High-Performance Signal Processing Architectures for Digital Aperture Array Telescopes

Richard Paul Armstrong

St. Peter's College
Astrophysics
University of Oxford

A thesis submitted to the
Mathematical, Physical and Life Sciences Division
for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term 2011

High-Performance Signal Processing Architectures for Digital Aperture Array Telescopes

Richard Paul Armstrong
St. Peter's College, University of Oxford
A thesis submitted for the degree of Doctor of Philosophy in Trinity Term 2011

Abstract

An instrument with the ability to image neutral atomic hydrogen (HI) to cosmic redshift will allow the fundamental properties of the Universe to be more precisely determined; in particular the distribution, composition, and evolutionary history of its matter and energy. The Square Kilometre Array (SKA) is a radio survey telescope conceived with this aim. It will have the observational potential for much further fundamental science, including strong field tests of gravity and general relativity, revealing the origin and history of cosmological re-ionisation and magnetism, direct measures of gravitational radiation, and surveys of the unmapped Universe. And it is the advance of instrumentation that will enable it.

This thesis makes three central contributions to radio instrumentation. Digital aperture arrays are a collector technology proposed for the key low- and mid- frequency ranges targeted by the SKA that have the potential to provide both the collecting area and field of view required for deep, efficient all-sky surveys of HI. The 2-Polarisations, All Digital (2-PAD) aperture array is an instrumental pathfinder for the SKA, novel in being a densely-spaced, wide-band aperture array that performs discrete signal filtering entirely digitally. The digital design of the 2-PAD radio receiver and the deployment of the aperture array and signal processing system at Jodrell Bank Radio Observatory is detailed in this thesis.

The problem of element anisotropy in small arrays, the atomic unit of the SKA station array, ultimately affects beam quality. Addressing this issue, a metaheuristic digital beam-shape optimisation technique is applied to a small beamformed array, and is shown to outperform traditional analytic solutions.

Digital processing for aperture arrays is challenging. A qualitative framework shows that energy, computational and communication requirements demand optimised processing architectures. A quantitative model reveals the physical limitations on architecture choice. An energy-optimised architecture, the IBM BIT integer array processor, is investigated in detail; a cycle-accurate architectural simulator and programming language are developed and used to build signal processing algorithms on the array architecture.

Contents

Acknowledgments

First and foremost, I would like to thank my supervisors. To Mike Jones, for invariably contributing clarifying perspective to every project, and for generously sharing of your deep understanding of radio instrumentation. To Kristian Zarb Adami, for your inimitable input into the finer details of instrumentation research. Thank you both for accepting me into your research group.

I express my gratitude to IBM Research, SKA South Africa and Research Councils UK for generous support through a Dorothy Hodgkin Postgraduate Fellowship.

Much of the work in this thesis would not have been possible in isolation. I thank all those with whom I have had the opportunity to work collaboratively: the entire 2-PAD team at the Universities of Manchester, Cambridge, and Oxford; Asim Kahn and Tony Brown at the University of Manchester; the Jodrell Bank Observatory group; and the CASPER collaboration at Berkeley and Cape Town.

I consider myself rarely fortunate to have met at Oxford those who have become both good friends as well as excellent colleagues. To Ian Heywood, Steve Rawlings, Aris Karastergiou, Danail Obreschkow, Danny Price, Jack Hickish, Griffin Foster, Charles Copley, Roger Deane, and Sascha Schediwy: I have entirely appreciated the opportunity to be able to know and to work with each of you. Particular thanks to Danny Price and Ian Heywood for solid suggestions to Chapters 3 and 1.

I would also especially like to thank Vanessa Ferraro-Wood for her warmth and excellence, and for always making the logistics of academic life seem effortless.

Thesis acknowlegements are at best an incongrous position to express deep gratitude for those who mean so much personally. Yet without them, doctoral work is uncontemplatable. I specifically mention my very first group of friends, Brenda, Simon, and Gillian, as well as those who have arrived more lately, Kieran and Justin, for a family life filled with love. To Kate, for enduring with sublime patience and grace. And lastly, to my parents, thank you for your love and support throughout.

Chapter 1

Introduction

The ability to map the Universe's most abundant baryonic constituent, neutral atomic hydrogen (HI), out to cosmological distances enables particularly rich science. Longitudinal measurements of the matter power spectrum — derived from the HI distribution as it traces both the luminous and non-luminous matter field — will tighten the constraints on cosmological parameters. HI also reveals the evolution of galaxies in the Universe, as well as the complex history of re-ionisation during the Epoch of Reionisation (EoR) as the first radiant objects re-ionise the neutral surrounding matter. Absorption spectra of HI radiation will illuminate the largely-uncharacterised inter-galactic medium (IGM). The evolution of the earliest collapsed bodies to form galaxies, superclusters and large scale structure in the cosmic filament will too be mapped through observation of HI radiation. In summary, HI tomography allows the large-scale properties of the Universe to be more precisely determined, particularly the amount, distribution and nature of its matter and energy, its age, and its evolutionary history.

Matter tomography motivates an HI survey instrument. Yet an instrument able to map neutral hydrogen to cosmological distances is capable of further transformational scientific contributions. Pulsar surveys of the Galaxy and beyond are likely to discover exceptional systems which test relativistic gravity in the ultra-strong-field limit. Measurement of the small variations in the time-of-arrival of pulses from thousands of millisecond pulsars should enable the detection of the stochastic background signature of gravitational waves. The evolution and formation of magnetic fields and their relation to structure formation is promised by all-sky Rotation Measure (RM)s of $\sim 10^8$ extragalactic sources. Light will be shed on the astrophysical conditions necessary for formation of earth-size ($M_{se} \lesssim 4M_{\oplus}$) planets. Finally, and not least of all, an instrument that

will survey as-yet-unexplored portions of frequency-, time- and area- space is certain to yield unexpected or even unconceived discoveries. Such an instrument, flexibly designed, is the tool that will provide some answers to the fundamental questions of tomorrow.

HI tomography on such a scale is only possible with an extremely sensitive radio instrument. In order to detect the vanishingly small electromagnetic signal produced by the hyperfine quantum ground-state transition of neutral hydrogen out to distances of $z \sim 1.5$ or higher will require an instrument with a collecting area of approximately a square kilometre, the ability to operate sensitively at low radio frequencies, and highly efficient survey capability. In this chapter, the scientific potential of a large HI survey telescope is expanded upon; the Square Kilometre Array (SKA) is introduced as the worldwide collaborative project able achieve these scientific goals.

Yet it is the development and optimisation of instrumentation that will ultimately allow such an instrument to be built. A novel type of radio receiver with the ability to combine a large Field of View (FoV), exquisite sensitivity and adaptive optics — the digital aperture array — is well-suited to simultaneously meet the SKA survey requirements. It is shown that much of the SKA's HI survey capability is provided by these digital aperture arrays, which will in turn uniquely position it to deliver promised transformational science and advances in fundamental physics.

The central concern of this thesis is the modeling, design and development of digital radio receiver architectures for the arrays of receiving antennae which will form the core of the radio instrument. This chapter lays the background for instrumentation research into digital aperture arrays in terms of the fundamental scientific goals of the Square Kilometre Array.

1.1 Cosmic Evolution with 21cm Tomography

The cosmic evolution of structure in the Universe is intricately associated with that of neutral atomic hydrogen. This most abundant constituent of the baryons is the most promising candidate to shed light on the evolution of structure from the very-nearly-isotropic state of the primordial photon-baryon fluid imprinted on the cosmic microwave background (CMB) through re-ionisation and eventually to the complex structure of the local Universe.

1.1.1 Galaxy Evolution

The formation and evolution of structure in the Universe occurred hierarchically. Small density perturbations in the photon-baryon fluid at recombination grew through gravitational instability over cosmic time. Gravitationally-bound structure merges hierarchically into the galaxy halos, galaxy clusters, and cosmic filaments observable in the local Universe.

This picture of cosmic evolution is based on two discrete observations; that of the very-nearly-isotropic state of the Universe imprinted on the CMB and that of the matter distribution in the local Universe. Yet the evolution from the former to the latter epochal state is constrained by little *direct* observational evidence.

Analytic models, on one hand, and simulations of bound objects (predominantly composed of non-baryonic cold dark matter (CDM)) on the other, are used to describe cosmic evolution. For example, the Press-Schechter formalism (Press and Schechter, 1974), an analytic mass distribution function of redshift for bound objects, is largely in agreement with semi-analytical simulations of structure evolution in the Millennium Simulation (Springel et al., 2005).

However, structure formation and evolution is also intricately related to the baryons. Neutral hydrogen, both atomic (HI) and molecular (H_2)[1], and Helium (He) have structure evolution rôles. Neutral hydrogen abundance over cosmic time influences star formation (e.g. Pei and Fall, 1995; Wolfe et al., 1986). The dynamical behaviour of the baryons in processes of star formation, photo-ionisation, accretion, and expulsion directly influences galaxy evolution itself. Knowledge of the properties of the neutral gas inside galaxies over significant cosmic time — particularly the total HI mass density as a function of redshift $\Omega_{\rm HI}(z)$, but also the mass density of neutral molecular gas $\Omega_{\rm H_2}(z)$, which is the material out of which stars eventually form[2] — is expected to shed light on complex baryonic interactions during galaxy formation (Rawlings et al., 2004).

Previous surveys have shed light on the HI properties of the local galaxies. The Parkes radio telescope has been used for low-redshift ($z \lesssim 0.05$) surveys of galaxy HI properties, particularly the galaxy HI Mass Function (HIMF). The results of the HI Parkes All-Sky Survey have enabled the HIMF of thousands of galaxies to be derived and $\Omega_{\rm HI}$ ($z \simeq 0$) to be determined by Zwaan et al. (2005).

Other telescopes used for local HIMF surveys include The Arecibo Radio Telescope, the Giant Meterwave

[1]Recent work has predicted a more complex interaction between atomic and molecular neutral hydrogen than has been assumed. Particularly, the global ratio of atomic to neutral gas $\Omega_{\rm H_2}/\Omega_{\rm HI}(z)$ is expected to *increase* with redshift (see Obreschkow and Rawlings, 2009b,a).

[2]Molecular hydrogen is difficult to detect directly; its high excitation temperature makes observation rare except for gas in hot or shocked environments, so carbon monoxide is used as a tracer (see Solomon and VandenBout, 2005; Heywood et al., 2011; Walter et al., 2011). Observation of the emission lines of molecular ^{12}CO is preferred and is thought to trace H_2 in high-redshift galaxies.

Radio Telescope (GMRT), the Very Large Array (VLA), and the Westerbork Synthesis Radio Telescope (WSRT). However, none of these surveys provide measurements of $\Omega_{\rm HI}$ beyond $z \gtrsim 0.05$.

The technique of 'HI spectral stacking' has been used to derive HI properties to deeper redshift by co-adding the rest-frame-normalised spectra of multiple galaxies. This increases marginally the stochastic galaxy redshift distance. Instrumental noise in the shifted spectra is expected to sum incoherently while the 21 cm signal sums coherently, increasing in signal-to-noise ratio with $\sqrt{N_{\rm obs}}$ (see for e.g. Chengalur et al., 2001; Lah et al., 2007, 2009; Freudling et al., 2011). This allows a deeper-redshift measurement of the statistical properties of galaxies.

However, the HI properties of galaxies over a more significant redshift range has yet to be observed directly; it is here that a large HI survey of local to distant galaxies will progress observational cosmology. A thorough 21cm HI mass function measurement of $\sim 10^9$ galaxies should allow $\Omega_{\rm HI}\ (z),\ (0 \leq z \lesssim 1.5)$ to be determined, which will enable stronger constraints on galaxy evolution scenarios over more than half of the age of the Universe and will illuminate the unique and critical rôle of neutral gas in galaxy evolution (Rawlings et al., 2004; Obreschkow and Rawlings, 2009a).

1.1.2 Dark Energy and Precision Cosmology

One of the defining features of the last 30 years of astronomy is that our understanding of the Universe's cosmology has become significantly more complete. A key observational advance is that several independent measurements show that the expansion of the Universe is accelerating. This motivates modification of current gravitational theory to include the existence of a material which begins (cosmologically) recently to dominate the expansion of the Universe. It is a central goal of modern cosmology to constrain the properties of this material, the 'dark energy'.

Our understanding of cosmology has also become more accurate. Measurement of the Cosmic Microwave Background (CMB) with the Wilkinson Microwave Anisotropy Probe (WMAP) has provided strong constraints on the cosmological parameters, including the dark energy density Ω_Λ, and enabled their values to be calculated to high accuracy (e.g. Spergel et al., 2003; Komatsu et al., 2011). CMB results agree with other, independent observational measurements of cosmological parameters such as supernovae surveys and baryon acoustic oscillations (BAO).

However, in order to enter the era of high-precision dark energy cosmology which demands cosmological parameter uncertainties of $< 1\%$ (see e.g. Rawlings et al., 2004; Blake et al., 2004), extra information is

required *in addition to* CMB measurements. There are two reasons for this. First, the measurement of the CMB temperature and polarisation power spectra from the surface of last scattering is a single measurement in time, originating early in the Universe's history. CMB missions like the Wilkinson Microwave Anisotropy Probe (WMAP) and Planck will provide increasingly precise measurements of the CMB power spectra[3], but will not on their own probe dark energy directly; dark energy has only a small influence on CMB measurements since its effect on the Universe becomes marked and direct only on matter clustering in the later Universe. Second, such results contain hidden theoretical priors (for instance, the assumption of early cosmic inflation) which, if relaxed, significantly increase uncertainties on other parameters (see Efstathiou, 2003).

