

Observational, digital readout, and calibration
techniques for studying the redshifted 21-cm signal of
hydrogen

Ian T. Hendricksen



McGill

Department of Physics, Trottier Space Institute
McGill University
Montréal, Québec, Canada

September 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Science in Physics

©2023 Ian T. Hendricksen

Abstract

Radio light emitted and absorbed from the hyperfine splitting of neutral hydrogen, known as the 21-cm signal, offers potentially revolutionary insights into the universe's history and composition. While the redshifted 21-cm signal is a single observational tool, the cosmology and astrophysics studied and the best choice for associated instrumentation varies immensely with redshift. The Mapper of the IGM Spin Temperature (MIST) experiment intends to measure the globally averaged 21-cm signal at low frequencies (approximately 25 to 125 MHz), where the conditions of the universe during the periods immediately before, during, and after the first stars and galaxies were born can be studied. At higher frequencies, the spatial fluctuations of the 21-cm signal can be measured through intensity mapping methods, using low spatial resolution observations of galaxies to study baryon acoustic oscillations (BAOs) and dark energy. The Canadian Hydrogen Observatory and Radio transient Detector (CHORD) is a next-generation radio interferometer currently under development and planned for construction at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, British Columbia. Sharing

science goals with the Canadian Hydrogen Intensity Mapping Experiment (CHIME), CHORD will build on its successes in intensity mapping, detection of fast radio bursts, and more, with significant improvements in collecting area, redundancy, and frequency coverage. The Deep Dish Development Array (D3A) located at the DRAO is a pathfinder for CHORD technologies, and currently consists of two three-meter dishes and three CHORD-like six-meter dishes. In this work, I present the measurement of the system temperature of D3A using observations of the radio galaxy Cygnus A and application of the interferometry concept of phase closure. I then discuss the requirements and possible designs of the digitization and channelization system of CHORD, known as the F-Engine. I present measurements characterizing the analog-to-digital converter (ADC) of the Xilinx Zynq Ultrascale+ ZU47DR radio frequency system-on-chip (RFSoc) field programmable gate array (FPGA) as a potential platform for the CHORD F-Engine. Finally, I present the Antenna Modeler and Performance Emulator (AMPERE) and the System for Optimizing Impedance Likelihood (SOIL), two complementary Python-based packages developed for MIST, which use antenna impedance measurements to characterize soil electrical parameters in order to develop an accurate beam model for the MIST antenna.

Abrégé

La lumière radio émise et absorbé par la transition hyperfine de l'hydrogène atomique, connue comme la raie à 21-cm, pourrait considérablement approfondir notre compréhension de l'histoire et de la composition de l'univers. Alors que la raie à 21-cm décalée vers le rouge est un outil d'observation unique, la cosmologie et l'astrophysique étudiées et le meilleur choix pour l'instrumentation associée varient énormément avec le décalage vers le rouge. Le expérience la Mapper of the IGM Spin Temperature (MIST) cherche à mesurer la moyenne globale de la raie à 21-cm à basse fréquence (approximativement 25-105 MHz dans le cas présent), où les conditions de l'univers pendant les périodes juste avant, pendant et après la naissance des premières étoiles et galaxies peuvent être étudiées. À de plus hautes fréquences, les fluctuations spatiales de la raie à 21-cm peuvent être mesurées grâce aux méthodes de cartographie de l'intensité, en utilisant des observations de galaxies à faible résolution spatiale pour étudier la structure à grande échelle de l'univers. L'Observatoire canadien de l'hydrogène et détecteur de transitoires radio (CHORD) est un interféromètre radio nouvelle génération en cours de développement, dont la construction

est prévue à l'Observatoire fédéral de radioastrophysique (OFR), près de Penticton en Colombie-Britannique. Partageant les objectifs scientifiques de l'Expérience canadienne de cartographie de l'intensité de l'hydrogène (CHIME), CHORD s'appuiera sur les succès de son prédécesseur en matière de cartographie de l'intensité de l'hydrogène, de détection de sursauts radio rapides, et plus encore, avec des améliorations significatives de la zone de collecte, de la redondance et de la couverture des fréquences. Le Deep Dish Development Array (D3A), situé à l'OFR, est un banc d'essai pour les technologies de CHORD, et comprend actuellement deux antennes paraboliques de trois mètres, et de trois antennes paraboliques de six mètres. Dans ce travail je présente la mesure de la température du système D3A en utilisant les observations de la radiogalaxie Cygnus A ainsi que l'application du concept de clôture de phase en interférométrie. J'aborde ensuite les exigences et conceptions possibles du système de numérisation et de canalisation de CHORD, connue comme le F-Engine. Je décris les mesures caractérisant le convertisseur analogique-numérique du système de radiofréquence sur puce (RFSoc) réseau de portes programmable sur site (FPGA) Xilinx Zynq Ultrascale+ ZU47DR comme une plateforme potentielle pour le F-Engine. Enfin, je présente le modélisateur d'antenne et émulateur de performance (AMPERE) ainsi que le système d'optimisation de la probabilité d'impédance (SOIL), deux package Python complémentaires développés pour MIST, qui utilisent les mesures de l'impédance de l'antenne pour caractériser les paramètres électriques du sol dans le but de précisément déterminer le faisceau de l'antenne.

Acknowledgements

I would first like to thank my supervisors Prof. Matthew Dobbs and Prof. Hsin Cynthia Chiang profusely for their guidance and support. This work (and indeed, all aspects of my research) is built upon their dedication to scientific excellency and academic integrity, and I am grateful that they continue to share this with me. I thank Cynthia for her constant enthusiasm and willingness to take a chance with me as my first supervisor during my undergraduate degree, and for continuing to guide me along this path today. I thank Matt for taking me on as a graduate student after working for only a short period in his lab as an undergraduate student, and for his passion for high-quality research that has motivated me to become a better scientist. I greatly appreciate everything they have done for me.

I would also like to thank Dr. Dallas Wulf for supporting my research with D3A over the past two years, and for his patient supervision of the work presented in Chapter 2. I also thank Vincent MacKay for insightful discussions related to the work presented in Chapter 2, and for helping me to refine the Abrégé. I am grateful to the CHORD and D3A collaborations for their support of my research.

I would also like to thank Dr. Jean-François (JF) Cliche not only for his guidance of the work presented in Chapter 3, but also for sharing with me his enthusiasm and passion for electronics and engineering. JF has been an incredible mentor for the past several years in electronics instrumentation, and I am deeply grateful for his support. I also thank Dr. Joshua Montgomery for help with the work presented in Chapter 3 and for my continued research with radio astronomy instrumentation. I thank t_0 for providing me access to their control and readout system and its support for my research with RFSocS.

Me gustaría agradecer al Dr. Raúl Monsalve para su supervisión de mi trabajo con el experimento MIST y toda la diversión que hemos compartido juntos (*“Honey, I’m home!”*). Raúl es uno de mis primeros mentores y un investigador intensamente apasionado y talentoso que me ha guiado con paciencia y sabiduría desde los primeros días de mi investigación. I also thank Prof. Ricardo Bustos and each member of the MIST collaboration for all they have done for me since I first joined the MIST team.

I thank the staff of Uapishka Station for their excellent accommodations and assistance during the first field deployment of MIST, and the Innu people for providing with us the opportunity to conduct research on their land.

I thank Prof. Laura Thomson for her incredible support as camp manager of the McGill Arctic Research Station (MARS), her vast enthusiasm for scientific research, and the many laughs we’ve shared. I thank the staff of the Polar Continental Shelf Program (PCSP) for their amazing work supporting scientific research in the Arctic. I also thank the people of

Resolute Bay, and indeed all the Inuit people for sharing with me something truly special by allowing me to explore and learn more about the people and their land. Qujannamiik.

I am deeply grateful to my mother and father, without whom I could never have imagined being where I am today. They are the foundation of the person I am, and I am immensely appreciative of everything they are, and all the things have done and continue to do for me. I love you both very much.

Finally, I thank Altaïr Truphème, who has been incredibly supportive and deeply compassionate throughout the writing of this thesis, and since I first met her. I am beyond words to express my gratitude for being with me when I needed her most, for everything that makes her unique, and for making me want to be a better person. Je t'aime beaucoup Altaïr.

Contents

List of Figures	xiv
List of Tables	xv
1 Introduction: Cosmology and Astrophysics at Radio Frequencies	3
1.1 A Brief History of the Universe	3
1.2 Cosmology with the 21-cm Signal	10
1.2.1 The Globally Averaged Signal	11
1.2.2 Spatial Fluctuations and Intensity Mapping	16
1.3 Fast Radio Bursts (FRBs)	19
1.4 Experiments Observing at Radio Frequencies	21
1.4.1 The Mapper of the IGM Spin Temperature (MIST)	23
1.4.2 The Canadian Hydrogen Observatory and Radio transient Detector (CHORD)	25
1.5 Thesis Outline	28

2	The Deep Dish Development Array (D3A)	30
2.1	D3A Technologies	31
2.2	D3A System Temperature	39
2.2.1	Principles of Interferometry	40
2.2.2	A Recipe to Measure T_{sys} from Phase Closure	44
2.3	Summary	63
3	Development of the CHORD F-Engine	65
3.1	CHORD F-Engine Designs	66
3.2	Characterization of the ZU47DR RFSoc ADC	69
3.3	The RFSoc as a Potential CHORD Platform	102
3.4	Summary	104
4	An Algorithm for Measuring Soil Electrical Properties from Antenna Impedance	107
4.1	Chromaticity Challenges for Global 21-cm Experiments	108
4.2	The Antenna Modeler and Performance Emulator (AMPERE)	111
4.3	The System for Optimizing Impedance Likelihood (SOIL)	125
4.3.1	Mathematical Background	125
4.3.2	Setup for Fitting Impedance Measurements at the Death Valley Site	129

4.3.3	Best-Fit Impedances and Soil Electrical Parameters for the Death Valley Site	131
4.4	Summary	139
5	Summary and Future Work	141
A	Diagrams of ADC Parameters	172

List of Figures

1.1	Overview of the 21-cm signal.	12
1.2	The MIST instrument at the McGill Arctic Research Station (MARS) with an inset map displaying previous MIST observing sites.	22
1.3	Overview of the CHORD instrument.	26
2.1	A photo of D3A at the DRAO with interferometer geometry shown in inset.	32
2.2	A simplified diagram of the D3A signal pipeline corresponding to the CHORD- like six-meter dishes.	33
2.3	Top: the D3A backend. Bottom: an ICE board.	37
2.4	A diagram of a two-element interferometer.	41
2.5	D3A squared gains $ g ^2$ as a function of frequency for each dish and polarization.	50
2.6	The normalized beams of each dish and polarization of D3A.	52
2.7	θ_{FWHM} for each dish and polarization of D3A as a function of frequency.	55
2.8	The beam solid angle of D3A as a function of frequency.	57

2.9	Top: The forward gain of D3A as a function of frequency. Bottom: The aperture efficiency η of D3A as a function of frequency.	59
2.10	T_{sys} of D3A as a function of frequency compared with the CHORD target T_{sys} of 30 K.	61
3.1	The Revision 1 CRS board from t_0 , SN004, equipped with a ZU47DR RFSoc.	70
3.2	A simplified diagram of the signal path for measurements of the ADC performance of the RFSoc.	71
3.3	An example of a single frame of data digitized by the RFSoc.	78
3.4	An example of the RFSoc FFT power spectrum with $f_c = 240.05$ MHz at approximately -35 dBFS.	79
3.5	The SNR and NSD of the RFSoc as a function of carrier frequency.	81
3.6	The power in the second- and third-order harmonics (H2 and H3) of the RFSoc as a function of carrier frequency.	83
3.7	The THD of the RFSoc as a function of carrier frequency.	85
3.8	The SINAD and ENOB of the RFSoc as a function of carrier frequency.	87
3.9	The SFDR of the RFSoc as a function of carrier frequency, including and excluding H2/3.	89
3.10	The maximum GTIS and OIS of the RFSoc as a function of carrier frequency.	90
3.11	The IM3 products of the RFSoc as a function of carrier power per input tone.	93
3.12	The DNL of the RFSoc.	97

3.13	The crosstalk leaked from CHx to CH0 of the RFSoc as a function of leakage frequency.	99
3.14	$ S_{11} $ for the CRS board analog inputs in the first Nyquist zone.	101
4.1	The effect of changing soil electrical parameters on the MIST antenna beam.	110
4.2	Diagram illustrating the AMPERE algorithm in the context of impedance interpolation.	112
4.3	Top: an example of the MIST blade dipole in FEKO. Bottom: diagrams of two soil models explored by MIST.	115
4.4	Top: all the simulated impedances for the one-layer soil model. Bottom: differences between simulated impedances for the one-layer soil model and interpolated impedances from AMPERE evaluated at the same parameters as the simulations.	118
4.5	Top: all the simulated impedances for the two-layer soil model. Bottom: differences between simulated impedances for the two-layer soil model and interpolated impedances from AMPERE evaluated at the same parameters as the simulations.	119
4.6	Histograms of the RMSE between simulated impedances for the one-layer soil model and impedances evaluated with AMPERE at the same parameters as the simulations.	121

4.7	Histograms of the RMSE between simulated impedances for the two-layer soil model and impedances evaluated with AMPERE at the same parameters as the simulations.	122
4.8	The best-fit impedance for the one-layer model compared with the measured impedance at Death Valley.	133
4.9	Corner plot corresponding to the one-layer model fit.	135
4.10	The best-fit impedance for the two-layer model compared with the measured impedance at Death Valley.	136
4.11	Corner plot corresponding to the two-layer model fit.	137
A.1	A diagram demonstrating intermodulation distortion.	173
A.2	A diagram demonstrating differential nonlinearity (DNL) in an ADC.	174

List of Tables

2.1	The approximate effective aperture A_e estimated from the FWHM of each dish and polarization.	54
3.1	The advantages and disadvantages of potential F-Engine designs for CHORD.	68
3.2	The resolution and sampling rates of the ADC platforms compared in this work.	104
3.3	Measured ADC quantities compared with quoted datasheet quantities and the EV8AQ160 ADC used on the ICE boards.	105
4.1	Left: the span of the simulation set corresponding to the one-layer soil model. Right: the span of the simulation set corresponding to the two-layer soil model.	116

List of Abbreviations

Abbreviation	Full Name	Chapter
ADC	Analog-to-Digital Converter	1, 2, 3, 5
ADU	Analog-to-Digital Unit	2
AMPERE	Antenna Modeler and Performance Emulator	1, 4, 5
BAO	Baryon Acoustic Oscillation	1, 5
BBN	Big Bang Nucleosynthesis	1
BIC	Bayesian Information Criterion	4
BOSS	Baryon Oscillation Spectroscopic Survey	1
eBOSS	extended Baryon Oscillation Spectroscopic Survey	1
CHIME	Canadian Hydrogen Intensity Mapping Experiment	1, 5
CHORD	Canadian Hydrogen Observatory and Radio transient Detector	1, 2, 3, 5
CMB	Cosmic Microwave Background	1
CRS	Control and Readout System	3
D3A	Deep Dish Development Array	1, 2, 3, 5
DM	dispersion measure	1
DNL	Differential Nonlinearity	3
DRAO	Dominion Radio Astrophysical Observatory	1, 2
EDGES	Experiment to Detect the Global EoR Signature	1
ENOB	Effective Number of Bits	3
FFT	fast Fourier transform	2
FLRW	Friedmann-Lemaître-Robertson-Walker (metric)	1
FPGA	Field Programmable Gate Array	1, 3
FRB	Fast Radio Burst	1, 5
f_c	Carrier frequency	3
f_{GTIS}	Gain-Time Interleaving Spur frequencies	3
f_{OIS}	Offset Interleaving Spur frequencies	3
f_{res}	Frequency resolution	3

Abbreviation	Full Name	Chapter
f_s	Sampling rate	3
FWHM	Full Width at Half Maximum	2
GTIS	Gain-Time Interleaving Spur	3
IGM	Intergalactic Medium	1, 5
IM3	Third-order intermodulation distortion	3
k_B	Boltzmann constant	1, 2
LNA	Low-Noise Amplifier	2
LSS	Large-Scale Structure	1, 5
LSB	Least Significant Bit	3
LPF	Low-Pass Filter	3
MARS	McGill Arctic Research Station	1
MIST	Mapper of the IGM Spin Temperature	1, 4, 5
NSD	Noise Spectral Density	3
n_f	Number of frequencies	4
n_p	Number of parameters	4
n_s	Number of samples per parameter	4
N_{sims}	Number of simulations	4
OIS	Offset Interleaving Spur	3
P_{fs}	Full-scale power	3
P_h	Total power in the harmonics	3
P_n	Total noise power	3
P_{nf}	Noise floor power	3
R_c	Cosine receiver response	2
R_s	Sine receiver response	2
RFI	Radio Frequency Interference	1, 2, 3
RFSoC	RF-System-on-Chip	1, 2, 3, 5
RMSE	Root-Mean-Square-Error	4
SDSS	Sloan Digital Sky Survey	1
SFDR	Spurious-Free Dynamic Range	3
SINAD	Signal-to-Noise-And-Distortion Ratio	3
SNR	Signal-to-Noise Ratio	3
SOIL	System for Optimizing Impedance Likelihood	1, 4, 5
THD	Total Harmonic Distortion	3
T_{sys}	System Temperature	1, 2, 3, 5
\mathcal{V}	Visibility	2
VLBI	Very Long Baseline Interferometry	1
WMAP	Wilkinson Microwave Anisotropy Probe	1
$\log(\mathcal{Z})$	Marginal likelihood	4
2dFGRS	Two-degree Field Galaxy Redshift Survey	1
6dFGS	Six-degree Field Galaxy Survey	1

Chapter 1

Introduction: Cosmology and Astrophysics at Radio Frequencies

1.1 A Brief History of the Universe

One of the most common aspects of our shared humanity across all cultures is our curiosity towards reality, nature, and our primordial beginnings. Whatever its origins, the current leading theory suggests that the universe began approximately 13.8 billion years ago with an event known as the Big Bang. Our modern understanding of the birth of the universe begins with Edwin Hubble in 1929. Hubble, building upon galaxy velocity measurements conducted by Milton Humason, observed that most galaxies are receding away from one another at a constant rate ([Hubble, 1929](#)), a phenomenon referred to as the Hubble-Lemaître law, or

simply Hubble's law. In the non-relativistic limit, the current recession velocity v of galaxies from one another in the absence of any comoving motion can be expressed as (Dodelson and Schmidt, 2021)

$$v = H_0 d, \quad (1.1)$$

where H_0 is the current recession rate, called the Hubble constant, and d is the physical distance between galaxies. This remarkable discovery suggested that the universe is not static — instead, it is expanding! The universe's expansion causes light to stretch as it travels through space, an effect known as redshift z defined by

$$1 + z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}}, \quad (1.2)$$

where λ_{obs} and λ_{emit} refer to the wavelength of light when it is observed and when it was emitted, respectively.

An expanding universe can be quantified by the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, which for a polar comoving coordinate system (r, θ, ϕ) can be generally expressed as

$$ds^2 = -c^2 dt^2 + a(t)^2 \left[dr^2 + S_\kappa(r)^2 d\Omega^2 \right], \quad (1.3)$$

which is valid on scales $\gtrsim 100$ Mpc. Here, s is the separation between two points in the space,

c is the speed of light, t is the cosmic time measured by an observer, and $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$.

The quantity $a(t)$ is known as the scale factor, and describes how distances change as a function of time due to the expansion of the universe, with the convention that $a = 1$ today. Note that the Hubble parameter¹ is defined by the scale factor as $H(t) \equiv \dot{a}/a$. The parameter $S_\kappa(r)$ quantifies any possible effect of curvature of the universe, dependent upon the curvature parameter κ :

$$S_\kappa(r) = \begin{cases} R \sin(r/R) & \kappa = +1 \text{ (Positive Curvature)} \\ r & \kappa = 0 \text{ (Flat)} \\ R \sinh(r/R) & \kappa = -1 \text{ (Negative Curvature)}. \end{cases} \quad (1.4)$$

where R is the radius of a spherical volume within the universe², and the value of κ indicates whether the universe is positively curved, flat, or negatively curved. In general, the scale factor a is dependent upon the energy density of the universe $\varepsilon(t)$, the curvature parameter κ , and the current radius of the observable universe R_0 . One of the principal goals in modern cosmology is to study $a(t)$ in order to gain insight into the overall composition and geometry of the universe.

The energy density $\varepsilon(t)$ of the universe is thought to be the sum of the energy from three dominant contributions: the energy density associated with radiation $\varepsilon_r(t)$, matter $\varepsilon_m(t)$,

¹The Hubble parameter at the current time is referred to as the Hubble constant H_0 .

²The radius R does not describe the size of the entire universe, which is unknown; rather, it can be thought of as a region within a larger volume with nothing particularly special or unique about it.

and the mysterious and poorly understood dark energy $\varepsilon_\Lambda(t)$. The equation of state $P = w\varepsilon$ for a given form of energy relates the energy density with the pressure P exerted by that form of energy. The term w is known as the equation of state parameter, which has a different value for each form of energy: $w_r = 1/3$ for radiation, $w_m = 0$ for matter, and $w_\Lambda < -1/3$ for dark energy. The special case $w_\Lambda = -1$ corresponds with a “cosmological constant” Λ , where

$$P_\Lambda = -\varepsilon_\Lambda = -\frac{c^2}{8\pi G}\Lambda, \quad (1.5)$$

where G is the gravitational constant. Note that the above expression implies that dark energy, characterized by Λ , exerts constant negative pressure! Einstein had famously introduced a cosmological constant Λ into the equations of general relativity such that they described a static universe, but later removed the term. When it was discovered that the expansion of the universe is accelerating, the cosmological constant Λ was reintroduced as a possible form of dark energy that remains unchanging with time.

An equation relating the parameters a , ε , κ , R_0 , and Λ that also accounts for the effects of general relativity is known as the Friedmann equation, which can be written as³

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3c^2}\varepsilon - \frac{\kappa c^2}{R_0^2}\frac{1}{a^2} + \frac{\Lambda}{3}, \quad (1.6)$$

³Note that the Friedmann equation and acceleration equation written here separate Λ from ε and P such that its inclusion is explicit, though in general the Λ term can be absorbed into ε and P .

which describes how a changes as a function of time. An equation describing the acceleration of a can be written as

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2}(\varepsilon + 3P) + \frac{\Lambda}{3}. \quad (1.7)$$

If the $\Lambda/3$ term dominates the right-hand side of the acceleration equation, it implies that the expansion of the universe is accelerating due to the negative pressure exerted by dark energy. Indeed, in the current age of the universe's history, this corresponds to reality.

The following description of the history of the universe is based on [Dodelson and Schmidt \(2021\)](#). When observed at large scales, on the order of 150 Mpc or larger, the universe is remarkably homogeneous and isotropic, an insight known as the cosmological principle. The current leading theory suggests that the cosmological principle can be explained by the theory of inflation: immediately following the Big Bang was a period of rapid, immense expansion, with the scale factor increasing by a factor of e^{60} over approximately 10^{-33} s. While the engine driving inflation is unknown, inflationary theories of cosmology can explain both the cosmological principle as well as the origins of the small-scale fluctuations that would later become the observed large-scale structure (such as galaxies, galaxy clusters, and filaments).

Following inflation was the radiation-dominated epoch, during which the universe was a hot, dense plasma where high-energy photons dominated, preventing nuclei from forming. During the radiation-dominated epoch, the coupled photon-baryon plasma experienced characteristic perturbations caused by overdensities in the plasma. The plasma

overdensities encouraged the accumulation of matter in overdensity regions. As the temperature of the overdensity region increased from the interaction of the condensed photon-baryon plasma, the pressure in the overdensity region would push the plasma outwards. The oscillations caused by the overdensity regions resulted in acoustic waves that traveled through the plasma. When the radiation decoupled from the baryons, the acoustic waves could no longer be supported, leaving behind regions of matter overdensity set by the sound horizon. The imprint of the acoustic waves on the distribution of matter are known as baryon acoustic oscillations (BAOs). The remaining matter overdensities would later manifest in the large-scale structure (LSS) of the universe. Since the scale of the oscillations is known, BAOs can be used as a standard cosmological ruler, which makes BAOs a useful tool for probing the effect of dark energy on the expansion of the universe.

The first detections of BAOs were made in 2005 by the Sloan Digital Sky Survey (SDSS) ([Eisenstein et al., 2005](#)) and the Two-degree-Field Galaxy Redshift Survey (2dFGRS) ([Cole et al., 2005](#)). The results were found to be consistent with those from experiments observing the CMB, such as measurements from the Wilkinson Microwave Anisotropy Probe (WMAP). Subsequent detections have been made by the Six-degree Field Galaxy Survey (6dFGS) ([Beutler et al., 2011](#)), the WiggleZ Dark Energy Survey ([Blake et al., 2011](#)), the Baryon Oscillation Spectroscopic Survey (BOSS) ([Anderson et al., 2012](#)), and the extended Baryon Oscillation Spectroscopic Survey (eBOSS) ([Zhao et al., 2022](#)).

As the universe cooled and photon energy decreased, light nuclei were finally able to

form, known as Big Bang Nucleosynthesis (BBN). At $z \sim 1100$, the plasma cooled enough to allow the first atoms to form, an event known as recombination. During the period of recombination, the photons became decoupled from the baryon gas, and were able to freely stream throughout the universe. We observe the imprint of the acoustic waves from the pre-decoupling period in the photon distribution today in the cosmic microwave background (CMB).

Following recombination, the universe entered a period known as the dark ages, in which the universe continued expanding as it gradually cooled. During the dark ages, the only source of radiation was from CMB photons, and neutral hydrogen dominated the baryonic matter in the universe. Eventually, matter began to cluster enough that temperatures needed for fusion could be reached, and the first stars were born in the cosmic dawn era. As matter began to condense into the first galaxies, the light from the first stars began to ionize the surrounding neutral hydrogen, forming bubbles of ionized hydrogen in a period known as reionization. Eventually, most of the matter in the universe would be ionized. Matter would be the dominant form of energy until it was eventually superseded by dark energy, causing the expansion of the universe to accelerate, which continues today.

Cosmologists are constantly searching for novel techniques to explore the most basic questions about our universe and understand its evolution. One of the most versatile tools expected to enable the next generation of cosmological discovery is the 21-cm transition of nuclear hydrogen.

1.2 Cosmology with the 21-cm Signal

The 21-cm signal refers to radio waves emitted or absorbed due to the hyperfine splitting of the hydrogen atom. The hyperfine splitting originates from the spin-spin coupling of the proton and electron of the hydrogen atom. Spin-spin coupling causes the hydrogen atom to have slightly greater energy when the spins of its proton and electron are parallel than when they are anti-parallel. The energy difference between the parallel and anti-parallel states⁴ is (Griffiths, 2004)

$$\Delta E_{21} = \frac{4g_p\hbar^4}{3m_p m_e^2 c^2 a^4} = 5.88 \times 10^{-6} \text{ eV}, \quad (1.8)$$

where g_p is the proton g -factor, \hbar is the reduced Planck constant, m_p is the proton mass, m_e is the electron mass, and a is the Bohr radius (not to be confused with the scale factor $a(t)$ from Equation 1.3). When the spin alignment spontaneously flips from parallel to anti-parallel as shown in Figure 1.1(a), a photon is emitted with energy ΔE_{21} , with wavelength

$$\lambda_{21} = \frac{hc}{\Delta E_{21}} \approx 21 \text{ cm}, \quad (1.9)$$

from which the name of the 21-cm signal originates. λ_{21} corresponds to $\nu_{21} = 1420$ MHz when emitted, and becomes redshifted due to the expansion of the universe according to

⁴Typically, the parallel and anti-parallel spins states are respectively referred to as the triplet and singlet states in the literature. I use the former notation for simplicity in this work.

Equation 1.2. By tuning our instruments to different redshifted wavelengths of the 21-cm signal, we can, in principle, study nearly any period of the post-recombination universe.

Broadly, the 21-cm signal can be studied in two ways: 1) by measuring the signal averaged across the entire sky, referred to as the globally averaged signal, and 2) studying the spatial fluctuations of the 21-cm signal. The instrumentation and analysis corresponding to these different techniques of observing the 21-cm signal are vastly different given the different systematics inherent to each type of measurement and the different redshift ranges in which they are typically studied. However, studies of the global signal and spatial fluctuations yield complementary science. I will present a brief overview of the key aspects of each technique and current measurement efforts in the following two sections.

1.2.1 The Globally Averaged Signal

The excitation temperature of the 21-cm signal, commonly referred to as the spin temperature T_s , is defined by the ratio of number densities of hydrogen atoms in the parallel and anti-parallel spin states, written as (Pritchard and Loeb, 2012)

$$\frac{n_1}{n_0} = \frac{g_1}{g_0} e^{T_*/T_s}, \quad (1.10)$$

where $g_1/g_0 = 3$ is the ratio of the statistical degeneracy factors of the parallel and anti-parallel states, and $T_* = \Delta E_{21}/k_b = 0.068$ K, where k_b is Boltzmann's constant. The spin temperature leads to the definition of the differential brightness temperature δT_b of the

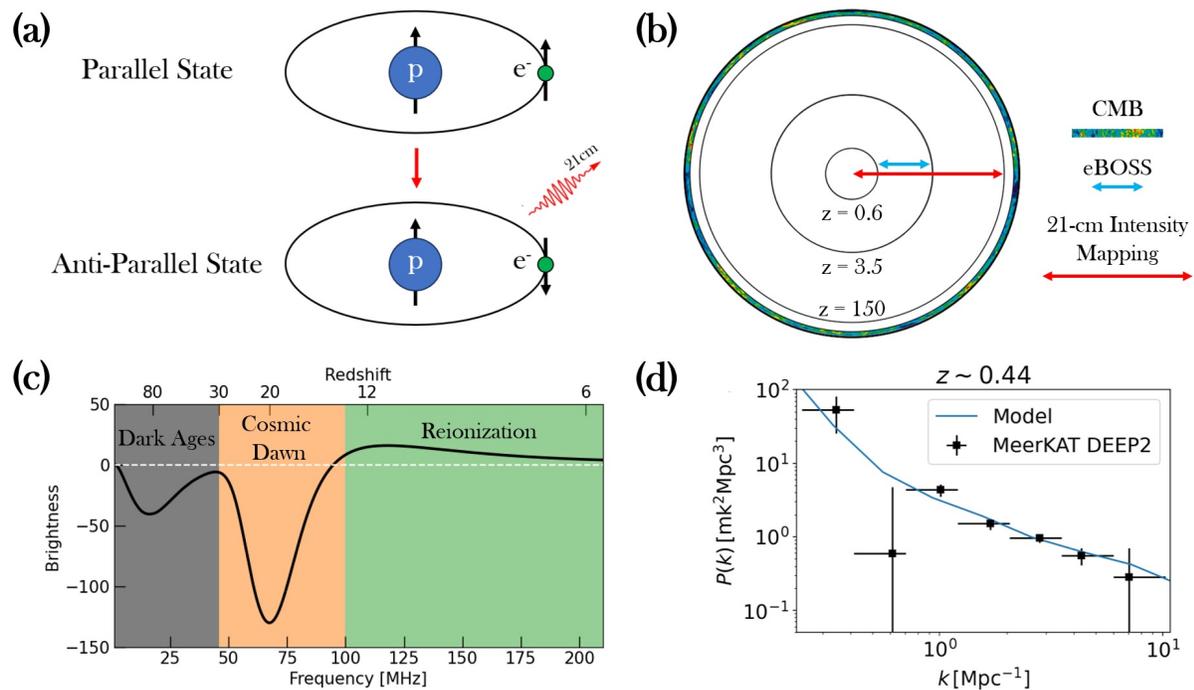


Figure 1.1: Overview of the 21-cm signal. **(a)** A diagram of 21-cm signal emission caused by a spin-flip transition. **(b)** A diagram of the approximate regions of the observable universe that have been probed by the Extended Baryon Oscillation Spectroscopic Survey (eBOSS) in comparison with the range that can be potentially surveyed through intensity mapping. Note that eBOSS does not survey the entire right ascension range. This diagram includes the redshift ranges corresponding to luminous red galaxies, emission line galaxies, and quasars measured by eBOSS^a. The script used to generate to-scale circles of specific redshifts was created by Prof. Adrian Liu. **(c)** a simulation of the global 21-cm signal using the default parameters of the ARES code (Mirocha, 2020). **(d)** The 21-cm power spectrum at $z \sim 0.44$ as measured by the MeerKAT telescope (Paul et al., 2023).

^aThe approximate redshift range derives from the technical details listed at <https://www.sdss4.org/surveys/eboss/>.

21-cm signal, which may be expressed as⁵

$$\delta T_{\text{b}} = \left(\frac{3\hbar c^3 A_{10}}{16k_{\text{b}}\nu_{21}^2} \right) \left[\frac{x_{\text{HI}}n_{\text{H}}}{(1+z)^2 dv_{\parallel}/dr_{\parallel}} \right] \left(\frac{T_{\text{s}} - T_{\gamma}}{T_{\text{s}}} \right) \quad (1.11)$$

where $A_{10} = 2.85 \times 10^{-15} \text{ s}^{-1}$ is the spontaneous 21-cm emission rate, x_{HI} is the fraction of neutral hydrogen, n_{H} is the number density of hydrogen atoms, $dv_{\parallel}/dr_{\parallel}$ is the gradient of the proper velocity v_{\parallel} along the line of sight distance r_{\parallel} , and T_{γ} is the background temperature (typically equivalent to the CMB temperature T_{CMB}). The global 21-cm signal refers to the monopole component of the differential brightness temperature δT_{b} , which can be studied across approximately $200 \gtrsim z \gtrsim 6$.

A model of the global 21-cm signal created using the default parameters of the Accelerated Reionization Era Simulations (ARES) code (Mirocha, 2020) is shown in Figure 1.1(c). The global 21-cm signal has three key features: the dark ages absorption feature, the cosmic dawn absorption feature, and the epoch of reionization emission feature. Insights into collisional coupling can be made through the dark ages absorption feature, where the spin temperature is coupled to the gas temperature. Studying the dark ages with the global 21-cm signal is crucial for precision cosmology measurements in that era, as there is considerable interest in measuring the dark ages power spectrum. The cosmic dawn absorption feature is caused by Ly α photons produced by the first stars, which promote 21-cm photon absorption via the

⁵Note that I explicitly refer to the *differential* brightness temperature δT_{b} since the global 21-cm brightness temperature is defined through the difference between the spin temperature T_{s} and the background temperature T_{γ} .

Wouthuysen-Field effect, and can thus be used to study the Ly α flux produced by the first stars. The emission feature, caused by heating of the intergalactic medium (IGM) by X-rays from the first stars, can be used to probe the history of reionization ([Liu and Shaw, 2020](#); [Pritchard and Loeb, 2012](#)).

In 2018, the Experiment to Detect the Global Epoch of Reionization Signature (EDGES) reported the detection of a flattened absorption feature in the differential brightness temperature centered at 78 MHz with an unexpectedly large amplitude of 500 mK ([Bowman et al., 2018](#)). The discovery of such an anomalously deep absorption feature has led to speculation over whether the measurement constitutes a true detection of the global 21-cm signal, which may require non-standard cosmology to explain ([Barkana, 2018](#)), or if the reported detection was biased by some systematic error ([Bradley et al., 2019](#)). In particular, the EDGES detection has led some to propose that the global 21-cm signal can be used to study dark matter, e.g., [Fraser et al. \(2018\)](#).

In 2022 however, the Shaped Antenna Measurement of the Background Radio Spectrum 3 (SARAS 3) team reported a non-detection of the EDGES signal in the 55–85 MHz band, and reject the EDGES best-fit model with 93.5% confidence ([Singh et al., 2022](#)). Therefore, it is crucial that independent measurements of the signal are made by other experiments. Other ground-based experiments hunting for the global 21-cm signal include ASSASSIN ([McKinley et al., 2020](#)), BIGHORNS ([Sokolowski et al., 2015](#)), CTP ([Nhan et al., 2019](#)), LEDA ([Price et al., 2018](#)), LWA-SV ([DiLullo et al., 2020](#)), PRI^ZM ([Philip et al., 2018](#)),

REACH (de Lera Acedo, 2019), and SITARA (Thekkepattu et al., 2022). Global signal experiments or concepts that will attempt to observe from space include DAPPER (Burns, 2021), DSL/Hongmeng (Shi et al., 2022), LuSEE Night (Bale et al., 2023), and PRATUSH⁶. In this thesis, I present work related to the Mapper of the IGM Spin Temperature (MIST), a ground-based global 21-cm experiment which is introduced in Section 1.4.1.

Since global 21-cm experiments are interested in the sky-averaged signal, instruments with high spatial resolution are unnecessary, and many experiments opt to use a single antenna with a wide beam. In addition, global 21-cm experiments do not require high frequency resolution, and thus do not need to integrate for very long to achieve sufficiently low noise. By using instruments with relatively coarse frequency resolution, radiometer noise can be integrated below that of the global 21-cm signal within a day.

However, the Galaxy emits synchrotron radiation at the same frequencies as the global 21-cm signal, whose brightness temperature is many orders of magnitude greater than the weak 21-cm signal. While techniques such as the use of polarization information have been proposed to tackle these foregrounds, they remain one of the most formidable challenges in global 21-cm cosmology. The minimization of instrumental systematics is another significant challenge for observations of the global 21-cm signal. Beam chromaticity can cause difficult-to-model structure in the measured spectrum. Further, since global signal experiments measure total power, spectral features caused by cable reflections or

⁶<https://www.rri.res.in/DISTORTION/pratush.html>

environmental effects can be misinterpreted as real detections of the features shown in Figure 1.1(c). Experiments aiming to detect the global 21-cm signal therefore require an acute understanding of instrumental systematics.

Further complicating measurements of the global 21-cm signal are the effects of radio frequency interference (RFI) and the ionosphere, which can cause time-varying corruption of data (Liu and Shaw, 2020). As such, global 21-cm experiments need to observe from remote locations, and ideally observe at night when the ionosphere is less active.

1.2.2 Spatial Fluctuations and Intensity Mapping

The spatial fluctuations of the 21-cm signal can in principle allow us to trace the distribution of matter over vast timescales, from the dark ages to the present. Using the spatial fluctuations of the 21-cm signal as a proxy for matter distribution suggests the possibility for cosmological probes, such as studies of dark energy using BAOs (Liu and Shaw, 2020).

Studying the distribution of matter through the 21-cm signal can be achieved through the technique of intensity mapping. As opposed to methods for mapping individual galaxies through optical instruments, which require high spatial resolution, intensity mapping abandons the need for high spatial resolution by instead focusing on lower spatial modes. For example, the BAO scale is of order degree angular size on the sky, which can be probed using relatively small dishes or baselines (Ansari et al., 2012). Each voxel

measured with intensity mapping can thus have hundreds to thousands of galaxies inside. As discussed in the first section of this chapter, BAOs can be used to track the expansion history of the universe, and intensity mapping can thus help to improve our understanding of the nature of dark energy.

Unlike global 21-cm experiments that typically measure with a single wide-beam antenna, intensity mapping experiments require finer spatial resolution, which can be achieved through the method of interferometry. Radio interferometers can benefit from having much wider fields of view than optical telescopes while also observing across their entire bandwidths simultaneously, enabling them to more efficiently produce maps of the LSS. Further, since intensity mapping does not resolve individual galaxies, maps of the universe can be made at higher redshifts than optical surveys are capable. Figure 1.1(b) shows a graphic of the regions of the observable universe mapped by the Extended Baryon Oscillation Spectroscopic Survey (eBOSS) compared with the regions that can be potentially probed with intensity mapping. The above makes intensity mapping a potentially competitive tool in comparison with optical surveys ([Ansari et al., 2012](#); [Liu and Shaw, 2020](#)).

The technique of intensity mapping is still in relative infancy ([Chang et al., 2010](#)), and there are formidable challenges to overcome before it can be used for studies in cosmology. Similar to global 21-cm experiments, intensity mapping experiments need to contend with galactic foregrounds that vastly outweigh the weak 21-cm fluctuations by a factor of $\sim 10^5$. In addition, the signal processing requirements for intensity mapping experiments are

immense in comparison to global 21-cm experiments. The number of antennas can be on the order of hundreds or more, which must be instrumented with individual signal paths while maintaining a high degree of systematic control. Intensity mapping experiments also need to process data at colossal rates, on the order of many terabits per second, requiring custom-designed digital signal processing hardware like the ICE system (Bandura et al., 2016).

