON-4 November	
data, and the blue curves is simulated from LAB data.	136
5.4.4 Reconstructed map for PAON-4 at 1420MHz. A: input LAB+CygA map; B: re-	
constructed map from LAB+CyaA map after filtering; C: reconstructed map from	
November data after filtering.	137
5.4.5 Extracted map around the CygA source region (right ascension from 290° to 310° ;	
and declination from 46.73° to 34.73) from the reconstructed map. A : extracted	
map from input LAB+CygA map; B: extracted map from reconstructed map using	
LAB+CyaA map after filtering; C: extracted map from reconstructed map using	
Nov data after filtering.	138
5.4.6 Reconstructed map for PAON-4 at 1400MHz. A: input map from Haslam+point	
source at 1400MHz; B: reconstructed map from November data after filtering.	138
5.4.7 The comparison of the reconstructed maps from the simulations with a sharp and	
smooth threshold for PAON-4 at 1420MHz.	139
5.4.8 The comparison of the reconstructed maps from the data with a sharp and smooth	
threshold for PAON-4 at 1420MHz.	140
A.1.1 The beam patterns in Fourier space with $d_{sep} = 2m$, dish size $D = 1.75m$ and	
frequency $\nu = 1200$ MHz.	162
A.1.2 The $B * A$ matrix.	163
A.2.1 The response matrix $B * A$ for the case of 24 feeds layout in line along North-	
South direction, and each feed has a single dipole $\sim 0.4m$ long, and feeds are	
separated by 1.6m.	164
A.2.2 The diagonal element of the Fig A.2.1, blue and red correspond to the two fre-	1.6.4
quencies $\nu_1, \nu_2, \ldots, \nu_n$	164
A.2.3 The response matrix $B * A$ for the case of 24 feeds layout in line along North-	
South direction, each feed is made of 4 packed dipoles, with the output signal	164
corresponding to the sum of individual dipole signals.	164
A.2.4 The sketch of dish array configurations.	160
A.2.5 The response matrix $B * A$ for 12 dish, with uniform spacing and 1 pointing.	16/
A.2.0 The response matrix $B * A$ for 12 dish, with uniform spacing and 5 pointing.	10/
A.2.7 The response matrix $B * A$ for 12 dish, with random spacing and 1 pointing.	10/
A.2.6 The response matrix $D * A$ for 12 dish, with random spacing and 5 pointing	108
B.0.1 Block Diagram for the map-reconstruction code	170

List of Tables

1.1	Cosmological Parameter Values	26
2.1	Coordinates of the circular array antenna.	49
3.1	The SVD computation for a few <i>m</i> -modes in the PAON-4 case.	63

Introduction

For thousands of years, with the enlightenment of human cognition, our view of cosmos has evolved from myth to philosophical reflections, and then to scientific theories or models based on observational facts and constructed in mathematical frameworks. As our observational technology advanced, we penetrates into ever deeper space. Since the 20th Century, a modern cosmological theory based on Einstein's general relativity theory of spacetime has been developed. In the last decade, a new era of precision cosmology dawned, as large amount of high precision cosmological observational data are obtained. These include Cosmic Microwave Background (CMB) data from the COBE, WMAP and Planck satellites and many ground-based or balloon-borne experiments; Type Ia supernovae observed by a number of sky surveys, as well as galaxy redshift surveys such as the 2dF and Sloan Digital Sky Survey (SDSS).

Amongst these observations, the galaxy redshift surveys probe the large scale structure (LSS) of the Universe and the evolution of galaxies. The LSS data can be used to measure cosmological parameters with high precision, e.g. from the shape of power spectrum one can measure the relative density of the cold dark matter and the masses of neutrinos, and through the baryon acoustic oscillation (BAO) the cosmological expansion history and equation of state of dark energy. By testing the primordial non-Gaussianity, it could also be used to probe the initial conditions and the very early history of the Universe, and the growth of structure could be used to test modified gravity theories. Besides the fundamental physics and cosmological models, the galaxy redshift surveys also enables study of how environments affect formation and evolution of galaxies.

Complementary to optical surveys, measurement of the neutral hydrogen (HI) distribution through its 21 cm line radiation may also be used to study the statistical properties of LSS in the Universe. This provides a tracer of the large scale structure which differs from that of optical surveys. The radio observation is also potentially less affected by atmospheric absorption at high redshifts. A drawback of the radio observation is that due to its long wavelength, the angular resolution is relatively low for large area surveys. However, the intensity mapping technic which is to survey the LSS with low angular resolution without resolving individual galaxies has been proposed as an efficient and economical way to map large volumes of the universe using the redshifted 21 cm emission. Such cosmological surveys would not only cover the large volumes of the universe, but also out to very high redshifts towards the observation of the Epoch of the Reionization (EoR). However, compared to the various foreground radiation such as the galactic synchrotron emission, galactic free-free emission, and the extragalactic radio sources, the 21cm signal is very weak. Although the 21cm signal and the various foregrounds have different characters, e.g. the

former is stochastic along the line of the sight (l.o.s.), while the latter are highly correlated along the l.o.s., so in principle the signal could be extracted from the foregrounds, but in reality this is a very challenging task. Currently, many EoR and intensity mapping experiments are trying various techniques to detect the 21cm signal.

My thesis research is a part of these efforts. The PAON and Tianlai projects are the French and Chinese pilot experiments for intensity mapping of LSS with radio interferometer arrays. The work presented here focuses on sky map reconstruction algorithms from visibilities for transit observations. This is the basic function for a radio array, and the first step towards 21cm measurement.

This dissertation is organized as follows. In the first chapter below, we briefly review some basic concepts of standard cosmology model, including the thermal history, structure growth, and several methods which can be used to probe the Dark Energy. In the second chapter, we review the neutral hydrogen (HI) content and its evolution in the Universe, and briefly introduce some recent HI surveys projects. Then we describe the HI intensity mapping concept and some EoR and intensity mapping projects. Finally, we introduce the PAON4 and Tianlai projects in some detail. In the third chapter, we presents an overview of the map making algorithm from full east-west transit (24 hours) visibilities or interferometric observations. In the fourth chapter, we discuss the application of the method to the PAON-4 telescope which is a 4 antennae test interferometer and the optimization of the array configuration. We present the expected beam shapes and the noise power spectrum for PAON-4 compared with a regular 2×2 array and a single dish telescope. Then, we also present the comparison of beam and noise power spectrum for several array configurations that we have considered for the Tianlai 16-dish pathfinder array, as well as a short discussion of array sensitivities to extragalactic and cosmological 21-cm signals. In the last chapter, we will illustrate the development of the PAON-4 and give some detailed description of the PAON-4 electronic and acquisition system. The PAON-4 have made some test observations last year, and we describe here the calibration procedure of the time ordered data (TOD), and present the first map made from these observations.

Chapter 1

Cosmology

1.1 Friedmann equations

In 1929, Hubble found that the recession speed of galaxies is proportional to its distance, this is the so called Hubble's law, the proportional constant $H_0 \equiv 100h \,\mathrm{km \, s^{-1} Mpc^{-1}}$ is called the Hubble constant. The expansion of the universe, a concept originated from the general theory of relativity, provides a physical explanation for the Hubble's law. If we adopt the basic hypothesis that the universe is homogeneous and isotropic on large scales (the cosmological principle), we could describe the 4-dimensional space-time with the Friedman-Lemaître-Robertson–Walker (FLRW) metric,

$$ds^{2} = c^{2}dt^{2} - a^{2}(t)\left(\frac{dr^{2}}{1 - kr^{2}} + r^{2}d\Omega^{2}\right)$$
(1.1.1)

where c is the speed of light. k = +1, -1, 0 corresponds to closed, open, and flat geometries, respectively. a(t) is the scale factor of the universe with cosmic time t, and (r, Ω) are comoving coordinates. By substituting the metric Eq. (1.1.1) into the Einstein field equations, we obtain the Friedmann equations, which describes the dynamical evolution of the background universe.

$$(\frac{\dot{a}}{a})^2 = -\frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2}$$
(1.1.2)

$$\ddot{a} = -\frac{4\pi G}{3}(\rho + 3p)a + \frac{\Lambda}{3}a$$
 (1.1.3)

with

$$H \equiv \frac{\dot{a}}{a} \tag{1.1.4}$$

The dot represent a derivative with respect to cosmic time t. The quantity H, called the Hubble parameter, describe the expansion rate of the universe. G is gravitational constant. ρ and p are the cosmic energy density and pressure respectively. A number of observations has shown that the cosmic expansion is accelerating, indicating the presence of a new form of matter in the Universe,

Symbol	$\Omega_b h^2$	$\Omega_c h^2$	t_0
Value	0.02230 ± 0.00014	$0.1188 {\pm} 0.0010$	$13.799 \pm 0.021 \times 10^9$ years
Symbol	n_s	Δ_R^2	au
Value	0.9667 ± 0.0040	$2.441^{+0.088}_{-0.092} \times 10^{-9}$	$0.066 {\pm} 0.012$

 Table 1.1: Cosmological Parameter Values

i.e. the dark energy. The cosmological constant Λ is the simplest form of dark energy, though other dark models have also been proposed. The simplest Λ CDM model has six parameters: physical baryon density parameter $\Omega_b h^2$; physical dark matter density parameter $\Omega_c h^2$; the age of the universe t_0 ; scalar spectral index n_s ; curvature fluctuation amplitude Δ_R^2 ; and reionization optical depth τ . Current observations such as CMB anisotropy, the brightness/redshift relation for supernovae, and large-scale galaxy clustering including the baryon acoustic oscillation feature are in excellent agreement with the Λ CDM model, and can be used to fit the value of the model parameters. The parameter values listed below are obtained from the Planck CMB power spectra, in combination with lensing reconstruction and external data (+BAO+SN) [91].

We set the non-relativistic matter (baryon and cold dark matter) density $\Omega_m = \Omega_b + \Omega_c$, and then in terms of the today's values of the density parameters, the first Friedmann equation could be put in the following form:

$$\frac{H^2}{H_0^2} = \frac{\Omega_{r0}}{a^4} + \frac{\Omega_{m0}}{a^3} + \frac{\Omega_{k0}}{a^2} + \Omega_{\Lambda 0}$$
(1.1.5)

with $\Omega_{r0} + \Omega_{m0} + \Omega_{k0} + \Omega_{\Lambda 0} = 1$, where Ω_{r0} is the radiation density today; Ω_{m0} is the non relativistic matter density today; Ω_{k0} is the spatial curvature density today; $\Omega_{\Lambda 0}$ is the cosmological constant or vacuum density today. Different parameters combination of Ω_m , Ω_r , Ω_Λ represents different universe model. In the standard cosmological model or the Λ CDM cosmological model, the dark matter started to dominate at low redshifts. The detail information about the cosmological parameters could be found at [89].

1.2 Thermal history and Structure growth

If one accepts cosmic expansion, and assumes that the current physical laws are at work throughout the whole space time, then as we want back to early time, the density of the Universe must be increasingly denser. According to the standard cosmological model, the universe started from the Big Bang, which happened about 14 Gyr ago. The extrapolation however has to stop at a time when Universe was extremely hot and dense state where we run into a singularity, and the laws of physics we currently know would become invalid.

In this section, we will briefly overview the thermal history of the universe, which mainly introduce the major events in each evolution stage after the Big Bang (Fig. 1.2.1).



Figure 1.2.1: Thermal history of the universe.

1.2.1 Thermal history

Inflation is a theory that the early Universe expanded in accordance with an exponential laws. According to standard lore, the inflationary epoch may have started from 10^{-36} second after the Big Bang to sometime between 10^{-33} and 10^{-32} second. During this time the Universe expand at a very rapid rate. This theory solved the Horizon problem, the Flatness problem and the Monopole problem of the standard hot big bang model. After the inflation, the large potential energy of inflaton decayed into particles, reheating the Universe. During inflation, small quantum fluctuations are stretched to large scales, which later grow into the structure of today's universe. After the inflationary epoch, the Universe continue to expand, and the Baryogenesis, i.e. the production of asymmetry between matter and anti-matter, may took place during this period. As the density and temperature declines, the various fundamental interactions of nature decoupled sequentially. After the Hadron and Lepton epoch, the matter which remain coupled and in thermo-equilibrium are photons, neutrinos, electrons and positrons, and proton and neutrons.

The Big Bang Nucleosynthesis (BBN) happened within the first 3 minutes after the Big Bang, during this period the positrons annihilated, and neutrinos decoupled. The neutrons and some protons are combined to form light nuclei, the most abundant is ⁴He, as well as few ³He, ³H, ²H, ⁷Li. Heavier elements were only produced in later stages of Universe evolution, by the nuclear reactions in stellar cores and supernova explosions. It is worth noting that the light elemental abundance calculated by the standard cosmological model is in highly agreement to the observational result, which gave powerful evidence to the validity of the Big Bang model.

When temperature at $T = T_{eq}$, the Universe entered into a matter-dominated period from radiation-dominated period. Before T~4000K, because of the strong coupling among radiation

and free electrons, the universe was non-transparent. When the temperature continually went down to 4000K, most of the protons together with free electrons combined into neutral hydrogen atoms, and then the universe become transparent. Matter decoupled with radiation, thus forming the background radiation that suffusing the whole universe. This period was called Recombination. The cosmic microwave background (CMB) was the electromagnetic radiation left over from that period. The CMB is a prediction of the Big Bang theory, and after its discovery the Big Bang theory is widely accepted.

The Dark Age is the period between recombination and the formation of the first stars. Based on the standard cosmological model, during this time there is no stars or other structures, perhaps the only observable signal from this period is the hydrogen 21cm line. At its end, the first stars formed in the earliest nonlinear structure, and their first ray of radiation light up the gas around them, mark the end of the dark ages and the beginning of the Epoch of Reionization (EoR).

Reionization is the process that the baryonic matter, mostly hydrogen, gets re-ionized after the dark ages. With the increasing formation of stars and galaxies, more ionized regions (HII regions) appeared and grown larger, and these regions merged with each other, so more and more intergalactic medium was ionized. Eventually the intergalactic medium were totally ionized, only the neutral hydrogen in dense regions such as galaxies or mini-halos where the recombination rate is higher may remain neutral, and the epoch of reionization ended.

The CMB radiation is an emission of uniform, black body thermal energy coming from all parts of the sky. The anisotropy of the cosmic microwave background [111] could be divided into two types: primary anisotropy, due to effects that occur at the last scattering surface and before; and secondary anisotropy, due to effects such as interactions of the background radiation with hot gas or gravitational potentials, which occur between the last scattering surface and the observer. The primary anisotropy, though random at first sight, has rich structure in its angular power spectrum, which are generated by the acoustic oscillations and diffusion damping processes in the photon-baryon plasma, and give its characteristic structure of peaks. Fig 1.2.2 shows the angular power spectrum of the cosmic microwave background radiation temperature anisotropy. Roughly speaking, the position of the first peak can be used to determines the curvature of the universe. The height of the second peak is sensitive to the reduced baryon density, and the third peak can be used to get information about the dark-matter density.

1.2.2 Structure growth

We now consider how the density fluctuations in the universe grow under the gravity. We can investigate this problem by perturbation theory. In the perturbation theory, we assume a homogeneous background, then consider small fluctuation. The linear analysis is only applicable when the perturbations are small, the growth would be saturated when the perturbations become large. Taken the expansion into consideration, the gravitational instability becomes weaker. Using the approximation of the long wavelength and neglecting the pressure terms that are proportional to k^2 , from the Newtonian equations,

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\delta - 4\pi G\rho_0\delta = 0 \tag{1.2.1}$$



Figure 1.2.2: The power spectrum of the cosmic microwave background radiation temperature anisotropy in terms of the angular scale. The data shown comes from the WMAP (2006), Acbar (2004) Boomerang (2005), CBI (2004), and VSA (2004) instruments. The theoretical model (solid line) is also shown.

The solutions shows that the perturbations in the non-relativistic component does not grow much during the radiation dominated era. After matter domination δ grows with the scale factor a(t). The growth stops once the universe become cosmological constant dominated (in the flat ΛCDM model). A useful fit is given by Carroll et at.(1992):

$$\delta_+(z) \sim D(z) \sim \frac{g(z)}{1+z} \tag{1.2.2}$$

where

$$g(z) \approx \frac{5}{2} \frac{\Omega_m(z)}{\Omega_m(z)^{4/7} - \Omega_\Lambda(z) + [1 + \frac{\Omega_m(z)}{2}][1 + \frac{\Omega_\Lambda(z)}{70}]}$$
(1.2.3)

The matter perturbation modes which entered horizon before the matter-radiation equality would not grow, while the modes which entered the horizon after the matter-radiation equality could grow by $\delta \sim a$. As a result, the power spectrum of perturbation density acquires the shape which is $P(k) \sim k^n$ with $n \sim 0.95$ at scales greater than about $0.02h \,\mathrm{Mpc}^{-1}$, but decreases at $k > 0.02h \,\mathrm{Mpc}^{-1}$. Generally, one can define the transfer function T(k):

$$\Phi(k,a) = \Phi_p(k)T(k)g(a) \tag{1.2.4}$$

where

$$T(k) = \frac{\Phi(k,a)}{\Phi_{large}(k,a)}$$
(1.2.5)

Here Φ is the potential on very large scales. This definition ensures the normalization on very large scale T(k) = 1. We then have

$$\delta(k,a) = \frac{3}{5} \frac{k^2}{\Omega_m H_0^2} \Phi_p(k) T(k) D_1(a)$$
(1.2.6)

and

$$P(k,a) = 2\pi^2 \delta_H^2 \frac{k^n}{H_0^{n+3}} \left(\frac{D_1(a)}{D_1(a=1)}\right)^2$$
(1.2.7)

The matter transfer function can be calculated with the Boltzmann code. An often-used analytical fit for the adiabatic model is given by

$$T_k = \frac{\ln(1+2.34q)}{2.34q} [1+3.8q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4]^{-1/4}$$
(1.2.8)

where $q=\frac{k/hMpc^{-1}}{\Gamma}$ where $\Gamma=\Omega_m h$ is the shape parameter.

Eventually the perturbation in the matter density became so strong that the linear approximation break down, as shown vividly by the existence of stars and galaxies and galaxy clusters. It is believed that these structures were formed in a two-step process. First, in regions where the density was a little larger than average, the cold dark matter together with baryonic matter expanded more slowly than the universe, eventually reaching a minimum density and then recontract. If an overdense region was sufficiently large, then its baryonic matter collapsed along with its cold dark matter.

Then in a second stage, after this collapse, the baryonic matter lost its energy through radiative cooling, and condensed into protogalaxies consisting of clouds of gas that eventually form stars. While the cold dark matter particles would not lost their energy through radiative cooling, so they remained in large more-or-less spherical halos around these galaxies.

1.3 Cosmic probes of Dark Energy

Dark matter is an unidentified type of matter comprising approximately 27% of the mass and energy in the observable universe that is also made up by dark energy, baryonic matter (ordinary matter), and neutrinos. The name refers to the fact that it does not emit photons, thus being invisible to the entire electromagnetic spectrum. Although dark matter has not been directly observed, its existence and properties are inferred from its gravitational effects, such as the motions of stars, gas and galacties, gravitational lensing, its influence on the universe's large-scale structure, and its effects on the cosmic microwave background. The DETF report [2] provide an overview of Dark Energy program.

There are several methods could be used for the probing. One of the developed observation is optical surveys or projects. They could be separated into imaging (photometric) and spectroscopic surveys. For example, The Dark Energy Survey (DES), The Large Synoptic Survey Telescope (LSST), The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS),

1.3. COSMIC PROBES OF DARK ENERGY

the Baryon Oscillation Spectroscopic Survey (BOSS), and Dark Energy Spectroscopic Instrument (DESI), all of them are spectroscopic surveys. EUCLID is a space mission under the development by the European Space Agency (ESA) project, which is featuring visible and infra imaging, as well as limited spectroscopic capabilities. The radio telescopes is also popular in the recent decades, which we will present in the next chapter.

1.3.1 Type Ia Supernovae

Type Ia supernova is a type of supernova that results from the violent explosion of a white dwarf star in binary systems. A white dwarf gradually accretes mass from its binary companion. As its mass approaches the Chandrasekhar limit, the internal pressure increases, and at some point thermonuclear explosion occurs [53]. Due to the nearly uniform mass at the explosion, SNIa produces consistent peak luminosity, which can be further calibrated with the Phillips relationship of peak luminosity and light curve. Therefore, it can be used as standard candles to measure the luminosity distance $d_L(z)$ of the host galaxies [68, 86], thus providing a useful tool to trace the accelerating expansion history of the universe, and helps us better understand dark energy. Please see [64] for more detailed history about the discovery of the accelerating expansion of the universe.

In recent years, many research groups are conducting supernovae surveys, such as the Supernova Legacy Survey (SNLS) [12, 47], the ESSENCE (Equation of state: SupErNovae trace Cosmic Expansion) [76, 110], the Nearby Supernova Factory(NSF) [25], the Carnegie Supernova Project (CSP) [38], the Lick Observatory Supernova Search (LOSS) [67], and the Sloan Digital Sky Survey (SDSS) [56, 62], etc. Currently, in the SNIa observation, the systematic errors are the major confining factors to precisely measure the properties of the dark energy.

1.3.2 Galaxy Clusters

Galaxy clusters (CL) are gravitational bound structures that typically contain 50 to 1000 galaxies and have a diameter from 2 to 10 Mpc with typical masses ranging from 10^{14} to 10^{15} solar masses. They are the largest gravitationally bound objects in the universe until superclusters were discovered [66]. CL have been found in a number of observations, and can be detected through the following approaches [14]:

- Optical or infrared imaging and spectroscopy: GL are found and confirmed by searching for overdensities and finding several galaxies at a similar redshift. In comparison, infrared searches are more useful in finding higher redshift clusters.
- X-ray imaging and spectroscopy: Active galactic nucleus (AGN) are the brightest X-ray emitting extragalactic objects, so clusters are quite prominent in X-ray surveys.
- Radio spectroscopy: GL of radio sources are used as tracers of cluster location,
- Sunyaev-Zel'dovich effect [118]: The hot electrons in the intracluster medium distort the CMB radiation through inverse Compton scattering, and observed distortions of CMB spectrum are used to detect the density perturbations of the universe.

• Gravitational lensing: The observed distortions due to CL bending the light from distant galaxies can be used to model the distribution of dark matter in the cluster.

In principle, the number density of cluster-sized dark halos n(z, M) as a function of redshift z and halo mass M can be accurately predicted from N-body simulations. Comparing these predictions to large area cluster surveys can provide precise constraint on the cosmic expansion history [48].

1.3.3 Weak Lensing

Weak lensing (WL) is an intrinsically statistical measurement of the slight distortions of distant galaxies images, due to the gravitational bending of light by structures in the Universe. It provides a way to measure the masses of astronomical objects without requiring assumptions about their composition or dynamical state. There are mainly four kinds of weak lensing: (1) WL of clusters of galaxies [104], which can reveal the dark matter distribution and properties and be able to constrain cosmological parameters. (2) Galaxy-galaxy lensing [16], which is used to measure mass density profiles of galaxy lens, mass-to-light rations, galaxy mass evolution and different types of galaxies evolution. (3) WL of large scale structure [46], which will be explained in details in next paragraph. (4) WL also has an important effect on CMB [69] and diffuse 21cm line radiation, which suit to probe cosmology at higher redshifts.

WL of LSS produces an observable tiny distortions in background galaxies. This distortion effect is much more subtle than the effect from cluster and galaxy-galaxy lensing. So surveys must be deep and wide, and the image quality must be very good. In 2000, this kind of WL was first detected by four independent research groups [109, 8, 59, 106], and subsequent observations have put constraints on cosmological parameters which is particularly the dark matter density Ω_m and power spectrum amplitude. From then on, WL has grown to be an increasingly accurate probe of DE and DM, and can compete with other cosmological probes. The current surveys aim to map out the 3D distribution of mass, which provide a useful tool to probe dark energy by its influence on the growth of structure. Many experiments have focused on weak lensing to measure the properties of DE and DM, such as the Dark Energy Survey (DES [37]), Panoramic Survey Telescope and Rapid Response System (Pan-STARRS [58]), and Large Synoptic Survey Telescope (LSST [65]).

1.3.4 Baryon Acoustic Oscillations

Baryon Acoustic Oscillations (BAO) could be understood as an overdensity or clustering of visible baryonic matter at certain length scales, resulting from acoustic waves which propagated in the early universe [82]. Similar to SNIa, which provides a standard candle for astronomical observations, BAO provides a standard ruler for length scale in cosmology to explore the expansion history of the universe [32]. The length of this standard ruler (\sim 150Mpc in today's universe [30]) corresponds to the distance that a sound wave travels from a point source at the end of inflation to the decoupling. BAO has a characteristic imprint on the matter pawer spectrum [55]. So it can be measured by looking at the large scale structure of matter. The apparent size of the BAO measured

1.3. COSMIC PROBES OF DARK ENERGY

from astronomical observation leads to the measurements of the Hubble parameter and the angular diameter distance [24].

The BAO observations are less affected by astronomical uncertainties than other probes of dark energy. The acoustic signature of BAO have already been obtained in the galaxy power spectrum at low redshift. Using a dedicated 2.5-m wide-angle optical telescope, the Sloan Digital Sky Survey (SDSS [113]), was launched in 2000. During its first phase of operations from 2000 to 2005, the BAO acoustic peak in correlation function is firstly measured by SDSS LRG sample, which is shown in Figure 1.3.1. Then, in 2005, the survey entered a new phase, i.e. the SDSS-II survey [101]. In 3 years, it completed the observations of imaging half the northern sky and mapping the 3-dimensional clustering of one million galaxies and 100,000 quasars. In mid-2008, SDSS-III started to collect data, and continued until 2014. It consists of four surveys executed on the same 2.5m telescope. The SDSS-III's Baryon Oscillation Spectroscopic Survey (BOSS [28]) focuses on mapping the Universe on the largest scales, creating the largest volume threedimensional map of galaxies ever created. Data Release 13 is the latest public release of BOSS spectra. Data Release 9 is the first public release of BOSS spectra. Prior to that, Data Release 8 included all BOSS imaging. The latest generation is SDSS-IV planing observation from 2014 to 2020. It contains the Extended Baryon Oscillation Spectroscopic Survey (eBOSS [103]), which extend precision cosmological measurements to a critical early phase of cosmic history, and aim at create the largest volume survey of the Universe to date.

BAO scales can also be measured by 21cm emission from neutral hydrogen, which after reionization are predominantly hosted in galaxies. This will be presented in next chapter.



Figure 1.3.1: Large-scale redshift-space correlation function of the SDSS LRG sample from [32]. The error bars are from the diagonal elements of the mock-catalog covariance matrix; however, the points are correlated.

Chapter 2

HI Observations

2.1 Radio telescope

Electromagnetic radiation in the frequency range between 3kHz ($\lambda \sim 10^5$ m) 300 GHz ($\lambda \sim 10^{-3}$ m) are usually called radio waves. Radio waves from space were first detected by Karl Guthe Jansky in 1932 at Bell Telephone Laboratories. Compared with artificial sources, astronomical radio sources are usually very weak, so radio astronomy telescope require large collecting area and highly sensitive detectors, and radio astronomical observatories are preferentially located away from artificial radiation sources such as broadcasting radio, television, cellphone, etc. to reduce the radio frequency interference (RFI).

A radio telescope may consist either a single antenna (often a large parabolic dish) or an array of radio antennas whose signals are electronically synthesized.

2.1.1 Single Dish

The angular resolution of a dish antenna is determined by the ratio of the wavelength of the radio waves and the diameter of the dish $\delta \sim \frac{\lambda}{D}$. To achieve high angular resolution and sensitivity, many-large size dishes have been built. For example, the Green Bank Telescope, located in West Virginia, United States, is the largest fully steerable dish radio telescope (GBT, [71], the left panel of Fig2.1.1). It has about 105-meter unblocked aperture, and operates in the frequency range of 0.1–116 GHz. The Effelsberg Telescope is the largest steerable single dish telescope in Europe, and also has an aperture of about 100m [108]. The Parkes Radio Telescope [17], completed in 1961 is located at New South Wales, Australia. The primary observing instrument is the 64-meter movable dishes in the world. The Arecibo radio telescope [79], located in Arecibo, Puerto Rico, has a 305 m dish, which was the largest single-aperture telescope from its completion in 1963 until 2016 when the FAST was completed. The antenna is only steerable within an angle of about 20° of the zenith by moving the suspended feed antenna.

The world's largest filled-aperture radio telescope at present is the Five hundred meter Aperture Spherical radio Telescope (FAST, [70], the right panel of Fig2.1.1) built by China, which is


Figure 2.1.1: Left: The 100 meter Green Bank Telescope; Right: FAST.

completed in 2016. The 500-meter-diameter dish is built into a natural depression in the landscape in Guizhou province; the feed antenna is in a cabin suspended above the dish on cables. The active dish is composed of 4450 moveable panels controlled by a computer. By changing the shape of the dish and moving the feed cabin on its cables, the telescope can be steered to point to any region of the sky up to 40° from the zenith. Although the dish is 500 meters in diameter, only a 300 meter circular area on the dish is illuminated by the feed antenna at any given time, so the actual effective aperture is 300 meters. The science objectives of the FAST radio telescope are: (1) Large scale neutral hydrogen survey; (2) Pulsar observations; (3) Leading the international vary long baseline interferometry (VLBI) network; (4) Detection of interstellar molecules; (5) Searching the interstellar communication signals; (6) Pulsar timing arrays.

2.1.2 Interferometer Array

The size of single antenna is limited due to structural strength, hence its angular resolution is not high. One can however, use interferometry of an array of antenna elements to achieve much higher angular resolution. The array may consist dishes or dipoles or other forms of antenna. All of the antennas in the array observed the same object, and were usually connected with coaxial cable, waveguide, optical fiber, or other type of transmission line. This increases the total collected signal. Radio interferometers creates a combined telescope whose angular resolution is equivalent to the largest baseline between array elements, and images of sky can be made by synthesis of the signals received by the array.

In order to produce a high quality image, a large number of different separations (baselines) between different telescopes are required. For example, the Very Large Array (VLA, [85], Fig 2.1.2) near Socorro, New Mexico, comprises 27 25-meter radio telescopes in a Y-shaped array.



Figure 2.1.2: Very Large Array.

These formed 351 independent baselines at any moment, which achieves a resolution of 0.2 arc seconds at 3 cm wavelengths. Another example of radio telescope array is the Giant Metrewave Radio Telescope (GMRT, [3]), located near Pune in India, which is an array of thirty fully steerable parabolic radio telescopes of 45 metre diameter, observing at metre wavelengths. The GMRT were built to search for highly redshifted 21-cm line radiation from primordial neutral hydrogen clouds in order to determine the epoch of galaxy formation in the universe [61], and observe many different astronomical objects such as HII regions, galaxies, pulsars, supernovae, and Sun and solar winds.

2.2 21 cm line and HI evolution

2.2.1 Hydrogen 21cm line

Hydrogen is the most abundant and ubiquitous chemical element in the universe, making up about 75% of normal matter by mass. It provides a powerful medium to study the properties of the universe, for example, epoch of reionization and large scale structure. This thesis focuses on the observations of the 21 cm emission line of neutral hydrogen. The 21 cm line arises from transition between two hyperfine levels of the hydrogen atom 1s ground state. This hyperfine splitting are due to the interaction of the magnetic moments of the proton and the electron, and causes two distinct energy levels, whose energy difference is $\Delta E = 5.9 \mu \text{eV}$. The energy difference corresponds to electromagnetic radiation, whose wavelength is 21.1 cm and frequency is ~ 1420.4MHz. The probability of spontaneous transition is very low, which is about $2.9 \times 10^{-15} s^{-1}$ and the mean half-life of the this transition is around 10 million years. Transitions could also be induced by atomic collisions and interact with ambient radiation field, especially the ubiquitous CMB photons.

2.2.2 21cm line evolution

The ratio of the number densities n_i in the two hyperfine levels of hydrogen atoms is usually described by the spin temperature. We label the two hyperfine levels with subscripts 0 and 1, among which 0 is for the lower energy level.

$$\frac{n_1}{n_0} = 3\exp\left(-\frac{h\nu_{10}}{kT_{\rm S}}\right)$$

The spin temperature is determined primarily by three rival physical processes: (1) hydrogen's absorption or emission of CMB photon; (2) collisions between hydrogen atoms and between hydrogen atoms and free electrons; (3) scattering between hydrogen atoms and Ly α and Lyman series photons.

The rate of these processes is much faster than the de-excitation time, so that the spin temperature could be given by the equilibrium of these processes [35, 41].

$$T_{\rm S}^{-1} = \frac{T_{\rm CMB}^{-1} + x_{\alpha} T_{\alpha}^{-1} + x_c T_{\rm K}^{-1}}{1 + x_{\alpha} + x_c}$$

where T_{CMB} is the CMB temperature, T_{α} is the color temperature of the Ly α radiation field and T_K is the gas kinetic temperature. x_c , x_{α} are the coupling coefficients due to atomic collisions and scattering of Ly α photons respectively.

Because the CMB photons are everywhere, the differences between brightness temperature of the 21cm signal and CMB background [10] is usually measured, which is given by

$$\delta T_b = \frac{T_S - T_{\rm CMB}}{1+z} (1 - e^{-\tau})$$

$$\approx 25 x_{\rm HI} (1 + \delta_b) \left(\frac{\Omega_b h}{0.03}\right) \left(\frac{\Omega_m}{0.3}\right)^{-\frac{1}{2}} \left(\frac{1+z}{10}\right)^{\frac{1}{2}} \times \left(1 - \frac{T_{\rm CMB}}{T_{\rm S}}\right) {\rm mK} \quad (2.2.1)$$

here, $x_{\rm HI}$ is the neutral fraction of hydrogen, and δ_b is the fractional overdensity in baryons. If the spin temperature is lower than CMB temperature, the 21cm spectral line will appear as absorption signal. Contrarily, it will be emission line. If $T_s \gg T_{\rm CMB}$, the 21cm signals become independent of the spin temperature, and directly are proportional to the gas density and ionization fraction.

The 21 cm signal evolution history is listed as follows (the detail information could be found in [93]):

- $150 \leq z \leq 1100$: The Compton scattering of the residual free electron maintain thermal coupling of the gas and the CMB, and the high gas density leads to collisional coupling so that $T_S = T_K = T_{\text{CMB}}$ and there is no detectable 21 cm signal.
- $30 \leq z \leq 150$: The gas cools adiabatically with $T_K \propto a(t)^{-2} \propto (1+z)^2$ as universe expand, and CMB cools according to $T_{\rm CMB} \propto a(t)^{-1} \propto (1+z)$. The collisional coupling cause $T_S = T_K < T_{\rm CMB}$, and an early absorption signal could be probed.

2.3. HI GALAXY SURVEYS

- $20 \leq z \leq 30$: As the expansion continues, collisional coupling becomes ineffective due to the decreasing gas density. Radiative coupling to the CMB cause $T_S = T_{\text{CMB}}$ and there are also no 21 cm signal detected.
- 15 ≤ z ≤ 20: Once the first sources formed, they emit Lyα photons which couple the spin temperature and gas. In general, the gas was not heated, and its temperature is also lower than that of CMB. So that T_S ~ T_K < T_{CMB}, and there are strong absorption signals.
- $10 \leq z \leq 15$: As sources grow up, the gas is gradually heated. In this regime, the gas temperature is much higher than CMB. The $Ly\alpha$ coupling saturates, so fluctuations in the $Ly\alpha$ no longer affect the 21 cm signal. With $T_S \sim T_K \gg T_{\text{CMB}}$, the 21 cm signals gradually become strong emission signals.
- $6 \leq z \leq 10$: Continued heating drives $T_K \gg T_{\text{CMB}}$. However, the reionization process reduces the HI regions. The 21 cm emission signals begin to decrease. And they are close to zero at the end of reionization.
- $z \leq 6$: After reionization, the residual HI regions are in the form of interstellar clouds composed of neutral atomic hydrogen, as well as helium and other elements.

Fig 2.2.1 [40] shows the 21cm signal evolution. Certainly, the exact redshift in each stage depends on the detailed cosmological model, reionization and stellar formation models.

Theoretically, 21 cm signals could be researched for different physical processes. For $T_S \gg T_{\text{CMB}}$, the 21 cm signals are proportional to gas density and ionization degree, and largely independent of the spin temperature. This is important to the study of typicality of reionization. After reionization, residual HI mainly exist inside galaxy. 21 cm surveys could be used to research large scale structure of galaxies.

2.3 HI galaxy surveys

A number of HI galaxies surveys have been carried out with large radio telescopes, such as the the Arecibo, Nancay, Parkes and Green Bank telescopes. Results of these studies and data releases could be found partly in [51, 1, 36, 42]. For example, the HI Parks All Sky Survey (HIPASS) [75] were conducted between 1997 and 2002 with CSIRO's 64-m Parkes Telescope, covering 71% of the sky, and identified more than 5000 galaxies, which include several new galaxies, and discovered the Leading Arm of the Magellanic Stream and a few gas clouds devoid of stars. The Arecibo telescope, equipped with ALFA (Arecibo L-band Feed Array), conducted four extragalactic HI surveys with different survey area and depth: the Arecibo Legacy Fast ALFA survey(ALFALFA, [44]), the Arecibo Galactic Environment Survey (AGES, [7]), the ALFA Zone of Avoidance Deep Survey (ZOA, [52]), and the Arecibo Ultra Deep Survey (AUDS, [39]). Preliminary results of the the AUDS and ZOA surveys have been reported, and plans are made for more extensive works.

ALFALFA is an on-going second generation blind extragalactic HI survey exploiting Arecibo's superior sensitivity, angular resolution and digital technology to conduct a census of the local HI



Figure 2.2.1: The 21cm signal evolve with redshift, from [40]. In this figure, the highest redshift is 25, so the first stage is not shown. The different linetype present different parameter models, so the 21cm signal evolution varies. The fiducial model is shown by the solid curve. The long-dashed curve take the model with stronger escape probability of ionizing photon, and the universe is obviously ionized much earlier, even before heating, so there is no emission signals. The short-dashed curve assume increasing the heating rate, further decreasing the strength and duration of the absorption epoch. The dot-dashed curve shows the model with both greater escape probability of ionizing photon and greater the heating rate, so the emission signals appear earlier with short duration.



Figure 2.3.1: Plot of the normalised redshift distribution of the HIPASS and ALFALFA surveys as well as the distribution of the ALFALFA sources in the three, spatially separated strips.



Figure 2.3.2: Upper left: The ALFALFA HIMF; Bottom left: histogram of the number of HI sources; Right hand: histogram of the contribution to Ω_{HI} of galaxies binned by HI mass.

sources over a cosmologically significant volume. ALFALFA have detected more than 30,000 extragalactic HI line sources out to $z \sim 0.06$. ALFALFA is detecting HI masses as low as 10^6 solar masses and as high as $10^{10.8}$ solar masses with positional accuracies typically better than 20 arcsec, allowing immediate identification of the most probable optical counterpart to each HI detection.

AGES is searching for galaxies in a number of quite different areas of the local universe. These vary from the Virgo Cluster, where there are huge numbers of galaxies very close together, to the Local Void, a region where few galaxies are found. Their science goals include: the HI mass function in different environments, the contribution of neutral gas to the baryonic mass density, the nature and link between high velocity hydrogen clouds and dwarf galaxies, the identification of gaseous tidal features as signatures of galaxy interactions and mergers, the low column density extent of galaxies, a comparison with atomic hydrogen detected by QSO absorption lines, the identification of isolated neutral gas clouds, the spatial distribution and properties of HI-selected galaxies and comparisons with numerical models of galaxy formation.

Fig. 2.3.1 displays the normalized redshift distributions of the HIPASS and ALFALFA surveys as well as the redshift distributions in the three separate ALFALFA regions from [81]. In Fig. 2.3.2 [43], the upper left panel shows the ALFALFA HIMF. Inset are the best fit parameters of a Schechter model. The lower plot is a histogram of the number of HI sources, per mass bin, on a logarithmic scale. The right hand panel is histogram of the contribution to Ω_{HI} of galaxies binned by HI mass.

2.4 **Projects for Epoch of Reionization**

Between the times of recombination and reionization, the baryonic content of the Universe, mostly hydrogen, existed in a neutral phase. Detecting the 21 cm emission from this time, all the way through to the end of reionization, has been proposed as a powerful way of studying early structure formation. This period of the Universe's history corresponds to redshifts of $z \approx 30$ to $z \approx 6 \sim 12$, implying a frequency range of $50 \sim 200$ MHz. There are a number of EOR experiments underway hoping to make headway in this area in the near future.