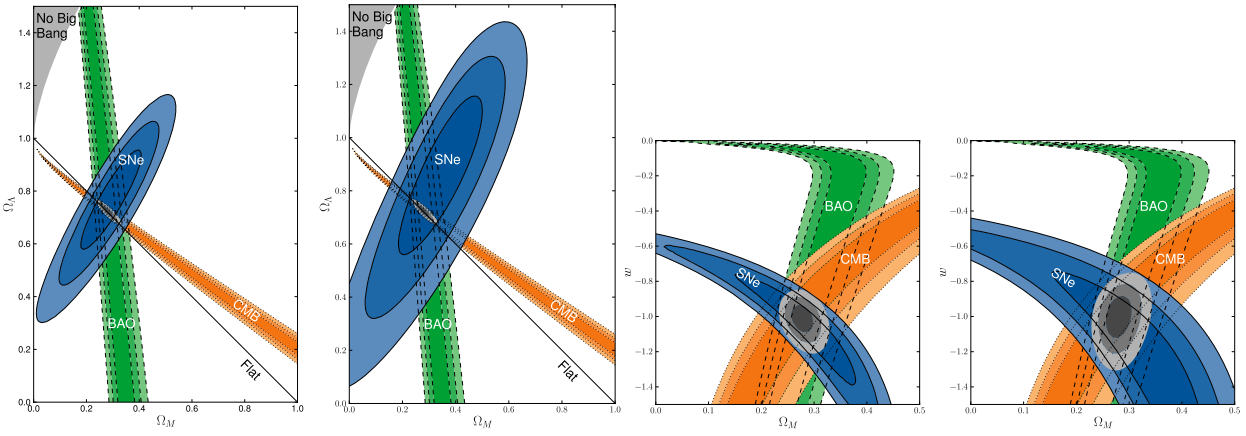
Independent constraints are placed on the dark energy density Ω_Λ by the use of high-redshift supernovae (SNe) surveys. Supernovae are expected to have a standard luminosity in outburst over cosmic volume and time (for e.g., see Hicken et al., 2009; Kowalski et al., 2008; Amanullah et al., 2010). A comparison of the observed brightness of such events is used to map the Universe to large redshift.

Cosmological parameters are also determinable due to the distribution of galaxies. The 2-point correlation function of galaxies (which are tracers of the mass density field) should contain the characteristic signature of the acoustic oscillations in the primordial fluid at the time of matter-energy decoupling (which correspond to features of approximately 100 Mpc at present epoch). The correlation function of matter in existing redshift galaxy surveys (for e.g. see Eisenstein et al., 2005; Percival et al., 2007) and stacked, deeper-redshift HI mass function (HIMF) measurements are used to confirm the existence of a characteristic scale of the acoustic horizon. Precise measurements of the dark energy require giga-parsec volumes to be probed, only possible with large number-count galaxy-redshift surveys (Eisenstein et al., 2005).

The central point is that a combination of CMB parameter measurements with a further, *independent*, measurement of the galaxy power spectrum $P(k)$ with different inherent parameter degeneracies, such as would be provided by a global HI survey to $z \sim 1.5$, would provide tighter, orthogonal constraints on cosmological parameter measurements and strengthen the precision of the measurement. Such precision is required for the important cosmological questions posed above to be addressed. In illustration of this effect, Figure 1.1 shows the improvement in measurements of the dark energy density, Ω_Λ and the total matter

[3]Planck is a third-generation space mission for CMB measurement aimed at (i) precision polarisation measurements and (ii) temperature measurement in the large-l tail where Silk damping (Silk, 1968) becomes significant rather than increasingly precise measurement of small-l acoustic peaks. Figure 1.2 shows the increase in accuracy over Wilkinson Microwave Anisotropy Probe (WMAP) at large multipoles. Indeed, current WMAP measurements are already cosmic variance limited for multipoles $l < 548$, and further precision is not possible (see Larson et al., 2011).

energy density, $\Omega_m = \Omega_b + \Omega_c + \Omega_v$, derived from the union of the three independent observations.



(a) (b)

Figure 1.1: (a) Contour lines of σ, 2σ and 3σ confidence level on Ω_Λ and Ω_m obtained from the union of CMB (Dunkley et al., 2009), BAO (Eisenstein et al., 2005) and the Union supernovae (SN) set (Kowalski et al., 2008) observational constraints with the dark energy equation of state $w = -1$ as a theoretical prior. (b) A plot of σ, 2σ and 3σ confidence level contours on w and Ω_m, for a flat Universe geometry showing individual constraints from CMB, BAO and the Union SN set, as well as the combined constraints. From Amanullah et al. (2010)

1.1.3 Mapping the Epoch of Re-Ionisation with 21-cm Tomography

The process of transition of the baryonic Universe from being largely neutral to being almost completely ionised by the first luminous objects is not fully understood. This is a critical period in the Universe's evolution history, yet its exact timeframe and transition history have yet to be directly observed.

Various cosmic reionisation scenarios are possible. The simplest of these is *instantaneous* reionisation: after the first electromagnetically radiant astrophysical objects in the Universe form, they ionise their immediate surrounding IGM. These ionised 'bubbles' expand to eventually join other ionised regions and thus fill the entire volume over time.

Other models are more likely. Existing observational constraints imply that the Universe may have had an extended (e.g. Dunkley et al., 2009) or complex (e.g. Furlanetto et al., 2006) history of reionisation, rather than a sudden, monotonic transition from a neutral to an ionised IGM. Spergel et al. (2007) model two-step reionisation consistent with current observational constraints; initial, partial reionisation (with ionised fraction $x_e < 1$) occurs at z_r and is followed by a second phase of complete reionisation, which completes at $z \gtrsim 6$. Furlanetto et al. (2006) review evidence for complex evolution, and physical theory which fits current,

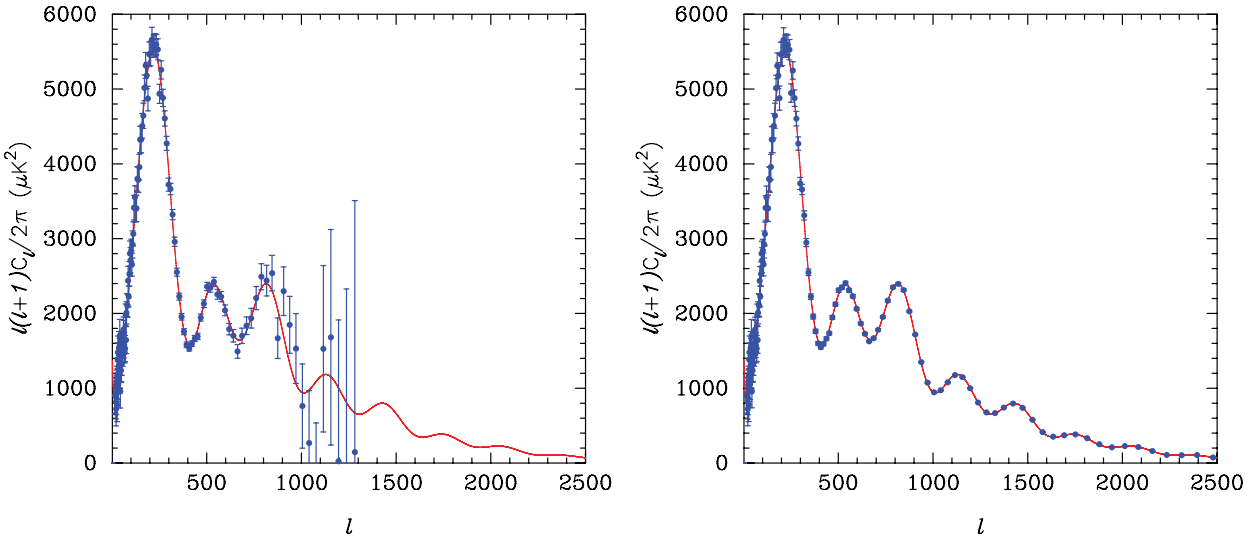


Figure 1.2: A comparison of the angular power spectrum provided by the Wikinson Microwave Anisotropy Probe (WMAP) satellite 4-year data (left panel) and simulated measurements with the Planck satellite (right panel). The red line shows the theoretical predictions of the concordance ΛCDM model. Both measurements are cosmic-variance-limited at angular scales $l \lesssim 530$. The sensitivity and angular resolution of Planck show improvements in large-angular-scale detail, visible in the right frame, which has been confirmed by the early data release (Ade et al., 2011). Figure from Aghanim et al. (2006)

conflicting observational constraints.

Three sets of observations delineate the reionisation timeframe. First, quasar spectra provide information about the later stages of reionisation. Characteristic neutral gas absorption features (Gunn-Peterson[4] absorption troughs, Gunn and Peterson 1965) are not evident in quasars observed at $z < 5.8$, but are evident for those at $z \sim 6.28$ (Becker et al., 2001; Fan et al., 2003), which implies that there is a significant cosmic HI fraction for $z > 6.28$. However, the usefulness of quasar spectral measurements is confined to epochs where the global neutral fraction is very low[5] ($\Omega_{\rm HI} < 10^{-3}$), and thus probe only the very latest stages of reionisation.

Second, the effect of reionisation can be detected on the CMB signal; polarisation anisotropies are created in the CMB signal due to Thomson scattering of microwave background photons by free electrons in ionised

[4]The Gunn-Peterson trough is a spectral feature of quasars due to the presence of neutral hydrogen in the IGM, predicted by Gunn and Peterson (1965). The Gunn-Peterson effect has demonstrated that the IGM is strongly ionised in the local Universe to at least redshift $z \sim 6.5$. Absorption lines in these spectra should appear at wavelengths less than that of the Lyα line for emission which encounters even small amounts of neutral hydrogen along its line-of-sight. The effect was confirmed by Becker et al. (2001) and Fan et al. (2003)

[5]The mean optical depth to Lyα absorption, τ_{es}, is found to be $6.45 \times 10^5 x_{\rm H} ((1+z)/10)^{3/2}$ by Furlanetto and Briggs (2004); thus, quasar absorption features are completely saturated for $\Omega_{\rm HI} < 10^{-3}$.

media. The WMAP temperature-polarisation cross-correlation (TE) data (e.g. Dunkley et al., 2009; Kogut et al., 2003) place an integral constraint on the total scattering optical depth, τ_s, of CMB photons. From these data, Kogut et al. (2003) estimate a reionisation timeframe $7 < z < 30$ (95% confidence). It is expected that the Planck space mission will provide even stronger constraints on CMB polarisation power spectra and thereby the optical depth to reionisation. However, an integral constraint such as optical depth is not able to differentiate non-monotonic redshift evolution of the neutral fraction, $\Omega_{\rm HI}\ (z)$; τ_s can at best *indirectly* constrain the redshift of reionisation and does not discriminate among complex reionisation models.

A third constraint is due to temperature-ionisation coupling. The widths of quasar Lyα absorption lines at $z = 3 - 3.5$ allow the IGM temperature to be determined. Theuns et al. (2002) use analogous He II reionisation-temperature dependence to predict HI reionisation-temperature dependence. These data imply hydrogen reionisation occurred *after* redshift $z \sim 9$.

Bounds on the reionisation timeframe from these three observations are $6 \simeq z_{\rm EOR} \simeq 20$, a redshift range $\Delta z_{\rm EOR}$ of ~ 14. This dictates a low-frequency bound for the redshifted HI line of 68 MHz (shifted from 1420 MHz at rest-frame). Reionisation detection through observation of the 21cm signal is a low-radio-frequency experiment.

The measurement of HI during the Epoch of Reionisation (EoR) will shed light on one of the most important processes in astrophysics, namely the formation of early electromagnetically luminescent objects in the Universe among underlying gravitational structure and the ionisation relationship of these objects with the diffuse, previously neutral IGM.

1.1.4 Hyperfine HI Transition Emission

So how, then, is neutral hydrogen observed? Neutral hydrogen (HI) undergoes hyperfine ground state splitting with probability $\zeta_{0-1} = 2.85 \times 10^{-15} s^{-1}$, emitting spectral line radiation at characteristic frequency of 1.42 GHz (or characteristic wavelength 21 cm). Such transitions occur between upper ($F = 1$) and lower ($F = 0$) excitation states of HI, and result in either radiation or absorption of energy. The effect was predicted by van der Hulst (1946) and first observed by Ewen and Purcell (1951). It is this weak and rare spectral line emission that directly traces the HI distribution.

The 21-cm line is sensitive to several properties of HI: its density δ, the global neutral fraction $\Omega_{\rm HI}$, and the 'spin temperature' of the HI cloud, $\rm T_s$.

Spin Temperature

The spin temperature of an HI cloud is the temperature at which its actual ratio of ground-state to excited-state atoms ($\frac{N_{ex}}{N_{\rm g}}$) would occur if the cloud were in a state of thermal equilibrium. The spin temperature of a gas is defined as the ratio of atoms in the upper to lower excitation by the Boltzmann exponential law (Abragam and Proctor, 1958),

$$\frac{N_{ex}}{N_{\rm g}} = \frac{p_{\rm ex}}{p_{\rm g}} \exp\left(-\frac{E_{\text{1-0}}}{kT_s}\right) \tag{1.1}$$

where $\frac{p_{\rm ex}}{p_{\rm g}}$ is the statistical probability ratio and $E_{\text{1-0}}$ is the transition energy level. $\rm T_s$ differs from the kinetic temperature at low gas densities. Here, the spin temperature is raised by resonant scattering of Lyman-alpha photons, which cause an equalisation of occupation ratio of upper to lower excitation states. In an HI cloud, the spin temperature may be expressed as

$$T_{\rm S} = \frac{T_{\rm CMB} + k_{\rm Ly\alpha} T_{\rm Ly\alpha}}{1 + k_{\rm Ly\alpha}}. \tag{1.2}$$

Hyperfine Excitation

Hyperfine excitation of HI is due to collisional-excitation (either atomic (H-H) or charged particle (H-e^-) collisions) or to the Wouthuysen-Field effect (after Wouthuysen, 1952; Field, 1958) where Lyα interaction connects $\rm T_s$ to the temperature of the radiation field (which is equal to the kinetic gas temperature $\rm T_k$ in most astrophysical systems). The difference in brightness temperature due to the hyperfine transition δT_b will appear against the CMB temperature $T_{\rm CMB}$ in either emission or absorption,

$$\delta T_b \approx \frac{T_{\rm s} - T_{\rm CMB}}{1+z}(1 - e^{-\tau_v}) \tag{1.3}$$

where τ_v is the optical depth of the hydrogen cloud. Equation 1.3 is re-formulated in terms of δ, the local overdensity relative to the mean HI density by Furlanetto et al. (2006)[6] to give

[6]Furlanetto et al. (2006) proceed as follows: the optical depth of the hydrogen cloud τ_v is

$$\begin{aligned}\tau_v &= \int \sigma_{01} \left(1 - e^{-E_{10}/k_B T_S}\right) \phi(v) n_0 \, \mathrm{d}s \\ &= \sigma_{01} \left(\frac{N_{\rm HI}}{4}\right) \left(\frac{hv}{k_{\rm B} T_{\rm S}}\right) \phi(v).\end{aligned}$$

where $N_{\rm HI}$ is the column density of HI and $\sigma_{01} = \frac{3c^2 \zeta_{01}}{8\pi v^2}$.

$$\delta T_b \approx 23(T_{\rm S} - T_{\rm CMB}) \left(\frac{1+z}{10}\right)^{\frac{1}{2}} (1+\delta)\Omega_{\rm HI} \text{ mK} \tag{1.4}$$

Essentially, equation 1.4 implies that the observability of the 21cm signal is dependent on both the spin temperature $T_{\rm S}$ and the local overdensity, δ.