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is a radio interferometer designed for intensity mapping from 400–800 MHz ($2.5 > z > 0.8$), and is located at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, BC, Canada. With the unusual design of a cylindrical reflector as opposed to parabolic dishes, CHIME has an extremely wide field of view (31000 deg^2), allowing it to observe 3/4 of the sky each day (Amiri et al., 2022). In 2022, CHIME reported the detection of the 21-cm power spectrum from $1.43 > z > 0.78$ by cross-correlating with the eBOSS survey, which constituted the first 21-cm intensity mapping measurements made with an interferometer (Amiri et al., 2023).

Paul et al. (2023) recently announced a detection of the cosmological power spectrum using intensity mapping techniques with the MeerKAT telescope. Notably, the detection presented by Paul et al. (2023) did not rely upon galaxy surveys, representing the first truly independent measurement of the power spectrum. The power spectrum measured by Paul et al. (2023) at $z \sim 0.44$ are shown in Figure 1.1(d). The MeerKAT detection is a pioneering

achievement for the field of intensity mapping, which is just beginning to demonstrate its capabilities as a novel tool for studying the LSS.

Other experiments having previously studied, currently studying, or planning to study the spatial fluctuations in the 21-cm signal include BINGO ([Abdalla et al., 2022](#)), GBT ([Wolz et al., 2021](#)), GMRT ([Pen et al., 2009](#)), HERA ([DeBoer et al., 2017](#)), HIRAX ([Newburgh et al., 2016](#)), LOFAR ([van Haarlem et al., 2013](#)), MWA ([Beardsley et al., 2019](#)), OVRO-LWA ([Eastwood et al., 2019](#)), PAPER ([Jacobs et al., 2015](#)), the Parkes Radio Telescope ([Anderson et al., 2018](#)), SKA ([and David J. Bacon et al., 2020](#)), and the Tianlai experiment ([Xu et al., 2015](#)).

1.3 Fast Radio Bursts (FRBs)

A key advantage of studying the universe at radio frequencies is that the same frequencies allow us to access other leading topics in astrophysics and cosmology. In particular, the phenomenon of fast radio bursts (FRBs) overlap with the frequencies targeted by experiments in 21-cm cosmology. FRBs are generally defined as colossal explosions of energy (on the order of 0.05–100 Jy) emitted over ms timescales, and are observed as radio waves ([Petroff et al., 2019](#)). FRB photons pass through media of varying electron density on their way to our detectors on Earth, which results in the dispersion of the light in the same manner as light passing through a prism. The amount of dispersion is quantified by the dispersion measure (DM), or the integrated electron density n_e along the line of sight l , given by ([Zhang, 2022](#))

$$\text{DM} = \int_0^{d_z} \frac{n_e}{1+z(l)} dl, \quad (1.12)$$

where d_z is the comoving distance to the FRB. The DM can be written as a sum of DMs from specific environments, including the ionosphere, interplanetary medium, interstellar medium, intergalactic medium, host galaxy, and local FRB environment. Researchers have found that most observed FRBs are of extragalactic origin. As such, FRBs have been proposed as a novel tool for studies in modern cosmology. In particular, since FRB photons interact with the IGM, where most of the baryonic matter in the universe is expected to reside, FRBs can be used to probe the baryon density Ω_b , or the fraction of baryons in the IGM f_{IGM} .

However, there exists a degeneracy between the redshift z of an FRB and cosmological parameters such as Ω_b and f_{IGM} , such that external information about one is required to measure another. In addition, since FRBs are typically poorly localized, one must make assumptions of the DM contribution from individual media, which could bias results. Therefore, experiments like CHIME are currently or planning on constructing “outriggers”, which are smaller copies of a core array positioned at continental baselines. Interferometry at such distances, known as Very Long Baseline Interferometry (VLBI), will enable these experiments to achieve mas localization of FRBs, which will help to identify FRBs within their local environments (Mena-Parra et al., 2022). Indeed, CHIME has already effectively revolutionized the field of FRB science, quadrupling the number of known FRBs with its first published catalog (CHIME/FRB Collaboration et al., 2021).

Not only are FRBs considered of cosmological interest due to their implied distance, but they are also of astrophysical interest due to their enormous energy budgets. Studies of the progenitor mechanism of FRBs have suggested that there may be multiple ways to produce an FRB. While [Andersen et al. \(2020\)](#) and [Bochenek et al. \(2020\)](#) reported the detection of a galactic FRB that was localized to a magnetar, [Kirsten et al. \(2022\)](#) reported the detection of an FRB localized within a globular cluster, which challenges progenitor theories involving young magnetars. Therefore, the FRB progenitor search also benefits from improvements in localization of FRBs, which can be enabled through VLBI.

1.4 Experiments Observing at Radio Frequencies

By observing the universe at different radio frequencies, we can tune our instruments to study a broad range of topics in astrophysics and cosmology. As discussed previously, the instrumentation and analysis methods vary widely upon the specific science goals targeted and the frequencies observed. In this section, I introduce two experiments in radio astrophysics and cosmology with distinct observational goals and instrumental designs: the Mapper of the IGM Spin Temperature (MIST) and the Canadian Hydrogen Observatory and Radio transient Detector (CHORD).

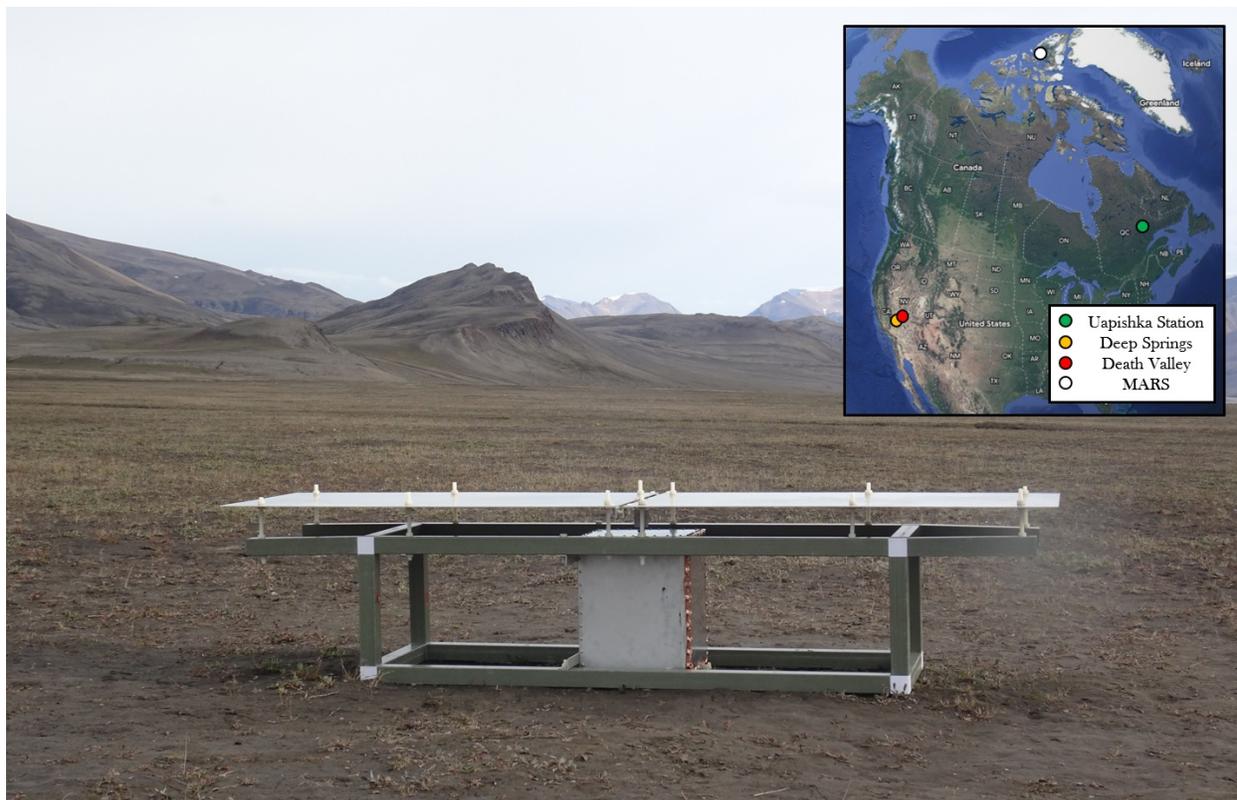


Figure 1.2: The MIST instrument at the McGill Arctic Research Station (MARS) with an inset map displaying MIST observing sites. The MIST instrument consists of a blade-dipole antenna mounted on a fiberglass frame. An RF-sealed receiver box sits immediately beneath the antenna, which houses the analog and digital signal processing electronics. The MIST instrument operates directly above the soil with no ground plane.

1.4.1 The Mapper of the IGM Spin Temperature (MIST)

The MIST experiment⁷ is a total-power radiometer seeking to study the early universe by measuring the global redshifted 21-cm signal from remote radio-quiet locations (Monsalve et al., 2023). The MIST instrument corresponds to a blade dipole antenna operating over approximately 25–105 MHz, or $55 > z > 13$, which includes both the end of the dark ages and cosmic dawn. MIST measures from remote locations because the 25–105 MHz band is dominated by RFI in regions of human development, including the FM band, television stations, airplane communication, and other contaminants. RFI at these frequencies can be orders of magnitude larger than even galactic radio emission, making detection of the global 21-cm signal unfeasible near areas of high population density. As such, MIST is designed to be compact, consume low power, and able to operate across a wide range of temperatures to enable observations from remote locations.

Each panel of the MIST blade dipole is 1.2 m by 0.6 m in size and are mounted on a fiberglass frame, which is effectively transparent at radio frequencies. The panel separation and flatness are adjustable to fine-tune the geometry and minimize associated systematic errors. The two panels are connected to a balun that sits above a receiver box, which contains all the electronic components of the signal processing pipeline aside from the balun.

The signals are filtered and amplified in the “analog stage”, which also contains a switching network for calibration purposes. The switching network allows for the

⁷<https://www.physics.mcgill.ca/mist/>

measurement of reflection coefficients (impedance) of the antenna and RF standards using an internal vector network analyzer (VNA), as well as the spectra of the antenna and a noise source. When the instrument is measuring spectra, the signals are digitized and channelized by an ADC-FPGA in the “digital stage”, which also operates a ROCK Pi X computer with a Linux operating system that collects and stores data. The analog and digital stages are separated into their own sealed Faraday cages within the receiver to prevent RFI contamination reaching the analog electronics. MIST operates with four 12 V lithium-ion batteries inside the receiver box which can provide power for up to two days of observations. The receiver box is sealed with dozens of screws in order to prevent RFI leakage from the internal components.

While some global 21-cm experiments implement ground planes beneath their antennas, it has been demonstrated that ground planes can introduce chromatic effects in the beam that are difficult to model ([Spinelli et al., 2022](#)). A key aspect of the MIST instrument is that it observes directly above soil with no ground plane, meaning that the MIST beam is sensitive to the electrical properties of the soil. Therefore, MIST operates in flat, open areas to minimize systematic effects associated with ground inhomogeneities. However, operating with no ground plane introduces a new challenge: MIST needs precise models of the soil electrical properties as a function of frequency to produce accurate beam models, which must be calibrated in order to recover the 21-cm signal.

MIST has so far observed from four locations in North America, in order of deployment

date: 1) Uapishka Station in Quebec, Canada in August 2021, 2) Deep Springs Valley in California, USA in May 2022, 3) Death Valley in Nevada, USA in May 2022, and most recently, 4) the McGill Arctic Research Station (MARS) on Axel Heiberg Island, Nunavut, Canada, with deployments to MARS in both the summer of 2022 and spring of 2023. MIST also intends to observe from the Atacama Desert in northern Chile. MIST is currently in the process of analyzing data collected at Deep Springs Valley, Death Valley, and MARS, with particular dedication to determining the electrical parameters of the soil at each site using various techniques in order to produce accurate beam models. In Chapter 4, I will describe an algorithm I have developed that fits the measured antenna impedance to models interpolated from simulations to extract the best-fit soil electrical parameters.

1.4.2 The Canadian Hydrogen Observatory and Radio transient Detector (CHORD)

CHORD⁸ is a next-generation radio interferometer currently in development and slated to begin observing midway through the decade (Vanderlinde et al., 2019). CHORD will observe across a wide frequency range covering 300–1500 MHz, which corresponds to $z < 3.7$ for the 21-cm line. Working alongside CHIME, CHORD will implement the best developments from CHIME while implementing novel technologies to enable it to be a world-leading observatory.

CHORD will be able to localize FRBs to within $\lesssim 10$ mas, which will enable CHORD

⁸<https://www.chord-observatory.ca/>

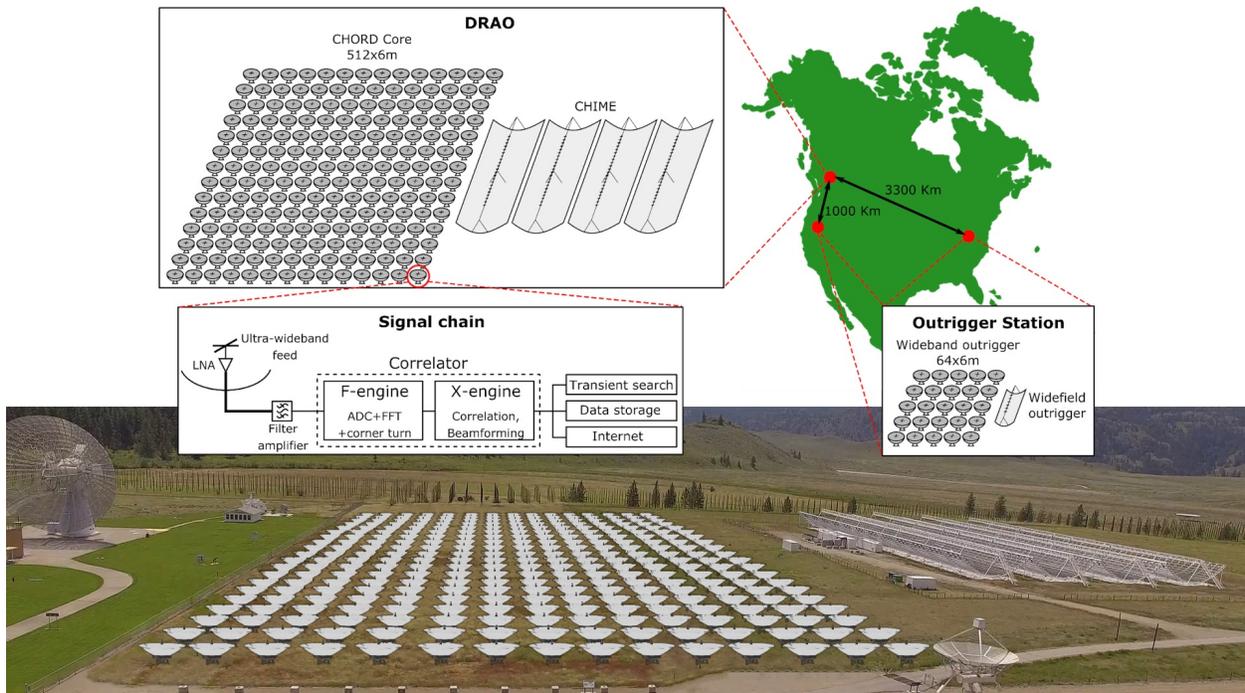


Figure 1.3: Overview of the CHORD instrument. The upper graphic displays the locations of the CHORD core array and outriggers in North America. The CHORD core consists of 512 six-meter dishes at the DRAO while each outrigger site consists of 64 six-meter dishes and a single CHIME-like reflector. The signal chain consists of wideband analog components (e.g., the feed and LNA), an F-X correlator, and several backends corresponding to different processing requirements for each science goal. Upper graphic designed by Juan Mena-Parra. Lower image is an artist's rendition of the CHORD core array at the DRAO, designed by Tracy Zhuo.

to study FRB source environments and use FRBs as cosmological probes. CHORD will also be an unparalleled instrument for mapping the distribution of matter throughout the universe. CHORD will produce the largest 3D map of the universe through measurements of the spatial fluctuations of the 21-cm signal. Using the method of intensity mapping, CHORD will study the effect of dark energy on the LSS and probe the BAO scale, offering insights into the conditions of the very early universe.

CHORD will consist of a core of 512 six-meter dishes situated immediately next to the site of the CHIME core at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, British Columbia, shown in an artist's rendition in the bottom of Figure 1.3. CHORD will also construct two outrigger stations to enable precision localization of FRBs, which will consist of 64 six-meter dishes and will be equipped with a single CHIME-like cylinder, which are currently under development to function as outriggers for CHIME.

Key aspects of the design of CHORD are a focus on redundancy and systematic control. A redundant array is one whose dishes are arranged along a grid, such that there are many repeated (redundant) baselines. Having redundant baselines can improve sensitivity to the spatial modes probed, as the visibilities of each element should be the same within noise. Implementing redundant calibration is advantageous because it makes few assumptions regarding the emission from the sky (Liu and Shaw, 2020). However, redundant arrays require a high degree of systematic control, on the order of one part in one thousand, to be considered truly redundant.

To achieve the science goals of CHORD and meet the redundancy and systematic control requirements for it to be a world-class instrument, CHORD has developed the Deep Dish Development Array (D3A), a testbed for CHORD technologies. D3A consists of two three-meter dishes and three CHORD-like six meter dishes located at the DRAO. D3A has or is expected to lead to key developments in the prototyping of the CHORD dishes, feed, low-noise amplifier (LNA), filtering and amplification chain, and digital readout system. I will provide a more detailed overview of D3A in Chapter 2.

1.5 Thesis Outline

This thesis addresses a wide variety of topics in 21-cm cosmology across a broad frequency range, spanning nearly the entire redshift range of $55 \gtrsim z$. I describe unique techniques in 21-cm cosmology instrumentation and analysis corresponding with the distinct science goals of CHORD and MIST. In Chapter 2, I describe the Deep Dish Development Array (D3A), a prototype interferometer for development of CHORD technologies located at the DRAO. I then present a measurement that I conducted of the system temperature T_{sys} of D3A in comparison of the CHORD target $T_{\text{sys}} = 30$ K. In Chapter 3, I compare two potential platforms for the F-Engine of CHORD, which will be responsible for the digitization, channelization, and corner-turn operation of antenna signals. I then present measurements that I made characterizing the analog-to-digital converter (ADC) of a third-generation Xilinx Ultrascale Zynq+ RF-System-on-Chip (RFSoc), whose

performance suggests that it is an attractive platform for the CHORD F-Engine. Finally, in Chapter 4, I discuss challenges and systematics introduced in the use of ground planes for global 21-cm experiments, and describe how MIST takes an alternative approach of using no ground plane and instead focusing on characterizing the soil electrical properties. I present the Antenna Modeler and Performance Emulator (AMPERE) and the System for Optimizing Impedance Likelihood (SOIL), two Python-based packages that I developed that work together to measure soil electrical parameters using measured and simulated impedances of the MIST antenna. I conclude with a summary of key results of this thesis and outline future work for each topic presented.

Chapter 2

The Deep Dish Development Array

(D3A)

The Deep Dish Development Array (D3A) is a pathfinder for CHORD technologies located near the proposed site of CHORD at the DRAO in Penticton, British Columbia. D3A is composed of three six-meter dishes and two three-meter dishes, observing at 400–1600 MHz and 400–800 MHz, respectively. The three-meter dishes correspond to the first demonstration of dish surface precision prior to the construction of the six-meter dishes and predate CHORD. In the measurements and discussion presented in this chapter, I will refer to only the six-meter dishes as they are similar to the design that will be used for CHORD dishes. The goal of D3A is to develop and demonstrate the performance and systematic control of the key technologies that will enable CHORD to meet its target specifications.

In Section 2.1, I provide an overview of the signal pipeline and key technologies developed with the six-meter dishes at D3A. In Section 2.2, I present a measurement of the system temperature T_{sys} of each six-meter dish and compare with CHORD’s target $T_{\text{sys}} = 30$ K.

2.1 D3A Technologies

Figure 2.1 shows the three six-meter dishes of D3A at the DRAO. The dishes are referred to as the “West,” “Mid,” and “East” dishes, corresponding to the labels shown in the main image and inset of Figure 2.1. The inset of Figure 2.1 also displays the geometry of the six-meter dishes, including the baselines, or the distances between each dish, as well as their arrangement with respect to the cardinal directions. The dishes are arranged along the EW direction for simplicity of analysis.

A simplified block diagram of the D3A pipeline is shown in Figure 2.2. In this section, I describe crucial elements of the pipeline and key developments of the technologies that will enable CHORD to be a leading instrument in the next generation of radio cosmology and astrophysics. Note that the signal processing pipeline of D3A has evolved significantly since it was first constructed; the pipeline presented here thus corresponds with the most up-to-date components as of the date of submission of this work.

While Condon, James J. and Ransom, Scott M. (2016) suggest that reasonable performance can be achieved with a dish surface *accuracy* of $\lambda_{\text{min}}/16$, CHORD requires a dish surface *precision* of 0.2 mm, or approximately $\lambda_{\text{min}}/1000$, to be able to implement the



Figure 2.1: A photo of the three six-meter dishes of D3A at the DRAO, referred to as the East, Mid, and West dishes from left to right. The blockhouse housing the D3A signal processing hardware is located at the far center-right in the main image and on the left in the inset. The inset satellite image displays the geometry of D3A, including the baselines between each dish and the cardinal coordinates; note that D3A is arranged along the EW axis. Main photo taken by Dr. Dallas Wulf. Inset satellite image captured with Google Maps.

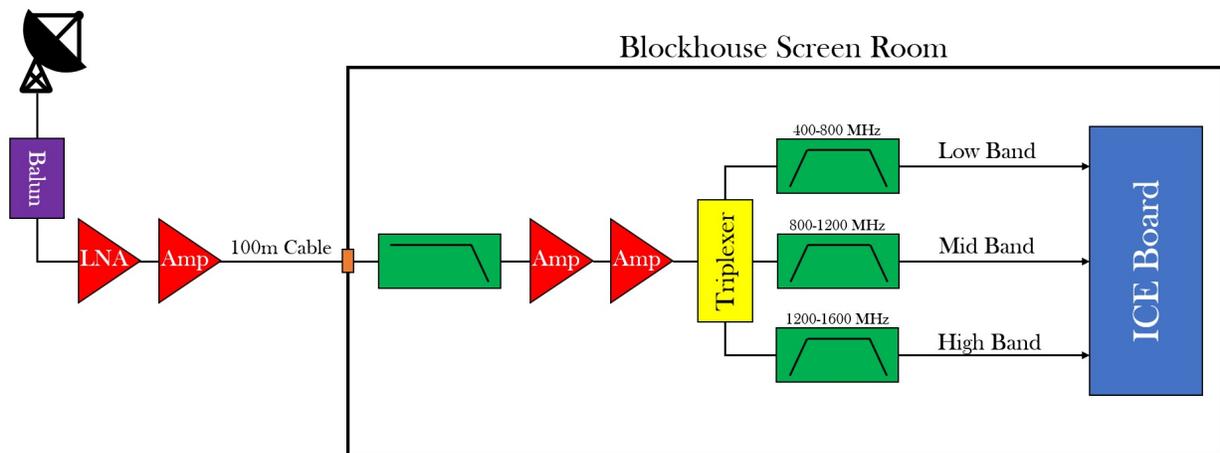


Figure 2.2: A simplified diagram of the D3A signal pipeline corresponding to the CHORD-like six-meter dishes. Bias tees are omitted for brevity; one bias tee is located on the backplane of the feed, and supplies the LNA and one amplifier, while the other bias tee supplies the two amplifiers within the blockhouse screen room. Note that this diagram corresponds to a single polarization on a single dish, and the frequency ranges of each band are approximate and do not correspond exactly with the 3 dB points of the band-pass filters. Individual component details can be found in [Pieters \(2021\)](#).

redundant calibration techniques described in Chapter 1 (Islam et al., 2022). Therefore, the dishes of D3A are made of a fiberglass composite material with an embedded metal mesh, and are built with a monolithic design, using a single mold to create the three dishes (Pieters, 2021). The use of a composite material in combination with a monolithic design improves the redundancy between each dish and accuracy control in comparison to other dish fabrication techniques. The composite material is effectively transparent at radio frequencies while the metal mesh acts as the reflector of the dish. The dishes are 2 mm thick, which means that the reflector follows the surface shape. As such, laser-tracker surveys of the dish surface can be made as an indirect measurement of the accuracy of the reflector in comparison with a perfect paraboloid. The dishes have been demonstrated to be accurate to the sub-mm levels required for CHORD (Islam et al., 2022). The ratio of the focal distance f to the diameter D of the dishes is $f/D = 0.25$, such that the focus coincides with the dish aperture plane. CHORD will correspond to $f/D = 0.21$, such that the focal plane sits slightly below the plane of the dish edges in order to reduce crosstalk between neighboring dishes and dish spillover effects.

Each dish is instrumented with a dual-polarization Vivaldi-like (MacKay et al., 2022) feed designed to operate over the wide 300–1500 MHz CHORD bandwidth. The feeds are cost-effective, at $\lesssim 75$ USD for large-scale production, and are approximately one-third the size of the longest observing wavelength. The feed has been optimized to achieve a high degree of performance across the wide CHORD bandwidth. MacKay et al. (2022) find that

simulations of the feed coupled to a dish with $f/D = 0.21$ have an aperture efficiency between 0.4 to 0.6, and its reflection coefficient is $\lesssim -10$ dB across the CHORD bandwidth. The simulations suggest a narrow beam, high forward gain, and sidelobe levels below -20 dB.

Following the feed is a printed-circuit board (PCB) microstrip balun that converts the differential signal from the antenna to a single-ended signal, with a low-noise amplifier (LNA) positioned immediately after. The feed impedance and beam shape is sensitive to the geometry of the balun and can be tuned to optimize performance (MacKay et al., 2022). The first active amplifier in the signal chain typically dominates the receiver temperature. The LNA serves to amplify the weak signal picked up by the antenna at the cost of introducing its own noise. Therefore, it is crucial that the LNA minimizes the amount of added noise to the system by being matched with the feed. D3A has implemented a custom-built LNA designed at the University of Calgary to meet the stringent requirements of the CHORD system temperature. The LNA has a reflection coefficient of $\lesssim -8$ dB from 320–1580 MHz, gain of 32 ± 1.2 dB, and can reach sub-20 K noise temperatures from 500–1400 MHz (Lai, 2022) without cryogenic cooling. Simulations including the noise contributions from the feed, LNA, back-end, ground spill, and sky have suggested that $T_{\text{sys}} < 30$ K across most of the CHORD bandwidth (MacKay et al., 2022).

After the LNA, the signal is further amplified by a ZX60-P103LN+ amplifier, and 100 meters of coaxial cable carries the signal to a screen room in the building adjacent to D3A, referred to as the blockhouse. RF-over-fiber (RFoF) cables have also been explored

as an alternative option for signal transmission at D3A. The coaxial cables from each dish enter the screen room through a bulkhead before being connected to the D3A rack. A copper mesh surrounding the screen room and the bulkhead connection both ensure that RFI from digital systems does not propagate outside of the room where it could be captured by the various instruments operating at the DRAO.

An image of the signal processing hardware used inside the screen room is shown in the top panel of Figure 2.3. Following the bulkhead, the signals pass through a low-pass filter (LPF) and two amplifiers before the signal is split into three copies by a triplexer. Each copy of the signal is filtered into a different Nyquist zone, which is necessary because the ADC samples at 800 MSPS and therefore cannot directly recover the full 400–1600 MHz band. These components are located in the middle of the top panel, mounted on plastic boards and inserted into a metal rack.

D3A uses two ICE boards for the digitization, channelization, and cross-correlation of inputs (Bandura et al., 2016). An example of an ICE board is shown in the bottom panel of Figure 2.3. The D3A ICE boards are installed at the top of the rack in the top panel. The ICE system is a low-cost and scalable radio astronomy signal processing platform that consists of FPGA-based motherboards, platform-specific daughter mezzanine boards, and high-density backplanes capable of 10 Gbit/s corner-turn operations. The ICE system is used in CHIME, the South Pole Telescope, and HIRAX, among other applications. Mounted on each ICE board at D3A are two daughter mezzanine boards each equipped with two

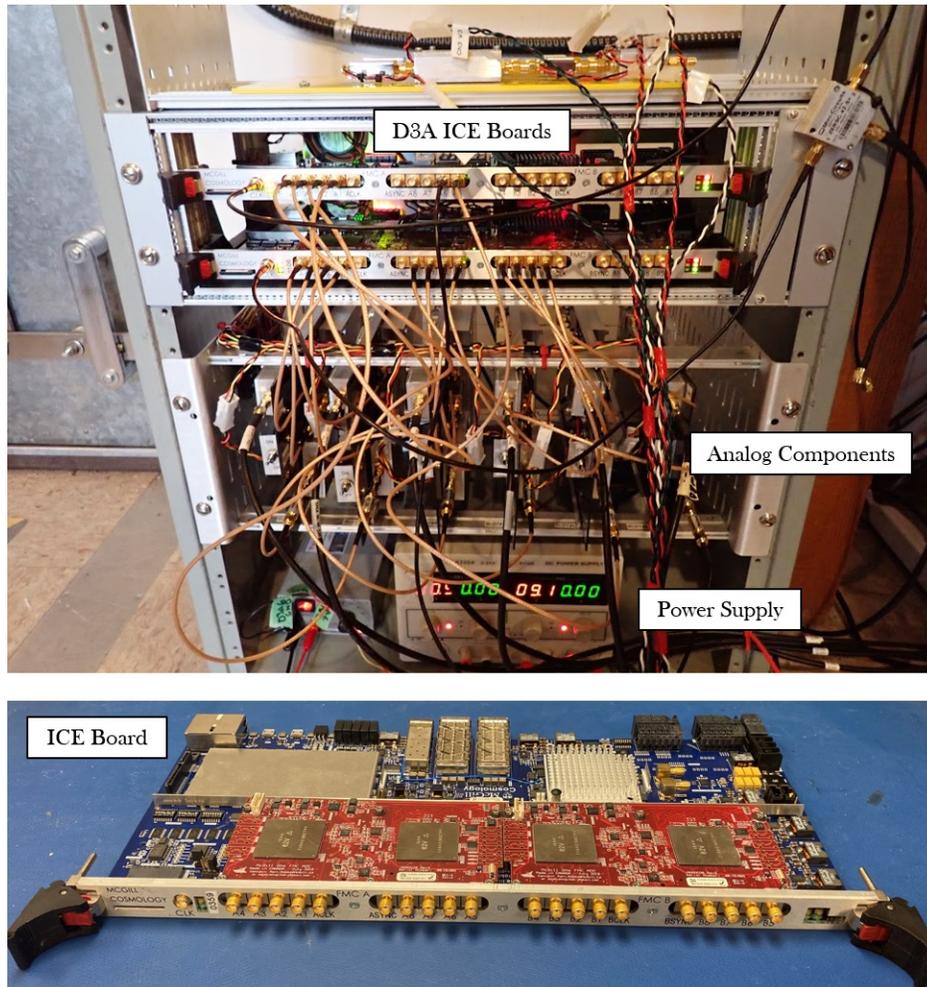


Figure 2.3: Top: the D3A backend inside the screen room of the blockhouse at the DRAO. The analog components, including filters, amplifiers, and triplexer are mounted on plastic boards and installed in the middle of the rack. The D3A ICE boards are mounted at the top of the rack and digitize, channelize, and correlate the analog signals. A power supply provides power to active analog components and the ICE boards. **Bottom:** an example of an ICE board (blue) with mounted ADC mezzanines (red). The FPGA is located underneath the white heat sink towards the center-right of the image.

Teledyne EV8AQ160 ADCs. D3A will be used to evaluate the performance of alternative digital systems.

Each ICE board is capable of processing up to 16 inputs at a sampling rate of 800 MSPS with 8 bits of resolution. The sampling rate of the EV8AQ160 ADCs does not meet the necessary 3 GSPS required by the Nyquist-Shannon sampling theorem to directly sample the full CHORD bandwidth. Therefore, D3A implements a triplexer in combination with band-pass filters to split the signal into the second, third, and fourth Nyquist zones defined by the 800 MSPS sampling rate. These zones are respectively referred to as the “low” band, covering 400–800 MHz, the “mid” band, covering 800–1200 MHz, and the “high” band, covering 1200–1600 MHz, as shown in Figure 2.2. Because D3A recovers data beyond the first Nyquist zone, it is crucial that the filters used for each band have sharp roll-off towards Nyquist frequencies in order to minimize aliasing. New filters are currently being designed in order to improve roll-off in comparison with the commercial filters currently used.

The digitized signals from the ADC mezzanines are sent to the FPGA on the ICE board, which computes the fast Fourier transform (FFT) of each channel using firmware developed for the CHIME F-Engine. After ten seconds of integrating each FFT, the FPGA then computes the auto- and cross-correlations for each input, and sends this data to the D3A control computer. The control computer runs the ICE system’s Python-based control code `pychfpga`¹, which communicates with the ICE boards and captures and processes data. D3A

¹<https://bitbucket.org/winterlandcosmology/pychfpga/src/master/>

requires two ICE boards to process all the data from the two three-meter dishes and the three six-meter dishes, each polarization, and each band. The ICE boards receive a clock signal and a GPS time signal from a Spectrum Instruments TM-4 GPS receiver which is used to synchronize the ADCs on each board.

D3A can be scheduled for observations remotely, and has observed typical radio interferometer calibrators such as the supernova remnant Cassiopeia A, the radio galaxy Cygnus A, and the Crab Nebula (Taurus A). The data saved by the control computer are automatically uploaded to the Cedar compute cluster operated by the Digital Research Alliance of Canada ([Digital Research Alliance of Canada, 2021](#)).

2.2 D3A System Temperature

In radio astronomy, power can be described with a noise temperature

$$T_{\text{N}} = \frac{P_{\nu}}{k_{\text{B}}}, \quad (2.1)$$

where P_{ν} is the total frequency-dependent power emitted or produced by a source (e.g., the sky, the receiver, the ground, etc.), and k_{B} is Boltzmann's constant. $P_{\nu} = k_{\text{B}}T$ is the low-frequency approximation of the thermal noise power per bandwidth of a resistor with temperature T . As such, T_{N} can be interpreted as the physical temperature of a matched (e.g., 50Ω) resistor whose thermally-generated power is the same as that emitted or produced

by a source.

The system temperature T_{sys} is then the total noise power collected from all sources referenced to the receiver input. In the context of CHORD, T_{sys} can be expressed as the sum of all noise temperatures from the dominant sources across the CHORD bandwidth, in approximate order of contribution, as

$$T_{\text{sys}} = T_{\text{r}} + T_{\text{sky}} + T_{\text{spill}}, \quad (2.2)$$

where T_{r} is the receiver noise temperature, with contributions from the LNA, feed, and backend, T_{sky} is the sky noise temperature, with contributions from the CMB, Milky Way, and atmosphere, and T_{spill} is the spillover noise temperature associated with the ground. T_{sys} is an important metric of an interferometer's sensitivity, especially for measuring low-power sources like the 21cm signal. With a three-element interferometer like D3A, one can take advantage of the phase closure of the baselines, which states that the vector sum of the baselines is equal to zero, in order to measure T_{sys} . In this section, I present a measurement of the D3A T_{sys} using an observation of Cygnus A conducted between approximately 04:00 to 08:00 on May 6, 2023.

2.2.1 Principles of Interferometry

This section presents an overview of the basic principles of interferometry, and provides the framework that leads to the derivation of T_{sys} from observational data. The derivations of

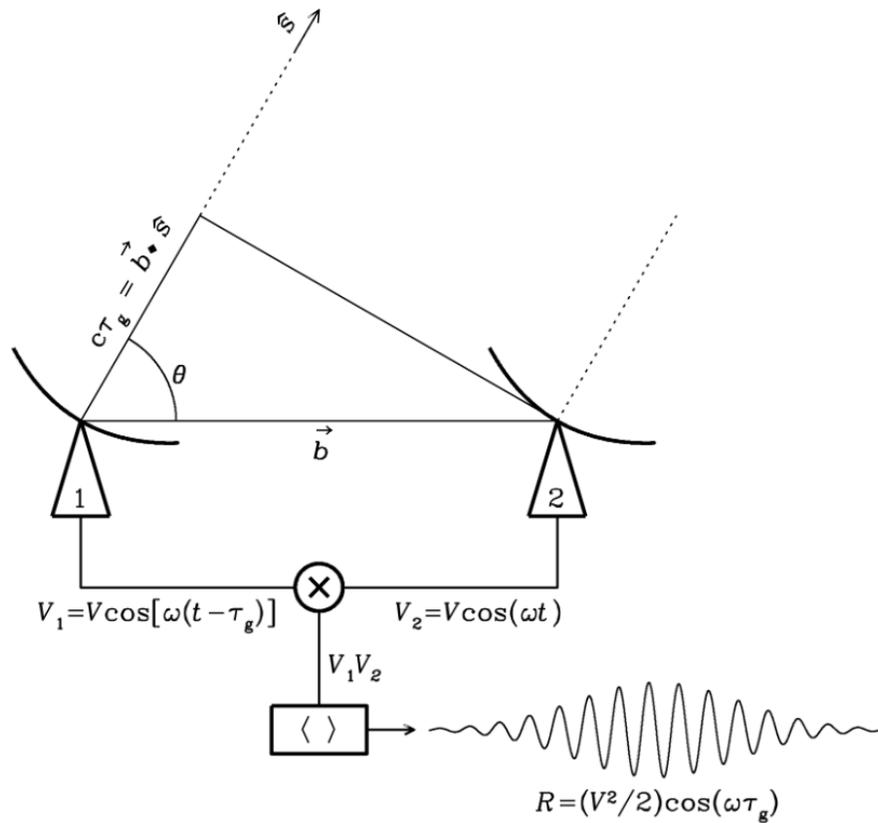


Figure 2.4: A diagram of a two-element interferometer. The two elements are separated by a baseline \vec{b} . A plane wave traveling along direction \hat{s} arrives at element 1 with a geometric delay defined by Equation 2.3 with respect to element 2. The voltage signals from each element are multiplied, or correlated, producing a cosine correlator response defined by Equation 2.5. Diagram from Condon, James J. and Ransom, Scott M. (2016).

key equations presented in this and the following sections are drawn from [Condon, James J. and Ransom, Scott M. \(2016\)](#) and [Thompson et al. \(2017\)](#). Figure 2.4 displays a cartoon of a two-element interferometer, where “element” refers to a single antenna mounted on a dish. The two elements are separated by a distance \vec{b} , referred to as the baseline. Each dish is pointed towards the sky along a direction \hat{s} . When a plane wave emitted by a point source with source intensity distribution $I(\hat{s})$ (in units of Jy sr⁻¹) travelling along \hat{s} with electric field plane perpendicular to \hat{s} (the plane of the page in Figure 2.4) arrives at the interferometer, element 1 absorbs radiation from the plane wave with a geometric delay

$$\tau_g = \frac{\vec{b} \cdot \hat{s}}{c}, \quad (2.3)$$

relative to element 2. As such, for a sinusoidal plane wave with amplitude V and angular frequency ω , the measured voltages at each element are

$$V_1 = V \cos(\omega(t - \tau_g)) \quad \text{and} \quad V_2 = V \cos(\omega t). \quad (2.4)$$

The two voltages are then multiplied and averaged to yield the cosine correlator response

$$R_c = \langle V_1 V_2 \rangle = \frac{V^2}{2} \cos(\omega \tau_g) = \frac{V^2}{2} \cos\left(\frac{\omega \vec{b} \cdot \hat{s}}{c}\right), \quad (2.5)$$

which is dependent upon the source intensity distribution $I(\hat{s})$ and the collecting area of each dish. As the point source transits across the sky as the Earth rotates, R_c varies sinusoidally,

and the sinusoids are referred to as fringes.

If instead our source is slightly spatially-extended (much smaller than the primary beamwidth), then the cosine correlator response is considered a sum of individual point sources according to

$$R_c = \int I(\hat{s}) \cos(2\pi\vec{b} \cdot \hat{s}/\lambda) d\Omega, \quad (2.6)$$

where Ω refers to the solid angle of the source. However, if our slightly extended source is composed of even and odd components, or $I = I_{\text{even}} + I_{\text{odd}}$, then R_c is only sensitive to the even components. As such, a phase delay of $\pi/2$ is introduced along the signal path of one element in order to be sensitive to the odd components, leading to the definition of the sine correlator response

$$R_s = \int I(\hat{s}) \sin(2\pi\vec{b} \cdot \hat{s}/\lambda) d\Omega. \quad (2.7)$$

A complex correlator combines cosine and sine correlators, and measures a quantity referred to as the visibility

$$\mathcal{V} = R_c - iR_s = \int I(\hat{s}) e^{2\pi\vec{b} \cdot \hat{s}/\lambda} d\Omega, \quad (2.8)$$

where the symbol \mathcal{V} is used to distinguish from the voltage V . The visibility is the basic measurement unit of interferometry, and enables sensitive measurements of sources that

would otherwise require enormous dishes in order to observe them at radio frequencies.