The Low-Frequency Array (LOFAR, [105]) is a large interconnected radio telescope using an array of simple omnidirectional antennas. It is built by ASTRON, the Netherlands Institute for Radio Astronomy and its international partners. The project is based on an interferometric array of radio telescopes using about 20,000 small antennas grouped in at least 48 stations, and observes below 250 MHz.

The Murchison Widefield Array (MWA, [57]) is a low-frequency radio telescope including 2048 dual-polarization dipole antennas and operating in the frequency range 80 - 300 MHz, which has been developed by an international collaboration, located in Western Australia. The scientific goals of the MWA include detecting neutral atomic Hydrogen emission from the cosmological EoR, to study radio transient phenomena, and so on.

The Precision Array for Probing the Epoch of Reionization (PAPER, [92]) is a low-frequency radio interferometer aimed toward making the first statistical detection of the 21 cm reionization

signal. It maps 21cm emission at high redshifts (z=7-12) to measure the power spectrum of fluctuations in the intergalactic medium introduced by the first luminous sources in the universe. PAPER has deployed a 64-antenna array in the Karoo reserve in South Africa, and a 32-antenna array near Green Bank. PAPER is currently expanding its South Africa array to 128 elements.

The 21 Centimetre Array (21CMA, [88]), is a radio telescope array designed to detect the earliest luminous objects in the universe, including the first stars, supernova explosions, and black holes. 21CMA consist an array of some ten-thousand log-periodic antennas spread over several square kilometers. The telescope is built on the high plateau of Ulastai in Xinjiang, China.

2.5 Intensity mapping and Dark Energy Observation

To observe large scale structure, the traditional way is to observe a large number of individual galaxies, and use their positions to derive the three dimensional map of galaxy distribution. This is also the method used in the optical galaxy redshift surveys. About 1% of the observable universe has been mapped by this way. Recently, a new strategy to map much larger volumes of the universe economically is proposed, which is to observe in low angular resolution and use the integrated emission of many galaxies without resolving individual ones. This is called "Intensity Mapping", and in fact very similar to the observation mode of the EOR experiments. After reionization, he remaining neutral hydrogen is stored in dense gas clouds, which is protected from ionizing UV radiation. These are predominantly hosted in galaxies, so the neutral hydrogen signal is effectively a tracer of the galaxy distribution, and also the underlying cosmic density field, with regions of higher density giving rise to a higher intensity of emission. Intensity fluctuations can therefore be used to reconstruct the power spectrum of matter fluctuations. The frequency of the emission line is redshifted by the expansion of the Universe, so by using radio receivers that cover a wide frequency band with many frequency channels, one can detect this signal as a function of redshift. This is similar in principle to a galaxy redshift survey, with the important distinction that galaxies do not need to be individually detected and measured, making intensity mapping a considerably faster method. Such cosmological surveys would be especially suitable for late time cosmological studies ($z \lesssim$ 3), in particular to constrain dark energy by measuring the geometry and expansion rate of the Universe using the Baryon Acoustic Oscillations (BAO). The growth rate of structure, useful for testing modifications to general relativity, can also be measured using Redshift Space Distortions (RSD). [87, 19, 6, 4, 95] In this scheme, the integrated radio emission of many HI clumps in cells of $\sim 10^3 \,\mathrm{Mpc}^3$ are measured without resolving of individual galaxies. Large wide-field radio telescopes, with an angular resolution of a fraction of a degree and a frequency resolution of $\lesssim\,1{\rm MHz}$ and sensitivities of $\lesssim\,1\,{\rm mK}$ per resolution element would be needed to observe the LSS, especially the BAO features.

The GBT (Fig 2.1.1), which we have described in section 2.1, is used for a single dish intensity mapping project. Another single dish project for intensity mapping is the BINGO (Baryon acoustic oscillations In Neutral Gas Observations, [11], please see Fig 2.5.1), which is a proposed 40 m single dish, whose main goal is to detect HI using the 21cm intensity mapping technique and measure the Baryon Acoustic Oscillation scale in the redshift range z = 0.12 - 0.48 (corresponding



Figure 2.5.1: BINGO.



Figure 2.5.2: CHIME.



Figure 2.5.3: HIRAX.

to a frequency range of 300 MHz). The static telescope will have two mirrors and around 50 pixels at an offset focus, operating as a transit telescope. It will scan approximately 2000 deg^2 in 1 year total observation time with an angular resolution $\theta_B \simeq 40 arcmin$ and $N_{beams} \sim 50$ feeds. Its system temperature is expected to be $T_{sys} \simeq 50 K$.

The Canadian Hydrogen Intensity Mapping Experiment (CHIME, [107], please see Fig 2.5.2) is an interferometric radio telescope in Canada, consisting of five parabolic, cylindrical reflectors and associated radio receivers and correlators. The structure is $100m \times 100m$ sensitive at 400 - 800 MHz. The telescope has no moving parts and will measure half of the sky each day as the Earth turns. CHIME will map the distribution of neutral Hydrogen over the redshift range from 0.8 to 2.5. These data will be used to produce a map of cosmic structure over the largest volume of the universe ever observed, with enough spatial resolution to discern the BAO that were imprinted on cosmic structure at early times. This three-dimensional map will enable cosmologists to better discern the cause of the Universe's accelerating expansion.

The Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX, [78], please see Fig 2.5.3) is a proposed array of 1024 6m fixed dishes that has received preliminary approval from the South African. The main goals of HIRAX include measurements of Baryon Acoustic Oscillations at redshift 0.8-2.5, searching for new pulsars and radio transients, and discovery of new neutral hydrogen absorbers.

2.6 PAON-4 and Tianlai projects

2.6.1 The PAON-4 array

The PAON-4 array is a small wide band test interferometer (L-band, 1250-1500 MHz) featuring four 5-metre diameter antenna, installed at the Nançay radio observatory in France (47°22′55.1"N, 2°11′58.7"E). PAON-4 has been designed and built within the BAORadio ¹ project in France [5]. It has a total collection area of $\sim 75 \text{ m}^2$ and 4 dual polarisation receivers. The dish pointing can be changed in declination through computer controlled electric jacks. PAON-4 can observe the sky in the declination range $12^{\circ} \leq \delta \leq 60^{\circ}$. At the end of the survey PAON-4 should be able to provide 3D sky maps (α, δ, ν) over $\sim 5000 \text{deg}^2$ of sky and $\sim 200 \text{ MHz}$ with a sensitivity of $\sim 50 \text{ mK}$ per $\sim 0.5 \times 0.5 \text{ deg}^2 \times 1 \text{ MHz}$ pixels (see section 4.1.3 for a detailed discussion of PAON-4 noise level).

The 36 visibilities (8 auto-correlations and 28 cross-correlations) are computed by the BAO-Radio electronic-acquisition system and written to disk with ~ 1 second time resolution. Tests observations with PAON-4 started in spring 2015 with the aim of evaluating the use of small dish arrays for intensity mapping and developing the calibration and map-making procedures for such instruments.

The physical diameter of the PAON-4 dishes are D = 5 m. We model the primary beam of the

¹https://groups.lal.in2p3.fr/bao21cm/



Figure 2.6.1: PAON-4 interferometer configuration

dish+feed as [15],

$$D(\gamma) \propto \frac{2 J_1[\pi(D_{\text{eff}}/\lambda) \sin \gamma]}{\pi(D_{\text{eff}}/\lambda) \sin \gamma},$$
(2.6.1)

where γ is the angle with respect to the reflector symmetry axis, $J_1(x)$ is the first order Bessel function, λ the wavelength and D_{eff} is the effective dish diameter illuminating the feed. Based on test observations, we have used an efficiency factor $\eta = 0.9$ for PAON-4, yielding an effective dish diameter $D_{\text{eff}} = \eta D = 4.5$ m. The single dish first null beam width (FNBM) is around $1.22\lambda/D_{\text{eff}} \sim 3.25^{\circ}$ at 1420MHz.

We studied a number of antenna arrangements before finally chose the PAON-4 configurations. In this thesis, we shall compare the results for the following configurations.

(a) The PAON-4 configuration. The configuration actually adopted for PAON4 is shown in figure 5.1.4. Three dishes are arranged at the vertices of an equilateral triangle with 12 m sides, one of its side is along the exact North-South line. The fourth dish is inside the triangle and on its West vertex bisector line, with its center 6 meter away from the West vertex. In addition to the auto-correlation signal, the PAON-4 configuration has 6 different baselines. There is no redundancy in baselines and visibilities, except for the auto-correlation signal.

(b) The 2 × 2 array. The four dishes are arranged into a 2 × 2 regular array. The sides of the square are aligned with the north-south (NS) and east-west (EW) directions, with antenna centres separated by d_{sep} . We use the same feed response and efficiency factor $\eta = 0.9$ as in the case of PAON-4. In addition to the autocorrelation beam, this configuration has 4 baselines, with a factor 2 redundancy for the NS and EW baselines, and the two diagonal baselines are also symmetric with respect to the meridian. (more on the mathematical implication of this property in the next section).

Actually we shall consider two 2×2 arrays:

(b1) a non-compact 2×2 array, with $d_{sep} = 14m$. This configuration is discussed first to help understand the mathematical tools introduced in next chapter, and how the characters of the array and survey would impact the reconstructed maps. However, the angular frequency (u, v) plane coverage is incomplete for this case, as it is in the spherical harmonic (ℓ, m) indices space. This

2.6. PAON-4 AND TIANLAI PROJECTS

is the reason behind the stronger mode mixing (frequency dependent beam) introduced by sparse arrays and the need to use compact arrays for intensity mapping.

(b2) A compact 2×2 array, where the dishes are separated by $d_{sep} = 7m$. This is the configuration which is compared with PAON-4, in terms of synthetised beam, transfer function and noise power spectrum.

(c) The large single dish. The third configuration compared to PAON-4 is a hypothetical single dish instrument with a diameter D = 15.5 m, covering more or less the complete PAON-4 array. Using again an $\eta = 0.9$ efficiency factor, the effective diameter for this single dish instrument would be $D_{\text{eff}} = 14 \text{ meter}$, yielding an angular resolution of $\sim 1^{\circ}$.

We consider a drift scan survey of a full east-west strip of sky, covering $\sim 27^{\circ}$ range in declination centered at the PAON-4 latitude ($\delta \simeq 47^{\circ}$) over a 6 month period. The survey will be composed of 25 constant declination scans, each shifted by 1 degree in declination. The central scan is targeted at the declination of $+47^{\circ}23'$ which is the latitude of location of PAON4. The two extreme scans are thus targeted at the declinations $+35^{\circ}23'$ and $+59^{\circ}23'$.

For comparison, we will also study the same survey for the compact 2×2 array case and the single dish case. For this exercise, the survey strategy of the 2×2 array would be assumed to be the same as the PAON-4 case. For the single dish case, given the higher angular resolution of the primary beam, the scan strategy is adapted by decreasing the declination step by a factor 3. The same sky area will be covered by 79 scans, each shifted by 0.33° .

2.6.2 The Tianlai project

The Tianlai project² is a 21cm intensity mapping experiment aimed at surveying the large scale structure and use its baryon acoustic oscillation (BAO) features to constrain dark energy models [20]. The current experiment is a pathfinder for testing the basic principles and key technologies of 21cm intensity mapping. The current pathfinder consists both a cylinder array and a dish array, located at a radio quiet site $(44^{\circ}10'47''N, 91^{\circ}43'36''E)$ in Hongliuxia, Balikun County, Xinjiang Autonomous Region in northwest China [21]. The construction of the Tianlai cylinder and dish pathfinder arrays have been completed at the end of 2015, the two arrays are now undergoing the commissioning process.

The Tianlai Dish Array

To explore the intensity mapping technique with interferometer array of dishes, the Tianlai experiment have a dish pathfinder array, with 16 dishes of 6m aperture, each equipped with a dual polarization receiver. The dish array is right next to the cylinder array (on its east side), and uses the same 288 core optical fibre cable for signal transportation. These dishes are also equipped with electronically controlled motor drives in the altitude-azimuth mount, which allow us to steer the dishes to point to almost any desirable directions above the horizon. These drives were not designed for tracking celestial targets with high precision. Instead, the regular observation mode we envisage for the dish array is to point the dishes along the meridian with a common elevation

²http://tianlai.bao.ac.cn



Figure 2.6.2: The regular or square array (Left) and the circular array (right) configurations for the 16 dishes.

(declination), and make drift scan observations. However, the azimuth drive allows more flexibility during commissioning and for test and calibration runs. It enables also the dish array to explore other scientific programs. The Tianlai array receivers are designed to have broad band response, except for a replaceable bandpass filter of 100 MHz bandwidth. The central frequency of the band is tuneable within the range of 500-1500 MHz. The visibilities (32 autocorrelations and 448 cross-correlation) for dish arrays are computed by the data-acquisition (DAQ) system and saved in hard drives.

The Tianlai dishes are D = 6m dishes. For the model calculations below, we shall assume the same primary beam model as in the PAON-4 case except that the aperture size is scaled, i.e. we will assume that $\eta = 0.9$, so $D_{\text{eff}} = 5.4$ m, and the primary beam (FNBW) for each dish is 2.73° at 1420 MHz and 3.1° at 1250 MHz.

Although we studied several antennae layouts for the 16-dish array, we will focus here on the discussion and comparison of the following two configurations (see Fig.2.6.2):

- (a) Regular Array. The antennas are positioned on the grid points of a 4×4 square, with the sides of the square aligned with the EW and NS directions, and separation for the nearest neighbour grid points to be $d_{sep} = 8.8$ m (in sec.4.2 we shall discuss the choice of d_{sep} in more details). The Regular Array have large number of redundant baselines, the total number of independent baselines is only 24 for each declination pointing.
- (b) Circular Array. This is the configuration we have adopted for the current Tianlai dish pathfinder array. One antenna is positioned at the centre, the remaining 15 antennas are arranged in two concentric circles around it. It is well known that the baselines of the circular array configurations are quite independent and have an overall wider coverage on the angular frequency (u, v) plane [102].. We have studied a number of circular arrangements,

Antenna	radius (m)	azimuth angle	x (EW) (m)	y (NS) (m)
A1	0	-	0	0
A2	8.8	0°	8.8	0
A3	8.8	60°	4.4	7.62
A4	8.8	120°	-4.4	7.62
A5	8.8	180°	-8.8	0
A6	8.8	240°	-4.4	-7.62
A7	8.8	300°	4.4	-7.62
A8	17.6	5°	17.53	1.53
A9	17.6	45°	12.45	12.45
A10	17.6	85°	1.53	17.53
A11	17.6	125°	-10.09	14.42
A12	17.6	165°	-17	4.56
A13	17.6	205°	-15.95	-7.44
A14	17.6	245°	-7.44	-15.95
A15	17.6	285°	4.56	-17
A16	17.6	325°	14.42	-10.09

Table 2.1: Coordinates of the circular array antenna.

with or without centre antenna, with one or two concentric rings, with different alignment between the inner and outer rings. From the investigations, we found that there is no significant difference in terms of map reconstruction performance between the different circular configurations that were studied. The final configuration we chose has 6 and 9 dishes in the inner and outer rings respectively, with radius given by the minimal separation of $d_{\rm sep} = 8.8 {\rm m}$ and $2 d_{\rm sep}$ respectively. The inner ring is symmetric with respect to the NS direction, i.e. if we place Antenna A1 at position (0,0), then antennae A2, ..., A7 are placed at radius of $d_{sep} = 8.8 {\rm m}$ and azimuth angles 0°, 60°, 120°, 180°, 240°, 300° respectively. The positions of the antennas on the outer ring are slightly rotated to accommodate the local terrain, and are located at azimuth angles 5°, 45°, 85°, 125°, 165°, 205°, 245°, 285°, 325° with respect to the N-S axis. The circular array has a large number of independent baselines, the total number of different baselines is 108, to be compared to 120, which is the maximum number of baselines, without any redundancies ($120 = 16 \times 15/2$). The polar and Cartesian coordinates of the antenna positions for this circular configuration are listed in table 2.1.

In fact, the Tianlai dishes are lightweight and the mount is detachable, so that in the future, we can also move the dishes to new configurations if this is deemed necessary or desirable. At present, we will carry observations with the current configuration, i.e. the circular array (b), but we shall also discuss also the computation results for the regular array (a).

We have considered the survey of a $\sim 33^{\circ}$ wide band in declination centered at the latitude of location of the Tianlai array, i.e. $\delta \sim 44^{\circ}10'$, slightly larger than the PAON-4 survey area described above. The survey would be composed of a total of 31 constant declination scans, each



Figure 2.6.3: The Tianlai Dish and Cylinder Array.

shifted by 1 degree in declination, with the two extreme scans targeted at the declinations $+29^{\circ}10'$ and $+59^{\circ}10'$.

At the end of such a survey, Tianlai should be able to provide 3D sky maps (α, δ, ν) over a area of more than 5000 deg^2 of sky and over $\sim 100 \text{ MHz}$ frequency band, with a sensitivity of $\sim 50 \text{ mK}$ per $\sim 0.25 \times 0.25 \text{ deg}^2 \times 1 \text{ MHz}$ pixels. Notice that the sensitivity of the Tianlai dish array is given for an angular resolution twice better than the PAON-4 case. Detailed discussion of Tianlai dish array sensitivity and noise power spectrum can be found in section 4.2.1. We consider also a survey of the polar cap area, for which a noise level ~ 3 times better ($\sim 17 \text{ mK}$) can be reached, over a sky area ten times smaller (see section 4.2.2).

The Tianlai Cylinder Array

The Tianlai cylinder pathfinder array has three adjacent cylindrical reflectors oriented in the North-South direction, each cylinder is 15 m wide and 40 m long. At present the cylinders are equipped with a total of 96 dual polarization receivers which do not cover the full length of the cylinders. In the future, the pathfiner instrument may be upgraded by simply adding more feed units and the associated electronics. The longer term plan is to expand the Tianlai array to full scale once the principle of intensity mapping is proven to work. The full scale Tianlai cylinder array would have a collecting area of $\sim (10^4 \text{ m})^2$, and $\sim 10^3$ receiver units. A forecast for its capability in measuring dark energy and constrain primordial non-Gaussianity can be found in [112]. In addition to the redshifted 21 cm intensity mapping observation, such surveys may also be used for other observations, such as 21cm absorber[114], fast radio burst (FRB) [73, 23], and electromagnetic counter part of gravitational wave events [34]. Fig 2.6.3 shows the Tianlai Dish and Cylinder array taken at last winter.

Chapter 3

Map making for transit interferometers

In this chapter we mainly discuss the specificities of making sky maps from visibilities obtained from a transit type interferometer, and the mathematical basis of the method. The outlines is as follows: first we present the definition of visibilities in section 3.1. Then, in section 3.2, we present the classical interferometry with rotation synthesis and sparse sampling of the (u, v) plane. In section A, we discuss the one-dimension reconstruction and the corresponding results. In section 3.3 and 3.4, we describe the formalism in the planar geometry (u, v) with u-mode decoupling and the spherical geometry (ℓ, m) with m-mode decoupling from transit interferometers respectively. We present the steps of solving the system in section 3.5. We also defined various indicators in section 3.6.

3.1 Visibilities

In this section we discuss the specificities of making sky maps from visibilities obtained from a transit type interferometer, and the mathematical basis of the method. These issues for transit interferometers and the separation of the inversion problem into independent sub-systems using m-mode decomposition in the spherical harmonic basis have already been discussed in [96]. However, the formalism described here as well as the corresponding software tools¹ have been developed independently, initially in flat sky approximation, and subsequently extended to spherical geometry following [96].

Throughout this section, we shall assume that individual antenna/feed responses, the array geometry and pointing directions are perfectly known. Moreover we will consider unpolarized sky emission, with brightness or temperature in the direction \vec{n} given by $I(\vec{n}) = E^* E = |E|^2$ where $E(\vec{n})$ is the complex scalar emission amplitude in a narrow, nearly monochromatic, frequency band. The method can be extended easily to the case of polarized sky, as shown by [97]. This is briefly discussed in section 3.7.

¹The code is written in C++ and uses the SOPHYA class library (http://www.sophya.org). The GIT repository will be available from https://gitlab.in2p3.fr/SCosmoTools/JSkyMap.

The Visibility $\mathcal{V}_{ij} \equiv \langle s_i^* \times s_j \rangle$ is the short time average cross correlation of output voltage from a pair of antennae or feeds s_i, s_j , located at positions $\vec{r_i}, \vec{r_j}$ with $\Delta \vec{r_{ij}} = \vec{r_j} - \vec{r_i}$:

$$s_i = \iint E(\vec{\hat{n}}) D_i(\vec{\hat{n}}) e^{i\vec{k}\cdot\vec{r_i}} d\vec{\hat{n}}$$
(3.1.1)

$$\mathcal{V}_{ij} = \iint I(\vec{\hat{n}}) D_i^*(\vec{\hat{n}}) D_j(\vec{\hat{n}}) e^{\vec{k} \cdot \Delta r_{ij}} d\vec{\hat{n}}$$
(3.1.2)

where D_i , D_j denotes the complex response function of each feed, $\vec{k} = -\frac{i2\pi\nu}{c}\vec{n}$ is electromagnetic wave vector at the observation frequency ν and c the speed of light. For arrays with identical feeds pointed to the same sky direction, $D_i(\vec{n}) = D_j(\vec{n}) = D(\vec{n})$, and the visibility expression reduces to

$$\mathcal{V}_{ij} = \iint I(\vec{\hat{n}}) L(\vec{\hat{n}}) e^{\vec{k} \cdot \Delta r_{ij}} d\vec{\hat{n}}$$
(3.1.3)

where $L(\vec{\hat{n}}) = D^* D$ is the antenna primary beam or response in intensity.

3.2 Classical radio interferometry

In what we refer here as *classical radio interferometry*, in the sense that it is familiar to the majority of radio astronomers, a set of identical antennae are used to observe a small region of sky, usually to obtain a high resolution image of a source. During the observation period all the antennae track the source, compensating the Earth rotation. The source intensity $I(\vec{n}, t)$ and beam response $L(\vec{n}, t)$ generally varies with time. However, even in the case of constant sources and constant telescope primary beams, the baseline delay $\vec{k} \cdot \Delta \vec{r}(t)$ would still vary with time, due to the rotation of the baseline generated by the rotation of the Earth with respect to the inertial frame of space, as shown in the variation of celestial coordinates of the baseline direction.

For observations with small field of view, it is possible to use the flat sky approximation in the vicinity of the source. For a coplanar array and using the small angle approximation (omitting the so called *w*-term), the visibility is given by

$$\mathcal{V}(u_0, v_0) = \iint I(\xi, \eta) \, L(\xi, \eta) \, e^{2i\pi(\xi u_0 + \eta v_0)} \, d\xi \, d\eta \tag{3.2.1}$$

where $(u_0, v_0) = (\Delta x / \lambda, \Delta y / \lambda)$ is the coordinates of the baseline vector in wavelength units, and ξ, η denotes the direction cosines of the baseline vector with respect to the reference point. The visibility in this approximation is simply the Fourier transform of the sky *seen* by a single antenna $I(\xi, \eta) \times L(\xi, \eta)$ for the angular frequencies (u_0, v_0) . Given the number of available baselines in a real array, and that the baselines of such an array are usually large compared to the antenna size, the (u, v) frequency plane is only sparsely sampled at any moment. However, each baseline changes as the antennae follow the source direction on the sky, the (u_0, v_0) follows an arcshaped track in the (u, v) plane, enhancing greatly the frequency plane sampling. It is possible to obtain a local sky map (dirty map) around the targeted position using an inverse Fourier transform. Additional processing is required to correct and compensate for the partial coverage of the angular frequencies. Iterative deconvolution algorithms, e.g. CLEAN [54, 22], are applied to recover the map of the sky [94]. Map of a large area of sky can be obtained by mosaicking of small areas[63, 74]. However, if the field of view is large, the *w*-term can not be neglected. A number of formalisms have been developed to deal with this, such as faceting [27], 3D Fourier transform [84], w-projection [26], A-projection[100], w-stacking [80], etc. Other refinement of the CLEAN method have also been developed, such as the the *software holography* [77] which can deal with direction-dependent beam effects in large field of view interferometer arrays. Its application to the analysis of MWA observations can be found in [98].

3.3 Non-tracking transit interferometers in the planar geometry

Before discussing the sky map reconstruction from visibilities in the true sky spherical geometry, we present first the formalism in the case of planar geometry. The spherical geometry is discussed in the next section.

As was explained earlier, a visibility \mathcal{V}_{ij} corresponds to cross correlation information from a pair of antennas separated by a baseline $\Delta \vec{r}_{ij} = (\Delta x, \Delta y, \Delta z)$ and the auto-correlation information, obtained from a single antenna is the visibility for a null $(\Delta \vec{r}_{ii} = (0, 0, 0))$ baseline. In the case of transit interferometers, visibilities will be recorded as a function of time or right ascension (α) as the earth rotation, pointing the instrument to different right ascension directions on the sky. In the case of dish instruments with pointing capabilities, the instrument field of view can be directed to a different declination, angle δ or β . For dish array transit instrument, we will assume that all antennae are pointing toward a common direction during a full east-west scan, the instrument observing a 360° stripe of sky along a common declination δ_k or β_k . The instrument scan strategy is thus defined by the set of K observed declinations (δ_k or β_k with k = 1...K), and the time, in days spent on each declination.

In the case of simplified planar geometry, we shall assume that the sky directions, defined by the two cartesian like coordinate angles (α, β) are defined over a rectangular angular area, $0 \le \alpha < 2\pi$ and $-\beta_{\max} \le \beta \le \beta_{\max}$. α correspond to the right ascension angle or φ in spherical coordinates $\vec{n} = (\theta, \phi)$, and β is an offset angle with respect to the central observed declination. The associated Fourier angular frequencies corresponding to the (α, β) coordinate system are denoted (u, v).

For simulations, we assume a known input sky map from which we compute visibilities. Instrument simulations and sky map reconstruction use both the angular plane (α, β) and the Fourier plane $(u, v) (I(\alpha, \beta) \rightarrow \mathcal{I}(u, v))$ where u, v are the Fourier angular frequencies associated to (α, β) .

The planar angular plane geometry would be a reasonable approximation for the case of instruments, observing a not too wide band of the sky ($\leq 20^{\circ}$) in equatorial or mid-latitude area of sky, far from the polar regions.

3.3.1 Convolution theorem for reconstruction

Starting from the general definition of visibilities given in section 3.1, we can write the expression of the visibility for a pair of antenna located in the same horizontal plane $\Delta z = 0$, with the baseline given by $(\Delta x, \Delta y, 0)$. It should be reminded that this expression of visibility correspond to a monochromatic, single frequency incoming wave, corresponding to wavelength λ :

$$\mathcal{V}_{ij} = \iint \sin\theta \, d\theta \, d\varphi \, I(\theta, \varphi) L(\theta, \varphi) \exp\left[2i\pi \sin\theta (\cos\varphi \frac{\Delta x}{\lambda} + \sin\varphi \frac{\Delta y}{\lambda})\right]$$
(3.3.1)

If we write this expression in planar, pseudo-cartesian coordinate system (α, β) and assuming that the antennae have a narrow field of view, so that the beam $L(\theta, \varphi)$ is non zero only for $\theta \ll 1$, or for $\alpha \ll 1$ and $\beta \ll 1$, so that that the small angle approximation can be used $(\sin \alpha \simeq \alpha, \sin \beta \simeq \beta)$:

$$\mathcal{V}_{ij} = \iint d\alpha \, d\beta \, I(\alpha, \beta) L(\alpha, \beta) \exp\left[2i\pi \left(\alpha \frac{\Delta x}{\lambda} + \beta \frac{\Delta y}{\lambda}\right)\right]$$
(3.3.2)

Using the properties of the Fourier transform, we can write also the expression of the visibility in the Fourier (u, v) plane:

$$\mathcal{V}_{ij} = \iint du \, dv \, \mathcal{I}(u, v) \, \mathcal{L}\left(u - \frac{\Delta x}{\lambda}, v - \frac{\Delta y}{\lambda}\right) \tag{3.3.3}$$

This is the mathematical expression of the well known result in interferometry, that a visibility corresponding to a baseline $(\Delta x, \Delta y)$ is a measurement of the sky Fourier modes $\mathcal{I}(u, v)$ centered around the angular frequency $(u_0, v_0) = \left(\frac{\Delta x}{\lambda}, \frac{\Delta y}{\lambda}\right)$, weighted by the beam angular frequency $\mathcal{L}(u, v)$.

So, if we have a set of visibility measurements of a given direction of the sky, an estimate of the sky convoluted by the primary antenna beam in this direction can be obtained by performing an inverse Fourier transform of the set of visibilities.

$$\hat{I}(\alpha,\beta) = \mathscr{F}(\mathcal{I}(u,v)\mathcal{L}(u,v)) = \mathscr{F}[\mathcal{V}_{ij}]$$
(3.3.4)

3.3.2 East-west transit observations

For transit instrument, we obtain a set of visibility $\mathcal{V}_{ij}(\alpha_p)$ as a function of the right ascension angle α_p , as the earth rotates and instrument observation direction scans the sky along constant declination directions, pointing toward the direction defined by α_p . In the expression of the visibility, the antennae primary beam response $L(\alpha, \beta)$ should be replaced by $L(\alpha - \alpha_p, \beta)$. Using the translation/phase shift property of the Fourier transform, and neglecting the sky drift during the integration time to obtain one visibility measurement, we can write the expression of the visibility expression in the Fourier plane as a function of α_p :

$$\mathcal{V}_{ij}(\alpha_p) = \iint du \, dv \, \mathcal{I}(u, v) \, \mathcal{L}\left(u - \frac{\Delta x}{\lambda}, v - \frac{\Delta y}{\lambda}\right) \, \exp\left(2i\pi\alpha_p u\right) \tag{3.3.5}$$

Assuming that the planar geometry correspond to a periodic sky along the two angular coordinates direction, the Fourier integral above should be replaced by a Fourier sum. The hypothesis of periodic sky is true along the right ascension α :

$$\mathcal{V}_{ij}(\alpha_p) = \sum_{u} \exp\left(-2i\pi\alpha_p u\right) \sum_{v} \mathcal{I}(u,v) \mathcal{L}\left(u - \frac{\Delta x}{\lambda}, u - \frac{\Delta y}{\lambda}\right)$$
(3.3.6)

Considering the set of measurements $\mathcal{V}_{ij}(\alpha_p)$ for all α_p values and taking into account the term $\exp(-2i\pi\alpha_p u)$, we identify the expression of the Fourier sum for the function $\mathcal{V}_{ij}(\alpha_p)$. By performing a Fourier transform over the set of $\mathcal{V}_{ij}(\alpha_p)$ measurements from a full east-west scan,

Fourier transform/FFT :
$$\mathcal{V}_{ij}(\alpha_p) \longrightarrow \widetilde{\mathcal{V}_{ij}}(u)$$

we obtain the values of sky modes weighted be the beam for each u-mode as a set of independent linear equations:

for each
$$u$$
: $\widetilde{\mathcal{V}_{ij}}(u) = \sum_{v} \mathcal{I}(u, v) \mathcal{L}\left(u - \frac{\Delta x}{\lambda}, v - \frac{\Delta y}{\lambda}\right)$

3.3.3 Instrument pointing in the north-south plane

As mentioned above, for instruments made of steerable dishes, or at least movable along the elevation axis (north-south plane), we can perform east-west transit along different constant declination scans, identified by the declination scan angle δ_k or β_k . One has to use a specific beam for each scan k, $L(\alpha, \beta)$ being replaced by $L_k(\alpha, \beta)$. If we neglect the change of baselines for $\beta_k \neq 0$, the tilted beam $L_k(\alpha, \beta)$ beam can be approximated by $L(\alpha, \beta - \beta_k)$, leading to a phase term $\exp(2i\pi\beta_k v)$ for the visibility expression in the Fourier plane. To solve for the sky, the set of visibilities from different baselines and different β_k should be gathered together for each u - mode.

Considering the above approximation for the tilted beam, the visibility term for each u-mode can be written as:

$$\widetilde{V_{ij}}(u,\beta_k) = \sum_{v} \exp(i2\pi v\beta_k) \mathcal{I}(u,v) \mathcal{L}(u - \frac{\Delta x}{\lambda}, v - \frac{\Delta y}{\lambda})$$
(3.3.7)

The matrix form is as follows,

$$\left(\begin{array}{c} \cdots \\ \widetilde{V_{ij}}(u,\beta_k) \\ \cdots \end{array}\right) = \mathbf{L} \times \left(\begin{array}{c} \cdots \\ \mathcal{I}(u,v) \\ \cdots \end{array}\right)$$

Now, the matrix $\mathbf{L} = [\mathcal{L}(u - \frac{\Delta x}{\lambda}, v - \frac{\Delta y}{\lambda}) \exp(i2\pi v\beta_k)]$ depend on instrument configuration and scan strategy.

In order to compute the sky map $I(\alpha, \beta)$, we have to compute the sky Fourier modes $\mathcal{I}(u, v0)$ by solving the above linear system. It should be noted however that the matrix L is in general non

invertible, except in the case of a full scan in the north-south direction, if we limit the sky modes in the Fourier plane up (u_{max}, v_{max}) to which the instrument is sensitive.

That's why we use the pseudo-inverse method see section 3.5 below. However, the size of the L matrix is quite large, and a direct computation would be difficult, or even impossible for large arrays. So we should make control them before computing directly. If we assume for example a 4×4 regular dish array and an single β_0 scan, we will have 24 baselines. For a sky map size (1500, 300) pixels covering (250°, 50°), the $\widetilde{V}_{ij}(u, \beta_0)$ vector size will be $\sim (2 \times 10^4)$, the L matrix size $\sim (2 \times 10^4, 2 \times 10^5)$ and $\mathcal{I}(u, v)$ size $\sim (2 \times 10^5)$. Even for this modest array size with a single scan, the matrices are already quite large for the system to be solved by directly. We can indeed use the u - mode decomposition to reduce the size of the systems to be solved, without any loss of information.

In addition, we keep only one $\widetilde{V_{ij}}(u, \beta_k)$ for a set of antennas pairs with the same baseline which could reduce the corresponding visibility noise accordingly as would be discussed later. Then, taking advantage of the case of full East-West (24 hours) sky coverage, we perform Fourier transform on time or right ascension dependent visibilities $V_{ij}(\alpha = \omega t)$ to compute m - mode visibilities $\widetilde{V_{ij}}(u)$. This allows us to split the large initial problem could into a set of smaller independent linear problems which numbers are equal to number of u modes $N = N_u$.

We illustrate those processes in the matrix form as follows,

$$\left(\begin{array}{c} \cdots\\ \widetilde{V_{ij}}(\beta_k)\\ \cdots \end{array}\right)_{u_i} = \mathbf{L}_{u_i} \times \left(\begin{array}{c} \cdots\\ \mathcal{I}(v)\\ \cdots \end{array}\right)_{u_i}$$

Now, the $\widetilde{V_{ij}}(\beta_0)_{u_i}$ vector size will be ~ (24), the \mathbf{L}_{u_i} matrix size ~ (24, 300) and $\mathcal{I}(u_i, v)$ size ~ (300), for 4 × 4 regular dish array and an single β_0 scan with same sky map size as above.

For each u, the matrices in detail is given by

$$\begin{pmatrix} V_{ij}(\beta_1) \\ V_{ij}(\beta_2) \\ \cdots \\ V_{ij}(\beta_m) \end{pmatrix}_{u_i} = \begin{pmatrix} \mathcal{L}_{ij}(v_1)e^{i2\pi\beta_1v_1} & \mathcal{L}_{ij}(v_2)e^{i2\pi\beta_1v_2} \cdots \mathcal{L}_{ij}(v_n)e^{i2\pi\beta_1v_n} \\ \mathcal{L}_{ij}(v_1)e^{i2\pi\beta_2v_1} & \mathcal{L}_{ij}(v_2)e^{i2\pi\beta_2v_2} \cdots \mathcal{L}_{ij}(v_n)e^{i2\pi\beta_2v_n} \\ \cdots \\ \mathcal{L}_{ij}(v_1)e^{i2\pi\beta_mv_1} & \mathcal{L}_{ij}(v_2)e^{i2\pi\beta_mv_2} \cdots \mathcal{L}_{ij}(v_n)e^{i2\pi\beta_mv_n} \end{pmatrix}_{u_i} \times \begin{pmatrix} \mathcal{I}(v_1) \\ \mathcal{I}(v_2) \\ \cdots \\ \mathcal{I}(v_n) \end{pmatrix}_{u_i}$$

And of course, the matrices will only have single pointing β_0 for cylinder case.

We have developed an algorithm using the pseudo inverse method of SVD to invert \mathbf{L}_{u_i} matrices which we could also apply to 2 dimension case. We use symbol $\widehat{\mathcal{I}}(u, v)$ as the observed or calculated sky map, then, ignoring the noise contribution for the sake of clarity, the estimated or computed sky Fourier modes can be written as a function of u - mode visibilities:

$$\left(\begin{array}{c} \dots \\ \widehat{\mathcal{I}}(v) \\ \dots \end{array}\right)_{u_i} = \mathbf{L}_{u_i}^{-1} \times \left(\begin{array}{c} \dots \\ \widetilde{V_{ij}}(\beta_k) \\ \dots \end{array}\right)_{u_i}$$

3.4. NON-TRACKING TRANSIT INTERFEROMETERS IN THE SPHERICAL GEOMETRY57



Recmap in rectangular geometry with full y scan, and 3dish



Figure 3.3.1: The reconstructed map in the planar geometry.

We compute vector $\{\widehat{\mathcal{I}}(v)\}_{u_i}$ for all of u mode which is independent with each other. Then we could assemble them together to get matrix $\widehat{\mathcal{I}}(u, v)$. Finally we could do an inverse Fourier transform to get the estimate of the sky map in angular domain $\widehat{I}(\alpha, \beta)$.

Fig. 3.3.1 shows the reconstructed map from a fiducial input map representing Bird's Nest Olympic stadium in Beijing (top panel)². We computed visibilities from a L-shaped 3 dish array, D=5m in diamater, using the autocorrelation signal and two cross-correlations, corresponding to the North-South baseline $d_{ns} = 5$ m, and a East-West baseline $d_{ew} = 5$ m. The input map size is (634,190), and its angular resolution ~ 0.57°. We have full β scan, and the number of β scan is about 190. The reconstructed map using the above u - mode decomposition is shown in the bottom panel.

3.4 Non-tracking transit interferometers in the spherical geometry

For interferometers operating in the transit mode, the baselines do not change with time in the ground coordinates, at least during an observation period spanning a sidereal day, but the visibilities recorded as a function of time correspond to observation of different parts of the sky. We will work in the equatorial coordinates, with right ascension α and declination δ . We also introduce the spherical coordinates (θ, φ), with $\theta = \pi/2 - \delta$ and $\varphi = \alpha$. The earth rotation makes the beams

²Beijing National Stadium https://en.wikipedia.org/wiki/Beijing.National_Stadium

time dependent and the effect corresponds to a shift of the beams $L_{ij}(\vec{n})$ by an offset angle $\alpha_p(t)$ along the right ascension direction:

$$\alpha_p(t) = \omega_e t \qquad t : \text{sidereal time}$$
 (3.4.1)

$$L_{ij}(\hat{n},t) = L_{ij}((\theta,\varphi),t) = L_{ij}(\theta,\varphi - \alpha_p(t))$$
(3.4.2)

where ω_e is the Earth angular rotation rate $(2\pi/23 \text{ hours } 56 \text{ minutes } 4 \text{ seconds})$.

In the celestial coordinates, the visibility of a baseline at any given time corresponds to the convolution of sky with the beam pattern for this baseline $L_{ij}(\vec{n}, t)$. Indeed, using discrete time and discrete angular directions on the sky and using [] to denote vectors, we can write the vector of visibilities for all baselines and for all observation times as a function of the unknown discretized sky $[I(\vec{n})]$ and the noise vector:

$$\left[\mathcal{V}_{ij}(t)\right] = \mathbf{L}_{ij}(t) \times \left[I(\vec{n})\right] + \left[n_{ij}(t)\right]$$
(3.4.3)

The beam matrix L encodes both the array response and the sky scan strategy,

$$\mathbf{L}_{ij}(t) \sim D_i^*(\vec{n}, t) D_j(\vec{n}, t) e^{i\vec{k}\cdot\Delta\vec{r}_{ij}}$$

Considering the visibilities for a single narrow frequency band, the L matrix has N_{pixel} columns, and N_t (number of time sample) $\times N_b$ (number of baseline) rows. N_{pixel} corresponds to the total number of pixels in sky. If far side lobes can be neglected, one can use a partial map of the sky, limited to the observed region, hence decreasing the N_{pixel} and the L matrix size, The determination of the unknown sky $I(\vec{n})$ is then the solution of a standard inverse linear problem. There are however two difficulties for solving the above equation. First, the dimension of the matrix L is very large, typically $10^5 \times 10^6$ for the current generation of experiments, and can reach $10^6 \times 10^7$ for the next generation experiments which are being planned, if the intensity mapping method proves successful. Indeed, the sky brightness unknown vector will have a size of 10^5 for a resolution of a fraction of a degree, determining the number of columns of the L matrix. CHIME and Tianlai will have $\sim 10^3$ baselines and $\gtrsim 10^3$ time samples over about 24 hours of observations, leading to $\gtrsim 10^6$ rows for L. Secondly, for many array configurations and sky observation strategies, the linear problem is under-determined and a solution can not be unambiguously determined.