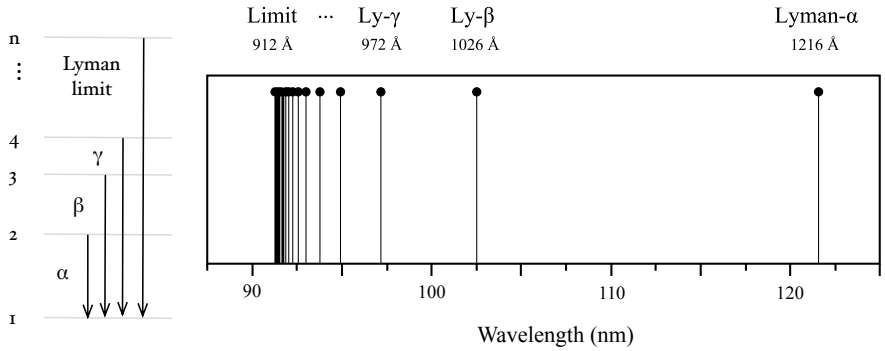


Figure 1.3: A sketch showing possible transitions in the hydrogen atom when an electron moves between energy levels. Transitions are grouped into series named after those who discovered them; the Lyman series and its resulting wavelengths for hydrogen is shown. Radiation from the first luminous objects Raman-scatters through the Lyα resonances and shifts neutral hydrogen between the hyperfine levels (F=0 and F=1). This process tends to bring the hydrogen spin temperature, $T_{\rm s}$ into equilibrium with the gas kinetic temperature T_k. This process is named after Wouthuysen (1952) and Field (1958).

The second term in Equation 1.2 is due to the Wouthysen-Field effect, where $k_{{\rm Ly}\alpha}$ is the coupling coefficient to the radiation field temperature $T_{{\rm Ly}\alpha}$ (Field, 1958).

In summary, variations in the observed differential brightness would depend on fluctuations in the density of neutral hydrogen $\Omega_{\rm HI}\,(z)$. It is this differential brightness temperature fluctuation as a function of redshift which is expected to contain the signature of cosmic reionisation.

1.2 Beyond Tomography: Further Science with a Radio Survey Telescope

A survey instrument with the sensitivity to map HI over a large redshift range has the ability to perform many other experiments. In this section, several experiments and surveys are described which are made possible by the low-frequency capability of a large radio survey telescope and which make the SKA such an important instrument.

1.2.1 Strong-Field tests of Gravity and General Relativity

Some rotating, collapsed neutron stars, called pulsars, emit pulsed radiation in the radio-frequency range (Hewish et al., 1969). These complex astrophysical systems have the potential to enable two experiments testing fundamental physics. An all-sky pulsar survey aimed at finding most observable Galactic pulsars and some bright extragalactic pulsars is likely to include the detection of at least one exceptional system, such as a pulsar-black hole binary. This system will test relativistic gravity in the ultra-strong-field limit, promising observational advances in the unification of quantum mechanics and relativistic gravity theories (see Kramer et al., 2004).

The second experiment is aimed at *direct* detection of the *stochastic* background signature of gravitational waves. The perturbation of space-time due to passing gravitational waves can be inferred by changes in the pulse arrival times of an array of millisecond pulsars under constant observation. The Earth and each of the set of pulsars in the 'timing array' are free masses influenced by changes in the space-time fabric; variations in pulse-arrival times should trace directly the variations in the fabric (Kramer, 2007; Hobbs et al., 2010). Such experiments may be among the most significant results promised by the SKA. Prime sensitivity is required in the 0.25–1.5 GHz range, especially of the crucial millisecond-period pulsars (Kramer et al., 2004).

1.2.2 Directed Phase-Space Surveys

Considering the large number, geographical extent and frequency range of modern astronomical electromagnetic instrumentation, perhaps what is most striking is the relatively small fraction of the sky under observation at any one time. The SKA will increase the fraction of the redshifted sky under direct observation over previous radio instruments due to a 50-fold increase in sensitivity and wide Field of View (FoV).

Exploration of the Unknown as a Targeted Design Philosophy

The history of radio astronomy suggests that observation often precedes and indeed stimulates theory, as evident in the discovery of non-thermal and synchrotron radiation (Burbidge, 1956), quasars, the cosmic microwave background (Penzias and Wilson, 1965), cosmic evolution, pulsars (Hewish et al., 1969), gravitational lensing, cosmic masers, molecular clouds, dark matter, and extrasolar planetary systems, which have resulted in the understanding of often entirely unexpected phenomena and systems[7]. With this in mind, *di-*

[7]See also the summary article by Wilkinson et al. (2004) for a history of increases in the astronomical observation space generating discovery.

rected efforts at monitoring the entire sky while exploring new regions of the time, frequency, and luminosity parameter space (which are inherently 'at odds' in an integrative instrument), are bound to yield discoveries. Indeed, it is well-recognised in the radio astronomy community (e.g. Cordes et al., 2004; Wilkinson et al., 2004) that targeted exploration of the parameter space spanned by time and luminosity will yield both a far greater number of known classes of radio sources as well as previously-unobserved and even unpredicted classes of objects.

Fast Radio Transients

Transient, energetic pulses are produced by extreme, and sometimes theoretically-challenging cosmic events. The short time duration of emission and surprisingly large energies (Hayashida et al. 1996; Linsley 1963) evidence the association of such events with extreme physical environments.

Many types of fast radio transients have been discovered serendipitously. Examples of unexplained observational phenomena which preceded the development of suitable theory and subsequent discovery of new classes of unexpected astrophysical objects include the discovery of pulsars (Hewish et al., 1969), Rotating Radio Transients (RRATs) (McLaughlin et al., 2006), magnetars (Cline et al., 1982) and Impulsive Extragalactic Transients (IETs) (Lorimer et al., 2007). It is expected that light will be shed on such extreme environments due to observation of existing and new short-time transient sources.

Directed Transient Searches

A large survey telescope is inherently a discovery machine; during the course of an all-sky galaxy HI survey, there will be opportunity to observe a large part of the missing phase space of radio transients. A measure of the efficiency of a blind survey telescope is given by Cordes (2007): a transient detection machine will take into account the (transient) source event rate, duration, and luminosity. Figure 1.4, from Cordes (2007), shows this relationship. It is the time-luminosity phase-space of known (left panel) and hypothetical (right panel) transient radio sources, identifying areas which are most favourable to exploration.

Essentially, a good transient-detection instrument has similar characteristics to a large HI survey instrument, in that both require good wide-field sensitivity. They differ in that transient machines have a dependence on instantaneous observed volume. Nevertheless, there is a real possibility of commensal transient observation during surveys.

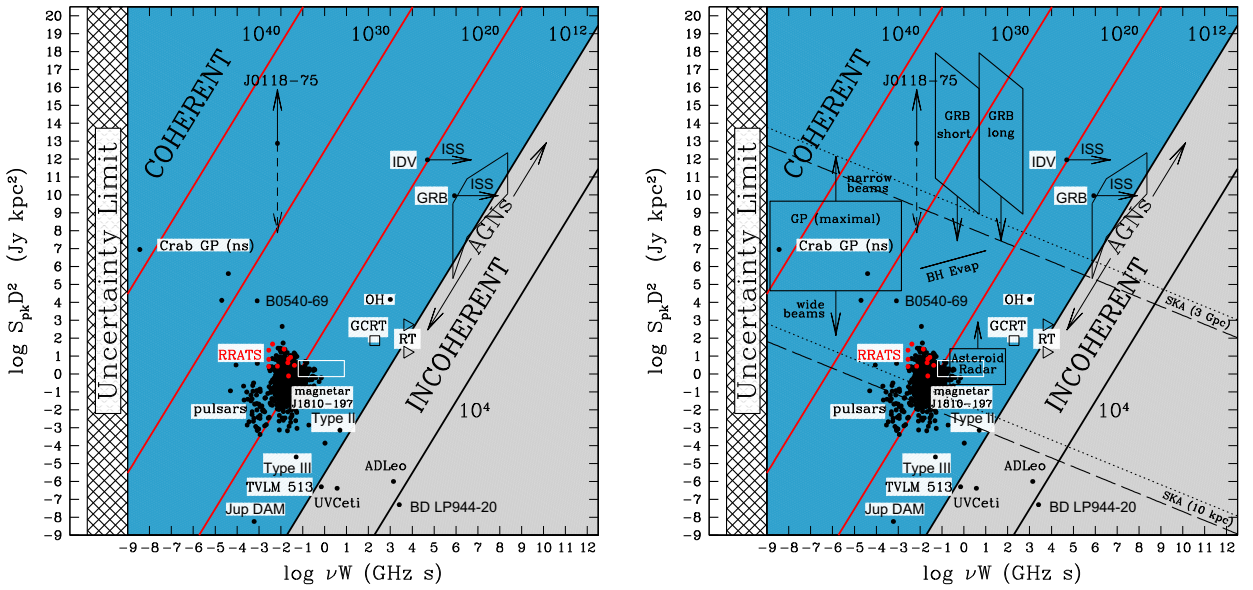


Figure 1.4: Time-luminosity phase space for radio transient sources. The left panel shows currently observed radio transients, while the right panel shows the same plot overlaid with likely areas in which as-yet undetected astronomical transients may be found. Each frame shows a log-log plot of the product of peak flux S_{pk} in Jy and the square of the observed area D^2 in kpc^2 as a function of the product of frequency v in GHz and pulse time W in s. The uncertainty limit occurs when the product of frequency and pulse time, $vW = 1$. From Cordes (2007)

1.2.3 Origin(s) and Evolutionary History of Cosmic Magnetism

The generation and cosmic evolution of magnetic fields in collapsed structure is poorly understood (e.g. Dolag et al., 2009; Gaensler et al., 2004). However, the dynamical behaviour of stars, galactic clouds and clusters in the local Universe is believed to be strongly coupled to these magnetic fields which pervade such structure (Chakrabarti et al., 1994; Miniati and Bell, 2011). A key aim in the field of cosmic magnetism is to further understanding of both the mechanisms and timescales for magnetic field generation in early and local structure as well as the evolution and persistence of magnetism over cosmic time.

Astrophysical magnetic fields are not *directly* observable. However, their properties are able to be derived through the polarisation characteristics of synchrotron radiation emission which is generated as relativistic charged particles move through them. Rotation Measure (RM) surveys of both Galactic (Beck and Gaensler, 2004) and extragalactic (Gaensler et al., 2004) sources will allow the determination of the critical relationship between structure formation and magnetic fields.

The key technological characteristic upon which both Galactic and extragalactic RM surveys rely is

excellent polarisation purity. Beck and Gaensler (2004) estimate that the great majority of polarised sources require at least −25dB of polarisation purity, and that some weakly polarised sources will require better than −40dB isolation between polarised channels for accurate RM measurements.

1.3 Future Radio Survey Instruments

The focus now shifts to the mechanism of observation. The central mechanism upon which the understanding of galaxy evolution and advances in precision cosmology and fundamental physics relies is an efficient, deep survey of the 21-cm radio sky. This requires, on the one hand, a Field of View (FoV), Ω, that is large enough to be able to cover the observable sky in a survey that takes no longer than a few years, and, on the other hand, exceptional sensitivity (the effective aperture area coupled with the system temperature, $A_{\rm eff}/T_{\rm sys}$) over a large frequency range in order to image the faint, redshifted 21-cm signal generated by cosmological HI.

In short, the efficiency of a survey is proportional to $\Omega \cdot \left(\frac{A_{\rm eff}}{T_{\rm sys}}\right)^2$. That is, the time required to map the entire sky to cosmic depth is dependent on the *sensitivity* — the square of the ratio of the effective area to the system temperature — as well as the Field of View (FoV). In this section, these requirements are developed in order to characterise the properties of future radio survey instruments.

1.3.1 Exquisite Sensitivity

Sensitivity requirements dictate a telescope with a very large collecting area. It was realised globally — and seemingly simultaneously — by the radio astronomy community that the required collecting area to image HI to cosmological distance ($z \sim 1.5$) is approximately 10^6 m^2 over a frequency range of ~ 0.6 to 1.42 GHz. There are multiple authors attributed with the original $\sim$ km^2 collecting area argument (see Wilkinson, 1991; Swarup, 1991; Dewdney and Landecker, 1991; Noordam et al., 1991), who highlight the sensitivity requirement for hydrogen surveys with a unified voice.

The requirement for tens of times larger collecting area than any current or planned radio telescope demands novel electromagnetic radiation collector technologies at all regions of the frequency range, but particularly at low radio frequencies (see, for e.g. Bell and Ekers, 2003; Schilizzi et al., 2007; Dewdney et al., 2010)[8].

[8]The series of memoranda at http://www.skatelescope.org/publications/ summarises a great fraction of the evolution of thought regarding novel collector technologies in future low-frequency radio instrumentation. The memoranda of Bell and

The important consideration from the perspective of system sensitivity is maintenance of a high ratio of effective area to system temperature over the entire observation range. The essential goal is keeping constant the ratio $\frac{A_{\rm eff}}{T_{\rm sys}} \gtrsim 2 \times 10^4$. However, no single collector technology is able to achieve this. Thus, different aperture techniques become appropriate at each of three major frequency ranges.