2.2.2 A Recipe to Measure T_{sys} from Phase Closure

In the previous section, an ideal interferometer with a simplified receiver was considered. In this section, I discuss the interferometer corresponding to the three six-meter dishes of D3A with realistic instrumental effects. I refer to quantities corresponding to the dishes of D3A with subscripts w, m, and e for the West, Mid, and East dishes, respectively. From the complex visibility measured for an observation of Cygnus A, I will derive an expression for T_{sys} , and present a measurement of the D3A T_{sys} .

For a real interferometer, the visibility \mathcal{V}_{wm} between the West and Mid dishes is also dependent upon the antenna beam pattern, the gain introduced by the analog amplification and filtering system, and the conversion of analog to digital signals. There are in fact two unique visibilities \mathcal{V}_{wm} : one for the X-polarization, and one for the Y-polarization. Note that while each polarization is ideally not sensitive to the other, some polarization leakage may exist. I will omit a subscript denoting the polarization of a given visibility in this section for simplicity, noting that it is implied that the term for a visibility is used for both polarizations.

I refer to the combination of a nominally time- and spatially-independent gain with the variable g , which corresponds to the total gain of all active and passive components along the signal pipeline. Note that in reality, g is not necessarily time-independent due

to temperature changes or other environmental effects, though in this analysis I will treat g as a constant. The beam $B(\hat{s})$ can be considered a spatially-dependent gain that may be assumed to vary approximately according to a Gaussian function. I will refer to g as “gain” and $B(\hat{s})$ as “beam” when necessary to avoid confusion, and in general, both terms should not be assumed to be equivalent for a given pair of dishes. Note that both g and $B(\hat{s})$ are also frequency-dependent quantities — I will omit frequency notation for simplicity throughout this section. Accounting for $B(\hat{s})$ in the visibility \mathcal{V}_{wm} leads to the expression

$$\mathcal{V}_{\text{wm}} = \int g_{\text{w}}^* g_{\text{m}} B_{\text{w}}^*(\hat{s}) B_{\text{m}}(\hat{s}) I(\hat{s}) e^{-i2\pi \vec{b}_{\text{wm}} \cdot \hat{s} / \lambda} d^2 \hat{s}, \quad (2.9)$$

where I have introduced terms for the gains g , noting that the subscripts denote the gain g and beam $B(\hat{s})$ associated with a given dish. In this expression, \mathcal{V}_{wm} is in units of analog-to-digital units squared (ADU^2), each g is in units of $\text{ADU Jy}^{-1/2}$, and each $B(\hat{s})$ is normalized and thus dimensionless. The unit ADU^2 results from the digitization of input signals, in which the signals are quantized according to digital codes, or ADU , and the correlation (multiplication) of these signals which squares the unit. Note that I have also converted the unit of integration from solid angle infinitesimal $d\Omega$ to Cartesian infinitesimal $d^2 \hat{s}$, which will help to simplify this expression. In order to measure T_{sys} , the gain g must first be determined from the signal pipeline to convert the visibilities from ADU^2 to Jy , and then determine the forward gain to convert from Jy to K by measuring the beam solid angle Ω .

A closed expression for the gains that can be solved with known quantities must first be

defined. I approximate the source intensity distribution $I(\hat{s})$ as a point-source² according to

$$I(\hat{s}) = S_0 \delta(\hat{s} - \hat{s}_0), \quad (2.10)$$

where S_0 is the known frequency-dependent flux density of Cygnus A in units of Jy derived from [Perley and Butler \(2017\)](#), and \hat{s}_0 is the direction of the Cygnus A. Note that while Cygnus A's position \hat{s}_0 is a fixed quantity when expressed in right ascension and declination, it is effectively a time-varying quantity due to the Earth's rotation. The visibility \mathcal{V}_{wm} can then be reduced to

$$\mathcal{V}_{\text{wm}} = g_{\text{w}}^* g_{\text{m}} B_{\text{w}}^*(\hat{s}_0) B_{\text{m}}(\hat{s}_0) S_0 e^{-i2\pi \vec{b}_{\text{wm}} \cdot \hat{s}_0 / \lambda}. \quad (2.11)$$

Similar expressions can be derived for V_{me} and V_{ew} . If S_0 , \hat{s}_0 , and the baselines are known, then there are three equations with six unknown quantities: g_{w} , g_{m} , g_{e} , B_{w} , B_{m} , and B_{e} .

The baselines could in principle be physically measured. However, one can take advantage of the principle of phase closure for a three-element interferometer, which states that the sum of the baseline vectors is zero, or

$$\vec{b}_{\text{wm}} + \vec{b}_{\text{me}} + \vec{b}_{\text{ew}} = 0. \quad (2.12)$$

²The assumption that Cygnus A is a point-source in the context of D3A can be justified by noting that the maximum extent of Cygnus A is 2 arcmin ([Perley and Butler, 2017](#)), which is smaller than the highest resolution θ of D3A as defined by the longest baseline \vec{b}_{ew} , which is $\theta \sim \lambda_{\text{max}} / |\vec{b}_{\text{ew}}| = 0.1875 \text{ m} / 39 \text{ m} = 16.5 \text{ arcmin}$.

This expression will allow for the elimination of the exponential terms from each expression, and reduce the equations to a simpler form that can be solved directly. One can write

$$\frac{\mathcal{V}_{\text{wm}}\mathcal{V}_{\text{ew}}}{\mathcal{V}_{\text{me}}^*} = \frac{[g_{\text{w}}^*g_{\text{m}}B_{\text{w}}^*(\hat{s}_0)B_{\text{m}}(\hat{s}_0)S_0e^{-i2\pi\vec{b}_{\text{wm}}\cdot\hat{s}_0/\lambda}][g_{\text{e}}^*g_{\text{w}}B_{\text{e}}^*(\hat{s}_0)B_{\text{w}}(\hat{s}_0)S_0e^{-i2\pi\vec{b}_{\text{ew}}\cdot\hat{s}_0/\lambda}]}{g_{\text{m}}g_{\text{e}}^*B_{\text{m}}(\hat{s}_0)B_{\text{e}}^*(\hat{s}_0)S_0e^{i2\pi\vec{b}_{\text{me}}\cdot\hat{s}_0/\lambda}}, \quad (2.13)$$

which reduces to

$$\frac{\mathcal{V}_{\text{wm}}\mathcal{V}_{\text{ew}}}{\mathcal{V}_{\text{me}}^*} = g_{\text{w}}^*g_{\text{w}}B_{\text{w}}^*(\hat{s}_0)B_{\text{w}}(\hat{s}_0)S_0e^{-i2\pi(\vec{b}_{\text{wm}}+\vec{b}_{\text{me}}+\vec{b}_{\text{ew}})\cdot\hat{s}_0/\lambda}. \quad (2.14)$$

Here is where the principle of phase closure helps to simplify the above expression — the complex exponential term contains Equation 2.12, such that the argument of the exponential is zero, and the expression becomes

$$\frac{\mathcal{V}_{\text{wm}}\mathcal{V}_{\text{ew}}}{\mathcal{V}_{\text{me}}^*} = |g_{\text{w}}|^2|B_{\text{w}}(\hat{s}_0)|^2S_0, \quad (2.15)$$

which can be rearranged as

$$|g_{\text{w}}|^2|B_{\text{w}}(\hat{s}_0)|^2 = \frac{1}{S_0} \frac{\mathcal{V}_{\text{wm}}\mathcal{V}_{\text{ew}}}{\mathcal{V}_{\text{me}}^*}, \quad (2.16)$$

where the values on the right-hand side are known or measured quantities. If the beam is then normalized at the time of transit of Cygnus A such that $|B_{\text{w}}(\hat{s}_0)|^2 = 1$, the gain can

then be solved according to

$$|g_w|^2 = \frac{1}{S_0} \frac{\mathcal{V}_{wm} \mathcal{V}_{ew}}{\mathcal{V}_{me}^*} \Big|_{\text{Transit}}, \quad (2.17)$$

where ‘‘Transit’’ refers to the transit time. Similar expressions can be derived for the Mid and East dishes.

The above expression assumes that $|B_w(\hat{s}_0)|^2 = 1$ at the time of transit, which is not necessarily true for a real beam. The boresight angle θ , or the angle between \hat{s}_0 and the geometric pointing vector of the dish \hat{s} , is indeed equal to 0 at the transit time. However, the beam center is not necessarily aligned along \hat{s} and may be offset by some angle θ_0 , such that $|B_w(\hat{s}_0)|^2 = 1$ at $\theta = \theta_0$ instead of at $\theta = 0$. This effectively means that the source may pass through the center of the beam before or after it transits. In fact, the source will not pass through the center of the beam at all if the dish is tilted at an incorrect angle.

Therefore, $|g_w|^2 |B_w(\hat{s}_0)|^2$ must be fit in order to solve for $|g_w|^2$, which makes fewer assumptions about exactly when $|B_w(\hat{s}_0)|^2 = 1$. I assume that the beam corresponds to a Gaussian function

$$|g_w|^2 |B_w(\theta)|^2 = |g_w|^2 \exp \frac{(\theta - \theta_0)^2}{2\sigma^2}, \quad (2.18)$$

where $|g_w|^2$ is the amplitude, θ is the boresight angle, θ_0 is the boresight angle of the beam center and σ is the standard deviation of the Gaussian. Note that $\theta_{\text{FWHM}} = 2\sqrt{2 \ln 2} \sigma$ is

the full-width half maximum (FWHM), or the width between the boresight angles where the normalized beam is equal to $1/2$. I fit the beams as a function of θ for each dish, frequency, and polarization, from which I recover the squared gains $|g|^2$ for each.

Figure 2.5 displays the squared gains $|g|^2$ for each dish and each polarization. It is expected that the gain varies smoothly as a function of frequency as it is representative of electronic components with smooth transfer functions. Since each frequency band has slightly different filter attenuation and amplifier gain, there may be distinct jumps from one band to the next. For example, the 1200–1600 MHz band of the Y-polarization gain is significantly higher than that of the X-polarization and the lower two bands for both polarizations (note that the axis is in log scale), which corresponds to additional amplifiers that have been added to that signal path. The effects of aliasing can also be seen around the Nyquist frequencies 1200 MHz and 1600 MHz, which manifests as additional power that leaks into a given band. Aliasing can result in undetermined fits for some frequency channels, seen as vertical spikes going off-plot towards the bottom, corresponding with extremely small values. Vertical spikes with high $|g|^2$ correspond with RFI.

I then normalize the beams by dividing Equation 2.16 by $|g|^2$, which for the West dish corresponds to

$$|B_w(\hat{s}_0)|^2 = \frac{1}{|g_w|^2 S_0} \frac{\mathcal{V}_{wm} \mathcal{V}_{ew}}{\mathcal{V}_{me}^*}. \quad (2.19)$$

Figure 2.6 displays the normalized beams of each dish and polarization as a function of

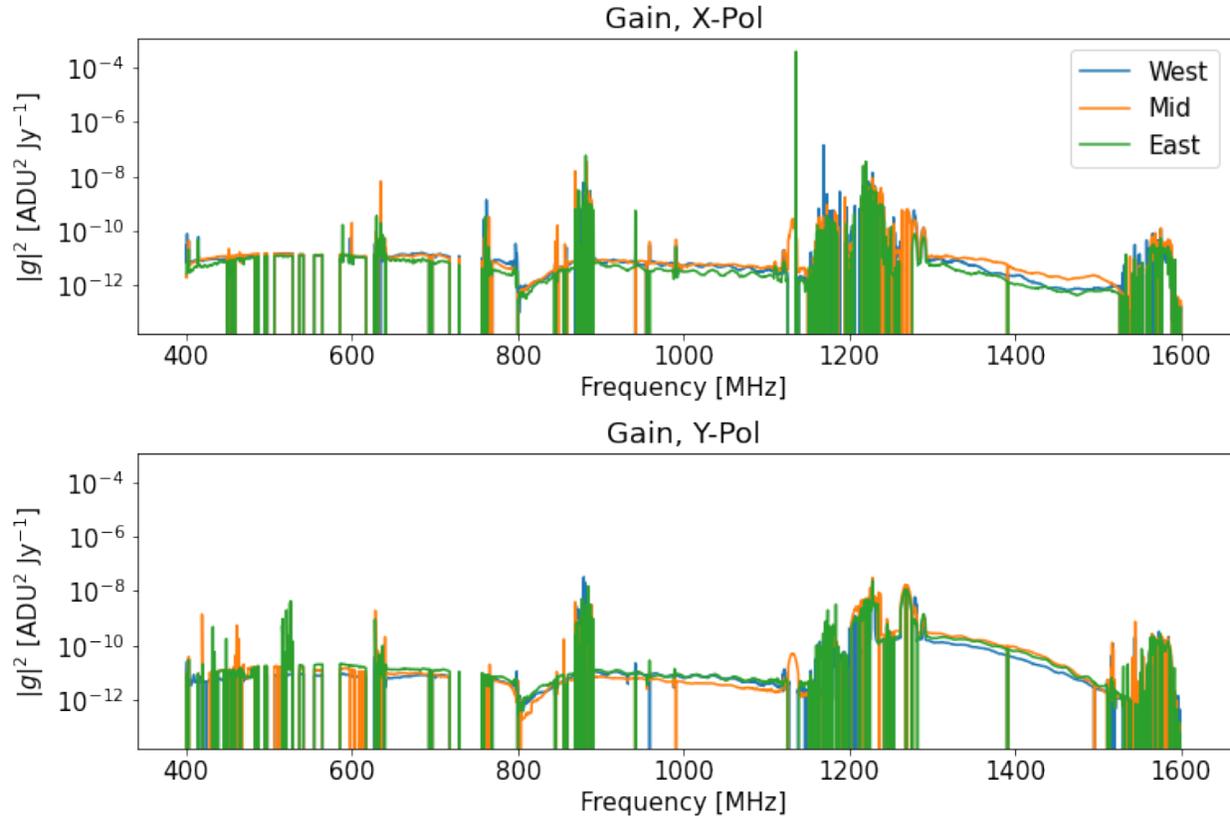


Figure 2.5: D3A squared gains $|g|^2$ as a function of frequency for each dish and polarization. $|g|^2$ varies smoothly (with small ripples from cable reflections) as a function of frequency, which is expected since the gain corresponds to the total frequency response of the components in the signal path, which have smooth transfer functions. $|g|^2$ is roughly consistent between the three dishes. The increased gain for each dish in the 1200–1600 MHz band of the Y-polarization results from additional amplifiers added to those signal paths.

boresight angle θ and frequency. White vertical stripes correspond to channels with RFI that saturate the gain axis. Aliasing is apparent from approximately 1100–1300 MHz as a result of the filters used after the triplexer as described in Section 2.1. Some sidelobes are observed in both polarizations, though they are more pronounced for the Y-polarization. The X-polarization beams are also slightly wider than those of the Y-polarization, which is explained in the discussion corresponding to Figure 2.7.

The gains $|g|^2$ can now be used to calibrate the autocorrelation visibilities \mathcal{V}_w , \mathcal{V}_m , and \mathcal{V}_e from ADU^2 to Jy. The autocorrelation visibility is the correlation of the voltage response of a single dish with itself. For example, the autocorrelation of the West dish for a transiting point-source can be derived from Equation 2.11 by replacing m with w according to

$$\mathcal{V}_w = |g_w|^2 |B_w(\hat{s}_0)|^2 S_0, \quad (2.20)$$

where I use only a single w for simplicity. The effect of the gain is calibrated by dividing by $|g_w|^2$, leaving us with an expression for \mathcal{V}_w in units of flux density:

$$\mathcal{V}_w [\text{Jy}] = \frac{\mathcal{V}_w [\text{ADU}^2]}{|g_w|^2 [\text{ADU}^2 \text{ Jy}^{-1}]} = |B_w(\hat{s}_0)|^2 S_0. \quad (2.21)$$

Now that the autocorrelation visibilities have been converted from ADU^2 to Jy, a similar conversion can be developed between flux density units in Jy to T_{sys} units in K. The conversion factor between flux density and T_{sys} is referred to as the forward gain. To

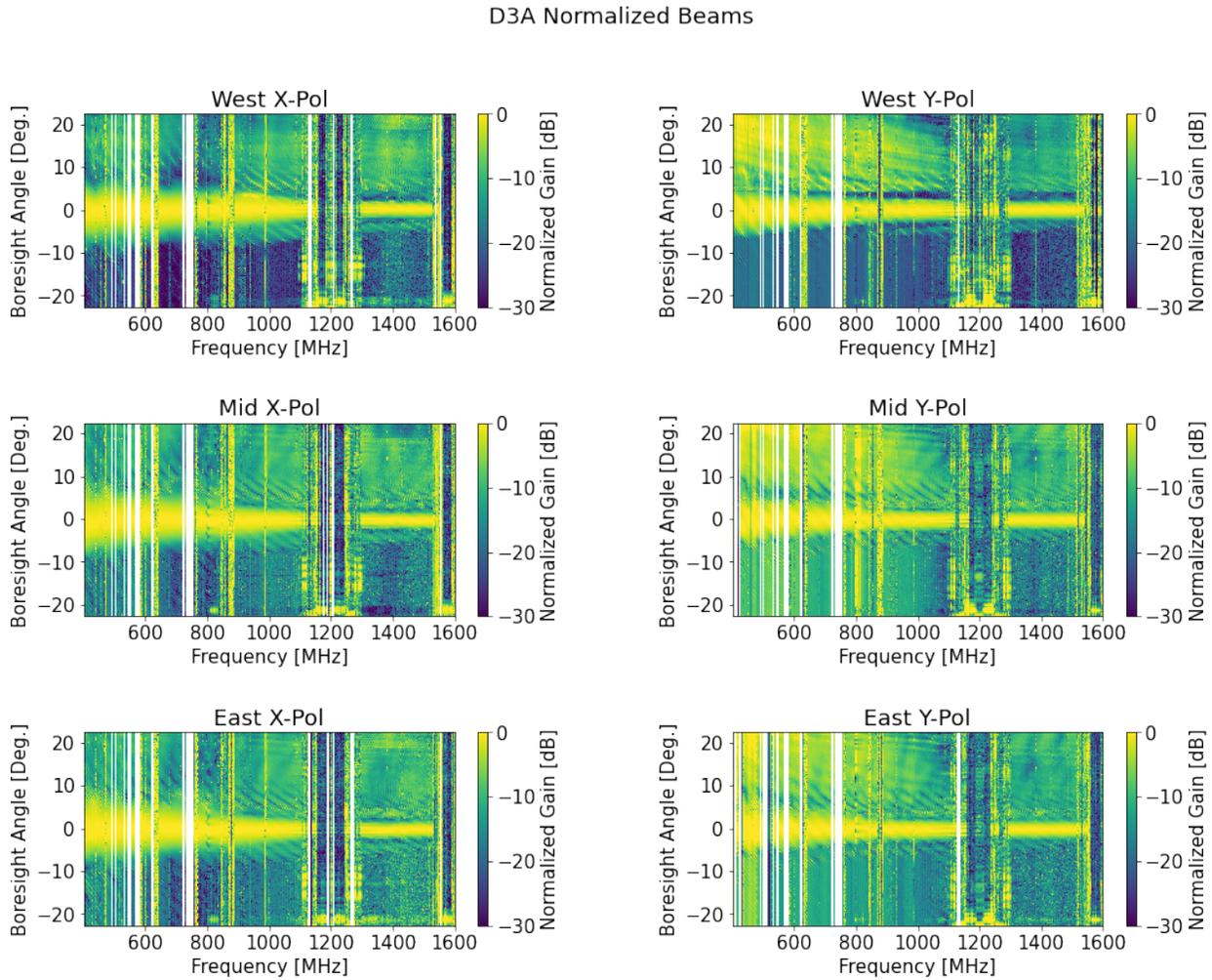


Figure 2.6: The normalized beams of each dish and polarization of D3A as a function of boresight angle and frequency. Aliasing is observed around the 1200 MHz and 1600 MHz Nyquist frequencies. The beams of the X-polarization are slightly wider than the beams of the Y-polarization due to an asymmetry in the radiation pattern of the feeds. The observed sidelobes are more pronounced for the Y-polarization, which is expected for the polarization corresponding to the narrower beam.

determine the forward gain, an estimate of the beam solid angle Ω is needed. In fact, the information to estimate Ω is already available from the best-fit θ_{FWHM} derived from the standard deviation of the Gaussian function in Equation 2.18.

Figure 2.7 displays θ_{FWHM} as a function of frequency for each dish and polarization. θ_{FWHM} decreases with frequency, which is expected since $\theta_{\text{FWHM}} \sim \lambda/D$ for a Gaussian beam, which is plotted as a dashed black line in the figure for comparison. In Table 2.1, I show the approximate effective apertures A_e for each dish, polarization, and frequency band. A_e is estimated from the best λ/A_e line that tracks the FWHM shown in Figure 2.7. While there is consistency in A_e between each dish and polarization towards higher frequencies, A_e can differ by over 1 m between the two polarization at lower frequencies, and there are inconsistencies of 0.6–1 m when comparing the West dish Mid and East dishes.

In the 400–800 MHz and 800–1200 MHz bands, the X-polarization FWHM is 1–2 degrees greater than the Y-polarization FWHM for each dish, while the difference in the 1200–1600 MHz band is a fraction of a degree. The differences are comparable with the simulations of the FWHM shown in Figure 13 of MacKay et al. (2022). While it is noted that the simulations correspond to $f/D = 0.21$, the behavior of the FWHM for $f/D = 0.25$ is expected to be consistent with the results shown in Figure 2.7.

The difference in θ_{FWHM} between each polarization is related to the design of the feed. MacKay et al. (2022) demonstrates that the radiation patterns of the feed along the E- and H-planes do not have equivalent widths. The feeds at D3A are aligned such that one

Band (MHz)	Effective Aperture A_e (m)					
	West, X	Mid, X	East, X	West, Y	Mid, Y	East, Y
400–800	4.6	4	4	6	5	5
800–1200	-	4.5	4.5	6	6	6
1200–1600	5.5	5.5	5.5	5.5	5.5	5.5

Table 2.1: The approximate effective aperture A_e corresponding to each dish and polarization. A_e is estimated from the λ/A_e line that best approximates the FWHM shown in Figure 2.7. The FWHM corresponding to the 800–1200 MHz band of the X-polarization of the west dish cannot be approximated by λ/A_e . In the lowest frequency band, A_e differs by approximately 0.6–1 m between the West dish in comparison to the Mid and East dishes. A_e is consistent between the dishes towards higher frequencies, with the exception of the X-polarization of the West dish in the 800–1200 MHz band. The inconsistency between the West dish in comparison with the Mid and East dishes indicates that D3A must focus on improving its redundancy in order to demonstrate that CHORD can meet the stringent requirements for 21-cm cosmology.

polarization is along the transit axis while the other is orthogonal to the transit axis. As such, one polarization probes a cut of the beam along the E-plane while the other probes the H-plane. Therefore, it is expected that θ_{FWHM} is greater for one polarization compared with the other.

For a symmetric Gaussian beam, the beam solid angle is

$$\Omega = \frac{\pi}{4 \ln 2} \theta_{\text{FWHM}}^2, \quad (2.22)$$

which assumes that θ_{FWHM} is equivalent for both polarizations. However, as shown in Figure 2.7, θ_{FWHM} is not necessarily the same for each polarization. Therefore, the beam solid angle

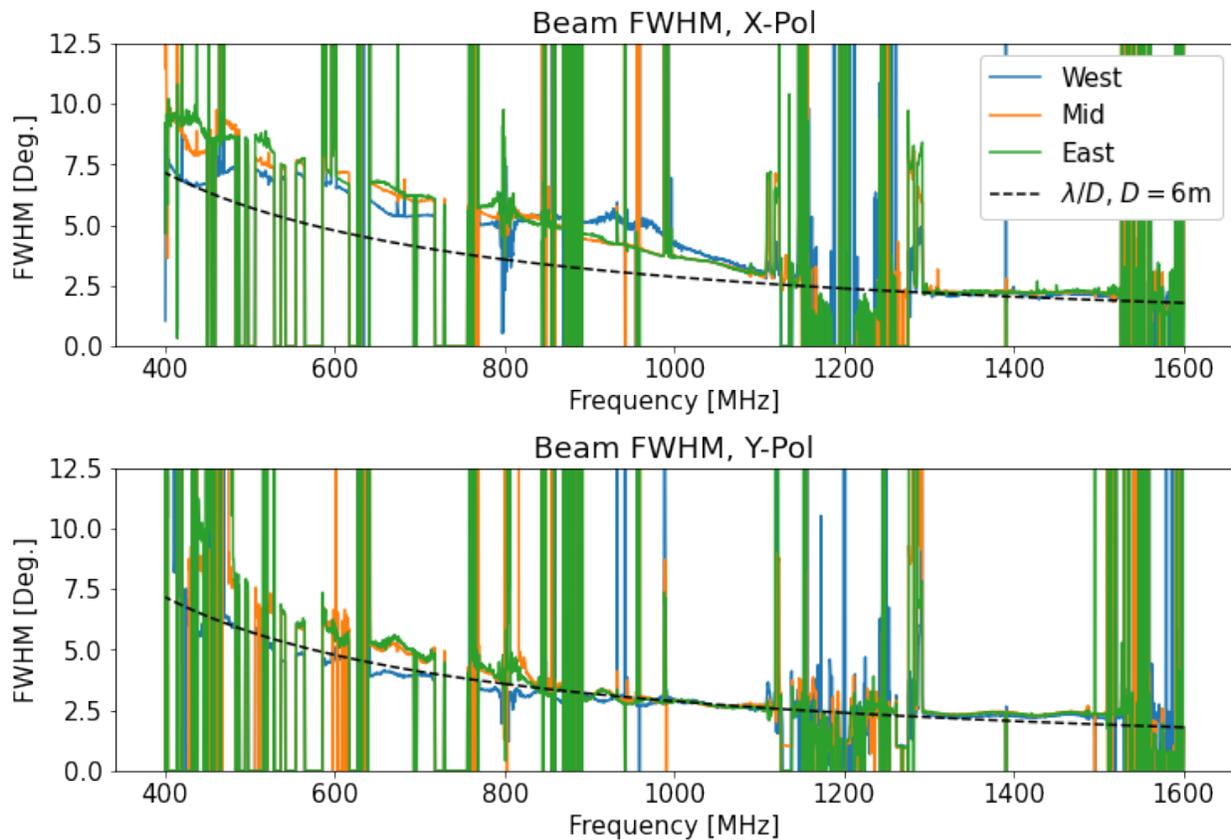


Figure 2.7: θ_{FWHM} for each dish and polarization of D3A as a function of frequency. $\theta_{\text{FWHM}} \sim \lambda/D$, corresponding to a Gaussian beam, is plotted for the geometric dish diameter $D = 6\text{ m}$ for comparison. The effective aperture A_e estimated from the λ/D line that best approximates the FWHM for each dish and polarization is shown in Table 2.1. θ_{FWHM} for the X-polarization is larger due to the asymmetry in the radiation patterns of the feeds (MacKay et al., 2022). However, there are notable distinctions when comparing θ_{FWHM} between each dish, which indicates systematic differences in the West dish.

Ω for an asymmetric or elliptical Gaussian beam can instead be expressed as

$$\Omega = \frac{\pi}{4 \ln 2} \theta_X \theta_Y, \quad (2.23)$$

where θ_X and θ_Y are the full-width half maxima of the two polarizations of the beam.

Figure 2.8 displays Ω for each dish calculated from Equation 2.23 and the theoretical Ω for a symmetric Gaussian beam with varying effective aperture A_e . Note that the theoretical Ω for a symmetric Gaussian beam derives from $\theta_{\text{FWHM}} \sim \lambda/D$. In the 400–800 MHz band, the Mid and East dishes can be best approximated by Ω corresponding to $A_e = 4.6$ m, but are more consistent with $A_e = 5.5$ m above ~ 950 MHz. The West dish can be best approximated by Ω corresponding to $A_e = 5.5$ m for most frequencies. At low frequencies, Ω is lower by approximately 0.002–0.01 sr for the West dish compared with the other two dishes, which results from θ_{FWHM} being smaller for both polarizations of the West dish.

The differences in θ_{FWHM} and Ω between dishes are a significant result, as D3A is designed to demonstrate that the technologies developed for CHORD can meet its redundancy and systematic control requirements. However, the differences observed when comparing the West dish with the Mid and East dishes are likely due to factors that can be controlled and thus improved. The West dish is instrumented with the first CHORD feed that was manufactured and assembled. It is known that the feed accumulated excess solder during assembly, which can add capacitance to the feed, altering its impedance and possibly the beam. In addition, the aluminum feeds are connected to the balun with an off-the-shelf gold

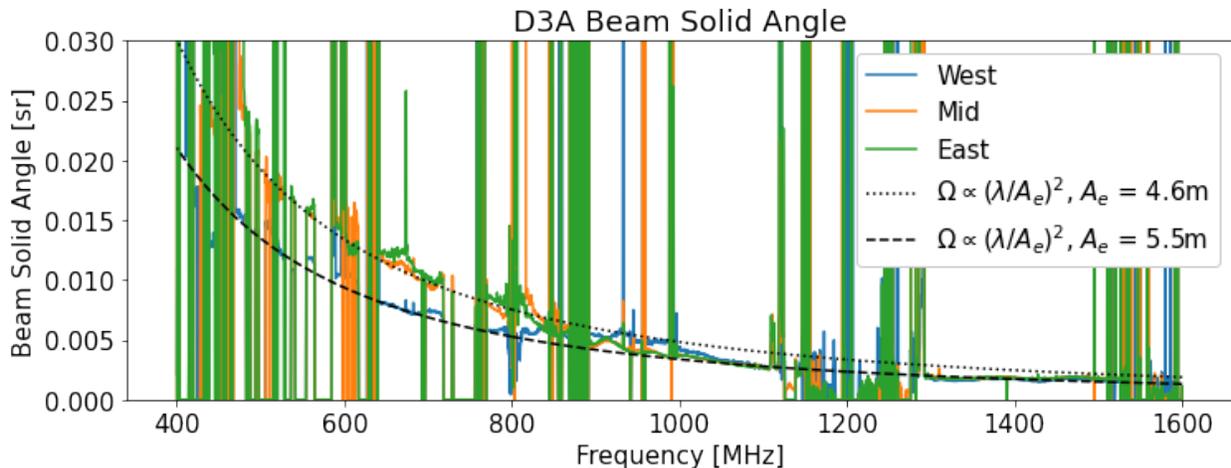


Figure 2.8: The beam solid angle of D3A as a function of frequency. Recall that I have assumed that the beam of each dish corresponds to an asymmetric Gaussian function. Ω for two symmetric Gaussian beams with varying A_e are plotted as dashed black lines for comparison. The Mid and West dishes can be approximated by Ω corresponding to a symmetric Gaussian beam with $A_e = 4.6$ m in the 400–800 MHz range, and $A_e = 5.5$ m above ~ 950 MHz. The West dish can be approximated by Ω corresponding to a symmetric Gaussian beam with $A_e = 5.5$ m for most frequencies. The differences between the West dish and the Mid and East dishes may result from the accumulation of solder on the feed of the West dish and the degradation of the balun connection to the feed.

pin, whose connection may have corroded since the assembly of the feed. However, the source of the differences is not yet confirmed. It is therefore crucial for the D3A collaboration to consider ways in which the redundancy can be improved.

The forward gain, or the ratio T_N/S , is an expression of a receiver's sensitivity to the flux density of a source. The forward gain is given by

$$\frac{T_N}{S} = \frac{\lambda^2}{2k_B\Omega}. \quad (2.24)$$

The forward gain is used to convert the autocorrelation visibilities from power units in Jy to T_{sys} units in K, which is expressed by the following for the West dish:

$$\mathcal{V}_w [\text{K}] = \frac{T_N}{S} \mathcal{V}_w [\text{Jy}] = \frac{\lambda^2}{2k_B \Omega} \mathcal{V}_w [\text{Jy}]. \quad (2.25)$$

The forward gain can also be used to estimate the aperture efficiency

$$\eta = \frac{A_e}{D} = \frac{2k_B T_N}{D S}. \quad (2.26)$$

The forward gain for each dish is shown in the top panel of Figure 2.9, along with the theoretical forward gain assuming $\Omega \propto (\lambda/D)^2$. The theoretical forward gain is approximately 1–5 mK Jy⁻¹ greater than the measured forward gain for non-RFI frequency channels (which in this plot is shown as empty frequency channels or vertical spikes, as the forward gain is inversely proportional to Ω). The forward gain of the West dish is inconsistent with the forward gains of the Mid and East dishes, which is expected considering the previous discussion of the inconsistency in the FWHM and beam solid angle.

The aperture efficiency η of the dishes is shown in the bottom panel of Figure 2.9. The West dish has a larger η , or equivalently a larger A_e , towards lower frequencies. A dish with a larger A_e is more sensitive, such that its measured noise temperature increases by a greater amount for the same source flux density; this is the essence of the forward gain, which itself

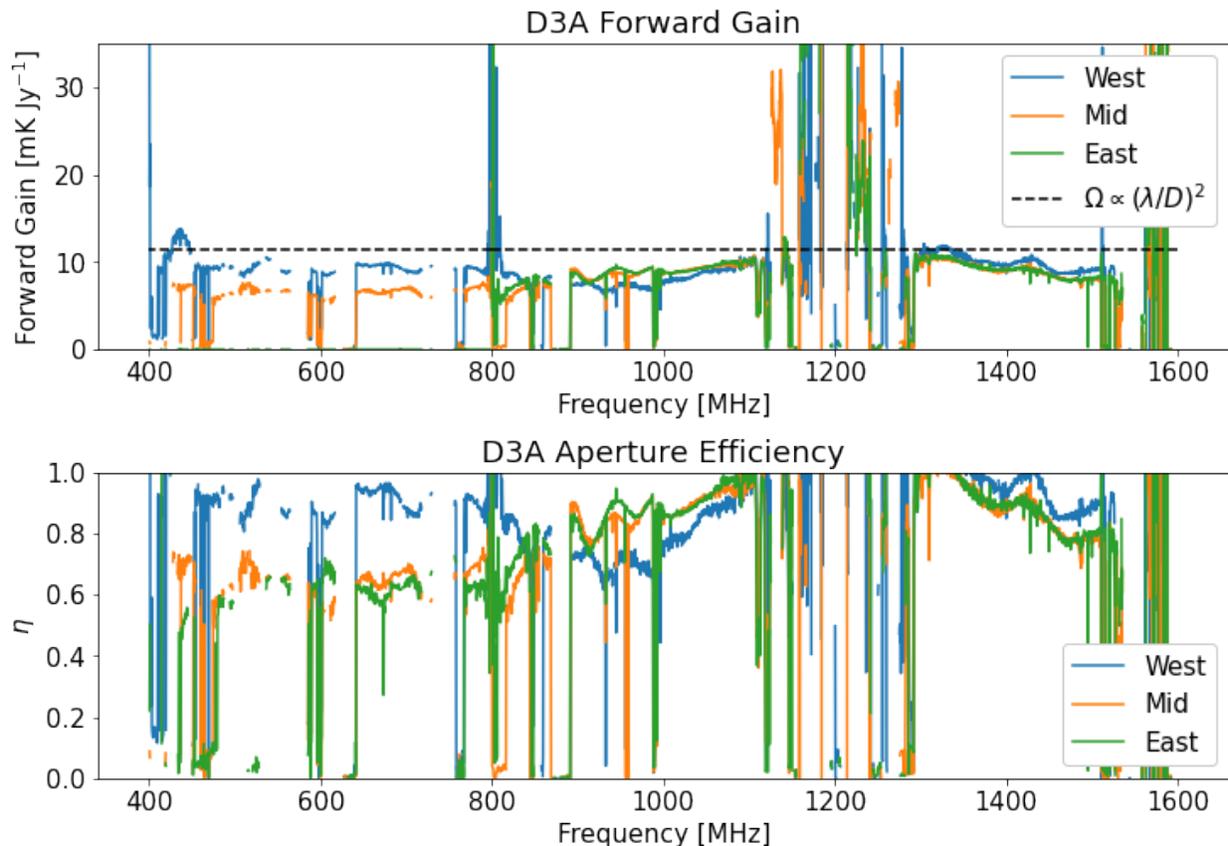


Figure 2.9: Top: The forward gain of D3A as a function of frequency. The theoretical forward gain for a symmetric Gaussian beam corresponding to $D = 6$ is 1–5 mK Jy⁻¹ greater than the measured forward gain. The forward gain of the West dish is inconsistent with the forward gains of the Mid and East dishes. **Bottom:** The aperture efficiency η of D3A as a function of frequency. η for the West dish is inconsistent with η for the Mid and East dishes. The behavior of η for each dish is roughly consistent with the effective aperture values presented in Table 2.1. The inconsistencies in the forward gain and η between the dishes may be due to differences in the assembly and degradation of the feed connection to the balun on the West dish. It may be necessary to reconsider other methods for connecting the feed to the balun if a faulty connection is found to be the source of the inconsistency.

is related to the size of the beam solid angle. However, the dish with the largest forward gain and η varies towards higher frequencies, with the Mid and East dishes showing better performance in the 800–1200 MHz band, before the West dish takes precedence again in the 1200–1600 MHz band for non-RFI and non-aliased channels. The behavior of the aperture efficiency is roughly consistent with the effective aperture values shown in Table 2.1.

With the forward gain in hand, I convert the autocorrelation visibilities \mathcal{V}_w , \mathcal{V}_m , and \mathcal{V}_e into units of K. The autocorrelation visibilities are time-dependent quantities that change according to time of day, solar activity, transient RFI, possible additional background sources, etc. Since T_{sys} as defined in Equation 2.2 consists of noise contributions from multiple elements and is a time-dependent quantity, the measured system temperature T_{sys} is defined here to correspond with T_{sys} at the earliest time of the observation. As such, T_{sys} is estimated during a period far from the transit of Cygnus A and sunrise, which can add additional noise power. However, it is noted that Cygnus A is close to the Galactic plane, which may lead to a larger T_{sys} than if I had chosen to observe a source further from the Galactic plane³.

The measured T_{sys} is plotted in Figure 2.10 for each dish and polarization along with the CHORD target specification of 30 K. The sky and LNA noise temperatures are expected to increase towards lower frequencies, which is consistent with Figure 2.10. While T_{sys} as plotted in Figure 2.10 is generally much larger than the 30 K CHORD target, it is not

³The choice to observe Cygnus A was determined based on the available bright sources that transit in the spring, newly instrumented electronics that were not present during earlier observations, and the small collecting area of D3A, among other limiting factors.