As already shown by [96], by working in the space of spherical harmonic coefficients and taking advantage of the full circle transit observation strategy foreseen for the intensity mapping experiments, the problem can be reduced to a much smaller set of independent linear systems, one for each spherical *m*-mode. The beam pattern associated to each visibility measurement (pair of antenna) is a complex function ($L_{ij}(\vec{n},t) \in \mathbb{C}$), and the baseline enters its expression through the phase factor. Its time dependence for transit observations is discussed below. Expanding in

3.4. NON-TRACKING TRANSIT INTERFEROMETERS IN THE SPHERICAL GEOMETRY59

spherical harmonics and omitting the time dependence of the beam,

$$I(\vec{n}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \mathcal{I}_{\ell,m} Y_{\ell,m}(\vec{n})$$
(3.4.4)

$$L_{ij}(\vec{n}) = D_i^*(\vec{n}) D_j(\vec{n}) e^{i\vec{k}\Delta r_{ij}^2}$$
(3.4.5)

$$= \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \mathcal{L}_{ij}(\ell, m) Y_{\ell,m}(\vec{\hat{n}})$$
(3.4.6)

The spherical harmonics $Y_{\ell,m}$ are defined through the Legendre associated polynomials $P_{\ell}^{m}(\vec{\hat{n}})$ for which we use the normalisation convention of [31],

$$Y_{\ell,m}(\vec{\hat{n}}) = \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_{\ell}^{m}(\cos\theta) e^{im\varphi}$$

The sky brightness temperature is real, for which the spherical harmonic coefficients satisfy the following symmetry relations,

$$I(\tilde{\hat{n}}) \in \mathbb{R} \to I^* = I \longrightarrow \mathcal{I}(\ell, -m) = (-1)^m \mathcal{I}^*(\ell, m).$$

Given the orthogonality of Spherical Harmonics when integrated over the whole sky, we can express the visibility for a given time t as a sum over the spherical harmonics coefficients. Expanding both $I(\vec{n})$ and $L_{ij}(\vec{n},t)$ in spherical harmonics, using the orthogonality and the above symmetry relation, we obtain

$$\mathcal{V}_{ij}(t) = \iint_{+\infty} I(\vec{\hat{n}}) L_{ij}(\vec{\hat{n}}, t) d\vec{\hat{n}}$$
(3.4.7)

$$= \sum_{m=-\infty}^{+\infty} \sum_{\ell=|m|}^{+\infty} (-1)^m \mathcal{I}(\ell,m) \mathcal{L}_{ij}(\ell,-m,t)$$
(3.4.8)

Notice that we have exchanged the order of the two sums, over ℓ and m. The spherical harmonics coefficients of the rotated/shifted beams can be written as:

$$\mathcal{L}_{ij}(\ell, m, t) = \mathcal{L}_{ij}^{0}(\ell, m) e^{-im\alpha_{p}(t)}$$
(3.4.9)

where $\mathcal{L}_{ij}^0(\ell, m)$ denotes the beam spherical harmonics coefficients for the reference (t = 0) pointing, i.e the antenna axis pointing toward $\alpha = 0$ right ascension. In the following, we will omit the ⁰ superscript in the beam coefficients. $\mathcal{L}_{ij}(\ell, m)$ denotes simply the beam for the reference right ascension $\alpha_p = 0$. The recorded visibilities as a function of right ascension α_p can then be expressed as:

$$\mathcal{V}_{ij}(\alpha_p) = \sum_{m=-\infty}^{+\infty} \sum_{\ell=|m|}^{+\infty} (-1)^m \mathcal{I}(\ell,m) \mathcal{L}_{ij}(\ell,-m) e^{im\alpha_p}$$
(3.4.10)

We recognise the expression as a Fourier transform for the periodic function $\mathcal{V}_{ij}(\alpha_p)$; as the feed response vanishes for large enough ℓ ($\mathcal{L}_{ij}(\ell, m) \to 0$ for $\ell > \ell_{\max}$), we can write the following relation satisfied by the visibility Fourier coefficients $\tilde{\mathcal{V}}_{ij}(m)$, computed from a set a regularly time sampled visibility measurements.

$$\tilde{\mathcal{V}}_{ij}(m) = \sum_{\ell=|m|}^{+\ell_{\max}} (-1)^m \mathcal{I}(\ell, m) \mathcal{L}_{ij}(\ell, -m)$$
(3.4.11)

The *m*-mode of the visibility for both positive and negative m $(\pm m)$ is given by sky spherical harmonics coefficients of the same *m*,

$$\tilde{\mathcal{V}}_{ij}(m) = \sum_{\ell=|m|}^{+\ell_{\max}} (-1)^m \mathcal{I}(\ell,m) \mathcal{L}_{ij}(\ell,-m)$$
(3.4.12)

$$\tilde{\mathcal{V}}_{ij}^{*}(-m) = \sum_{\ell=|m|}^{+\ell_{\max}} \mathcal{I}(\ell,m) \mathcal{L}_{ij}^{*}(\ell,m)$$
(3.4.13)

The full linear system of Eq. (3.4.3) can thus be decomposed into a set of much smaller $(10^3 \times 10^3)$ independent linear system, one for each m, with $m_{\text{max}} = \ell_{\text{max}}$. The beam matrix L has indeed a block diagonal structure in the harmonic space, which is schematically shown in Fig.3.4.1. Grouping all array baselines together in a vector, and taking into account the noise contribution, the visibility measurement equation in the Fourier space can be written in matrix form as:

$$\left[\tilde{\mathcal{V}}\right]_{m} = \mathbf{L}_{m} \times \left[\mathcal{I}(\ell)\right]_{m} + \left[\tilde{n}\right]_{m}$$
(3.4.14)

The sky spherical harmonics coefficient for a given m and for $m \le \ell \le \ell_{\max}$ are grouped in the sky vector $[\mathcal{I}(\ell)]_m$. We will consider only positive m values ($0 \le m \le \ell_{\max}$) for the linear systems defined above, the two visibility measurements for $\pm m$ of equations 3.4.12 and 3.4.13 will be represented by two rows of the matrix \mathbf{L}_m . This matrix will thus have ℓ_{\max} columns and $2 \times n_{\text{beams}}$ rows. The total number of beams n_{beams} will be more precisely defined in the next paragraph. The $[\tilde{n}]_m$ represent the noise contribution vector to the m-mode visibilities, corresponding to the Fourier transform of time domain noise.

For dish arrays, the instantaneous field of view is a small fraction of the whole sky, and a circular strip of sky along one of the latitude line can be obtained by carrying out transit observation for 23 hours 56 minutes 4 seconds continuously. By changing the elevation angle of the dish pointing, strips with different central declination can be obtained. For dish arrays, the effective number of beams would be equal to the number of different baselines times the number of constant elevation scans,

$$n_{\text{beams}} = N_b \times n_{\delta_n}$$

The beam for an antennae pair ij making constant elevation drift scan observation with declination



Figure 3.4.1: Schematic representation of m-mode decomposition in spherical harmonics space of the skymap reconstruction from visibilities.

 δ_p is

$$L_{ij}^{\delta_{p}} = D_{i}^{\delta_{p}}(\hat{\vec{n}}) D_{j}^{\delta_{p}*}(\hat{\vec{n}}) e^{i\vec{k}\cdot\Delta\vec{r}_{ij}}$$
(3.4.15)

$$= \sum_{lm} \mathcal{L}_{ij}^{\delta_p}(\ell, m) Y_{\ell m}(\hat{\vec{n}})$$
(3.4.16)

3.5 Solving the system

Now, the mainly problem is calculating the pseudo-inverse of the L_m matrix, which correspond to A_u matrix in planar geometry. They are same as each other in mathematical analysis. So we only present this steps just in spherical geometry.

The sky brightness temperature spherical harmonics coefficients can be estimated by solving each of the *m*-modes linear systems defined by Eq. 3.4.14. The \mathbf{L}_m matrix size is $2n_{\text{beams}} \times \ell_{\text{max}}$, with ℓ_{max} around few thousands for array sizes $\leq 100 \text{ m}$ and a number of beams up to a to a few thousands for the current generation of instruments. Although these systems are usually under-determined, the solution can formally be written as:

$$\left[\widehat{\mathcal{I}}(\ell)\right]_{m} = \mathbf{H}_{m}\left[\widetilde{\mathcal{V}}\right]_{m}$$
(3.5.1)

where [] are used to denote vectors and \mathbf{H}_m is the noise weighted Moore-Penrose pseudo-inverse of \mathbf{L}_m [9].

To make map from a given set of visibilities with noise, we look for a maximum likelihood solution. Here we assume that the noise on visibility measurement follows a Gaussian random process, with variance $\mathbf{N}_m = \langle [\tilde{n}]_m [\tilde{n}]_m^{\dagger} \rangle$. We consider moreover that noise is uncorrelated for different *m*-modes. This hypothesis is valid as long as the time domain noise is a Gaussian random process characterised by a power spectrum. The likelihood function of the measurement can be written as,

$$p(\mathcal{V}|\mathcal{I}) \propto \exp[-\frac{1}{2}(\mathcal{V} - \mathbf{L}\mathcal{I})^T \mathbf{N}^{-1}(\mathcal{V} - \mathbf{L}\mathcal{I})].$$

The solution is given by

$$\widehat{\mathcal{I}}_m = (\mathbf{L}_m^{\dagger} \mathbf{N}_m^{-1} \mathbf{L}_m)^{-1} \mathbf{L}_m^{\dagger} \mathbf{N}_m^{-1} \mathcal{V}_m \equiv \mathbf{H}_m \,\mathcal{V}_m \tag{3.5.2}$$

$$\mathbf{H}_m = (\mathbf{L}_m^{\dagger} \mathbf{N}_m^{-1} \mathbf{L}_m)^{-1} \mathbf{L}_m^{\dagger} \mathbf{N}_m^{-1}$$
(3.5.3)

Redundant baselines are counted once with their noise level being scaled accordingly, i.e. $\sigma_n^2 \propto N_{rb}^{-1}$, where N_{rb} denotes the number of redundant baselines (number of antennae pairs with the same baseline).

If we further assume that noise is uncorrelated between different baselines, the noise covariance matrix N_m for each *m* becomes diagonal. In this case, the computation can be further simplified as

$$\mathbf{H}_m = \left(\mathbf{N}_m^{-\frac{1}{2}} \mathbf{L}_m\right)^{-1} \mathbf{N}_m^{-\frac{1}{2}}$$
(3.5.4)

Here, $N^{-1} = \left(N^{-\frac{1}{2}}\right)^{\dagger} N^{-\frac{1}{2}}$, which could be obtained simply by taking the inverse of square root of each non-zero diagonal element. However, $\mathbf{N}_m^{-\frac{1}{2}} \mathbf{L}_m$ is often a non invertible matrix, so we have to use the pseudo-inverse method. Here, we use the Singular Value Decomposition (SVD) of a matrix, for which we need to compute the inverse of the matrix with the eigenvalues. Given a $m \times n$ real or complex matrix A, which can be factorized in the form

$$A = U\Sigma V^{\dagger} \tag{3.5.5}$$

where U is an $m \times m$ real or complex unitary matrix, V^{\dagger} (the conjugate transpose of V, or simply the transpose of V if V is real) is an $n \times n$ real or complex unitary matrix, and Σ is an $m \times n$ rectangular diagonal matrix which has at most min(m, n) non zero eigenvalues, denoted by Σ_{ii} . These diagonal entries Σ_{ii} are known as the singular values of A. A common convention is to list the singular values in descending order, then the diagonal matrix Σ is uniquely determined by A, though the matrices U and V are not unique. Using SVD, the pseudo-inverse of the matrix A is given by

$$B \equiv A^{-1} = V \bar{\Sigma}^{-1} U^{\dagger} \tag{3.5.6}$$

where $\bar{\Sigma}^{-1}$ is a $n \times m$ diagonal matrix, which is formed by replacing every nonzero diagonal entry Σ_{ii} by its reciprocal $1/\Sigma_{ii}$ and transposing the resulting matrix.

Due to limited numerical precision in the computation, even zero eigenvalues elements of $\Sigma_{i,i}$ will have some small non-zero value, which would give rise to large $\Sigma_{i,i}^{-1}$ and affect the result greatly if left unattended. The contribution of noise to the reconstructed sky modes are also large for the modes with small values. In order to limit the contribution of noise and to avoid numerical instabilities, we limit the value of eigenvalues, and the small eigenvalues of Σ are set to 0 before the inversion, and its inverse also set to 0 and ignored in subsequent computation. In practice, we set two threshold values for the diagonal elements. If the diagonal element $\Sigma_{i,i} < \Sigma_{0,0} \times \epsilon_r$

m modes	frist eigenvalue	thresholds	non-zero eigenvalues
1	5.15878	0.103176	53
50	1.93873	0.0387746	71
100	1.71407	0.0342814	51
150	1.38509	0.0277018	45
200	1.56662	0.0313324	41
250	1.39035	0.0278069	28
300	0.604846	0.0120969	16
350	0.0449966	0.01	4

Table 3.1: The SVD computation for a few *m*-modes in the PAON-4 case.

or $\Sigma_{i,i} < \epsilon_a$, we set $1/\Sigma_{i,i} \to 0$, where $\Sigma_{0,0}$ is the largest eigenvalue, and ϵ_r, ϵ_a are small value ratios which define the relative and absolute thresholds for defining non zero eigenvalues Σ_{ii} when computing their inverse $1/\Sigma_{ii}$. We experimented with different values of ϵ_r, ϵ_a , finally choosing $\epsilon_r = 0.02$ and $\epsilon_a = 0.01$ for most map reconstructions shown here, in order to avoid contaminating the maps by modes with large noise.

The above threshold is a sharp cut in the eigenvalues when computing the pseudo-inverse. This leads to some striping and oscillations in the reconstructed sky maps. Then we introduce a smooth threshold, for which the eigenvalues are smoothly decreasing for elements larger than the relative or absolute threshold. In chapter 5, we will show the comparison of reconstructed map with sharp and smooth threshold.

A review on the Moore-Penrose pseudo inverse computation and properties can be found in [9]. Once all the sky spherical harmonics coefficients are determined by solving all the *m*-modes systems, we can compute the sky map $\hat{I}(\vec{n})$ by performing an inverse Spherical Harmonics Transform (SHT) on the estimated spherical modes coefficients $[\hat{\mathcal{I}}(\ell)]_m$.

Table 3.1 shows a few examples of the SVD computation for the pseudo-inverse $(\mathbf{N}_m^{-1/2}\mathbf{L}_m)^{-1/2}$ matrix for the PAON4 configuration. In this table we show for the selected m mode the largest eigenvalue, the threshold value, and the number of non-zero eigenvalues, i.e. the number of eigenvalues which are above the threshold (eigenvalues below threshold are set to zero to avoid contamination by numerical error). The threshold is different for each case, because the largest eigenvalue is different, and we see that among the examples listed here the threshold values change, implying that most are determined by the relative criterion, i.e. that the eigenvalue must be less than 0.02 of the largest eigenvalue. As m increases, the non-zero modes decreases, eventually reaching 0.01, which is the absolute threshold adopted here. We have also tried to vary the threshold a bit, such variations affect the number of non-zero modes, but not by much as long as we try to keep the numerical solution stable.

3.6 Various indicators

3.6.1 Reconstructed maps and PSF

The L_m matrices depend only on the array configurations or baselines, individual antenna beams and the scanning strategy, so H_m also depends on these. It does depend on the noise covariance matrix structure, but not on its values. For instance, the H_m remains unchanged if we change the total survey duration, or the system temperature for all feeds, without changing the array configuration (baselines and number of redundant baselines), or the scanning strategy, i.e. how the fraction of the total survey time spend on each declination. So, once the H_m are computed, we can apply them to different input visibilities to reconstruct different sky maps:

- To obtain the instrument response to a point source, which corresponds to the PSF (Point Spread Function) or the instrument beam, we reconstruct the maps from mock visibilities computed from input maps containing point sources at different declinations. The PSF is independent of the right ascension, but varies for different declinations.
- Starting from an input sky map, we can compute its decomposition into spherical harmonics (I(l, m)) using the SHT. Then using the L_m matrices computed by the map-making tools, we can compute the visibility matrices, with or without adding noise. Applying the H_m to a set of such mock visibility data, we can reconstruct the sky maps as seen by an transit interferometric array. For demonstrations in this thesis, we have used the Leiden-Argentina-Bonn(LAB) survey [60] for the sky emission at 21-cm (Galactic HI). The LAB data has been used to create spherical maps at several frequencies, suitable for processing by our software tools. The sky emission brightness map at 21-cm (1420.4 MHz) in equatorial coordinates used in this thesis is shown in the top panel of figure 3.6.1. For the radio continuum which is dominated by the Galactic synchrotron emission at the relevant frequencies, we have used full sky maps generated by the Global Sky Model (GSM) [29]. We will consider the continuum radiation at 1250 MHz (bottom panel of Fig.3.6.1), so we could see the system response and map making at different frequencies.
- We can also compute pure noise maps, if the input visibility vectors contain contribution from noise only. These pure noise maps can be used to compute survey noise power spectrum, as an alternative to using the noise covariance matrix (see section 3.6.3 below). To limit statistical fluctuations, we generated 50 random noise maps from noise-only visibilities for computing noise power spectra.

We have used spherical maps with HEALPix pixelization scheme [45] for the reconstructed maps presented in this thesis, although two other pixelization schemes are currently provided by SOPHYA and could be used by the map making software. We have checked that the results are not sensitive to ℓ_{max} and the corresponding HEALPix n_{side} parameter as long as the map resolution is at least a factor 2 higher than the synthesised beam resolution:

$$\ell_{\rm max} \gtrsim \frac{2\pi \, D_{\rm array}}{\lambda}$$



Figure 3.6.1: Input maps for our simulation. Top: the sky brightness temperature at 1420.4 MHz (21-cm) from the LAB survey; Bottom: the sky brightness temperature at 1250MHz (radio continuum) computed using the GSM model.

where $D_{\rm array}$ is the diameter of the disk covering the full array. For PAON-4 with $D_{\rm array} \sim 18 \,\mathrm{m}$, $\ell_{\rm max} = 750$ and $n_{\rm side} = 256$ would be more than enough for reconstructing maps. However, The Tianlai circular array configuration with $D_{\rm array} \sim 40 \,\mathrm{m}$ requires $\ell_{\rm max} \gtrsim 1200$. We have thus used $\ell_{\rm max} = 1500$ and HEALPix $n_{\rm side} = 512$, corresponding to a pixel resolution of $\sim 6.9 \,\mathrm{arcmin}$, for most of the results presented in this thesis.

3.6.2 Instrument response and transfer function

The *m*-mode reconstruction matrix $\mathbf{R}_m \equiv (\mathbf{H}_m \mathbf{L}_m)$ tells us how the estimated sky spherical harmonics coefficients $(\hat{\mathcal{I}}(\ell, m))$ are related to the true sky ones $(\mathcal{I}(\ell, m))$

$$\left[\widehat{\mathcal{I}}(\ell)\right]_{m} = \mathbf{R}_{m} \left[\mathcal{I}(\ell)\right]_{m}$$
(3.6.1)

Ideally, if $\mathbf{R}_m = \mathbf{I}$ where \mathbf{I} is the identity matrix, then we would be able to recover the sky spherical harmonic *m*-mode completely from the observations. However, in reality this is not possible. Although each *m* mode is measured independently for a full circle transit observation, for each given *m* the different ℓ coefficients are still correlated, the physical measurement data is a mix of different ℓ mode contributions. The \mathbf{R}_m matrix gives the window function in ℓ -space for the estimated sky. We can define the core response matrix \mathbf{R} by extracting the diagonal terms from individual \mathbf{R}_m matrices:

$$\mathbf{R}(\ell,m) = \mathbf{R}_m(\ell,\ell)$$

For reconstruction, the $\mathbf{R}(\ell, m)$ is insufficient and the original \mathbf{R}_m matrices are needed, but the $\mathbf{R}(\ell, m)$ matrix can give some idea of how well an (ℓ, m) mode is measured with the given array, so it can help us to see the effectiveness of our reconstruction in the (ℓ, m) space.

We can further compress the response function by computing the transfer function, which is defined by the average over the m-modes from the response matrix \mathbf{R} ,

$$T(\ell) = \langle |\mathbf{R}(\ell, m)| \rangle_m \tag{3.6.2}$$

Let's consider visibilities corresponding to an input white noise map, without any additional noise $(\sigma_{\text{noise}} = 0)$

$$\langle |\mathcal{I}(\ell,m)|^2 \rangle = C^{\mathrm{in}}(\ell) = \mathrm{const} \langle \mathcal{I}(\ell,m) \left(\mathcal{I}(\ell',m') \right)^* \rangle = \delta_{\ell\ell',mm'} C^{\mathrm{in}}(\ell)$$

if we reconstruct the map from such visibilities and compute the reconstructed map power spectrum, we can write it as:

$$\langle \left[\widehat{\mathcal{I}}(\ell) \right]_m \left[\widehat{\mathcal{I}}(\ell') \right]_m^{\dagger} \rangle = \langle \mathbf{R}_m \left[\mathcal{I}(\ell)^{in} \right]_m \left[\mathcal{I}'(\ell)^{in} \right]_m^{\dagger} \mathbf{R}_m^{\dagger} \rangle$$

where \dagger denotes Hermitian conjugate (transpose and complex conjugate). Noting that \mathbf{R}_m are projector matrices.

$$\begin{aligned} \mathbf{R}_m^{\dagger} &= \mathbf{R}_m \\ \mathbf{R}_m^{\dagger} \mathbf{R}_m &= \mathbf{L}_m^{\dagger} \mathbf{H}_m^{\dagger} \mathbf{H}_m \mathbf{L}_m = \mathbf{H}_m \mathbf{L}_m = \mathbf{R}_m \end{aligned}$$

and that for a white noise input map, the covariance matrix in spherical harmonics space is proportional to the identity matrix I:

$$\langle \left[\mathcal{I}(\ell) \right]_m \left[\mathcal{I}(\ell') \right]_m^{\dagger} \rangle = \text{const} \times \mathbf{I}$$

we obtain that:

$$\langle |\widehat{\mathcal{I}}(\ell,m)|^2 \rangle = \mathbf{R}_m(\ell,\ell) \, C^{\mathrm{in}}(\ell) = \mathbf{R}(\ell,m) \, C^{\mathrm{in}}(\ell)$$

So, if we compute the reconstructed map power spectrum by averaging $|\widehat{\mathcal{I}}(\ell, m)|^2$ over all mmodes, the ratio of the reconstructed map angular power spectrum, to the input map flat angular power spectrum would be equal to the transfer function defined above:

$$C^{\rm rec}(\ell) = \langle |\widehat{\mathcal{I}}(\ell, m)|^2 \rangle_m \longrightarrow T(\ell) = \frac{C^{\rm rec}(\ell)}{C^{\rm in}(\ell)}$$

where $C^{\text{rec}}(\ell)$ is the power spectrum of the reconstructed map, computed from visibilities corresponding to the observation of a white noise sky, and $C^{\text{in}}(\ell) = \text{const}$ denotes the input sky flat power spectrum. The computation of the transfer function from reconstructed map power spectrum proves easier to use when additional filtering in the (ℓ, m) plane or masking in angular space is applied after the $\hat{\mathcal{I}}(\ell, m)$ computation stage.

If we consider a masked sky map, a spherical map where pixels outside the observed area are put to zero, the computed variance of pixel values is lowered by a factor $\sim f_{\rm sky}$, where $f_{\rm sky} = \Omega_{\rm obs}/(4\pi)$ is the observed fraction of the full sky area. We expect thus to obtain transfer functions with levels close to $f_{\rm sky}$.

It should be noted that the cosmological signal is characterised by its 3D power spectrum P(k), which determines the signal power spectrum $C^{\text{sig}}(\ell)$ for frequency shells reconstructed by the map making process. The transfer function can be used to compute the *observed signal* power spectrum for each frequency shell, $C^{\text{obs}}(\ell) = T(\ell) \times C^{\text{sig}}(\ell)$. In the absence of foregrounds, the comparison between the expected observed signal power spectrum $C^{\text{obs}}(\ell)$ and the noise power spectrum (section 3.6.3 below) is the main tool to estimate the ability of a given instrument to measure a signal characterised by its power spectrum.

3.6.3 Error covariance matrix and noise power spectrum

If we consider the reconstruction of sky spherical harmonics coefficients from pure noise visibilities ($\tilde{\mathcal{V}}_{ij} = \tilde{n}_{ij}$), the covariance matrix $\mathbf{Cov}_m(\ell_1, \ell_2)$ of the estimator $\hat{\mathcal{I}}(\ell, m)$ for each mode m can be computed from the \mathbf{H}_m matrix and the noise covariance matrix:

$$\begin{split} \mathbf{N}_m &= \left[\tilde{\mathcal{V}}_{ij} \right]_m \cdot \left[\tilde{\mathcal{V}}_{ij} \right]_m^{\dagger} \\ \mathbf{Cov}_m(\ell_1, \ell_2) &= \left\langle \left[\widehat{\mathcal{I}}(\ell) \right]_m \cdot \left[\widehat{\mathcal{I}}(\ell) \right]_m^{\dagger} \right\rangle \\ &= \mathbf{H}_m \, \mathbf{N}_m \, \mathbf{H}_m^{\dagger} \end{split}$$

The covariance matrix is not diagonal, especially due to partial sky coverage in declination. However, if we ignore this correlation and use the diagonal terms only for each m mode, we can gather them together to create the $\sigma_{\mathcal{I}}^2(\ell, m)$ variance matrix. This matrix informs us on how well each (ℓ, m) mode is measured. This noise variance matrix can then be used in subsequent processing steps, for example to throw out modes with large errors, or by applying weights inversely proportional to the error variance. We can even compress further this information by computing the noise power spectrum:

$$\sigma_{\mathcal{I}}^2(\ell, m) = \mathbf{Cov}_m(\ell, \ell) \tag{3.6.3}$$

$$C^{\text{noise}}(\ell) = \langle \sigma_{\mathcal{I}}^2(\ell, m) \rangle_m \tag{3.6.4}$$

As pointed out in the previous section, we can also compute noise maps by applying the \mathbf{H}_m matrices to noise only visibility. We can then use the reconstructed maps to compute the noise power spectrum, which is identical to the one obtained directly from Eq.(3.6.4). However, the noise maps can still prove useful for computing the noise power spectrum when further filtering in angular or spherical harmonics space is applied.

As mentioned earlier, the \mathbf{H}_m matrix does not change if the visibility noise matrix is scaled. The map making is thus performed with a value of visibility time sample noise $\sigma_{\text{noise}} = 1$ K. To compute the noise level for a given survey, the noise covariance matrix is rescaled by the effective σ_{noise} . To compute this value, we take into account the system temperature T_{sys} , the total survey time t_{survey} and the frequency band $\Delta \nu$ for each sky map. The number n_t of time samples for the visibilities for a 24 hours constant declination scan is fixed by the maximum value of m, with $n_t = 2m_{\text{max}}$, as the m-modes visibility vector and time indexed visibilities are related by an FFT. m_{max} is itself equal to the ℓ_{max} which has to be chosen so that $\ell_{\text{max}} \gtrsim \frac{2\pi D_{\text{array}}}{\lambda}$. For the survey strategies discussed in this thesis, we have distributed the observation time evenly among all constant declination scans. The effective σ_{noise} for measured visibility time samples can then be written as a function of integration time per time sample t_{int} :

$$\sigma_{\text{noise}}^2 = \frac{2T_{\text{sys}}^2}{t_{\text{int}}\,\Delta\nu} \tag{3.6.5}$$

$$t_{\rm int} = \frac{t_{\rm survey}/{\rm Days}}{n_{\delta_n}} \times \frac{24 \times 3600}{2 \, m_{\rm max}}$$
(3.6.6)

As pointed out above, the variance of masked map is lower by a factor $f_{\rm sky}$ compared to the corresponding full map. In order to make the noise power spectra comparable for the different configurations, all the noise power spectra shown in this thesis are rescaled according to $C^{\rm noise}(\ell) \times (1/f_{\rm sky})$, where $C^{\rm noise}(\ell)$ is computed from eq. 3.6.4.

3.7. EXTENSION TO POLARISATION

3.6.4 Filtering in (ℓ, m) space and angular masking

Once the sky spherical harmonic coefficients are computed by the map making process described above, it is possible to apply additional filters, either in the Fourier space ((ℓ, m) plane) or angular space. These filters can be used for example to decrease the noise level in the final sky map, by ignoring or damping modes with high noise, the noise variance matrix being the key tool to design such filters. Another possible application would be to correct the instrument response frequency dependence, in which case the **R** matrix or the full response matrix **R**_m would be the key tool.

Finally, optimal filters for component separation and cosmological signal extraction could be designed by the simultaneous use of the instrument response, noise covariance matrix, statistical knowledge of the signal and the foregrounds. The discussion of such optimal filtering methods is beyond the scope of this thesis. Here, we will apply a simple mask with sharp edges in declination to define precisely the fiducial sky region, and we also consider two simple noise-reduction filters described below.

The first simple filter we consider is $W_1(\ell, m)$, which uses the noise variance matrix $\sigma_{\mathcal{I}}^2$.

$$W_1(\ell,m) = \begin{cases} 1, & \text{if } \sigma_{\mathcal{I}}^2(\ell,m) < \sigma_{\text{thr}}^2; \\ \frac{1}{\sigma_{\mathcal{I}}^2(\ell,m)}, & \text{if } \sigma_{\mathcal{I}}^2(\ell,m) > \sigma_{\text{thr}}^2. \end{cases}$$

The threshold σ_{thr}^2 is defined as $K \sigma_{\min}^2$ where σ_{\min}^2 denotes the minimum value of the noise variance matrix, and K a constant factor. A value of K = 50 has been used for the examples shown in this thesis. This filter suppresses modes with very large errors.

A second filter we consider is the weight function $W_2(\ell)$, independent of m. This second weight function is used to reject high noise modes at high ℓ near the edge of instrument sensitivity region, and also the low ℓ modes when the autocorrelation signal is not used:

$$W_{2}(\ell) = \begin{cases} \left(1 + e^{\frac{\ell - \ell_{A}}{\Delta \ell_{A}}}\right)^{-1}, & \text{With - AutoCorr;} \\ \left(1 + e^{\frac{\ell - \ell_{A}}{\Delta \ell_{A}}}\right)^{-1} \times \left(1 + e^{\frac{\ell_{B} - \ell}{\Delta \ell_{B}}}\right)^{-1}, & \text{No - AutoCorr.} \end{cases}$$

The filter parameter ℓ_A , $\Delta \ell_A$, ℓ_B , $\Delta \ell_B$ are determined empirically. For the case of PAON-4, $\ell_A = 440$, $\Delta \ell_A = 15$, $\ell_B = 90$, $\Delta \ell_B = 10$. For the Tianlai circular dish array case ($\ell_A = 1050$, $\Delta \ell_A = 15$, $\ell_B = 120$, $\Delta \ell_B = 10$). Fig. 3.6.2 shows the weight function $W_2(\ell)$ for these two cases.

3.7 Extension to polarisation

The formalism presented before deals with unpolarised sky signal which is suitable for the description of the cosmological 21-cm emission. However, it is well known that most RFI are strongly polarised, as well as components of the foregrounds emission, in particular emission from compact radio sources [33] or the synchrotron emission of our own galaxy. For polarised emission,



Figure 3.6.2: Weight function $W(\ell)$ applied to computed $\mathcal{I}(\ell, m)$. The black curve suit for PAON4 with autocorrelation, while the red curve suit for PAON4 without autocorrelation. The blue curve is for Tianlai dish array without autocorrelation.

the Faraday rotation due to the interstellar magnetic field imprints frequency dependent structures on the polarised emission, increasing foreground separation difficulty. The reconstruction of the polarised sky emission maps is thus mandatory for intensity mapping project. In this section we describe briefly how the method described in this thesis and the corresponding code can be extended to handle reconstruction of polarised brightness maps from polarised visibility signals. We largely follow the results given in the reference [97].

Polarisation characterises the vectorial nature of electromagnetic (EM) radiation, representing a fundamental property separate from its frequency and intensity. The polarisation of an antenna refers to the orientation of the electric field of the radio wave with respect to the Earth's surface and is determined by the physical structure of the antenna and by its orientation. We assume that each antenna is equipped with dual polarisation receivers, measuring two orthogonal linear polarisations (\hat{x}, \hat{y}) of the incoming electromagnetic field [18], for example a component \hat{x} parallel to the horizontal plane \hat{y} parallel to the meridian plane. The measured electric signal for each polarisation is a combination of the corresponding projection of the electric field contributions coming from different and incoherent directions of the sky. The polarization state of electromagnetic waves is often described using a 4-element column vector corresponding to the Stokes parameters $S = (I, Q, U, V)^T$ where superscript T denotes the matrix transpose. If e_x and e_y denotes the two electric field components transverse to the line of sight, one gets

$$I = \langle e_x e_x^* \rangle + \langle e_y e_y^* \rangle \qquad \qquad Q = \langle e_x e_x^* \rangle - \langle e_y e_y^* \rangle \\ U = \langle e_x e_y^* \rangle + \langle e_y e_x^* \rangle \qquad \qquad V = -i(\langle e_x e_y^* \rangle - \langle e_y e_x^* \rangle) \qquad (3.7.1)$$

where the $\langle \rangle$ denotes a time average, and we have omitted the direction $(\vec{\omega})$ dependence for simplicity. The visibilities \mathcal{V}_{p_i,p_j} have to be computed for all signal pairs, (p_i, p_j) indices identifying

3.7. EXTENSION TO POLARISATION

the antenna pair (i, j), as well as the polarisation probe x or y. The full set of visibilities \mathcal{V}_{p_i,p_j} can be split in two sets: x and y polarisations auto and cross correlations $\mathcal{V}_{ij}^{xx}, \mathcal{V}_{ij}^{yy}$ and cross polarisation visibilities $\mathcal{V}_{ij}^{xy}, \mathcal{V}_{ij}^{yx}$. For an array with N dual polarisation receivers, there will be a total of $2N^2$ visibilities, corresponding to 2N autocorrelations for the x and y polarisation signals, $\frac{N(N-1)}{2}$ cross correlations visibilities for each of xx and yy polarisation signal pairs, and $\frac{N^2}{2}$ visibilities for each of the cross polarisation xy and yx pairs.

$$\mathcal{V}_{p_i,p_j} = \left[\mathcal{V}_{ij}^{xx}; \mathcal{V}_{ij}^{yy}; \mathcal{V}_{ij}^{xy}; \mathcal{V}_{ij}^{yx} \right]$$
(3.7.2)

$$p_i = \{(i, x), (i, y)\} \qquad p_j = \{(j, x), (j, y)\}$$
(3.7.3)

The generalization of Eq. 3.1.2 reads

$$\mathcal{V}_{p_i p_j} = \iint \sum_{a} \mathbf{L}^a_{p_i p_j}(\vec{\omega}) S_a(\vec{\omega}) \, d\vec{\omega}$$
(3.7.4)

where the sum on index a runs over the four Stokes parameters. The four beams

$$\mathbf{L}_{p_i p_j} = \left\{ L_{p_i p_j}^I, L_{p_i p_j}^Q, L_{p_i p_j}^U, L_{p_i p_j}^V \right\}$$

are a generalisation of the beam pattern of Eq.(3.4.6) that takes into account the response of each of the two linear polarisation probes of the feed, including the response to the incoming electric field signal, as well as all possible leakage sources from one polarisation to the other. It might include other effects impacting polarisation measurement, such as polarisation direction rotation or leakage due to the atmosphere and/or earth magnetic field. The Stokes parameters decomposition requires spin-weighted spherical harmonics [115] with spin-0 for I and V and spin-2 ($_{s}Y_{\ell m}$) for Q and U. From the two real quantities U and Q, we define two complex linear combinations, and corresponding spherical harmonics coefficients: $\bar{Q} = (Q + iU)/2$ and $\bar{U} = (Q - iU)/2$:

$$I(\vec{\omega}) \xrightarrow{_{0}Y_{\ell m}} \mathcal{I}_{\ell m} \qquad \qquad \bar{Q}(\vec{\omega}) \xrightarrow{_{+2}Y_{\ell m}} \bar{Q}_{\ell m}$$
$$\bar{U}(\vec{\omega}) \xrightarrow{_{-2}Y_{\ell m}} \bar{\mathcal{U}}_{\ell m} \qquad \qquad V(\vec{\omega}) \xrightarrow{_{0}Y_{\ell m}} \mathcal{V}_{\ell m} \qquad (3.7.5)$$

The angular responses of the polarised beam **L** may also be decomposed in spherical harmonics. As for the Stokes parameters, we define linear combination of Q and U beams: $L^{\bar{Q}} = L^Q - iL^U$ and $L^{\bar{U}} = L^Q + iL^U$ to match the definition of \bar{Q} and \bar{U} respectively.

$$L_{p_ip_j}^{I} \xrightarrow{0Y_{\ell m}} \mathcal{L}_{p_ip_j;\ell m} \qquad \qquad L_{p_ip_j}^{\bar{Q}} \xrightarrow{-2Y_{\ell m}} \mathcal{L}_{p_ip_j;\ell m}^{\bar{Q}}$$

$$L_{p_ip_j}^{\bar{U}} \xrightarrow{+2Y_{\ell m}} \mathcal{L}_{p_ip_j;\ell m}^{\bar{U}} \qquad \qquad L_{p_ip_j}^{V} \xrightarrow{0Y_{\ell m}} \mathcal{L}_{p_ip_j;\ell m}^{V} \qquad (3.7.6)$$

Using the symmetry property of spin-2 spherical harmonics ${}_{s}Y_{l,m}^{*} = (-1)^{s+m} {}_{-s}Y_{l,-m}$, and the orthogonality of the spin-weighted spherical harmonics, the extension of Eq.(3.4.8) reads

$$\mathcal{V}_{p_i p_j} = \sum_{ml} (-1)^m \left(\mathcal{L}_{p_i p_j;\ell,-m}^I \mathcal{I}_{\ell m} + \mathcal{L}_{p_i p_j;\ell,-m}^{\bar{Q}} \bar{\mathcal{Q}}_{\ell m} + \mathcal{L}_{p_i p_j;\ell,-m}^{\bar{U}} \bar{\mathcal{U}}_{\ell m} + \mathcal{L}_{p_i p_j;\ell,-m}^V \mathcal{V}_{\ell m} \right) \quad (3.7.7)$$
It is convenient to decompose \overline{Q} and \overline{U} with the gradient (E) and curl (B) components and the corresponding beams $\mathcal{L}_{p_i p_j}$

$$\bar{\mathcal{Q}}_{\ell m} = -\mathcal{E}_{\ell m} - i\mathcal{B}_{\ell m} \qquad \qquad \mathcal{L}_{p_i p_j;\ell m}^{\bar{Q}} = \left(-\mathcal{L}_{p_i p_j;\ell m}^E + i\mathcal{L}_{p_i p_j;\ell m}^B\right)/2 \\
\bar{\mathcal{U}}_{\ell m} = -\mathcal{E}_{\ell m} + i\mathcal{B}_{\ell m} \qquad \qquad \mathcal{L}_{p_i p_j;\ell m}^{\bar{U}} = \left(-\mathcal{L}_{p_i p_j;\ell m}^E - i\mathcal{L}_{p_i p_j;\ell m}^B\right)/2 \qquad (3.7.8)$$

Then,

$$\mathcal{V}_{p_i p_j} = \sum_{m\ell} (-1)^m \left(\mathcal{L}_{p_i p_j;\ell,-m}^I \mathcal{I}_{\ell m} + \mathcal{L}_{p_i p_j;l,-m}^E \mathcal{E}_{\ell m} + \mathcal{L}_{p_i p_j;l,-m}^B \mathcal{B}_{\ell m} + \mathcal{L}_{p_i p_j;l,-m}^V \mathcal{V}_{\ell m} \right) \quad (3.7.9)$$

As all Stokes parameters are real functions then $\bar{U}^*(\vec{\omega}) = \bar{Q}(\vec{\omega})$ and this leads the relation in harmonic space $\bar{Q}_{\ell m} = (-1)^m \bar{U}^*_{l,-m}$ and to relations which extend the case of $\mathcal{I}_{\ell m}$ as

$$X_{l,-m} = (-1)^m X_{\ell m}^* \quad \mathbf{X} \in \{\mathcal{I}, \, \mathcal{E}, \, \mathcal{B}, \, \mathcal{V}\}$$
(3.7.10)

So, one can extend both the Fourier decomposition Eq. (3.4.11) as well as the positive and negative m-mode separation Eqs. (3.4.12), (3.4.13).

$$\tilde{\mathcal{V}}_{p_i p_j}(m) = \sum_{\ell=|m|}^{+\ell_{max}} \sum_{\mathcal{X}} (-1)^m \mathcal{L}_{p_i p_j; l, -m}^X \mathcal{X}_{\ell m}$$
(3.7.11)

$$\tilde{\mathcal{V}}_{p_i p_j}^*(-m) = \sum_{\ell=|m|}^{+\ell_{max}} \sum_{\mathcal{X}} \mathcal{L}_{p_i p_j; l, m}^{X*} \mathcal{X}_{\ell m}$$
(3.7.12)

with $\mathcal{X} = \mathcal{I}, \mathcal{E}, \mathcal{B}, \mathcal{V}$.