At the lowest radio frequencies ($\lesssim 0.25$ GHz), the noise contribution from confusing synchrotron sources is large (i.e the background sky is 'bright') and may be approximated as being proportional to the inverse square of frequency. From an instrumentation perspective, this contribution is unavoidable (i.e astrophysical) and dominates the system noise temperature $T_{\rm sys}$. Figure 1.8 shows the sky noise temperature contribution to $T_{\rm sys}$ in the radio regime. In order to maintain constant sensitivity, low frequency radio instruments should have an effective area which increases quadratically with decreasing frequency in compensation for the increasing sky temperature.

However, sky noise drops rapidly with increasing frequency, so that, in the mid-frequency radio range ($0.3 \lesssim f \lesssim 20$ GHz), it is a less dominant contributor to total system temperature. Here, reduction of the instrumental contribution to $T_{\rm sys}$ is the most significant, and is thus a key area of optimisation.

1.3.2 Large Field of View

All-sky surveys must be able to be completed in a few years. In order to achieve this requires not just exceptional signal sensitivity, but also a large Field of View (FoV).

However, large fields of view have issues for image quality, since there will be many more confusing foreground sources. This has implications for the instrumentation design to be able to handle astrophysical 'foreground' subtraction. Image quality requirements are motivated by galaxy evolution; in order to detect signals in the presence of sources up to seven orders of magnitude greater in brightness dictates an image dynamic range of $\sim 10^7$.

Another consideration is that some experiments, such as galaxy surveys, may not actually require resolved detail of smaller than $10''$ at $z = 1.5$ (HI line at 0.57 GHz), which corresponds to a cosmic volume of 85 kpc at $z = 1.5$, or an interferometric baseline of 10 km (Wright et al., 2006).

In order to efficiently achieve the required survey science (i.e. a complete all-sky HI measurement for galaxy evolution, dark energy parameter measurement, EoR evolution, all-sky rotation measures, and all-sky pulsar surveys) in reasonable time requires exceptional image quality in low-frequency, low-resolution images

Ekers (2003), Schilizzi et al. (2007) and Dewdney et al. (2010) are good reference points.

(Wright et al., 2006).

1.4 The Square Kilometre Array

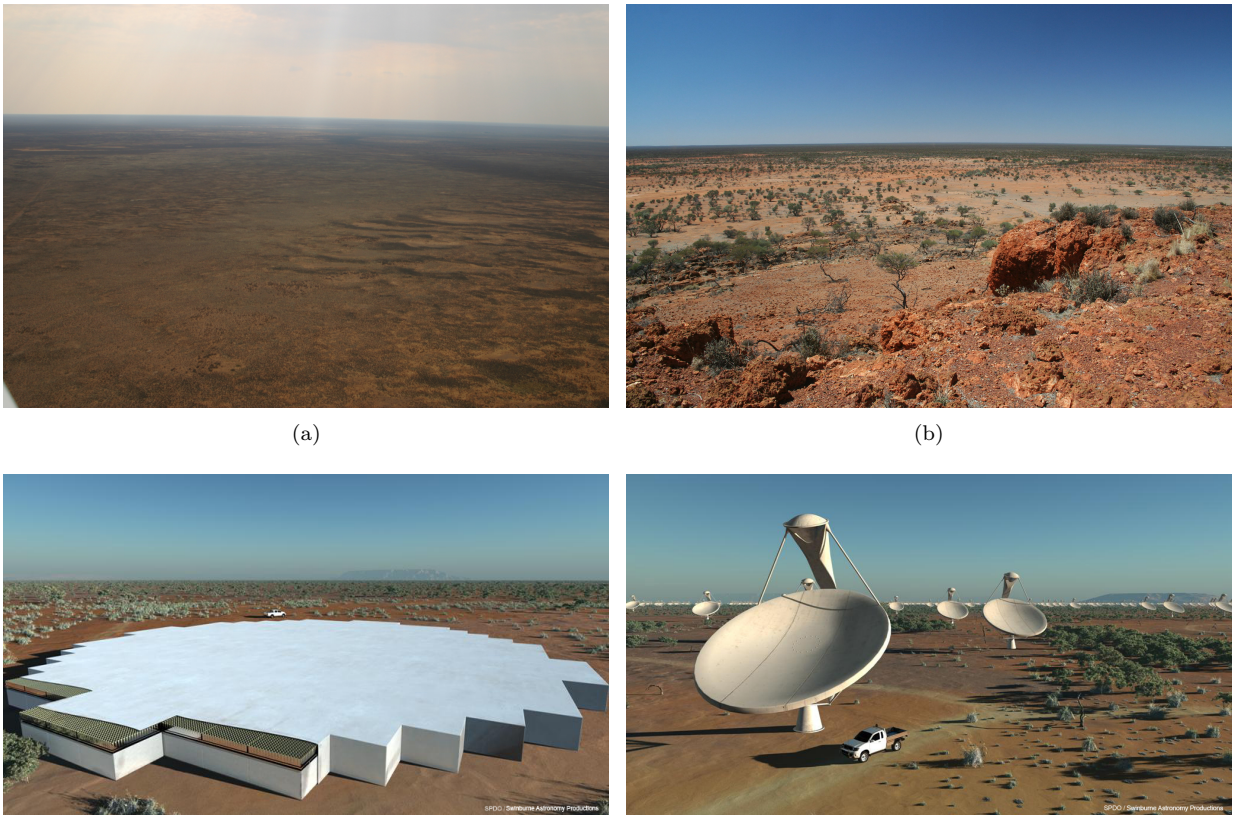
The SKA is a joint project by the international radio astronomy community to build a radio telescope with the sensitivity and survey capability to image neutral hydrogen to cosmological redshift. The SKA will enable hitherto unrealised cosmological surveys in unexplored regions of the radio parameter space; it is certain to generate new discovery.

The key scientific goals for the SKA are those of the large radio survey instrument which have been laid out in this chapter. To summarise, science targets include and go beyond those related to tomography of neutral hydrogen; in addition to a billion galaxy neutral hydrogen survey to $z \sim 1.5$ and direct imaging of the 21-cm reionisation signal, periodic and transient surveys of the Galaxy and beyond are likely to discover exceptional systems which test relativistic gravity in the ultra-strong-field limit. The stochastic background signature of gravitational waves can be inferred from variations in the pulse-arrival time from highly-accurate, short-period pulsars. The evolution and formation of magnetic fields and their relation to structure formation is promised by measurement of the magnetic fields around many extragalactic sources. At the high-frequency limit of the SKA, light will be shed on the astrophysical conditions necessary for formation of earth-size planets. Finally, and not least of all, an instrument that will survey as-yet-unexplored portions of frequency-, time- and area- space is certain to yield unpredicted discoveries. The SKA is the tool that will provide answers to some of the future fundamental questions.

The SKA will be located in a remote desert site in either South Africa or Australia[9] to minimise the effects of human-generated signals. Figure 1.6 shows a 'station' (groups of collecting area) distribution scenario for both candidate sites, showing the sheer geographical extent of the synthesised collecting area[10] of the proposed instrument.

From an instrumentation perspective, the SKA will be extremely challenging, since much of the proposed collector technology has not been used on such a large scale previously. Particularly for the focus of this thesis, the SKA is a science instrument with *unprecedented* real-time signal processing requirements. It will require the processing capability in order to synthesise large effective antennas and many electronic beams from millions of small elements, each producing gigahertz-bandwidth signals. These beams will be combined

[9] Again, www.skatelescope.org is an essential reference for the current decisions regarding instrument siting.
[10] For an in-depth treatise of the technique of aperture synthesis in radio astronomy, see Thompson et al. (2001)



(a) (b)

(c) (d)

Figure 1.5: A selection of images of the Square Kilometre Array (SKA) site. 1.5(a) is an image showing the Canarvon site in the Karoo region of South Africa. 1.5(b) shows the Murchison Radio-astronomy Observatory in the Western Australian desert. 1.5(c) is an artistic representation of an aperture array station. The station diameter is approximately 65m and the receivers are cover with a radio-transparent weather covering. 1.5(d) shows one of the candidate parabolic dishes (with an offset-Gregorian design) in the dense central core of the array.

to form a single interferometric instrument with collector stations spanning baselines of up to thousands of kilometres.

1.4.1 Collector Technologies

It is planned that the SKA will be able to cover a frequency range of approximately 70MHz to 10GHz, with a goal to provide even greater range. To be able to cover such a large frequency range and geographical area will require heterogenous aperture technologies; that is, the instrument will almost certainly be composed of multiple antenna technologies, each optimised to emission characteristics in a particular frequency range.

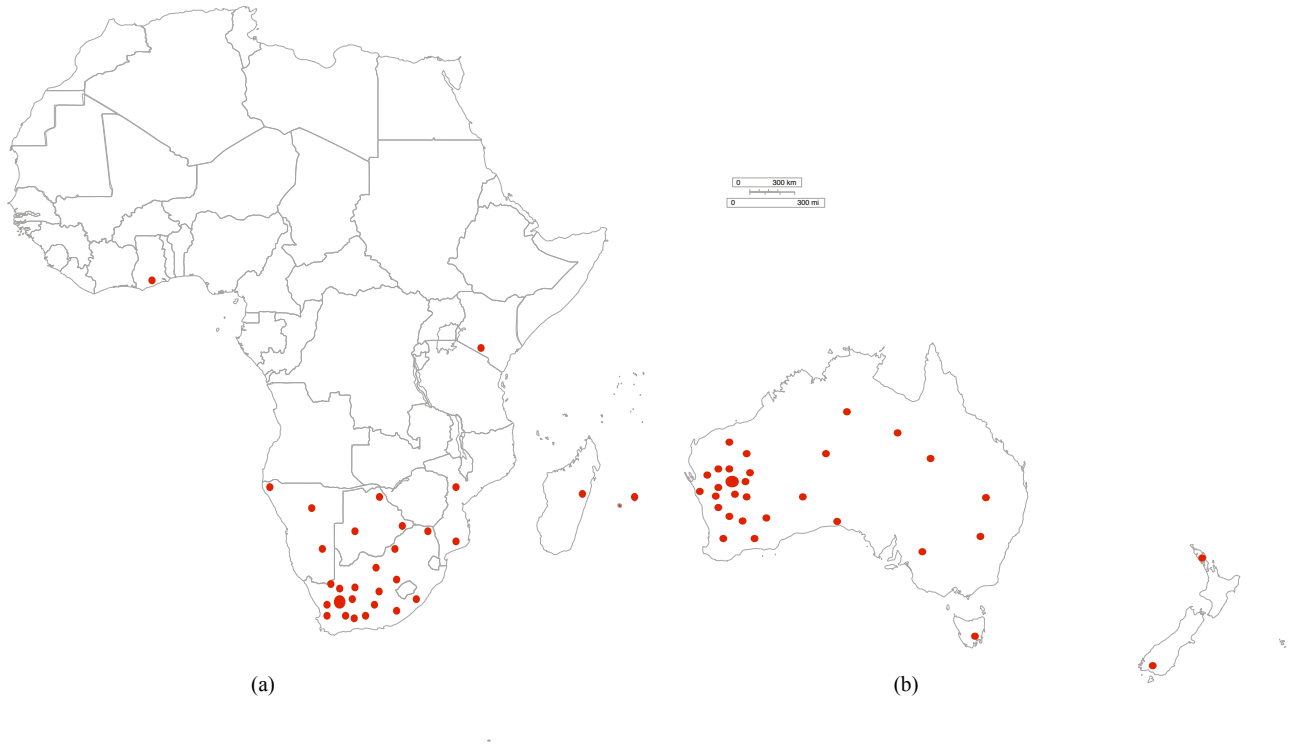


Figure 1.6: A scenario for station distribution for the Square Kilometre Array candidate sites in (a) Africa and (b) Australia. The central core will contain $\sim$ 50% of the total collecting area distributed within a radius of 5km for point-source sensitivity, a further 25% out to 200km, and the final 25% will stretch out to more than 3,000km; *continental* baselines which provide exceptional resolution.

The instrument is expected to be built in two phases, with the first phase being designed and constructed from 2016 – 2019. The first stage is to include low-frequency arrays and some of the total eventual dish contingent. The second phase will include the mid-frequency, dense aperture array and multi-pixel feeds for the dish array. In the expected phased construction scenario (e.g. Dewdney et al., 2010), sparsely-spaced aperture arrays[11] and parabolic dish arrays will operate at the low ($\sim$ 70 MHz – 450 MHz) and high (0.7 – 10 GHz) radio frequency ranges respectively. Densely-spaced, electromagnetically-coupled aperture arrays (element spacing less than the Nyquist spacing) will be part of the second phase of construction and will operate at the mid radio frequency range. Table 1.1 summarises these proposed antenna technologies, including their frequency range and current estimates for the total area contribution of each antenna technology to the final instrument.

[11] Sparse aperture arrays are electronically phased arrays where the receiving elements are not placed such as to fully sample the incoming wave-front and generally have little mutual electromagnetic coupling

Table 1.1: Antenna technologies proposed for the Square Kilometre Array (SKA) and their approximate area contributions and frequencies in the final (and Phase 1) instrument (Dewdney et al., 2010).

Technology	Approximate Area (m^2)	Frequency Coverage (GHz)
Parabolic Dishes ($\sim$ 15 m)	$\sim 10^5$ (44179)	0.7-10 (0.45 – 3)
Dense aperture array	$0.57 - 0.71 \times 10^6$ (0)	0.3 – 0.8 (n/a)
Sparse aperture array	1.3×10^6 (1.3×10^6)	0.07 – 0.45 (0.07 – 0.45)

1.4.2 Aperture Arrays

The frequency range below 1.42 GHz is critical for key Square Kilometre Array (SKA) science. Galaxy evolution and dark energy experiments rely on the deep-redshift galaxy survey. Strong-field tests of gravity and general relativity (GR) with ms-pulsar arrays relies centrally on mid-frequency observations which avoid attenuation by Galactic dust at higher SKA frequencies and severe dispersion due to free electrons at lower frequencies. Maintaining effective area at low frequencies is important. If the A_{eff} required at low radio frequencies (0.07 – 0.7 GHz) were to be provided by a filled aperture, extremely large physical collecting surfaces would required. Providing this effective area with the large parabolic antennas becomes impractical for lower frequencies. A solution to the collector technology problem in the low radio range are 'aperture arrays': arrays composed of many simple receivers, which may be combined coherently and adaptively in order to provide the required collecting area without some of the physical implementation disadvantages of parabolic collectors.