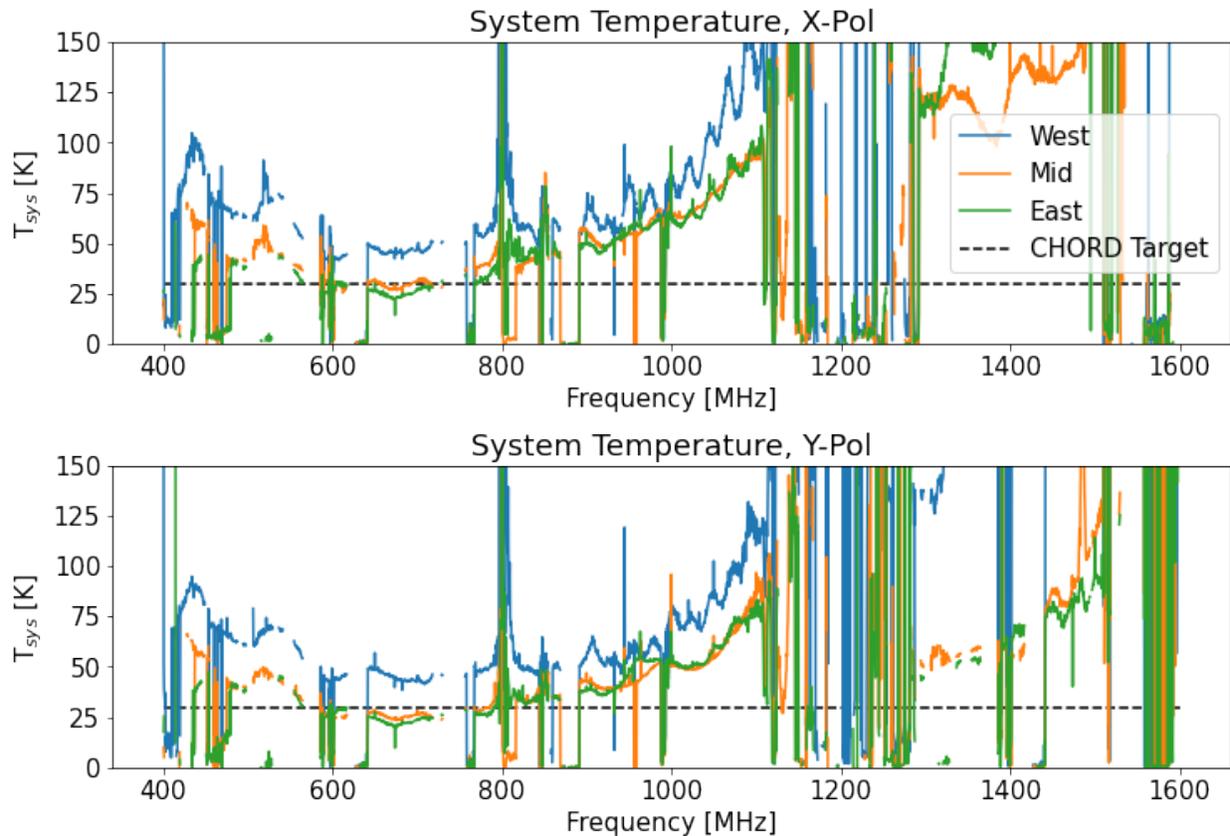


Figure 2.10: T_{sys} of D3A as a function of frequency compared with the CHORD target T_{sys} of 30 K. T_{sys} is roughly consistent with the simulated T_{sys} presented in MacKay et al. (2022) in the 400–800 MHz band. Aliased noise power drives T_{sys} near Nyquist frequencies. Most of the channels in the 800–1200 MHz band appear to be affected by aliasing, which results in increasing T_{sys} towards the 1200 MHz. The filters used in the two polarizations of the 1200–1600 MHz band attenuate the signal within the band towards higher frequencies, which likely explains the higher T_{sys} . T_{sys} in the 1200–1600 MHz band is lower by over 100 K for the Y-polarization than for the X-polarization due to additional amplifiers added in the signal paths for the Y-polarizations. The system temperature can be improved by implementing band-pass filters with sharper roll-off to mitigate aliasing, or by using a different digital system that can sample at a higher rate, and therefore does not require band-pass filtering.

necessarily representative of the T_{sys} expected for CHORD due to significant aliasing and in-band attenuation from the band-pass filters.

The presence of aliasing is most prominently observed as large spikes in T_{sys} for the frequency channels around the Nyquist frequencies 800 MHz, 1200 MHz, and 1600 MHz. Many channels in the 800–1200 MHz band appear to be corrupted by aliased noise power from higher frequencies, which drives up the system temperature. Indeed, the ZX75BP-1062-S+ filters used for the 800–1200 MHz band do not attenuate below 3 dB until 1320 MHz, or below 20 dB until 1535 MHz ([Mini-Circuits, b](#)). Therefore, T_{sys} grows towards the Nyquist frequency at 1200 MHz. ZX75BP-1307-S+ filters are used for the 1200–1600 MHz band of the X-polarization, which attenuate by 3.5 dB at 1520 MHz and by 20 dB at 1710 MHz ([Mini-Circuits, c](#)). ZX75BP-1350-S+ filters are used for the 1200–1600 MHz band of the Y-polarization, which attenuate by 3 dB at 1455 MHz and by 20 dB at 1572 MHz ([Mini-Circuits, d](#)). In both cases, some in-band signal power is attenuated by the filters, which likely causes increasing T_{sys} towards higher frequencies in each band.

Towards lower frequencies, T_{sys} approaches 30 K slightly after 600 MHz, which roughly corresponds with the simulated T_{sys} presented in [MacKay et al. \(2022\)](#). As such, with improved band-pass filters to mitigate aliasing and in-band attenuation, it is expected that T_{sys} will demonstrate improved performance. However, as will be discussed in the following chapter, if an alternative platform for the digital signal processing hardware at D3A is used, no band-pass filtering will be necessary, which will eliminate the challenges presented with

the current hardware.

T_{sys} is consistently higher for the West dish compared to the Mid and East dishes for nearly all frequency channels which corresponds to its higher forward gain. Recall that the inconsistency seen in the West dish in comparison to the other two dishes may be due to differences in the feed of the West dish. The most notable difference when comparing the two polarizations is seen in the 1200–1600 MHz band, where T_{sys} for the Y-polarization is nearly 100 K lower than the X-polarization. This is likely due to additional amplifiers added to the 1200-1600 MHz band for the Y-polarization, which serves to reduce the noise temperature contributions of “upstream” electronics.

2.3 Summary

In this chapter, I presented the Deep Dish Development Array (D3A), a pathfinder for the development of CHORD technologies. I then presented measurement of the system temperature T_{sys} of D3A by taking advantage of the phase closure of a three-element interferometer. While the measured T_{sys} can be greater by 100 K in the 800–1200 MHz and by over 150 K in the 1200–1600 MHz bands (depending upon the polarization and dish) compared with the CHORD target T_{sys} of 30 K, it is likely due to aliasing and in-band attenuation caused by the band-pass filters used for these bands. Towards lower frequencies, the T_{sys} behavior is roughly in agreement with that presented in [MacKay et al. \(2022\)](#), suggesting that improvements in the filtering at D3A is necessary for a more

accurate measurement of T_{sys} . However, the need for band-pass filtering can be completely eliminated by using an alternative digital signal processing platform, which will be described in the next chapter. It was also found that there are significant differences when comparing the performance of the West dish with the Mid and East dishes. However, it is noted that the West dish is instrumented with the first feed manufactured and assembled for D3A, and therefore may have slight differences compared with the feeds of the other two dishes. As such, D3A will focus on improving its redundancy and systematic control in order to meet CHORD requirements.

Chapter 3

Development of the CHORD

F-Engine

The CHORD F-Engine corresponds to the hardware, firmware, and software responsible for the digitization, channelization, and corner-turn operation of signals from the upstream analog system. It is inspired by the design of the CHIME F-Engine. The F-Engine must be capable of processing in real-time the colossal data rates required by CHORD, which is expected to be on the order of tens of terabits per second. The F-Engine must also be cost-effective, scalable, compact, and capable of science-grade digital signal processing. In this chapter, I present two possible designs for the CHORD F-Engine, and discuss the ADC performance of a potential alternative platform for the CHORD F-Engine.

3.1 CHORD F-Engine Designs

There are currently two potential designs for the CHORD F-Engine: 1) an ICE¹ system-based F-Engine, similar to that of CHIME (Bandura et al., 2016), and 2) an RF-System-on-Chip (RFSoc) based design. Table 3.1 presents a comparison of the two systems.

The ICE-based system would be similar in design to the current signal processing pipeline of the Deep Dish Development Array (D3A), a testbed for the development of CHORD technologies described in Chapter 2. For CHORD however, the ADC would be modified to sample at 1.2 GSPS rather than 0.8 GSPS, and use an analog diplexer to split the signal into two frequency bands that are separately digitized, similar to that described in Chapter 2. The first band would be filtered by a 600–1200 MHz bandstop filter, a 300 MHz high-pass filter, and a 1500 MHz low-pass filter, and would recover the 300–600 and 1200–1500 MHz regions of the full CHORD bandwidth through aliasing. The second would be filtered by a 600–1200 MHz bandpass filter and recover that frequency range, again through aliasing of the second Nyquist zone.

An ICE-based F-Engine offers the benefit of requiring minimal modifications to the hardware of an existing system that has already been demonstrated to process data at the level of scientific quality, as well as meeting the other criteria discussed at the beginning of this chapter. However, the necessity of band-splitting presents the challenge of maintaining filter stability for accurate removal of aliased power, which requires careful design of the

¹Note that ICE is not an acronym.

analog filters. As such, some frequency channels around the Nyquist frequencies will need to be discarded in such a design. The lost bandwidth will depend on filter stability and the precision required for each science case. In addition, the resolution of the ADC offers a limited dynamic range to prevent saturation due to increased RFI at the DRAO, which could lead to further discarded data.

The RFSoc-based system would be a novel design based on the AMD Zynq Ultrascale+ RFSoc Field Programmable Gate Array (FPGAs). The ADCs of the RFSocs developed by AMD have superior documented performance in comparison with the Teledyne EV8AQ160 ADCs used on the ICE ADC mezzanines ([AMD Inc., 2022a](#); [Teledyne, 2019](#)). Notably, the RFSocs are capable of sampling at greater than 3 GSPS, the Nyquist-Shannon sampling criterion for the CHORD bandwidth, using interleaved sampling. Interleaved sampling implements multiple ADC cores sampling at the same rate, but offset in phase from each other, such that the samples from each ADC can be combined to achieve a higher effective sampling rate. Stringent calibration on the phase, gain, and offset of the individual ADCs with respect to one another is crucial for accurate digitization of the input signal. A key aspect of the higher sampling rate offered by the RFSoc is that the band-pass filtering needed for the ICE-based F-Engine becomes unnecessary; a single anti-aliasing filter would be needed for each input. Therefore, there would be fewer components that would need to be characterized in the signal path, which leads to a smoother frequency response, a crucial requirement for 21-cm cosmology.

System	Advantage	Disadvantage
ICE	Lower risk (demonstrated system)	Requires band-splitting
RFSoc	Directly sample full CHORD band	Higher risk (novel system)

Table 3.1: The advantages and disadvantages of potential F-Engine designs for CHORD.

In addition, the latest-generation RFSocS can sample with a resolution up to 14-bits; this increases the dynamic range of the ADC, reducing the gain requirements of the analog system, helping to minimize non-linear behavior in the amplifiers, and reducing the possibility of saturation of the ADC from RFI. The RFSocS also offer increased FPGA resources, with approximately 1.25 times the block random access memory (BRAM) and 2 times the digital signal processing (DSP) resources for the ZU47DR model in comparison with the Kintex 7 420T FPGAs used on ICE boards ([AMD Inc., 2022b](#); [Xilinx Inc., 2020](#)). The ZU47DR also has an additional 22.5 Mb of high density “ultra”-RAM (URAM) which is not included in the Kintex 7 models. The additional resources, for example, can be applied to increased FFT length to improve frequency resolution, and increased number of polyphase filter bank (PFB) taps to reduce leakage into adjacent frequency bins.

Despite the wide benefits offered by an RFSoc-based system, the performance of the RFSoc ADCs must be verified independently to determine their suitability for the CHORD F-Engine. Further, a custom hardware, firmware, and software platform tailored to the requirements of a radio astrophysics and cosmology experiment is needed. In the next section, I will present measurements characterizing the ADC performance of a candidate RFSoc platform for CHORD.

3.2 Characterization of the ZU47DR RFSoc ADC

Measurement Setup

In this section, I characterize the performance of the dual ADC tile ZU47DR RFSoc, capable of simultaneously processing 8 ADC channels at up to 5 GSPS with 14 bits of resolution. I use a t_0 Control and Readout System (CRS) implemented with a ZU47DR as a representative potential RFSoc platform for CHORD. The CRS is a printed circuit board (PCB) designed for the next generation of precision signal processing hardware for radio cosmology and astrophysics. The CRS board is displayed in Figure 3.1, with the ZU47DR placed slightly left-of-center under a black heat sink, and the analog inputs along the bottom right, where passive components are contained beneath RF shields. I will be sampling at $f_s = 3$ GSPS, and characterize the ADC performance using channel ADC1, whose input is the furthest right on the bottom row of SMA connectors in Figure 3.1. I will refer to ADC1 as channel 0 or CH0, with a similar indexing for other ADC channels.

Figure 3.2 displays a diagram of the measurement setup. To generate RF tones, I use a RIGOL DSG836 signal generator (RIGOL Technologies Inc., 2019) for each test, and synchronize the clocks of the signal generator and the CRS board. In certain cases where a second signal generator is needed, I use a RIGOL DSG821 to generate an additional tone. An SLP-1650+ DC-1.4 GHz low-pass filter (LPF) is used for Figures 3.4 to 3.10 in order to reduce high-frequency noise and harmonics from aliasing into the first Nyquist zone (0-

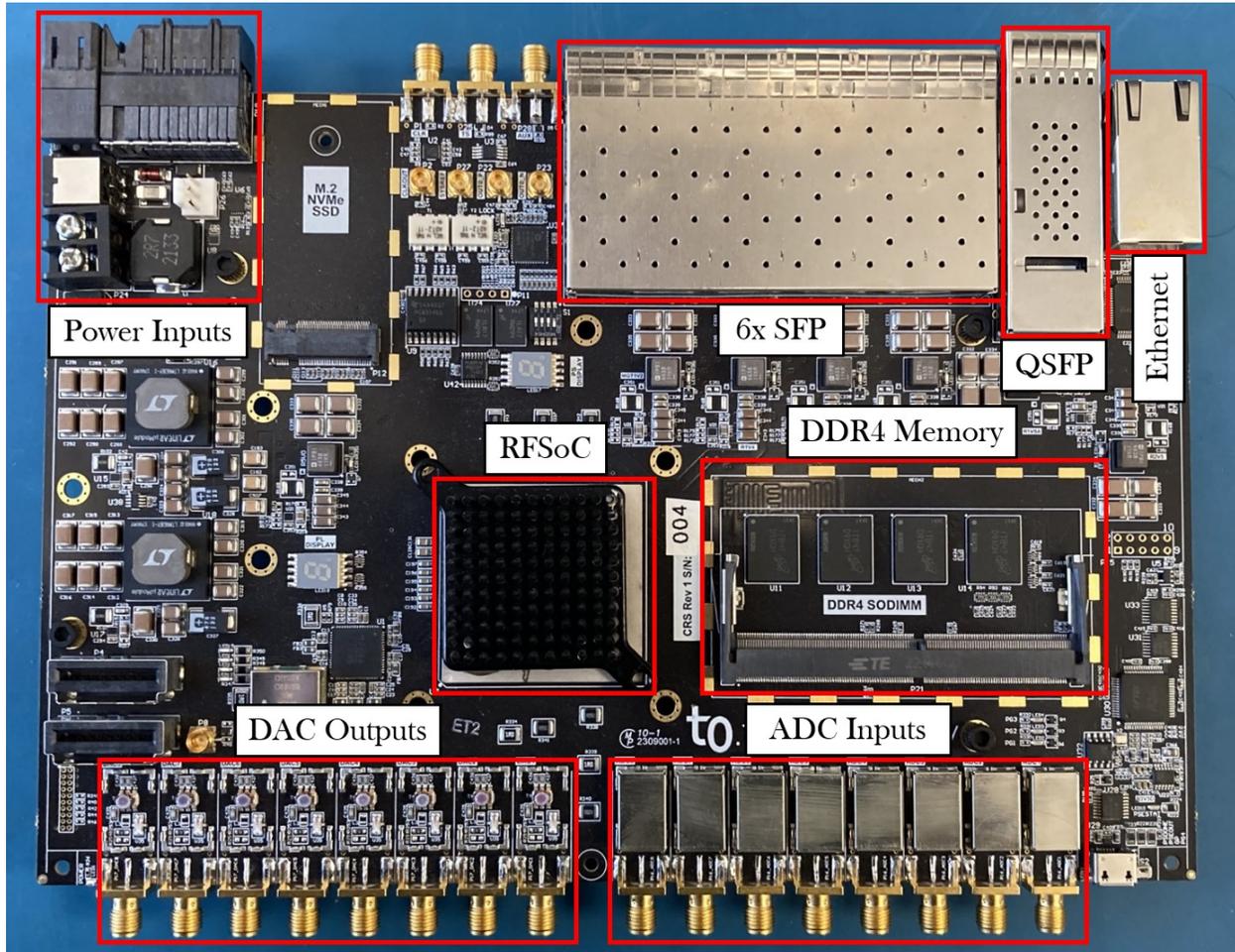


Figure 3.1: The t_0 CRS board, Revision 1, SN004, equipped with a ZU47DR RFSoc. RF shields have been added around components along the signal path of the ADC inputs, which includes an LFCN-9170+ DC–9.17 GHz LPF, an MABA-011118 balun, and other passive components. Power connections are adaptable for benchtop setups and backplane connections. An ethernet connector enables communication with the ZU47DR. A cage of six 25G small form-factor pluggable (SFP) connectors and a single 100G Quad-SFP (QSFP) connector can be used for data offloading.

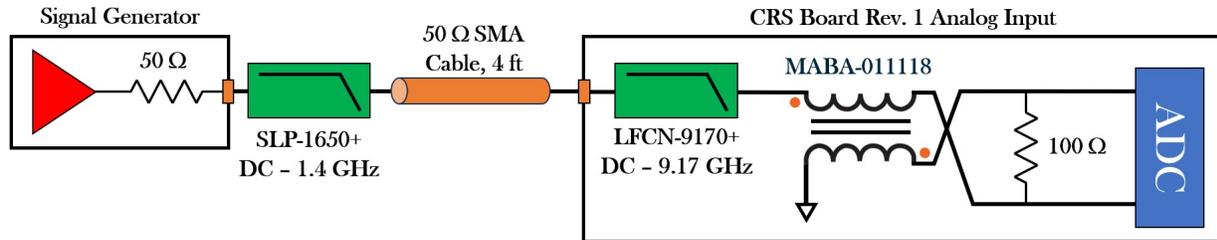


Figure 3.2: A simplified diagram of the signal path for measurements of the ADC performance of the RFSoc. Note that the SLP-1650+ LPF is only used for the measurements in Figures 3.4 to 3.10. Changes to the measurement setup for Figures 3.11 to 3.14 are clarified in the discussion corresponding to each figure.

1500 MHz). The signal generator has a typical level uncertainty of 0.5 dB in the power of the carrier tone. The second-order harmonics of the signal generator range from approximately -35 to -60 dBc, while the third-order harmonics range from approximately -45 to -75 dBc.

A 50 Ω SMA cable is used to connect the output of the signal generator (or SLP-1650+ DC-1400 MHz LPF) to the analog SMA input of the CRS board. The signal passes through an LFCN-9170+ DC-9.17 GHz LPF ([Mini-Circuits, a](#)) before being converted from a 50 Ω single-ended signal to a 100 Ω differential signal using a MABA-011118 balun ([MACOM Technology Solutions Inc.](#)). The differential signal is then fed to ADC. I capture data using `pychfpga` ([Bandura et al., 2016](#)), the Python-based control software used for the ICE system, which has been upgraded by Dr. Jean-François Cliche for use with RFSocs. Data is captured in “frames”, which is defined below along with other definitions that will be used throughout this chapter:

- **Frame:** A frame is defined to be $M = 16,384$ ADC samples. For a single spectrum,

I capture 96 frames and compute the FFT power for each frame individually using `numpy`, and then integrate all 96 spectra to approach a flat noise floor. The resulting integrated spectrum then has $M/2 = 8192$ frequency bins, with a frequency resolution f_{res} given by

$$f_{\text{res}} = \frac{f_s}{M} = \frac{3 \text{ GSPS}}{2^{14}} = 183.105 \text{ kHz.} \quad (3.1)$$

- **Noise Floor P_{nf} :** The noise floor power P_{nf} (or simply noise floor) is the median value of an integrated spectrum, which for a spectrum containing a carrier frequency and noise only provides an estimation of the noise power per frequency bin. The median is taken as opposed to the average because the average value is biased by stronger tones that can artificially increase the noise floor estimate. If a frequency bin is not identified as a peak (see below), it is considered a noise bin, and the power in that bin is assumed to be equal to the noise floor.
- **Total Noise P_{n} :** The total noise power P_{n} (or total noise) is the sum of the power of all noise bins. The total noise can equivalently be expressed as the undoing of the processing gain introduced by the finite number of frequency bins from the FFT, and may be determined from the noise floor P_{nf} as ([Kester and Bryant, 2005](#))

$$P_{\text{n}} = 10 \log_{10} \left(\frac{P_{\text{nf}}}{M/2} \right) = P_{\text{nf}} \text{ (dBFS)} + 10 \log_{10} \left(\frac{M}{2} \right). \quad (3.2)$$

The magnitude of P_n is equivalent to the SNR for an ADC dominated by quantization noise. Note that the above equation assumes that all frequency bins correspond to noise bins, while a real ADC will have power in the carrier tone, harmonics, distortion products, and other spurious tones. However, the majority of frequency bins correspond to noise bins for a digitized sine wave, so the total noise may be considered an upper limit.

- **Peak:** A peak in a given integrated spectrum is defined as any frequency bin with power $P > P_{\text{nf}} + T$, where T is some threshold, typically defined to be 2 dB. Note that the choice of T is arbitrary and that the number of identified peaks is dependent on T . The value of T is chosen to be low enough to capture as many spurs as possible while ensuring noise bins are not mistaken for peaks. A peak can be the carrier tone, harmonics, interleaving spurs, spectral leakage bins, or unidentified spurs.
- **Full-scale:** The full-scale power P_{fs} is calibrated by injecting a sine wave into an ADC input and checking that the maximum and minimum peaks of wave exercise nearly all 2^{14} digital codes. In practice, a wave that is considered full-scale is one that approaches -8192 to 8191 LSBs, and might be from 50 to 100 LSBs from true full-scale in order to avoid saturation of the signal. I find that full-scale can be achieved with sine wave with $f_{\text{cal}} = 240.05126953125$ MHz at 5.6 dBm fed into a 50Ω SMA cable. This frequency is chosen to be as close as possible to the only in-band carrier frequency quoted by the datasheet (240 MHz, the next being 1.9 GHz), and such that an integer number

of cycles fits within a single frame (avoiding the need for windowing). Note that I quote the full-scale power from the perspective of the signal generator. The datasheet states that full-scale corresponds to $1 V_{PPD}$ at 100Ω at the ADC input, equivalent to 1 dBm (note that the ADC input on the ZU47DR is a differential input) (AMD Inc., 2022a). Approximately 4.6 dBm is lost from the signal generator. 1.7 dBm is lost from the signal generator output to the ADC input due to the SMA cable, the low-pass filter (LPF) used in Figures 3.4 to 3.10, the balun, and other analog components. The other 2.9 dBm is lost internally, possibly due to firmware-controlled attenuation. Note that the full-scale power increases slightly with frequency (around 1–2 dB), but that I always calibrate at f_{cal} . The power spectra are recorded in dBFS units, which corresponds to a ratio between the power at a given frequency to the power in the carrier tone; 0 dBFS is equivalent to full-scale power, and all other values are negative.

The ZU47DR datasheet measures most ADC metrics with carrier tones at -1 dBFS, and quotes these values in dBc units, which corresponds to a ratio between the carrier tone power and some other signal. However, at and near full-scale, the signal generator contributes significant noise power, increasing the noise floor by nearly 15 dB, as well as introduces non-uniform noise floor shape. As such, the signal generator dominates over the ADC noise when injecting full-scale signals at -1 dBFS. In order to avoid biasing the measurements presented by the signal generator noise, I will instead inject carrier tones at -35 dBFS such that the noise is dominated by the ADC rather than the signal generator. While it is possible that

certain spurs or harmonics may grow with increased carrier power, many signal processing systems in radio astronomy maintain low-power input signals to ADCs in order to prevent RFI from saturating the ADC. Therefore, the performance metrics measured with carrier tones at -35 dBFS is more representative of the expected performance of the RFSoc for radio astronomy applications.

I convert datasheet values quoted in dBc to dBFS such that a direct comparison can be made between my measurements and the datasheet (as the low-power carrier will cause measurements made in dBc to be much smaller than datasheet values). Note that this typically only changes values quoted in dBc by 1 dB, since most datasheet values were measured with a -1 dBFS carrier tone (e.g., 74 dBc is equivalent to -75 dBFS in this case). In addition, the datasheet only quotes values for $f_c = 240$ MHz within the first Nyquist zone. For datasheet values presented, if the quoted value is approximately equivalent between the first two carrier frequencies measured (240 MHz and 1.9 GHz), I assume that the value is constant across the first Nyquist zone. Otherwise, I only quote the value at 240 MHz in comparison with my own measurements. $f_c \sim 240$ MHz is always measured in order to provide a direct comparison in such cases.

The datasheet suggests that the RFSoc ADC has superior performance compared with EV8AQ160 ADCs used in the ICE system. Therefore, the measured ADC performance should be consistent with the datasheet, and demonstrate improvements in comparison to the EV8AQ160. In this section, I only compare the measured performance with the datasheet,

and then compare with the EV8AQ160 in Section 3.3.

Time-Domain Measurements

In the upper panel of Figure 3.3, I present a single frame of a raw ADC timestream of a “slow” sine wave with carrier frequency $f_c = 3000.183105$ MHz at -35 dBFS. Note that the DC–1400 MHz LPF is not present in this measurement. This carrier frequency is equivalent to $f_s + f_{\text{res}}$, meaning that the carrier tone shows up in the spectrum at $f_{\text{res}} = 0.183105$ MHz due to aliasing (Oppenheim and Schaffer, 2010). A higher-frequency carrier tone must be aliased rather than simply injecting a carrier tone at f_{res} because the ADC input is AC-coupled, meaning it behaves as a low pass filter, attenuating the power at f_{res} severely. This frequency is selected since a single period of the wave fits within a single frame, allowing me to easily observe the noise of the sine wave.

I then fit this quantized sine wave to an ideal floating-point sine wave, and subtract the raw data from the best fit in order to observe the residuals, presented in the lower-left panel. I then compute a histogram of these residuals, which is plotted in the lower-right panel. I fit a Gaussian to the histogram, and find mean $\mu = -0.06$ least significant bits (LSBs, also referred to as digital codes) and standard deviation $\sigma = 5.71$ least significant bits. One estimate the full-scale signal-to-noise ratio (SNR) from the standard deviation (which is equivalent to the RMS noise power) according to (Kester and Bryant, 2005)

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_{\text{fs}}^2}{\sigma^2} \right) = 20 \log_{10} \left(\frac{2^{n-1}}{\sigma\sqrt{2}} \right) = 20 \log_{10} \left(\frac{2^{13}}{5.71\sqrt{2}} \right) = 60.1 \text{ |dBFS|}, \quad (3.3)$$

where $\sigma_{\text{fs}} = 2^{n-1}/\sqrt{2}$ is the RMS power for a full-scale sine wave with an amplitude exercising all 2^n digital codes, dependent only upon the ADC resolution $n = 14$ bits. This expression for the SNR assumes that RMS noise power is dominated by quantization noise only. Note that I quote the SNR in absolute dBFS, or |dBFS|, to distinguish from dBFS (which is typically a negative value).

Signal-to-Noise Ratio (SNR) and Noise Spectral Density (NSD)

One can also determine the SNR directly by measuring the power spectrum. Figure 3.4 presents an example power spectrum of a 240.05 MHz carrier tone at approximately -35 dBFS. The carrier tone, harmonics, interleaving spurs (defined below), and datasheet noise floor are displayed, which is consistent with the measured noise floor. I check if any of the identified peaks as defined earlier occur at the first 100 harmonics of the carrier tone. The gain-time interleaving spurs (GTIS) caused by the gain-time (or gain-phase) calibration occur at frequencies f_{GTIS} defined by (AMD Inc., 2022a)

$$f_{\text{GTIS}} = K \left(\frac{f_s}{8} \right) \pm f_c \text{ for } K = 0, 1, \dots, 8. \quad (3.4)$$

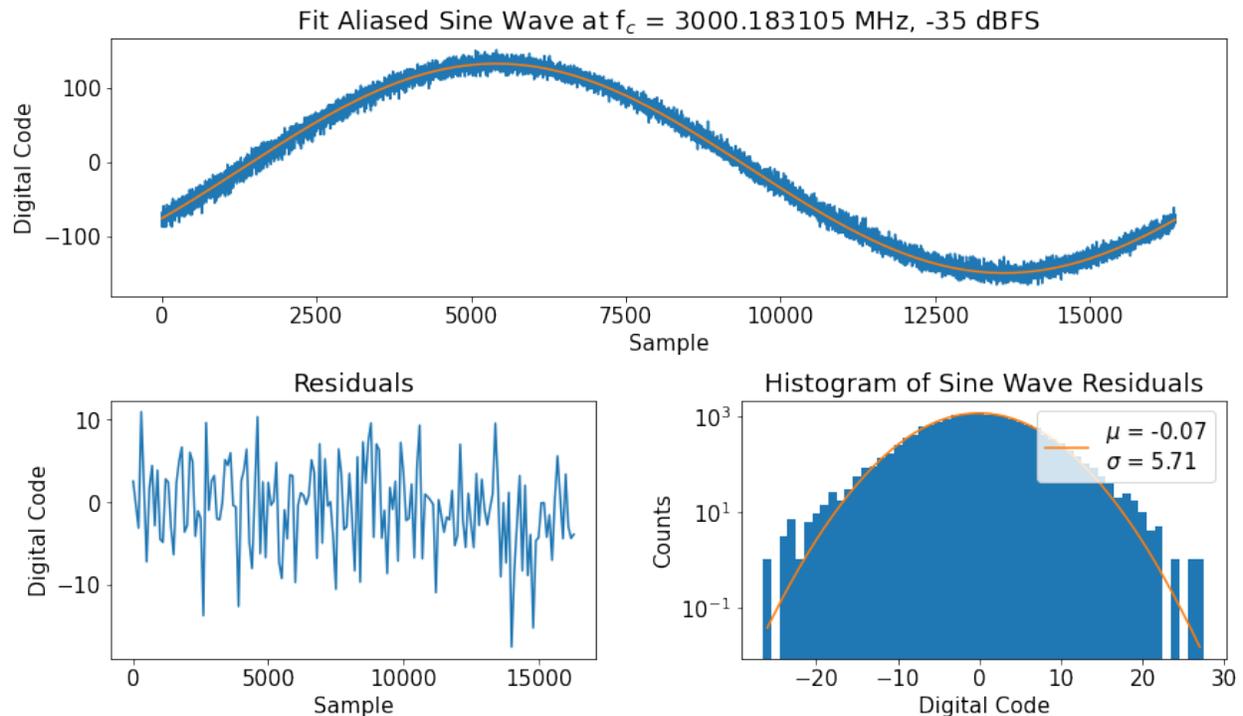


Figure 3.3: An example of a single frame of data digitized by the RFSoc. The top panel displays the captured data in blue, which represents a sine wave at -35 dBFS with $f_c = 3000.183105$ MHz, corresponding to $f_s + f_{\text{res}}$, which is aliased to $f_{\text{res}} = 0.183105$ MHz such that one full cycle fits within a single frame. Also shown in the top panel in orange is the best-fit floating point sine wave. The residuals between the digitized data and best-fit sine wave are shown in the bottom-left panel. Note that every 100^{th} residual is shown to allow the variation in the distribution to be seen. A histogram of these residuals is shown in the bottom-right panel, and a Gaussian function is fit to the distribution of counts. The residuals are Gaussian-distributed out to approximately 3σ . From the standard deviation σ of the best-fit Gaussian function, or equivalently the RMS noise power of the digitized sine wave, the SNR can be estimated according to Equation 3.3. I find $\text{SNR} = 60.1$ [dBFS], which is consistent with direct measurements of the SNR from the spectra and the datasheet

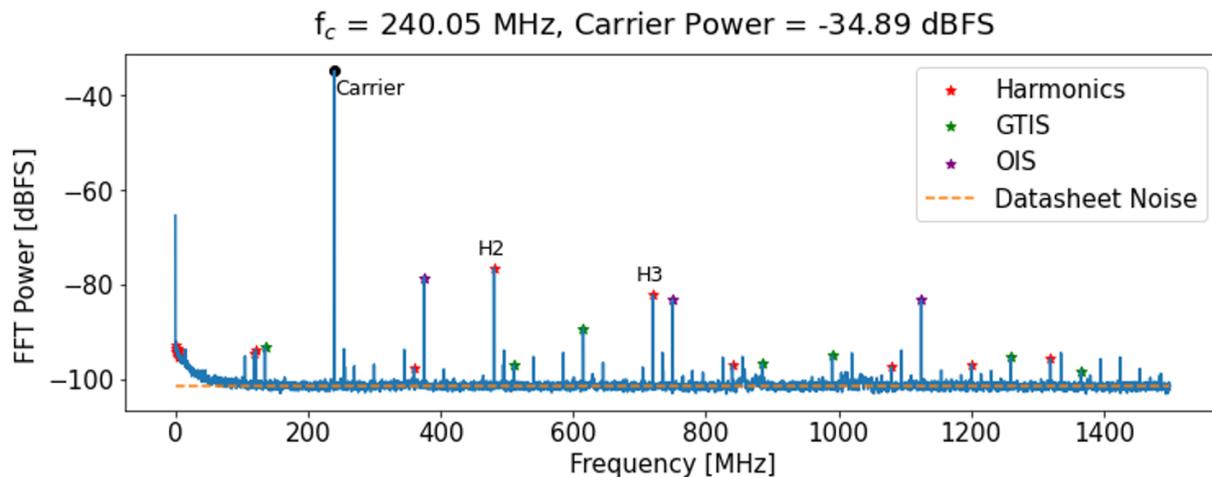


Figure 3.4: An example of the RFSoc FFT power spectrum with $f_c = 240.05$ MHz at approximately -35 dBFS. The carrier, harmonics of the carrier (the first 100 harmonics are checked), gain-time interleaving spurs (GTIS) defined by Equation 3.4, and offset interleaving spurs (OIS) defined by Equation 3.5 are identified. The datasheet noise is converted from the NSD according to the sampling rate f_s and the number of points per frame M , which agrees well with the FFT noise floor. Measurements of the SNR, the power in the harmonics, and the power in the interleaving spurs is detailed in subsequent figures.

Likewise, the offset interleaving spurs (OIS) caused by the offset calibration occur at frequencies f_{OIS} defined by (AMD Inc., 2022a)

$$f_{\text{OIS}} = K \left(\frac{f_s}{8} \right) \text{ for } K = 0, 1, \dots, 8. \quad (3.5)$$

Note that most harmonics and some interleaving spur frequencies are aliased back into the first Nyquist zone.

The full-scale signal-to-noise ratio (SNR) of an ADC is the ratio between the full-scale

power P_{fs} and the total noise P_{n} measured from the power spectrum, given by (Xilinx Inc., 2019)

$$\text{SNR} = 10 \log_{10} \left(\frac{P_{\text{fs}}}{P_{\text{n}}} \right) = 0 - P_{\text{n}} \text{ (dBFS)}, \quad (3.6)$$

where, by definition, $P_{\text{fs}} = 0$ dBFS. I find that $\text{SNR} = 62.3$ |dBFS| is the average value for carrier frequencies in the first Nyquist zone (0–1500 MHz). Another metric for quantifying the noise is the noise spectral density (NSD), or the noise power contained in a 1 Hz frequency bin. The NSD is given by (Xilinx Inc., 2019)

$$\text{NSD} = 10 \log_{10} \left(\frac{P_{\text{n}}}{f_{\text{s}}/2} \right) = P_{\text{n}} \text{ (dBFS)} - 10 \log_{10} \left(\frac{f_{\text{s}}}{2} \right). \quad (3.7)$$

I find that the average $\text{NSD} = -154$ dBFS/Hz for carrier frequencies in the first Nyquist zone. The SNR and NSD as a function of f_{c} are displayed in Figure 3.5, which demonstrate consistency with the datasheet values (AMD Inc., 2022a). Note that the datasheet only quotes the NSD, which I have converted to full-scale SNR. The SNR and NSD provide an estimate of the dynamic range of the ADC assuming that the SNR is noise-limited only.

Second- and Third-Order Harmonic Distortion

However, as seen in Figure 3.4, the dynamic range is also dependent upon the power in harmonics and spurs. The power in the x^{th} harmonic H_x is given by

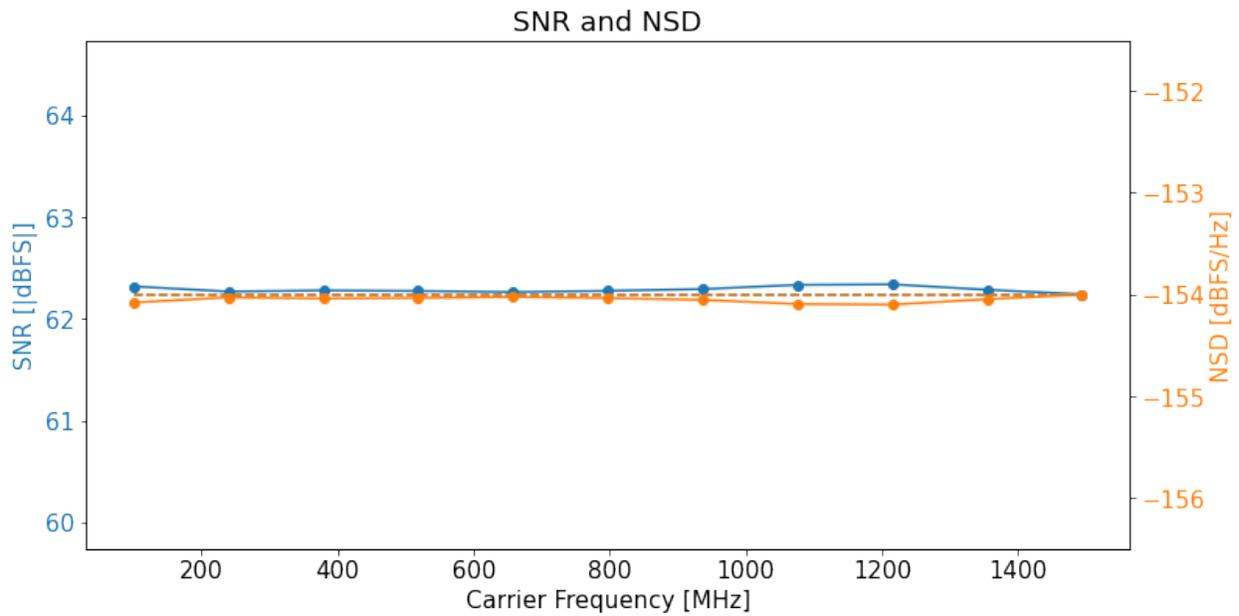


Figure 3.5: The SNR and NSD of the RFSoc as a function of carrier frequency, with average values of 62.3 |dBFS| and -154 dBFS/Hz, respectively. Datasheet values for each metric are shown as dashed lines in the color of that metric. Note that the SNR is quoted in absolute dBFS, or |dBFS|, to distinguish from dBFS. The datasheet SNR is not explicitly quoted; instead, I calculate P_n from the datasheet NSD and use Equation 3.6 to compute the SNR. Both the SNR and NSD vary by a tiny amount with frequency and are consistent with the datasheet.

$$H_x = 10 \log_{10} \left(\frac{P_{hx}}{P_{fs}} \right) = P_{hx} \text{ (dBFS)}, \quad (3.8)$$

which can be contributed to by both the signal generator and the ADC. The power in the second- and third-order harmonics H2 and H3, respectively, are presented as a function of f_c in Figure 3.6, as well as datasheet values at 240 MHz. I find that H2 ranges from -73.3 to -95.2 dBFS while H3 ranges from -82.6 to -99 dBFS. Note that this measurement was made with a DC–1400 MHz LPF between the signal generator and the analog input to suppress high-frequency (> 1400 MHz) harmonics generated by the signal generator. As a result, I measure $H2 > H3$ at lower f_c in contrast with the datasheet because the second-order harmonics of the signal generator are greater than the third-order harmonics (RIGOL Technologies Inc., 2019), and signal generator harmonics for low f_c are not suppressed by the LPF. This means that the H2 and H3 values measured at higher frequencies are more representative of the harmonics generated by the ADC, since any harmonics generated by the signal generator at those frequencies are suppressed by the LPF. H2 and H3 are thus consistent with the datasheet.