Extending our map making software to perform computation for the polarised case would be rather straightforward, except maybe for the computation of the polarised beams $\left\{L_{p_ip_j}^I, L_{p_ip_j}^Q, L_{p_ip_j}^U, L_{p_ip_j}^V\right\}$, from individual feed polarised beam responses. The implementation of the extension is postponed to future work.

Chapter 4

Application to different dense configurations

In this chapter, we apply the formalism developed in the last section to both dish array and cylinder array configuration.

4.1 Application to PAON-4

In this section, we apply the formalism developed in the last section to the PAON-4 case, and also compare it with the compact 2×2 array and large single dish.

However, in order to provide clues for understanding a given instrument response and the impact of various parameters, we first discuss the features of response matrices \mathbf{R}_m , \mathbf{R} using some specific cases, in particular a non-compact 2×2 array. This will also illustrate why we choose the very compact array layout for PAON-4 and Tianlai. Before applying the full 2D reconstruction method to PAON-4 and Tianlai configurations, we have trid to illustrate some of the features associated with interferometric reconstruction from visibilities in Appendix A.

4.1.1 Beams and response matrix features

The antenna pair beam patterns are the key elements to understand the complete instrument response. We shall thus present briefly the characteristics of a few beams before discussing the instrument response matrix features. Figure 4.1.1 shows the beam patterns $\mathcal{L}_{\ell,m}$ for a few configurations (baselines and declinations) for dishes with effective diameter $D_{\text{eff}} = 4.5 \text{ m}$ and $\lambda = 21 \text{ cm}$. If the antenna is pointing toward the north (or south) pole, and assuming a axisymmetric feed response around the dish axis, the beam in the (ℓ, m) plane would only have non zero coefficients for m = 0 and for $\ell \leq 2\pi D_{\text{eff}}/\lambda$. However, if the antenna axis is pointed to another direction, then the beam will have non zero coefficients for $m \neq 0$. This can be seen for example on the topleft panel of fig. 4.1.1, which represent the single dish (auto-correlation) beam with the antenna axis in the equatorial plane ($\delta = 0$). We see that the beam pattern covers a triangular shaped area



Figure 4.1.1: The beam patterns in spherical harmonics $\mathcal{L}_{\ell,m}$ with dish size D = 4.5m. Top left: auto-correlation located at equator; top right: cross-correlation beam from a pair of dish in North-South direction with $d_{sep} = 7m$ located at equator; bottom left: cross-correlation beam from a pair of dish in East-West direction with $d_{sep} = 7m$ located at equator; bottom right: cross-correlation beam from a pair of dish in East-West direction with $d_{sep} = 7m$ located at declination $\delta = 60^{\circ}$.



Figure 4.1.2: Examples of \mathbf{R}_m matrix. Here we assume that the band-like region between $35^{\circ}23'$ and $59^{\circ}23'$ is surveyed. Top: the \mathbf{R}_m matrix for an auto-correlation of a single dish with $D_{\text{eff}} = 4.5m$. Top Left: m = 0; Top Center: m = 50; Top Right: m = 100. Bottom: the \mathbf{R}_m matrix for an auto-correlation of a pair of dishes in EW baseline and $d_{\text{sep}} = 7m$. Bottom Left: m = 50; Bottom Center: m = 250.

in the (ℓ, m) space, due to pointing at the equator, the beam coverage extends to the maximum allowed m value, i.e. $m = \ell$. For pointing to an arbitrary direction defined by the declination δ , the bound would actually be $m = \ell \cos \delta$. As expected, we can see also that the beam falls off beyond $\ell_{max} \sim 2\pi D_{\text{eff}}/\lambda \sim 135$.

For cross correlations, we expect the beam center to be given by $(\ell, m) = (2\pi |\vec{u}|, 2\pi u \cos \delta)$, where $\vec{u} \equiv (u, v, w) = (\Delta x/\lambda, \Delta y/\lambda, \Delta z/\lambda)$. For an antenna pair separated by an east-west baseline, the beam has a crescent shape, with $m \approx \ell \cos \delta$ and centered at $(\ell_0 \sim 2\pi d_{\text{sep}}/\lambda, m \sim \ell_0 \cos \delta)$. This is shown on the bottom left and bottom right panels, for an east-west baseline with $d_{\text{sep}} = 7m$ and for declinations $\delta = 0^\circ$ and $\delta = 60^\circ$ respectively. By contrast the northsouth baseline (top right panel) is mildly sensitive to the sky intensity variations along the East-West direction; the beam pattern is centered at m = 0, but is shifted along the ℓ direction to $\ell_0 \sim 2\pi d_{\text{sep}}/\lambda \sim 210$, and the extension along m-direction $(-m_{max} < m < +m_{max})$ is given by the dish size $m_{max} \sim 2\pi D_{\text{eff}}/\lambda \sim 135$.

To gain a better sense of the reconstruction, we plot in Fig.4.1.2 the \mathbf{R}_m matrix for a few m

values, and for a sky survey area similar to the PAON-4 case, i.e $35^{\circ} \lesssim \delta \lesssim 59^{\circ}$. The top panels show the case of a single dish ($D_{\text{eff}} = 4.5\text{m}$) auto-correlation, with m = 0 (top left), m = 50 (top center) and m = 100 (top right). For the m = 0 case, in principle the $\mathbf{R}_m(\ell, \ell')$ matrix can have infinite range of ℓ, ℓ' , but at large ℓ, ℓ' the beam response function $\mathcal{L}_{\ell,m}$ becomes very small, and in our inversion procedure such modes would be below the threshold and cut off. We see that the $\mathbf{R}_0(\ell, \ell')$ effectively cuts off for $\ell, \ell' > 135$. For the m = 50 and m = 100 cases, the ℓ and ℓ' are limited to within $m \leq \ell, \ell' \leq m/\cos \delta$, where in this case $\delta \sim 59^{\circ}$. For very large m values, the beam response function $\mathcal{L}_{\ell,m}$ becomes very small, so that all such \mathbf{R}_m matrix would become zero.

The bottom panels of Fig. 4.1.2 show the \mathbf{R}_m matrix for the case of a single EW baseline of $d_{\text{sep}} = 7\text{m}$, with m = 50 (bottom left), m = 100 (bottom center) and m = 250 (bottom right). Here we note that for m = 0 mode the \mathbf{R}_m matrix would be zero, as the response function for this baseline is non-zero for a spot along the line $m = \ell \cos \delta$ (see fig. 4.1.1). For $D_{\text{eff}} = 4.5\text{m}$ and $d_{\text{sep}} = 7\text{m}$, the beam response is negligible out the ℓ -range $50 \leq \ell \leq 350$. Again, for given m within this range, the distribution of ℓ and ℓ' is limited by $m < \ell, \ell' < m/\cos \delta$.

In both cases, we see that R_m matrices have band structures along the diagonal, with width increasing at high m. When the survey area increases and a larger declination band is observed, the width of the band along the diagonal of the response matrices decreases. For a full sky coverage, the response matrix becomes diagonal.

Below, we shall analyse the compressed response matrix **R** which gathers the diagonal terms of \mathbf{R}_m (eq. 3.6.2). As an example, let us consider a 2 × 2 array here, where four dishes of 5 m aperture are arranged on the four corners of a square, with the side length of the square to be 15m. However, the visibility of the two diagonal baselines are related by complex conjugate in the (ℓ, m) space:

 $V_{SE-NW}(\ell,m) = V_{SW-NE}(\ell,-m) = V_{SW-NE}^{\dagger}(\ell,m),$

so in Fig.4.1.3 we will see these two baselines appear to occupy the same region in the (ℓ, m) space.

To show how $\mathbf{R}(\ell, m)$ matrix will look like for the transit observation of a narrow strip along a constant declination line, as would be achieved by a single pointing of the dish array, and show also the effect of observations at different declinations, we plot the \mathbf{R} matrix in logarithmic color scale in fig.4.1.3, for two constant declination scans, one at $\delta = \delta_1 = 35^{\circ}23'$ and another at $\delta = \delta_2 = 59^{\circ}23'$. These two declinations correspond to the edges of the sky region that would be covered by the PAON-4 observations. We can easily see on the figure two set of covered regions, corresponding to the two declinations. Each baseline for each pointing covers one distinct region in the (ℓ, m) space, as expected from the beam shapes discussed above (fig. 4.1.1).

We can distinguish four pieces regions in figure 4.1.3: the wing-shaped regions near the origin $(\ell, m) = (0, 0)$ are derived from the auto-correlation signal, with the two intensive stripe of $m = \ell \cos \delta$ for the two declinations, the $\delta = \delta_1 = 35^{\circ}23'$ on the *outer* side.

The region around $(\ell = 450, m = 0)$ obviously corresponds with the North-South baseline with $d_{\rm sep} = 15$ m. Here the two pointings are largely coincident with each other, except that the $35^{\circ}23'$ one extends further in the *m* direction. We also see that the region spread is $\Delta \ell \sim 150$, so $\Delta \ell / \ell_c \sim D/d_{\rm sep}$. The EW baseline corresponds to the two narrow strips centered at the same ℓ



Figure 4.1.3: The **R** matrix for 4 dishes with 15m baselines and pointing at $35^{\circ}23'$ and $59^{\circ}23'$.

but with $m = \ell \cos \delta$, with some fringes within (i.e. $m < \ell \cos \delta$) region. The diagonal baselines correspond to the region with same m as the EW baseline but larger ℓ . As we noted earlier, in this case the visibilities of the two diagonal baselines are exactly complex conjugates, so in the $\mathbf{R}(\ell, m)$ matrix they occupied the same region.

Having discussed the compressed response matrix for the two scan case, we turn now our attention to the case of the survey of a continuous band from $\delta_1 = 35^{\circ}23'$ to $\delta_2 = 59^{\circ}23'$ with the same array configuration (2 × 2 and $d_{sep} = 15$ m). The corresponding $\mathbf{R}(\ell, m)$ matrices for two observing frequencies, 1420MHz and 1250 MHz, are represented in figure 4.1.4. Here the color scales are linear. With the wider band of sky, the $\mathbf{R}(\ell, m)$ matrix can be regarded as the superposition of the individual narrow strips, the covering of the sky by the individual baselines are still distinctly seen. We also note the region due to autocorrelation shows a "highlighted" region between $m = \ell \cos 35^{\circ}23'$ and $m = \ell \cos 59^{\circ}23'$, but there is also some finite values within the region $m < \ell \cos 59^{\circ}23'$ due to the superposition.

At the two frequencies, the general shape of the **R** matrices are similar, but shifted in position. This is expected, as the (ℓ, m) individual beam positions and extensions vary as $1/\lambda$. With a separation distance $d_{sep} \sim 3D$, there are large uncovered regions in the (ℓ, m) plane, which is a problem for 21cm intensity mapping observations. Indeed, the number of (ℓ, m) modes simultaneously measured at different frequency would be smaller than a configuration covering the plane fully, which would lead to stronger beam frequency dependence and mode mixing.

4.1.2 PAON4 beam and (ℓ, m) plane response

Traditionally, interferometer arrays are employed to achieve high angular resolution, which requires long baselines. However, as shown above, with long baselines there are inevitably *holes* on the (u, v) or (ℓ, m) plane which will not be covered during observations, hence the corresponding Fourier modes are not measured, as demonstrated in Fig. 4.1.4, where the shortest baselines



Figure 4.1.4: The $\mathbf{R}(\ell, m)$ for 4 dishes with 15m baselines observing at 1420 MHz (top) and 1250 MHz (bottom) for the survey of the region with $35^{\circ}23'^{\circ} < \delta < 59^{\circ}23'^{\circ}$.



Figure 4.1.5: Comparison of the PAON-4 synthetic beam(left panel) with that of the compact 2×2 array (center panel) and the D=15.5 m single dish (right panel).

are 15m. For sparse images, e.g. a sky dominated by point sources, good image reconstruction might still be achievable. However, for reconstructing the diffuse intensity distribution such as the 21cm, this will be a major obstacle, especially because at different frequencies ν , the missing or unobserved modes would be different, making it hard to separate the cosmological 21cm signal from the strong continuum foreground. If the baselines are sufficiently short, then at least within certain spatial frequency ranges, the (ℓ, m) plane sampling would be complete, and a better sky reconstruction becomes possible. That's why we shall consider more compact arrays below.

We note that the separation between the PAON-4 dishes are small, only slightly longer than the closest-packed configuration. We also compare it with a compact 2×2 regular array with $d_{sep} = 7m$ and a D=15.5-meter diameter single dish (D_{eff} = 14 m).

For the PAON-4 and a survey composed of 25 constant declination scans, a total of 175 (150 without auto-correlation) beams are used for the map making, compared to 125 (100 without auto-correlation) for the 2×2 case. The survey for the large single dish is composed of 79 constant declination scans, corresponding to a total of 79 beams for the survey.



Figure 4.1.6: Comparison of R matrix for PAON-4 (left), 2×2 (center) and D=15.5 m single dish ($D_{eff} = 14 \text{ m}$) configuration (right)



Figure 4.1.7: The matrix **R** with (top) and without (bottom) auto-correlations for the PAON-4 case, with 1420 MHz (Left), 1250 MHz (Middle) and 1200 MHz (Right) frequencies.

We calculate the beam, or the response to a point source from the full sky reconstruction, as described in section 3.6.1 for the PAON-4, 2×2 and large single dish configurations. The reconstructed beam depends slightly on the source declination, the beams shown here correspond to the central declination of observation, i.e. 47° N. Figure 4.1.5 shows the 2D beams for the PAON-4 case (left panel) compared with the two other configurations: the 2×2 array (center panel) and the D=15.5 m single dish (right panel). One can see that the PAON-4 beam has a hexagonal symmetry, generated from the product of triangular symmetry and reflection symmetry, which is much more circular in shape than the 2×2 (configuration b), and its resolution is also slightly better than the D=15.5 m single dish ($D_{eff} = 14$ m).

Figure 4.1.6 shows the response matrix $\mathbf{R}(\ell, m)$, defined in equation (3.6.2) for PAON-4, compared with the compact 2×2 array and the single dish configurations. In last chapter, we have already discussed the behaviour of the $\mathbf{R}(\ell, m)$ matrix for individual baseline and for individual pointing, as well as the cases of an 2×2 array with longer baselines. With insight gained from that exercise, we can analyse the \mathbf{R} matrices here, which to a good approximation is the linear superposition of different baselines and individual pointings.

The $\mathbf{R}(\ell, m)$ for the single dish is very simple, which has a triangular shape. In the present case it extends to higher ℓ, m , with the maximum determined by $\ell_{\text{max}} = 2\pi D_{\text{eff}}/\lambda \approx 420$ but otherwise similar to the auto-correlation for the small dish discussed before. The triangle is bounded by $m = \cos 35^{\circ}23'\ell \approx 0.83\ell$, and inside the triangle there is an inner boundary at $m = \cos 59^{\circ}23'\ell \approx 0.51\ell$, which is marked by a clear inner boundary, and the $\mathbf{R}(\ell, m)$ is largest within the two boundaries. However, the superposition of the stripes also fill up the region with $m < \cos 59^{\circ}23'\ell \approx 0.51\ell$. This is reasonable, for the modes in this part of (ℓ, m) space are basically modes along the NS direction, for which the information is available with the superposition of many narrow strips along the latitude, but would not be available for a single narrow strip.

Compared with the non-compact 2×2 array with longer (15m) baselines, the compact 2×2 array considered here has shorter baselines, so that the region for the NS and EW baselines overlap with each other and also with the auto-correlation part. The diagonal baselines are also there, with m similar to the EW baseline but with larger ℓ . Thanks to the overlap, now the (ℓ, m) modes up to certain (ℓ_{\max}, m_{\max}) can be measured completely with this array configuration.

For the PAON-4 array, there are six independent baselines, so the coverage in the (ℓ, m) space is more complicated, but generally the regions are better covered, and actually extend to higher ℓ values and larger area in the (ℓ, m) plane. The little hole in the region $(150 < \ell < 250, m \sim 0)$ is due to the lack of a short north-south baseline with $d \sim 6m$, and the uncovered area around $(\ell \sim 400, m \sim 120)$ is due the lack of a baseline close to $d_{NS} \sim d_{EW} \sim 6m$.

In Fig.4.1.7, we plot the **R** matrix for the PAON-4 array with three different frequencies: 1420 MHz, 1250 MHz and 1200 MHz, so as to have a better idea of how the mode reconstruction changes at the different frequencies. The general shape of the covered (ℓ, m) regions are similar, but at lower frequency, due to the lower angular resolution, the covered regions are smaller.

80



Figure 4.1.8: Comparison of the error variance matrix for PAON-4 (left), 2×2 (center) and D=15.5 m single dish ($D_{eff} = 14 \text{ m}$) configuration (right)

4.1.3 PAON4 noise power spectrum and transfer function

The auto-correlation signals are usually not used in interferometric observations. Indeed, the autocorrelation signals are very sensitive to the variation of noise level, which can easily swamp the sky signal. Only for the case of white stationary noise, the noise term in the auto-correlation contribution can be subtracted. For cross-correlations, the average contribution remains zero unless there is a correlated noise source between the two different receivers: $\langle n_i n_i \rangle \propto T_{sys}^2$; $\langle n_i n_j \rangle = 0$, $i \neq j$. But as shown in Fig.4.1.7, without the auto-correlations, low spatial frequency modes are not sampled, the reconstruction can not be done at small (ℓ, m) . Using small separation between the dishes will help reduce the unobserved modes in the low (ℓ, m) region. This is one of the reasons why we use small dishes, and close packed array configurations, which also avoids the incomplete coverage at higher (ℓ, m) .

We will discuss here the computed noise power spectrum and transfer functions for PAON-4 and the two other configurations, for map reconstruction when taking into account auto-correlation visibilities (With-Autocorrelation) or ignoring these visibilities (No-Autocorrelation).

The noise level for visibility time samples are computed according to the equations (3.6.5) and (3.6.6). As stated earlier, we have chosen $\ell_{max} = 1500$ and HEALPix $n_{side} = 512$ for the reconstructed maps shown here. We consider thus that we have $n_t = 2m_{max} = 3000$ visibility time samples over 24 hours. For a total survey duration of $t_{survey} = 175$ days (6 months), each δ – scan would be repeated 7 times, leading to a total integration time of $t_{int} \sim 201.6$ s per time sample. Assuming a system temperature $T_{sys} = 100$ K and $\Delta \nu = 1$ MHz frequency bin width, we obtain a noise level of 9.96 mK for the PAON-4 and 2×2 cases.

For the single dish case, each of the 79 scans would be repeated twice, leading to a total survey duration of 158 days, slightly shorter than the PAON-4 and 2×2 array survey time. Each time sample would have a total integration time of ~ 58 s, leading to a noise level of ~ 18 mK for the single dish case.

Figure 4.1.8 shows the error covariance matrix, defined in equation (3.6.3), for the PAON-4 (left), 2×2 (middle) and 14m single dish (right) survey. In these maps, we have set a very



Figure 4.1.9: Comparison of the transfer function $T(\ell)$ (left panel) and the noise power spectrum $C^{\text{noise}}(\ell)$ (right panel) for PAON-4 array in red, compact 2×2 array in blue and the single dish in black. Solid lines: map-making with auto-correlations, Dashed lines: map-making without auto-correlations.



Figure 4.1.10: Comparison of the transfer function $T(\ell)$ after filtering and masking (left panel) and noise power spectrum $C^{\text{noise}}(\ell)$ after filtering and masking (right panel) for PAON-4 in red, compact 2×2 in blue and single dish in black. Solid lines: map-making with auto-correlations, Dashed lines: map-making without auto-correlations.



Figure 4.1.11: Histograms of reconstructed noise map pixel values for PAON-4. Top panel: pixellation with HEALPix $n_{side}=256$, black curve: the raw map; red curve: filtered map with $\ell_{max}\sim420$. Bottom panel: blue curve: filtered map with $n_{side}=128$ and $\ell_{max}\sim375$; magenta curve: $n_{side}=64$ with $\ell_{max}\sim190$.

large error for all points with no data at all, so that the regions which are well-measured (low noise) are represented by dark color. We can see that there are some similarities in the distribution with the $\mathbf{R}(\ell, m)$ matrix, but unlike the \mathbf{R} matrix distribution, which is fairly smooth, here we see a lot more variations in the distribution, as weights also goes into the number of redundant baselines. The single dish survey is still the simplest, it achieves uniformly low-noise measurement for $\ell < 300$, but the error blows up at $\ell > 350$. Interestingly, there is a fairly large error at the line $m = \ell \cos 59^{\circ}23'$. The (ℓ, m) modes at $m > \ell \cos 59^{\circ}23'$ are measured primarily along the EW direction within the survey region, and those modes with $m < \ell \cos 59^{\circ}23'$ are measured primarily along the remaining the NS direction. However, the modes at $m = \ell \cos 59^{\circ}23'$ have relatively large error.

For the two arrays, the distribution is fairly complicated, though we can see where the **R** matrix is substantial, the measurement error is also relatively small. The region around the center (ℓ, m) value for each baseline is a basin of low measurement error, separated by some watersheds with larger errors. This is easy to see in the case of the 4 baselines of the 2×2 array (recall that the two diagonal baselines coincides each other on the (ℓ, m) map, when restricted to $m \ge 0$ half plane). For the case of the PAON-4 array there are six independent baselines, so the distribution is even more complicated but can also be identified. As in the single dish case, along the watershed at $m = \ell \cos 59^{\circ}23'$ the error is somewhat larger. The interferometers do probe higher ℓ region, up to $\ell \sim 500$ in the case of PAON-4, but there are also some regions with no measurement. The small value at the edge of large ℓ (bottom of the figure) is however an artefact of the computation procedure: as the elements of the **A** matrix goes to zero, the corresponding elements in the pseudo-inverse **B** are also set to zero, and as a result we see the response function and error drop to small values at the edges, but this does not affect our final estimation of the measurement error.

In the left panel of Fig.4.1.9 we plot the power transfer functions computed from the **R** matrix. As expected from the map-making algorithm used here, the transfer function is nearly constant with respect to ℓ , up to some ℓ_{max} between 300 and 400, above which it drops precipitously. The single dish transfer function is very smooth, the two arrays have some small features, the PAON-4 a slightly smoother than the 2×2 array. The value of the plateau region in the transfer function (~ 0.16) is determined by the fraction of mapped sky f_{sky} , which is given by $f_{sky} = \frac{1}{2} (\cos \theta_1 - \cos \theta_2)$, where θ_1 and θ_2 define the declination range of the surveyed region.

The noise power spectrum can be calculated from the covariance matrix, and is plotted in the right panel of Fig.4.1.9. For the single dish configuration noise power spectrum increases smoothly with angular frequency (ℓ), up to $\ell = 400$, at which point it drops as the transfer function vanishes. For the two arrays, without autocorrelations (dashed lines), the noise power spectrum is zero at $\ell \leq 100$ ($\theta \geq 2^{\circ}$). The noise power spectrum for a reconstruction taking into account autocorrelations is also shown, for which the small ℓ power is non-zero. The noise power generally increases with ℓ until ℓ_{max} , where they also vanishes along with the transfer function. Again, the PAON-4 and 2 × 2 configurations present some structures in the noise power spectrum $C^{\text{noise}}(\ell)$ as in the transfer function. However, the $C^{\text{noise}}(\ell)$ curve for PAON-4 is smoother than that of the 2 × 2 (b) configuration.

As discussed in section 3.6.4, additional filtering and masking can be applied after the determination of the sky spherical harmonic coefficient to improve the result. This is illustrated in figure 4.1.10, where the border pixels outside of the band $34^{\circ}23' < \delta < 60^{\circ}23'$ are masked out,

4.2. APPLICATION TO THE TIANLAI DISH ARRAY

and a smoothing filter function $W_2(\ell)$ is applied to suppress some modes with high noises. The transfer function for the 2 × 2 configuration present a dip around $\ell = 150$, due to the high noise modes around the ($\ell = 150, m = 100$) visible in the error variance matrix. This filtering enables us to decrease significantly the noise level in the reconstructed maps, as shown in the map pixel value histograms (Fig.4.1.11).

In Fig. 4.1.11 we compare the reconstructed temperature map pixel value histograms for reconstruction from pure noise visibilities. The HEALPix pixellation scheme is used to define the pixels, in which the total number of the pixels for the whole sky is $12n_{side}^2$, and the angular size of each pixel is $\sim \sqrt{\pi/3} n_{side}^{-1}$ rad. So for $n_{side} = 256$, the pixel size is about $\left(\frac{1}{4}\right)^2 \text{ deg}^2$. The black curve shows the pixel value distribution from the reconstructed map without any additional processing (raw map). To reduce the noise, we may filter out modes with large noises. The narrower red curve shows the distribution of a filtered map, which is obtained by applying the angular plane masking and (ℓ, m) filtering with $\ell_{\rm max} \sim 420$. The half width of the distribution decreases from $146 \,\mathrm{mK}$ to $55 \,\mathrm{mK}$ after the filtering, but at the expense of reduced (ℓ, m) coverage. However, note the pixel size is still smaller than the reconstructed beam width, so the neighbouring pixels may be correlated. In the bottom panel, we show the pixel distribution for larger pixel sizes. The blue curve represents the case of $n_{side}=128$ (with $\ell_{max}\sim375$), and the magenta curve represents the case of $n_{side} = 64$ (with $\ell_{max} \sim 190$) respectively. The angular masking are applied in both cases. The width of the distribution decreases to 45 mK for the $n_{side} = 128$, corresponding to 0.5 degree pixels, and 15 mK for the $(n_{side} = 64)$, for 1 degree pixels. These values are consistent with simple estimates based on the expected noise in visibilities and the reconstructed beam size.

4.2 Application to the Tianlai Dish Array

The method and analysis criteria presented above can also be applied to the Tianlai experiment, and the insight gained from the analysis of the PAON-4 array in the previous section would be useful for the understanding the bigger Tianlai 16-dish array response. As in the PAON-4 case, we also have studied a number of configurations but we will focus here only on two configurations, the regular 4×4 layout and a circular array layout, the latter (circular) being the current Tianlai configuration 16-dish array configuration. The importance of carrying intensity mapping observations with compact interferometric arrays, in order to ensure a complete sampling of the (u, v) plane or (ℓ, m) space have already been discussed in section 4.1.1. We have thus limited our investigations to relatively compact dish array layouts.

Before discussing the beam pattern and map-making capability, we first consider the blockage of the antenna with different separation scale d_{sep} . For simplification, we treat the antennas as circular dishes of 6m diameter placed on the same horizontal plane, aligned in the North-South direction, and then we define the geometric blocking factor, which is the overlapped projected cross section divided by the total area of the dishes. This ignores the effect of diffraction and multiple overlaps, but is easy to compute. The blocking factor as a function of separation scale d_{sep} and zenith angle are shown in Fig.4.2.1 for the regular (left) and circular (right) arrays. The



Figure 4.2.1: The blocking factor for the regular 4×4 configuration (left) and the circular configuration (right) as a function of separation scale d_{sep} and antenna pointing angle with respect to the zenith.



Figure 4.2.2: Comparison of Tianlai 16-dish configuration (circular) with a regular 4×4 array (center). The beam for the Tianlai 16-dish circular configuration, without the autocorrelation signal is shown on the right panel. $4.4 \times 4.4 \text{ deg}^2$ high resolution area extracted from the reconstructed maps, centered on a point source position.

regular array generally have worse blocking as the antennas are aligned in the north south direction, while for the circular array most dishes are not aligned along the NS axis.

Obviously at small zenith angle, i.e. when the dishes are pointed toward the zenith, they will not block each other whatever the d_{sep} value is. As the zenith angle increases, the dishes may be partially blocked by its neighbours to the north or south, though for each dish the blocking is different. For the circular configuration with the minimum 6 m separation, only at zenith angles $< 47^{\circ}$ we can achieve blocking factor < 10%. However, when d_{sep} increases, the same blocking factor can be achieved at larger zenith angles. For example, at $d_{sep} = 9m$, we read from the figure that < 10% blocking factor can be achieved for observations with zenith angle up to 70°, which is more or less the maximum zenith angle observation foreseen with Tianlai. To minimize the ground preparation work, we have tried a few different values of d_{sep} around 9m and calculated the positions of the antennas in the circular array configuration. Finally the value $d_{sep} = 8.8m$, has been chosen, which correspond to the current configuration.

4.2.1 The Tianlai dish array mid-latitude survey

Fig. 4.2.2 shows the reconstructed 2D synthesized beams, i.e. the reconstructed map for a point source, for the rectangular array (Left panel), the circular array with auto-correlation included (Centre panel) and the circular array without including the auto-correlation (Right panel). The rectangular array exhibits a strong cross-shaped grating type pattern in its beam, while for the circular array, the beam is nearly circular-symmetric. The beam formed with and without auto-correlation signals are similar, with only subtle differences on large scales, so that using only the cross-correlation signals would not much affect the observations except on very large angular scales.

We plot in figures.4.2.3 the compressed response matrix $\mathbf{R}(\ell, m)$ (Eq. 3.6.2) for the 16-dish (a) regular 4×4 array (top panels) and (b) circular configuration (bottom panels) and for two frequencies: 1420 MHz (left panels) and 1250 MHz (right panels). As expected, the general shape of the (ℓ, m) space distribution is similar for the two frequencies, but with a wider area covered in the (ℓ, m) space at the higher 1420 MHz frequency. The **R** matrices have a triangular shape with a cut off for $\ell \leq 70$, as here we have used only the cross-correlations and neglected the errorprone auto-correlations. We note that the triangles are bounded by $m \gtrsim \cos 29.10^{\circ}\ell \approx 0.87\ell$ as in the 4-dish cases, and also the watershed along the line $m = \cos 59.10^{\circ}\ell \approx 0.51\ell$, but the overall distribution is much more uniform than the 4-dish cases, as we have many more dishes and baselines for either configurations.

The regular array is very similar to the 2×2 array discussed earlier. Now we have both short baselines in the compact 2×2 array and the long baselines in the non-compact 2×2 array. Again, the "slanted" baselines, i.e. the baselines not in the due NS or due EW direction are symmetric with respect to the NS line, so that each one would have a counter part with the same NS component but inverted EW component, and the visibilities of these pairs would be complex conjugates so that on the $\mathbf{R}(\ell, m)$ matrix they will coincide. In spite of the increase in the number of baselines, the coverage in the (u, v) plane or (ℓ, m) space is not very dense, we can see clearly some grid patterns in the **R** matrix, e.g. at $(\ell, m) = (185, 94), (414, 94), (668, 94), (414, 282), (556, 282), (764, 282)$



Figure 4.2.3: Comparison of the response matrix $\mathbf{R}(\ell, m)$ for the 4 × 4 regular 16 dish configuration (top) and the Tianlai circular 16 dish configuration (bottom) at 1420MHz (left) and 1250MHz (right).



Figure 4.2.4: Reconstructed map for the Tianlai-16 dish configuration (bottom) at 21 cm, and the original 21 cm map (top) from the LAB survey. Modes with $\ell \leq 100$ have been filtered out of the LAB map. Only a band in declination $20^{\circ} < \delta < 70^{\circ}$ is shown, the planned survey area is $30^{\circ} < \delta < 60^{\circ}$ as discussed in the text.

for the 1420MHz case, and the slightly shifted counterparts in the 1250 MHz case.

In the Tianlai circular configuration there are not many due EW or NS baselines, and the outer rings are not symmetric with respect to the NS line either. So the circular configuration of the Tianlai array has a fairly smooth response in (ℓ, m) space. Compared with regular configuration, the circular configuration has larger overall coverage of the (ℓ, m) plane. The regular array coverage extends down to $\ell \sim 100$, but at the large ℓ values ($\ell > 800$), it does not cover all m values, only the region 300 < m < 650 are covered, while smaller and larger values of m are not covered, and even in the covered part we can see the coverage is fairly non-uniform. The circular configuration, on the other hand, shows fairly uniform coverage up to $\ell \sim 950$, as we would expect from its more uniformly distributed baselines. Also, compared with the 1420 MHz case, the covered (ℓ, m) region is smaller for the 1250 MHz case, reflecting the fact that at the lower frequency, the angular resolution is lower. However, it still provides good coverage for a large part of the (ℓ, m) space, so that reliable measurements of the same spherical harmonics modes could be performed at both frequencies.

The bottom panel of Fig. 4.2.4 shows the reconstructed map at 1420.4 MHz derived from simulated Tianlai 16-dish array observation with the configuration and survey strategy defined above. As the auto-correlation visibilities have been ignored, the reconstructed map looks somewhat different from the original LAB map (figure 3.6.1), as large scale structures are not visible. However, if we apply a high pass filter defined in section 3.6.4 to the LAB map:

$$W_3(\ell) = \left(1 + e^{\frac{\ell_B - \ell}{\Delta \ell_B}}\right)^{-1}, \qquad \ell_B = 65, \Delta \ell = 10$$

we obtain a map which is very close to the *No-Autocorrelation* reconstructed map. The high-pass filtered LAB map is shown in the top panel of figure 4.2.4, and we see that the reconstruction



Figure 4.2.5: Comparison of the error variance matrix for Tianlai-16 dish circular configuration (right) with a regular 16 dish 4×4 array (left) at 1420MHz.



Figure 4.2.6: The transfer function for the 16-dish regular (left) and circular (right) configurations. The curves are purple: 1420MHz, red: 1250MHz, green: 1200MHz.

4.2. APPLICATION TO THE TIANLAI DISH ARRAY

quality is very good within the surveyed region, though of course at the boundary, it is degraded.

The survey around the zenith is composed of 31 constant declinations scans. We consider that each δ -scan would be repeated 7 times, corresponding to a total survey duration of $31 \times 7 = 217$ days. The total integration time for each visibility time sample would be $t_{int} \sim 201.6$ s for $n_t = 2m_{max} = 3000$. Using Eq.(3.6.5) and Eq.(3.6.6), and assuming a system temperature $T_{sys} = 100$ K, and $\Delta \nu = 1$ MHz, we obtain a visibility noise level of $\sigma_{noise} \sim 9.96$ mK, similar to the PAON-4 case.

Figure 4.2.5 shows the corresponding error variance matrix (Eq. 3.6.3) in the (ℓ, m) basis for the rectangular and circular array at 1420 MHz. We have already seen in the previous cases that this variance matrix has more non-uniformity than the response **R** matrix. This is also true in the present case, especially for the configuration (a), where we can clearly see *islands* with relatively large errors in the (ℓ, m) space, which are located exactly at the grid patterns of the **R** matrix. However, even in this case, we can see that the errors are much smaller than the 4-dish case, as we have many more baselines. The circular configuration, on the other hand, have a much larger and more uniform *sea* of low error region, though at large ℓ corresponding the longer baselines, the errors become somewhat larger. Notice that we can still see an *island chain* along the $m = 0.51\ell$ line, but not as strongly peaked as in the PAON-4 case.

In Fig.4.2.6, we plot the transfer function $T(\ell)$ (without auto-correlation) for the (a) 16 dish regular array (left) and (b) 16 dish circular array (right), with the purple, red and green curves for 1420MHz, 1250 MHz and 1200 MHz observations respectively. The transfer function is generally flat over the effective ℓ range for the arrays, with some slight wiggles. For the circular array the wiggles are even less pronounced than the regular array. It should also be noted that the wiggling structure shifts with frequency. These wiggling features can affect the BAO power spectrum measurement, and have to be taken into account in the final analysis. If we can determine the transfer function exactly, and the error variance is smaller than the signal, it is still possible to correct for it. It is however highly desirable to select configurations minimising wiggles and non uniformities in the response matrix and transfer function when designing the instrument and the survey strategy.

In Fig. 4.2.7 we plot the noise power spectrum, again without using the auto-correlation data. We have seen that the transfer functions are shifted slightly at different frequencies, here we plot for a single frequency the measured noise power spectrum for the regular (red) and circular (green) array, as well as the filtered spectrum for the regular (orange) and circular (blue) curves.

The histogram distribution of map pixel values reconstructed from pure noise visibilities are shown on Fig.4.2.8 for the 16-dish regular configurations, and Fig.4.2.9 for the 16-dish circular configurations at 1420.4MHz. We have performed map reconstruction varying the map pixel resolution, through the HEALPix n_{side} parameter, with n_{side} = 512 for the top panel and 256 for the bottom panel respectively. The latter has a resolution of ~ 13.7', which is about half of the synthesized beamwidth. For the raw maps, i.e. the reconstructed maps without additional processing, the circular configuration has slightly broader pixel distribution than the regular configuration, due to a wider coverage of the (ℓ , m) space. The temperature variance is ~ 202 mK for the circular array and ~ 159 mK for the regular array. The histograms in red show the pixel value distributions after applying the angular mask and the (ℓ , m) filtering, as shown by the red histograms, the circular configuration pixel noise level (58 mK) fares better than the regular configuration (72 mK).



Figure 4.2.7: Comparison of the noise power spectrum $C^{\text{noise}}(\ell)$. The green and blue curves correspond to the 16-dish circular layout, while red and orange curves correspond to the regular 4×4 array. Blue and orange noise power spectrum have been obtained after filtering in (ℓ, m) space and masking.



Figure 4.2.8: Histograms of map pixel values for the 4×4 regular array. Top panel: $n_{side} = 512$. Black curve: raw map; red curve: filtered map with $\ell_{max} \sim 1100$. Bottom panel: $n_{side} = 256$; blue curve: filtered map with $\ell_{max} \sim 780$; magenta curve: filtered map with $\ell_{max} \sim 420$.



Figure 4.2.9: Same as Fig.4.2.8, but for 16 dish circular array.

To compare with the PAON-4 case, we consider the pixellation with $n_{side} = 256$, which is the same as the highest resolution used for the PAON-4 array, the variance is 52 mK for filtering at $\ell_{max} = 780$ (blue) and 22 mK for $\ell_{max} = 490$ (magenta) for the regular array, and 18mK and 10 mK respectively for the circular array. The circular configuration produce much better result than the PAON-4 for the same pixel size (55 mK), which is expected as the Tianlai dish array have many more dishes.

From these comparisons we see that the circular array have better coverage of the (ℓ, m) space, resulting in a higher reconstruction quality and lower noise power spectrum. That's why we have currently chosen this configuration for the Tianlai array. However, there are also advantages in the regular array. For example, the regular array allows the redundant baseline calibration. The problem of calibration is not considered in the present chapter, we will deal with it in a future study, one should however bear in mind this advantage of the regular array when comparing it with the circular array.

4.2.2 The Tianlai dish array polar cap survey

Up to now, we have mainly discussed the observation of mid-latitude bands of sky. The high latitude or polar cap region is both interesting and have some specialities in its observation. From the perspective of an observer on the ground, the sky rotates around the celestial pole, if one points the telescope to the pole, the same point will be observed during all times, so that much deeper exposure of this small region can be achieved within a relatively short time. It might therefore



Figure 4.2.10: Tianlai 16-dish polar cap survey. Top: the response matrix $\mathbf{R}(\ell, m)$; Bottom: The error covariance matrix.

prove interesting to carry out the observation first in the polar cap region.

We note that in the polar cap region, the beam of the telescope may cover all right ascension angles at the same time. The HEALPix pixelisation might not be the best scheme to represent the polar cap. We have used here higher resolution HEALPix maps with $n_{side} = 1024$ to avoid distorsions near the pole, while keeping the same spherical harmonics $m_{max} = \ell_{max} = 1500$ in the calculation.