An aperture array[12] is a collection of physically-fixed, wide-beamwidth elements, where the outputs of constituent elements are combined to form a larger synthetic aperture in such a way as to direct the shaped maximum power response toward a particular spatial direction. A simple physical analogy is that the electronic combination of signals with beamforming techniques performs the same spatial filtering task as a parabolic dish does in a traditional radio telescope. However, aperture arrays need not be mechanically steered, and the number of concurrent pointings ('beams') achievable is limited only by the capacity of electronic systems. Figure 1.7 illustrates the basic principle of the beamforming operation: an incremental electronic delay is added to the output of each antenna in order to coherently interfere incoming radiation from the intended direction θ, and incoherently interfere radiation originating from all other directions.

[12]There is some ambiguity surrounding the terms "phased array" and "aperture array". The former implies that constituent array elements are individually modulated in phase to form a coherent receiver. The latter implies that the array is also a fully-adaptive optical aperture able to respond to known instrumental and transmission media instability. Both terms describe the SKA technology. The use of each term is context-specific, but they can usually be used interchangeably.

Beamformed receiver arrays have appeared recently in radio telescope instruments. The low-frequency array (LOFAR) is the largest of these; it is a primarily interferometric instrument composed of several arrays of beamformed receivers, distributed over Northern and Central Europe[13].

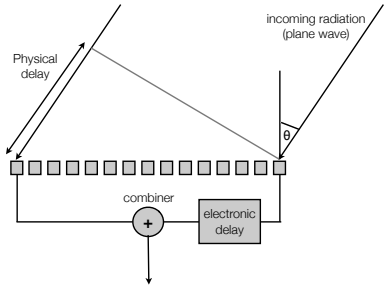


Figure 1.7: A schematic illustrating the principal of operation of a beamformer. The signal received at the rightmost element is delayed relative to the signal at the leftmost element. Similarly, an incremental delay is added to each of the other outputs in turn (not shown). Coherent interference concentrates signals from direction θ. Different delay configurations allow beams to be formed in different directions

Aperture arrays have other significant advantages over dish-based systems at low radio frequencies. Unlike a parabolic dish, the antenna elements of a phased array are not physically moved to scan the sky, removing the need for large mechanical and electromechanical control systems. Arbitrarily large apertures may be synthesised from relatively low-cost elements, allowing a large effective area, A_{eff}, to be realised. The cost of dishes has been found to scale as a power law of the dish diameter, $D^{\alpha_{\text{dish}}}$ $(2.5 \leq \alpha_{\text{dish}} \leq 3.0)$[14]. However, the scaling exponent of cost of aperture arrays α_{array} is likely to be closer to ~ 2 (*caveat lector* – the exact cost scaling function has yet to be determined for large-scale array instruments).

Maintenance of pointing accuracy and structural stability against gravitational deformations and wind loading mean that *steerable* dishes with diameters > 100m become impractical and are unlikely ever to be constructed. Structural problems are largely avoided in aperture arrays, since they are fixed on the ground. Third, there is an inverse relationship between required FoV, which is necessitated by survey time requirements, and aperture diameter, which is necessitated by sensitivity requirements (see Section 1.3.1). Subject to the computational cost of each beam, phased arrays allow an arbitrary-sized Field of View (FoV) to be filled with as many beams as required, and varied as a function of frequency.

At higher frequencies, however, with the number of elements required per unit area increasing as f^2, the

[13] see www.lofar.org

[14] The cost scaling function function for radio astronomy parabolic dishes is modelled as a power law, $C = D^{\alpha_{\text{dish}}} + \epsilon$, where α_{dish} is estimated to be 2.5 (Robertson, 1992), 2.7 (Wilkinson, 1991) and 3.0 (Weinreb and D'Addario, 2001)

cost effectiveness of arrays falls rapidly and parabolic dishes almost certainly remain the collector technology of choice.

For these reasons, it is in the sub-GHz regime that the phased array finds its critical application. A mid-frequency array promises to provide the sensitivity, speed and flexibility required for the planned radio surveys in the scientifically important frequency range below 1 GHz. This ability is not able to be provided by dish-based systems within the required costs. Digital systems have further advantages. Rapidly-decreasing costs of digital computation promise that aperture arrays might meet the cost, sensitivity and system-temperature constraints of neutral hydrogen (HI) surveys. Thus, the development of a low-noise aperture array core and associated signal processing systems is critical to the success of this survey instrument.

1.4.3 Receiver Element Spacing

The attraction of sparse-spacing of aperture arrays is twofold. They are able to provide an effective collecting area which is proportional to λ^2 (an antenna has an average effective collecting area $\langle A_{\rm eff}\rangle$ which is dependent on observation wavelength; i.e. $\langle A_{\rm eff}\rangle \propto \frac{\lambda^2}{4\pi}$). This means that there is comparatively more collecting area per receiver chain. Since the important aspect of collecting area of an aperture is its angular size in wavelengths, at low frequencies this results in great efficiency compared to parabolic dishes or even densely-spaced array collectors. Sparse arrays also provide greater resolution since the sampled aperture is larger.

In a densely-spaced array (Nyquist spacing at $\lambda_{\rm max}$, the design frequency), there is a decrease in the effective aperture at frequencies lower than $\lambda_{\rm max}$, since the aperture is oversampled at all frequencies $\lambda < \lambda_{\rm max}$. This does not imply that Nyquist spacing is misguided, simply that the effective area no longer scales with frequency. SKA collector sparseness is sensibly transitioned at the frequency at which the sky temperature levels out (i.e. ~ 0.25 GHz): sparse arrays will operate at low frequencies and will be gradually replaced with Nyquist-spaced arrays at approximately the sky temperature 'knee frequency', which is shown in Figure 1.8.

1.4.4 Digital Aperture Arrays

Thermal noise generated in a radio receiver is a central concern in instruments operating at frequencies above ~ 0.25GHz. In this frequency regime, the dominance of the total system temperature by astrophysical sources – predominantly non-thermal (synchrotron) radiation – begins to diminish. Figure 1.8 shows the sky brightness temperature as a function of frequency. In radio telescopes operating above ~ 0.25 GHz, instrumental noise, quantified as the total system equivalent noise temperature $T_{\rm sys}$, will be the most significant

contributor to the total noise budget.

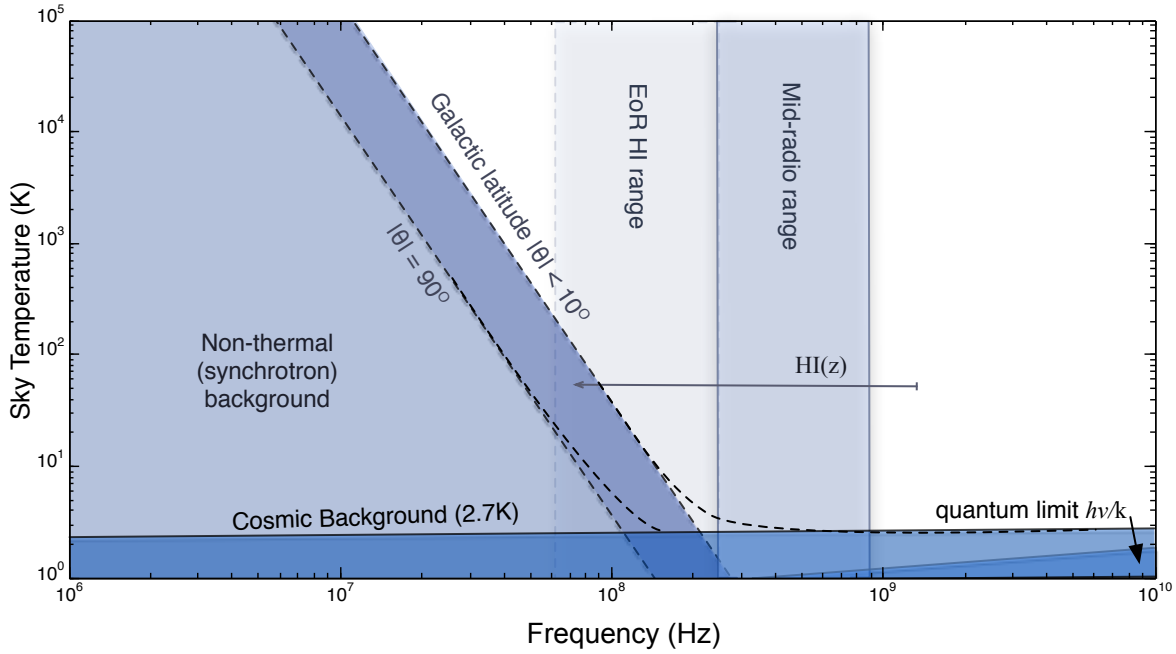


Figure 1.8: Noise temperature of the sky in the radio regime. At frequencies below ~ 0.25 GHz, the brightness is dominated by non-thermal (synchrotron) emission. The atmosphere is not an important attenuator at SKA frequencies and the strong molecular lines of O_2, OH and H_2O do not occur before 22 GHz. Receiver noise is the most significant contributor to the total system noise in the 0.25 – 20 GHz range. The importance of the mid-radio range to the measurement of $\Omega_{\rm HI}\ (z)$ is shown. Adapted from Kraus (1966) and Rohlfs and Wilson (1986).

The hardware implementation of beamforming may of be realised by various permutations of radio frequency (RF) electronics (power combiners, analogue phase-shifters and amplifiers) and digital beamforming techniques. Beamformers composed of analogue elements are subject to the known variability of their components. Digital beamformers solve some of the noise temperature issues, yet are still subject to variability of the components in the receiver chain, particularly the low-noise amplifier (LNA) and digitiser. After digitisation, digital systems do not vary with environmental factors, while analogue systems do.

Furthermore, digital systems provide the ability, during the *non-reversible* beamforming process, to correct the bandpass (frequency response), to shape the spatial response, and even to correct for ionospheric instability[15].

[15]The ionosphere may be treated as an anisotropic 'lens' that distorts the received image. If it can be correctly modelled (this is a large assumption) optical aberration correction may then be performed post-detection (e.g. Kedar et al., 2003), or in the case of digital beamformed arrays, correction can also be performed during the actual beamforming process at very little extra *computational* cost, which saves a large amount of processing energy.

However, it is perhaps easier to manage thermal effects in the receivers at the feeds of parabolic dishes. This is because the receivers at the feeds of large dishes are a single, central target for combatting radiometer noise. Cooling just the radiometer at the focus pays dividends over the entire surface area of the dish. However, aperture arrays are inherently distributed, with up to hundreds of radiometers per square metre spread over the entire aperture. This presents a logistical difficulty for cryogenic cooling for radiometer noise reduction. Because of this, aperture arrays are likely to be uncooled. Because of this, aperture elements are particularly sensitive to all other receiver noise contribution. Thus, characterisation of their systematic error is critical.

The central issue affecting large digital arrays is that they require efficient, high-performance, well-characterised, and inexpensive receivers. This includes efficient high-performance signal processing systems to process wide-band digital signals, as well as exceptional control of station beam-shape to isolate desired spatial frequencies in the presence of bright sources in the beam side-lobes. In this thesis it is argued that the central limiting factor on the adoption of fully-digital phased array is the cost of power and that from the perspective of performance, digital arrays may be made superior to their analogue equivalents.

1.5 Thesis Layout

The scientific motivations and broad technical challenges for a large radio survey telescope known as the SKA have been presented in this chapter. It is shown that sensitivity requirements motivate novel collector technologies at low radio frequencies. Digital aperture arrays are a promising collector technology in order to achieve the sensitivity required of SKA surveys.

The work described in this thesis is fundamentally concentrated on instrumentation research for digital aperture arrays. There are three central issues facing such development; a demonstration of the feasibility of the digital aperture array technology, provision of sufficient beam quality which is comparable to or better than parabolic dishes, and provision of the extreme amount of signal processing required to enable the practical use of large arrays of phased collectors in an energy-efficient way. This thesis makes a contribution to each of these three areas.

In Chapter 2, a model of astronomical signals and the effects of digital processing is developed. In particular, a methodology is developed to calculate the true quantisation noise for particular input signals, replacing the approximate, incorrect model ubiquitously applied. It is shown that discretisation is a non-

linear operation with no general closed-form analytic description, which has severe issues for integrative radiometers which use digital arithmetic in the receiver architecture (of which nearly all new radio telescopes will be). Chapter 3 presents a mathematical description and techniques of temporal and spatial digital signal filtering for the two chapters which follow.

in Chapter 4, the design and development of the digital beamforming system for the 2-PAD radio instrument is described. 2-PAD is unique in being the first mid-frequency, wide-band aperture array composed of entirely digital spatial filtering and beamforming[16]. The digital beamformer is developed with programmable hardware and open algorithm libraries. The beamformer has been deployed and tested at the array site at Jodrell Bank. The chapter presents some results which were first published in Armstrong et al. (2010) and Armstrong et al. (2009).

Chapter 5 is concerned with the problem of apodisation and element anisotropy in small arrays[17]. A digital beam-shape optimisation technique is developed in collaboration and is implemented for small beamformed arrays. It is shown that a meta-heuristic computational method, particle-swarm optimisation, is able to outperform analytic methods of element current distribution. Both the beam-width and sidelobe levels are improved compared to both the unweighted case and to traditional analytic solutions.

Chapter 6 and 7 describe work concerned with high-performance signal processing. Digital processing for aperture arrays places stringent requirements on potential processing platforms. Both computational and energy-efficiency requirements dictate the use of optimised processing architectures. Chapter 7 investigates one such architecture — the IBM BIT integer array processor — in detail. I have developed a full cycle-accurate architectural simulator and domain-specific programming language which are used to build algorithms on the array architecture. This is made available as an open-source project. Armstrong and Jones (2010) describes an initial performance analysis performed with the simulation system. Timing statistics from this work allow the absolute architectural efficiency of the IBM BIT to be derived, which uniquely position it in the 'processing-architecture space' in a sparsely-populated region of programmability, energy-efficiency and high-performance.