Total Harmonic Distortion (THD)

The harmonics produced by the ADC are not limited to second- and third-order, as seen in Figure 3.4. The total power in the harmonics P_h is the RMS sum (or root-sum-square) of the individual harmonic components, which may be written as (Kester and Bryant, 2005)

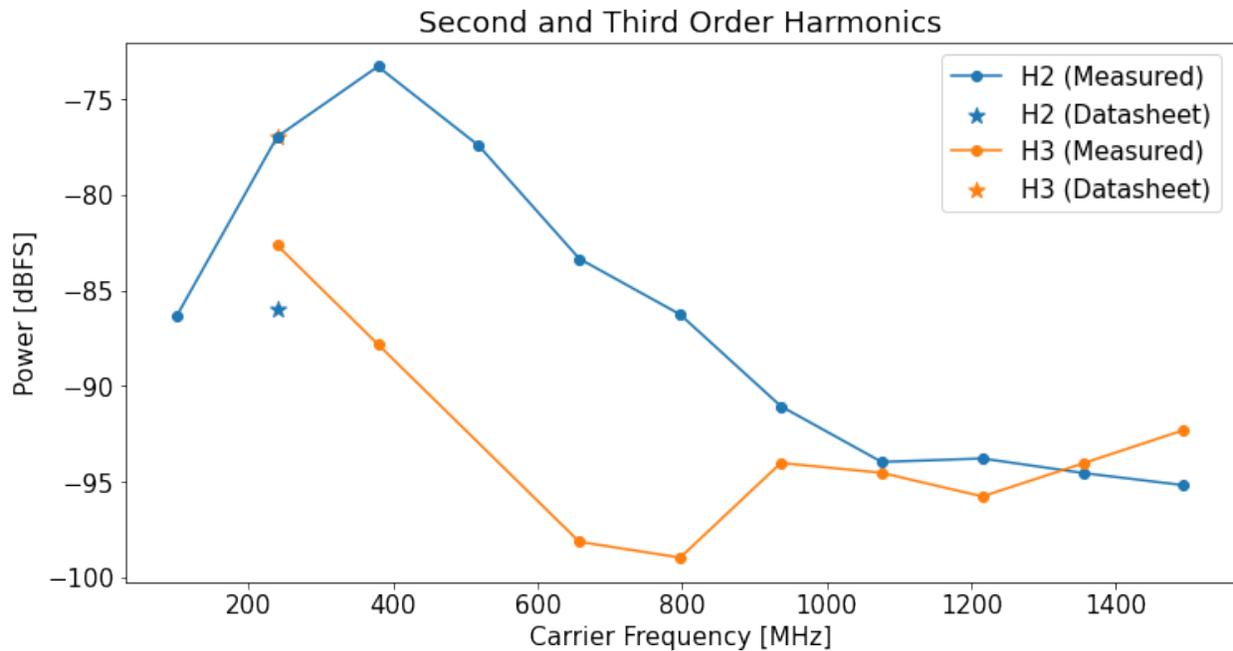


Figure 3.6: The power in the second- and third-order harmonics (H2 and H3) of the RFSoc as a function of carrier frequency. Measured values are shown as circles while datasheet values at 240 MHz are shown as stars (note that the datasheet H3 is slightly obscured by the measured H2). The signal generator has greater power in the second-order harmonics than the third-order harmonics, which dominate towards lower f_c because they are not suppressed by the DC–1400 MHz LPF. The values of H2 and H3 towards higher f_c are thus more representative of the harmonics produced by the ADC, and are consistent or improved in comparison with datasheet values.

$$P_h = \sqrt{P_{h2}^2 + P_{h2}^2 + \cdots + P_{hN}^2} \quad (3.9)$$

for $N-1$ total harmonics. The total harmonic distortion (THD) is then the ratio of P_h to the full-scale power, given by (Kester and Bryant, 2005)

$$\text{THD} = 10 \log_{10} \left(\frac{P_h}{P_{fs}} \right) = P_h \text{ (dBFS)}. \quad (3.10)$$

In Figure 3.7, I display the measured THD, which ranges from -73.3 to -89.7 dBFS. The measured THD implies that the total power of all identified harmonics is less than that of the total noise, which is approximately -62.3 dBFS according to the SNR and NSD measurements presented in Figure 3.5. However, individual harmonics are still more powerful than the noise in a given frequency bin. To mitigate the additional power from the harmonics, CHORD can maintain a signal power that is greater than the power in the largest harmonic. Also recall that the power in the second- and third-order harmonics in these measurements is dominated by the signal generator.

Signal-to-Noise-and-Distortion Ratio (SINAD) and Effective Number of Bits (ENOB)

As such, the the total power in the harmonics is key to measuring the signal-to-noise-and-distortion ratio (SINAD), which is the ratio of the total distortion power P_d plus the total

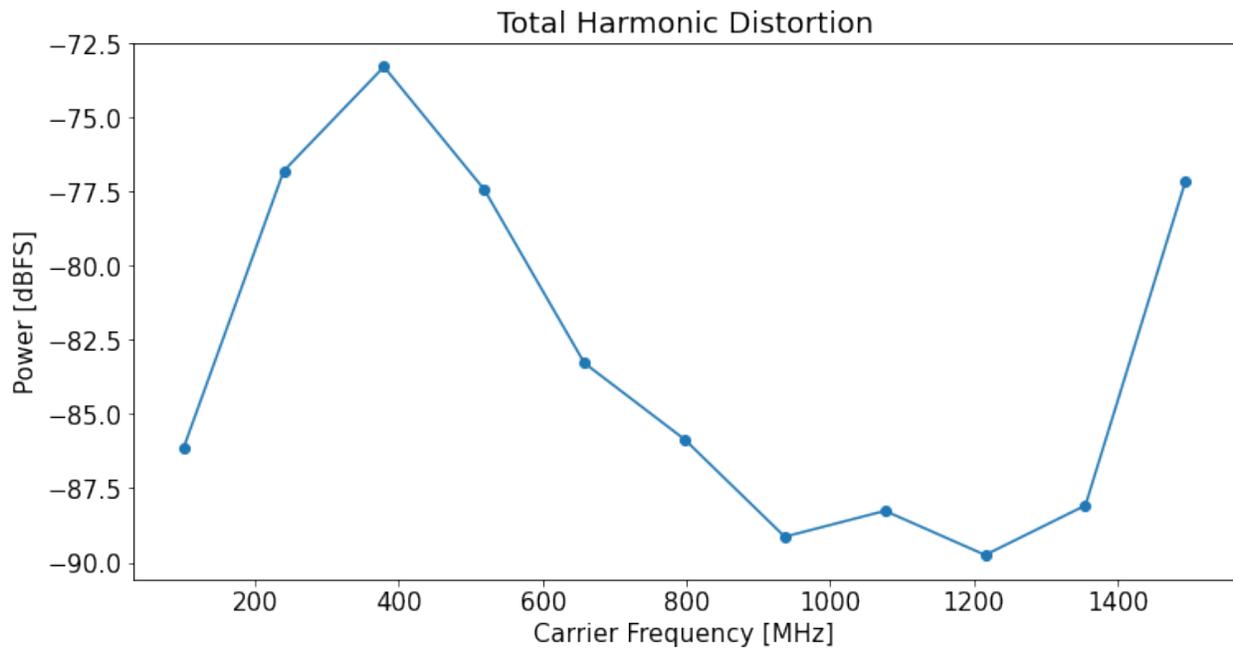


Figure 3.7: The THD of the RFSoc as a function of carrier frequency. At lower frequencies, the second- and third- order harmonics from the signal generator are not suppressed by the LPF; higher-frequency values of the THD are thus more representative of the harmonic distortion produced by the ADC. Note that the high THD at the far right is measured close to $f_s/2 = 1500$ MHz. The THD values presented here suggest that the ADC is dominated by the noise power rather than harmonic distortion power.

noise power P_n to the full-scale power, given by (Xilinx Inc., 2019)

$$\text{SINAD} = 10 \log_{10} \left(\frac{P_{fs}}{P_n + P_d} \right). \quad (3.11)$$

The SINAD is thus the dynamic range of the ADC considering both noise and harmonic distortion. For a typical ADC, one can make the approximation that all distortion products are contained in the harmonics, or $P_d = P_h$. However, since the interleaving of multiple ADCs produces additional distortion, the power in interleaving spurs must be considered in the SINAD, which show up at the frequencies defined by Equations 3.4 and 3.5. If it is assumed that all other distortion products are time-varying RFI, the total distortion power P_d may be expressed as the RMS sum of the power in the harmonics and all identified interleaving spurs.

The SINAD can be equivalently expressed by the effective number of bits (ENOB), given by (Xilinx Inc., 2019)

$$\text{ENOB} = \frac{\text{SINAD} - 1.76}{6.02}, \quad (3.12)$$

which is derived from the expression of the SNR for an ideal ADC, which is dependent only upon the ADC resolution. The ENOB may be interpreted as the ADC resolution required to produce the measured SINAD assuming that the ADC is only dominated by quantization noise.

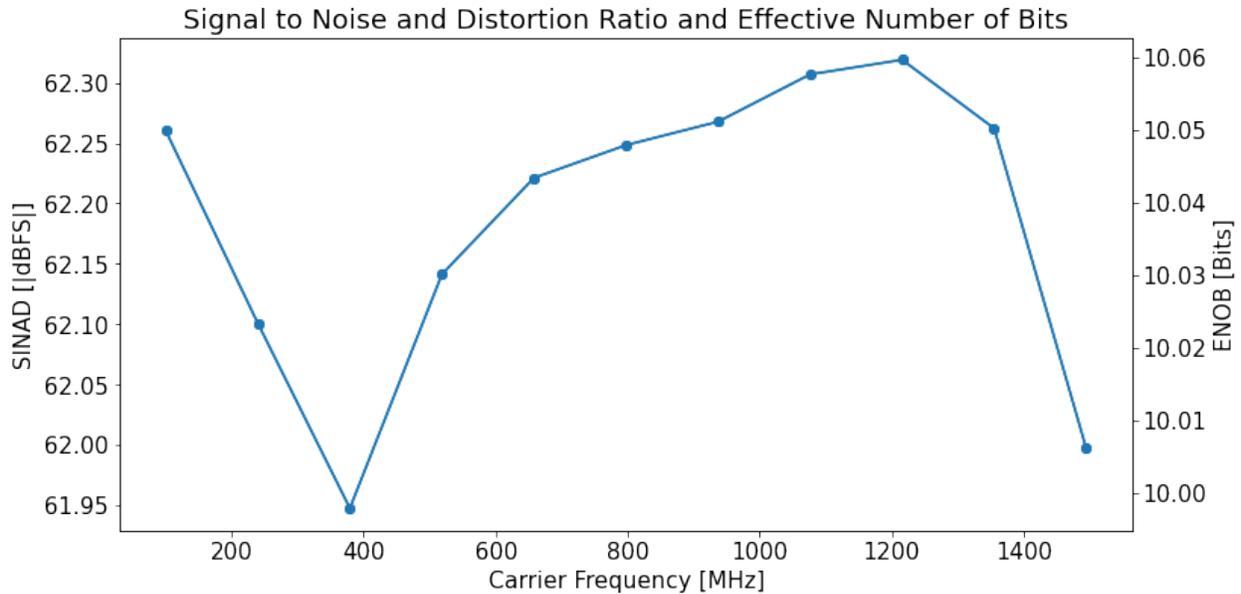


Figure 3.8: The SINAD and ENOB of the RFSoc as a function of carrier frequency. The SINAD is found to be nearly equivalent to the measured SNR, indicating that the total noise outweighs the combined total distortion power in the harmonics plus the interleaving spurs. As such, the ENOB agrees with that implied by the datasheet, which is about 10 bits if the SNR is computed from the NSD quoted in the datasheet, and then replace the SINAD with the datasheet SNR in Equation 3.12 (which also assumes that the quoted value is not dominated by distortion).

The measured SINAD and the ENOB in the first Nyquist zone are displayed in Figure 3.8. I find that the average SINAD is 62.2 [dBFS], and the average ENOB is 10 bits. The ENOB is approximately 4 bits less than the ADC resolution, which is 14 bits. The SNR computed from the NSD in the datasheet is approximately 62.2 [dBFS], meaning that the SINAD and ENOB are both consistent with the implied datasheet SNR and ENOB.

Spurious-Free Dynamic Range (SFDR)

The spurious-free dynamic range (SFDR) is the ratio of the power in the largest spur to the full-scale power. The SFDR is important in characterizing the effective dynamic range of the ADC, as the largest spur does not necessarily occur at the same frequency as a function of f_c and time. Therefore, the SFDR is representative of the highest additional power that may be added to a particular frequency bin. Figure 3.9 displays the SFDR as a function of f_c , both including and excluding H2 and H3. At lower frequencies, the SFDR is dominated by either H2 or H3. At higher frequencies, a different spur defines the SFDR, which is likely related to the LPF suppressing higher-frequency harmonics for high f_c . The SFDR excluding H2 and H3 ranges from -78.5 to -89.5 dBFS, or an average value of -86.3 dBFS, which is consistent with the datasheet.

Interleaving Spurs

The ADC implements an active calibration of the gain, phase (time), and offset of each ADC core such that their combined samples behaves effectively like a single ADC sampling at a higher frequency. The active calibration results in “interleaving spurs” that occur at defined frequencies which add power to specific frequency bins. The frequencies f_{GTIS} corresponding with gain-time interleaving spurs (GTIS) caused by the gain-time calibration are defined by Equation 3.4. The frequencies f_{OIS} corresponding with offset interleaving spurs (OIS) caused by the offset calibration are defined by Equation 3.5.

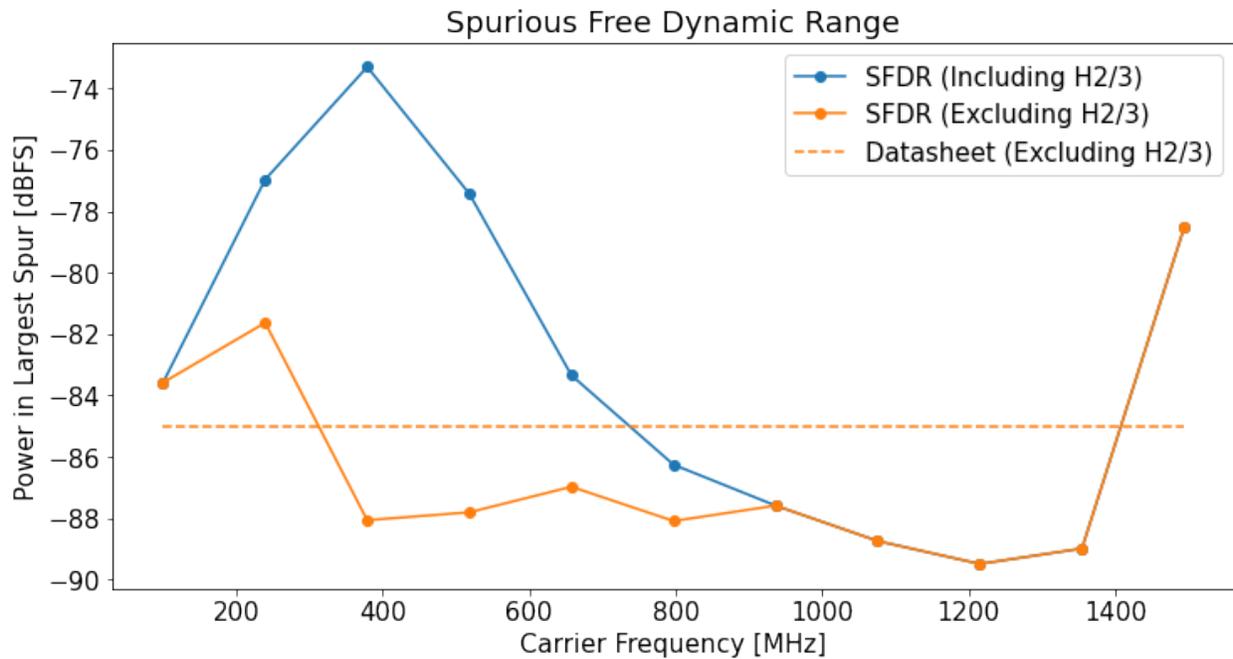


Figure 3.9: The SFDR of the RFSoc as a function of carrier frequency, including H2 and H3 in blue and excluding H2 and H3 in orange. Towards higher frequencies, the SFDR in both cases is set by an interleaving spur rather than H2 or H3, while at lower frequencies, the SFDR is typically equivalent to H2 because of the large harmonics from the signal generator. The SFDR excluding the second- and third-order harmonics is consistent with the datasheet.

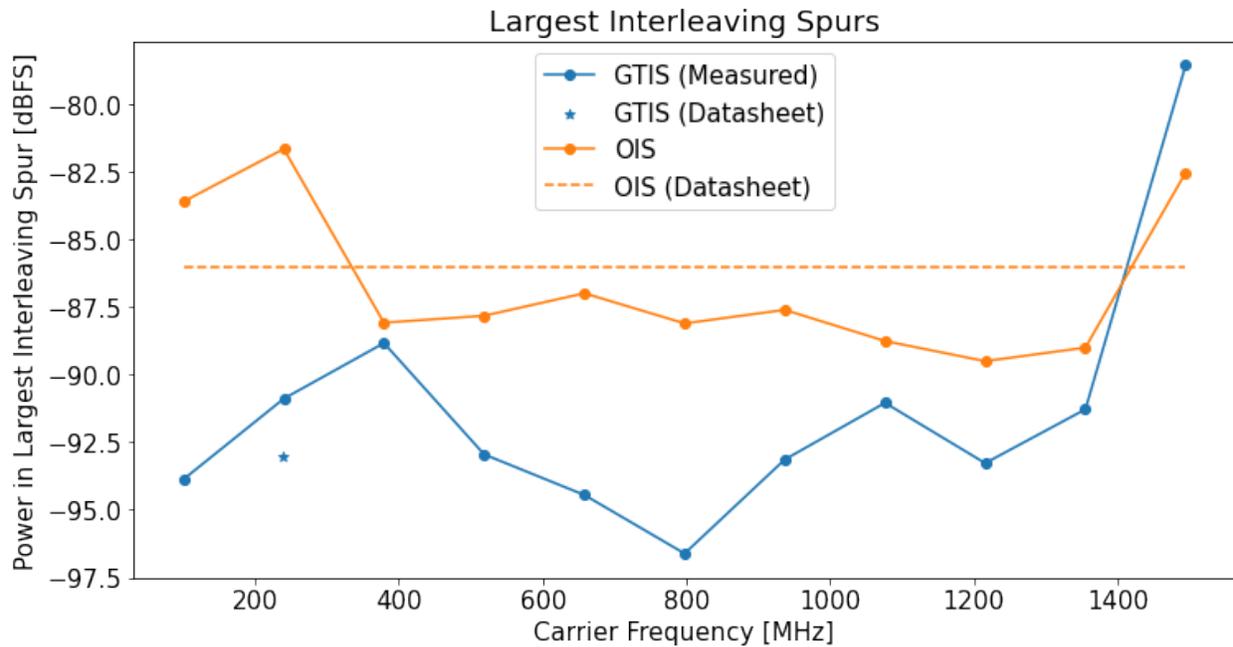


Figure 3.10: The maximum GTIS and OIS of the RFSoc as a function of carrier frequency. The average GTIS is -91.3 dBFS while the average OIS is -86.7 dBFS, which are both consistent with the datasheet.

The power at frequencies f_{GTIS} and f_{OIS} are recorded, and the largest GTIS and OIS are presented in Figure 3.10 as a function of f_c along with datasheet values at 240 MHz. I find that the GTIS range from -78.5 to -96.6 dBFS, or an average of -91.3 dBFS, while the OIS range from -81.6 to -89.5 dBFS, or an average of -86.7 dBFS. While the power in the largest interleaving spurs are greater for the measured values than the datasheet values at 240 MHz, the average GTIS and OIS are consistent with the datasheet.

Third-Order Intermodulation Distortion (IM3)

The third-order intermodulation distortion (IM3) is a measurement of the power in the tones that arise from the mixing of two carrier tones injected into the same ADC. A diagram demonstrating intermodulation distortion is displayed in Figure A.1 in Appendix A. Intermodulation distortion derives from the nonlinearity in the transfer function of a real ADC, which can be approximated as a polynomial series. If the input to an ADC is the sum of two sine waves with frequencies f_1 and f_2 , it can be shown using trigonometric identities that the third-order component of the polynomial transfer function produces additional tones in the power spectrum at $2f_1 - f_2$ and $2f_2 - f_1$. As shown in Figure A.1, other intermodulation distortion products are generated as well, though the third-order intermodulation distortion products are those typically quoted in datasheets. Intermodulation distortion products are of particular concern in radio astronomy, as two RFI tones with small Δf can mix and add distortion power to other frequency bins.

To conduct this measurement, I use the DSG836 signal generator and an additional RIGOL DSG821 signal generator to generate the second tone. The two signals are combined using a Z99SC-62-S+ RF splitter/combiner, which has ~ 22.5 dB of isolation at around 240 MHz. The splitter/combiner serves to both combine the two tones as well as prevent leakage from one signal generator to another, which can cause additional intermodulation distortion. The signal generators generate tones at 229.98046875 MHz and 250.1220703125 MHz, such that the center frequency is approximately 240 MHz and the

tones are approximately separated by $\Delta f = 20$ MHz, as specified by the datasheet. Note that I have also included a SLP-300+ DC-270 MHz LPF at the output of each signal generator in order to suppress the second order harmonics, which can mix with the carriers and add additional intermodulation distortion power at $2f_1 - f_2$ and $2f_2 - f_1$. The power per carrier tone is swept from -8 dBFS to -6 dBFS such that the growth in power of the IM3 products with carrier power is observed, and so that measurements are conducted at around -7 dBFS, the carrier power used in the datasheet.

Figure 3.11 displays the maximum and minimum IM3 as a function of power per carrier tone. The maximum IM3 always occurs at 250.1220703125 MHz and the minimum IM3 always occurs at 229.98046875 MHz. The datasheet quotes a maximum IM3 of -72 dBc² and a typical value of -80 dBc, while I measure an average maximum value of -65 dBc and an average minimum value of -78 dBc for the carrier powers probed. While the minimum IM3 is consistent with the typical datasheet value, the maximum IM3 is 7 dB greater than the datasheet maximum. The difference between the measured and datasheet values may be partially due to the splitter/combiner, which can allow some fraction of the power of each carrier to leak into the other signal generator and produce additional IM3 distortion. However, CHORD can maintain a signal power greater than the power in the highest IM3 distortion product in order to mitigate added power from IM3 distortion.

²Note that while dBc is typically quoted as the power ratio between the carrier to another value (positive dBc), the datasheet quotes dBc as the power ratio between some value and the carrier (negative dBc).

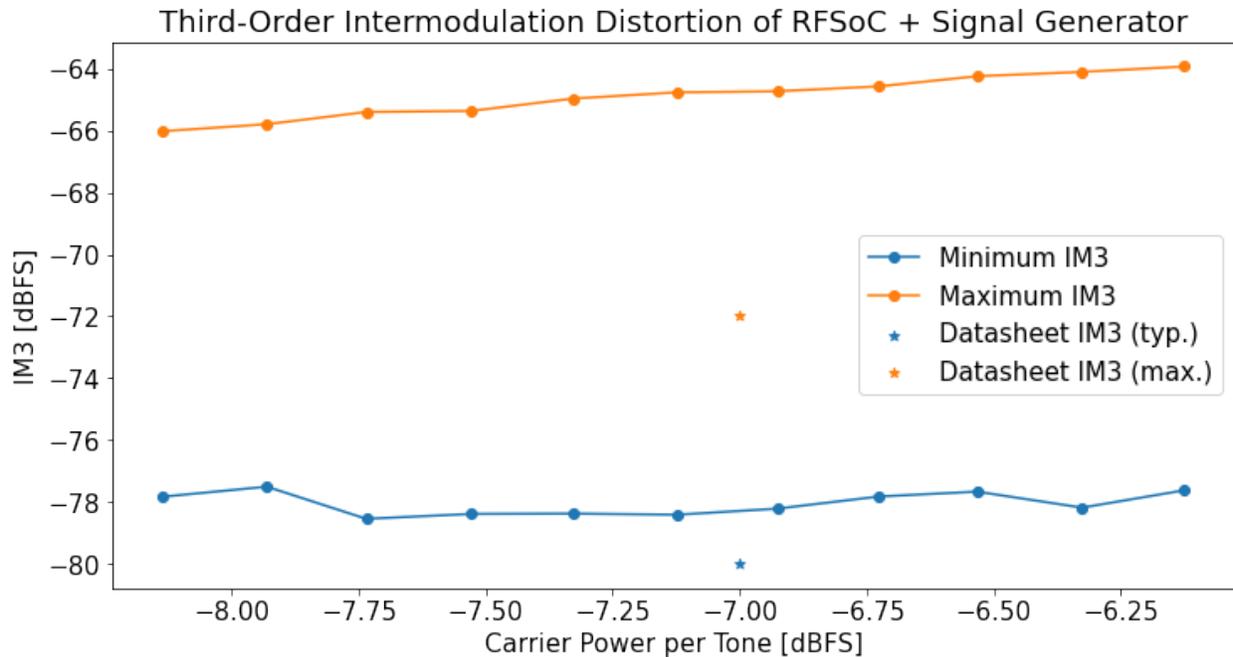


Figure 3.11: The IM3 products of the RFSoc as a function of carrier power per input tone. While the minimum measured IM3 is roughly consistent with the datasheet IM3, the maximum measured IM3 is 7 dB greater than the datasheet maximum IM3. The discrepancy in the maximum IM3 could be due to the imperfect isolation of the Z99SC-62-S+ splitter/combiner, resulting in some leakage from one signal generator to another that results in additional intermodulation distortion. However, CHORD can maintain a signal power greater than the power in the highest IM3 distortion product in order to mitigate IM3 distortion, and monitor frequency bins expected to be contaminated with IM3 distortion power from known RFI tones.

Differential Nonlinearity (DNL)

The differential nonlinearity (DNL) of an ADC is a measure of the width of the voltage range corresponding to individual digital codes. A diagram demonstrating DNL is displayed in Figure A.2 in Appendix A. An ideal ADC has zero DNL, meaning that the width of the voltage range for every individual code is equivalent. For example, an ADC with a full-scale analog input voltage of 1 V that samples with a resolution of $n = 2$ bits will allocate 0.25 V per digital code, since there are 2^n digital codes. However, real ADCs often have varying voltage ranges for each digital code due to manufacturing imperfections, which means that some digital codes may be over- or under-represented when digitizing analog signals.

Since the voltage range corresponding to each digital code cannot be measured directly, I instead use a digitized sine wave to measure the DNL of the ZU47DR. A sine wave is an ideal waveform to use for measurement of the DNL because its Fourier transform is narrowly defined and can be produced by signal generators with less distortion than other waveforms. A histogram of counts per digital code is then computed, which is proportional to the DNL of the ADC. For example, a digital code with a positive DNL will count more samples than expected for an ideal ADC, as that digital code would correspond with a wider voltage range, and count some samples that should correspond with a neighboring digital code.

A sine wave is generated with $f_c = 573.12986341$ MHz at full-scale, and 1000 frames are captured. f_c , the carrier power, and the number of frames are chosen such that all digital codes are exercised (for example, a frequency that is a multiple of f_{res} may miss some

digital codes, arbitrarily affecting the DNL). The histogram of counts per digital code is computed and normalized to the total number of counts. Each frame is fit to a sine wave with fixed frequency³ f_c , and the best-fit sine wave for each frame is quantized assuming it is processed by an ideal 14-bit ADC, whose normalized histogram is also computed. These two histograms are presented in the top panel of Figure 3.12. The DNL is then computed according to

$$\text{DNL} = \frac{\text{Normalized Measured Counts} - \text{Normalized Simulated Counts}}{\text{Normalized Simulated Counts}}, \quad (3.13)$$

which is presented in the lower panel of the same figure along with the datasheet-specified maximum DNL. Note that this measurement is not reliable at extreme digital codes because while the ideal sine wave has a constant amplitude, the amplitude of the sine wave digitized by the RFSoc is not necessarily as stable. A sine wave will have more counts towards extreme digital codes, such that over- or under-representation of counts caused by amplitude fluctuations will be more pronounced.

The measured DNL nearly always falls within the range specified by the datasheet, except towards high digital codes, meaning that the ADC behaves linearly within the specified tolerance. A low DNL is important in radio astronomy applications, since most receivers are designed such the RMS power of the input is far below full-scale in order to accomodate

³Note that the clocks of the signal generator and CRS board are synchronized.

for transient RFI that may saturate the ADC. A result of having low-power input signals is that they only exercise a small portion of the total number of digital codes. If there is a high DNL within this range, a significant portion of the digitized signal will be distorted. The measured DNL indicates that CHORD can supply a low-power input signal to the RFSoc and expect essentially no added distortion from differential nonlinearities in the ADC.

Crosstalk

Crosstalk is a crucial measurement in the characterization of radio cosmology and astrophysics instrumentation, as data analysis is conducted through the cross-correlation of voltage signals from different antennas. If one antenna picks up a signal received by an adjacent antenna, the signals will be correlated and thus amplified when the cross-correlation is computed between the two antennas.

The crosstalk is measured by injecting a signal into one ADC channel with frequency f_c and measuring the power at f_c in another ADC channel. The DC–1400 MHz LPF is not present in this measurement. In Figure 3.13, I present the crosstalk from ADC channels 1, 3, and 7 onto channel 0. Channel 0 is injected with a -35 dBFS carrier at fixed frequency in order to exercise a small number of digital codes (to prevent issues with interleaving calibration), while a “leakage tone” at -1 dBFS is individually injected into the other channels. The crosstalk power observed in channel 0 is measured as the leakage tone is swept in frequency across the first Nyquist zone.

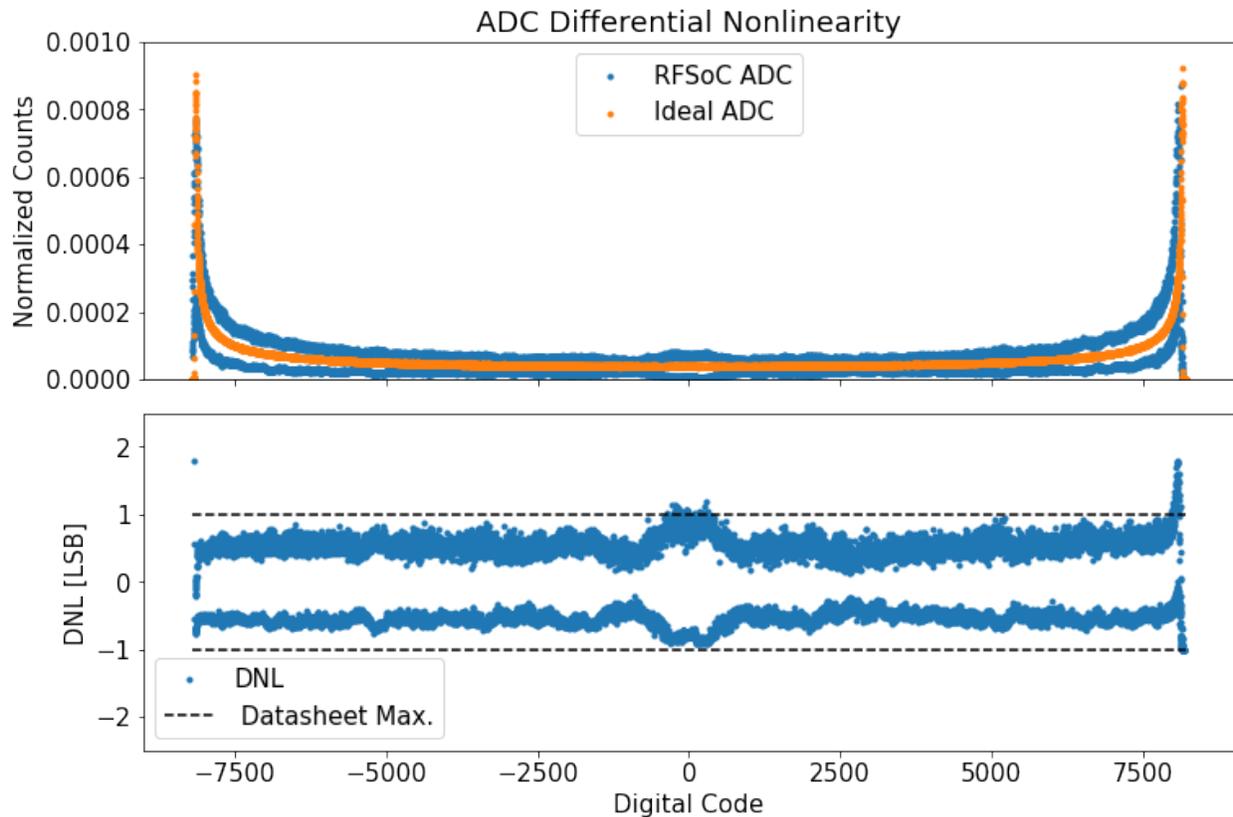


Figure 3.12: Top: in blue, a histogram of counts for 1600 frames of a sine wave digitized by the RFSoc with $f_c = 573.12986341$ MHz at full-scale. In orange, a simulated sine wave with the same amplitude and frequency digitized by an ideal ADC. **Bottom:** the DNL of the RFSoc, computed from the histograms in the top panel and Equation 3.13. The DNL for the majority of digital codes is well-contained within the specified datasheet maximum. A gap in the DNL is observed, which may be an intentional design by the manufacturer. Note that this test is not reliable for extreme digital codes (e.g., the large spike seen towards the far-right) because the amplitude of the sine wave digitized by the RFSoc is not necessarily as stable as the amplitude of the ideal sine wave, and a sine wave will inherently have more counts towards extreme digital codes. The measured DNL means that CHORD can supply a low-power input signal to the RFSoc, and expect that the digitized signal is essentially free from distortion due to differential nonlinearity.

I find that the crosstalk from channel 1 to channel 0 represents the highest crosstalk, which is expected due to their proximity, and is frequency-dependent (possibly due to some reactive coupling between the two analog traces that behaves like a high-pass filter). The maximum crosstalk is slightly above the datasheet maximum (or worst-case) value of -76 dBFS. However, note that the crosstalk is not only dependent upon crosstalk within the RFSoc, but all the components between the SMA input and the RFSoc as well, which is different between the CRS board and the board used in the datasheet. Therefore, the discrepancy between the measured and datasheet crosstalk is expected and is likely caused by the different platforms used. The crosstalk from channels 3 and 7 onto channel 0 are far below that of channel 1 and the datasheet value, which is likely due to the larger trace separation between these channels on the CRS board. As such, the worst-case crosstalk measured is approximately -72 dBFS, comparable to the datasheet worst-case value, and the crosstalk is reduced significantly with increasing channel separation. Therefore, if CHORD alternates input signal polarizations between adjacent channels on the CRS board, crosstalk due to the ADC and CRS analog path will be highly suppressed.

S_{11}

The S_{11} of a device is a measure of the reflected power due to a mismatch in the impedance of the device with respect to some nominal or reference impedance. The S_{11} expressed in dB units may be written as (Pozar, David M., 2012)

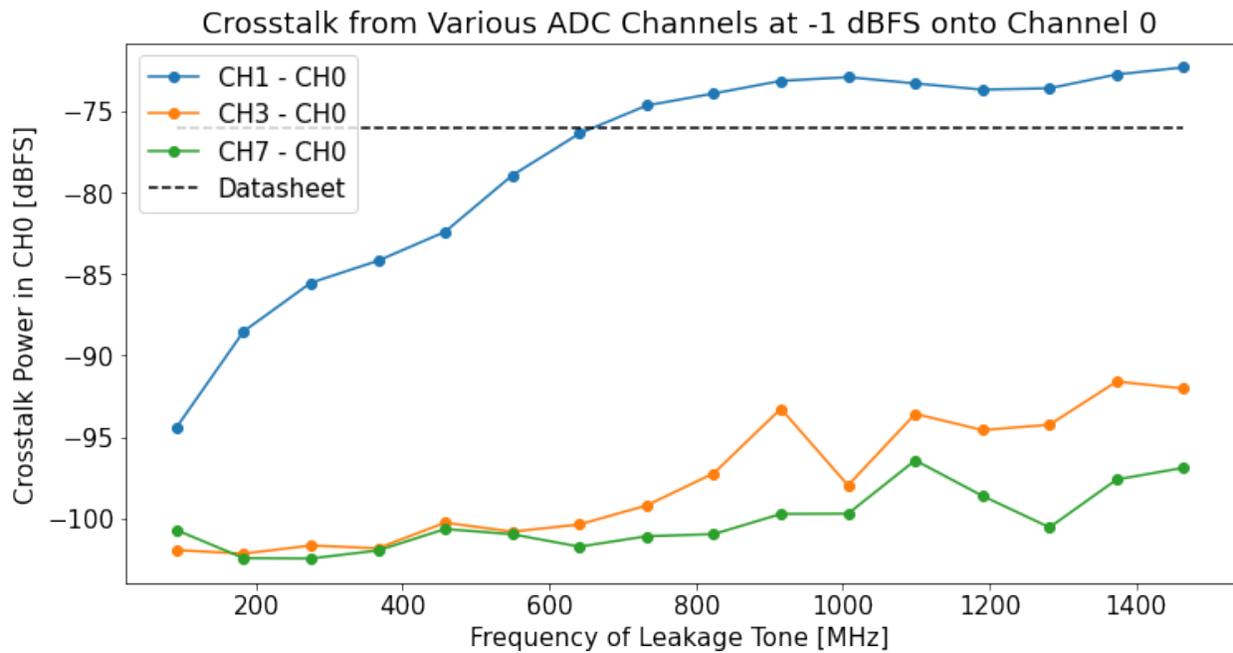


Figure 3.13: The crosstalk leaked from CHx to CH0 of the RFSoc as a function of leakage frequency. As expected, crosstalk between adjacent channels is the most significant. The observed increase in crosstalk for CH1 to CH0 is expected due to capacitive coupling between the two analog signal paths that behaves as a high-pass filter. The crosstalk from more distant channels is far below the worst-case value specified by the datasheet, and is close to the noise floor. As such, CHORD can highly suppress crosstalk by alternating input signal polarizations between adjacent channels.

$$S_{11} \text{ [dB]} = 20 \log_{10}(\Gamma) = 20 \log_{10} \left(\frac{Z_L - Z_0}{Z_L + Z_0} \right), \quad (3.14)$$

where Γ is the input reflection coefficient of the device (note that as defined, $S_{11} = \Gamma$ when the S_{11} is expressed as a complex number), Z_L is the load impedance, and Z_0 is the reference impedance, which is typically 50Ω for RF electronics. $S_{11} = 0$ dB occurs when the load impedance $Z_L \rightarrow \infty$, indicating that all the incident power is reflected, and $S_{11} \rightarrow -\infty$ dB in the ideal case when $Z_L \rightarrow Z_0$, indicating maximum power transfer and zero reflected power. Note that while in general, S_{11} is a complex number, the magnitude $|S_{11}|$ is a more relevant quantity than the phase, so I will refer only to the magnitude in the following discussion. In radio astronomy, keeping $|S_{11}|$ as small as possible is important to minimize the amount of power that is sent back up the analog amplification and filtering chain. Any power reflected into this chain from the ADC input may cause distortion and nonlinear behavior in the amplifiers, which would eventually be digitized by the ADC and corrupt data.

The $|S_{11}|$ measured at the input of each ADC channel of the CRS board is presented in Figure 3.14 across the first Nyquist zone. I find that $|S_{11}| \lesssim -9$ dB across the CHORD bandwidth, suggesting that $\lesssim 13\%$ of the input power is reflected by the analog input. The datasheet quotes $|S_{11}| < -14$ dB, or $\lesssim 4\%$ of the input power is reflected. Note that my measurement of the $|S_{11}|$ is measured from the perspective of the analog input, meaning that there are additional contributions to the $|S_{11}|$ along the analog path to the ADC. For example, the $|S_{11}|$ of the LFCN-9170+ LPF shown in Figure 3.2 is approximately -15

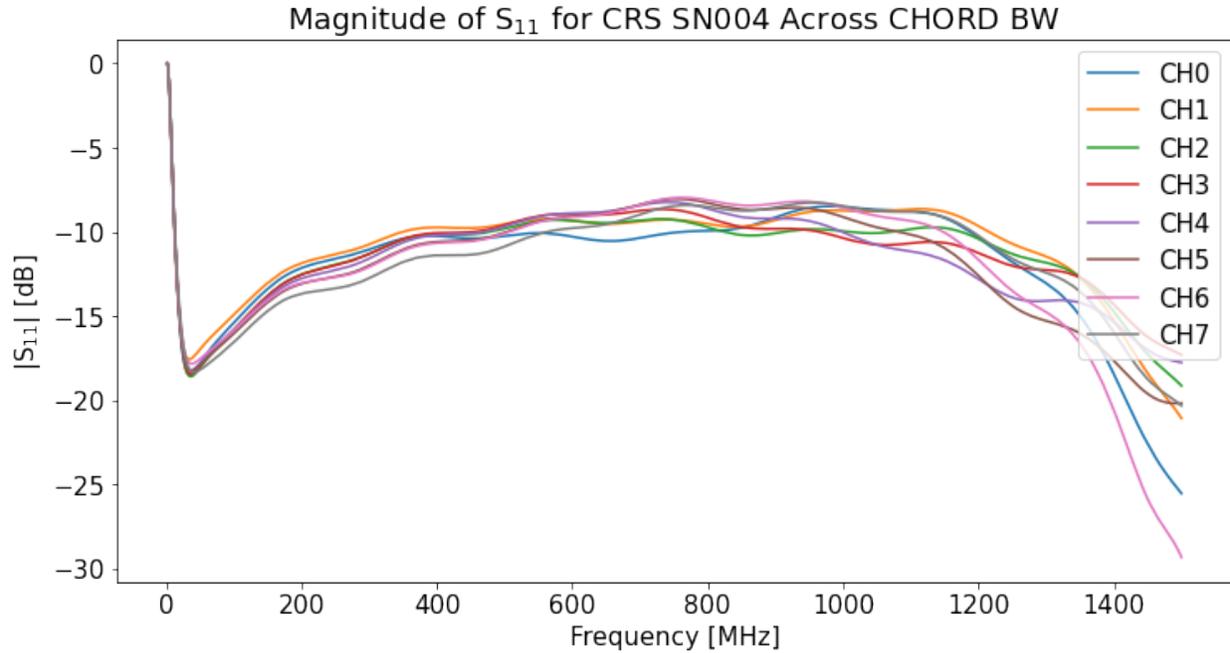


Figure 3.14: $|S_{11}|$ for the CRS board analog inputs in the first Nyquist zone (the CHORD bandwidth is from 300–1500 MHz). $|S_{11}| \lesssim -9$ dB in the CHORD bandwidth, suggesting $\lesssim 13\%$ of the input power is reflected. While the datasheet quotes $|S_{11}| < -12$ dB, it is noted that this discrepancy is likely due to differences in PCB trace routing or components used along the analog input path. The target S_{11} for CHIME was -10 dB.

to -17 dB from 1000–1500 MHz⁴ (Mini-Circuits, a), while the $|S_{11}|$ of the MABA-011118 balun is typically -15 dB from 0–1500 MHz (MACOM Technology Solutions Inc.). While the datasheet does not specify the platform used to characterize the ADC, the discrepancy between measured and datasheet $|S_{11}|$ is likely related to differences in the analog signal path.