For the polar cap, we consider a survey of sky starting from the north celestial pole ($\delta = 90^{\circ}$) and down to the latitude of $\delta = 75^{\circ}$, through a total of 16 scans, each shifted by 1° in declination, so that the sky region with $75^{\circ} < \delta < 90^{\circ}$ is observed. Here we assume a total duration survey comparable with the mid-latitude case, though the latitude range covered is about half of the mid-latitude survey. The actual area is even smaller at such high latitudes ($\sim 600 \text{ deg}^2$), compared to

4.2. APPLICATION TO THE TIANLAI DISH ARRAY



Figure 4.2.11: Left: the input polar cap map after a high pass filtering of $\ell > 100$. Right: The reconstructed map from mock visibilities corresponding to the Tianlai 16-dish polar cap survey. The polar cap region with a radius of 20° is represented, while the lowest declination scan of the survey is at 15° .

~ 7000 deg², so that we expect a much deeper survey, with reconstructed map noise level around ~ 3 times lower for the polar cap survey. As an example, consider a simple plan of making a scan at one declination each day, going through the 16 declinations, and then repeat the whole process 14 times. Each declination will be observed 14 times, twice the number of mid-latitude case, requiring $16 \times 14 = 224$ days, about 7.5 months to carry the full survey. The integration time per sample would then be $t_{int} \sim 403.2s$, and assuming a system temperature of $T_{sys} = 100$ K, and $\Delta \nu = 1$ MHz, we obtain a visibility noise level of $\sigma_{noise} \sim 7$ mK using Eq.(3.6.5) and Eq.(3.6.6), $\sqrt{2}$ times lower than the mid latitude case.

In Fig.4.2.10 we plot the $\mathbf{R}(\ell, m)$ matrix (top) and error covariance matrix (bottom) for the polar region. We see that the $\mathbf{R}(\ell, m)$ is restricted to relatively small m for the polar cap survey, due to high values of observed declinations. However, modes up to $\ell \sim 1200$ are still measured. This smaller range of m does not hamper the reconstruction of the map near the polar region, because here the temperature variations are indeed described by the smaller m modes.

As in the mid-latitude case, without using the auto-correlations, we will miss the small ℓ modes, as a result the reconstructed map will be somewhat different from the original. However, if we run the input map through the high pass filter described above, then we see that the reconstructed map resembles this high-pass filtered map as shown in Fig.4.2.11, except at the borders, beyond the surveyed region at $\delta \leq 15^{\circ}$.

In the left panel of Fig. 4.2.12, we plot the transfer function $T(\ell)$. The red curves are the "raw"



Figure 4.2.12: The transfer function $T(\ell)$ (left panel) and noise power spectrum $C^{\text{noise}}(\ell)$ (right panel) for the Tianlai 16-dish polar cap survey.

results obtained from the **R** matrix directly, with the solid line and dashed line showing the cases with and without using the auto-correlation respectively. The blue curves are the transfer function $T(\ell)$, after filtering and masking, obtained by applying the filter function $W_1(\ell)$ and $W_2(\ell)$ defined above to suppress some high noise modes, and masking out the border pixels outside of the band $0^\circ < \theta < 16^\circ$ from the reconstructed map. After these processing, the wiggles amplitude in the transfer function is reduced, and a very smooth transfer function is obtained. We see that the transfer function still extend to $\ell \sim 1250$, even for low *m*-modes, again showing that the map and power spectrum reconstruction at the polar region can be done as well as in the mid-latitudes, while benefiting from the longer integration time for the polar region.

The noise power spectrum $C^{\text{noise}}(\ell)$ are calculated from the covariance matrix or the maps reconstructed from pure noise visibilities with filtering and masking. Figure 4.2.12 show the obtained $C^{\text{noise}}(\ell)$ curves for the polar cap survey: we see the noise power is reduced thanks to the filtering, and reasonable measurements can be made up to $\ell < 1000$.

Figure 4.2.13 shows the distribution of temperature map pixel values (in Kelvin) reconstructed from pure noise visibilities for the polar cap survey. The black curve in the top panel shows the pixel value distribution from the raw map reconstructed from $\hat{\mathcal{I}}(\ell, m)$ without any additional processing, which has a noise standard deviation $\sigma_{noise} \approx 70$ mK. If we apply the angular mask and (ℓ, m) filtering with $\ell_{max} \sim 1100$, the pixel noise is reduced to 18 mK, as shown by the red histogram. These noise values have been obtained for HEALPix maps with $n_{side} = 512$, corresponding to a pixel resolution of $\sim 0.1 \text{ deg}$. In the bottom panel of Fig.4.2.13 we show the result for larger pixels with $n_{side} = 256$. The blue curve represents the case with filtering at $\ell_{max} \sim 780$, and the magenta curve for the filtering with $\ell_{max} \sim 420$, leading to noise standard deviations $\sigma_{noise} = 7.45$ mK and 4.3 mK respectively. We obtain noise levels significantly lower than the mid-latitude case, and comparing $\sigma_{noise}^{pol.cap}$ values to the corresponding mid-latitude values of 18 mK and 11 mK respectively, we see that the gain in noise level is compatible with the



Figure 4.2.13: Histograms of reconstructed map pixel values for the polar cap survey. Top panel: $n_{side} = 512$. Bottom panel: $n_{side} = 256$. Black curve: raw map; red curve: filtered map with $\ell_{max} \sim 1100$. Bottom panel: $n_{side} = 256$. Blue curve: filtered map with $\ell_{max} \sim 780$; magenta curve: filtered map with $\ell_{max} \sim 420$.



Figure 4.2.14: Noise power spectrum for the Tianlai mid-latitude survey and PAON-4, compared to Galactic 21 cm signal power spectrum. The power spectrum of the LAB map si shown in black, and the red curve corresponds to its extension to higher resolution ($\ell > 750$) using GALFA. The blue curve is the noise power spectrum for the PAON-4 case, and the magenta curve represent the Tianlai 16-dish array mid-latitude survey. The noise power spectrum have been computed without the auto-correlation signals.

expectations from the ratio of the surveyed area, $\propto \sqrt{f_{sky}} \sim \sqrt{\frac{7000 \text{deg}^2}{600 \text{deg}^2}} \sim 3.$

4.2.3 Tianlai 16-dish array sensitivity

Figure 4.2.14 shows the comparison of the Milky May 21-cm power spectrum with the Tianlai 16-dish and PAON-4 noise power spectra. The 21-cm power spectrum is derived from the LAB (for $\ell < 750$) survey and GALFA survey data [83] survey (for $\ell > 750$). We have rescaled the angular power spectra with sky coverage fraction, i.e. what is plotted is $C_{\text{map}}(\ell) * (1/f_{sky})$, where $C_{\text{map}}(\ell)$ is the map raw, uncorrected angular power spectrum. The blue curve in Fig. 4.2.14 is the noise power spectrum for PAON-4, and the magenta curve is the noise power spectrum of the Tianlai 16-dish circular array. We see that the noise power spectrum for both PAON-4 and Tianlai 16-dish array are well below the Galactic HI power, so both should be able to measure the Galactic HI without difficulty.

Both the PAON-4 and Tianlai-16 dish array are small prototype arrays, their sensitivities are not sufficient to detect the neutral hydrogen in the large scale structure within reasonable time. In order to detect the latter, arrays with many more elements are required. To see this, we compare the noise angular power spectrum with the expected signal from the large scale structure $C^{\text{sig}}(\ell) = C^{21}(\ell) \times T(\ell)$ at 1050 MHz (z = 0.35), taking into account the transfer function $T(\ell)$. We have assumed a global neutral hydrogen relative density of $\Omega_{\text{HI}}b = 0.62 \times 10^{-3}$ [99] with bias factor b = 1 to compute the expected cosmological 21-cm signal $C^{21}(\ell)$. To simplify the computation



Figure 4.2.15: The expected angular noise power spectra $C^{\text{noise}}(\ell)$ at 1050 MHz for a 2 year survey of a 20° band in declination for square arrays with 16 (black), 64 (brown), 144 (yellow), 256 (blue) dishes with $T_{sys} = 100$ K. The green curve correspond to the 256 dish array with $T_{sys} = 50$ K and the red curve shows the expected cosmological signal at 1050 MHz, filtered by the instrument response $C^{21}(\ell) \times T(\ell)$.

and also make the comparison easier, we consider several regular square arrays, with 16, 64, 144 and 256 dishes. In Fig.4.2.15 we plot the forecasted noise angular power spectra for these four configurations, respectively 4×4 , 8×8 , 12×12 , and 16×16 D=6-m dish arrays. We have considered a survey covering a 20° band in declination ($35^{\circ}10' < \delta < 53^{\circ}10'$), composed of 13 constant declination scans, each shifted by 1.5 degrees, with a longer total survey duration (2 years). We have also shown the $C^{\text{noise}}(\ell)$ for 16×16 array with a $T_{sys} = 50$ K, which could be considered as the design target. We see that the noise angular power spectra drops steadily with the array size, scaling roughly as $1/N_{\text{dishes}}$. Note that this is the angular power spectrum, and only for a single frequency bin. The signal will be further boosted by combining different frequency bins, or computing a 3D power spectrum, noise being uncorrelated between frequency bins, while the LSS signal is correlated. However, the presence of foregrounds due to the Galactic synchrotron and the radio sources introduces correlation along the frequency axis or redshift, making the 21-cm LSS signal extraction much more challenging.

However, the Tianlai 16-dish array survey should be sensitive enough to detect extra galactic HI clumps. The total 21-cm powe r flux Φ_{21} received on Earth from an atomic hydrogen clump with mass $M_{\rm HI}$, at a luminosity distance d_L can be written as [13]:

$$\Phi_{21} \simeq 20 \times 10^{-20} \frac{M_{\rm HI}}{10^9 M_{\odot}} \left(\frac{1 {\rm Mpc}}{d_L}\right)^2 ~({\rm W/m^2})$$

Assuming that the clump 21-cm emission frequency dispersion is below 1 MHz ($\Delta v \leq 200 \text{ km/s}$), we can convert the power flux Φ_{21} into temperature excess $\Delta T_{21}^{\text{pix}}$ in map pixels covering a solid angle $\delta \Omega \simeq 0.25^2 \text{ deg}^2$, corresponding to $n_{\text{side}} = 256$, $\ell_{\text{max}} \simeq 800$ and $\Delta \nu = 1$ MHz in fre-



Figure 4.3.1: Regular configuration of the cylinder array.

quency:

$$\Phi_{21} \times \frac{\lambda^2}{4 \,\delta\Omega} \sim k_B \,\Delta T_{21}^{\text{pix}} \,\Delta\nu \; ; \qquad k_B = 1.38 \times 10^{-23} \,\text{J/K}$$
$$\Delta T_{21}^{\text{pix}} \sim \Phi_{21} \times \frac{\lambda^2}{4 \,\delta\Omega} \times \frac{1}{k_B \Delta\nu} \sim 0.4 \times 10^{20} \,\Phi_{21} \; (\text{K})$$

We can then write the excess temperature $\Delta T_{21}^{\rm pix}$ due to an HI clump in $\sim 0.25^2 \, \rm deg^2 \times 1 MHz$ pixels as:

$$\Delta T_{21}^{\text{pix}} \sim 8 \times \left(\frac{M_{\text{HI}}}{10^9 M_{\odot}}\right) \left(\frac{1 \text{Mpc}}{d_L}\right)^2 \quad \text{K}$$
 (4.2.1)

The Tianlai 16-dish polar cap survey should reach a noise level of $\sigma_{\text{noise}} \sim 7.5 \,\mathrm{mK}$ for map pixels $\sim 0.25^2 \,\mathrm{deg}^2 \times 1 \mathrm{MHz}$ and $\ell_{\max} \simeq 800$, compared to $\sim 20 \,\mathrm{mK}$ for the Tianlai mid-latitude survey. If we consider a $3\sigma \simeq 25 \,\mathrm{mK}$ detection threshold, we see that HI clumps with masses $\sim 3 \times 10^8 M_{\odot}$ would be detected up to $d_L \lesssim 10 \,\mathrm{Mpc}$ or $\sim 3 \times 10^9 M_{\odot}$ up to $d_L \lesssim 30 \,\mathrm{Mpc}$. Based on the HIPASS survey [119] and ALFALFA survey [72] results, the HI mass function is about $dn/d \ln M_{\mathrm{HI}} \sim 10^{-1.4} (\,\mathrm{Mpc}/h)^{-3}$ and fairly flat in this mass range, so we estimate that the Tianlai-16 dish array should be able to detect $\sim 10^2$ such clumps in a survey covering $f_{\mathrm{sky}} \sim 16\%$ of the sky in the polar cap area.

4.3 Application to the Tianlai Cylinder array

For the Tianlai cylinder array, the simplest arrangement of the existing 96 feeds is to have 32 feeds on each cylinder, regularly spaced so that on each cylinder the feeds form a uniformly spaced linear array. Two such configurations are considered here:

4.4. THE REGULAR ARRAY CONFIGURATION

- Regular 1. The feed spacing is taken to be $d_{sep} = 0.4$ m, which is about one wavelength at the observation frequency of 750 MHz. In this configuration, the feeds occupies only 12.4 m of the total 40 m length of the cylinder, as as shown on Fig4.3.1.
- Regular 2. The feed spacing is taken to be $d_{sep} = 0.8m$, about twice the wavelength at the cylinder,

One may also consider configurations with irregular positioning of the feeds to reduce grating lobes. In this chapter we consider a very simple extension: on each cylinder the feeds still forms a uniform linear array, but the number of feeds and hence the spacing of the array is different on each cylinder. We have total 96 feeds at the present time. Marking the cylinders from East to West as cylinder1, cylinder2 and cylinder3 respectively, we consider the following configurations:

- Irregular 1. This is the first irregular cylinder array with number of feeds on each cylinder 31, 32 and 33 respectively. The feeds occupy 12.4 m along NS direction on each cylinder. The feed spacing would be $d_{\text{sep}} = 0.413m$ for cylinder1, $d_{\text{sep}} = 0.4m$ for cylinder2 and $d_{\text{sep}} = 0.388m$ for cylinder3.
- Irregular 2. This is the second irregular cylinder array with number of feeds on each cylinder 31, 32 and 33 respectively, but the feeds occupy 24.8 m along NS direction on each cylinder. The feed spacing would be $d_{sep} = 0.827m$ for cylinder1, $d_{sep} = 0.8m$ for cylinder2 and $d_{sep} = 0.775m$ for cylinder3 for this

To simulate the map making process, we use an input map based on the Global Sky Model (GSM) [29], shown in Fig. 4.3.2. The map is obviously dominated by the radiation from the galactic plane, which is mostly synchrotron emission from galactic cosmic ray electrons. For the computations carried out in this work, we have used HEALPix [45] to pixellate the celestial sphere, with $n_{side} = 512$. In our spherical harmonics transformation we take $\ell_{max} = 1500$, which is sufficient for the angular resolution of Tianlai pathfinder cylinder array. figures

4.4 The regular array configuration

The primary beams for each feed on the cylinders is narrow in the East-West (EW) direction and wide in the North-South (NS) direction, as determined by the cylinder reflector curvatures. We model the primary beam of a single feed on associated with a cylindrical reflector as

$$D(\alpha,\beta) \propto \frac{\sin(\alpha \pi (L_x/\lambda))}{\alpha \pi (L_x/\lambda)} \frac{\sin(\beta \pi (L_y/\lambda))}{\beta \pi (L_y/\lambda)}$$
(4.4.1)

where (α, β) are the two angles with respect to the feed axis, along the EW and NS planes respectively. λ is the wavelength, L_x and L_y are the effective feed sizes along the EW and NS planes. We take $L_y = 0.3$ m for the Tianlai cylinder feeds, and $L_x = 13.5$ m corresponding to an illumination efficiency of 0.9 for the feed on the 15m wide cylinder. These parameters gives a beam width $\sim 100^\circ$ in the North-South direction, and $\sim 2^\circ$ in the East-West direction at 750 MHz, The actual



GSM map @ 750 MHz (Temperature in Kelvin)

Figure 4.3.2: GSM (Global Sky Model) map at 750 MHz, used as the true sky for the reconstruction with the Tianlai cylinder configuration.

values will be obtained by fitting the real observational data, these are heuristic values but should be sufficient for our estimations here. The primary beam is shown as the left panel of Fig. 4.4.1.

For uniformly spaced linear arrays, grating lobes appear when the spacing is larger than half wavelength $(d_{sep} > \lambda/2)$. This is because the phase factor $\exp(i2\pi d_{sep}\sin\theta/\lambda)$ is periodic with respect to $\sin\theta$, and when $d_{sep} > \lambda/2$ the maximum appears more than once. We show the synthetic beam for the regular case 1 and regular case 2 in the central and right panels respectively in Fig. 4.4.1. These are obtained by making the full synthesis of a point source image located at the latitude of the array, i.e. $44^{\circ}10'$. As we can see in the figure, there are strong grating lobes along the NS direction in the synthetic beams. The position of the *n*th order grating lobe is $\sim n\lambda/d_{sep}$. At 750MHz, the positions are $\pm 57.3^{\circ}$ for the regular case 1 ($d_{sep} = 0.4$ m) and $\pm 28.6^{\circ}$ for the regular case 2 ($d_{sep} = 0.8$ m). In addition, there are also primary beam side lobes along both the NS and EW direction, these are less prominent in their level and has smaller periods.

To have a better understanding of the synthetic beams in the spherical harmonic space, let us consider the beams of a single pair of receivers. In Fig.4.4.2 we show the beam patterns function for four cases: the auto-correlation (top left), and the cross-correlations for a due EW baseline between two cylinders (top right), for a due NS baseline (bottom left), and a SE-NW baseline (bottom right). By definition, only the region $-\ell < m < \ell$ has valid values. In the dish case, the auto-correlation covers a triangular region with the top at the origin $(\ell, m) = (0, 0)$, two sides and extends along $m = \pm \ell cos\delta$ where δ is declination of the observation, and up to $\ell_{max} = 2\pi D/\lambda$ where D is the effective aperture. The auto-correlation in the cylinder case is very different, taking up a butterfly shape. This is because the cylinder primary beam is asymmetric in the NS and EW direction. As described in Eq. (4.4.1), along the NS direction which corresponds to $m \sim 0$ the primary beam has very low resolution, while along the EW direction the cylinder primary beam is



Figure 4.4.1: The primary beam (left) and synthetic beams for the regular 1 (center) and regular 2 (right) configurations.



Figure 4.4.2: The beam patterns in spherical harmonics $\mathcal{L}_{\ell,m}$ with size $L_x = 13.5$ m, $L_y = 0.3$ m and centered at latitude 44.15°. Top Left: auto-correlation of a feed; Top Right: cross-correlation for a EW baseline with $d_{sep} = 15$ m; Bottom Left: cross-correlation beam for a NS baseline with $d_{sep} = 12$ m; Bottom Right: cross-correlation for a SE-NW baseline with $(\Delta x, \Delta y) = (15m, 12m)$.

about ~ 2° at 750 MHz, which corresponds to $\ell \sim 2\pi L_x/\lambda \sim 210$. Indeed, the figure shows that the auto-correlation function extends substantially to $\ell \sim 210$ along the two wings. Also, since the cylinder has almost the whole observable sky in its field of view, which includes the equator, the $m = \pm \ell$ is saturated.

For the cross-correlations, the beam pattern centers at $(\ell, m) \sim (2\pi |\vec{u}|, 2\pi u)$ as expected, where $\vec{u} \equiv (u, v, w) = (b_x, b_y, b_z)/\lambda$. So the EW baseline is centered near $m \sim \ell$, while NS centered near m = 0, with $\ell \sim 2\pi b/\lambda$. Note that here we are plotting only positive part of the baseline in one direction, so for the EW baseline the beam is on the m > 0 side. If we are to plot the reverse direction, it would appear on the symmetric position at m < 0.

Figure 4.4.3 shows the response matrix $\mathbf{R}(\ell, m)$ for the two regular configurations at frequency 750MHz. In last section, we noted that for each baseline the \mathbf{R} matrix has a certain distribution centered at $(\ell, m) = 2\pi b/\lambda$, where b is the baseline length. The m position depends on both the EW component of the baseline and the declination of the strip to be observed. For an array with many baselines the \mathbf{R} matrix are well described by the superposition of these individual baselines. For the cylinder array, the field of view is not limited to a narrow strip, but a hemisphere or even larger spherical zone. As such, the cylinder baseline would only be bounded by $m = \ell$. In the cylinder case, the \mathbf{R} matrix at m = 0 is significant up to $\ell \sim 190(380)$ for the Regular 1 (Regular 2) case, which corresponds to the modes probed by the maximum baseline along one cylinder. The longest baseline of the array is however the diagonal ones, i.e. the baselines from North/South end of the East cylinder to the South/North end of the West cylinder, so the \mathbf{R} matrix distributes mostly



Figure 4.4.3: Comparison of the R matrix for regular 1(left) and regular2 (right) configurations.



Figure 4.4.4: Comparison of the error variance matrix for regular 1(left) and irregular2 (right) configurations.

on a band along $m = \ell$, with some extension to higher ℓ in the region between m = 0 and $m = \ell$, forming a shark fin shape. The region near $m \sim 0$ is limited to relatively small ℓ due to the fact that our NS baselines are shorter, especially for the Regular 1 case. Future extensions which fill up the remaining part of the cylinder would help improve the m = 0 region. In the Regular 2 case, larger part of the (ℓ, m) space are covered than the Regular 1 case, but here the receivers are spread more widely, reducing the density of the baseline coverage, so here there are more apparent non-uniformity, as shown by the vertical stripes at m = 120 and 350. These can be understood as follows: as shown in Fig.4.4.2, each baseline is sensitive to some part of the (ℓ, m) space. The part of (ℓ, m) space which is not covered by baselines in the array would not be well reconstructed. As the cylinder array are aligned along the three cylinders, we can expect that the m value centered at 0, 235, 470 will be covered, while regions between these, centered at m = 120 and 350 are not well covered and may have large errors. Furthermore, if looking carefully, some fringes near m = 0 can also be seen, which may be due to the grating lobes.



Figure 4.4.5: Reconstructed sky map for the Tianlai cylinder configuration at 750 MHz. Top: Regular 2 configuration; Bottom: Regular 1 configuration. The input map is the GSM map at 750 MHz.

106

4.5. THE IRREGULAR ARRAY CONFIGURATION

Figure 4.4.4 shows the corresponding error covariance matrix in the (ℓ, m) basis at 750 MHz. For the Regular 2 case the error is particularly large, but even for the Regular 1 case, the errors are also relatively large at these m values. The error values at other regions are relatively small. Additionally, in the Regular 2 case, near m = 0 there is relatively large error and also the error shows some rapid modulation in ℓ . These fringes are similar to the ones appeared in the **R** matrix at the same positions, and are due to the strong grating lobes.

In Fig.4.4.5 we show the reconstructed map at 750MHz derived from simulated regular cylinder array observation ignoring the instrument noise. The top figure use the Regular 1 configuration, and the bottom figure use the Regular 2 configuration. Comparing with the original map Fig. 4.3.2, there are spurious features appearing in the reconstructed map. This is very obvious for the Regular 2 case, and also present in the Regular 1 case (e.g. the bright spot at $(270^\circ, 54^\circ)$). These are produced by the grating lobes of the brighter sources such as the Galactic plane and strong point sources, and the Regular 2 case is worse than the Regular case 1. Because of such spurious features, one can not use array with such configurations to make reliable sky survey.

4.5 The irregular array configuration

As we saw in the last section, spurious images appeared in the reconstructed maps of the regular array due to the presence of grating lobes. To avoid this problem, one could adopt spacings less than half wavelength, or employ non-uniform spacing in the linear array. However, at the wavelength of our observation, it is not practical to have spacing less than half wavelength. There are many possible non-uniform spacing schemes, here we choose a very simple one: adopting slightly different spacings on the three different cylinders. So we take the same total length on the three cylinders, but place 31, 32, and 33 feeds on each cylinder, so that the unit separations are different in each case. We choose the same two total lengths as the regular cases described in last section. So for the Irregular 1 case, the basic spacings are $d_{sep} = 0.413, 0.4, 0.388m$ for the three cylinders respectively, with a total length of 12.4m; for the Irregular 2 case, the basic spacings are 0.827, 0.8 and 0.775 m respectively, with a total length of 24.8m.

There are still some degeneracies in the north-south baseline. For instance, there are 30, 31, 32 instances of $d_{sep} = 0.413, 0.4, 0.388$ m NS baselines in the Irregular 1 configuration, respectively. Nevertheless, for the whole array there are NS baselines of different lengths. The slightly different positioning of the receivers also creates baselines which deviates from the EW direction to different degrees. The whole set up allows wider and more uniform coverage on the (ℓ, m) -plane.

Fig. 4.5.1 shows the synthesized beam for the two irregular cases. Here we see that the level of grating lobes is greatly reduced. Whereas in Fig.4.4.1 we can see clearly the sharp grating lobes at 28° for Regular 2 configuration and at 57° for both Regular 1 and Regular 2 configurations, in Fig. 4.5.1 at these angles the lobes are barely visible. Of course, there are still the primary beam side lobes, but these are generally much smaller. Here we note that the Irregular 2 lobes are weaker than the Irregular 1 lobes.

We plot in Figure 4.5.2 the compressed response matrix $\mathbf{R}(\ell, m)$ for the Irregular1 (left) and Irregular2 (right) configurations at 750MHz. As expected, the general shapes of the (ℓ, m) space


Figure 4.5.1: The synthetic beam for Tianlai cylinder irregular configuration.



Figure 4.5.2: Comparison of the R matrix for the Irregular 1 (left) and Irregular2 (right) configurations.



Figure 4.5.3: Comparison of the error variance matrix for irregular configurations.

distribution are similar for the two cases, but with a wider area covered in the (ℓ, m) space for the Irregular 2 configuration due to the larger array sizes. The broad outline of the shapes in this figure are also similar to those in Fig.4.4.3, but here the distribution is more smooth and uniform due to the more widely spread-out (ℓ, m) coverage in the irregular configurations. The features at m = 120 and 350 in Irregular 2 configuration are much less prominent than in the Regular 2 case.

Figure 4.5.3 shows the corresponding error covariance matrix in the (ℓ, m) basis. Here the regions of larer error are spread out more widely, but the error value at the maximum is much reduced when compared with the regular configurations. The Irregular 1 case has smaller errors than the Irregular 2 case, as the baselines are more concentrated in the former case which helps reducing the errors.

Figure 4.5.4 shows the simulated reconstruction map at 750 MHz with the Irregular 1 (top) and Irregular 2 (bottom) configurations. We can see that in both cases, the reconstruction works relatively well, the spurious features shown in Fig.4.4.5 are absent in these figures, and most features of the original map are well produced. There are still some regions where the reconstruction shows some artefacts, e.g. the stripes at $(350^\circ, 60^\circ)$ and $(190^\circ, 12^\circ)$ in the Irregular 1 map, and the stripes south of the equator in the Irregular 2 map. However, the overall quality for the two maps are good.

In Figure 4.5.5, we plot the power spectrum transfer functions $T(\ell)$ (left panel) and the noise power spectrum (right panel) for the Irregular 1 and Irregular 2 configurations. Here we have masked out the border pixels outside the band $0^{\circ} < \theta < 105^{\circ}$ which are not well constructed, and suppressed (ℓ, m) modes with large errors by applying a weight proportional $\sigma_{\mathcal{I}}^{-2}(\ell, m)$ to all modes which have error larger than $K\sigma_{min}^2$, where σ_{min} is the minumum value of the noise covariance matrix, and for the threshold value we choose K = 50. The transfer function decreases toward higher ℓ , but it is generally smooth, though there are curvatures at certain ℓ s. The Irregular 1 configuration has higher response at lower ℓ , but decreases more rapidly at higher ℓ as expected, as its baselines are concentrated in smaller regions and sensitive more to the larger angular scales. For the noise power spectrum, we see that the Irregular 1 configuration achieved lower noise power than the Irregular 2 configuration. In both cases the noise power spectrum show several peaks and



Figure 4.5.4: Reconstructed sky map for the Irregular 1 (top) and Irregular 2 (bottom) configurations at 750 MHz.



Figure 4.5.5: Comparison of the transfer function $T(\ell)$ (left panel) and the noise power spectrum $C^{\text{noise}}(\ell)$ (right panel) for the Irregular 1 and Irregular 2 configurations.

4.5. THE IRREGULAR ARRAY CONFIGURATION

troughs, which are due to the different density of baselines on the (ℓ, m) plane. We also draw the expected large scale structure 21cm signal power on the same plot, where we assume cosmology from [90], and for neutral hydrogen we adopt $\Omega_{HI}b = 0.62 \times 10^{-3}$ [99]. The 21cm signal is only a few times the noise. Note that this is for the detection at a single frequency, we will have more frequency data, but at the same time there are also complications of foreground removal and calibration, which are beyond the scope of the present work. Considering these factors, we see that detecting the 21cm signal would be a great challenge.

112 CHAPTER 4. APPLICATION TO DIFFERENT DENSE CONFIGURATIONS

Chapter 5

PAON-4 description and first maps

5.1 PAON-4 detailed description

Since 2007, le Laboratoire de l'accélérateur linéaire (LAL, CNRS / Université Paris Sud), the Research Institute on the Fundamental Laws of the Universe (Irfu, CEA) and three laboratories of l'Observatoire de Paris (GEPI, LESIA and USN), collaborate within the BAORadio project to develop the methods and techniques necessary for the cosmological readings of the HI line.

5.1.1 PAON-2

The PAON (PAraboles à l'Observatoire de Nançay) project began with a version with two parabolas of three meters in diameter (PAON2), which is shown in Fig5.1.1. This small interferometer worked for four months from the end of September 2012 and allowed to qualify the site and study the noise of the antennas through observations of radio sources such as Cygnus A.

The 23^{rd} Nov. 2012 an observation of CygA transit has been performed by a rather long data acquisition. Both cross-correlations between the aligned polarizations are shown on Fig5.1.2. The run in total represents 4h (x-axis). The 250MHz total band is along the y-axis of each image starting on top-left corner at 1250MHz. The main CygA fringes are localized roughly in the bin range [500, 1000] of the image while we see also human induced signals before and after the transit concentrated around 1275 ± 10 MHz (ie. the franges spacing do correspond to Sky rotation).

After 4 months of operations, PAON2 has reached its main objectives. Observations of bright sources such as the sun of Cygnus A allowed for fringe fitting and antenna lobe determination. A first measurement of $T_{\text{sys}/\eta}$ has also been made, where T_{sys} is the system noise level or temperature, and η is the antenna efficiency, the ratio of the effective area to the geometric area of the reflector. This is shown in figure 5.1.3, in reasonable agreement with expectations.

5.1.2 PAON-4

L'interféromètre PAraboles à l'Observatoire de Nançay (PAON), inaugurated in April 2014, is one of the components of this project which aims to evaluate the viability of an interferometer using



Figure 5.1.1: PAON2.



Figure 5.1.2: Cross-correlations between aligned polarizations during the observation of CygA transit.



Figure 5.1.3: Cross-correlations between aligned polarizations during the observation of CygA transit.

small parabolas about five meters in diameter. The instrument envisaged for cosmological surveys could indeed be a dense network, composed of a few hundred antennas of this type, spread over a surface of one hectare.

In mid-2013, the collaboration decided to design four new parabolas which constitute the current version of the instrument, PAON4. Funding was largely provided by the LAL and the Observatoire de Paris; The mechanics is the responsibility of the GEPI, the frontal electronics was developed by the Irfu while the LAL (Serdi) was in charge of the digital part and the acquisition.

The PAON-4 array is a small wide band test interferometer (L-band, 1250-1500 MHz) featuring four 5-meter diameter antenna, installed at the Nancay radio observatory in France (located at the latitude of $47^{\circ}22'55.1$ " North and the longitude of $2^{\circ}11'58.7$ " East). PAON-4 has been designed and built within the BAORadio ¹ project in France. We have studied several configurations for PAON-4 shown in Fig5.1.4, from which we actually adopte configuration 4. Three dishes are arranged at the vertices of an equilateral triangle with 12 m sides, one of its side is along the exact North-South line. We note that the first dish is inside the triangle, with its center 6 meter away from the West vertex. The second dish is the West vertex one, the third dish is the South-West vertex one, and the fourth dish is the North-West vertex one. In addition to the auto-correlation signal, the PAON-4 configuration has 6 different baselines. There is no redundancy in baselines and visibilities, except for the auto-correlation signal. It has a total collection area of ~ $75m^2$ and 4 dual polarisation receivers. The dish pointing can be changed in declination through computer controlled electric jacks, and can point up to 35° toward the South and 13° toward the North with respect to the vertical axis. PAON-4 can observe the sky in the declination range $12^{\circ} \lesssim \delta \lesssim 60^{\circ}$.

¹https://groups.lal.in2p3.fr/bao21cm/



Figure 5.1.4: PAON-4 interferometer configuration



 $\begin{array}{l} \mbox{Synchrotron map @ 400 MHz - Eq. Coordinates (ra,dec)} \\ \mbox{PAON4 accessible sky region} \\ \mbox{38 S < Elevation < 15 N} \rightarrow 10 < \delta < 60 \ \mbox{at Nançay} \end{array}$

Figure 5.1.5: The PAON-4 accessible sky area.



Figure 5.1.6: DIGFFT board (4 analog channels sampled at 500 MSample/s with 8 bits dynamic range) capable of digital filtering (FFT) developed for the BAORadio project.

The figure 5.1.5 shows the PAON-4 accessible sky area in the synchrotron map at 400MHz where the other area is shadowed.

The PAON-4 electronic and acquisition system consists of the following main components:

- Two Low Noise Amplifiers (LNA), one for each polarisation, located on each feed.
- Due to the transmission loss of the RF signals in the cables, a second stage amplifier is added, located on each antenna mount.
- An analog electronic box with Local Oscillator and mixer, located on the central antenna mount will shift the received RF frequency band 1250-1500MHz to the lower frequency 0-250MHz band.
- Analog signals in the band 0-250 MHz are transmitted from the central antenna electronic box to the Digital electronic crate hosting the digitisers (ADC boards) in the Embrace container over coax cables, with lengths around 50 meters. Please see figures 5.1.6.
- ADC boards sample the input signal at 500 MHz, with 8bits dynamic range, to covering 250 MHz band. They can output either the raw waveforms, or the fourier coefficients, after performing FFT on the FPGA. Each ADC board can handle up to 4 analog RF signals, performing an FFT on each digitization frame of 8192 time samples, sampled at 500 MHz or every 2 ns, leading to 4096 frequency components.
- Digital streams of raw waveforms or FFT coefficients are transferred on optical fibers to the acquisition computer cluster, which is located in the Nançay computer building, at a distance of about 300 meters.



CABLAGE du 8 Fevrier 2017 (carte ADC en mode RAW)

Figure 5.1.7: The current PAON-4 cabling scheme (figure provided by Claude Pailler).

• The mixer and ADC board configuration and the antenna pointing are remotely controlled through a computer (PC) located in the Embrace container. This PC is connected through the network to the acquisition cluster, more specifically to its master controller node. Each antenna is equipped with a micro-controller to handle the mechanical motion along the elevation axis. These micro-controllers are also accessible through ethernet.

In the initial PAON-4 setup, the Acquisition computer cluster which runs a software correlator had 5 nodes. Two front-end nodes (bao8/9) receive digital streams from ADC boards over 4 fibers, which carried each waveforms or FFT coefficients for 2 analog signals, 8 in total corresponding to the 2 polarisations of each feed. Two nodes (bao5/6) were dedicated to the computation of visibilities. Each front-end machine divides the full 250 MHz spectral band in several sub bands, after FFT on the ADC boards, or on the front-end nodes on signal waveforms, dispatching each sub band to one of the correlation computer nodes. In the current configuration, the full 250 MHz is divided into two 125 MHz sub-bands, which are processed by the two correlation computing nodes. A master node (bao3) controls the overall acquisition process.

The acquisition cluster has been upgraded early 2017 and has now four front-end nodes, receiving each data from a single optical fibre (digital streams of 2 analog signals). Each of the front-end computers can handle raw signal waveforms for the two analog signals, performing software FFT and dispatching Fourier coefficients to correlation computing nodes. The current PAON-4 cabling scheme, up to the front-end nodes is represented on figure 5.1.7.

Correspondance des lignes (0 à 35) de la matrice complexe mvis_Obj et des polarisations corrélées								
H: polarisation horizontale V: polarisation verticale (ex : matrice ligne 13 : 2H3V cross-corrélation entre polarisations horizontale antenne 2 et verticale antenne 3)								
(1,1)1H1H 🤅								
(1,2)1H2H 1	(2,2)2H2H 8							
(1,3)1H3H 2	(2,3)2H3H 9	(з,з)знзн 15						
(1,4)1H4H 3	(2,4)2H4H 10	(3,4)3H4H 16	(4,4)4H4H 21					
(1,5)1H1V 4	(2,5)2H1V 11	(3,5)3H1V 17	(4,5)4H1V 22	(5,5)1V1V 26				
(1,6)1H2V 5	(2,6)2H2V 12	(3,6)3H2V 18	(4,6)4H2V 23	(5,6)1V2V 27	(6,6)2V2V 30			
(1,7)1H3V 6	(2,7)2H3V 13	(3,7)3H3V 19	(4,7)4H3V 24	(5,7)1V3V 28	(6,7)2V3V 31	(7,7)3V3V 33		
(1,8)1H4V 7	(2,8)2H4V 14	(3,8)3H4V 20	(4,8)4H4V 25	(5,8)1V4V 29	(6,8)2V4V 32	(7,8)3V4V 34	(8,8)4V4V 35	

Auto et cross corrélations PAON4

Figure 5.2.1: The current arrangement of visibilities over the matrix rows (*Table provided by Claude Pailler*).

5.2 PAON-4 observations and data

The PAON-4 observations are carried in constant declination transit mode. The antennae are pointed toward a common declination direction and data us acquired over a certain duration, usually several hours when scanning around a source, or full east-west (EW) 24 hours scans.

36 visibilities are computed from the 8 RF signals which correspond to the two polarisation directions, denoted H and V, for the four feeds. The eight signals are thus identified by the antenna number and the polarisation direction: 1H, 1V, 2H, 2V ... The set of 36 visibilities include the 8 auto-correlations and 28 cross correlations. There are 6 different, non redundant baselines in PAON4, leading to 6 HH, 6V V and 16 HV cross-correlations. The cross-correlations are save as visibility matrices, with 36 rows and 4096 columns corresponding the the frequencies spanning the 1250-1500 MHz band, leading 61 KHz maximum frequency resolution. The current arrangement of visibilities over the matrix rows are shown in figure 5.2.1.

PAON-4 visibility matrices are currently saved either as FITS files (a single matrix per file), or the SOPHYA PPF format, single or multiple matrix per file. Each matrix has associated time

- DATEOBS = 2015-03-12T10:55:0.0 (T=string) / Date, Time corresponding to TimeTagFirst
- DELTIME = 0.9901946160000002743 (T=double) / visib cumul time (seconds)
- FirstFC = 0 (T=long int) / First FrameCounter
- FirstTT = 0 (T=long int) / First TimeTag
- LastFC = 6118 (T=long int) / Last FrameCounter
- LastTT = 123774327 (T=long int) / Last TimeTag
- MeanFC = 3216.5027343749998181 (T=double) / Mean FrameCounter
- MeanTT = 63334551.396484375 (T=double) / Mean TimeTag
- NPAQSUM = 5120 (T=long int) / Number of paquets summed

Figure 5.2.2: The example of visibility matrix/FITS file header.

information as illustrated in the figure 5.2.2. "DATEOBS" is the computer system time, with a precision of 0.1 second, which is enough given the instrument angular resolution. Data paquets are tagged by the ADC boards using a FramceCounter, identifying each digitisation frame, and a 125 MHz clock (TimeTag), marking precisely the relative time between the digitisation paquets. Each visibility matrix header includes the FramceCounter/TimeTag information corresponding to the first and last digitisation frames that has been averaged.

PAON-4 test observations have been started in spring 2015, mostly used for debugging and understanding RFI, beams and system. We focus here on the last data sets taken in September and November 2016, with ADC boards performing the FFT. The figure5.2.3 below summarises the observations towards Cygnus A and Cassiopeia A (CasA) performed with PAON-4 in the last quarter of 2016, before upgrading the acquisition cluster and modifying the ADC board acquisition mode from FFT to RAW. All the observations have been managed by Claude Pailler (LAL) and Qizhi Huang (LAL/NAOC) has performed the data cleaning and calibration and provided some of the figures included in this chapter. The CygA data taken in november has lower level of RFI and the results presented here concern thus only the November CygA data.