Chapter 8 draws some conclusions about the work done in this thesis. There are significant theoretical challenges to the correct modeling of quantisation distortion of digital signal processing and digitisation. Non-ideal elemental responses are able to be corrected with optimisation algorithms, improving the uniformity and quality of the sub-unit beam response. All-digital beamformer arrays are shown to be a viable technology,

[16] see 2-pad.physics.ox.ac.uk/.
[17] which will be the atomic processing unit of SKA station arrays

as demonstrated by a prototype array. Digital computation has a large energy cost; the immense number of processing operations demanded by digital aperture arrays can not be fulfilled by general-purpose processing architectures within energy (and thus cost) budgets. Efficient, high-performance architectures must be sought. Despite such challenges, the scientific potential of digital aperture arrays means that overcoming these issues in the near future is important.

Chapter 2

Digital System Noise

Digital systems are a central part of current and future radio instrumentation. Signal processing operations are often performed with low-bit-width digital arithmetic. Thus, it is required that the effect of digital processing be accurately described and then extracted if astronomical signals are to be uniquely identifiable in data products. This has been recognised since the early work on digital radio instrumentation (e.g. Weinreb, 1963). Analytical expressions for the digital corruption in single-quantisation digital systems exist. However, the difficulty in accurately describing *compound* digital processing means that digitally-induced error is described only by approximate models.

Discretisation theory, including sampling theory and the statistical quantisation theory are developed in this chapter in order to meaningfully analyse the systematic effects of both digitisation and digital computation on astronomical signals.

The intention of performing such analysis is that it would then inform a cascading digital computation model which would allow the quantisation noise to be correctly expressed or bounded for an arbitrary digital system, and thereby removed or mitigated. However, it is shown in this chapter that the discretisation error may be quite pathological. The statistical distribution of the quantisation error is not in general able to be precisely expressed and it co-varies with the input signal.

2.1 Introduction

Continuous time, continuous amplitude quantities are represented in digital systems by finite time-resolution, finite bit-width approximations. Such discrete signals are both *sampled in time* and *quantised in amplitude*.

In order to account for the induced error of finite-bit-width digitisers and systems, a common strategy is to use an independent quantisation noise model of the process. This model is descriptive only when the quantisation noise signal, called here $d(s,t)$, is indeed independent of the data signal $s(t)$ (i.e. when $d(s,t) = d(t)$). This chapter develops theory in order to analyse under which conditions such models are appropriate. A central result of such analysis is that signals generated by underlying processes with Gaussian statistics are in fact *not* modelled exactly by a pseudo-noise model, and simplistic corrections to binned representations of normal distributions (Sheppard, 1898) are in general an invalid correction.

However, before the merits or otherwise of digital processing models may be profitably discussed, it is important to specify the nature of desired astronomical signals, and how indeed they are to be measured.

2.2 Astronomical Signal Model

Astronomical signals received by a radiometer are generated by processes which may be described *statistically* by stochastic random variables with the properties of stationarity, ergodicity, and Gaussian distribution statistics. The exact signal value is not deterministically calculable, but the statistical properties of the signal are.

2.2.1 Stationary, Ergodic, Gaussian Random Variables

A received signal $s(t)$ may be described by its statistical properties. The vast majority of astronomical sources fall into a class that is correctly modelled within the finite observation bandwidth as being generated by a zero-mean Gaussian noise process. The normally distributed (Gaussian) probability distribution is a good model of the statistics of underlying processes of thermal and synchrotron astronomical emission (see for e.g. Thompson et al., 2001) as well as those of receiver-generated noise. The Probability Density Function (PDF) of the signal is a measure of the probability that at an arbitrary time the function has a value of x:

$$f_s(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \tag{2.1}$$

where the area of the PDF is:

$$\int_{-\infty}^{\infty} f_s(x)\mathrm{d}x = 1$$

This class of probability density functions (PDFs) are fully described by their mean μ_s and variance σ_s^2

$$\mu_s = \langle s \rangle = \int_{-\infty}^{\infty} s f_s(x)\mathrm{d}x = 0,$$

$$\sigma_s^2 = \langle s^2 \rangle - \langle s \rangle^2 = \int_{-\infty}^{\infty} s^2 f_s(x)\mathrm{d}x.$$

2.2.2 Deviations from the Stationary Gaussian Signal Model

However, and evidently, real systems will have many exceptions to the stationary Gaussian signal model. On the one hand, astronomical sources may be observed to vary in flux due to inherent underlying processes. For example, a signal from a pulsar is best described by a non-stationary Gaussian process, and Rickett (1975) and Cordes (1975) show that amplitude-modulated shot noise is a better model of such emission. These are *desired* effects which are also a deviation from the stationary signal model.

On the other hand, there also exist deviations peculiar to the observation which are generally *undesired*. Extraneous or unwanted signals such as those generated by human communication, radar, weather-monitoring, and a multitude other processes grouped as radio-frequency interference (RFI) have complex frequency and temporal structure. Knowledge of this frequency and temporal structure allows the contaminating signals to be removed, with care[1], from the data. Thus, particularly for wide frequency bands, system designs must take into account these intermittent, powerful signals which could cause system saturation and severe nonlinearity.

2.3 Signal Sensing Techniques

The central task of radio instrumentation is to uniquely identify astronomical signals. Sensing techniques, particularly detection, beamforming and interferometry are the radiometric tools which enable and improve

[1]Radio-Frequency Interference (RFI) mitigation strategies may be classified as either 'fight' or 'flight'. The simplest and most common of the mitigation tactics is temporal 'data blanking' during periods of intense interference (flight), ranging to more complex frequency-space techniques (fight). Care is merited in this process and other interference mitigation processes since they alter the statistics of the measurement. This is particularly critical in, for example, minute HI power spectrum ($P(k)$) measurements (see e.g. Furlanetto et al., 2006), as mitigation processes may contribute to the data variance in frequency (and thus, redshift) space.

unique signal detection in the presence of unavoidable systematic and environmental noise.

2.3.1 Detection

Under most conditions at radio frequencies (except for conditions of low-pressure and high kinetic energies), the Raleigh-Jeans approximation to the Planck law for the radiation spectral distribution of a black-body in the regime where $hv/kT \ll 1$ is $B(v,T) = kT\frac{2v^2}{c^2}$, is a good relation between radiative power and black-body temperature. A black-body at a temperature of T produces thermal noise with a power spectral density over the finite bandwidth of observation which may be approximated as being of uniform amplitude kT, when $v \ll c$. Many early radio telescope receiver devices consist of square-law detection circuits (Rohlfs and Wilson, 1986, see e.g.), which enables measurement of signal power, and through computation of the autocorrelation, the power spectral density. The output of such a detector is related to the *power* of the input signal:

$$y(t) = as^2(t) \tag{2.2}$$

A continuum measurement is a quantification of the total power due to an astronomical body within the finite bandwidth of observation. However, there is a power contribution to this measurement from other sources local and cosmic, including the receiver itself. Receiver analysis relies on the work of Johnson (1927) and Nyquist (1928b); a resistive circuit element transfers an average power per unit bandwidth of $\mathrm{P} = kT_N$. That is, the amount of power generated by a resistive load, an appropriate model for most of the components of a radiometer, is proportional to its temperature. Thus, over narrow observation bandwidths, both circuit elements and astronomical sources produce a signal with power proportional to their temperature. The equivalent temperature of these contributions, each referred to the input is expressed as a total system temperature,

$$T_{sys} = T_a + T_r + T_o. \tag{2.3}$$

T_a is the equivalent temperature due to the source, $T_r = T_a + T_d$ is the receiver temperature which includes the digital (T_d) and analogue (T_A) system. T_o is the contribution from other sources that may interfere with the observation which may include (depending on the desired quantity to be observed) the primordial radiation background, ground spillover noise, atmospheric thermal emission noise and other effects.

2.3.2 Relation between Digital Computation and Noise Temperature

If the exact change in Signal-to-Noise Ratio (SNR) of a particular digital computation and the input signal power is known, it may be used to relate the computation to an equivalent temperature in the above receiver model. The equivalent physical temperature for a known spectral noise power N_{di} of a digital computation may be defined as

$$\begin{aligned} T_{di} &= \int_{f_L}^{f_H} \frac{N_{di}(f)}{k} \\ &= \int_{f_L}^{f_H} \frac{P_s(f)}{k \cdot \mathrm{SNR}_{di}(f)} \end{aligned} \tag{2.4}$$

where SNR_{di} is the signal-to-noise ratio of the digital computation, and k is Boltzmann's constant. This result allows digital systems to form part of the standard radiometer noise model. It is also a general result which relies only on knowledge of the digital process signal-to-noise ratio.

Quantisation noise equivalent temperature is often calculated with the Sheppard approximation (Sheppard, 1898) to the quantisation noise, $N_{di} = \frac{\Delta^2}{12}$. In this case the integral in 2.4 evaluates to $\frac{\Delta^2}{2B \cdot k \cdot 12}$, where B is the bandwidth. Using this independent quantisation noise model[2], an ideal $m = 16$-bit digitiser operated over a $B = 0.2$ GHz bandwidth has an equivalent noise temperature of

$$T_{di} = \frac{2^{-2m}}{2B \cdot k \cdot 12} = 3514.9 \text{ K}. \tag{2.5}$$

2.3.3 Fundamental Noise Limit of a Radiometer

The accuracy and precision with which the desired quantity of T_a may be measured is limited by variation in the measurement of total power (of equivalent temperature T_{sys}) by the receiver. This fluctuation, ΔT, is given by the radiometer equation as

$$\Delta T = \frac{K T_{sys}}{\sqrt{\Delta v \tau}} \tag{2.6}$$

[2]It is noted that this calculation relies on the quantisation noise being independent and normally distributed. Accepting this assumption, quantisation noise is simply an equivalent thermal contribution which occurs deep in the radiometer chain and thus has a negligible effect on the system noise. If the assumption of Gaussian ity is not correct, however, then the thermal contribution model is incorrect.

where K is a constant, peculiar to a specific receiver architecture (K is equal to unity in the case of an ideal square-law detection circuit — Rohlfs and Wilson 1986), and the fluctuation in T_{sys} is smoothed by the square root of the number of independent measurements, the bandwidth-integration time product $\Delta v\tau$. Thus, in order to detect extremely faint source temperature fluctuations, it is desirable to maximise observed bandwidth, which requires wide-band receiver architectures, and to maximise integration times, which requires excellent system stability and control of accumulating error.

System stability is compromised heavily by receiver gain fluctuations, since the increase in antenna temperature due to the source as measured by a receiver is indistinguishable from system gain fluctuations.

$$\Delta T = KT_{sys}\frac{1}{\sqrt{\Delta v\tau}}\frac{\Delta G}{G} \tag{2.7}$$

Gain instability has traditionally been combated by switching the receiver input with a known-temperature reference noise source (Dicke, 1946). Since in this scheme half the time is spent observing a calibration source, the observation time for a specified sensitivity is increased. Other schemes exist which make continuous comparison between the signal and calibration source.

2.3.4 Spectroscopic Observation

Measurement of the fine spectral distribution of astrophysical emission allows a quantification of the total emitted power per incremental frequency channel.

Spectral details of the emission contain important information. Spectral lines are narrow frequency signal components which appear in either emission or absorption as a consequence of interaction between a quantum system (atomic or molecular) and a photon, and reveal information about dynamical atomic and molecular gas masses, temperatures and pressures in galaxies and other populations of gas.

For example, observations of pulsars and other transient signals that undergo frequency dispersion in the inter-stellar medium (ISM) (for galactic sources), or even the poorly-understood IGM (for the much rarer class of observable extra-galactic transient sources) require fine frequency measurement in order to reconstruct the original emission pulse profile. Short-time pulses in particular are excellent probes of the gas and magnetic fields through which they traverse to reach the earth's surface (Cordes et al., 2004).

Spectroscopy also has the potential to reveal information on a cosmological scale. For example, the neutral hydrogen spectral line is used to measure the volumetric distribution of matter in the Universe by using this

principal baryonic constituent as a tracer of the matter field. Lyα absorption lines from distant quasars allow the total mass distribution to be inferred as a function of redshift. This allows the acoustic oscillations of the baryons (BAO) to be measured across cosmological epoch, constraining fundamental physical parameters.

Furthermore, spectroscopic measurements additionally enable the removal of RFI from corrupted frequency ranges, which is not possible with a total power measurement.

2.3.5 Coherent Spatial Filtering: Beamforming

Beamforming is a technique that uses both temporal and spatial information to separate signal components arriving at an array of antennas. Spatial filtering may be performed in several ways: continuously, in which a focussing (parabolic) surface is used to coherently collect the incoming radiation at its focus, or discretely, where samples are taken of the electric field over the spatial aperture. If the spatial sampling of the receiver aperture is discrete, the processor which performs the spatial filtering is called a 'beamformer.'

A classification of astronomical beamformers is made according to whether the incoming wavefront is 'Nyquist sampled' or not, and are called dense and sparse arrays respectively; if spaced more than $\lambda/2$, the origin of an impinging signal may not be uniquely (unambiguously) resolved. Spatial samples that are not separated by a minimum of $\frac{\lambda}{2}$ do not contain independent information. Aliased beam components that result from spatial undersampling are termed 'grating lobes.' Narrowband, discrete spatial sampling is analogous to the well-established techniques of temporal sampling, and similar arguments for preservation of signal information content hold. Spatial filtering is returned to in Chapter 3.

2.3.6 Interferometry

Interferometry is a signal *detection* process that records cross-correlation function of electromagnetic waves received at physically separate receivers. Phase-coherent measurements of electromagnetic waves are superimposed in order derive the mutual coherence function across the aperture and thus to extract the angular distribution of the source radiation. The technique of aperture synthesis is central to modern radio astronomy, providing the resolution required to distinguish angular detail of electromagnetic radiation.

The theory of aperture synthesis is covered in detail in, for example, Thompson et al. (2001) and Rohlfs and Wilson (1986). Interferometric theory is not developed further here, being beyond the scope of the thesis, but it is noted that aperture arrays will form part of a larger interferometric instrument in the Square Kilometre Array (SKA).