⁴The LFCN-9170+ datasheet does not quote the return loss below 1000 MHz.

3.3 The RFSoc as a Potential CHORD Platform

In this section, I compare the measured quantities to their corresponding values measured for the EV8AQ160 ADC used on the ICE board, or those quoted in the datasheet if there is no corresponding EV8AQ160 value. Table 3.3 presents all quantities measured in this chapter alongside those of the datasheet and EV8AQ160, where applicable. The table also shows an evaluation of whether a given value measured for the RFSoc is improved with respect to the corresponding EV8AQ160 value, or is consistent with the datasheet value if no EV8AQ160 value exists.

The SNR of the ZU47DR is approximately 16 dB greater than that of the EV8AQ160. The measured noise properties of the ADC (SNR and NSD) therefore offer a significantly improved dynamic range for a radio astronomy receiver to mitigate ADC saturation from transient RFI compared with the EV8AQ160.

The highest-power second-order harmonic is approximately 15 dB greater than the datasheet value at 240 MHz; however, this maximum second-order harmonic is quoted for $f_c \sim 400$ MHz. Recall that the signal generator produces large harmonics itself and that the second-order harmonics dominate over the third-order harmonics, both of which are only suppressed below 3 dB by the DC–1400 MHz LPF for $f_c \gtrsim 700$ MHz. In Figure 3.6, the harmonics decrease with f_c , suggesting that high-frequency measurements of the harmonics are more representative of the harmonics produced by the ADC only, in which case, the harmonics are consistent or improved in comparison with the datasheet values.

An RFSoc used to digitize amplified signals from an antenna will be limited by its own harmonics plus that of upstream amplifiers rather than a signal generator. The RFSoc ADC has high enough dynamic range to observe even low-power harmonics; as such, amplifiers must be as spectrally pure as possible. Nevertheless, the average total harmonic distortion is more than 25 dB lower than that of the EV8AQ160.

Similar to the SNR, the SINAD is improved by 16–17 dB over the EV8AQ160, and the ENOB is improved by 3 bits. In addition, the highest measured SFDR is 24 dB better than that of the EV8AQ160. The improved SINAD, ENOB, and SFDR further support the superior dynamic range for the RFSoc compared with the EV8AQ160.

The datasheet gain-time interleaving spurs are contained within the measured range, but the largest gain-time interleaving spur is 14.5 dB greater than the datasheet value. The measured offset interleaving spurs are greater than the datasheet value by 4.4 dB at worst, and lower than the datasheet value by 3.5 dB at best, with the average value consistent with the datasheet value.

As discussed earlier, the measured third-order intermodulation distortion may be degraded by the additional mixing from the signal generator. However, the measured third-order intermodulation distortion shows significant improvement compared with the EV8AQ160, and is found to be 16 dB lower. Though the EV8AQ160 has better differential nonlinearity than the RFSoc by approximately a factor of 4, both platforms contain the differential nonlinearities within ± 1 least significant bits. Since the input signal to the

RFSoc will be amplified to use more bits than if the EV8AQ160 was used, the effective distortion from differential nonlinearity will be lower for the RFSoc.

The ZU47DR shows vastly superior crosstalk mitigation compared with the EV8AQ160, which is lower by 26 dB when comparing with crosstalk measurements on the same EV8AQ160 chip. Even when the EV8AQ160 chip is on separate mezzanines, the crosstalk on adjacent channels of the ZU47DR is still 6 dB lower. Analog signal path differences likely explains the difference between the measured, datasheet, and EV8AQ160 $|S_{11}|$ values, which can always be improved through changes to the analog path or a matching network at the front of the analog input. Overall, the RFSoc ADC outperforms the EV8AQ160 in nearly all metrics.

Quantity	RFSoc (This Work)	RFSoc (Datasheet)	EV8AQ160
n (bits)	14	14	8
f_s (GSPS)	3	5	0.8

Table 3.2: The resolution and sampling rate of the ADC platforms compared in this work.

3.4 Summary

In this chapter, I presented two potential platforms for the CHORD F-Engine: 1) an ICE-system-based F-Engine, and 2) an RFSoc-based F-Engine. The measurements discussed in this chapter suggest that the RFSoc ADC offers significant performance improvements compared with the EV8AQ160 ADC used on the ICE boards.

Quantity	Measured	Datasheet	EV8AQ160	Improved or consistent?
SNR (dBFS)	62.3 [†]	62.2 [‡]	46.3 to 46.5	✓
NSD (dBFS/Hz)	-154 [†]	-154	-132.5 to -132.7	✓
H2 (dBFS)	-73.3 to -95.2	-88, -72	-	✓
H3 (dBFS)	-82.6 to -99	-78, -66	-	✓
THD (dBFS)	-73.3 to -89.7	-	-48 to -53	✓
SINAD (dBFS)	62.2	-	44 to 45	✓
ENOB (Bits)	10	10 [‡]	6.9 to 7.2	✓
SFDR (dBFS)	-73.3 to -89.5	-	-49 to -54	✓
SFDR _{xH2/3} (dBFS)	-78.5 to -89.5	-85	-	✓
GTIS (dBFS)	-78.5 to -96.6	-93	-	✓
OIS (dBFS)	-81.6 to -89.5	-86	-	✓
IM3 (dBc)	-65, -78 [*]	-75, -80 ^{**}	-48	✓
DNL (LSB)	~ ± 1	± 1	-0.19 to -0.24	Note ^a
Crosstalk (dBFS)	< -72	< -76	< -46 (same chip) < -66 (diff. mezz.)	✓ ✓
S ₁₁ (dB)	< -9	-12	< -14 (400–800 MHz)	Note ^b

Table 3.3: Measured ADC quantities compared with quoted datasheet quantities and the EV8AQ160 ADC used on the ICE boards. “Improved” refers to a comparison with the EV8AQ160 while “consistent” refers to a comparison with the datasheet. The resolution and sampling rate of each ADC are shown in Table 3.2. Values for the EV8AQ160 ADC are quoted from an internal CHIME memo characterizing the unit. Datasheet values quoted are typical values at 240 MHz and 1.9 GHz, unless the value at both frequencies is the same (note that 1.9 GHz is within the second Nyquist zone for measured quantities, but within the first Nyquist zone for datasheet quantities). A single datasheet value indicates that values at 240 MHz and 1.9 GHz are equivalent. Some measured quantities are not quoted by the datasheet or the EV8AQ160. Measured quantities H2 through SFDR_{xH2/3} (SFDR excluding H2 and H3) are quoted for the range of measured values across the first Nyquist zone. [†]Averaged across the first Nyquist zone (0–1500 MHz). The SNR quoted is the directly measured average value from Figure 3.5 rather than that calculated in Equation 3.3. [‡]Calculated from the datasheet NSD, assuming that noise domination is from quantization only. ^{*}The measured average maximum and minimum IM3, respectively. ^{**}The quoted maximum and typical IM3, respectively.

^aWhile the DNL is smaller for the EV8AQ160, the RFSoc has a resolution of 14 bits while the EV8AQ160 has a resolution of 8 bits. Both platforms ensure essentially no distortion from differential nonlinearity.

^bThe |S₁₁| discrepancy is reflective of the differences in platforms used to measure each device. The |S₁₁| can be improved by changing components or adding a matching network.

Future work will be dedicated to implementing the CRS board at D3A, and performing analyses similar to that discussed in Chapter 2 to validate its on-sky performance. As mentioned in Chapter 2, a digital signal processing platform that can sample at the CHORD Nyquist sampling frequency will eliminate the need for band-pass filtering. Therefore, a more representative measurement of T_{sys} can in principle be made with the RFSoc.

Likewise, by not relying on band-pass filtering to recover the entire CHORD band, using the RFSoc means there are less analog components that need to be characterized in the signal chain. As such, a signal chain using the RFSoc leads to a smoother gain, which is crucial for 21-cm cosmology. However, since the performance of the ADC itself may vary from chip to chip, each ADC will need to be characterized independently to gauge the redundancy of the RFSoc. To digitize all 1024 analog signals for CHORD, 128 RFSoc-based circuit boards are required. One possibility to characterize the performance of each would be to develop quality control tests that can be performed by the PCB manufacturer, and then perform all or a subset of the measurements detailed in this chapter on each board.

The channelization firmware to process digitized samples with 14-bit resolution and at 3 GSPS is also in development, as well as the hardware to implement the corner-turn operation. As such, an F-Engine platform based around the RFSoc is quickly becoming an attractive alternative to enable CHORD to meet its desired performance specifications.

Chapter 4

An Algorithm for Measuring Soil

Electrical Properties from Antenna

Impedance

The sensitivity needed to detect the global redshifted 21-cm signal require precise calibration of instruments and systematic control ([Liu and Shaw, 2020](#)). Chromatic effects associated with the spatially-dependent antenna beam are a particularly important source of concern for global 21-cm experiments. In order to reduce beam chromaticity effects, the MIST experiment implements a unique instrument design corresponding to a blade dipole sitting directly above the ground with no reflective ground plane. However, the effect of the soil electrical properties need to be understood in order to accurately account for their effect

on the beam. In this chapter, I present an algorithm to measure soil electrical properties using antenna impedance, with the goal of using this information to inform simulations of the MIST beam model.

4.1 Chromaticity Challenges for Global 21-cm Experiments

Antennas with a chromatic (frequency-dependent) beam can couple with foregrounds and introduce non-smooth spectral structure, which can prove challenging to model at the level required for detection of the global 21-cm signal ([Mozdzen et al., 2016](#); [Singh et al., 2018](#); [Vedantham et al., 2014](#)). Therefore, detailed knowledge of the antenna beam is crucial to extract the global 21-cm signal¹. Most global 21-cm experiments correspond to ground-based, single-antenna instruments. Many of these antennas operate above a ground plane (or ground screen), which is a metal sheet or mesh of various size and geometry lying beneath an antenna. The function of a ground plane is to increase the gain of the antenna in the direction of the sky (similar to a parabolic reflector) as well as obscure the ground from the antenna, as the soil electrical properties of the ground can introduce beam chromaticity.

However, due to their finite size, ground planes have been shown to introduce additional chromatic effects in the beam that are distinct from those caused by the ground.

¹Note that in this chapter, I only focus on the extraction of the cosmic dawn feature of the global 21-cm signal.

For example, [Spinelli et al. \(2022\)](#) showed that increasing the size of ground planes can introduce more complex frequency structure while not adequately suppressing the effect of soil electrical properties. Likewise, the effects of ground plane-induced chromaticity can mimic the expected absorption feature in the global 21-cm signal ([Bradley et al., 2019](#)). These complications motivate global 21-cm instrument designs that do not incorporate ground planes².

As such, the MIST experiment observes without a ground plane. With a carefully selected site, MIST can in principle operate above a continuous plane that extends beyond the far field of the antenna, and avoid the chromatic effects resulting from using a ground plane. However, the lack of a ground plane introduces contrasting systematic effects and calibration challenges. Without a ground plane, a significant fraction of the beam solid angle is intercepted by the ground. The instrument therefore has reduced sky response and increased sensitivity to the electromagnetic properties of the underlying soil.

Figure 4.1 displays the simulated directivity of the MIST blade dipole antenna at zenith as a function of frequency for a variety of soil models ([Monsalve et al., 2023](#)). This figure illustrates that the beam is strongly dependent upon the soil electrical properties. Therefore, MIST must devote careful attention to the characterization of soil models, and ideally implement multiple methods to compare models derived from different methods. Note that while the electrical properties of the soil can introduce chromatic effects to the

²Instrument designs incorporating monopole antennas have also been explored as an alternative to dipoles operating above ground planes.

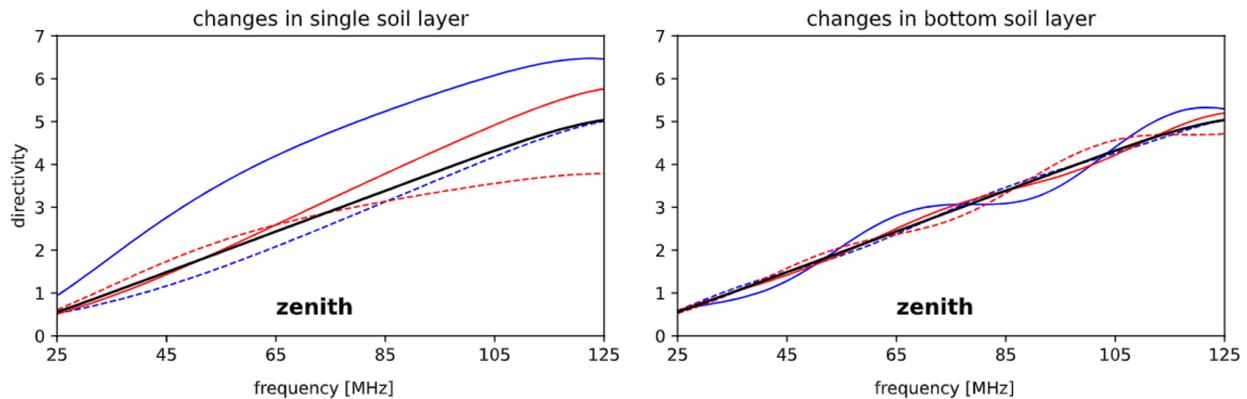


Figure 4.1: The effect of changing soil electrical parameters on the MIST antenna beam. The simulated directivity of the MIST antenna at zenith is plotted as a function of frequency for varying soil electrical parameters. The left panel corresponds to one-layer soil models while the right corresponds to two-layer soil models. Figure adapted from [Monsalve et al. \(2023\)](#), which contains further details for each simulation.

beam, the effect is distinct from that of a ground plane, which means that the MIST beam is associated with entirely different systematics from other experiments. The advantage of operating directly above soil is that the sharp electrical discontinuities associated with ground planes are eliminated.

MIST uses electromagnetic modeling software, namely Feko³, in order to generate models of the beam based on antenna geometry and environment. Therefore, MIST requires on-site measurements of the soil electrical properties to inform Feko models of the instrument environment. There are a number of soil electrical characterization techniques developed by geophysicists that MIST has or is currently exploring, including electrode arrays to determine soil resistivity and ground-penetrating radar (GPR) to map spatial variations in the soil

³<https://altair.com/feko>

(Reynolds, 2011). I have developed a novel technique that uses on-site measurements of the antenna impedance to estimate the best-fit soil electrical properties, which I describe in the following sections.

4.2 The Antenna Modeler and Performance Emulator (AMPERE)

The Antenna Modeler and Performance Emulator (AMPERE) is a Python-based package that interpolates various antenna quantities generated from simulations. In MIST, AMPERE is used to interpolate the impedance of the MIST blade dipole by using simulations of the impedance generated in the Feko simulation software. In this section, I describe the AMPERE algorithm and present evaluations of its accuracy.

Figure 4.2 displays a diagram demonstrating the AMPERE algorithm. For simplicity, the algorithm is discussed in the context of a single soil layer with conductivity σ and relative permittivity ε . First, a set of impedances of the MIST blade dipole antenna are simulated, as shown in the left column. Individual lines correspond to the antenna impedance for a single set of soil parameters. Next, the impedance as a function of soil parameters is fit to an n_p -dimensional polynomial, where n_p is the number of parameters for a given soil model. It is important to note that the order of the polynomial can be chosen by the user. The best-fit coefficients can then be used to evaluate the impedance of the antenna with a new

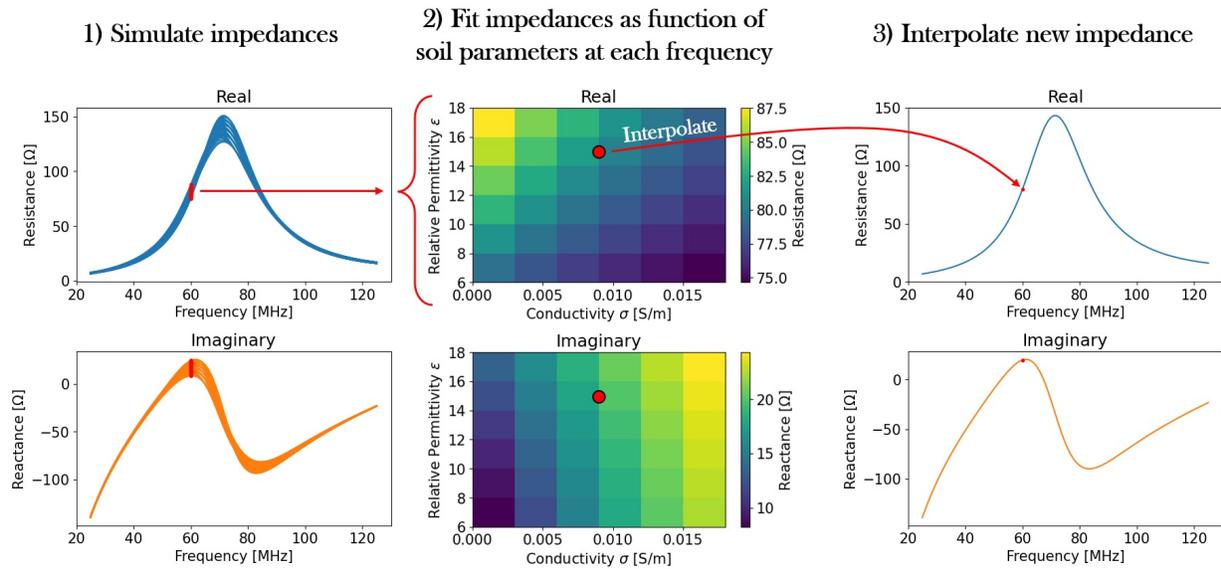


Figure 4.2: Diagram illustrating the AMPERE algorithm in the context of impedance interpolation. 1) A set of impedances is simulated in an EM simulation software. 2) At each frequency, the impedance as fit as a function of soil parameters. 3) The best-fit parameters are then used to interpolate for a new impedance corresponding to some set of soil parameters lying within the soil parameter space corresponding to the simulated impedances. This is done for both real and imaginary components of the impedance.

set of soil parameters, shown as a red dot in the center column of the figure. This process is repeated for each frequency, producing a new interpolated impedance, as shown in the right plots. Note that the real and imaginary parts of the impedance are fit for separately.

In this section, I will introduce two sets of simulations: one corresponding to a one-layer soil model, and one corresponding to a two-layer soil model. Each simulation set is designed to model the soil properties of the MIST observing site in Death Valley, Nevada, USA. I will describe the metrics used to evaluate the accuracy of the interpolations using these simulation sets in this section, and then apply the simulations in the following section to fit for the impedance of data collected at Death Valley.

Figure 4.3 shows the MIST blade dipole as implemented in Feko in the top panel (borrowed from [Monsalve et al. \(2023\)](#)). The panel length is 1.2 m, the panel width is 0.6 m, and the panel height is 0.52 m. The simulated antenna is composed of two thin panels for the blade dipole, the balun (located between the panels and receiver), and the MIST receiver. The holes in the panels are used to mount the panels to the fiberglass frame using nylon screws and washers. The bottom panel of Figure 4.3 displays the two soil models, where the antenna in the top panel sits directly above the surface. The one-layer soil model is composed of a semi-infinite plane of soil with conductivity σ and relative permittivity ε that extends down to an infinite depth. The one-layer model thus has $n_p = 2$ soil parameters. The two-layer soil model begins with the first layer, which is an infinite plane extending from the surface to a depth d with conductivity σ_1 and relative

permittivity ε_1 . The second layer is an infinite plane extending from depth d to infinity with conductivity σ_2 and relative permittivity ε_2 . The two-layer model thus has $n_p = 5$ soil parameters.

The simulation set corresponding to the one-layer model is composed of $N_{\text{sims}} = (n_s)^{n_p} = 36$ individual simulated impedances, where $n_s = 6$ is the number of samples per parameter. The soil parameters are arranged along an evenly-spaced grid corresponding to each combination of samples for the two soil parameters (note that AMPERE does not require samples to be arranged along a grid, but is done here for ease of automatic generation of a large number of simulations). The ranges of samples for each parameter, the span, are displayed in the left of Table 4.1 for the one-layer model.

The simulation set corresponding to the two-layer model is composed of $N_{\text{sims}} = 1024$ individual simulated impedances, where $n_s = 4$ samples per parameter. As with the one-layer model, the parameters of the two-layer model are arranged along an evenly-spaced grid. The spans of each parameter for the two-layer model are displayed in the right of Table 4.1.

The interpolation accuracy is assessed by comparing the interpolated and directly simulated impedance values. That is, if a given one-layer simulated impedance corresponds to the soil parameters $(\sigma_i, \varepsilon_i)$, then an impedance is interpolated with soil parameters $(\sigma_i, \varepsilon_i)$ and compared with the associated simulation. The difference between simulated and interpolated impedances with the same soil parameters is computed for each combination of soil parameters within a given simulation set. This exercise is done for

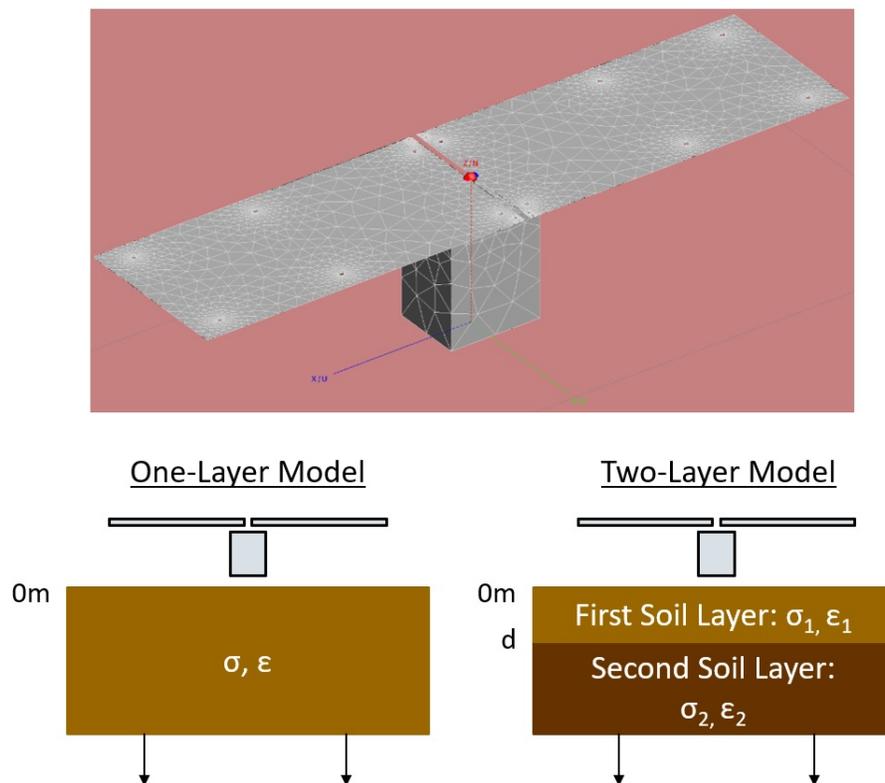


Figure 4.3: Top: an example of the MIST blade dipole in FEKO. The model consists of the two antenna panels, the receiver box, the balun (underneath the panels and above the receiver box), and two wires connected to an excitation port between the panels. The MIST instrument sits directly above the soil. Image borrowed from [Monsalve et al. \(2023\)](#), which contains further details. **Bottom:** diagrams of two soil models implemented in Feko.

One-Layer Model		Two-Layer Model	
Soil Parameter	Span	Soil Parameter	Span
σ [S/m]	0.085–0.135	σ_1 [S/m]	0.09–0.15
ε	1–4	σ_2 [S/m]	0–0.015
		ε_1	1–1.3
		ε_2	68–80
		d [m]	1.4–2.6

Table 4.1: **Left:** the span of the simulation set corresponding to the one-layer soil model. The parameters are arranged in an evenly-spaced grid, and there are 6 samples per parameter. **Right:** the span of the simulation set corresponding to the two-layer soil model. The parameters are arranged on an evenly-spaced grid, and there are $n_s = 4$ samples per parameter.

multiple sets of polynomial orders, which is used to evaluate the optimal polynomial order used when fitting impedances measured in field deployments. “Polynomial order” is defined as the set of polynomial orders used to fit each parameter. For example, the polynomial order for a one layer-model fit with a second-order polynomial for each parameter is [2, 2].

The top of Figure 4.4 shows the overplotted simulations for the one-layer soil model. The bottom of the figure shows the absolute differences between simulated and interpolated impedances evaluated with the same soil parameters for the one-layer soil model for the three best sets of polynomial orders. Note that the y-axis scale is different for each column, and the differences for each set of parameters are overplotted. The difference in the real and imaginary parts are shown in separate plots. The region of values within one standard deviation of the impedance differences is shown in cyan. The accuracy of the interpolations varies with frequency-dependent structure as well as between different

sets of parameters. Notably, the magnitude of the differences is highly dependent upon the polynomial order. The differences corresponding to third-order polynomial interpolations are orders of magnitude smaller than the first- and second-order polynomial interpolations. Further, it is found that applying polynomial orders above third-order leads to differences that are greater by many orders of magnitude compared with those shown in Figure 4.4.

The top of Figure 4.5 shows the overplotted simulations for the two-layer soil model. The bottom of the figure shows the absolute differences between simulated and interpolated impedances evaluated with the same soil parameters for the two-layer soil model for the three best sets of polynomial orders. Similar to the results shown in Figure 4.4, the absolute difference is dependent upon the polynomial order chosen. The polynomial order corresponding to [2, 1, 1, 1, 1] leads to the lowest differences and standard deviation for the two-layer model. Other combinations of polynomial orders are comparable to the absolute differences when using all first-order polynomials (left panel in Figure 4.4, or are larger by many orders of magnitude. It is likely that poor behavior is observed for some polynomial orders because they do not accurately represent the simulated data.

The differences between simulations and interpolations that are shown above are quantified with the root-mean-square error (RMSE)

$$\text{RMSE}_{\text{real}} = \sqrt{\frac{1}{N_{\text{sims}}} \sum_f (\text{Re}\{Z_{\text{interp}}(f)\} - \text{Re}\{Z_{\text{sim}}(f)\})^2}, \quad (4.1)$$

where Z_{interp} refers to the interpolated impedance for a given set of parameters shared with

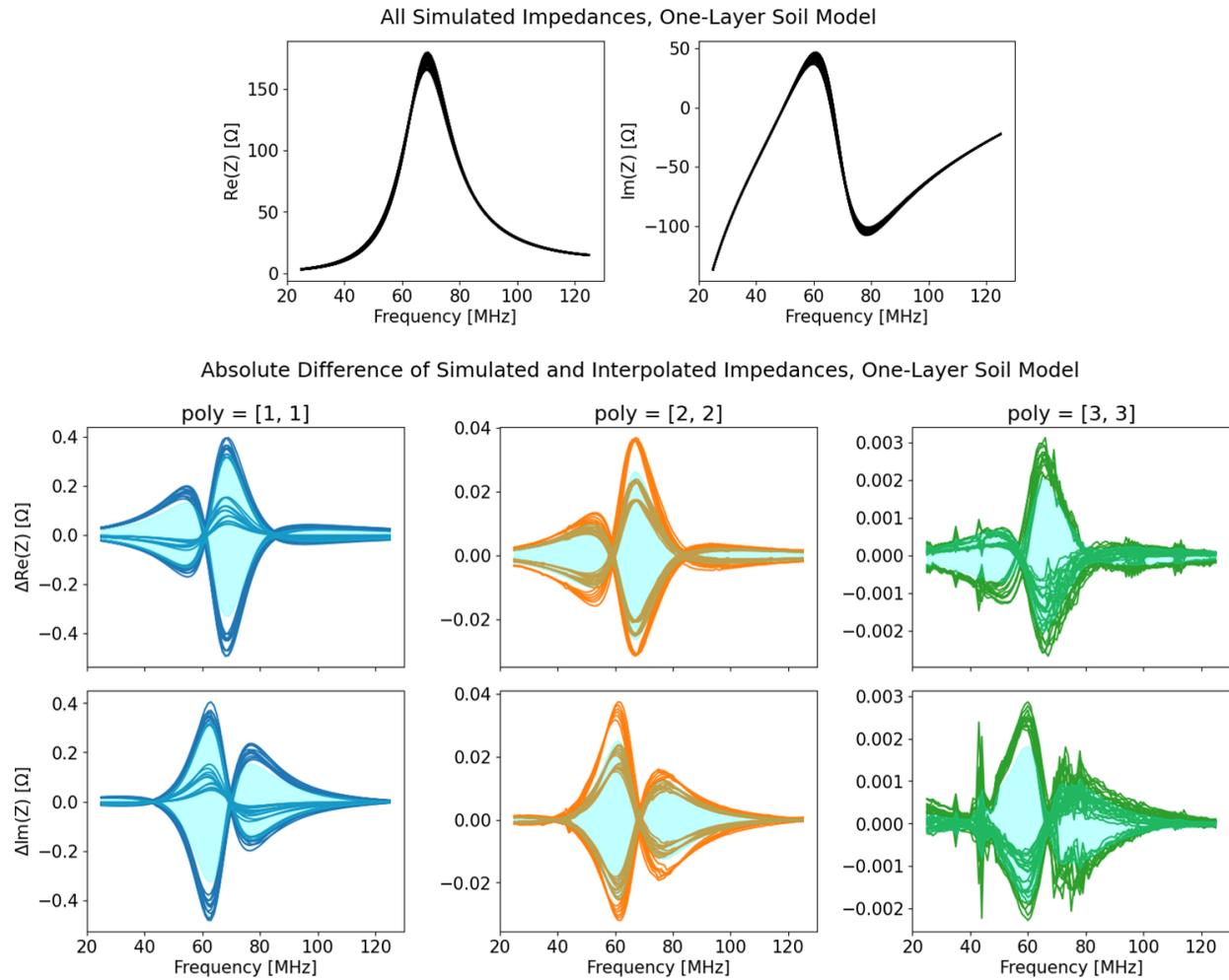


Figure 4.4: Top: all the simulated impedances for the one-layer soil model. Bottom: differences between simulated impedances for the one-layer soil model and impedances evaluated with AMPERE at the same parameters as the simulations. Note that the y-axis scales are different for each column. The region of values within one standard deviation of the impedance differences is shown in cyan. The third-order polynomial interpolations show the highest accuracy in comparison with simulations.

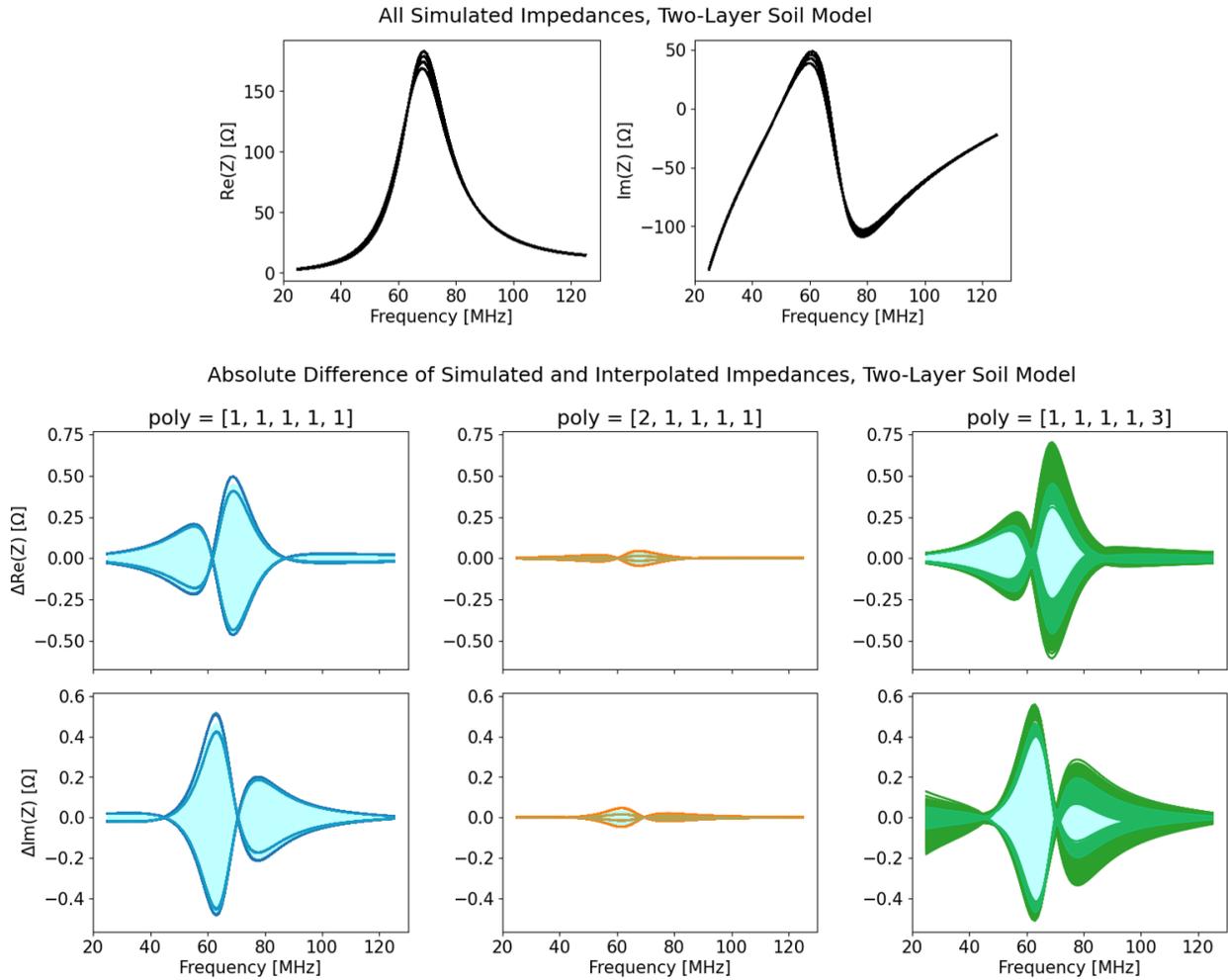


Figure 4.5: **Top:** all the simulated impedances for the two-layer soil model. **Bottom:** differences between simulated impedances for the two-layer soil model and impedances evaluated with AMPERE at the same parameters as the simulations. The region of values within one standard deviation of the impedance differences is shown in cyan. The polynomial order corresponding to [2, 1, 1, 1, 1] leads to the highest accuracy in comparison with simulations.

that of Z_{sim} , which refers to a given simulated impedance. Similar to the previous plots, the RMSE is computed for the real and imaginary parts separately to observe whether there are significant differences in the behavior between the two parts of the impedance. Computing the histogram of the RMSEs can be used to understand the overall accuracy of interpolations for a given polynomial order and determine the optimal order to use for impedance fitting.

Histograms of the RMSEs computed for the one-layer soil model are shown in Figure 4.6 for varying polynomial order. The RMSE counts for the first-order polynomial interpolations have the widest distribution, with most RMSE counts in the 0.1–0.16 Ω bins. While the RMSE counts for the second- and third-order polynomial are comparable, the counts for the third-order polynomial RMSEs are all contained within the 0–0.01 Ω bin, which suggests that third-order polynomials lead to the most accurate interpolations for the one-layer soil model, which is consistent with the results from Figure 4.4.

Histograms of the RMSEs computed for the two-layer soil model are shown in Figure 4.7 for varying polynomial order. The polynomial orders corresponding to [1, 1, 1, 1, 1] and [1, 1, 1, 1, 3] have comparable RMSE distributions, though the distribution for [1, 1, 1, 1, 3] is slightly wider. The RMSEs for the polynomial order corresponding to [2, 1, 1, 1, 1] are contained within the 0–0.025 Ω bins and are approximately an order of magnitude lower than the other two polynomial orders, which is consistent with the results shown in Figure 4.5.

It is clear from the results presented in this section that the interpolations for the one-

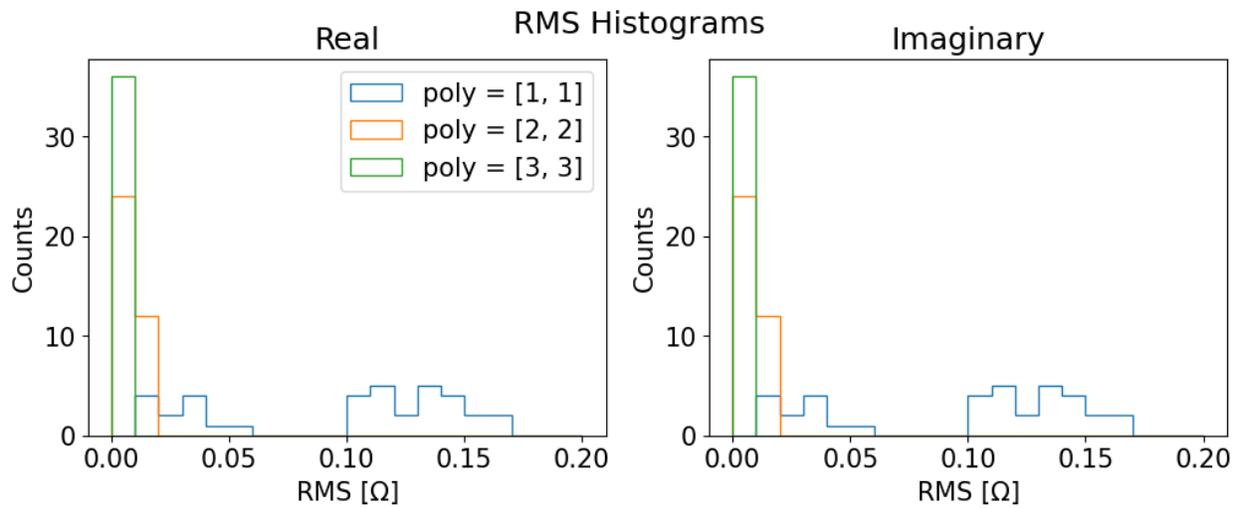


Figure 4.6: Histograms of the RMSE between simulated impedances for the one-layer soil model and impedances evaluated with AMPERE at the same parameters as the simulations. The histogram of the RMSE corresponding to third-order polynomial interpolations has more counts in lower RMSE bins than the other two polynomial orders, which is in agreement with the results from Figure 4.4.

layer model can be orders of magnitude more accurate than those of the two-layer model for the best polynomial order. The accuracy of interpolations corresponding to a given set of simulations is dependent (among other factors) upon the number of samples per parameter. Recall that the one-layer simulations have $n_s = 6$ samples per parameter while the two-layer simulations have $n_s = 4$ samples per parameter. However, the number of simulations N_{sims} increases dramatically with additional parameters, as $N_{\text{sims}} = (n_s)^{n_p}$. For example, for a two-layer set of simulations with $n_s = 6$ samples per parameter, $(n_s)^{n_p} = 6^5 = 7776$ simulations. The 1024 simulations for the two-layer model took several days to produce, meaning that increasing the number of samples further extends the simulation time from days to weeks, which is an inefficient timescale for analysis with AMPERE. Therefore, simulation sets with more parameters are limited by the practicality of producing a large number of simulations.

It is unknown to what depth the MIST antenna is sensitive, and it could in principle be sensitive to three or more layers. While the number of soil layers is dependent upon the specific observing site, two-layer soil models can reasonably approximate the soil structure in certain cases. In particular, the MARS site during the summer is composed of two layers: an “active” layer of unfrozen soil lies above a permafrost layer, which remains frozen year-round.