We consider a drift scan survey of a full east-west strip of sky, covering about 12° range in declination centered at the latitude ($\delta \sim 40.73^{\circ}$) over half month at Nov, 2016. The survey will be composed of 11 constant declination scans, each shifted by 1 degree in declination. So, the PAON-4 would be carried in transit mode, assuming that all of the antennae are pointing to a given direction. Visibilities will be recorded as a function of time or right ascension as the sky moves due to the earth rotation.

For default, we use data set CasAnov16 in this chapter, because Cassiopeia A, Cygnus A and Crab all have data on nov16. Number of files/timebin of CasAnov16 is 12000 (0-11999). Here we show the information/header contained in the vismtx_0_0.ppf.

For PAON-4, there are 4 different auto-correlation channels for each polarisation. The data analysied here correspond to the H polarisation. They are more or less similar except their amplitude, so we will only show one of them. In Fig5.2.4, we plot the auto-correlation of the raw

5.2. PAON-4 OBSERVATIONS AND DATA

			-										
Nom Réf.	Z	Δz	Début	Méridien	Fin	Notes	Nom Réf.	Z	Δz	Début	Méridien	Fin	Notes
CygA6sep16	-6,65°	0°	17 h 00	20 h 45	17 h 00	lf=10Gb	CygA17nov16	-6,65°	0°	13 h 30	16 h 02	13 h 30	If=1Gb Pqts erronés F3 : 117
CygA7sep16	-5,65°	+1°	17 h 30	20 h 41	17 h 30	If=1Gb 6 paquets erronés sur F2	CygA18nov16	-5,65°	+1°	14 h 00	15 h 58	14 h 00	If=1Gb Pqts erronés F2 : 1 F3 : 8
CygA8sep16	-7,65°	-1°	18 h 00	20 h 37	18 h 00	lf=10Gb	CygA19nov16	-7,65°	-1°	14 h 45	15 h 54	14 h 45	lf=1Gb Pqts erronés F2 : 1 F3 : 386
CygA9sep16	-4,65°	+2°	18 h 30	20 h 33	18 h 30	lf=1Gb	CygA21nov16	-4,65°	+2°	10 h 00	15 h 46	10 h 00	If=1Gb Pqts erronés
CygA11sep16	-8,65°	-2°	16 h 00	20 h 26	16 h 00	lf=1Gb							F3.7
CygA12sep16	-3,65°	+3°	16 h 30	20 h 22	16 h 30	lf=1Gb	CygA22nov16	-8,65°	-2°	10 h 30	15 h 42	10 h 30	lf=1Gb
CygA13sep16	-9,65°	-3°	17 h 00	20 h 18	17 h 00	If=10Gb	CygA23nov16	-3,65°	+3°	11 h 00	15 h 39	11 h 00	lf=1Gb Pqts erronés F3 : 78
CygA14sep16	-6,65°	0°	17 h 30	20 h 14	17 h 30	lf=10Gb	CygA24nov16	-9,65°	-3°	11 h 30	15 h 35	11 h 30	If=1Gb Pqts erronés F3 : 4
CygA15sep16	-5,65°	+1°	18 h 00	20 h 10	18 h 00	lf=10Gb	CygA25nov16	-2,65°	+4°	12 h 00	15 h 31	12 h 00	If=1Gb Pqts erronés F2 : 3 F3 : 1
CygA16sep16	-7,65°	-1°	18 h 30	20 h 06	18 h 30	lf=10Gb	CygA26nov16 EN COURS	-10,65°	-4°	12 h 30	15 h 27	12 h 30	lf=1Gb

Décalages CygA

(du 6 au 16 septembre 2016)

Décalages CygA

(novembre 2016)

Points communs visicalc : 8000 (NPaqSum=8064) visicompzm : validé

version mfacq : v10r1

Points communs visicalc : 3000 trigger : 4 kHz visicompzm : validé taux de perte sur bao8/bao9 : ≈ 4% version mfacq : v10r1 matrices de visibilité par seconde sur bao5/bao6 : %1,25 (if=1Gb) %1,23 (if=10Gb)

CasA30sep16

11,42°

0°

(du 26 au 30 septembre 2016)

trigger : 4 kHz

Ζ Nom Réf. Δz Début Méridien Fin Notes CasA26sep16 0° 21 h 00 lf=10Gb 11.42° 22 h 50 1 h 00 CasA27sep16 12,42° +1° 21 h 00 22 h 46 1 h 00 lf=1Gb -1° lf=10Gb CasA28sep16 10,42° 21 h 00 22 h 42 1 h 00 CasA29sep16 11,42° 0° 21 h 00 22 h 38 1 h 00 lf=10Gb

22 h 34

1 h 00

lf=1Gb

Décalages CasA

Points communs

21 h 00

visicalc : 8000 (NPaqSum=8064) trigger : 4 kHz visicompzm : validé taux de perte global : $1,96 \le tx \le 10,96\%$ matrices de visibilité par seconde sur bao5/bao6 : ≈0,47

Figure 5.2.3: The summaries of the observations towards Cygnus A and Cassiopeia A (CasA) performed with PAON-4 in the last quarter of 2016 (Tables provided by Claude Pailler).

121



Figure 5.2.4: The auto-correlation of the raw TOD at 1400MHz (top panel) and 1420MHz (bottom panel).



Figure 5.2.5: The cross-correlation of the raw TOD at 1400MHz (left panel) and 1420MHz (right panel).

5.3. DATA PROCESSING AND MAP MAKING

TOD (Time-Ordered Data) at 1400MHz (top panel) and 1420MHz (bottom panel). Between 16 and 15 o'clock, there are two peaks at the top panel. The first peak is mainly due to the transit of the Cygnus A and the second one is mainly from the Cygnus X. For the bottom panel, around 16.5 o'clock and 1 o'clock, the two peaks come from the HI emission of the Milk Way. There are a lot of radio interference existing by the shape of sharp peaks in the range from 3 to 4 o'clock in both, which also dispersedly exist in other place. Except these, the emissions mainly come from the synchrotron radiation. We only show the real part of the TOD, because the imag part of the auto-correlation are zero. Fig5.2.5 show the cross-correlation of the raw TOD at 1400MHz (left panel) and 1420MHz (right panel), which correspond to the baseline (10.37m, 6m). We could obviously see the interferometric fringe of the Cygnus A.

5.3 Data Processing and Map making

5.3.1 Data cleaning and calibration procedure

Before using the visibility time streams $V_{ij}(t)$ to reconstruct sky maps, the data has to be cleaned to remove RFI, and calibrated in phase and gain. The data shown here has been processed through a processing pipeline, developed by Qizhi Huang, which the following steps:

- determination of frequency dependent gain $g(\nu)$, and time dependent gain G(t) for each analog input signal
- RFI cleaning
- Phase calibration
- Absolute gain or radiometric calibration, to convert visibility signals in arbitrary units (AU) into physical sky brightness in temperature
- Extraction of main instrument characteristics: system temperature Tsys, main dish lobes, baselines...

We describe briefly the above steps below, before using the cleaned calibrated visibility streams to reconstruct sky maps.

The data analysed here has been acquired with the FFT performed on the ADC board, which is the source of an artifact on the computed visibilities. The average value of the visibilities (real/imaginary parts) are expected to the zero when dominated by noise, while the data here exhibits a non zero real parts for visibilities. In order to correct this effect, the visibility data is convolved with a window function to determine the baseline or offset that has to be subtracted from the data.



Figure 5.3.1: A: The $g(\nu)$ by auto-correlation for the third antenna panel of PAON-4. B: Time dependent gain G(t).

Gain calibration

we assume that the Gain can be divide into two part, one is frequency dependent gain, the other is time dependent gain $g(\nu)$:

$$G(t,\nu) = G(t) \cdot g(\nu) \tag{5.3.1}$$

 $g(\nu)$ is obtained from the auto-correlation spectra, computed from the clean part of the spectra, far from point sources, galactic plane crossings and time intervals affected by strong RFI. Assuming the the time dependent gain G(t) do not vary over short time intervals of few minutes, we compute the frequency dependent gain $g(\nu)$ by averaging the auto-correlation spectra over short time intervals. The frequency dependent gains $g(\nu)$ are all normalised.

To determine the time dependent gain G(t), we assume that the auto-correlation signals are dominated by noise fluctuations with constant noise temperature T_{sys} . We consider then that the time variations of the auto-correlation signal fluctuations $\sigma(t)$ can be written as the product of the time dependent gain G(t) and a fixed value dispersion level σ_0 : $\sigma(t) = G(t) \times \sigma_0$

Fig5.3.1A show $g(\nu)$ calculated every 10 minutes. There are totally 66 curves in this figure, and the curves are similar. It means that at least in one day, $g(\nu)$ is stable. Not only for this, we also calculate $g(\nu)$ for different days and on different sources, then the $g(\nu)$ is still stable. Finally, we consider that $g(\nu)$ is stable at any time for this november observation. Fig5.3.1B show time dependent gain G(t), which decreases from morning to afternoon and increases in the evening. It varies with the environment temperature and is also different for other days. Lower temperature leads to larger gain, while higher temperature leads to lower gain. G(t) will affect the profile of the data along times axis significantly. When fitting the fringes and making the sky map, we must calibrate G(t) carefully. Fig5.3.2 shows the 2D time-frequency map for cross-correlation (6m,0).

Radio frequency interference (RFI) can be many times stronger than the astronomical signal. Except placing filters in hardware, we mainly deal with RFI within the observed frequency



Figure 5.3.2: 2D time-frequency map for cross-correlation (6m, 0).



Figure 5.3.3: Data Masking for identify and remove the RFI.

bandwidth by advanced algorithms in software. The software try to find those data in time space which are contaminated by the interfering sources. Subsequently, those data are removed in further analysis, which is operated by calculating the moving average to get the sigma of the residual data smaller than the threshold and leave the Gaussian noise fluctuation. We refer to this process as data mask/flag. Fig. 5.3.3 shows the RFI masking at 1420.4 MHz for the auto-correlation signal 3H. Strong RFI are shown as blue spikes, while the red curve shows the signal after masking/cleaning.

Fig5.3.4 shows the corrected TOD RFI cleaning. The left two panels are auto-correlation for different frequency, corresponding to Fig5.2.4 which are raw TOD. The right panels are cross-correlation for frequency 1420MHz, corresponding to Fig5.2.5. We only show this frequency, due to the curves similar for frequency 1400MHz except the amplitude.

Phase calibration

Due to the system effect, for example, different cable lengths, the received phases have been shifted asynchronous for cross-correlations. Then, the phase calibration should be considered. After gain calibrating, we could determine the phase through the source transit. So, phase calibration will be done by fitting the corresponding fringe. There is linear relationship between shifted phase and frequency. These phases also satisfy the follow relationship:

$$\Delta \Phi_{ij} = \Delta \Phi_{ik} + \Delta \Phi_{kj} = \Delta \Phi_{ik} - \Delta \Phi_{jk} = \Delta \Phi_{kj} - \Delta \Phi_{ki}$$
(5.3.2)



Figure 5.3.4: Corrected TOD using Gain

For antenna 3 and 4, the EW component of their baseline is *zero*. We should not see the fringe and difficult to get phase Φ_{34} . However, we could use the relationship $\Delta \Phi_{34} = \Delta \Phi_{31} - \Delta \Phi_{41} = \Delta \Phi_{32} - \Delta \Phi_{42}$ to predict Φ_{34} and calibrate the correspond data by this method.

Absolute radiometric calibration

Now, let us consider the auto-correlation signals after gain calibration. The 8 auto-correlation should be more or less similar in shape. However, the system temperature are not uniform for different channel. Time ordered visibility data has also to be converted and reordered following the right ascension (RA) angular coordinate. In 1420.4 MHz, the strongest emissions are the galactic 21cm emission and galactic synchrotron emission, while other diffuse emissions only account for less than 1%. Therefore, we use the LAB data (21cm) and GSM (almost the synchrotron) to do amplitude calibration of auto-correlations. It is worth noting that before using LAB and GSM we need to fit their resolutions same as PAON4. Then, we use the peak around RA=300deg to calibrate the scale from Arbitrary unites into Temperature Kelvin. Fig5.3.5 shows the auto-correlation with amplitude calibration, and the sigma noise and system temperature for each channel.

Instrument characteristics

We use the cleaned, gain and phase calibrated visibility signals during bright source transit to derive instrument geometric parameters, as well as noise temperature. Indeed, fitting the phase variations with time during the source transit for each baseline and frequency bin, we are able to compute the delay (phase shift) due to cables and electronics, as well as the East-West baseline. Using similar methods, it is also possible to determine the North-South baselines. Assuming the first antenna located at (0,0), the antenna position is given by antenna2 (-5.993m,-0.001m), antenna3 (4.38m,-5.996m), and antenna4 (4.38m,5.995m). Using auto-correlation signals during the transit, we could obtain the beam in EW direction. If the pointing is exactly on source, we could determine the beam angular size, which is the effective dish diameter illuminating the feed.



Figure 5.3.5: Left: The auto-correlation with amplitude calibration. Right: sigma noise and system temperature for each channel.

So we introduce an efficiency factor η . By observing a point source, for example, CygA, we could calculate that the antenna effective diameter is around 4.4m. The efficiency factor is $\eta \sim 0.88$ for PAON-4.

As mentioned earlier, the cleaning and calibration pipeline has been developed by Qizhi Huang, as part of his PhD work. A more in depth discussion of the method and the results will be included in his thesis.

5.3.2 Quick Map Making

The quick map making method presented here is based on the beam forming technic. It is possible to form synthetic beams by combining signals from several antennae. This method can work either by performing analog combination of narrow band waveforms, with delays and gain terms, or through digital beam forming, computing linear combination of signal Fourier coefficients from a set of antennae. We show below that such the signal intensity within the synthetised beams can also be computed using the visibilities.

Let us first consider the simple case of two antenna signals S_1 , S_2 , which are complex numbers, corresponding for example to the Fourier amplitude for a given frequency, after dong FFT on the analog signal coming out of each antenna, we combine the two signals using two complex coefficients c_1 , c_2 . Denoting the sky intensity by I, we have $I_1 = |S_1|^2$, $I_2 = |S_2|^2$. We write the combined signal as $S = c_1S_1 + c_2S_2$ and $I = |S|^2 = S^*S$. Developing the above expression, one obtain: $I = |c_1|^2I_1 + |c_2|^2I_2 + c_1c_2^*V_{12} + (c_1c_2^*V_{12})^*$, where I_1 and I_2 correspond to the autocorrelation term and $V_{12} = S_1S_2^*$ is the cross-correlation between the two antenna. Please note that the visibilities V_{ij} computed by the software correlator correspond indeed to $S_iS_j^*$ where i, j are the two antenna numbers.

Considering a set of signals s_i from N antennae located in positions $\vec{r_i}, i = 1...N$. We assume here that the s_i signals correspond to the complex Fourier coefficient at frequency ν and wavelength λ of the RF signal collected by the antenna *i*. We compute the linear combination S

of the signals using a set of complex coefficients $c_i \in \mathbb{C}$:

$$S = \sum_{i=1}^{N} c_i s_i \qquad c_i \in \mathbb{C} \quad c_i = \rho_i e^{i\varphi_i} \quad a_i, \varphi_i \in \mathbb{R}$$

Computing the intensity of the RF signal in the synthetic beam, we obtain:

$$I = S^* S = \sum_{i=1}^{N} \sum_{j=1}^{N} c_i^* s_i^* \times c_j s_j$$
$$I = \sum_{i=1}^{N} |c_i|^2 |s_i|^2 + \sum_{i=1}^{N} \sum_{j=1+1}^{N} c_i^* c_j s_i^* s_j + c_i c_j^* s_i s_j^*$$

remembering that the visibilities correspond to the time average of auto and cross correlations, the average intensity can be computed using the visibilities V_{ij} .

$$V_{ij} = \langle s_i^* \, s_j \rangle$$

$$\langle I \rangle = \sum_{i=1}^N |c_i|^2 \, V_{ii} + \sum_{i=1}^N \sum_{j=1+1}^N (W_{ij} + W_{ij}^*) \qquad W_{ij} = (c_i c_j^*) \, V_{ij}$$

The first sum correspond to auto-correlations terms, while the second term correspond to crosscorrelations. Note also that the visibilities computed by the PAON-4 correlator correspond to the complex conjugate of the above expression $\langle s_i s_j^* \rangle$.

For an incoming electromagnetic wave characterised by the wave-vector $\vec{k} = \frac{2\pi}{\lambda} \vec{u}_k$, where \vec{u}_k is the unit vector along the incident wave direction. The wave phase θ_i at the position \vec{r}_i of the antenna (i) can be written as:

$$\exp\left(2\pi\nu\,t - \vec{k}\cdot\vec{r_i}\right) = \exp\left(2\pi\nu\,t + \theta_i\right) \longrightarrow \theta_i = -\frac{2\pi}{\lambda}\,\vec{u_k}\cdot\vec{r_i}$$

To form a beam in the direction \vec{u}_k , one has to compensate the phases θ_i , by using weights c_i with the opposite phase values. For a set of identical antennae, we can have equal gain term $\rho_i = \rho$:

$$c_i = \rho e^{-i\theta_i} = \rho e^{i\vec{k}\cdot\vec{r_i}}$$

$$c_i^* c_j = \rho^2 e^{i\vec{k}\cdot(-\vec{r_i}+\vec{r_j})} = \rho^2 e^{i\phi_{ij}}$$

$$\phi_{ij} = \vec{k}\cdot(\vec{r_j}-\vec{r_i}) = \frac{2\pi}{\lambda} \vec{u_k}\cdot(\vec{r_j}-\vec{r_i})$$

In the case of the QuickMap sky map reconstruction method from transit visibilities discussed here, we need only to form beams along several directions in the meridian plane (vertical plane along the north-south direction). We shall thus neglect the effect of sky drift along the right



Figure 5.3.6: Phase correction term.

ascension (RA) during the visibility averaging time interval of a few seconds, corresponding to about an arcmin on sky, as it is the case of any map making method from transit visibilities.

The sky direction at a given RA in the meridian plane will be identified by the angle β with respect to the vertical axis, positive toward the north. The declination angle δ on the sky will thus be given by $\delta =$ latitude $+ \beta$. The dish or antenna common pointing direction in the meridian plane during a constant declination scan (k) will be denoted β_k . The antenna positions will be referenced with respect to a cartesian coordinate system Oxyz, with the Ox axis along the West-East direction, positive toward East, Oy axis along the South-North direction, positive toward North, and the Oz axis along the vertical, positive for increasing heights. If we assume that all antenna lie in the same xOy plane (same height), the phase correction term expression reduces to (figures 5.3.6):

$$\phi_{ij} = -\frac{2\pi}{\lambda} \sin\beta \times (\vec{r_j} - \vec{r_i})_y = -\frac{2\pi}{\lambda} \sin\beta \times (y_j - y_i)$$

$$\phi_{ij} = -\frac{2\pi}{\lambda} \sin\beta (\Delta y)_{ij} \qquad (\Delta y)_{ij} = d_{NS}$$

where $(\Delta y)_{ij} = d_{NS}$ is the baseline projected along the north-south direction.

Using each constant declination scan along β_k , we can form a single dimensional map along RA by applying the weights $\rho = 1$ and $\phi_{ij} = -\frac{2\pi}{\lambda} \sin\beta - \frac{2\pi}{\lambda} \sin\beta_k (\Delta y)_{ij}$. We can also form beams along shifted directions $\beta = \beta_k + \delta\beta$ by applying the corresponding phase compensation terms, and correcting for the beam response $\rho = L(\delta\beta)/L(0)$, where L denotes the single dish beam response in intensity in angular domain.

When we have several constant declination scans, we combine the result from two adjacent scans for reconstructing maps along *beta* or declination values lying in between two scans: $\beta_k < \beta < \beta_{k+1}$, using suitable weights.



Figure 5.3.7: Quick map making at 1420Mhz; Top: all observed region. Left: extracted from top figures from $RA=55^{\circ}$ to $RA=95^{\circ}$; Right: extracted from top figures from $RA=280^{\circ}$ to $RA=325^{\circ}$.

A more detailed description of the linear combinations for reconstructing maps from PAON-4 data using the QuickMap algorithm is given below.

The reconstructed maps has rectangular shape, 24 hour (=360°) in RA with angular resolution $\sim 0.25^{\circ}$. Considering that we have constant declination scans every 1° along β or elevation, the output map will have the corresponding declination (First case) (a ROW in the map \rightarrow constant declination), and three interpolation $\delta\beta = +0.25^{\circ}, +0.5^{\circ}, +0.75^{\circ}$ between two scans (Second case). At the two edges of the map, we can add two $\delta\beta$ scans $0.25^{\circ}, 0.5^{\circ}$ using the edge scan (Third case).

It is worth noting that we have to take into account the "w-term" which correspond to the phase difference between antenna no sitting in the same plane perpendicular to the line of sight. If $\beta_0 = (\beta_{\text{PAON}-4} - \beta_{\text{source}})$ is the line of sight direction with respect to the zenith, and assuming that all antenna are in the same horizontal plane, we will have a term (projected baseline) like $d_{NS} \sin(\beta_0)$. This term would of course be present even in the First case, i.e. reconstruction along a given scan direction.

First case

For map rows corresponding to the scan declinations, we use only the visibility data from the corresponding scan. The map pixel would be obtained simply by adding antennae signals with $c_i = 1$ (same weight and no phase shift). That should correspond more or less to summing all the visibilities.

We write the combined signal as

$$S = S_1 + S_2 + e^{i\frac{2\pi}{\lambda}(-6)\sin(\beta_0)}S_3 + e^{i\frac{2\pi}{\lambda}6\sin(\beta_0)}S_4$$
(5.3.3)

Here, the phase $e^{i\frac{2\pi}{\lambda}(-6)\sin(\beta_0)}$ and $e^{i\frac{2\pi}{\lambda}6\sin(\beta_0)}$ appear due to w-term. For PAON-4 case, center dish is considered as reference which is marked as the first antenna, so $d_{NS} = 0m$ for the second antenna which is sat in westward direction corresponding to the first one; $d_{NS} =$

5.3. DATA PROCESSING AND MAP MAKING

-6m for the third antenna in southeastward direction; and $d_{NS} = 6m$ for the fourth antenna in northeastward direction. $\beta_0 = 0$ for zenith direction. We define that β_0 is negative degree as antennae point to north direction corresponding to zenith and β_0 is positive degree as pointing to the south.

Assuming $\varphi = \frac{2\pi}{\lambda} 6\sin(\beta_0)$, then

$$I = S^*S$$

= $I_1 + I_2 + I_3 + I_4 + V_{12} + V_{21} + e^{i\varphi}V_{13} + e^{-i\varphi}V_{31} + e^{-i\varphi}V_{14} + e^{i\varphi}V_{41}$
+ $e^{i\varphi}V_{23} + e^{-i\varphi}V_{32} + e^{-i\varphi}V_{24} + e^{i\varphi}V_{42} + e^{-2i\varphi}V_{34} + e^{2i\varphi}V_{43}$ (5.3.4)

There are 4 auto-correlations (I_1, I_2, I_3, I_4) , and 6 cross-correlations $(V_{12}, V_{13}, V_{14}, V_{23}, V_{24}, V_{34})$. All these terms are present in the *I* expression. We expect the four auto-correlation to be the same, or very close, if correctly calibrated. For V_{ij} and V_{ji} , they are complex conjugate relationship $V_{ij} = (V_{ji})^*$, in other words, we get $2 * real(V_{ij})$ which ensures that the final results *I* is real.

Second case

For interpolated map rows between two sky scans, we combine the visibilities from the two β_0 scans. Equal weight $|c_i|$ are applied to the two sets for 0.5^o interpolation, while the closest β_0 scan has higher weight for 0.25^o and 0.75^o scans. Of course the c_i weights/coefficients have phases to compensate for the phase differences of the sky signals reaching the antennae from declinations shifted with respect to the antennae pointing directions and existing only in the NS direction.

Now, we write the signals for the two scans at the interpolated pointing $(\beta_0 + \beta_1)$,

$$S_1^{\star} = S_1 + S_2 + e^{i\frac{2\pi}{\lambda}[(-6)\sin(\beta_0 + \beta_1)]}S_3 + e^{i\frac{2\pi}{\lambda}[6\sin(\beta_0 + \beta_1)]}S_4$$
(5.3.5)

$$S_2^{\star} = S_1' + S_2' + e^{i\frac{2\pi}{\lambda}[(-6)\sin(\beta_0' + \beta_2)]}S_3' + e^{i\frac{2\pi}{\lambda}[6\sin(\beta_0' + \beta_2)]}S_4'$$
(5.3.6)

Here, the phase corresponding to β_0 also come from w-term. We assume that the pointing S_1 is toward the north compared with pointing S_2 , then $\beta'_0 = \beta_0 + 1^o$. The phase corresponding to β is the phase shift for interpolated pointing. The value of β could be 0.25^o , 0.5^o , or 0.75^o , with $\beta_2 = \beta_1 - 1^o$. The interpolated declination is $\beta_{\text{PAON}-4} - (\beta_0 + \beta_1)$.

Assuming $\varphi_1 = \frac{2\pi}{\lambda} [6\sin(\beta_0 + \beta_1)]$ and $\varphi_2 = \frac{2\pi}{\lambda} [6\sin(\beta'_0 + \beta_2)]$, then

$$I_{1}^{\star} = I_{1} + I_{2} + I_{3} + I_{4} + V_{12} + V_{12}^{*} + e^{i\varphi_{1}}V_{13} + [e^{i\varphi_{1}}V_{13}]^{*} + e^{-i\varphi_{1}}V_{14} + [e^{-i\varphi_{1}}V_{14}]^{*} + e^{i\varphi_{1}}V_{23} + [e^{i\varphi_{1}}V_{23}]^{*} + e^{-i\varphi_{1}}V_{24} + [e^{-i\varphi_{1}}V_{24}]^{*} + e^{-2i\varphi_{1}}V_{34} + [e^{-2i\varphi_{1}}V_{34}]^{*} (5.3.7)$$

$$I_{2}^{\star} = I_{1}^{\prime} + I_{2}^{\prime} + I_{3}^{\prime} + I_{4}^{\prime} + V_{12}^{\prime} + V_{12}^{\prime}^{*} + e^{i\varphi_{2}}V_{13}^{\prime} + [e^{i\varphi_{2}}V_{13}^{\prime}]^{*} + e^{-i\varphi_{2}}V_{14}^{\prime} + [e^{-i\varphi_{2}}V_{14}^{\prime}]^{*} + e^{i\varphi_{2}}V_{23}^{\prime} + [e^{i\varphi_{2}}V_{23}^{\prime}]^{*} + e^{-i\varphi_{2}}V_{24}^{\prime} + [e^{-i\varphi_{2}}V_{24}^{\prime}]^{*} + e^{-2i\varphi_{2}}V_{34}^{\prime} + [e^{-2i\varphi_{2}}V_{34}^{\prime}]^{*} (5.3.8)$$

In fact, we compute I_I^* and I_2^* for the two scans, introducing the phase term for V_{13} , V_{14} , V_{23} , V_{24} , V_{34} . Then, we sum I_I^* and I_2^* with the two weights (c_1, c_2) to get the interpolated intensity,

$$I = c_1^2 I_1^{\star} + c_2^2 I_2^{\star} \tag{5.3.9}$$

The exactly weights coefficient c_i is given by,

• for
$$\beta_1 = 0.5^o$$
 and $\beta_2 = -0.5^o$, $c_1 = c_2 = 0.7$

- for $\beta_1 = 0.25^o$ and $\beta_2 = -0.75^o$, $c_1 = 0.94$, $c_2 = 0.33$ (closed to north pointing)
- for $\beta_1 = 0.75^o$ and $\beta_2 = -0.25^o$, $c_1 = 0.33$, $c_2 = 0.94$ (closed to south pointing)

Third case

For the map rows beyond the edge scan, we use only the corresponding edge scan, applying the phase term compensating the shift in NS between the sky direction and antennae pointing directions. Now, the signals at the interpolated pointing is given by,

$$S = S_1 + S_2 + e^{i\frac{2\pi}{\lambda}[(-6)\sin(\beta_0 + \beta)]}S_3 + e^{i\frac{2\pi}{\lambda}[6\sin(\beta_0 + \beta)]}S_4$$
(5.3.10)

Assuming $\varphi = \frac{2\pi}{\lambda} [6\sin(\beta_0 + \beta)]$, then the intensity I with the weight c could be written as

$$I = c^{2}S^{*}S$$

= $c^{2}\left\{I_{1} + I_{2} + I_{3} + I_{4} + V_{12} + V_{12}^{*} + e^{i\varphi}V_{13} + [e^{i\varphi}V_{13}]^{*} + e^{-i\varphi}V_{14} + [e^{-i\varphi}V_{14}]^{*} + e^{i\varphi}V_{23} + [e^{i\varphi}V_{23}]^{*} + e^{-i\varphi}V_{24} + [e^{-i\varphi}V_{24}]^{*} + e^{-2i\varphi}V_{34} + [e^{-2i\varphi}V_{34}]^{*}\right\}$ (5.3.11)

The weights coefficient c_i is given by,

- for north edge with $\beta = -0.5^{\circ}, c = 0.7$
- for north edge with $\beta = -0.25^{\circ}$, c = 0.94
- for south edge with $\beta = 0.5^{o}$, c = 0.7
- for south edge with $\beta = 0.25^{\circ}$, c = 0.94

Fig5.3.7 shows the Quick map making for 1420MHz from data of Nov, 2016.

5.4 Instrument and Map making performance

Using the PAON4 calibrated November 2016 data, we have reconstructed sky map from visibilities at two frequencies: 1420.4 and 1400 MHz. It is important to compare the obtained maps with the known maps, for example, the Leiden-Argentina-Bonn(LAB) map at 21cm and Haslam map at other frequencies. When computing simulated visibilities from maps, we should make sure that the input maps have the same frequency band and map compatible angular resolution corresponding to the observed data. Here, we use the updated PAON-4 November observation data, which have visibility data binned in 512 frequency channels, each spanning 480 KHz from 1250 MHz to 1500 MHz. We select two frequency 1420.4MHz and 1400MHz to show the results. The effective diameter of the four antennas is about 4.4 meters. The instrument resolution is around 1° for PAON-4. So the map resolution $n_{side} = 256$ is enough for spherical maps with HEALPix pixelisation scheme.

In the observed region, the CygA source is the brightest source, which is not only important for calibration but also when comparing data with the simulated signals. We will compare the autocorrelation signals and cross-correlation signals between the PAON-4 data and the simulated data from corresponding known map during transiting this source. It should be stressed that in order to obtain meaningful results, the CygA source extension being significantly below the map pixel resolution, it should be included as a point source, i.e. a single pixel in the HEALPix spherical maps. That means that we can not use directly maps provided by modeling softwares as GSM, as point sources appears extended in these maps. This will lead to mistakes at the amplitude and the fringe shape when we compute the auto-correlation signals and cross-correlation signals. So we have used input maps without point source, and then add the CygA as a point source with the corresponding temperature, which is about 1400K for $\lambda = 21$ cm and could be calculated by following fomulae. The CygA is a double source, but the separation is much smaller than our map pixel size. We have modeled it as a single point source of 1500 Jy at 1400 MHz. The flux (in Jansky) is converted to pixel temperature excess T_{pix} using the formulae:

$$T_{\rm pix} = \frac{Flux \times \lambda^2}{k_B \, 2\delta\Omega_{\rm pix}} = \frac{Flux \times 10^{-26}}{1.38 \times 10^{-23}} \times \frac{\lambda^2}{2\delta\Omega_{\rm pix}}$$

We have taken the average of 7 maps from LAB maps temperature, each covering ~ 40 kHz in order to obtain a 21 cm temperature map with a bandwidth compatible with the PAON4 data analysed here. The synchrotron map at 1400 MHz has been obtained using the desourced, destripped Haslam map ² [50, 49] extrapolated to from the 408 MHz to about 1400 MHz, using a spectral index $\beta \sim 2$, which correspond to a factor ~ 10 decrease in temperature.

Here the LAB map for the sky emission at 21cm (Galactic HI) is shown in equatorial coordinates in figure 5.4.4A for the declination range 62° to 28° . For PAON-4 November data, a total of 11 constant declination scans are available, covering the angular offset range -5 deg to +5 deg with respect to the CygA source declination, or -1.65° to -11.65° with respect to the zenith, given the PAON-4 site declination of 47.38°. PAON-4 pointing accuracy is around 0.2° , and the exact values of pointed declination are not yet available.

5.4.1 Comparing data and simulations for the auto and cross correlation signals

Before reconstructing map, we compare the observed visibilities with simulated auto-correlation cross-correlation signals computed using LAB+CygA as an input map. The data calibration is reasonable, but it is not prefect. The local ground of the PAON-4 has not been fattened before installing the antennae and we can expect currently unknown height differences between the antennae of up to ~ 25 cm.

Fig5.4.1 shows the comparison of data and simulated visibilities for the auto-correlation signals for the eleven scans at frequency 1420.4 MHz. The black curve corresponds simulated visibilities

²https://https://www.nasa.gov/goddard



Figure 5.4.1: The auto-correlation signals comparison of PAON-4 data (red curve) and simulated visibilites (black curve) for the eleven scans at frequency 1420.4 MHz.



Figure 5.4.2: The cross-correlation signal comparison between simulated signals (black curve) from LAB+CygA map and the PAON-4 observed visibilities (red curve) at 1420.4 MHz. The red curves show the cross-correlation from PAON-4 November data, and the black curves is simulated from LAB data.

from LAB map, while the red curves show the observed PAON-4 visibilities. There are two peaks in all subfigures, which is mostly due the 21cm emission from the two crossing of the Milk Way around right ascension of $\alpha \sim 85^{\circ}$ and $\alpha \sim 290^{\circ}$. The contribution of CygA source cannot be easily be distinguished from the Milky Way 21 cm emission around $\alpha \sim 290^{\circ}$. We see that there is reasonable agreement between PAON-4 calibrated data and visibilities computed from an input map using the map reconstruction software. It should however be stressed that the simulated visibilities shown here are obtained going through the complete map reconstruction in the spherical harmonics space (ℓ, m) . Indeed, we proceed as follow to compute these simulated visibilities:

- 1. We compute the auto and cross-correlation beams, and convert them into beams in (ℓ, m) space.
- 2. We apply also spherical harmonics transform to the input map to obtain the sky representation $a_{\ell m}^{\text{sky}}$ in (ℓ, m) space.
- 3. For each m-mode, we compute the beam and sky convolution in (ℓ, m) space for each beam to obtain the m-mode visibility vectors $\tilde{V}_{ij}(m)$
- 4. We apply a 1-D Fourier transform to the m-mode visibility vector $V_{ij}(m)$ for each beam (or baseline) to compute the visibilities in angular space, $V_{ij}(\alpha)$ which is then compared to the observed calibrated visibilities We could find most of the auto-correlation signals are reasonable.

Except the last two subfigures, corresponding to declination offsets with respect to CygA of -4 and -5 degrees, the two sets of visibilities (PAON4 data and simulated ones) are compatible. For the declination ofsset -4° , the peak in the observed visibility is narrower around right ascension 300° , compared to the simulated one, and for the last offset value of -5° , the amplitude of the peak around right ascension 90° is higher for the observed visibility. These differences might be due to some calibration problems, such as wrong time dependent gain corrections.

Fig 5.4.2 shows the comparison of cross-correlation signals from PAON-4 November 2016 observations and the simulated visibilities from LAB+CygA map, at 1420.4 MHz and the elevation -6.65° , with the antennae pointing toward CygA declination ($\delta = 40.73^{\circ}$). The data for this elevation has been collected on November 17th, 2016. The red curves correspond to the cross-correlation from PAON-4 November data, while the black curves correspond to visibilities simulated from LAB+CygA map.

We can see a good data-simulation agreement for the first baseline (6m,0), while a shift in time (or phase shift) can be seen for the other four baselines (4.39m, \pm 6m) and (10.39m, \pm 6m). This shift can be compensated by a additional phase of $\sim \pi$. As there is an uncertainty of ~ 25 cm on height (z-coordinate) of the antennae, we have tried to compensate for this phase shift by changing the z-coordinate of antennae 3 and 4. we leave the antennae 1 and 2 at $z_1 = z_2 = 0$, while we shift by 10 cm the height of antennae 3 and 4 $z_3 = +10$ cm $z_4 = -10$ cm. We obtain thus the seven following baselines for the auto and 6 cross-correlations:



Figure 5.4.3: The corrected cross-correlating signals for the pointing of 0° at 1420MHz for baseline with Δz . The red curves show the cross-correlation from PAON-4 November data, and the blue curves is simulated from LAB data.

AutoCor	:	(0 m, 0 m, 0 m)
1H2H	:	(-6 m, 0 m, 0 m)
1H3H	:	(+4.39 m, -6 m, +0.1 m)
1H4H	:	(+4.39 m, +6 m, -0.1 m)
2H3H	:	(+10.39 m, -6 m, +0.1 m)
2H4H	:	(+10.39 m, +6 m, -0.1m)
3H4H	:	(0, +12 m, -0.2 m)

Fig 5.4.3 shows the comparison of the six visibilities signals, real and imaginary parts, for the PAON-4 observations and simulated ones, after the correction of antennae z-positions. PAON-4 data is shown in red, and simulated visibilities in blue. We still see a discrepancy for the last 3H4H signal, which is a pure NS baseline. More investigations are necessary to identify the source of this problem. We have thus used the auto-correlation and the first 5 cross-correlations for the maps reconstructed from PAON-4 data here.

5.4.2 Reconstructed maps from PAON4 data using m-mode decomposition

We proceed now to the reconstruction of sky maps sky map from PAON-4 data, using corrected baselines with Δz terms. The reconstruction process has been discussed in full detail in chapter



Figure 5.4.4: Reconstructed map for PAON-4 at 1420MHz. A: input LAB+CygA map; B: reconstructed map from LAB+CyaA map after filtering; C: reconstructed map from November data after filtering.

3. It can be decomposed in the following steps: We compute the beams in (ℓ, m) plane for each baseline and elevation, convert the visibilities as a function of right ascension $V_{ij}(\alpha)$ to the m-mode visibilities $\tilde{V}_{ij}(m)$, using an FFT. The map is then the reconstructed in spherical harmonicas space and the sky map in angular space is then computed using an inverse spherical harmonic transform. The eigenvalue thresholds when computing the pseudo inverse will have an effect on the map reconstruction. By lowering the threshold, we reconstruct more modes, but at the expense of extra noise. This has been discussed in Sec. 3.5.

There are also the problem of stripe due to the sharp edge of m-mode in Alm space. We would like to have compromise between the noise and the level of structures. After trying several values for the thresholds, the maps shown here have been computed using the values $(thr_{ratio} = 0.05, and thr_{abs} = 0.01)$.

Before computing the sky map in angular space, we can have an additional filtering step in (ℓ, m) plane, as discussed in chapter 3. We have applied an (ℓ, m) plane filter with $\ell_{\rm cut} \sim 300$ to obtain the filtered maps shown here.

Fig 5.4.4 shows the reconstructed map from simulated visibilities LAB+CygA map, as well as the one reconstructed from PAON4 November 2016 data, using the auto-correlation and the 5 cross correlations 1H2H, 1H3H, 1H4H, 2H3H, 2H4H. Fig 5.4.5 shows the extracted map from each panel of the figure 5.4.4, where the CygA source can easily be identified.

For the radio continuum which is dominated by the Galactic synchrotron emission at the relevant frequencies, we have used full sky maps extrapolated from Haslam map at 408 MHz, with the addition of a point source at the CygA position. This synchrotron map at 1400 MHz is shown in Fig 5.4.6 A, and the reconstructed map from PAON-4 data at 1400 MHz is shown on the lower panel of this same figure. There is a reasonable agreement between reconstructed maps from PAON-4 data and the 21 cm or synchrotron maps.



Figure 5.4.5: Extracted map around the CygA source region (right ascension from 290° to 310°; and declination from 46.73° to 34.73) from the reconstructed map. A : extracted map from input LAB+CygA map; B: extracted map from reconstructed map using LAB+CyaA map after filtering; C: extracted map from reconstructed map using Nov data after filtering.



Figure 5.4.6: Reconstructed map for PAON-4 at 1400MHz. A: input map from Haslam+point source at 1400MHz; B: reconstructed map from November data after filtering.



Figure 5.4.7: The comparison of the reconstructed maps from the simulations with a sharp and smooth threshold for PAON-4 at 1420MHz.