2.4 Analysis of Sampling and Quantisation of Astronomical Signals

Modern radio instruments are designed to measure extremely small signal perturbations. Precise knowledge of each noise process and systematic instrumental effect is required in next-generation digital radio instruments — particularly the SKA — which will rely centrally on finite-bit-width arithmetic signal processing. For instance, HI fluctuation experiments demand a close understanding of error contributions so that the excess data variance hoped to provide evidence of re-ionisation evolution in the EoR. Signal perturbation must be known to have been generated astronomically and not to have been generated by other processes such as RFI or its statistical mitigation artefacts, ionospheric fluctuation or, most critically to this chapter, incorrectly-characterised digital instrumental effects.

Digital quantisation and computation rounding are nonlinear processes which degrade signal integrity. They are especially significant when a large number of digital processing operations are required to take place before signal detection.

2.4.1 Previous Quantisation Analysis in Correlator Design

The effect of quantisation in digital radio instrumentation is analysed by Weinreb (1963) for the single-bit digital autocorrelation spectrometer. In particular, the ability to recover the true signal autocorrelation function exactly from a sinusoidal function of the measured autocorrelation, $\rho_x(\tau) = \sin\left[\frac{\pi}{2}\rho_y(\tau)\right]$ is demonstrated, following Van Vleck (1943). Quantisation noise for multi-bit systems has been considered from the perspective of its effect on the correlation functions of stochastic variables by Burns and Yao (1969), Hagen and Farley (1973), Jenet and Anderson (1998) and Thompson et al. (2001).

However, the Van Vleck relation, and similar relations for multi-bit correlator permutations (e.g. Hagen and Farley 1973) fails to be accurate when many intervening calculations occur prior to detection. Modern digital spectrometers, beamformers and correlators must therefore compensate for systematic, pre-detection digital quantisation distortion.

The Friis-depth[3] of the digital system in a radiometer has led its equivalent temperature to be regarded as

[3]The Friis-depth, $\mathfrak{F}_i \in \mathbb{N}_1$, is the index of the ith term in the Friis noise-figure cascade equation, $F_{sys} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_M - 1}{\prod_{m=1}^{M-1} G_m}$, with F_i the noise factor and G_i is the power gain of the i-th component in the radiometer chain.

negligible, especially for bit-widths $n_b \geq 8$ (e.g. Rohlfs and Wilson, 1986; Thompson et al., 2001). However, noise contribution is only able to modelled as a fixed noise temperature if it is indeed of uniform spectral power density. Non-linear quantisation processes have spectral components that do not combine incoherently, since quantisation spectra do not generally follow those of the Gaussian distribution. Furthermore, even though a single quantisation may not dominate the signal, the significant number of operations required in finite-bit-width digital systems means that the quantisation error for even moderate to large bit-widths (> 16) becomes significant. The analysis developed in this chapter is that of pre-detection (or pre-correlation) operations of signal filtering, beamforming, beam-shaping and equalisation.

2.5 Sampling Theory

Signal discretisation is an inherently non-linear operation and thus must incorporate information loss. A discrete signal measurement attempts to preserve in a finite number of samples of finite precision the entire information content of the original signal. There exist techniques which ensure that all discretisation non-linearity can be confined either to the finite-precision time measurements or to the finite-precision amplitude measurements.

The theory of sampling provides sample-timing bounds for which information preservation is possible. The sample-separation criterion is attributable to many authors. Shannon (1949) state the sampling theory formally and give a proof. It was, however, stated as being generally applied in the field earlier by Nyquist (1928a) in the theorem of distortionless transmission for application to telegraphy as well as in parallel but unconnected work by Kotelnikov (1933). Earlier still, it was developed under the mathematics fold in the *Interpolation Theory of a Cardinal Function* by Whittaker (1915) and later refined in Whittaker (1935). It is central to modern signal processing, and is known variously as some combination of the surnames of Nyquist, Shannon, Whittaker and Kotelnikov, or simply as *the sampling theorem*.

2.5.1 The Non-Linearity Dilemma

This early work in signal information theory has shown that by fixing the sample rate such that sampling occurs at known, periodic instants of minimum guaranteed regularity, and ensuring that the amplitudes of a sampled signal are represented *precisely*, the information content of the sampled signal can be shown to be preserved. The satisfied sampling theorem allows the the non-linearity of discretisation (the general

'discretisation error') to be shifted to that incurred due to the finite-precision representation of amplitude levels.

The other possibility in performing signal discretisation is that amplitude level thresholds are fixed at a guaranteed maximum separation and precise samples in time be taken when the continuous signal amplitude crosses a particular level. This will also allow perfect signal reconstruction as long as the amplitude levels are not spaced by more than certain separation, but with the 'opposite' requirement of precise time samples. Pursuing this approach would shift the discretisation error to that incurred through finite-precision representation of time samples.

Signal processing theory is built on the former assumption: that time samples are exactly periodic and precise, while amplitude measurements are of finite precision. Neither technique avoids information loss; quantisation loss is inevitable. The non-linearity inherent to discretisation is not avoided by following the sampling theorem, but rather confined to the finite-precision amplitude representation problem.

2.5.2 Frequency-Domain Description of Sampling

Sampling theory is well-understood and the foundation of modern signal processing. This section is included for completeness in the discussion of digital representation of continuous quantities. Consider a time-varying continuous signal $x(t)$ at the digitiser input sampled at uniform time intervals of T seconds. This process is represented mathematically by multiplying a sequence of Dirac delta functions (referred to as an impulse carrier), each with area T,

$$c(t) = \sum_{k=-\infty}^{\infty} T\delta(t - kT)$$

with the original function $s(t)$ to give

$$s(t) \cdot c(t) = \sum_{k=-\infty}^{\infty} s(kT)T\delta(t - kT)$$

Expressed as a complex Fourier series, the impulse carrier is

$$c(t) = \sum_{n=-\infty}^{\infty} e^{jn\Omega t}$$

where the Fourier coefficient of each term in the series is given by

$$\frac{1}{T}\int_{-T/2}^{T/2} c(t)e^{-jn\Omega t}dt = \frac{1}{T}\int_{-T/2}^{T/2} T\delta(t)e^{-jn\Omega t}dt = 1$$

(i.e. it has a value of unity for each term of the series). The Fourier transform of this series is then

$$\mathcal{F}\{c(t)\} = \mathcal{F}\left\{\sum_{k=-\infty}^{k=\infty} T\delta(t-kT)\right\} = \sum_{k=-\infty}^{k=\infty} T\delta(\omega - n\Omega)$$

By the convolution theorem, the product $s(t)\cdot c(t)$ in the time domain is equivalent to the Fourier-domain *convolution* of $S(j\omega)$ and $C(j\omega)$,

$$S(j\omega) \star C(j\omega) = \mathcal{F}\left\{s(t)c(t)\right\}$$
$$= \sum_{n=-\infty}^{n=\infty} S(j\omega - jn\Omega) = S_s(j\omega)$$

Where the sampled signal is the frequency-domain convolution of the input signal and a function that is a set of sampling impulses.

2.5.3 A Sufficient Criterion for Signal Recovery: The Band-limited Sampling Theorem

In the simplest expression, the *Sampling Theorem* asserts that it is possible to recover a time-series signal exactly from its sampled representation if the frequency replicas of $S(j\omega)$ in the sampled frequency-domain signal $S_s(j\omega)$ do not overlap. Shannon (1949) formulates the sampling theorem as follows:

Sampling Theorem *If a function contains no frequencies higher than W cycles per second (cps) it is completely determined by giving its ordinates at a series of points spaced $\frac{1}{2W}$ seconds apart.*

By definition, the Fourier transform of $s(t)$, $S(\omega)$ has no frequency components outside of the band W.[4]

[4]This derivation has proceeded for a 'base-band' signal: a signal that has frequency components confined to the contiguous band which extends from 0 Hz. The sampling theorem is readily generalisable to *band-limited* signals (those with frequency components confined to a contiguous frequency range not necessarily starting at 0 Hz). Recent work (see for e.g. Candes et al., 2005, and references therein) shows that the band need not even be contiguous. It is possible to recover sparse signals (i.e. those with sparse frequency-domain support) with high accuracy from a small number of stochastic measurements. If this 'compressed-sensing' technique is implemented systematically in the temporal domain, and the information content matches a sparse frequency support basis, then there are implications for the energy-efficiency of sampling and subsequent processing processes.

Then

$$\begin{aligned} s(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega t}d\omega \\ &= \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} S(\omega)e^{j\omega t}d\omega \end{aligned} \tag{2.8}$$

Let $t = \frac{n}{2W}, t \in integers$. Then

$$s\left(\frac{n}{2W}\right) = \frac{1}{2\pi}\int_{-2\pi W}^{2\pi W} S(\omega)e^{j\omega\frac{n}{2W}}d\omega \tag{2.9}$$

It can be seen that the left side of Equation 2.9 contains the values of $s(t)$ at the sampling instances, and the right side is the Fourier series expansion of $S(\omega)$.

2.6 Statistical Quantisation Theory

Quantisation is a non-linear mapping of the signal voltage to a discrete set of amplitude levels for storage within a *finite* number of bits on a digital system. It is performed when an analogue signal is transferred to a digital computer, and also each time a calculated quantity is stored back in the representation bit-width. Importantly, such quantisation error is not actually independent of the input signal; it is in general *deterministically related* to the input signal.

The simplicity of the description and formulation of the quantisation process belies its theoretical complexity. Signal quantisation is a non-linear functional mapping from the continuous-amplitude domain to the discrete-amplitude domain. Analysis of non-linear operations such as quantisation is problematic. Quantisation is, however, a linear operation on the Probability Density Function (PDF) of a given signal. Linear analysis is enabled through operations upon the statistical distribution of the signal; the PDF of the quantised signal may be obtained through exclusively linear operations on the PDF of the continuous-amplitude signal (see, e.g. Bennett, 1948; Widrow, 1956). This enables some useful results, unobtainable otherwise, but does not generally give exact expressions for the quantisation distortion. The central tenet upon which statistical quantisation analysis is based is that although quantisation acts in a non-linear manner on signals, it acts in a linear manner on their corresponding statistical distributions.

Uniform Quantisation

A quantiser may be described as a non-linear, single-valued function of an input variable. In general, n-bit quantisation is a non-invertible mapping of the signal value $s(\mathrm{n})$ into a set of indices $i \in \{0, 1, \ldots, N-1\}$. The continuous real line of signal amplitudes is divided into disjoint sets $\{\mathcal{R}_j\}_{j=0}^{N-1}$. Each set is a monotonic section of the real values which are taken on by the continuous input signal.

The index i maps invertibly to a single output code $c_i \in \{c_k\}_{k=0}^{N-1}$ if the input signal at the quantisation time falls into the region $\mathcal{R}_i$. If the monotonic sections have equal amplitude span, the quantiser is known as a uniform quantiser. If, further, the analogue input values fall within the 'non-overloaded-range' of the digitiser (i.e. within specified input signal power limits) , the quantisation deviation will never be greater than $\Delta/2$ (where Δ is the quantum size), and will be approximately uniformly distributed in this range for a large-amplitude, *non-zero-frequency* input.

The difference between the input signal, s, and the quantised output, s', is the quantisation error signal, d. Figure 2.1 schematically shows a uniform quantiser and the relation between input and output variables.

Quantisation Level Optimisation

Quantisation levels need not be uniform. For instance, in order to minimise signal distortion or maximise Signal-to-Noise Ratio (SNR), either the input threshold levels or the output level codes respectively may be dynamically set. Since the desired quantity of an observation is a time series of power spectral measurements, unbiased power measurements are best acquired using power-optimised levels (Jenet and Anderson, 1998).

Other level optimisation targets different aspects of the measurement. Max (1960) develop a numerical method to calculate *voltage*-optimised signal levels. Lloyd (1982) develop a least-squares quanta-spacing scheme for both Gaussian and Laplacian input signals. Arbitrary level thresholding is achieved equivalently by signal pre-distortion followed by uniform quantisation.

2.6.1 Quantisation Analysis in Linear Signal Processing Systems

Signal quantisation also occurs each time an arithmetic operation is performed upon a digital value. The original signal is re-quantised as many times as the signal is scaled and then stored; in particular it is re-quantised after each addition or multiplication.

A multi-stage signal processing quantisation model describes the effective non-linear quantisation from

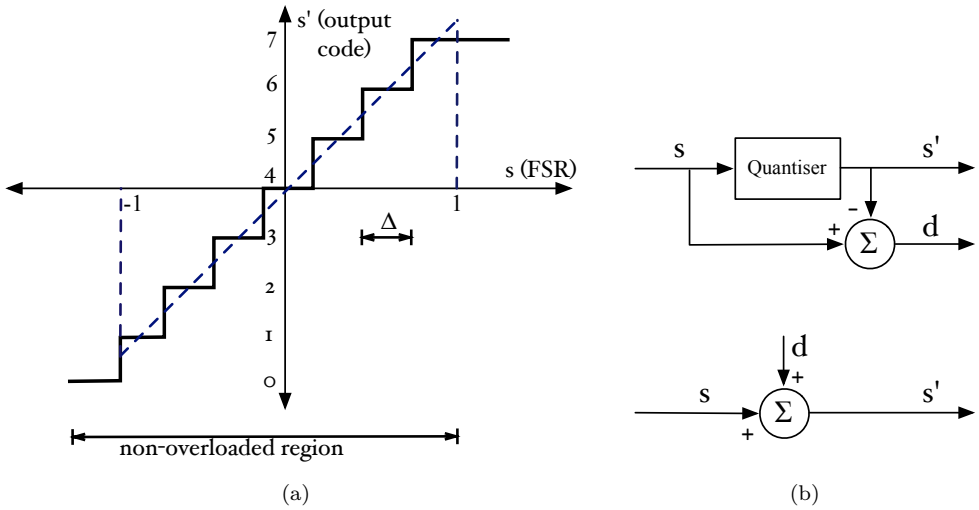


Figure 2.1: Schematic of an ideal uniform quantiser. 2.1(a) shows the output digital code as a function of the input analogue level. Since there is not a level-change at the zero analogue input level, this class of quantiser is referred to a 'mid-tread' quantiser. If, as is natural in binary digital systems, output codes are in units of 2^{M}, $(M \in \mathbb{N})$, this gives rise to an inequality in output code range for negative and positive input values. On the other hand, 'mid-rise' quantisers have the output level step shifted by a half quantum such that a level change occurs at the zero analogue level and an equal number of positive and negative input value output codes. Uniform quantisation suits signals with unknown distribution statistics. Optimisation of quantisation levels is justified when properties of the input signal are known (see e.g. Max, 1960; Lloyd, 1982; Jenet and Anderson, 1998). Figure 2.1(b) shows the relationship between the input signal, $s(t)$, the quantised signal, s', and the quantisation noise, d. The quantisation signal is shown as an additive signal that is deterministically related to the input signal and thus generally correlated with it.

the perspective of the input signal. Initial quantisation occurs in the Q_0 digitisation operation. Subsequent scaled additions and scalar multiplications are modelled as a gain stage proceeded by a re-quantisation. The output signal s'^N is modified by the cumulative effective quantisation and gain of all the quantisation stages $\prod_{i=1}^{N} Q_{i-1}G_i$, as shown in Figure 2.2.