However, the limitations in the number of simulations that can be produced in a reasonable amount of time also have consequences for the number of layers in a given soil model. Each additional layer adds a minimum of three parameters (conductivity, permittivity, and depth) to the simulations. For example, a three-layer model has $n_p = 8$

parameters: three conductivities, three permittivities, and two depths (i.e., the depth of the interface between each layer). For $n_s = 4$ samples per parameter for a three-layer model, $(n_s)^{n_p} = 4^8 = 65536$ simulations are required, which may take weeks to produce. Further, the soil is currently modeled as being composed of layers that are homogeneous and infinite in the horizontal direction. It is possible that there are changes in the soil electrical properties along the horizontal direction, which would further complicate the soil model and require more simulations⁴. Any number of layers greater than two in the manner discussed in this section are thus left to future exploration.

As previously discussed, the polynomial order chosen is also linked to the accuracy of the interpolations. The optimal polynomial order may be chosen by comparing the absolute differences and RMSEs corresponding to different polynomial orders, which are then implemented when fitting to the measured impedance. The optimal polynomial order is not necessarily the highest order — for example, applying the one-layer model with fourth-order polynomial interpolations results in deviations an order of magnitude greater than the first-order polynomial case. It is likely that in such a case, the model is over-parameterized, as the order of the polynomials is approaching the number of samples per parameter $n_s = 6$. Also note that standard polynomials may not be the best functions to accurately model the simulations, which could also lead to misleading results from impedance fits. Therefore, other types of functions will be explored in the future.

⁴In addition, Feko does not provide the capability to generate soil models that vary along the horizontal direction.

4.3 The System for Optimizing Impedance Likelihood (SOIL)

The System for Optimizing Impedance Likelihood (SOIL) is a companion package to AMPERE and is used to determine the soil electrical properties of MIST observing sites by fitting AMPERE models to measured on-site impedances. SOIL manages impedance fits and provides the user with various options in the fitting algorithm, including the log-likelihood function, polynomial order, data variance, and more. In this section, I will discuss key aspects of SOIL and present example impedance fits to the measured MIST blade dipole impedance at Death Valley.

4.3.1 Mathematical Background

SOIL implements the `ultranest` Bayesian inference package to determine best-fit soil parameters and compare different soil models (Buchner, 2021). `ultranest` uses the MLFriends Monte Carlo nested sampling algorithm, which shrinks the prior range to iterate towards improved log-likelihood values (Buchner, 2014, 2019). Within the framework of SOIL, `ultranest` is used to explore a prior space typically (though not required to be) defined by the span of each parameter in a given simulation set. For each `ultranest` iteration, a set of parameters is selected according to the MLFriends algorithm, and SOIL calls AMPERE to interpolate for an impedance corresponding to those

parameters. The interpolated impedance then becomes the model in the log-likelihood function $\ln(\mathcal{L})$, defined by

$$\ln(\mathcal{L}) = -\frac{1}{2}(Z_{\text{meas}} - Z_{\text{interp}})^T \Sigma^{-1} (Z_{\text{meas}} - Z_{\text{interp}}), \quad (4.2)$$

where Z_{meas} is the measured impedance at a MIST observing site, Z_{interp} is an interpolated model impedance from AMPERE⁵, and Σ is the error matrix, further discussed below.

Note that each impedance Z in this case does not refer to a complex array, since the log-likelihood is not defined as a complex value — instead, each Z is typically a concatenation of the real impedance and the imaginary impedance (though the user has the option to use only the real or imaginary parts if so desired). Concatenating the real and imaginary parts of the impedance in the log-likelihood means that both are considered in evaluating the goodness of fit, which makes the log-likelihood a more robust metric for comparing a given impedance model to measured data. All results presented in this section correspond to a log-likelihood with concatenated real and imaginary parts of the impedance.

The user has several available options for how errors are treated. There are nominally two sources of error to consider in the log-likelihood: errors associated with the measured impedance data Σ_{meas} , and errors associated with AMPERE interpolations Σ_{interp} . In general, Σ can be expressed as the quadrature sum of these two error matrices, or

⁵Note that Z_{interp} in Equations 4.1 and 4.2 have slightly different definitions.

$$\Sigma = \Sigma_{\text{meas}} + \Sigma_{\text{interp}}, \quad (4.3)$$

which is considered a “quadrature” sum since each matrix contains variance and possibly covariance values. The simplest treatment of the errors in SOIL is to assume that Σ_{meas} is constant as a function of frequency, and $\Sigma_{\text{interp}} = 0$. In this case, Σ_{meas} is a diagonal matrix whose elements correspond to a constant variance σ_{meas}^2 . The user also has the option to use the covariance matrix of a given set of data, and can use only the diagonal elements or the full matrix if so desired. The use of the data covariance matrix in the error matrix Σ has not been explored in sufficient detail and is thus left for future work.

The interpolation error matrix Σ_{interp} corresponds to the covariance of the interpolated impedances evaluated with the same soil parameters as those of the simulation set; that is, the interpolated impedances shown in Figures 4.4 and 4.5 in Section 4.2. Σ_{interp} can then be expressed as

$$\Sigma_{\text{interp}} = \frac{1}{N_{\text{sims}} - 1} (\Delta Z - \overline{\Delta Z})^T (\Delta Z - \overline{\Delta Z}) \quad (4.4)$$

where I divide by $N_{\text{sims}} - 1$ rather than N_{sims} since a degree of freedom is lost due to the subtraction of the means. ΔZ is a matrix containing the differences between the each interpolated and simulated impedances, and $\overline{\Delta Z}$ is matrix whose rows contain the average value across all ΔZ . Σ_{interp} thus contains information on the accuracy of AMPERE at each

frequency for a specific set of simulation sets and polynomial order. Note that, as with the log-likelihood, the impedance differences ΔZ used to calculate the covariance matrix correspond to the concatenation of the real and imaginary parts of the impedance.

As `ultranest` is a Bayesian inference package, there are two relevant metrics that SOIL uses for model comparison: the marginal likelihood⁶, or “evidence” \mathcal{Z} , quoted in this work as $\log(\mathcal{Z})$, and the Bayesian information criterion (BIC). The marginal likelihood is (Buchner, 2014)

$$\mathcal{Z} = \int \mathcal{L}(\theta)p(\theta)d\theta, \tag{4.5}$$

where θ corresponds to a given model parameter and $p(\theta)$ refers to the prior distribution for θ . Therefore, \mathcal{Z} is the integral of the log-likelihood over the parameter space of θ scaled by the prior distribution. The marginal likelihood \mathcal{Z} is a useful quantity because it can be used to compare difference models by to evaluate the Bayes factor K , or the ratio $K = \mathcal{Z}_1/\mathcal{Z}_2$, where \mathcal{Z}_1 corresponds to model 1 and \mathcal{Z}_2 corresponds to model 2. Note that this assumes equal prior probability for each model. A value of $K > 1$ more strongly supports model 1 to represent the measured data while $K < 1$ more strongly supports model 2.

The BIC can be expressed as (Wit et al., 2012)

$$\text{BIC} = n_p \ln(n_d) - 2 \ln(\mathcal{L}) \tag{4.6}$$

⁶The symbol for the marginal likelihood \mathcal{Z} is used to distinguish from the symbol for the impedance Z .

where n_d is the number of data points. When comparing models, the model with lower BIC is more strongly supported. Note that $\chi^2 = -2\ln(\mathcal{L})$, so the BIC is proportional to χ^2 .

4.3.2 Setup for Fitting Impedance Measurements at the Death Valley Site

Before presenting fit results using the one-layer and two-layer simulation sets described in the previous section, there are several important clarifications that must be made about the fits presented here. The simulation sets have been informed by previous sets of simulations, which have helped to develop a better understanding of the required spans for each parameter and the number of samples per parameter. The simulations used here do not necessarily reflect the final simulation set for the Death Valley site. The polynomial orders correspond to those resulting in the lowest deviations from the simulated impedances in Figures 4.4 and 4.5, or equivalently, the polynomial orders whose RMSE histograms have the majority of counts in low RMSE bins. The optimal polynomial orders used for the impedance fits in this section therefore correspond to third-order polynomials for each parameter in the one-layer model and [2, 1, 1, 1, 1] for the two-layer model. Determining the polynomial order that results in the most accurate impedances is a crucial step in reducing systematic errors that may be introduced into the impedance fits.

The error matrix Σ corresponds to the sum of data and interpolation errors according to Equation 4.3, where Σ_{meas} is assumed to be a diagonal matrix composed of a constant data

variance $\sigma_{\text{meas}}^2 = 1 \Omega^2$, and Σ_{interp} is computed according to Equation 4.4. The priors are uniformly distributed and bounded by the spans of each parameter in a given simulation set (that is, the full range of each parameter may be explored by the nested sampling algorithm).

The appearance of an apparent additional reactance in the Feko simulations has been discovered, which results in disagreement between the best-fit and measured impedance towards higher frequencies. The disagreement is caused by the different ways in which the wires connected to the panels are accounted for in the measured data and the simulations. The antenna impedance as measured by the VNA in the MIST receiver is affected by the impedance of the balun, which is connected to each panel by a copper wire. The antenna impedance presented in this section has had the contribution from the balun removed, which also accounts for the wires that connect the balun to the panels. As such, the measured antenna impedance is nearly perfectly referenced to the point where the wires connect to the panels.

However, the Feko simulations do not include the wires connecting the panels to the balun — instead, there are two wires connected to a voltage source (or excitation port) in between the two panels (Monsalve et al., 2023). This is done because it is simplest to have the excitation ports in the same location in Feko. This introduces an additional positive reactance from the wires which is present in the simulated (and therefore interpolated) impedances, but not the measured impedances.

Therefore, an inductive reactance $X_L = 2\pi fL$ is subtracted from Z_{interp} in the log-

likelihood in Equation 4.2, and the inductance L is considered an additional fit parameter, with a prior range from 1 μH to 1 nH. The inductance L is fit for instead of accounted for analytically because the connection of the wires from the balun to the panels is not always equivalent to laboratory measurements of the balun, which may change the measured impedance. Introducing L helps to improve the agreement between the best-fit and measured impedances, as well as to reduce bias introduced to the soil parameters from the inductance.

It is also important to note that the MIST instrument measures the antenna impedance once approximately every two hours. Measuring the impedance over such a timescale is useful for quantifying any changes in the impedance as a function of time, which could occur due to environmental changes that may cause, for example, fluctuations in the soil parameters due to rain or excessive heat. The fits presented in this chapter correspond to a single measurement of the antenna impedance. Fits in comparison with the average impedance for a set of data have also been explored, and the possibility of creating time-varying soil models has been considered.

4.3.3 Best-Fit Impedances and Soil Electrical Parameters for the Death Valley Site

Figure 4.8 displays the best-fit impedance for a one-layer soil model compared with the measured impedance at the Death Valley site. As previously mentioned, this fit corresponds to third-order polynomials for each parameter, inductive reactance correction, and an error

model considering both data error and interpolation error. I find that $\log(\mathcal{Z}) = -602$ and $\text{BIC} = 1190$. The best-fit conductivity σ is approximately 0.1 S/m. Note that the blue line in Figure 4.1 corresponds with $\sigma = 0.1$ S/m, and has significantly higher directivity at the beam zenith compared with the other lines, which correspond to values of σ that are one to two orders of magnitude lower. Therefore, an antenna measuring above a single soil layer with $\sigma = 0.1$ S/m is less sensitive deeper into the soil.

The relative permittivity ε is approximately 1, which is unusual for soil as $\varepsilon = 1$ is the relative permittivity of air. The inductance L is consistent with the self inductance of a 2 cm wire of 1 mm thickness, which corresponds to the wires connecting the panels to each excitation port in the Feko simulations (Monsalve et al., 2023). While the residuals at higher and very low frequencies are close to 0, those in transition periods of the impedance are larger, particularly between 50–70 MHz. The parameters and residuals may indicate that the one-layer model does not sufficiently model the measured impedance, or that the current model is not accounting for some unknown systematic.

Figure 4.9 displays the corner plot for the one-layer model fit shown in Figure 4.8. The sample distribution for the relative permittivity ε is clustered at the lower limit of the prior range, indicating that a value consistent with that of air is preferred. It is possible that the low relative permittivity is due to as yet unknown systematic effects, or that the effective relative permittivity of the soil is in fact consistent with air. The distributions of the conductivity σ and inductance L are approximately Gaussian-distributed. The parameters

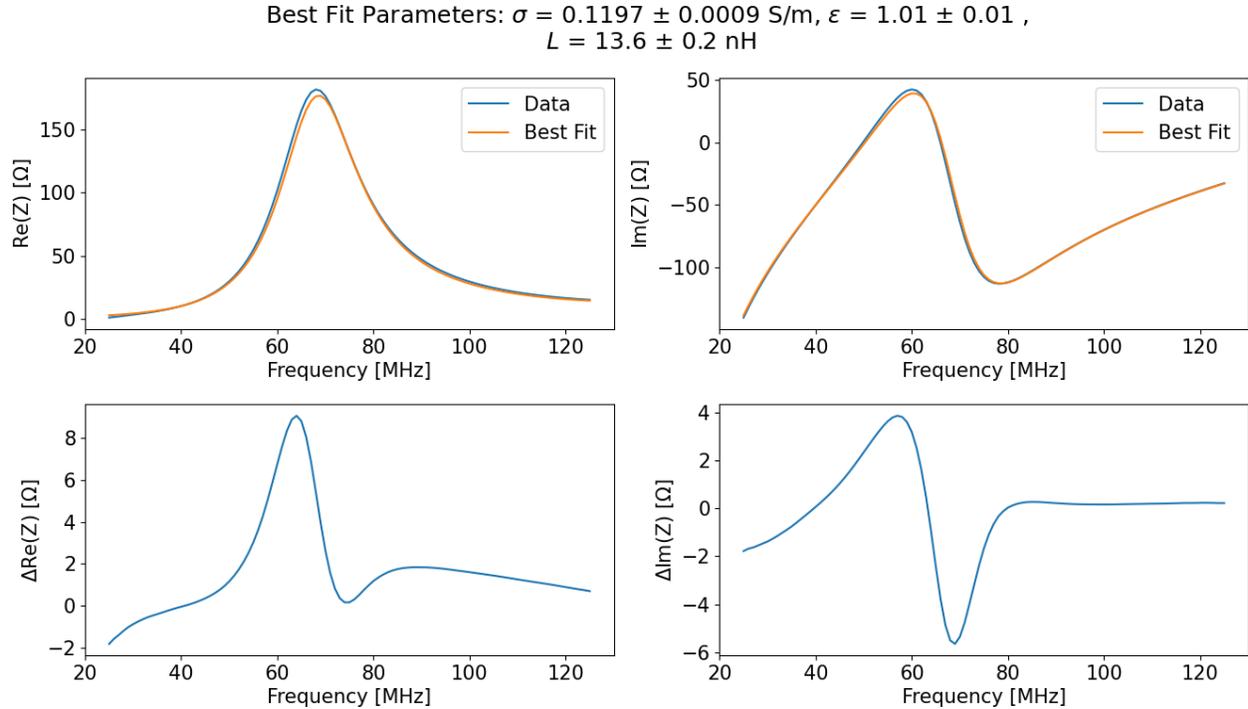


Figure 4.8: The best-fit impedance for the one-layer model compared with the measured impedance at Death Valley. The value of the conductivity σ corresponds with that of the blue line in Figure 4.1, meaning that more of the beam is directed towards the sky than lower σ values, such that the antenna is less sensitive to deeper soil. The relative permittivity ϵ is consistent with air, possibly indicating that there may be systematic effects that are not accounted for in the simulations. The inductance L is consistent with the expected inductance of the wires connecting the panels to the excitation port in the simulations. I find $\log(\mathcal{Z}) = -602$ and $\text{BIC} = 1190$ for this fit.

appear uncorrelated with each other.

Figure 4.10 displays the best-fit impedance for a two-layer model compared with the measured impedance at the Death Valley site. The polynomial order used for the impedance interpolations corresponds to $[2, 1, 1, 1, 1]$, and both interpolation and data errors are accounted for as with the one-layer model. The best-fit conductivity σ_1 and relative permittivity ε_1 of the first layer and the inductance L are consistent with the best-fit parameters for the one-layer result. The residuals are also consistent between the one-layer and two-layer fits, which may indicate that the parameters of the first layer dominate the effect on the impedance. I find that $\log(\mathcal{Z}) = -597$ and $\text{BIC} = 1200$ for this fit.

Figure 4.11 displays the corner plot corresponding to the two-layer model fit shown in Figure 4.10. The sample distributions corresponding to correlations between σ_1 , ε_1 , and L are consistent with the distributions shown for the one-layer fit in Figure 4.9. The best-fit conductivity σ_2 , relative permittivity ε_2 , and depth d of the second layer correspond to the center of the spans for each parameter, indicating that the parameters have not converged.

The Bayes factor for the two models is $K = \mathcal{Z}_1/\mathcal{Z}_2 = 0.007$, where \mathcal{Z}_1 and \mathcal{Z}_2 correspond to the marginal likelihoods for the one-layer and two-layer models, respectively. $K = 0.007$ more strongly supports the two-layer model as the best model for the data, though K does not necessarily suggest that the specific two-layer model used here is the *best* model. I find that the $\text{BIC} = 1190$ for the one-layer model, which corresponds to $\chi^2 = 1174$, and

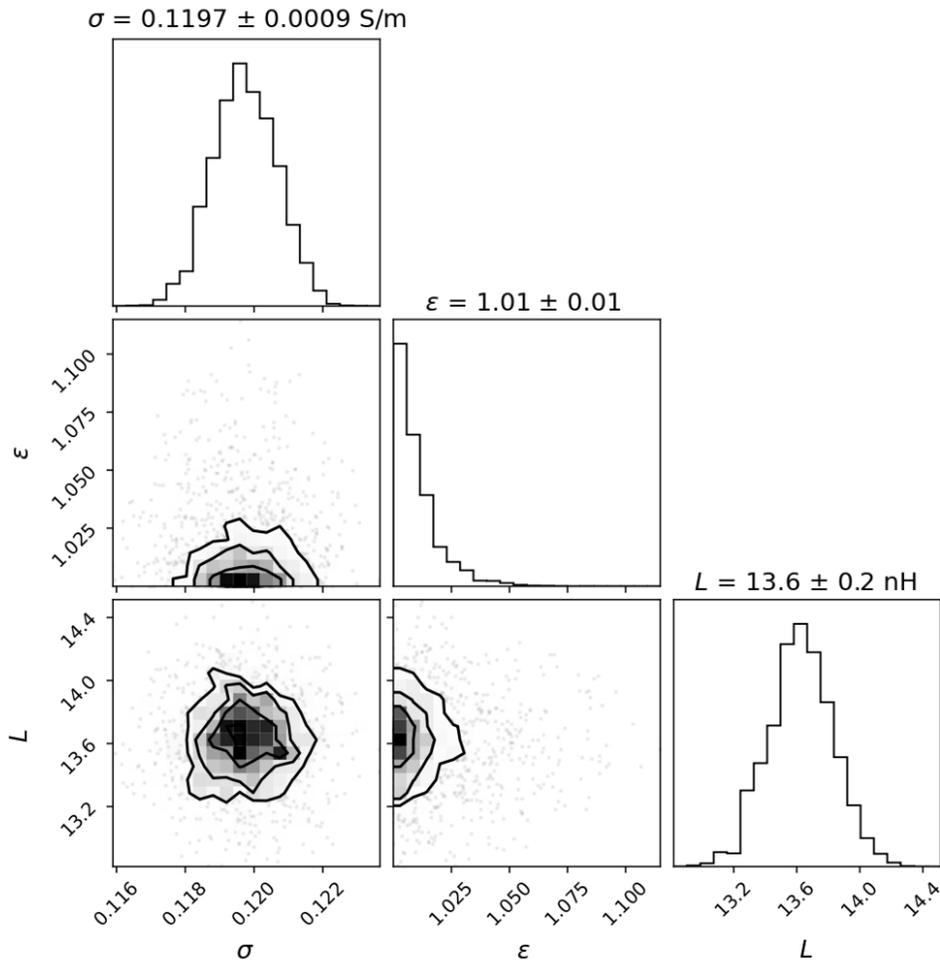


Figure 4.9: Corner plot corresponding to the one-layer model fit. The distributions of the samples for the conductivity σ and the inductance L appear Gaussian-distributed. The samples for the relative permittivity ϵ are clustered at 1, which is consistent with the relative permittivity of air. Such a low permittivity could be indicative of systematic effects that have not been accounted for, or that the effective relative permittivity of the soil is indeed consistent with air. The parameters appear uncorrelated with each other.

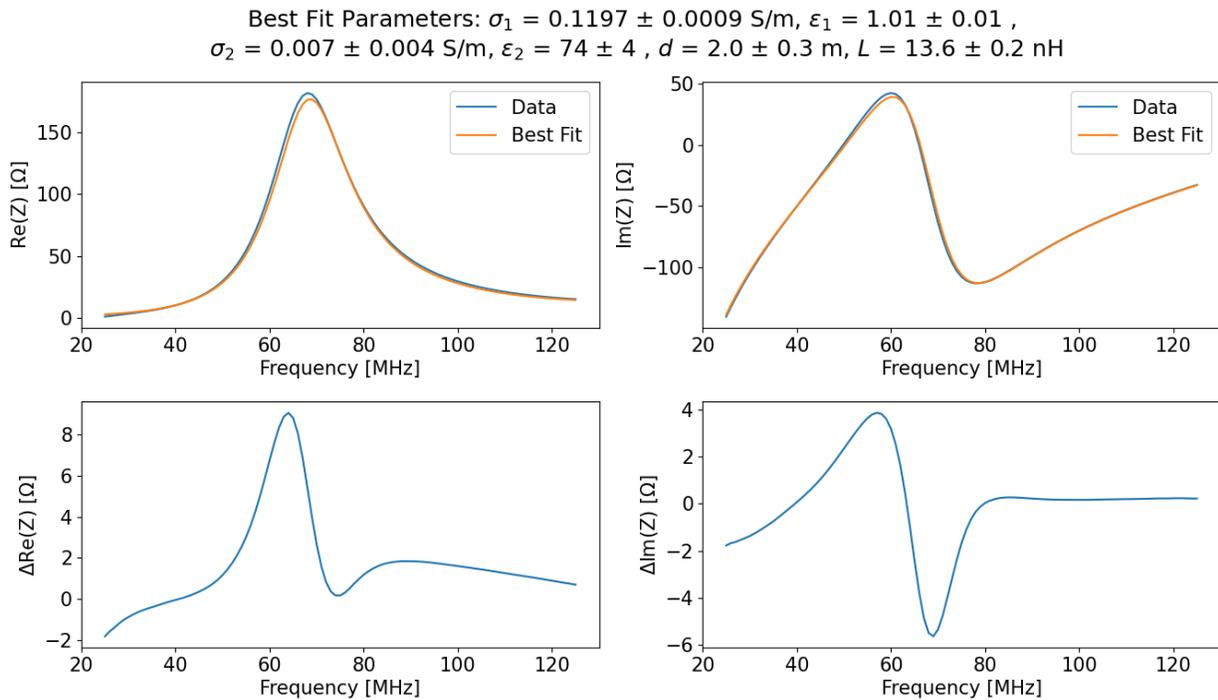


Figure 4.10: The best-fit impedance for the two-layer model compared with the measured impedance at Death Valley. The best-fit conductivity σ_1 and relative permittivity ϵ_1 of the first layer and the inductance L are consistent with the best fit values for the one-layer model. The best-fit conductivity σ_2 , relative permittivity ϵ_2 , and depth d of the second layer correspond to the center of the spans for each parameter. I find $\log(\mathcal{Z}) = -597$ and $\text{BIC} = 1200$ for this fit.

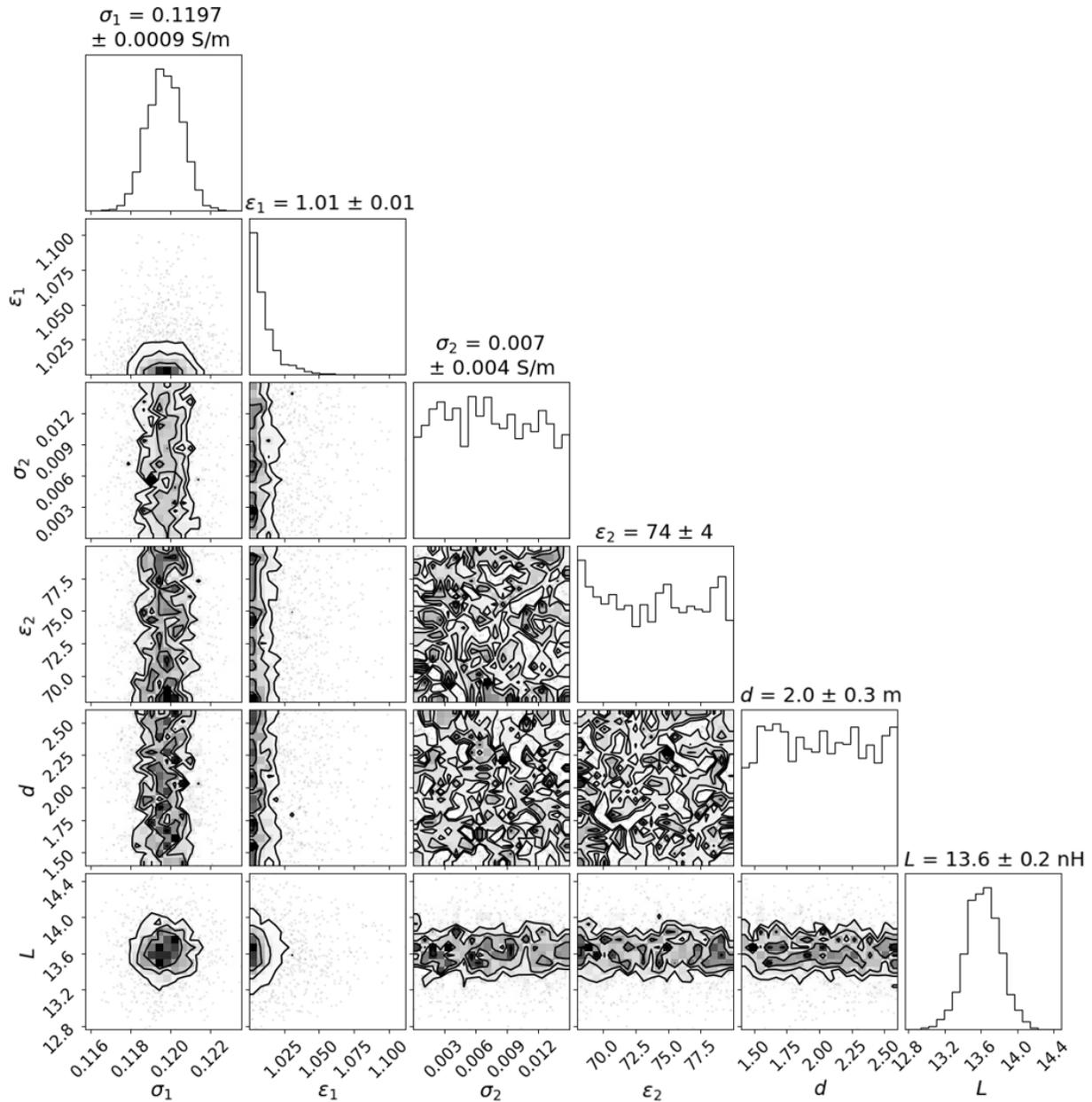


Figure 4.11: Corner plot corresponding to the two-layer model fit. The distributions of the correlations for σ_1 , ϵ_1 , and L are similar to those for the distributions of σ , ϵ , and L shown in Figure 4.9, while the distributions for σ_2 , ϵ_2 , and d explore the full parameter space, suggesting that these parameters have not converged. As such, the results for the second-layer parameters are not reliable estimates of the second-layer soil properties.

BIC = 1200 for the two-layer model, which corresponds to $\chi^2 = 1173$. As such, there is only marginal improvement in the best-fit impedance for the two-layer model. The best-fit parameters of the second layer in the two-layer model did not converge, and therefore cannot be considered an accurate model of the second layer (if there is indeed a second layer). The span of the second layer relative permittivity has previously been expanded, which resulted in similar values for the best-fit parameters.

It is possible that the best-fit impedance and first-layer parameters for the two-layer model are so consistent with the parameters of the one-layer model because the effect of the first layer dominates, such that the beam is only minimally affected by the second layer. Recall that a conductivity $\sigma = 0.1$ S/m for a one-layer soil model corresponds with the blue line in the left panel of Figure 4.1. The directivity at zenith for $\sigma = 0.1$ S/m is greater by 1–3 than directivities corresponding with values of σ that are orders of magnitude lower. Therefore, $\sigma = 0.1$ S/m directs more of the beam towards the sky, meaning that it is less sensitive to deeper soil. As such, if the first-layer conductivity dominates, it will prove challenging to recover the second-layer parameters using measurements of the MIST antenna. It is therefore crucial to use other techniques to measure the soil electrical properties in order to compare with the results from SOIL and understand the soil environment. For example, at the Deep Springs and Death Valley sites, only four-wire measurements have been conducted. At MARS, the soil has been measured with GPR, and impedance measurements have been made using a separate configuration of the MIST antenna and a small monopole antenna.

AMPERE and SOIL can be applied to these impedance measurements to see if the best-fit soil parameters are consistent between different antennas.

4.4 Summary

In this chapter, I discussed challenges related to beam chromaticity for global 21-cm experiments, and how the use of ground planes can produce difficult-to-model effects in the beam. The MIST experiment implements the novel approach of operating without a ground plane, dedicating calibration efforts to understanding the effect of soil electrical properties on the beam. I then introduced the AMPERE and SOIL packages, which work together to fit the measured impedance of the MIST antenna using EM simulations. By determining the parameters corresponding to the best-fit impedance, a model for the properties of the soil at each MIST observing site can in principle be estimated. The best soil model can then be used to calculate a beam model for the MIST antenna.

I presented the best-fit impedance and soil parameters corresponding to one-layer and two-layer soil models for the Death Valley site, and found that the two-layer model is preferred according to the Bayes factor $K = 0.007$. However, the first-layer soil parameters of the two-layer model are consistent with the soil parameters of the one-layer model, and the parameters for the second layer of the two-layer model did not converge. The results may suggest that the first layer dominates the effect on the impedance, which prevents convergence of the second-layer parameters. Other soil characterization techniques such as

four-wire methods and ground penetrating radar can provide independent comparisons of results between different methods, which will help to inform the models used in AMPERE and SOILand gain more insight into the soil environment. Future work will be dedicated to conducting similar analyses for other MIST observing sites, improving the simulation models and methods used in AMPERE and SOIL, and investigating other soil analysis techniques.

Chapter 5

Summary and Future Work

The 21-cm signal of neutral hydrogen is a novel tool well-suited to enable the next generation of cosmological study. The monopole component of the 21-cm signal, referred to as the global 21-cm signal, can be used to study the conditions of the early universe, corresponding to the periods before, during, and after the birth of the first luminous objects. Two experiments having claimed detection of the global 21-cm signal present conflicting results, making follow-up measurements by independent experiments crucial to understanding the exact nature of the early universe. The spatial fluctuations in the 21-cm signal can also be studied in principle from the dark ages to the present. The spatial fluctuations of the 21-cm signal trace the distribution of matter in the universe, and can thus be used to measure the evolution of the large scale structure. In particular, by using intensity mapping methods to study 21-cm signal fluctuations as a proxy for the distribution of matter, one can make sensitive probes

of dark energy. The instruments and analysis methods vary widely depending on which aspect of the 21-cm signal is being observed, although they both face similar significant challenges in characterizing large Galactic foregrounds and extreme control of instrument systematics. In this thesis, I have presented research related to several topics in 21-cm cosmology, corresponding to experiments measuring different aspects of the 21-cm signal and spanning nearly the entire redshift range below $z = 55$.

I first discussed research for the Canadian Hydrogen Intensity Mapper and Radio transient Detector (CHORD), a next-generation radio interferometer that will study fluctuations in the 21-cm signal at $z < 3.7$. With a focus on redundancy and systematic control, CHORD will make the largest 3D map of the universe to date through the technique of intensity mapping, using baryon acoustic oscillations (BAOs) as a standard ruler to monitor the expansion rate of the universe to probe the nature of dark energy. CHORD will also observe tens of thousands of fast radio bursts (FRBs), and enable localization precision to within $\lesssim 10$ mas using outrigger stations separated from the core array by continental baselines.

In Chapter 2, I presented the Deep Dish Development Array (D3A), a testbed for the prototyping of CHORD technologies, including the CHORD dishes, feed, LNA, and digital readout system. I then presented a measurement of the D3A system temperature using the principle of phase-closure for a three-element interferometer. I found that the West dish has a consistently higher system temperature across the D3A bandwidth than the

Mid and East dishes, and in general is inconsistent with the performance of the other two dishes. The differences observed with the West dish are of particular concern, as the goal of D3A is to demonstrate that the technologies developed for CHORD can meet the CHORD redundancy and systematic control requirements. It is possible that the differences result from the fact that the West dish is equipped with the first CHORD feed manufactured, where it accumulated excess solder during assembly, or that the pin connecting the balun to the feed has degraded over time. Future work will be dedicated to understanding the source of the differences between the dishes and improving their redundancy.

I also found that the D3A system temperature is above the 30 K CHORD target by tens of K in most of the D3A band. Recall that the analog-to-digital converter (ADC) at D3A samples at 800 MSPS, meaning that the signal has to be split into different Nyquist zones using band-pass filters to recover the full D3A band. In the 800–1200 MHz band, the high system temperature is likely caused by additional aliased noise power from out-of-band frequencies due to the slow roll-off of the band-pass filter. In the 1200–1600 MHz band, the band-pass filter attenuates in-band signals towards higher frequencies, which increases the system temperature contribution from upstream components.

However, the system temperature in the low band (400–800 MHz) is more consistent with what is expected from simulations: the system temperature rises below ~ 600 MHz as the sky becomes brighter, but above ~ 600 MHz, the system temperature is consistent with 30 K for the Mid and East dishes. If the assumption is valid that the high system temperature

observed in the 800–1200 MHz and 1200–1600 MHz bands is caused by the filters, better performance is expected if the current filters are replaced by filters with sharper roll-off at the Nyquist frequencies at the cut-off frequencies of each band. While the additional power from aliasing can in principle be accounted for, a superior measurement can in principle be made with an RF-System-on-Chip (RFSoc), which can sample at the Nyquist sampling rate for D3A and therefore recover the entire band without the need for band-splitting. The signal processing firmware for the RFSoc is currently in development, and an RFSoc will be instrumented at D3A to make improved measurements of its performance.

In Chapter 3, I discussed two potential designs for the CHORD F-Engine, which will be responsible for the digitization, channelization, and corner-turn operation of input signals. One design would implement slight changes to the ICE system used in the CHIME F-Engine, and would require band-pass filtering to recover the full CHORD band similar to the method currently used at D3A. However, band-pass filtering can introduce undesirable effects that require careful attention. As such, another possible design for the CHORD F-Engine is to use a platform based on an RFSoc, which does not require band-pass filtering and can recover the full CHORD band directly, leading to a smoother and more stable gain in the signal chain. However, the RFSoc is an entirely new system that requires independent characterization of its performance to demonstrate its viability as a CHORD F-Engine platform.

I presented measurements characterizing the ADC of the AMD Zynq Ultrascale+ ZU47DR RFSoc mounted on a t_0 Control and Readout System (CRS) board. I find that

the RFSoc ADC generally demonstrates consistency with the datasheet, and vastly outperforms the EV8AQ160 ADC used on the ICE boards. The RFSoc ADC has a greater dynamic range by nearly 20 dB, lower distortion, and improved or consistent linearity. In particular, the RFSoc demonstrates superior crosstalk performance, which is crucial for digital systems in radio astronomy that correlate signals from multiple antennas. I find that the worst RFSoc crosstalk is better by 26 dB when comparing with the crosstalk measured between two channels on the same ADC chip in the ICE system, and by 6 dB when comparing with the crosstalk measured between two channels on separate mezzanines in the ICE system. Note that the crosstalk improves with channel separation, so CHORD can mitigate ADC crosstalk even further by alternating antenna polarizations for each ADC input. As previously mentioned, the RFSoc firmware is currently under development, and future work will be dedicated to demonstrating its on-sky performance of the RFSoc at D3A.

In this thesis, I also presented research for the Mapper of the IGM Spin Temperature (MIST), which aims to detect the global 21-cm signal at $55 > z > 13$ from remote radio-quiet locations. In Chapter 4, I discussed chromaticity challenges for global 21-cm experiments, and how the presence of ground planes can lead to difficult-to-model structure in the antenna beam. As such, the MIST experiment implements the novel approach of operating its instrument directly above the soil with no ground plane, which has the advantage of removing the sharp electrical discontinuities associated with ground planes.

However, by operating directly above the soil, the MIST beam becomes coupled to the soil electrical properties, which are different for each observing site. Therefore, MIST dedicates effort to exploring techniques for characterizing the soil electrical properties associated with each observation site in order to inform simulations of the MIST beam about the antenna environment.

I presented the Antenna Modeler and Performance Emulator (AMPERE), a Python-based package that interpolates various antenna quantities using EM simulations. I then evaluated the performance of AMPERE with two sets of simulations of the MIST antenna impedance generated using the Feko software: one set corresponding to a one-layer soil model, and another set corresponding to a two-layer soil model. I find that the accuracy of the interpolations is dependent upon the order of the polynomials used to fit each parameter, and that the accuracy improves by orders of magnitude with more samples per parameter. However, there are practical limitations associated with generating simulations with more samples for parameters as the number of soil layers increases. By choosing the optimal polynomial order, I have demonstrated that AMPERE can achieve accuracy within a fraction of 1Ω for impedances oscillating over a range of approximately 150–200 Ω .

I then introduced the System for Optimizing Impedance Likelihood (SOIL), a companion package to AMPERE. SOIL manages fits to the measured impedance of the MIST antenna using the `ultranest` Bayesian inference package and models generated by AMPERE. I presented the best-fit impedance and soil parameters for observations conducted at Death

Valley, USA in May 2022 using one- and two-layer soil models and the optimal polynomial orders for each simulation set.

I found that while the two-layer model is slightly preferred, the soil parameters associated with the second layer did not converge, and the parameters of the first layer were consistent with the parameters for the one-layer model. It is possible that the MIST beam at the Death Valley site was not sensitive deep into the soil. The best-fit conductivity was approximately 0.1 S/m, which significantly increases the directivity towards the zenith compared with conductivity values that are orders of magnitude lower. Further, the best-fit relative permittivity of the one-layer model and the top layer of the two-layer model was found to be consistent with the relative permittivity of air, which may suggest that there are unknown systematics in the instrument or the simulations that need to be addressed. Therefore, the MIST experiment will dedicate future research to improving the simulations used in AMPERE and SOIL, and explore other techniques and analysis methods for characterizing the soil electrical properties.

Bibliography

E. Abdalla, E. G. M. Ferreira, R. G. Landim, A. A. Costa, K. S. F. Fornazier, F. B. Abdalla, L. Barosi, F. A. Brito, A. R. Queiroz, T. Villela, B. Wang, C. A. Wuensche, A. Marins, C. P. Novaes, V. Liccardo, C. Shan, J. Zhang, Z. Zhang, Z. Zhu, I. Browne, J. Delabrouille, L. Santos, M. V. dos Santos, H. Xu, S. Anton, R. Battye, T. Chen, C. Dickinson, Y.-Z. Ma, B. Maffei, E. J. de Mericia, P. Motta, C. H. N. Otobone, M. W. Peel, S. Roychowdhury, M. Remazeilles, R. M. Ribeiro, Y. Sang, J. R. L. Santos, J. F. R. dos Santos, G. B. Silva, F. Vieira, J. Vieira, L. Xiao, X. Zhang, and Y. Zhu. The BINGO project. *Astronomy & Astrophysics*, 664:A14, Aug. 2022. doi: 10.1051/0004-6361/202140883. URL <https://doi.org/10.1051/0004-6361/202140883>.

AMD Inc. *Zynq UltraScale+ RFSoc Data Sheet: DC and AC Switching Characteristics*. AMD Inc., 2022a. URL <https://docs.xilinx.com/r/en-US/ds926-zynq-ultrascale-plus-rfsoc>.

AMD Inc. *Zynq UltraScale+ RFSoc Product Data Sheet: Overview (DS889)*. AMD Inc., 2022b. URL <https://docs.xilinx.com/v/u/en-US/ds889-zynq-usp-rfsoc-overview>.