5.4.3 Effect of eigenvalue threshold on reconstructed maps

In section 3.5, we have discussed the solution of the reconstruction, and two kinds of the thresholds on the eigenvalues. For the sharp kind of threshold, it is the sample case, but will leads to some striping and oscillations in the reconstructed sky maps. So we have introduced a smooth kinds of the thresholds to avoid the striping effect. Here, for the sharp kind of the threshold, we take the relative threshold = 0.05 and absolute threshold = 0.02. For the smooth kind of the threshold, we use two mathematical expressions to explain what we do: $thr1 = min(\epsilon_r \times \Sigma_{00}, \epsilon_a)$, and $thr2 = \alpha \times thr1$, with $\alpha = 0.5$ here. When computing the inverse of the diagonal eigenvalue matrix, we apply the following rules:

- for $\Sigma_{ii} < thr1$, we take simply computed $1/\Sigma_{ii}$.
- for $thr2 < \Sigma_{ii} < thr1$, we limit the impact of noise by taking 1/thr1.
- for $\Sigma_{ii} < thr2$, 1/thr1 goes smoothly to zero.

Fig5.4.7 and Fig5.4.8 shows the comparison of the reconstructed maps from the simulations and from the data with a sharp and smooth threshold for PAON-4. If we compare the A1 with A2, or A3 with A4 from the simulation, we could find obviously effect of the reducing of the stripes. However, this decrease of striping is not clearly visible for the PAON-4 Nov data case,



Figure 5.4.8: The comparison of the reconstructed maps from the data with a sharp and smooth threshold for PAON-4 at 1420MHz.

comparing the B1 with B2. The reconstructed maps from the data show indeed a level of striping significantly larger than the ones in simulations. These stripes are likely due to other effects. One possible source is the pointing errors. Our scan directions are known to 0.1-0.2 degrees, and that might be responsible for some the spurious structures we see in the reconstructed maps.

Summary and Outlooks

Transit observation is a very efficient way to survey large area of sky. In such observations, the antennas are fixed or changed only the declination of their pointing, while the sky drifts across their field of view as the Earth rotates. As the instrument is fixed on ground, the instrument response and ground pickups are stable, and the mechanical design could be greatly simplified, and data analysis may also take advantage of such simplicity.

Here we proposed a map-making algorithm based on spherical harmonic transformation. We show that the reconstruction of the whole sky can be decomposed into a number of *m*-modes, each can be computed independently, and the numerical complexity of the problem can be reduced by several orders of magnitude. We have also developed an efficient and flexible code for parallel computing, which can complete the map reconstruction from the visibilities obtained in transit observations of single dish or interferometer arrays.

In this dissertation we investigated the cases of dish array and cylinder array, for several different configurations and survey strategies. We simulated the time-ordered visibility data, and reconstructed sky maps from such data. We also computed the response matrix, transfer function and noise covariance matrix for the different array configurations, which can help us to understand the function of performance of the arrays. The configurations we considered include the following:

(1) the relative simple case of 4 dishes, which can be arranged in a 2×2 regular array, or the triangular arrangement used in the actual PAON-4 test array. We showed the importance of having short baselines and dense array configuration for reconstruction, which provides complete coverage in the (ℓ, m) space without holes. This is important because in the 21cm intensity mapping or epoch of reionization observations, foreground have to be subtracted, and incomplete coverage in the (ℓ, m) would result in unsolvable mode-mixing problem. We also showed that with large number of independent (non-redundant) baselines, more uniform coverage on the (ℓ, m) plane can be obtained.

(2) we then discussed the case of 16 dishes, including a 4×4 regular configuration, and the nearly double concentric ring configuration used n the Tianlai dish array. The Tianlai dish array configuration was chosen, so as to provide good (ℓ, m) coverage while at the same time minimize blocking factor during observations with large zenith angle. We computed the transfer function and noise angular power spectrum, and found that the circular configuration used by the Tianlai array gives better coverage in the (ℓ, m) space, and correspondingly also good angular resolution and symmetric beam, though the regular 4×4 array does have the advantage that redundant baseline calibration could be used in its observation.

We also compared the case of mid-latitude observation to that of polar cap observation. Our algorithm naturally handles the case of polar cap observation. For a given period of observation time, in the polar cap region higher sensitivity can be achieved, at the price of observing smaller survey area. In the polar cap area smaller *m*-modes are probed, but generally the sensitivity range of ℓ is stilled determined by the long baselines of the array.

The sensitivity of the Tianlai 16-dish array is not high enough to detect the cosmic HI signal, but it could measure the Galactic HI signal up to $\ell < 1000$. Beyond the Milky Way, it may also detect extragalactic HI clumps with mass $10^9 M_{\odot}$ up to 30 Mpc.

(3) We then applied the same method to the Tianlai cylinder array pathfinder. At present, the cylinder array have a total of 96 dual polarization feeds on three north-south oriented cylinders, each cylinder is 15m wide and 40m long, the length of the cylinder allows adding more feeds. We considered two basic element spacings, one about one wavelength, another about two wavelength. On each cylinder the feeds are arranged as a uniform spacing linear array. However, if the spacing is greater than half-wavelength, grating lobes would appear for linear arrays. So we considered two types of arrangement plans: one is what we called the "regular array", where on each of the three cylinders we have equal number (32) of feeds, all with the same spacing; another which we called "irregular array" have 31, 32, 33 feeds on the three cylinders respectively, on each cylinder the feeds are arranged with the same spacing, but on the three cylinders and spacing is different, though the total length (longest baseline) are the same. We showed that with the regular array, there are grating lobes and strong sources on the sky may introduce ghost images in the reconstructed map, but with the irregular array, the grating lobes are essentially eliminated and good reconstruction could be made.

We gone through the same analysis with the cylinder array, and obtained the synthesis beam, response matrix and noise covariance matrix, and the transfer function, we also computed the noise angular power spectrum. We found that for a system temperature of $T_{sys} = 50$ K and a bandwidth of 1MHz, the angular power spectrum of the cosmic 21cm signal is slightly larger than the noise, detecting such a signal would be a difficult challenge, but the detection may still be achieved by combined analysis for multiple frequency channels.

Finally, this code is applied to analyze the data obtained in the survey observation conducted by PAON-4. For the real data collected in such observation, we first make quality check of the data, and identified and removed RFIs, then we calibrated the instrument phase and gain of the data by using strong point sources Cas A and Cyg A. The calibration result is presented, and we finally obtained the calibrated map and compared it with known sky map.

This dissertation presents a whole set of computation formalism for transit observations, which is applicable to single dishes and interferometer arrays, including dishes, cylinders, or other types of antennas. The formalism could be extended to treat polarizations. However, it is still limited in scope. The map making is performed with a single frequency, multi-frequency joint analysis and foreground subtraction is not covered, and only the simplest treatment of calibration in the real data is investigated, a self-consistent formalism for calibration and map-making is yet to be developed. These problems needs to be solved for a high precision observation of the neutral hydrogen.

Bibliography

- [1] M. Aaronson, J. Huchra, J. R. Mould, R. B. Tully, J. R. Fisher, H. van Woerden, W. M. Goss, P. Chamaraux, U. Mebold, B. Siegman, G. Berriman, and S. E. Persson. A catalog of infrared magnitudes and H I velocity widths for nearby galaxies. *ApJS*, 50:241–262, November 1982.
- [2] A. Albrecht, G. Bernstein, R. Cahn, W. L. Freedman, J. Hewitt, W. Hu, J. Huth, M. Kamionkowski, E. W. Kolb, L. Knox, J. C. Mather, S. Staggs, and N. B. Suntzeff. Report of the Dark Energy Task Force. *ArXiv Astrophysics e-prints*, September 2006.
- [3] S. Ananthakrishnan. The Giant Meterwave Radio Telescope / GMRT. Journal of Astrophysics and Astronomy Supplement, 16:427, 1995.
- [4] R. Ansari, J. E. Campagne, P. Colom, J. M. Le Goff, C. Magneville, J. M. Martin, M. Moniez, J. Rich, and C. Yèche. 21 cm observation of large-scale structures at z ~ 1. Instrument sensitivity and foreground subtraction. A&A, 540:A129, April 2012.
- [5] R. Ansari, J.-E. Campagne, P. Colom, C. Magneville, J.-M. Martin, M. Moniez, J. Rich, and C. Yèche. BAORadio: A digital pipeline for radio interferometry and 21 cm mapping of large scale structures. *Comptes Rendus Physique*, 13:46–53, January 2012.
- [6] R. Ansari et al. "Reconstruction of HI power spectra with radio-interferometers to study dark energy". In *Preprint*, 2009.
- [7] R. Auld, R. F. Minchin, J. I. Davies, B. Catinella, W. van Driel, P. A. Henning, S. Linder, E. Momjian, E. Muller, K. O'Neil, S. Sabatini, S. Schneider, G. Bothun, L. Cortese, M. Disney, G. L. Hoffman, M. Putman, J. L. Rosenberg, M. Baes, W. J. G. de Blok, A. Boselli, E. Brinks, N. Brosch, J. Irwin, I. D. Karachentsev, V. A. Kilborn, B. Koribalski, and K. Spekkens. The Arecibo Galaxy Environment Survey: precursor observations of the NGC 628 group. *MNRAS*, 371:1617–1640, October 2006.
- [8] D. J. Bacon, A. R. Refregier, and R. S. Ellis. Detection of weak gravitational lensing by large-scale structure. *MNRAS*, 318:625–640, October 2000.
- [9] J. C. A. Barata and M. S. Hussein. The Moore-Penrose Pseudoinverse: A Tutorial Review of the Theory. *Brazilian Journal of Physics*, 42:146–165, April 2012.
- [10] R. Barkana and A. Loeb. In the beginning: the first sources of light and the reionization of the universe. *Phys. Rep.*, 349:125–238, July 2001.
- [11] R. A. Battye, M. L. Brown, I. W. A. Browne, R. J. Davis, P. Dewdney, C. Dickinson, G. Heron, B. Maffei, A. Pourtsidou, and P. N. Wilkinson. BINGO: a single dish approach to 21cm intensity mapping. *ArXiv e-prints*, September 2012.
- [12] S. Baumont, C. Balland, P. Astier, J. Guy, D. Hardin, D. A. Howell, C. Lidman, M. Mouchet, R. Pain, and N. Regnault. PHotometry Assisted Spectral Extraction (PHASE) and identification of SNLS supernovae. A&A, 491:567–585, November 2008.
- [13] J. Binney and M. Merrifield. *Galactic Astronomy*. 1998.
- [14] S. Borgani. Cosmology with Clusters of Galaxies. In M. Plionis, O. López-Cruz, and D. Hughes, editors, A Pan-Chromatic View of Clusters of Galaxies and the Large-Scale Structure, volume 740 of Lecture Notes in Physics, Berlin Springer Verlag, page 24, 2008.
- [15] M. Born and E. Wolf. *Principles of Optics*. October 1999.
- [16] T. G. Brainerd, R. D. Blandford, and I. Smail. Weak Gravitational Lensing by Galaxies. *ApJ*, 466:623, August 1996.
- [17] C. Brüns, J. Kerp, L. Staveley-Smith, U. Mebold, M. E. Putman, R. F. Haynes, P. M. W. Kalberla, E. Muller, and M. D. Filipovic. The Parkes H I Survey of the Magellanic System. *A&A*, 432:45–67, March 2005.
- [18] E. Carretti, S. Cortiglioni, C. Sbarra, and R. Tascone. Antenna instrumental polarization and its effects on E- and B-modes for CMBP observations. *A&A*, 420:437–445, June 2004.
- [19] T.-C. Chang, U.-L. Pen, J. B. Peterson, and P. McDonald. Baryon Acoustic Oscillation Intensity Mapping of Dark Energy. *Physical Review Letters*, 100(9):091303, March 2008.
- [20] X. Chen. The Tianlai Project: a 21CM Cosmology Experiment. International Journal of Modern Physics Conference Series, 12:256–263, March 2012.
- [21] X. Chen. The Tianlai 21cm intensity mapping experiment. *IAU General Assembly*, 22:2252187, August 2015.
- [22] B. G. Clark. An efficient implementation of the algorithm 'CLEAN'. A&A, 89:377, September 1980.
- [23] L. Connor, H.-H. Lin, K. Masui, N. Oppermann, U.-L. Pen, J. B. Peterson, A. Roman, and J. Sievers. Constraints on the FRB rate at 700-900 MHz. *ArXiv e-prints*, February 2016.
- [24] A. Cooray, W. Hu, D. Huterer, and M. Joffre. Measuring Angular Diameter Distances through Halo Clustering. *ApJ*, 557:L7–L10, August 2001.

- [25] Y. Copin, N. Blanc, S. Bongard, E. Gangler, L. Saugé, G. Smadja, P. Antilogus, G. Garavini, S. Gilles, R. Pain, G. Aldering, S. Bailey, B. C. Lee, S. Loken, P. Nugent, S. Perlmutter, R. Scalzo, R. C. Thomas, L. Wang, B. A. Weaver, E. Pécontal, R. Kessler, C. Baltay, D. Rabinowitz, and A. Bauer. The Nearby Supernova Factory. *New Astron. Rev.*, 50:436– 438, June 2006.
- [26] T. J. Cornwell, K. Golap, and S. Bhatnagar. The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm. *IEEE Journal of Selected Topics in Signal Processing*, 2:647–657, November 2008.
- [27] T. J. Cornwell and R. A. Perley. Radio-interferometric imaging of very large fields The problem of non-coplanar arrays. A&A, 261:353–364, July 1992.
- [28] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey, R. H. Barkhouser, J. E. Bautista, A. Beifiori, A. A. Berlind, V. Bhardwaj, D. Bizyaev, C. H. Blake, M. R. Blanton, M. Blomqvist, A. S. Bolton, A. Borde, J. Bovy, W. N. Brandt, H. Brewington, J. Brinkmann, P. J. Brown, J. R. Brownstein, K. Bundy, N. G. Busca, W. Carithers, A. R. Carnero, M. A. Carr, Y. Chen, J. Comparat, N. Connolly, F. Cope, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, T. Delubac, R. de Putter, S. Dhital, A. Ealet, G. L. Ebelke, D. J. Eisenstein, S. Escoffier, X. Fan, N. Filiz Ak, H. Finley, A. Font-Ribera, R. Génova-Santos, J. E. Gunn, H. Guo, D. Haggard, P. B. Hall, J.-C. Hamilton, B. Harris, D. W. Harris, S. Ho, D. W. Hogg, D. Holder, K. Honscheid, J. Huehnerhoff, B. Jordan, W. P. Jordan, G. Kauffmann, E. A. Kazin, D. Kirkby, M. A. Klaene, J.-P. Kneib, J.-M. Le Goff, K.-G. Lee, D. C. Long, C. P. Loomis, B. Lundgren, R. H. Lupton, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. McDonald, I. D. McGreer, R. G. McMahon, O. Mena, J. Miralda-Escudé, A. D. Montero-Dorta, F. Montesano, D. Muna, A. D. Myers, T. Naugle, R. C. Nichol, P. Noterdaeme, S. E. Nuza, M. D. Olmstead, A. Oravetz, D. J. Oravetz, R. Owen, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. K. Parejko, I. Pâris, W. J. Percival, I. Pérez-Fournon, I. Pérez-Ràfols, P. Petitjean, R. Pfaffenberger, J. Pforr, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, R. Rebolo, J. Rich, G. T. Richards, C. M. Rockosi, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martin, L. Samushia, A. G. Sánchez, C. Sayres, S. J. Schmidt, D. P. Schneider, C. G. Scóccola, H.-J. Seo, A. Shelden, E. Sheldon, Y. Shen, Y. Shu, A. Slosar, S. A. Smee, S. A. Snedden, F. Stauffer, O. Steele, M. A. Strauss, A. Streblyanska, N. Suzuki, M. E. C. Swanson, T. Tal, M. Tanaka, D. Thomas, J. L. Tinker, R. Tojeiro, C. A. Tremonti, M. Vargas Magaña, L. Verde, M. Viel, D. A. Wake, M. Watson, B. A. Weaver, D. H. Weinberg, B. J. Weiner, A. A. West, M. White, W. M. Wood-Vasey, C. Yeche, I. Zehavi, G.-B. Zhao, and Z. Zheng. The Baryon Oscillation Spectroscopic Survey of SDSS-III. AJ, 145:10, January 2013.
- [29] A. De Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich. A model of diffuse Galactic radio emission from 10 MHz to 100 GHz. *MNRAS*, 388:247–260, July 2008.

- [30] S. Dodelson. *Modern cosmology*. 2003.
- [31] J.R. Driscoll and D.M. Healy. Computing fourier transforms and convolutions on the 2-sphere. *Advances in Applied Mathematics*, 15(2):202 250, 1994.
- [32] D. J. Eisenstein, I. Zehavi, D. W. Hogg, R. Scoccimarro, M. R. Blanton, R. C. Nichol, R. Scranton, H.-J. Seo, M. Tegmark, Z. Zheng, S. F. Anderson, J. Annis, N. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, A. Connolly, I. Csabai, M. Doi, M. Fukugita, J. A. Frieman, K. Glazebrook, J. E. Gunn, J. S. Hendry, G. Hennessy, Z. Ivezić, S. Kent, G. R. Knapp, H. Lin, Y.-S. Loh, R. H. Lupton, B. Margon, T. A. McKay, A. Meiksin, J. A. Munn, A. Pope, M. W. Richmond, D. Schlegel, D. P. Schneider, K. Shimasaku, C. Stoughton, M. A. Strauss, M. SubbaRao, A. S. Szalay, I. Szapudi, D. L. Tucker, B. Yanny, and D. G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633:560–574, November 2005.
- [33] J. S. Farnes, B. M. Gaensler, and E. Carretti. A Broadband Polarization Catalog of Extragalactic Radio Sources. ApJS, 212:15, May 2014.
- [34] L. Feng, R. Vaulin, and J. N. Hewitt. Detectability of Late-Time Radio Afterglows from Compact Binary Coalescence. *ArXiv e-prints*, May 2014.
- [35] G. B. Field. Excitation of the Hydrogen 21-CM Line. *Proceedings of the IRE*, 46:240–250, January 1958.
- [36] J. R. Fisher and R. B. Tully. Neutral hydrogen observations of a large sample of galaxies. *ApJS*, 47:139–200, December 1981.
- [37] B. Flaugher. The Dark Energy Survey. In APS April Meeting Abstracts, March 2012.
- [38] G. Folatelli, M. M. Phillips, C. R. Burns, C. Contreras, M. Hamuy, W. L. Freedman, S. E. Persson, M. Stritzinger, N. B. Suntzeff, K. Krisciunas, L. Boldt, S. González, W. Krzeminski, N. Morrell, M. Roth, F. Salgado, B. F. Madore, D. Murphy, P. Wyatt, W. Li, A. V. Filippenko, and N. Miller. The Carnegie Supernova Project: Analysis of the First Sample of Low-Redshift Type-Ia Supernovae. AJ, 139:120–144, January 2010.
- [39] W. Freudling, L. Staveley-Smith, B. Catinella, R. Minchin, M. Calabretta, E. Momjian, M. Zwaan, M. Meyer, and K. O'Neil. Deep 21 cm H I Observations at $z \approx 0.1$: The Precursor to the Arecibo Ultra Deep Survey. *ApJ*, 727:40, January 2011.
- [40] S. R. Furlanetto. The global 21-centimeter background from high redshifts. *MNRAS*, 371:867–878, September 2006.
- [41] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Phys. Rep.*, 433:181–301, October 2006.
- [42] R. Giovanelli and M. P. Haynes. Redshift surveys of galaxies. ARA&A, 29:499-541, 1991.

- [43] R. Giovanelli and M. P. Haynes. Extragalactic HI surveys. A&ARv, 24:1, December 2016.
- [44] R. Giovanelli, M. P. Haynes, B. R. Kent, P. Perillat, A. Saintonge, N. Brosch, B. Catinella, G. L. Hoffman, S. Stierwalt, K. Spekkens, M. S. Lerner, K. L. Masters, E. Momjian, J. L. Rosenberg, C. M. Springob, A. Boselli, V. Charmandaris, J. K. Darling, J. Davies, D. Garcia Lambas, G. Gavazzi, C. Giovanardi, E. Hardy, L. K. Hunt, A. Iovino, I. D. Karachentsev, V. E. Karachentseva, R. A. Koopmann, C. Marinoni, R. Minchin, E. Muller, M. Putman, C. Pantoja, J. J. Salzer, M. Scodeggio, E. Skillman, J. M. Solanes, C. Valotto, W. van Driel, and L. van Zee. The Arecibo Legacy Fast ALFA Survey. I. Science Goals, Survey Design, and Strategy. AJ, 130:2598–2612, December 2005.
- [45] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622:759–771, April 2005.
- [46] J. E. Gunn. On the Propagation of Light in Inhomogeneous Cosmologies. I. Mean Effects. *ApJ*, 150:737, December 1967.
- [47] J. Guy, M. Sullivan, A. Conley, N. Regnault, P. Astier, C. Balland, S. Basa, R. G. Carlberg, D. Fouchez, D. Hardin, I. M. Hook, D. A. Howell, R. Pain, N. Palanque-Delabrouille, K. M. Perrett, C. J. Pritchet, J. Rich, V. Ruhlmann-Kleider, D. Balam, S. Baumont, R. S. Ellis, S. Fabbro, H. K. Fakhouri, N. Fourmanoit, S. González-Gaitán, M. L. Graham, E. Hsiao, T. Kronborg, C. Lidman, A. M. Mourao, S. Perlmutter, P. Ripoche, N. Suzuki, and E. S. Walker. The Supernova Legacy Survey 3-year sample: Type Ia supernovae photometric distances and cosmological constraints. *A&A*, 523:A7, November 2010.
- [48] Z. Haiman, J. J. Mohr, and G. P. Holder. Constraints on Cosmological Parameters from Future Galaxy Cluster Surveys. *ApJ*, 553:545–561, June 2001.
- [49] C. G. T. Haslam, U. Klein, C. J. Salter, H. Stoffel, W. E. Wilson, M. N. Cleary, D. J. Cooke, and P. Thomasson. A 408 MHz all-sky continuum survey. I - Observations at southern declinations and for the North Polar region. A&A, 100:209–219, July 1981.
- [50] C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 MHz all-sky continuum survey. II The atlas of contour maps. *A&AS*, 47:1, January 1982.
- [51] M. P. Haynes and R. Giovanelli. Neutral hydrogen in isolated galaxies. IV Results for the Arecibo sample. *AJ*, 89:758–800, June 1984.
- [52] P. A. Henning, C. M. Springob, F. Day, R. Minchin, E. Momjian, B. Catinella, E. Muller, B. Koribalski, K. Masters, C. Pantoja, M. Putman, J. L. Rosenberg, S. Schneider, and L. Staveley-Smith. The ALFA Zone of Avoidance Survey. In R. Minchin and E. Momjian, editors, *The Evolution of Galaxies Through the Neutral Hydrogen Window*, volume 1035 of *American Institute of Physics Conference Series*, pages 246–248, August 2008.

- [53] W. Hillebrandt and J. C. Niemeyer. Type IA Supernova Explosion Models. ARA&A, 38:191–230, 2000.
- [54] J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417, June 1974.
- [55] J. A. Holtzman. Microwave background anisotropies and large-scale structure in universes with cold dark matter, baryons, radiation, and massive and massless neutrinos. *ApJS*, 71:1–24, September 1989.
- [56] J. A. Holtzman, J. Marriner, R. Kessler, M. Sako, B. Dilday, J. A. Frieman, D. P. Schneider, B. Bassett, A. Becker, D. Cinabro, F. DeJongh, D. L. Depoy, M. Doi, P. M. Garnavich, C. J. Hogan, S. Jha, K. Konishi, H. Lampeitl, J. L. Marshall, D. McGinnis, G. Miknaitis, R. C. Nichol, J. L. Prieto, A. G. Riess, M. W. Richmond, R. Romani, M. Smith, N. Takanashi, K. Tokita, K. van der Heyden, N. Yasuda, and C. Zheng. The Sloan Digital Sky Survey-II: Photometry and Supernova IA Light Curves from the 2005 Data. *AJ*, 136:2306–2320, December 2008.
- [57] D. C. Jacobs, B. J. Hazelton, C. M. Trott, J. S. Dillon, B. Pindor, I. S. Sullivan, J. C. Pober, N. Barry, A. P. Beardsley, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. Kim, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. The Murchison Widefield Array 21 cm Power Spectrum Analysis Methodology. *ApJ*, 825:114, July 2016.
- [58] N. Kaiser. The Pan-STARRS Survey Telescope Project. In Advanced Maui Optical and Space Surveillance Technologies Conference, page E9, 2007.
- [59] N. Kaiser, G. Wilson, and G. A. Luppino. Large-Scale Cosmic Shear Measurements. ArXiv Astrophysics e-prints, March 2000.
- [60] P. M. W. Kalberla, W. B. Burton, D. Hartmann, E. M. Arnal, E. Bajaja, R. Morras, and W. G. L. Pöppel. The Leiden/Argentine/Bonn (LAB) Survey of Galactic HI. Final data release of the combined LDS and IAR surveys with improved stray-radiation corrections. *A&A*, 440:775–782, September 2005.
- [61] V. K. Kapahi and S. Ananthakrishnan. Astronomy with the Giant Metrewave Radio Telescope (GMRT). *Bulletin of the Astronomical Society of India*, 23:265, September 1995.

- [62] R. Kessler, A. C. Becker, D. Cinabro, J. Vanderplas, J. A. Frieman, J. Marriner, T. M. Davis, B. Dilday, J. Holtzman, S. W. Jha, H. Lampeitl, M. Sako, M. Smith, C. Zheng, R. C. Nichol, B. Bassett, R. Bender, D. L. Depoy, M. Doi, E. Elson, A. V. Filippenko, R. J. Foley, P. M. Garnavich, U. Hopp, Y. Ihara, W. Ketzeback, W. Kollatschny, K. Konishi, J. L. Marshall, R. J. McMillan, G. Miknaitis, T. Morokuma, E. Mörtsell, K. Pan, J. L. Prieto, M. W. Richmond, A. G. Riess, R. Romani, D. P. Schneider, J. Sollerman, N. Takanashi, K. Tokita, K. van der Heyden, J. C. Wheeler, N. Yasuda, and D. York. First-Year Sloan Digital Sky Survey-II Supernova Results: Hubble Diagram and Cosmological Parameters. *ApJS*, 185:32–84, November 2009.
- [63] J. Kim. Direct reconstruction of spherical harmonics from interferometer observations of the cosmic microwave background polarization. *MNRAS*, 375:625–632, February 2007.
- [64] R. P. Kirshner. Foundations of supernova cosmology, page 151. 2010.
- [65] V. L. Krabbendam, J. H. Burge, C. F. Claver, B. Cuerden, W. Davison, W. J. Gressler, J. Kingsley, H. M. Martin, D. R. Neill, S. Olivier, D. Phillion, J. Sebag, D. Sweeney, and LSST Collaboration. LSST Telescope Design Developments. In *American Astronomical Society Meeting Abstracts*, volume 37 of *Bulletin of the American Astronomical Society*, page 1205, December 2005.
- [66] A. V. Kravtsov and S. Borgani. Formation of Galaxy Clusters. *ARA&A*, 50:353–409, September 2012.
- [67] J. Leaman, W. Li, R. Chornock, and A. V. Filippenko. Nearby supernova rates from the Lick Observatory Supernova Search - I. The methods and data base. *MNRAS*, 412:1419–1440, April 2011.
- [68] B. Leibundgut, R. Schommer, M. Phillips, A. Riess, B. Schmidt, J. Spyromilio, J. Walsh, N. Suntzeff, M. Hamuy, J. Maza, R. P. Kirshner, P. Challis, P. Garnavich, R. C. Smith, A. Dressler, and R. Ciardullo. Time Dilation in the Light Curve of the Distant Type IA Supernova SN 1995K. *ApJ*, 466:L21, July 1996.
- [69] A. Lewis and A. Challinor. Weak gravitational lensing of the CMB. *Phys. Rep.*, 429:1–65, June 2006.
- [70] D. Li. Summary of The FAST Project. In L. Qain and D. Li, editors, Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015, volume 502 of Astronomical Society of the Pacific Conference Series, pages 93–97, February 2016.
- [71] Y. Maan, A. A. Deshpande, V. Chandrashekar, J. Chennamangalam, K. B. Raghavendra Rao, R. Somashekar, G. Anderson, M. S. Ezhilarasi, S. Sujatha, S. Kasturi, P. Sandhya, J. Bauserman, R. Duraichelvan, S. Amiri, H. A. Aswathappa, I. V. Barve, G. Sarabagopalan, H. M. Ananda, C. Beaudet, M. Bloss, D. B. Dhamnekar, D. Egan, J. Ford, S. Krishnamurthy,

N. Mehta, A. H. Minter, H. N. Nagaraja, M. Narayanaswamy, K. O'Neil, W. Raja, H. Sahasrabudhe, A. Shelton, K. S. Srivani, H. V. Venugopal, and S. T. Viswanathan. RRI-GBT Multi-band Receiver: Motivation, Design, and Development. *ApJS*, 204:12, January 2013.

- [72] A. M. Martin, E. Papastergis, R. Giovanelli, M. P. Haynes, C. M. Springob, and S. Stierwalt. The Arecibo Legacy Fast ALFA Survey. X. The H I Mass Function and Ω_{-} H I from the 40% ALFALFA Survey. *ApJ*, 723:1359–1374, November 2010.
- [73] K. Masui, H.-H. Lin, J. Sievers, C. J. Anderson, T.-C. Chang, X. Chen, A. Ganguly, M. Jarvis, C.-Y. Kuo, Y.-C. Li, Y.-W. Liao, M. McLaughlin, U.-L. Pen, J. B. Peterson, A. Roman, P. T. Timbie, T. Voytek, and J. K. Yadav. Dense magnetized plasma associated with a fast radio burst. *Nature*, 528:523–525, December 2015.
- [74] J. D. McEwen and A. M. M. Scaife. Simulating full-sky interferometric observations. MN-RAS, 389:1163–1178, September 2008.
- [75] M. J. Meyer, M. A. Zwaan, R. L. Webster, L. Staveley-Smith, E. Ryan-Weber, M. J. Drinkwater, D. G. Barnes, M. Howlett, V. A. Kilborn, J. Stevens, M. Waugh, M. J. Pierce, R. Bhathal, W. J. G. de Blok, M. J. Disney, R. D. Ekers, K. C. Freeman, D. A. Garcia, B. K. Gibson, J. Harnett, P. A. Henning, H. Jerjen, M. J. Kesteven, P. M. Knezek, B. S. Koribalski, S. Mader, M. Marquarding, R. F. Minchin, J. O'Brien, T. Oosterloo, R. M. Price, M. E. Putman, S. D. Ryder, E. M. Sadler, I. M. Stewart, F. Stootman, and A. E. Wright. The HIPASS catalogue I. Data presentation. *MNRAS*, 350:1195–1209, June 2004.
- [76] G. Miknaitis, G. Pignata, A. Rest, W. M. Wood-Vasey, S. Blondin, P. Challis, R. C. Smith, C. W. Stubbs, N. B. Suntzeff, R. J. Foley, T. Matheson, J. L. Tonry, C. Aguilera, J. W. Blackman, A. C. Becker, A. Clocchiatti, R. Covarrubias, T. M. Davis, A. V. Filippenko, A. Garg, P. M. Garnavich, M. Hicken, S. Jha, K. Krisciunas, R. P. Kirshner, B. Leibundgut, W. Li, A. Miceli, G. Narayan, J. L. Prieto, A. G. Riess, M. E. Salvo, B. P. Schmidt, J. Sollerman, J. Spyromilio, and A. Zenteno. The ESSENCE Supernova Survey: Survey Optimization, Observations, and Supernova Photometry. *ApJ*, 666:674–693, September 2007.
- [77] M. F. Morales and M. Matejek. Software holography: interferometric data analysis for the challenges of next generation observatories. *MNRAS*, 400:1814–1820, December 2009.
- [78] L. B. Newburgh, K. Bandura, M. A. Bucher, T.-C. Chang, H. C. Chiang, J. F. Cliche, R. Davé, M. Dobbs, C. Clarkson, K. M. Ganga, T. Gogo, A. Gumba, N. Gupta, M. Hilton, B. Johnstone, A. Karastergiou, M. Kunz, D. Lokhorst, R. Maartens, S. Macpherson, M. Mdlalose, K. Moodley, L. Ngwenya, J. M. Parra, J. Peterson, O. Recnik, B. Saliwanchik, M. G. Santos, J. L. Sievers, O. Smirnov, P. Stronkhorst, R. Taylor, K. Vanderlinde, G. Van Vuuren, A. Weltman, and A. Witzemann. HIRAX: a probe of dark energy and radio transients. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 9906 of *Proc. SPIE*, page 99065X, August 2016.

- [79] C. P. O'Dea and H. E. Payne. An Arecibo Search for Neutral Hydrogen with Broad Line Widths in Clusters of Galaxies. In *Bulletin of the American Astronomical Society*, volume 23 of *BAAS*, page 1338, September 1991.
- [80] A. R. Offringa, B. McKinley, N. Hurley-Walker, F. H. Briggs, R. B. Wayth, D. L. Kaplan, M. E. Bell, L. Feng, A. R. Neben, J. D. Hughes, J. Rhee, T. Murphy, N. D. R. Bhat, G. Bernardi, J. D. Bowman, R. J. Cappallo, B. E. Corey, A. A. Deshpande, D. Emrich, A. Ewall-Wice, B. M. Gaensler, R. Goeke, L. J. Greenhill, B. J. Hazelton, L. Hindson, M. Johnston-Hollitt, D. C. Jacobs, J. C. Kasper, E. Kratzenberg, E. Lenc, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, N. Kudryavtseva, D. Oberoi, S. M. Ord, B. Pindor, P. Procopio, T. Prabu, J. Riding, D. A. Roshi, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and C. L. Williams. WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy. *MNRAS*, 444:606–619, October 2014.
- [81] S. S. Passmoor, C. M. Cress, and A. Faltenbacher. Clustering of H I galaxies in the H I Parkes All Sky Survey and Arecibo Legacy Fast ALFA Survey. *MNRAS*, 412:L50–L54, March 2011.
- [82] P. J. E. Peebles and J. T. Yu. Primeval Adiabatic Perturbation in an Expanding Universe. *ApJ*, 162:815, December 1970.
- [83] J. E. G. Peek, C. Heiles, K. A. Douglas, M.-Y. Lee, J. Grcevich, S. Stanimirović, M. E. Putman, E. J. Korpela, S. J. Gibson, A. Begum, D. Saul, T. Robishaw, and M. Krčo. The GALFA-HI Survey: Data Release 1. *ApJS*, 194:20, June 2011.
- [84] R. A. Perley. Imaging with Non-Coplanar Arrays. In G. B. Taylor, C. L. Carilli, and R. A. Perley, editors, *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, page 383, 1999.
- [85] R. A. Perley, P. J. Napier, and B. J. Butler. The Expanded Very Large Array: goals, progress, and plans. In J. M. Oschmann, Jr., editor, *Ground-based Telescopes*, volume 5489 of *Proc. SPIE*, pages 784–795, October 2004.
- [86] S. Perlmutter, S. Gabi, G. Goldhaber, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, R. Pain, C. R. Pennypacker, I. A. Small, R. S. Ellis, R. G. McMahon, B. J. Boyle, P. S. Bunclark, D. Carter, M. J. Irwin, K. Glazebrook, H. J. M. Newberg, A. V. Filippenko, T. Matheson, M. Dopita, and W. J. Couch. Measurements of the Cosmological Parameters Ω and Λ from the First Seven Supernovae at z ≥ 0.35. *ApJ*, 483:565–581, July 1997.
- [87] J. B. Peterson, K. Bandura, and U. L. Pen. The Hubble Sphere Hydrogen Survey. astroph/0606104, June 2006.
- [88] J. B. Peterson, U.-L. Pen, and X.-P. Wu. The Primeval Structure Telescope. *Modern Physics Letters A*, 19:1001–1008, 2004.

- [89] Planck Collaboration, R. Adam, N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, A. Bonaldi, L. Bonavera, J. R. Bond, J. Borrill, F. R. Bouchet, M. Bucher, C. Burigana, E. Calabrese, J.-F. Cardoso, J. Carron, H. C. Chiang, L. P. L. Colombo, C. Combet, B. Comis, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, J. Delabrouille, E. Di Valentino, C. Dickinson, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, F. Elsner, T. A. Enßlin, H. K. Eriksen, E. Falgarone, Y. Fantaye, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, A. Gruppuso, J. E. Gudmundsson, F. K. Hansen, G. Helou, S. Henrot-Versillé, D. Herranz, E. Hivon, Z. Huang, S. Ili, A. H. Jaffe, W. C. Jones, E. Keihänen, R. Keskitalo, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, A. Lähteenmäki, J.-M. Lamarre, M. Langer, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. López-Caniego, Y.-Z. Ma, J. F. Macías-Pérez, G. Maggio, A. Mangilli, M. Maris, P. G. Martin, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M.-A. Miville-Deschênes, D. Molinari, A. Moneti, L. Montier, G. Morgante, A. Moss, P. Naselsky, P. Natoli, C. A. Oxborrow, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, L. Patrizii, O. Perdereau, L. Perotto, V. Pettorino, F. Piacentini, S. Plaszczynski, L. Polastri, G. Polenta, J.-L. Puget, J. P. Rachen, B. Racine, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, M. Rossetti, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, G. Sirri, R. Sunyaev, A.-S. Suur-Uski, J. A. Tauber, M. Tenti, L. Toffolatti, M. Tomasi, M. Tristram, T. Trombetti, J. Valiviita, F. Van Tent, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, A. Zacchei, and A. Zonca. Planck intermediate results. XLVII. Planck constraints on reionization history. ArXiv e-prints, May 2016.
- [90] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. A&A, 571:A16, November 2014.
- [91] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13, September 2016.
- [92] J. Pober, A. Parsons, D. Backer, R. Bradley, C. Parashare, N. Gugliucci, E. Benoit, J. Aguirre, D. Jacobs, D. Moore, C. Carilli, J. Manley, and C. van der Meere. The Precision Array for Probing the Epoch of Reionization. In American Astronomical Society Meeting Abstracts #217, volume 43 of Bulletin of the American Astronomical Society, page 432.06, January 2011.

152

- [93] J. R. Pritchard and A. Loeb. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, August 2012.
- [94] R. J. Sault and T. A. Oosterloo. Imaging algorithms in radio interferometry. *ArXiv e-prints*, January 2007.
- [95] H.-J. Seo, S. Dodelson, J. Marriner, D. Mcginnis, A. Stebbins, C. Stoughton, and A. Vallinotto. A Ground-based 21 cm Baryon Acoustic Oscillation Survey. *ApJ*, 721:164– 173, September 2010.
- [96] J. R. Shaw, K. Sigurdson, U.-L. Pen, A. Stebbins, and M. Sitwell. All-sky Interferometry with Spherical Harmonic Transit Telescopes. *ApJ*, 781:57, February 2014.
- [97] J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen. Coaxing cosmic 21 cm fluctuations from the polarized sky using m -mode analysis. *Phys. Rev. D*, 91(8):083514, April 2015.
- [98] I. S. Sullivan, M. F. Morales, B. J. Hazelton, W. Arcus, D. Barnes, G. Bernardi, F. H. Briggs, J. D. Bowman, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. deSouza, D. Emrich, B. M. Gaensler, R. Goeke, L. J. Greenhill, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, E. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, N. Udaya Shankar, K. S. Srivani, J. Stevens, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. Fast Holographic Deconvolution: A New Technique for Precision Radio Interferometry. *ApJ*, 759:17, November 2012.
- [99] E. R. Switzer, K. W. Masui, K. Bandura, L.-M. Calin, T.-C. Chang, X.-L. Chen, Y.-C. Li, Y.-W. Liao, A. Natarajan, U.-L. Pen, J. B. Peterson, J. R. Shaw, and T. C. Voytek. Determination of z 0.8 neutral hydrogen fluctuations using the 21 cm intensity mapping autocorrelation. *MNRAS*, 434:L46–L50, July 2013.
- [100] C. Tasse, S. van der Tol, J. van Zwieten, G. van Diepen, and S. Bhatnagar. Applying full polarization A-Projection to very wide field of view instruments: An imager for LOFAR. A&A, 553:A105, May 2013.
- [101] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, R. C. Nichol, J. P. Ostriker, N. Padmanabhan, W. J. Percival, D. J. Schlegel, D. P. Schneider, R. Scoccimarro, U. Seljak, H.-J. Seo, M. Swanson, A. S. Szalay, M. S. Vogeley, J. Yoo, I. Zehavi, K. Abazajian, S. F. Anderson, J. Annis, N. A. Bahcall, B. Bassett, A. Berlind, J. Brinkmann, T. Budavari, F. Castander, A. Connolly, I. Csabai, M. Doi, D. P. Finkbeiner, B. Gillespie, K. Glazebrook, G. S. Hennessy, D. W. Hogg, Ž. Ivezić, B. Jain, D. Johnston, S. Kent, D. Q. Lamb, B. C. Lee, H. Lin, J. Loveday, R. H. Lupton, J. A. Munn, K. Pan, C. Park, J. Peoples, J. R. Pier,

A. Pope, M. Richmond, C. Rockosi, R. Scranton, R. K. Sheth, A. Stebbins, C. Stoughton, I. Szapudi, D. L. Tucker, D. E. vanden Berk, B. Yanny, and D. G. York. Cosmological constraints from the SDSS luminous red galaxies. *Phys. Rev. D*, 74(12):123507, December 2006.

- [102] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr. Interferometry and Synthesis in Radio Astronomy, 2nd Edition. 2001.
- [103] J. Tinker and SDSS-IV Collaboration. The Start of SDSS-IV and eBOSS. In American Astronomical Society Meeting Abstracts, volume 225 of American Astronomical Society Meeting Abstracts, page 125.06, January 2015.
- [104] J. A. Tyson, F. Valdes, and R. A. Wenk. Detection of systematic gravitational lens galaxy image alignments - Mapping dark matter in galaxy clusters. *ApJ*, 349:L1–L4, January 1990.
- [105] M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, M. Brentjens, A. Nelles, R. Beck, H. Falcke, R. Fender, J. Hörandel, L. V. E. Koopmans, G. Mann, G. Miley, H. Röttgering, B. W. Stappers, R. A. M. J. Wijers, S. Zaroubi, M. van den Akker, A. Alexov, J. Anderson, K. Anderson, A. van Ardenne, M. Arts, A. Asgekar, I. M. Avruch, F. Batejat, L. Bähren, M. E. Bell, M. R. Bell, I. van Bemmel, P. Bennema, M. J. Bentum, G. Bernardi, P. Best, L. Bîrzan, A. Bonafede, A.-J. Boonstra, R. Braun, J. Bregman, F. Breitling, R. H. van de Brink, J. Broderick, P. C. Broekema, W. N. Brouw, M. Brüggen, H. R. Butcher, W. van Cappellen, B. Ciardi, T. Coenen, J. Conway, A. Coolen, A. Corstanje, S. Damstra, O. Davies, A. T. Deller, R.-J. Dettmar, G. van Diepen, K. Dijkstra, P. Donker, A. Doorduin, J. Dromer, M. Drost, A. van Duin, J. Eislöffel, J. van Enst, C. Ferrari, W. Frieswijk, H. Gankema, M. A. Garrett, F. de Gasperin, M. Gerbers, E. de Geus, J.-M. Grießmeier, T. Grit, P. Gruppen, J. P. Hamaker, T. Hassall, M. Hoeft, H. A. Holties, A. Horneffer, A. van der Horst, A. van Houwelingen, A. Huijgen, M. Iacobelli, H. Intema, N. Jackson, V. Jelic, A. de Jong, E. Juette, D. Kant, A. Karastergiou, A. Koers, H. Kollen, V. I. Kondratiev, E. Kooistra, Y. Koopman, A. Koster, M. Kuniyoshi, M. Kramer, G. Kuper, P. Lambropoulos, C. Law, J. van Leeuwen, J. Lemaitre, M. Loose, P. Maat, G. Macario, S. Markoff, J. Masters, R. A. McFadden, D. McKay-Bukowski, H. Meijering, H. Meulman, M. Mevius, E. Middelberg, R. Millenaar, J. C. A. Miller-Jones, R. N. Mohan, J. D. Mol, J. Morawietz, R. Morganti, D. D. Mulcahy, E. Mulder, H. Munk, L. Nieuwenhuis, R. van Nieuwpoort, J. E. Noordam, M. Norden, A. Noutsos, A. R. Offringa, H. Olofsson, A. Omar, E. Orrú, R. Overeem, H. Paas, M. Pandey-Pommier, V. N. Pandey, R. Pizzo, A. Polatidis, D. Rafferty, S. Rawlings, W. Reich, J.-P. de Reijer, J. Reitsma, G. A. Renting, P. Riemers, E. Rol, J. W. Romein, J. Roosjen, M. Ruiter, A. Scaife, K. van der Schaaf, B. Scheers, P. Schellart, A. Schoenmakers, G. Schoonderbeek, M. Serylak, A. Shulevski, J. Sluman, O. Smirnov, C. Sobey, H. Spreeuw, M. Steinmetz, C. G. M. Sterks, H.-J. Stiepel, K. Stuurwold, M. Tagger, Y. Tang, C. Tasse, I. Thomas, S. Thoudam, M. C. Toribio, B. van der Tol, O. Usov, M. van Veelen, A.-J. van der Veen, S. ter Veen, J. P. W. Verbiest, R. Vermeulen, N. Vermaas,

C. Vocks, C. Vogt, M. de Vos, E. van der Wal, R. van Weeren, H. Weggemans, P. Weltevrede, S. White, S. J. Wijnholds, T. Wilhelmsson, O. Wucknitz, S. Yatawatta, P. Zarka, A. Zensus, and J. van Zwieten. LOFAR: The LOw-Frequency ARray. *A&A*, 556:A2, August 2013.

- [106] L. Van Waerbeke, Y. Mellier, T. Erben, J. C. Cuillandre, F. Bernardeau, R. Maoli, E. Bertin, H. J. McCracken, O. Le Fèvre, B. Fort, M. Dantel-Fort, B. Jain, and P. Schneider. Detection of correlated galaxy ellipticities from CFHT data: first evidence for gravitational lensing by large-scale structures. A&A, 358:30–44, June 2000.
- [107] K. Vanderlinde and Chime Collaboration. The Canadian Hydrogen Intensity Mapping Experiment (CHIME). In *Exascale Radio Astronomy*, volume 2, April 2014.
- [108] R. Wielebinski, N. Junkes, and B. H. Grahl. The Effelsberg 100-m Radio Telescope: Construction and Forty Years of Radio Astronomy. *Journal of Astronomical History and Heritage*, 14:3–21, March 2011.
- [109] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell'Antonio, and G. Bernstein. Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405:143–148, May 2000.
- [110] W. M. Wood-Vasey, G. Miknaitis, C. W. Stubbs, S. Jha, A. G. Riess, P. M. Garnavich, R. P. Kirshner, C. Aguilera, A. C. Becker, J. W. Blackman, S. Blondin, P. Challis, A. Clocchiatti, A. Conley, R. Covarrubias, T. M. Davis, A. V. Filippenko, R. J. Foley, A. Garg, M. Hicken, K. Krisciunas, B. Leibundgut, W. Li, T. Matheson, A. Miceli, G. Narayan, G. Pignata, J. L. Prieto, A. Rest, M. E. Salvo, B. P. Schmidt, R. C. Smith, J. Sollerman, J. Spyromilio, J. L. Tonry, N. B. Suntzeff, and A. Zenteno. Observational Constraints on the Nature of Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey. *ApJ*, 666:694–715, September 2007.
- [111] E. L. Wright. Theoretical Overview of Cosmic Microwave Background Anisotropy. *Measuring and Modeling the Universe*, page 291, 2004.
- [112] Y. Xu, X. Wang, and X. Chen. Forecasts on the Dark Energy and Primordial Non-Gaussianity Observations with the Tianlai Cylinder Array. *ApJ*, 798:40, January 2015.
- [113] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F.

Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120:1579–1587, September 2000.

- [114] H.-R. Yu, T.-J. Zhang, and U.-L. Pen. Method for Direct Measurement of Cosmic Acceleration by 21-cm Absorption Systems. *Physical Review Letters*, 113(4):041303, July 2014.
- [115] M. Zaldarriaga and U. Seljak. All-sky analysis of polarization in the microwave background. *Phys. Rev. D*, 55:1830–1840, February 1997.
- [116] J. Zhang, R. Ansari, X. Chen, J.-E. Campagne, C. Magneville, and F. Wu. Sky reconstruction from transit visibilities: PAON-4 and Tianlai dish array. *MNRAS*, 461:1950–1966, September 2016.
- [117] J. Zhang, S.-F. Zuo, R. Ansari, X. Chen, Y.-C. Li, F.-Q. Wu, J.-E. Campagne, and C. Magneville. Sky reconstruction for the Tianlai cylinder array. *Research in Astronomy and Astrophysics*, 16:158, October 2016.
- [118] P. Zhang, U.-L. Pen, and B. Wang. The Sunyaev-Zeldovich Effect: Simulations and Observations. ApJ, 577:555–568, October 2002.
- [119] M. A. Zwaan, M. J. Meyer, L. Staveley-Smith, and R. L. Webster. The HIPASS catalogue: Ω_{HI} and environmental effects on the HI mass function of galaxies. *MNRAS*, 359:L30–L34, May 2005.

Appendices

Appendix A

One-dimension reconstruction

This appendix summarises the early work I did, when I started my PhD work at LAL, in order to understand the key features of reconstructing sky maps from visibilities, specially the impact of combination of multiple pointings in declination and irregular array spacing along the north-south direction. The formalism used here is a simplified version of the one presented in chapter 3, section 3.3 in planar geometry, restricted to one dimension, either along the east-west (right ascension) or the north-south (declination) direction. The main reason to have included this material here is the insight that can be gained from the analysis of the response matrix, relating the reconstructed sky modes to the true ones, in the simple cases where the two angular directions are decoupled. we are interested in evaluating the ability of an instrument setup and scan strategy to measure single u or v modes, or a relatively narrow band in the angular frequency plane.

The notation I have used in this appendix corresponds to my original notation which differs slightly from the one used throughout chapter 3. Indeed, the projection matrix relating visibilities to sky modes, corresponding is denoted A here instead of L in chapter 3, and its pseudo-inverse is denoted B instead of H.

A.1 East-West direction

We discuss first the case of East-West direction. The suitable case is that all of the antennas are pointing at same right ascension. Here, we set up the pointing upward perpendicular to the East-West direction. Then, the instantaneous visibility is just a complex number.

$$\mathcal{V}_{ij} = \int I(\alpha) L(\alpha) e^{i2\pi\alpha \frac{\Delta x}{\lambda}} d\alpha \qquad (A.1.1)$$

where α is angular coordinate along East-West direction, and Δx is the baseline of a pair of antennas, along this direction.

Actually, the pointing are shifted along EW direction as sky rotates. If we consider that the axis of the antennas is in the direction $\alpha(t)$ and in the approximation of a small lobe, we can

perform it as a function of time due to the earth rotation,

$$\mathcal{V}_{ij}(\alpha(t)) = \int I(\alpha + \alpha(t)) L(\alpha) e^{i2\pi\alpha \frac{\Delta x}{\lambda}} d\alpha$$

= $\int du \mathcal{I}(u) \mathcal{L}_{ij}(u - u_0) e^{i2\pi\alpha(t)u}$ (A.1.2)

where $\mathcal{I}(u)$ and $\mathcal{L}_{ij}(u-u_0)$ are the Fourier transform of sky intensity and antenna lobe for angular frequency/wavenumber $u_0 = \frac{\Delta x}{\lambda}$. If u_0 is zero, the Fourier lobe corresponds auto-correlation. If not, it should be cross correlation.

As sky rotate, a phase factor $\exp(2i\pi\alpha(t)u)$ is appeared in Fourier dominance. We consider that the sky is periodic. So we could replace Fourier transform by discrete Fourier transform.

$$\mathcal{V}_{ij}(\alpha(t)) = \sum_{u_i} \mathcal{I}(u_i) \mathcal{L}_{ij}(u_i - u_0) e^{i2\pi\alpha(t)u_i}$$
(A.1.3)

The equation A.1.3 is linear system of equations. The measurements are visibilities, which are marked as known quantities. Assuming that the antenna lobe for one dish with diameter D use approximation $L(\alpha) = [\frac{\sin(\pi(D/\lambda)\alpha)}{\pi(D/\lambda)\alpha}]^2$ in angular plane corresponding to triangle shape $\bigwedge_{[\pm D/\lambda]}(u)$ in Fourier plane. All of the antennas are pointing to a given direction without moving, so the lobe would be also known quantities, and the phase factors are the function of time due to sky rotation. The remains are Fourier sky map $\mathcal{I}(u)$ which we name for "observed sky" and prepare to calculate from above linear system of equations. The equationA.1.4 is shown again in matrix form as follows,

$$\begin{pmatrix} V_{ij}^{\alpha_1} \\ V_{ij}^{\alpha_2} \\ \cdots \\ V_{ij}^{\alpha_m} \end{pmatrix} = \begin{pmatrix} \mathcal{L}_{ij}^{u_1} e^{i2\pi\alpha_1 u_1} & \mathcal{L}_{ij}^{u_2} e^{i2\pi\alpha_1 u_2} \cdots & \mathcal{L}_{ij}^{u_n} e^{i2\pi\alpha_1 u_n} \\ \mathcal{L}_{ij}^{u_1} e^{i2\pi\alpha_2 u_1} & \mathcal{L}_{ij}^{u_2} e^{i2\pi\alpha_2 u_2} \cdots & \mathcal{L}_{ij}^{u_n} e^{i2\pi\alpha_2 u_n} \\ \cdots \\ \mathcal{L}_{ij}^{u_1} e^{i2\pi\alpha_m u_1} & \mathcal{L}_{ij}^{u_2} e^{i2\pi\alpha_m u_2} \cdots & \mathcal{L}_{ij}^{u_n} e^{i2\pi\alpha_m u_n} \end{pmatrix} \times \begin{pmatrix} \mathcal{I}(u_1) \\ \mathcal{I}(u_2) \\ \cdots \\ \mathcal{I}(u_n) \end{pmatrix}$$
(A.1.4)

To illustrate the reconstruction convenient, the matrix $[\mathcal{L}_{ij}(u_i - u_0) \exp(i2\pi\alpha u_i)]$ is symbolized as Matrix A. The next step should be solving this linear system of equation.

For a set of visibility measurements for regularly spaced values of right ascension angles α_j covering the full range of right ascension values, the phase terms $\exp(2i\pi\alpha_j u_i)$ correspond to a Fourier series transform. The **A** matrix above could then be factorise as the beam part times a DFT (Discrete Fourier Transform) part. The DFT part can be inverted directly, leading to the separation of the problem into a set of separate linear systems, one for each u - mode.

A.1.1 Pseudo-inverse of matrix A

In linear algebra, the singular value decomposition (SVD) is a factorization of a real or complex matrix, with many useful applications in signal processing and statistics. The SVD could be used for computing pseudo-inverse of a matrix.

A.1. EAST-WEST DIRECTION

The singular value decomposition of an $m \times n$ real or complex matrix A is a factorization of the form

$$A = U\Sigma V^*$$

where U is $m \times m$ real or complex unitary matrix, Σ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and V* (the conjugate transpose of V, or simply the transpose of V if V is real) is an $n \times n$ real or complex unitary matrix. The diagonal entries Σ_{ii} of Σ are known as the singular values of A. A common convention is to list the singular values in descending order. In this case, the diagonal matrix Σ is uniquely determined by A (though the matrixes U and V do not know.)

The pseudo-inverse of the matrix A using SVD is given by

$$B = A^{-1} = V\Sigma^+ U^* \tag{A.1.5}$$

where Σ^+ is formed by replacing every nonzero diagonal λ_i entry by its reciprocal $1/\lambda_i$ and transposing the resulting matrix. We use the symbol B matrix as the pseudo-inverse of the matrix A.

Due to limited numerical precision in the computation, even zero elements of $\Sigma_{i,i}$ will have some small non-zero value, which would give rise to large $\Sigma_{i,i}^{-1}$ and affect the result greatly if left unattended. To ensure the stability of the computation, the small eigenvalues of Σ are set to 0 before the inversion, and its inverse also set to 0 and ignored in subsequent computation. In practice, we set some threshold value for the diagonal elements. If the diagonal element $\Sigma_{ii} < \Sigma_{00} \times$ threshold_ratio, we will put it to 0, corresponding Σ^+ element put to zero.

Now, the reconstructed sky could be expressed as follows,

$$\left(\begin{array}{c} \cdots\\ \widetilde{\mathcal{I}(u_i)}\\ \cdots \end{array}\right) = B \times \left(\begin{array}{c} \cdots\\ V_{ij}(a_t)\\ \cdots \end{array}\right) = B \times A \times \left(\begin{array}{c} \cdots\\ \mathcal{I}(u_i)\\ \cdots \end{array}\right)$$

here, the matrix $[\mathcal{I}(u_i)]$ is the reconstructed sky or "observed sky" and the matrix $[\mathcal{I}(u_i)]$ is the true sky. The matrix B * A will tell us how well we can recover the sky and how well the mode mixing is under control. If the matrix B * A is identity matrix, the reconstruction would be perfect. At the currently step, we set the eigenvalue threshold ratio as 0.005 which will be discussed in detail for 2D reconstruction.

A.1.2 Some results of response matrix B * A (EW)

For the East-West direction, the results will be similar for both dish and cylinder case, given similar dimension along the EW axis. So, here, we only show the results for dish array configuration. There are 4 dishes layout in line along East-West direction with uniform spacing $d_{sep} = 2m$, dish size D = 1.75m and frequency $\nu = 1200$ MHz. Then, the beam shape in Fourier space is show as Fig A.1.1.

We are mainly concerning about the B * A matrix, which we name it as response matrix. As remarked earlier, the beam pointing are shifted due to sky rotate. Here, we consider that there



Figure A.1.1: The beam patterns in Fourier space with $d_{sep} = 2m$, dish size D = 1.75m and frequency $\nu = 1200$ MHz.

are totally 32 sky pointings covering 2π radians period with dish immobilized. Based on these conditions, we could compute the B * A matrix, please see Fig A.1.2. We could find that the B * A matrix are identity matrix within the $u_{max} \sim 31$ which depend on array resolution. It means that we would be able to recover the sky Fourier u-mode completely from the observations within the telescopes resolution. And there are no mode mixing between different u-modes with full sky coverage.

It is worth noting that we have chosen compact array layout. If not, there are some gaps between different *u*-modes, which we will discuss in next paragraphs.

A.2 North-South direction

We have discussed the East-West direction reconstruction. Now, we investigate the North-South direction reconstruction, which correspond to the declination direction with angle β denoting it n the planar, pseudo-cartesian geometry (see chapter 3, section 3.3). There are no pointing shifted effect along the North-South direction, and the practical situation, i.e. survey strategy, is different for dish and cylinder case. For dish array, we could select several pointings, and each time, all



Figure A.1.2: The B * A matrix.

of the antennas point to same direction or declination. However, for cylinder array, There are no other choice, except pointing upward perpendicular to the North-South earth surface. Both of them correspond to partial sky coverage.

So, according to the pointing effect, we could express the visibilities as a function of pointing β_p , angle in the North-South plane (or the elevation angle)

$$\mathcal{V}_{ij}(\beta_p) = \int I(\beta + \beta_p) L(\beta) e^{i2\pi\delta \frac{\Delta y}{\lambda}} d\beta$$

= $\int dv \mathcal{I}(v) \mathcal{L}_{ij}(v - v_0) e^{i2\pi\beta_p v}$ (A.2.1)

where δ is angular coordinate along North-South direction, and Δy is the baseline of a pair of antennas. $\mathcal{I}(v)$ and $\mathcal{L}_{ij}(v - v_0)$ are also the Fourier transform of sky intensity and antenna lobe for angular frequency/wavenumber $v_0 = \frac{\Delta y}{\lambda}$.

We consider that the sky is also periodic along North-South direction. So we could replace Fourier transform by discrete Fourier transform.

$$\mathcal{V}_{ij}(\delta_p) = \sum_{v_i} \mathcal{I}(v_i) \,\mathcal{L}_{ij}(v_i - v_0) e^{i2\pi\beta_p v_i} \tag{A.2.2}$$

The equation A.2.2 is linear system of equations. These are similar as East-West direction case, and its matrix form are same as equation A.1.4. We also assume that the antenna lobe for one dish with diameter D use approximation $L(\delta) = [\frac{\sin(\pi(D/\lambda)\delta)}{\pi(D/\lambda)\delta}]^2$ in angular plane corresponding to triangle shape $\Lambda_{[\pm D/\lambda]}(v)$ in Fourier plane. The next steps are calculating the pseudo-inverse of matrix $A = [\mathcal{L}_{ij}(v_i - v_0) \exp(i2\pi\delta_p v_i)]$, using SVD method, and computing the response matrix B * A.



Figure A.2.1: The response matrix B * A for the case of 24 feeds layout in line along North-South direction, and each feed has a single dipole $\sim 0.4m$ long, and feeds are separated by 1.6m.



Figure A.2.2: The diagonal element of the Fig A.2.1, blue and red correspond to the two frequencies ν_1 , ν_2 .



Figure A.2.3: The response matrix B * A for the case of 24 feeds layout in line along North-South direction, each feed is made of 4 packed dipoles, with the output signal corresponding to the sum of individual dipole signals.

A.2. NORTH-SOUTH DIRECTION

A.2.1 Some results of response matrix B * A (NS)

The question we would like to answer here is how close is the matrix B*A to the identity, or at least a diagonal matrix. The results presented here will bring some answers to this question. In this case, given the different beam patterns for dish and feeds on cylinders along the north-south direction, we have to study each case separately. The results are thus shown to both dish and cylinder case.

Cylinder case

First, we consider two kinds of cylinder array configurations, with single pointing. The first case contains 24 feeds layout in line along North-South direction, and each feed have single dipole $\sim 0.4m$ long, separated by 1.6m. We consider two frequencies: $\nu_1 \sim 1300$ MHz and $\nu_2 = 1.12 \times \nu_1$.

Fig A.2.1 give us the information of response matrix B * A for first cylinder case. We could see that there are gap between different v-modes and these gaps are shifted for different frequency. The Fig A.2.2 show the diagonal element of the Fig A.2.1, which shows clearly that the gaps are shifted. This kind of Fig A.2.1 will lead to stronger beam frequency dependence and mode mixing.

So we consider the second kinds of cylinder case which also have 24 feeds layout in line along North-South direction, but each feed is made of 4 packed dipoles, 0.4m long with the output signal corresponding to the sum of individual dipole signals, and the feeds are separated by 1.6m. Fig A.2.3 illustrate the information of response matrix B * A for compact cylinder array case. We can see that the gaps are significantly smaller.

Dish case

We discuss now two dish array configurations. We consider an array of 12 dishes in line along NS direction. The first configuration correspond to uniform spacing, while the second configuration corresponds to the random spacing, as sketched in FigA.2.4. We that the single dish beam width is ~ $6rad^{-1}$ and the constant spacing of ~ $8rad^{-1}$ in Fourier space for the regular array. As for the cylinder case, we consider two frequencies nu_1 and $\nu 2 = 1.12 \times \nu 1$. The input sky is supposed to have 30^o extension along the north-south (NS) direction. We have also considered two scanning strategies, one made of a single pointing, and a second one composed of 5 pointings $\beta_k, k = 1 \dots 5$ regularly spaced over the 30^o of the sky. The four configurations (C1,C2,C3,C4) discussed here are thus as follows:

- C1: 12 dish, with uniform spacing and 1 pointing, corresponding to Fig A.2.5.
- C2: 12 dish, with uniform spacing and 5 pointing, corresponding to Fig A.2.6.
- C3: 12 dish, with random spacing and 1 pointing, corresponding to Fig A.2.7.
- C4: 12 dish, with random spacing and 5 pointing, corresponding to Fig A.2.8.

In the FigA.2.5, we see that the B*A matrix starts to have some blocks along the diagonal. Individual v - modes are not measured any more, and each sky v - mode estimate is a linear



APPENDIX A. ONE-DIMENSION RECONSTRUCTION

Figure A.2.4: The sketch of dish array configurations.

combination of true sky modes, as expected from a partial sky coverage (single pointing. The are also some gaps, due to dish spacing being larger than the dish size. We can see on the right panel of this figure the blocks are slightly shifted along the diagonal, as the frequency is increased from ν_1 to ν_2 . This will cause some mode mixing which we would like to avoid.

By increasing the sky coverage with multiple pointing, we see that we have been able to decrease the size of the blocs along the diagonal of the B*A matrix, getting closer to a diagonal matrix. The B*A matrix for the two frequencies for configuration C2, regular array and 5 β_k pointings are shown in figure A.2.6.

We prefer to know the mode mixing effect for multiply pointing case. FigA.2.6 shows the matrices B * A for 5 pointings. We could see also extending for larger frequency. However, there are no more mode mixing for more pointing case.

Fig A.2.7 and Fig A.2.8 shows the matrices B * A for configurations C3 and C4, corresponding to 12 dishes random spacing, respectively with 1 pointing and 5 pointings and for two frequencies. We see that the response matrix B*A gets much closer to a diagonal matrix, meaning that we can reconstruct single v - modes, minimising the mode mixing. Considering the results shown in figure A.2.8 the combination of irregular spacing with several pointings provides a response matrix close to a diagonal matrix, at least for small u-modes.

In conclusion, for East-West direction, due to the earth rotation, we have access to a full sky coverage with regular sampling of the right ascension angles. All sky u - modes accessible to the instrument are thus On the other hand for north-south (NS) direction, due to the partial sky coverage, the recovered v-modes correspond to linear combination of true sky modes. However, combining non uniform spacing and multiple pointings with compact array configurations, we can enhance the sky map reconstruction quality, and decrease the mode mixing.



Figure A.2.5: The response matrix B * A for 12 dish, with uniform spacing and 1 pointing.



Figure A.2.6: The response matrix B * A for 12 dish, with uniform spacing and 5 pointing.



Figure A.2.7: The response matrix B * A for 12 dish, with random spacing and 1 pointing.



Figure A.2.8: The response matrix B * A for 12 dish, with random spacing and 5 pointing.

Appendix B

Map making software

The map reconstruction software and associated tools (**JSkyMap**) are written in C++ and use the SOPHYA (SOftware for PHYsics Analysis)¹ class library. SOPHYA is a collection of C++ classes designed to ease data analysis software development and provide the following services to the map making software:

- Input/Output services in different formats, including ASCII, FITS and the native PPF SO-PHYA binary format
- Various standard numerical analysis algorithms, including FFT and linear algebra and interface to LAPACK
- Several map pixelisation in spherical geometry, including the HEALPix format [45]
- Spherical Harmonics Transform
- Classes to perform parallel computation

The SOPHYA documentation and instructions to get and build the library and associated tools is available on the web site: http://www.sophya.org . The source code itself is available on gitalb: https://gitlab.in2p3.fr/SOPHYA . The JSkyMap software is also available from gitalb: https://gitlab.in2p3.fr/SCosmoTools/JSkyMap

We have developed the sky map reconstruction code both in rectangular geometry ((u, v)) plane) and spherical geometry $((\ell, m))$ plane). The rectangular geometry can be used when observing a narrow band in declination, at low declinations. We have performed a number of cross checks, for example computing visibilities in spherical geometry and performing the reconstruction in rectangular geometry. The code has similar structure in the two geometry. For the sake of clarity, we present here only the map reconstruction code in spherical geometry. The software is organised around few main classes:

¹http://www.sophya.org



Figure B.0.1: Block Diagram for the map-reconstruction code

- BeamTP and BeamLM classes which represent the beam for a single feed or a pair of feeds in angular ($\vec{\omega} = (\theta, \varphi)$) domain in spherical geometry and in the spherical Harmonics coefficient domain ((ℓ, m) plane). The BeamVis class computes the (ℓ, m) plane response for a pair of feeds/antenna given the the baseline, and the array position in latitude.
- The template class PseudoInverse provides the specific services to handle the computation of \mathbf{B}_m matrices and the noise covariance matrices.
- Utility classes and function to handle the computation of the set of baselines from the antenna positions in an array.
- The JSphSkyMap is the main class in the reconstruction code. It computes the A_m and B_m matrices, starting from a set of beams in the (ℓ, m) plane corresponding to an instrument layout and sky scanning strategy. It provides also methods to computes mock visibility data, given an input sky, as well as methods to reconstruct the sky from visibilities. This class implements parallelism at the level of A_m and B_m matrices for different ℓ .

The overall functional structure of the **JSkyMap** code is shown in Fig.B.0.1. The major computation steps are listed below:

1. We compute first the baselines from the array configuration, i.e. the coordinates of the array elements. To simplify numerical handling, we replace redundant baselines by a single beam, scaling the noise level accordingly $\propto 1/n_{rb}$, where n_{rb} is the number of antenna pairs with the exact same baseline.

- 2. We compute then the beams in the (ℓ, m) plane from the baselines and scan strategy (the observed declinations on sky). As the beam computation involves Spherical Harmonics Transform (SHT) which is computation intensive, multi-threaded computation has been implemented for this step.
- 3. One can then compute the A_m and the visibilities noise covariance matrices for each *m*-mode and then the corresponding B_m matrix, using the pseudo-inverse computation. The computation for theses steps benefits also from multo-threaded implementation, taking advantage of *m*-mode parallelism.
- 4. Mock visibilities (with or without noise) can be computed using the A_m matrices.
- 5. Finally \mathbf{B}_m matrices can be used to compute estimated Sky spherical harmonics coefficients from mock or observed visibilities, as well as the corresponding error covariance matrix. The sky map is recovered using an inverse SHT.

Appendix C

The two published papers

Two papers have been published in MNRAS (Monthly Notices of the Royal Astronomical Society) and RAA (Research in Astronomy and Astrophysics) based on the work presented in chapter 3 and 4 of this dissertation.

- Sky reconstruction from transit visibilities: PAON-4 and Tianlai dish array. MNRAS, Volume 461, Issue 2, p.1950-1966 [116]
- Sky reconstruction for the Tianlai cylinder array. RAA, Volume 16, Issue 10, article id. 158 [117]

The first page of each of the two publications, with the list of authors and the abstract are inserted here.

Sky reconstruction from transit visibilities: PAON-4 and Tianlai Dish Array

Jiao Zhang^{1,2,3}, Reza Ansari² *, Xuelei Chen^{1,3,4}, Jean-Eric Campagne²,

Christophe Magneville⁵, and Fengquan Wu¹

¹Key Laboratory of Computational Astrophysics, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

²Université Paris-Sud, LAL, UMR 8607, F-91898 Orsay Cedex, France & CNRS/IN2P3, F-91405 Orsay, France

³University of Chinese Academy of Sciences, Beijing 100049, China

⁴Centre for High Energy Physics, Peking University, Beijing 100871, China

⁵CEA, DSM/IRFU, Centre d'Etudes de Saclay, F-91191 Gif-sur-Yvette, France

Accepted 2016 June 15. Received 2016 June 14; in original form 2016 March 27

ABSTRACT

The spherical harmonics *m*-mode decomposition is a powerful sky map reconstruction method suitable for radio interferometers operating in transit mode. It can be applied to various configurations, including dish arrays and cylinders. We describe the computation of the instrument response function, the point spread function (PSF), transfer function, the noise covariance matrix and noise power spectrum. The analysis in this paper is focused on dish arrays operating in transit mode. We show that arrays with regular spacing have more pronounced side lobes as well as structures in their noise power spectrum, compared to arrays with irregular spacing, specially in the north-south direction. A good knowledge of the noise power spectrum $C^{\text{noise}}(\ell)$ is essential for intensity mapping experiments as non uniform $C^{\text{noise}}(\ell)$ is a potential problem for the measurement of the HI power spectrum. Different configurations have been studied to optimise the PAON-4 and Tianlai dish array layouts. We present their expected performance and their sensitivities to the 21-cm emission of the Milky Way and local extragalactic HI clumps.

Key words: techniques: interferometric – methods: data analysis – methods: numerical – cosmology: observations – (cosmology:) large-scale structure of Universe – radio lines: galaxies

1 INTRODUCTION

Measurement of the neutral hydrogen (HI) distribution through its 21-cm line radiation is a powerful method for studying the statistical properties of Large Scale Structure (LSS) in the Universe, complementary to optical surveys. However, given the very faint radio brightness of typical HI clumps, detection of individual galaxies in 21-cm at cosmological distances ($z \gtrsim 1$) requires very large collecting areas, around ~ km². Moreover, extracting cosmological information from LSS requires the observation of large volumes of universe to probe long wavelength modes with sufficient precision in order to be competitive with the optical galaxy surveys. In recent years, the intensity mapping technique has been suggested as an efficient and economical way to map large volumes of the universe using the HI 21-cm emission. Such cosmological surveys would be especially suitable for late time cosmological studies ($z \lesssim 3$), in particular to constrain dark energy through the Baryon Acoustic

Oscillations (BAO) and Redshift Space Distortions (RSD) measurements (Peterson et al. 2006; Chang et al. 2008; Ansari et al. 2009; Ansari et al. 2012b; Seo et al. 2010). In this scheme, the integrated radio emission of many HI clumps in cells of ~ 10^3 Mpc³ is measured without detection of individual galaxies. Large wide-field radio telescopes, with an angular resolution of a fraction of a degree and a frequency resolution of \lesssim 1MHz and sensitivities of \lesssim 1 mK per resolution element would be needed to observe the LSS, especially the BAO features.

Several groups throughout the world are aiming to carry such surveys. A number of projects with single dishes (possibly equipped with multi-beam receivers) and interferometer arrays have been proposed. Single dish intensity mapping surveys have been carried out on existing telescopes such as the Green Bank Telescope (GBT) (Chang et al. 2010; Switzer et al. 2013; Masui et al. 2013) and construction of dedicated instruments are being planned such as the BINGO (BAO from Integrated Neutral Gas Observations) project which is a single dish radio telescope equipped with an array of feeds in the focal plane (Battye et al. 2013; Dickinson 2014). The interferometer arrays include CHIME (Canadian Hydrogen Map-

^{*} E-mail:ansari@lal.in2p3.fr

Sky reconstruction for the Tianlai cylinder array

Jiao Zhang^{1,2,3}, Shifan Zuo^{1,3}, Reza Ansari², Xuelei Chen^{1,3,4}, Yichao Li^{1,3}, Fengquan Wu^{1,3}, Jean-Eric Campagne², Christophe Magneville⁵

- ¹ Key Laboratory of Computational Astrophysics, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China
- ² Université Paris-Sud, LAL, UMR 8607, F-91898 Orsay Cedex, France & CNRS/IN2P3, F-91405 Orsay, France
- ³ University of Chinese Academy of Sciences, Beijing 100049, China
- ⁴ Center for High Energy Physics, Peking University, Beijing 100871, China
- ⁵ CEA, DSM/IRFU, Centre d'Etudes de Saclay, F-91191 Gif-sur-Yvette, France

Abstract In this paper, we apply our sky map reconstruction method for transit type interferometers to the Tianlai cylinder array. The method is based on the spherical harmonic decomposition, and can be applied to cylindrical array as well as dish arrays and we can compute the instrument response, synthesised beam, transfer function and the noise power spectrum. We consider cylinder arrays with feed spacing larger than half wavelength, and as expected, we find that the arrays with regular spacing have grating lobes which produce spurious images in the reconstructed maps. We show that this problem can be overcome, using arrays with different feed spacing on each cylinder. We present the reconstructed maps, and study the performance in terms of noise power spectrum, transfer function and beams for both regular and irregular feed spacing configurations.

Key words: Cosmology: observation, HI intensity mapping; Method: transit telescope; map making

1 INTRODUCTION

The determination of the neutral hydrogen (HI) distribution from 21cm line observation is an important method to study the statistical properties of Large Scale Structures in the Universe. The intensity mapping technic is an efficient and economical way to map the universe using the (HI) 21cm emission, which is suitable for late time cosmological studies (($z \leq 3$)), specially for constraining dark energy models through baryon acoustic oscillation (BAO) features (Peterson et al. 2006; Chang et al. 2008; Ansari et al. 2009; Ansari et al. 2012; Seo et al. 2010; Gong et al. 2011). So that large wide-field and wide band radio telescopes would be needed to achieve rapidly the observation of large volumes of the universe. Several

UNIVERSITE PARIS-SACLAY

Titre : Reconstruction de cartes à partir des observations d'interféromètres radio en mode transit pour les expériences de cartographie d'intensité à 21 cm - Application à Tianlai et PAON-4

Mots-clés : cosmologie, énergie noire, émission 21 cm (HI), interférométrie radio, cartographie d'intensité, reconstruction de cartes en harmoniques sphériques

Résumé : L'analyse des propriétés statistiques de la distribution de la matière dans le cosmos (Grandes Structures, LSS or Large Scale Structure) est l'une des principales sondes cosmologiques qui permettent l'étude du modèle standard cosmologique, en particulier les paramètres caractérisant la matière noire et l'énergie noire. Les Oscillations Acoustiques Baryoniques (BAO's) sont l'une des mesures qui peuvent être extraites de l'étude de la distribution de matière à grande échelle (LSS).

L'observation de la distribution cosmique de la matière à partir de l'émission à 21 cm de l'hydrogène atomique neutre (HI) est une nouvelle méthode, complémentaire des relevés optiques pour cartographier la distribution de la matière dans le cosmos.

La méthode de cartographie d'intensité (Intensity Mapping) a été proposée depuis moins d'une dizaine d'années comme une méthode efficace pour cartographier en trois dimensions l'émission radio à 21 cm. Elle n'implique en particulier pas la détection des objets individuels (galaxies), et peut donc être effectué avec des instruments plus modestes en taille que ceux comme SKA ou FAST qui sont conçus pour détecter les galaxies à 21 cm à des distances cosmologiques.

Des interféromètres radio utilisant un ensemble de réflecteurs cylindriques ou paraboliques fixes, observant le

ciel en mode transit sont adaptés à la cartographie d'intensité. Le mode d'observation spécifique de ce type de radio télescope en cartographie d'intensité est étudié dans le cadre de ce travail de thèse. On montre en particulier qu'une méthode spécifique de reconstruction des cartes du ciel à partir des visibilités peut-être appliquée aux observations de ces interféromètres fonctionnant en mode transit. Cette méthode correspond à la décomposition en modes m des harmoniques sphériques et est très performante pour la reconstruction de grandes zones du ciel observées en mode transit.

Un code de reconstruction fondé sur ce principe a été développé, ainsi que différents critères de comparaison des performances instrumentales, comme le lobe d'antenne synthétisé, le spectre de bruit sur les cartes reconstruites et la réponse globale de l'instrument dans le plan (l,m) des harmoniques sphériques.

La méthode a été appliquée à différentes configurations des interféromètres composés de réflecteurs paraboliques ou cylindriques dans le cadre des projets PAON-4 et Tianlai. Outre l'optimisation des configurations des interféromètres Tianlai et PAON-4, le travail présenté inclut une première application de la méthode aux données PAON-4.

Title: Map making from transit interferometers observations for 21cm Intensity Mapping experiments Application to Tianlai and PAON-4

Keywords: cosmology, dark energy, 21cm emission (HI), radio interferometry, intensity mapping, spherical harmonic maps reconstruction

Abstract: The analysis of the statistical properties of the distribution of matter in the cosmos (LSS or Large Scale Structure) is one of the main cosmological probes that allow the study of the cosmological standard model, in particular the parameters characterizing dark matter and dark energy. Baryonic Acoustic Oscillations (BAO's) are one of the measurements that can be extracted from the study of matter distribution in large-scale structure (LSS).

The observation of the cosmic distribution of the matter from neutral atomic hydrogen (HI) 21 cm emission is a new method, complementary to the optical observation to map the distribution of matter in the cosmos.

In the last decade, the Intensity Mapping method has been proposed as an effective method for mapping the 21cm radio emission in three dimensions. In particular, it does not require the detec- tion of individual objects (galaxies), and can therefore be performed with instruments smaller in size than those such as SKA or FAST, which are designed to detect 21 cm galaxies at cosmological distances. A radio interferometer using a set of fixed cylindrical or parabolic reflectors observing the sky in transit modes are suitable instruments for intensity mapping surveys. The specific observational mode from this type of radio telescope by intensity mapping is studied in the context of this thesis. We show in particular that a specific sky maps reconstruction method from the visibilities can be applied to the observations of these interferometers operating in transit mode. This method corresponds to the m-modes decomposition of the spherical harmonics and is very efficient for the reconstruction of large sky areas observed in transit mode.

A reconstruction code based on this principle has been developed, as well as different crite- ria for the comparison of instrumental performances, such as the synthesized antenna lobe, the noise spectrum of the reconstructed maps and the overall instrument response in the the spherical harmonics (l,m) plane.

The method has then been applied to different configurations of interferometers composed of parabolic or cylindrical reflectors in the PAON-4 and Tianlai projects. In addition to optimizing the Tianlai and PAON-4 interferometer configurations, the work presented here includes a first application of the method to the PAON-4 data.

Université Paris-Saclay

Espace Technologique / Immeuble Discovery Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France