M. Amiri, K. Bandura, A. Boskovic, T. Chen, J.-F. Cliche, M. Deng, N. Denman, M. Dobbs, M. Fandino, S. Foreman, M. Halpern, D. Hanna, A. S. Hill, G. Hinshaw, C. Höfer, J. Kania, P. Klages, T. L. Landecker, J. MacEachern, K. Masui, J. Mena-Parra, N. Milutinovic, A. Mirhosseini, L. Newburgh, R. Nitsche, A. Ordog, U.-L. Pen, T. Pinsonneault-Marotte, A. Polzin, A. Reda, A. Renard, J. R. Shaw, S. R. Siegel, S. Singh, R. Smegal, I. Tretyakov, K. V. Gassen, K. Vanderlinde, H. Wang, D. V. Wiebe, J. S. Willis, and D. Wulf. An Overview of CHIME, the Canadian Hydrogen Intensity Mapping Experiment. *The Astrophysical Journal Supplement Series*, 261(2):29, July 2022. doi: 10.3847/1538-4365/ac6fd9. URL <https://doi.org/10.3847/1538-4365/ac6fd9>.

M. Amiri, K. Bandura, T. Chen, M. Deng, M. Dobbs, M. Fandino, S. Foreman, M. Halpern, A. S. Hill, G. Hinshaw, C. Höfer, J. Kania, T. L. Landecker, J. MacEachern, K. Masui, J. Mena-Parra, N. Milutinovic, A. Mirhosseini, L. Newburgh, A. Ordog, U.-L. Pen, T. Pinsonneault-Marotte, A. Polzin, A. Reda, A. Renard, J. R. Shaw, S. R. Siegel, S. Singh, K. Vanderlinde, H. Wang, D. V. Wiebe, and D. Wulf. Detection of Cosmological 21 cm Emission with the Canadian Hydrogen Intensity Mapping Experiment. *The Astrophysical Journal*, 947(1):16, Apr. 2023. doi: 10.3847/1538-4357/acb13f. URL <https://doi.org/10.3847/1538-4357/acb13f>.

- and David J. Bacon, R. A. Battye, P. Bull, S. Camera, P. G. Ferreira, I. Harrison, D. Parkinson, A. Pourtsidou, M. G. Santos, L. Wolz, F. Abdalla, Y. Akrami, D. Alonso, S. Andrianomena, M. Ballardini, J. L. Bernal, D. Bertacca, C. A. P. Bengaly, A. Bonaldi, C. Bonvin, M. L. Brown, E. Chapman, S. Chen, X. Chen, S. Cunnington, T. M. Davis, C. Dickinson, J. Fonseca, K. Grainge, S. Harper, M. J. Jarvis, R. Maartens, N. Maddox, H. Padmanabhan, J. R. Pritchard, A. Raccanelli, M. Rivi, S. Roychowdhury, M. Sahlén, D. J. Schwarz, T. M. Siewert, M. Viel, F. Villaescusa-Navarro, Y. Xu, D. Yamauchi, and J. Zuntz. Cosmology with Phase 1 of the Square Kilometre Array Red Book 2018: Technical specifications and performance forecasts. *Publications of the Astronomical Society of Australia*, 37, 2020. doi: 10.1017/pasa.2019.51. URL <https://doi.org/10.1017/pasa.2019.51>.
- B. C. Andersen, K. M. Bandura, M. Bhardwaj, A. Bij, M. M. Boyce, P. J. Boyle, C. Brar, T. Cassanelli, P. Chawla, T. Chen, J.-F. Cliche, A. Cook, D. Cubranic, A. P. Curtin, N. T. Denman, M. Dobbs, F. Q. Dong, M. Fandino, E. Fonseca, B. M. Gaensler, U. Giri, D. C. Good, M. Halpern, A. S. Hill, G. F. Hinshaw, C. Höfer, A. Josephy, J. W. Kania, V. M. Kaspi, T. L. Landecker, C. Leung, D. Z. Li, H.-H. Lin, K. W. Masui, R. Mckinven, J. Mena-Parra, M. Merryfield, B. W. Meyers, D. Michilli, N. Milutinovic, A. Mirhosseini, M. Münchmeyer, A. Naidu, L. B. Newburgh, C. Ng, C. Patel, U.-L. Pen, T. Pinsonneault-Marotte, Z. Pleunis, B. M. Quine, M. Rafiei-Ravandi, M. Rahman, S. M. Ransom, A. Renard, P. Sanghavi, P. Scholz, J. R. Shaw, K. Shin, S. R. Siegel, S. Singh, R. J. Smegal,

- K. M. Smith, I. H. Stairs, C. M. Tan, S. P. Tendulkar, I. Tretyakov, K. Vanderlinde, H. Wang, D. Wulf, A. V. Zwaniga, and T. C. Collaboration. A bright millisecond-duration radio burst from a galactic magnetar. *Nature*, 587(7832):54–58, Nov 2020. ISSN 1476-4687. doi: 10.1038/s41586-020-2863-y. URL <https://doi.org/10.1038/s41586-020-2863-y>.
- C. J. Anderson, N. J. Luciw, Y. C. Li, C. Y. Kuo, J. Yadav, K. W. Masui, T. C. Chang, X. Chen, N. Oppermann, Y. W. Liao, U. L. Pen, D. C. Price, L. Staveley-Smith, E. R. Switzer, P. T. Timbie, and L. Wolz. Low-amplitude clustering in low-redshift 21-cm intensity maps cross-correlated with 2dF galaxy densities. *Monthly Notices of the Royal Astronomical Society*, 476(3):3382–3392, May 2018. doi: 10.1093/mnras/sty346.
- L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, A. S. Bolton, J. Brinkmann, J. R. Brownstein, A. Burden, A. J. Cuesta, L. A. N. da Costa, K. S. Dawson, R. de Putter, D. J. Eisenstein, J. E. Gunn, H. Guo, J.-C. Hamilton, P. Harding, S. Ho, K. Honscheid, E. Kazin, D. Kirkby, J.-P. Kneib, A. Labatie, C. Loomis, R. H. Lupton, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, C. K. McBride, K. T. Mehta, O. Mena, F. Montesano, D. Muna, R. C. Nichol, S. E. Nuza, M. D. Olmstead, D. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. Parejko, I. Paris, W. J. Percival, P. Petitjean, F. Prada, B. Reid, N. A. Roe, A. J. Ross, N. P. Ross, L. Samushia, A. G. Sanchez, D. J. Schlegel, D. P. Schneider, C. G. Scoccola, H.-J. Seo, E. S. Sheldon, A. Simmons, R. A. Skibba, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magana, L. Verde, C. Wagner, D. A. Wake, B. A. Weaver,

- D. H. Weinberg, M. White, X. Xu, C. Yeche, I. Zehavi, and G.-B. Zhao. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 427(4):3435–3467, Dec. 2012. doi: 10.1111/j.1365-2966.2012.22066.x. URL <https://doi.org/10.1111%2Fj.1365-2966.2012.22066.x>.
- R. Ansari, J. E. Campagne, P. Colom, J. M. L. Goff, C. Magneville, J. M. Martin, M. Moniez, J. Rich, and C. Yè che. 21 cm observation of large-scale structures at $z \sim 1$. Instrument sensitivity and foreground subtraction. *Astronomy & Astrophysics*, 540: A129, Apr. 2012. doi: 10.1051/0004-6361/201117837. URL <https://doi.org/10.1051%2F0004-6361%2F201117837>.
- S. D. Bale, N. Bassett, J. O. Burns, J. D. Jones, K. Goetz, C. Hellum-Bye, S. Hermann, J. Hibbard, M. Maksimovic, R. McLean, R. Monsalve, P. O’Connor, A. Parsons, M. Pulupa, R. Pund, D. Rapetti, K. M. Rotermund, B. Saliwanchik, A. Slosar, D. Sundkvist, and A. Suzuki. LuSEE ’Night’: The Lunar Surface Electromagnetics Experiment. arXiv 2301.10345, 2023.
- K. Bandura, A. N. Bender, J. F. Cliche, T. de Haan, M. A. Dobbs, A. J. Gilbert, S. Griffin, G. Hsyu, D. Ittah, J. M. Parra, J. Montgomery, T. Pinsonneault-Marotte, S. Siegel, G. Smecher, Q. Y. Tang, K. Vanderlinde, and N. Whitehorn. ICE: A Scalable, Low-Cost FPGA-Based Telescope Signal Processing and Networking System. *Journal of*

- Astronomical Instrumentation*, 05(04):1641005, 2016. doi: 10.1142/S2251171716410051.
URL <https://doi.org/10.1142/S2251171716410051>.
- R. Barkana. Possible interaction between baryons and dark-matter particles revealed by the first stars. *Nature*, 555(7694):71–74, Mar. 2018. doi: 10.1038/nature25791. URL <https://doi.org/10.1038/nature25791>.
- A. P. Beardsley, M. Johnston-Hollitt, C. M. Trott, J. C. Pober, J. Morgan, D. Oberoi, D. L. Kaplan, C. R. Lynch, G. E. Anderson, P. I. McCauley, S. Croft, C. W. James, O. I. Wong, C. D. Tremblay, R. P. Norris, I. H. Cairns, C. J. Lonsdale, P. J. Hancock, B. M. Gaensler, N. D. R. Bhat, W. Li, N. Hurley-Walker, J. R. Callingham, N. Seymour, S. Yoshiura, R. C. Joseph, K. Takahashi, M. Sokolowski, J. C. A. Miller-Jones, J. V. Chauhan, I. Bojičić, M. D. Filipović, D. Leahy, H. Su, W. W. Tian, S. J. McSweeney, B. W. Meyers, S. Kitaeff, T. Vernstrom, G. Gürkan, G. Heald, M. Xue, C. J. Riseley, S. W. Duchesne, J. D. Bowman, D. C. Jacobs, B. Crosse, D. Emrich, T. M. O. Franzen, L. Horsley, D. Kenney, M. F. Morales, D. Pallot, K. Steele, S. J. Tingay, M. Walker, R. B. Wayth, A. Williams, and C. Wu. Science with the Murchison Widefield Array: Phase I results and Phase II opportunities. *Publications of the Astronomical Society of Australia*, 36:e050, Dec. 2019. doi: 10.1017/pasa.2019.41.
- F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson. The 6dF Galaxy Survey: baryon acoustic oscillations and

- the local Hubble constant. *Monthly Notices of the Royal Astronomical Society*, 416(4): 3017–3032, July 2011. doi: 10.1111/j.1365-2966.2011.19250.x. URL <https://doi.org/10.1111%2Fj.1365-2966.2011.19250.x>.
- C. Blake, E. A. Kazin, F. Beutler, T. M. Davis, D. Parkinson, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, D. Croton, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I. hui Li, B. Madore, D. C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee. The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations. *Monthly Notices of the Royal Astronomical Society*, 418(3):1707–1724, Oct. 2011. doi: 10.1111/j.1365-2966.2011.19592.x. URL <https://doi.org/10.1111%2Fj.1365-2966.2011.19592.x>.
- C. D. Bochenek, V. Ravi, K. V. Belov, G. Hallinan, J. Kocz, S. R. Kulkarni, and D. L. McKenna. A fast radio burst associated with a Galactic magnetar. *Nature*, 587(7832): 59–62, Nov. 2020. doi: 10.1038/s41586-020-2872-x.
- J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen, and N. Mahesh. An absorption profile centred at 78 megahertz in the sky-averaged spectrum. *Nature*, 555 (7694):67–70, Mar 2018. ISSN 1476-4687. doi: 10.1038/nature25792. URL <https://doi.org/10.1038/nature25792>.
- R. F. Bradley, K. Tauscher, D. Rapetti, and J. O. Burns. A Ground Plane Artifact that

- Induces an Absorption Profile in Averaged Spectra from Global 21 cm Measurements, with Possible Application to EDGES. *The Astrophysical Journal*, 874(2):153, Apr. 2019. doi: 10.3847/1538-4357/ab0d8b. URL <https://doi.org/10.3847%2F1538-4357%2Fab0d8b>.
- J. Buchner. A statistical test for nested sampling algorithms. *Statistics and Computing*, 26(1-2):383–392, Sept. 2014. doi: 10.1007/s11222-014-9512-y. URL <https://doi.org/10.1007%2Fs11222-014-9512-y>.
- J. Buchner. Collaborative Nested Sampling: Big Data versus Complex Physical Models. *Publications of the Astronomical Society of the Pacific*, 131(1004):108005, Aug. 2019. doi: 10.1088/1538-3873/aae7fc. URL <https://doi.org/10.1088%2F1538-3873%2Faae7fc>.
- J. Buchner. Ultranest – a robust, general purpose bayesian inference engine. arXiv 2101.09604, 2021.
- J. O. Burns. Transformative science from the lunar farside: observations of the dark ages and exoplanetary systems at low radio frequencies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2188):20190564, 2021. doi: 10.1098/rsta.2019.0564. URL <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2019.0564>.
- T.-C. Chang, U.-L. Pen, K. Bandura, and J. B. Peterson. An intensity map of hydrogen 21-cm emission at redshift $z \sim 0.8$. *Nature*, 466(7305):463–465, July 2010. doi: 10.1038/nature09187.

- CHIME/FRB Collaboration, M. Amiri, B. C. Andersen, K. Bandura, S. Berger, M. Bhardwaj, M. M. Boyce, P. J. Boyle, C. Brar, D. Breitman, T. Cassanelli, P. Chawla, T. Chen, J. F. Cliche, A. Cook, D. Cubranic, A. P. Curtin, M. Deng, M. Dobbs, F. A. Dong, G. Eadie, M. Fandino, E. Fonseca, B. M. Gaensler, U. Giri, D. C. Good, M. Halpern, A. S. Hill, G. Hinshaw, A. Josephy, J. F. Kaczmarek, Z. Kader, J. W. Kania, V. M. Kaspi, T. L. Landecker, D. Lang, C. Leung, D. Li, H.-H. Lin, K. W. Masui, R. McKinven, J. Mena-Parra, M. Merryfield, B. W. Meyers, D. Michilli, N. Milutinovic, A. Mirhosseini, M. Münchmeyer, A. Naidu, L. Newburgh, C. Ng, C. Patel, U.-L. Pen, E. Petroff, T. Pinsonneault-Marotte, Z. Pleunis, M. Rafiei-Ravandi, M. Rahman, S. M. Ransom, A. Renard, P. Sanghavi, P. Scholz, J. R. Shaw, K. Shin, S. R. Siegel, A. E. Sikora, S. Singh, K. M. Smith, I. Stairs, C. M. Tan, S. P. Tendulkar, K. Vanderlinde, H. Wang, D. Wulf, and A. V. Zwaniga. The First CHIME/FRB Fast Radio Burst Catalog. *The Astrophysical Journal Supplement Series*, 257(2):59, Dec. 2021. doi: 10.3847/1538-4365/ac33ab.
- S. Cole, W. J. Percival, J. A. Peacock, P. Norberg, C. M. Baugh, C. S. Frenk, I. Baldry, J. Bland-Hawthorn, T. Bridges, R. Cannon, M. Colless, C. Collins, W. Couch, N. J. G. Cross, G. Dalton, V. R. Eke, R. D. Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, K. Glazebrook, C. Jackson, A. Jenkins, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, B. A. Peterson, W. Sutherland, and K. T. and. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 362(2):505–534, Sept. 2005. doi:

10.1111/j.1365-2966.2005.09318.x. URL <https://doi.org/10.1111%2Fj.1365-2966.2005.09318.x>.

Condon, James J. and Ransom, Scott M. *Essential Radio Astronomy*. Princeton University Press, 2016. ISBN 978-0691137797.

E. de Lera Acedo. REACH: Radio Experiment for the Analysis of Cosmic Hydrogen. In *2019 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pages 0626–0629, 2019. doi: 10.1109/ICEAA.2019.8879199.

D. R. DeBoer, A. R. Parsons, J. E. Aguirre, P. Alexander, Z. S. Ali, A. P. Beardsley, G. Bernardi, J. D. Bowman, R. F. Bradley, C. L. Carilli, C. Cheng, E. de Lera Acedo, J. S. Dillon, A. Ewall-Wice, G. Fadana, N. Fagnoni, R. Fritz, S. R. Furlanetto, B. Glendenning, B. Greig, J. Grobbelaar, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, M. Kariseb, S. A. Kohn, T. Lecalake, A. Liu, A. Loots, D. MacMahon, L. Malan, C. Malgas, M. Maree, Z. Martinot, N. Mathison, E. Matsetela, A. Mesinger, M. F. Morales, A. R. Neben, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, R. Sell, C. Smith, A. Syce, M. Tegmark, N. Thyagarajan, P. K. G. Williams, and H. Zheng. Hydrogen Epoch of Reionization Array (HERA). *Publications of the Astronomical Society of the Pacific*, 129(974):045001, Apr. 2017. doi: 10.1088/1538-3873/129/974/045001.

- Digital Research Alliance of Canada. Cedar, 2021. <https://docs.alliancecan.ca/wiki/Cedar/en> [Accessed: June 9, 2023].
- C. DiLullo, G. B. Taylor, and J. Dowell. Using the Long Wavelength Array to Search for Cosmic Dawn. *Journal of Astronomical Instrumentation*, 09(02), June 2020. doi: 10.1142/s2251171720500087. URL <https://doi.org/10.1142/s2251171720500087>.
- S. Dodelson and F. Schmidt. *Modern Cosmology*. Academic Press, 2nd edition, 2021. ISBN 9780128159484.
- M. W. Eastwood, M. M. Anderson, R. M. Monroe, G. Hallinan, M. Catha, J. Dowell, H. Garsden, L. J. Greenhill, B. C. Hicks, J. Kocz, D. C. Price, F. K. Schinzel, H. Vedantham, and Y. Wang. The 21 cm Power Spectrum from the Cosmic Dawn: First Results from the OVRO-LWA. *The Astronomical Journal*, 158(2):84, July 2019. doi: 10.3847/1538-3881/ab2629. URL <https://doi.org/10.3847/1538-3881/ab2629>.
- 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. *The Astrophysical Journal*, 633(2): 560–574, Nov. 2005. doi: 10.1086/466512. URL <https://doi.org/10.1086%2F466512>.
- S. Fraser, A. Hektor, G. Hütsi, K. Kannike, C. Marzo, L. Marzola, A. Racioppi, M. Raidal, C. Spethmann, V. Vaskonen, and H. Veermäe. The EDGES 21 cm anomaly and properties of dark matter. *Physics Letters B*, 785:159–164, Oct. 2018. doi: 10.1016/j.physletb.2018.08.035. URL <https://doi.org/10.1016%2Fj.physletb.2018.08.035>.
- D. J. Griffiths. *Introduction to Quantum Mechanics*. Pearson Education, 2nd edition, 2004.
- E. Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173, 1929. doi: 10.1073/pnas.15.3.168. URL <https://www.pnas.org/doi/abs/10.1073/pnas.15.3.168>.
- M. N. N. Islam, D. Ölçek, G. Lacy, H. C. Chiang, and R. Hellyer. Errors in Deep Dish Development Array (6m) construction and metrology steps. In H. K. Marshall, J. Spyromilio, and T. Usuda, editors, *Ground-based and Airborne Telescopes IX*, volume 12182 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 121820F, Aug. 2022. doi: 10.1117/12.2629423.
- D. C. Jacobs, J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. S. Ali, J. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, P. Klima, A. Liu, D. H. E. MacMahon, J. R. Manley, D. F. Moore, I. I. Stefan, and W. P. Walbrugh.

- MULTIREDSHIFT LIMITS ON THE 21 cm POWER SPECTRUM FROM PAPER. *The Astrophysical Journal*, 801(1):51, Mar. 2015. doi: 10.1088/0004-637x/801/1/51. URL <https://doi.org/10.1088%2F0004-637x%2F801%2F1%2F51>.
- W. Kester and W. Bryant. *The Data Conversion Handbook*. Newnes, Norwood, MA, 2005. ISBN 0750678410.
- F. Kirsten, B. Marcote, K. Nimmo, J. W. T. Hessels, M. Bhardwaj, S. P. Tendulkar, A. Keimpema, J. Yang, M. P. Snelders, P. Scholz, A. B. Pearlman, C. J. Law, W. M. Peters, M. Giroletti, Z. Paragi, C. Bassa, D. M. Hewitt, U. Bach, V. Bezrukovs, M. Burgay, S. T. Buttaccio, J. E. Conway, A. Corongiu, R. Feiler, O. Forssén, M. P. Gawroński, R. Karuppusamy, M. A. Kharinov, M. Lindqvist, G. Maccaferri, A. Melnikov, O. S. Ould-Boukattine, A. Possenti, G. Surcis, N. Wang, J. Yuan, K. Aggarwal, R. Anna-Thomas, G. C. Bower, R. Blaauw, S. Burke-Spolaor, T. Cassanelli, T. E. Clarke, E. Fonseca, B. M. Gaensler, A. Gopinath, V. M. Kaspi, N. Kassim, T. J. W. Lazio, C. Leung, D. Z. Li, H. H. Lin, K. W. Masui, R. Mckinven, D. Michilli, A. G. Mikhailov, C. Ng, A. Orbidans, U. L. Pen, E. Petroff, M. Rahman, S. M. Ransom, K. Shin, K. M. Smith, I. H. Stairs, and W. Vlemmings. A repeating fast radio burst source in a globular cluster. *Nature*, 602(7898):585–589, Feb. 2022. doi: 10.1038/s41586-021-04354-w. URL <https://doi.org/10.1038%2Fs41586-021-04354-w>.
- C. Lai. 300-1500 MHz Broadband LNA for the Canadian Hydrogen Observatory and

- Radio-transient Detector, 2022. Available at <https://prism.ucalgary.ca/items/e6ff3082-fa5c-4600-97c3-a1635699aea9>.
- A. Liu and J. R. Shaw. Data Analysis for Precision 21 cm Cosmology. *Publications of the Astronomical Society of the Pacific*, 132(1012):062001, June 2020. doi: 10.1088/1538-3873/ab5bfd.
- V. MacKay, M. Lai, P. Shmerko, D. Wulf, L. Belostotski, and K. Vanderlinde. Low-cost, Low-loss, Ultra-wideband Compact Feed for Interferometric Radio Telescopes. *arXiv e-prints*, art. arXiv:2210.07477, Oct. 2022. doi: 10.48550/arXiv.2210.07477.
- MACOM Technology Solutions Inc. *MABA-011118*. MACOM Technology Solutions Inc. URL <https://cdn.macom.com/datasheets/MABA-011118.pdf>.
- B. McKinley, C. M. Trott, M. Sokolowski, R. B. Wayth, A. Sutinjo, N. Patra, J. NambissanT., and D. C. X. Ung. The All-Sky SignAl Short-Spacing INterferometer (ASSASSIN) – I. Global-sky measurements with the Engineering Development Array-2. *Monthly Notices of the Royal Astronomical Society*, 499(1):52–67, Sept. 2020. doi: 10.1093/mnras/staa2804. URL <https://doi.org/10.1093%2Fmnras%2Fstaa2804>.
- J. Mena-Parra, C. Leung, S. Cary, K. W. Masui, J. F. Kaczmarek, M. Amiri, K. Bandura, P. J. Boyle, T. Cassanelli, J.-F. Cliche, M. Dobbs, V. M. Kaspi, T. L. Landecker, A. Lanman, and J. L. Sievers. A Clock Stabilization System for CHIME/FRB Outriggers.

- The Astronomical Journal*, 163(2):48, Jan. 2022. doi: 10.3847/1538-3881/ac397a. URL <https://doi.org/10.3847%2F1538-3881%2Fac397a>.
- Microchip Technology Inc. ADC Differential Nonlinearity. Technical report, 2021. URL <https://microchipdeveloper.com/adc:adc-differential-nonlinearity>. Accessed: June 27, 2023.
- Mini-Circuits. *LFCN-9170+*. Mini-Circuits, a. URL <https://www.minicircuits.com/pdfs/LFCN-9170+.pdf>.
- Mini-Circuits. *ZX75BP-1062-S+*. Mini-Circuits, b. URL <https://www.mini-circuits.com/pdfs/ZX75BP-1062-S+.pdf>.
- Mini-Circuits. *ZX75BP-1307-S+*. Mini-Circuits, c. URL <https://www.minicircuits.com/pdfs/ZX75BP-1307-S+.pdf>.
- Mini-Circuits. *ZX75BP-1350-S+*. Mini-Circuits, d. URL <https://www.minicircuits.com/pdfs/ZX75BP-1350-S+.pdf>.
- J. Mirocha. ARES: Accelerated Reionization Era Simulations. Astrophysics Source Code Library, record ascl:2011.010, Nov. 2020.
- R. A. Monsalve, V. Bidula, R. Bustos, C. H. Bye, H. C. Chiang, M. Diaz, X. Guo, I. Hendricksen, E. Horneker, F. Lucero, H. Mani, F. McGee, F. P. Mena, M. Pessoa,

- G. Prabhakar, O. Restrepo, J. L. Sievers, and N. Thyagarajan. Mapper of the IGM Spin Temperature (MIST): Instrument Overview. 2023. [Manuscript in preparation].
- T. J. Mozdzen, J. D. Bowman, R. A. Monsalve, and A. E. E. Rogers. Limits on foreground subtraction from chromatic beam effects in global redshifted 21 cm measurements. *Monthly Notices of the Royal Astronomical Society*, 455(4):3890–3900, Feb. 2016. doi: 10.1093/mnras/stv2601.
- 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 H. J. Hall, R. Gilmozzi, and H. K. Marshall, editors, *Ground-based and Airborne Telescopes VI*, volume 9906 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 99065X, Aug. 2016. doi: 10.1117/12.2234286.
- B. D. Nhan, D. D. Bordenave, R. F. Bradley, J. O. Burns, K. Tauscher, D. Rapetti, and P. J. Klima. Assessment of the Projection-induced Polarimetry Technique for Constraining the Foreground Spectrum in Global 21 cm Cosmology. *The Astrophysical Journal*, 883(2):

- 126, Sept. 2019. doi: 10.3847/1538-4357/ab391b. URL <https://dx.doi.org/10.3847/1538-4357/ab391b>.
- A. V. Oppenheim and R. W. Schaffer. *Discrete-Time Signal Processing*. Pearson, 3rd edition, 2010. ISBN 978-0131988422.
- S. Paul, M. G. Santos, Z. Chen, and L. Wolz. A first detection of neutral hydrogen intensity mapping on Mpc scales at $z \approx 0.32$ and $z \approx 0.44$. arXiv 2301.11943, 2023.
- U.-L. Pen, T.-C. Chang, C. M. Hirata, J. B. Peterson, J. Roy, Y. Gupta, J. Odegova, and K. Sigurdson. The GMRT EoR experiment: limits on polarized sky brightness at 150 MHz. *Monthly Notices of the Royal Astronomical Society*, 399(1):181–194, Oct. 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.14980.x. URL <https://doi.org/10.1111/j.1365-2966.2009.14980.x>.
- R. A. Perley and B. J. Butler. An Accurate Flux Density Scale from 50 MHz to 50 GHz. *The Astrophysical Journal Supplement Series*, 230(1):7, May 2017. doi: 10.3847/1538-4365/aa6df9.
- E. Petroff, J. W. T. Hessels, and D. R. Lorimer. Fast radio bursts. *The Astronomy and Astrophysics Review*, 27(1):4, May 2019. ISSN 1432-0754. doi: 10.1007/s00159-019-0116-6. URL <https://doi.org/10.1007/s00159-019-0116-6>.
- L. Philip, Z. Abdurashidova, H. C. Chiang, N. Ghazi, A. Gumba, H. M. Heilgendorff,

- J. Hickish, J. M. Jáuregui-García, K. Malepe, C. D. Nunhokee, J. Peterson, J. L. Sievers, V. Simes, and R. Spann. Probing Radio Intensity at high-Z from Marion: 2017 Instrument. arXiv 1806.09531, 2018.
- E. Pieters. Prototyping and testing analog electronics and dishes for HIRAX and CHORD, 2021. Available at <https://escholarship.mcgill.ca/concern/theses/6682x895k?locale=en>.
- Pozar, David M. *Microwave Engineering*. John Wiley & Sons, Inc., 4th edition, 2012.
- D. C. Price, L. J. Greenhill, A. Fialkov, G. Bernardi, H. Garsden, B. R. Barsdell, J. Kocz, M. M. Anderson, S. A. Bourke, J. Craig, M. R. Dexter, J. Dowell, M. W. Eastwood, T. Eftekhari, S. W. Ellingson, G. Hallinan, J. M. Hartman, R. Kimberk, T. J. W. Lazio, S. Leiker, D. MacMahon, R. Monroe, F. Schinzel, G. B. Taylor, E. Tong, D. Werthimer, and D. P. Woody. Design and characterization of the large-aperture experiment to detect the dark age (LEDA) radiometer systems. *Monthly Notices of the Royal Astronomical Society*, May 2018. doi: 10.1093/mnras/sty1244. URL <https://doi.org/10.1093/mnras/sty1244>.
- J. R. Pritchard and A. Loeb. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, Aug. 2012. doi: 10.1088/0034-4885/75/8/086901.
- J. M. Reynolds. *An Introduction to Applied and Environmental Geophysics*. Wiley-Blackwell, West Sussex, UK, 2011. ISBN 978-0-471-48535-3.

- RIGOL Technologies Inc. *DSG800 Data Sheet*. RIGOL Technologies Inc., 2019.
URL <https://beyondmeasure.rigoltech.com/acton/attachment/1579/f-0664/1/-/-/-/-/DSG800%20Data%20Sheet.pdf>.
- Y. Shi, F. Deng, Y. Xu, F. Wu, Q. Yan, and X. Chen. Lunar Orbit Measurement of the Cosmic Dawn's 21 cm Global Spectrum. *The Astrophysical Journal*, 929(1):32, Apr. 2022. doi: 10.3847/1538-4357/ac5965. URL <https://dx.doi.org/10.3847/1538-4357/ac5965>.
- S. Singh, R. Subrahmanyan, N. U. Shankar, M. S. Rao, B. S. Girish, A. Raghunathan, R. Somashekar, and K. S. Srivani. SARAS 2: a spectral radiometer for probing cosmic dawn and the epoch of reionization through detection of the global 21-cm signal. *Experimental Astronomy*, 45(2):269–314, Apr. 2018. doi: 10.1007/s10686-018-9584-3.
- S. Singh, J. Nambissan T., R. Subrahmanyan, N. Udaya Shankar, B. S. Girish, A. Raghunathan, R. Somashekar, K. S. Srivani, and M. Sathyanarayana Rao. On the detection of a cosmic dawn signal in the radio background. *Nature Astronomy*, 6(5):607–617, May 2022. ISSN 2397-3366. doi: 10.1038/s41550-022-01610-5. URL <https://doi.org/10.1038/s41550-022-01610-5>.
- M. Sokolowski, S. E. Tremblay, R. B. Wayth, S. J. Tingay, N. Clarke, P. Roberts, M. Waterson, R. D. Ekers, P. Hall, M. Lewis, M. Mossammaparast, S. Padhi, F. Schlagenhauer, A. Sutinjo, and J. Tickner. BIGHORNS - broadband instrument

- for global Hydrogen Reionisation signal. *Publications of the Astronomical Society of Australia*, 32, 2015. doi: 10.1017/pasa.2015.3. URL <https://doi.org/10.1017/pasa.2015.3>.
- M. Spinelli, G. Kyriakou, G. Bernardi, P. Bolli, L. J. Greenhill, A. Fialkov, and H. Garsden. Antenna beam characterization for the global 21-cm experiment LEDA and its impact on signal model parameter reconstruction. *Monthly Notices of the Royal Astronomical Society*, 515(2):1580–1597, July 2022. doi: 10.1093/mnras/stac1804. URL <https://doi.org/10.1093/mnras/stac1804>.
- Teledyne. *EV8AQ160 QUAD ADC Datasheet DS0846*. Teledyne, 2019. URL <http://www.e2v.com/resources/account/download-datasheet/2291>.
- J. N. Thekkepattu, B. McKinley, C. M. Trott, J. Jones, and D. C. X. Ung. System design and calibration of SITARA—a global 21 cm short spacing interferometer prototype. *Publications of the Astronomical Society of Australia*, 39:e018, 2022. doi: 10.1017/pasa.2022.13.
- A. R. Thompson, J. M. Moran, and J. Swenson, George W. *Interferometry and Synthesis in Radio Astronomy, 3rd Edition*. 2017. doi: 10.1007/978-3-319-44431-4.
- 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 Bemmelen, 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. Brüggemeyer, 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. Doorduyn, 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. Griebmeier, 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. Juetten, 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. *Astronomy & Astrophysics*, 556:A2, Aug. 2013. doi: 10.1051/0004-6361/201220873.
- K. Vanderlinde, A. Liu, B. Gaensler, D. Bond, G. Hinshaw, C. Ng, C. Chiang, I. Stairs, J.-A. Brown, J. Sievers, J. Mena, K. Smith, K. Bandura, K. Masui, K. Spekkens, L. Belostotski, M. Dobbs, N. Turok, P. Boyle, M. Rupen, T. Landecker, U.-L. Pen, and V. Kaspi. The Canadian Hydrogen Observatory and Radio-transient Detector (CHORD). In *Canadian Long Range Plan for Astronomy and Astrophysics White Papers*, volume 2020, page 28, Oct. 2019. doi: 10.5281/zenodo.3765414.
- H. K. Vedantham, L. V. E. Koopmans, A. G. de Bruyn, S. J. Wijnholds, B. Ciardi, and M. A. Brentjens. Chromatic effects in the 21 cm global signal from the cosmic dawn. *Monthly Notices of the Royal Astronomical Society*, 437(2):1056–1069, Jan. 2014. doi: 10.1093/mnras/stt1878.

- E. Wit, E. van den Heuvel, and J.-W. Romeijn. ‘All models are wrong...’: an introduction to model uncertainty. *Statistica Neerlandica*, 66(3):217–236, 2012. doi: <https://doi.org/10.1111/j.1467-9574.2012.00530.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9574.2012.00530.x>.
- L. Wolz, A. Pourtsidou, K. W. Masui, T.-C. Chang, J. E. Bautista, E.-M. Müller, S. Avila, D. Bacon, W. J. Percival, S. Cunnington, C. Anderson, X. Chen, J.-P. Kneib, Y.-C. Li, Y.-W. Liao, U.-L. Pen, J. B. Peterson, G. Rossi, D. P. Schneider, J. Yadav, and G.-B. Zhao. H scpi/scp constraints from the cross-correlation of eBOSS galaxies and Green Bank Telescope intensity maps. *Monthly Notices of the Royal Astronomical Society*, 510(3):3495–3511, Dec. 2021. doi: 10.1093/mnras/stab3621. URL <https://doi.org/10.1093/mnras/stab3621>.
- Xilinx Inc. Understanding Key Parameters for RF-Sampling Data Converters. Technical report, 2019. URL <https://docs.xilinx.com/v/u/en-US/wp509-rfsampling-data-converters>.
- Xilinx Inc. *7 Series FPGAs Data Sheet: Overview (DS180)*. Xilinx Inc., 2020. URL https://docs.xilinx.com/v/u/en-US/ds180_7Series_Overview.
- Y. Xu, X. Wang, and X. Chen. Forecasts on the Dark Energy and Primordial Non-Gaussianity Observations with the Tianlai Cylinder Array. *The Astrophysical Journal*, 798(1):40, Jan. 2015. doi: 10.1088/0004-637X/798/1/40.

- B. Zhang. The Physics of Fast Radio Bursts. arXiv 2212.03972, 2022. URL <https://doi.org/10.48550/arXiv.2212.03972>.
- C. Zhao, A. Variu, M. He, D. Forero-Sánchez, A. Tamone, C.-H. Chuang, F.-S. Kitaura, C. Tao, J. Yu, J.-P. Kneib, W. J. Percival, H. Shan, G.-B. Zhao, E. Burtin, K. S. Dawson, G. Rossi, D. P. Schneider, and A. de la Macorra. The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: cosmological implications from multitracer BAO analysis with galaxies and voids. *Monthly Notices of the Royal Astronomical Society*, 511 (4):5492–5524, Feb. 2022. doi: 10.1093/mnras/stac390. URL <https://doi.org/10.1093/mnras/stac390>.

Appendix A

Diagrams of ADC Parameters

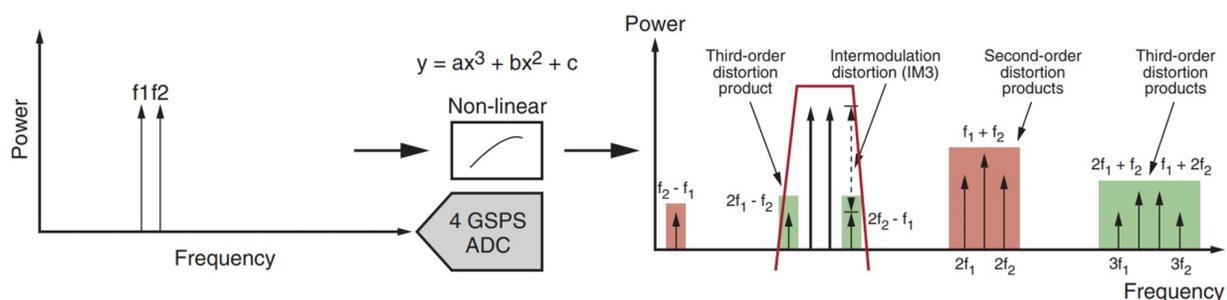


Figure A.1: A diagram demonstrating intermodulation distortion. In the left plot, two analog tones f_1 and f_2 are digitized by a real ADC with a polynomial transfer function. In the right plot, intermodulation distortion products show up in the digital spectrum at the frequencies shown in the right plot due to trigonometric identities of raising the sum of two sine waves to a given power. Many of the intermodulation distortion products The third-order intermodulation distortion (IM3) products are typically quoted in datasheets, though many other products will be in-band for CHORD depending on the frequencies f_1 and f_2 . Intermodulation distortion is an important ADC metric for radio astronomy, as it can cause power from RFI spikes to leak into other frequency bins. Figure from [Xilinx Inc. \(2019\)](#).

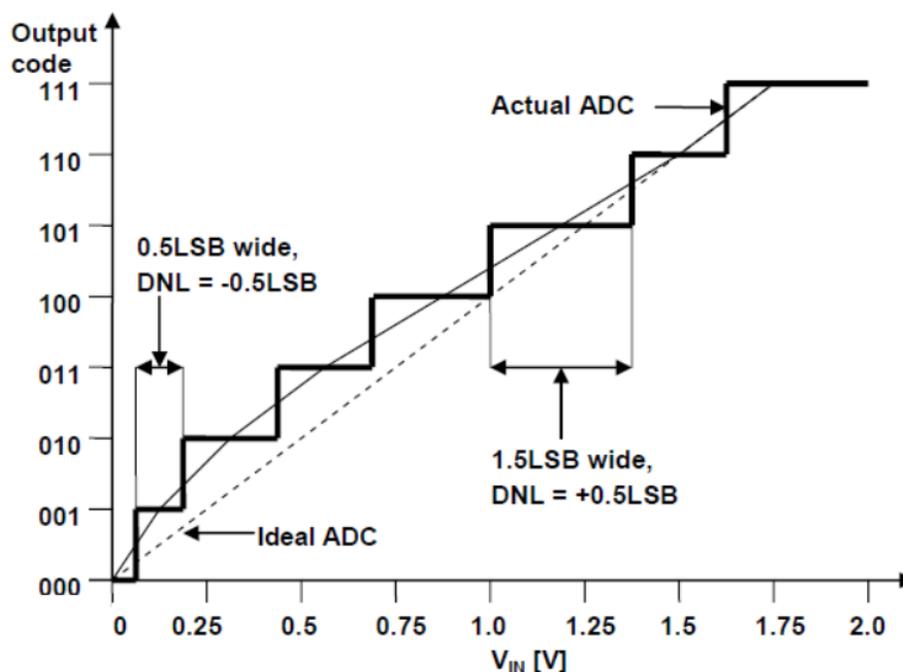


Figure A.2: A diagram demonstrating differential nonlinearity (DNL) in an ADC. An ADC digitizes signals by mapping a digital code (y-axis) to a given voltage range with width ΔV (x-axis). An ideal ADC has a linear transfer function, meaning that the width ΔV is constant for each voltage range corresponding to a digital code. A real ADC does not necessarily have an equivalent width ΔV^* that is greater or less than ΔV . The DNL is the difference between the ideal and real widths $\Delta V - \Delta V^*$, and is expressed in units of least significant bits. The DNL is a key metric for radio astronomy, as signal levels are typically kept at lower power in order to accommodate RFI spikes. Low power signals will only exercise a few LSBs, meaning that the effects of DNL are more prominent for radio astronomy applications. Figure from [Microchip Technology Inc. \(2021\)](#).