2.6.2 The Independent Quantisation Noise Model

It is almost ubiquitous in signal processing analysis both within and beyond radio astronomy to model the quantisation process by representing the output of the quantiser as a sum of the input signal and an independent Gaussian noise signal which represents the quantisation noise addition (e.g. Hagen and Farley, 1973; Gray, 1990; Proakis and Manolakis, 1996; Thompson et al., 2001).

This is called here the Independent Quantisation Noise (IQN) model. The practice dates to work on statistical moments of signal distributions by Sheppard (1898) for corrections to 'binned' or histogram

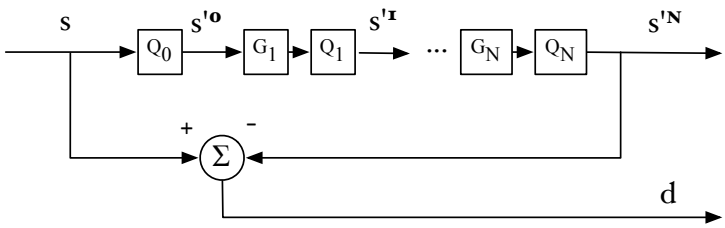


Figure 2.2: Schematic of an otherwise linear signal processing operation from the perspective of non-linear quantisation operations performed on the input signal. Gain stages collapse to pure quantisation stages: a gain of 2^n applied before a 2^m bit quantisation is equivalent to a 2^{n-m} bit quantisation. Gain loses meaning for $m \geq n$ and is replaced by a multiplication by 0.

approximations to normal (Gaussian) probability distributions. It was seen to be an equivalent process to integer rounding in early quantisation analysis by Bennett (1948).

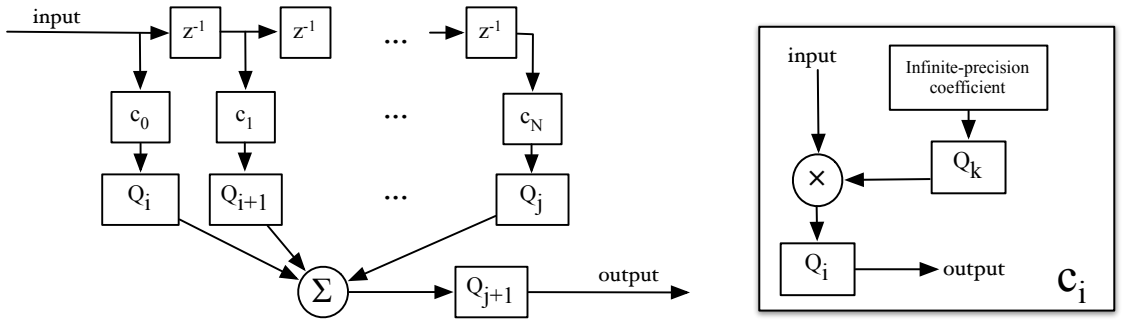
In the model, independent, zero-mean quantisation noise is uncorrelated with the input signal s and is uniformly distributed over the range of a quanta $(-\Delta/2, \Delta/2)$. The expected value of the sequence, $E\{d\} = 0$, the mean square of d, $E\{d^2\} = \sigma_d^2 = \dfrac{\Delta^2}{12}$, mean fourth is $E\{d^4\} = \dfrac{\Delta^4}{80}$ and so on. Odd higher-order moments are zero. Since $d(n)$ is an ergodic, stationary white sequence, $E\{d(n)d(m)\} = 0, (n \neq m)$.

2.6.3 Linear Signal Processing Analysis with the Independent Quantisation Noise Model

Using the Independent Quantisation Noise (IQN) model, the approximate quantisation error of a processing operation may be estimated by treating each quantisation step as an independent, additive noise process. This is the model adopted by most quantisation analysis.

The general Finite Impulse Response (FIR) filter is taken as an example. Each arithmetic operation is an effective re-quantisation from the intermediate result to the representation bit-width. Figure 2.3 shows the architectural structure of the FIR filter consisting of inherent arithmetic quantisation. Each quantisation operation adds to the total processing noise. The output noise power of the filter is then estimated as

$$\begin{aligned} P_{\text{f\_noise}} &= \frac{1}{12}\left\{\text{coefficient quantisation} + \text{tap quantisation} + \text{sum quantisation}\right\} \\ &= \frac{1}{12}\left\{\sum_{c=i}^{j} Q_c^2 + \sum_{t=k}^{l} Q_t^2 + 1\right\}. \end{aligned} \tag{2.10}$$



(a) (b)

Figure 2.3: Schematic of an finite impulse response filter structure showing quantisation contributions. The signal is re-quantised in each filter tap after multiplication with the filter coefficient. It is further re-quantised after the summation of all filter taps before production of the output sample. This is shown in 2.3(a). 2.3(a) shows filter coefficients modelled as infinite-precision numbers which are quantised to the filter bit-width. This quantisation also contributes to the final quantisation noise.

where the Q_c is coefficient quantisation and Q_t is quantisation due to tap multiplication re-quantisation. Tap quantisation is determined by both the signal bit-width and the filter coefficient normalisation. The output Signal-to-Noise Ratio (SNR) of the filter is

$$\begin{aligned} \mathrm{SNR_f} &= \frac{P_\mathrm{s}}{P_{\mathrm{f\_noise}}} \\ &= \frac{12\kappa(2^m)^2}{\sum_{c=i}^{j} Q_c^2 + \sum_{t=k}^{l} Q_t^2 + 1} \\ &= \frac{12\kappa(2^{2m})}{N_c + N_t + 1}. \end{aligned} \tag{2.11}$$

where m is the bit-width and K is a constant dependent on the input signal power. Particularly, $\kappa = \frac{1}{2}$ for a sinusoidal input.

2.7 Physical Digitisers

Every physical digitiser (a device which performs both sampling and quantisation) is inherently noisy to some extent. By this it is meant that the exact time at which a sample is measured as well as the exact signal level range which is assigned to a given quanta is subject to noise corruption as a function of frequency,

input level and environmental factors. Deviations from the ideal are generally measured for a particular digitiser, which characterise the deviation of a particular n-bit digitiser from the ideal in terms of noise performance, spectral contamination, and differential and integral nonlinearity. Of course, an *ideal* n-bit digitiser *also* corrupts the signal in a predictable way, defined generally by the quantisation theory. Initial digitisation often represents the most significant contribution to the system quantisation noise, when signal processing is performed with a much greater number of bits than used in digitisation. In this section, a mathematical framework for digitisers is introduced for both the real and the ideal digitiser. In particular, standard performance criteria are introduced and are related to radiometer modeling framework of equivalent receiver temperature characterisation.

2.7.1 Ideal Digitiser

An n-bit digitiser, $\mathcal{D}_n[s]$, is a non-invertible mapping of the signal value, $s(n)$, into a set of indices $i \in \{0, 1, \ldots, N-1\}$. The continuous signal amplitude set is divided into disjoint sets $\{\mathcal{S}_k\}_{k=0}^{N-1}$. If the input signal at the quantisation time falls into the region $\mathcal{S}_i$, it is mapped to an index i. In turn, i maps invertibly to a single output code, $x_i \in \{x_k\}_{k=0}^{N-1}$. A digitiser is modelled mathematically as:

$$i = r_p \left\{ \left(\frac{2^N}{\mathrm{FSR}} \right) \cdot (V_{\mathrm{in}} - V_{\mathrm{low}}) \right\} \tag{2.12}$$

where FSR is the full-scale range of the digitiser ($\mathrm{FSR} = V_{\mathrm{high}} - V_{\mathrm{low}}$) and $r_p\{\cdot\}$ is an arithmetic rounding operation that is specific to the type of digitiser used. Particularly, it computes the 'floor' function ($r_{\mathrm{floor}}\{x\} \triangleq x_k - \mathrm{mod}\{x_k, 1\}$) for a mid-rise digitiser and the 'round' function ($r_{\mathrm{round}}\{x\} \triangleq \mathrm{int}\{x_k\}$) for the mid-tread digitiser.

2.7.2 Characterisation Standards

Characterisation standards provide a quantification of the deviation of particular digitiser implementations from the ideal. Physical digitisers are composed of particular architectures — such as the sigma-delta modulator, which are further complicated since they use feedback quantisation loops (e.g. Gray et al., 1989; Chang and Bibyk, 1999) — as well as variable analogue measurement components. IEEE (1994) is a standard for measurement of this characteristic digitiser error based on a monotone input signal. Isolated frequency-specific performance is then extrapolated to wideband performance, which can be problematic in that this is

not a specification of the device under actual signal conditions, and it has been shown that the digitisation process is both mathematically and practically different for these input signals.

Digitiser performance metrics related to the total noise power are the Signal-to-Noise-and-Distortion Ratio (SINAD) and its pseudo-inverse, the Effective Number-Of-Bits (ENOB). The SINAD is the ratio of the root-mean-square (rms) of a sinusoidal input signal, $A \cdot \sin(2\pi f_0 t + \phi) + C,\ (A \simeq \text{FSR})$, to the rms of the noise

$$\text{SINAD} = 20\log_{10}\frac{A}{\sqrt{2 \cdot P_{\text{noise}}}} \tag{2.13}$$

where the factor of $\sqrt{2}$ is due to the sinusoid possessing half of the signal power of the complex exponential. If the input signal is shown to cause independent quantisation noise, the ideal digitiser is able to be modelled by IQN, and will have a noise power of $\Delta^2/12$ and rms $\Delta/\sqrt{12} = \dfrac{\text{FSR}}{\sqrt{12} \cdot 2^n}$. Thus, the ideal SINAD becomes

$$\text{SINAD}_{\text{ideal}} = 20\log_{10}\frac{\sqrt{12} \cdot A}{\sqrt{2} \cdot q} = 20\log_{10}\frac{\sqrt{6} \cdot 2^n A}{\text{FSR}}$$

$$= 20\log_{10}\sqrt{6} \cdot 2^n, \quad \text{where } \zeta = \frac{A}{FSR} \triangleq 1), \tag{2.14}$$

which equates to approximately 7.782 dB of distortionless range for the single-bit digitiser which increases by 6.021 dB per bit as the number of bits n increases. The ENOB of the digitiser is the inverse of this relationship,

$$\text{ENOB} = \frac{\log_2 10}{20}\text{SINAD} + \log_2\frac{1}{\sqrt{6} \cdot \zeta}. \tag{2.15}$$

These identities may be used to derive the equivalent thermal noise temperature of a digital process. Then the equivalent temperature of the digitiser with known SINAD $\mathcal{S}_d$ is

$$\begin{aligned} T_q &= \frac{P_{\text{noise}}}{k} \\ &= \frac{A^2}{k} 10^{(-\mathcal{S}_d/10)} \end{aligned} \tag{2.16}$$

and for the ideal quantiser case,

$$T_q = \frac{A^2}{k\sqrt{6}\zeta 2^n} \tag{2.17}$$

Other metrics characterise device consistency through input signal power and frequency range . Differential and integral measures of digitiser non-linearity display the worst-case deviation from ideal output code

behaviour. Spectral performance of the digitiser is specified by the spurious free dynamic range (SFDR) and the Total Harmonic Distortion (THD), which are ratios of the input signal to, respectively, the most powerful harmonic and the average of the first N harmonics.

2.8 Statistical Properties of the Quantisation Noise

The Independent Quantisation Noise (IQN) model is a simplifying model that has been used prevalently to model the effects of integer rounding. Its history is traced to the analysis of Sheppard (1898) on numerical signal moment corrections for quantised probability distributions.

Despite its ubiquity in quantisation noise models, an IQN model is not generally correct. This was known from early in the history of quantisation analysis. Clavier et al. (1947) show that a sinusoidal input signal will not have a white quantisation noise spectrum and derive an approximate result (since no closed-form expression is possible) showing clear deviation from the white noise case. Bennett (1948) later derives an approximate error function for a Gaussian input signal, which again is not white in the spectral domain. Bennett concludes, however, that the quantisation noise is *approximately* independent if

- the signal amplitude does not exceed the quantiser limits
- the quantiser has a *large* number of levels
- the probability distribution of the input signal is smooth
- the signal is not of zero frequency

Widrow (1956) give a statistical analysis of quantisation operation through a probabilistic treatment of the input signal. The results of this work provide conditions on the probability distribution of the input signal under which the IQN model is correct. Again it is shown that IQN does not model quantisation noise in general. In particular, an input signal with Gaussian distribution satisfies none of the quantisation conditions and is not correctly modelled by the IQN model. However, Widrow confirms Bennett's heuristic prediction of approximate independence of a quantised Gaussian process under the same conditions with a statistical analysis. Widrow concludes that the deviation between the IQN and the true quantisation noise when the input signal has a Gaussian probability distribution is very small when the quantisation step size is less than about half of one standard deviation of the input, and in such cases IQN is a reasonable model[5].

[5]Analysis is also performed for other input signals, using numerical approximations. Gray (1990) and Widrow and Kollar

Vardeman (2005) is less generous to the justification of the IQN model, stating that it is fundamentally unjustified in its application to integer-rounding moment corrections. Vardeman shows that it is precisely incorrect to apply Sheppard's corrections to integer rounding of a continuos signal, and that there is no mathematical relationship whatever between the two. Sheppard's corrections are shown to be applicable to integer rounding only in certain highly-contrived situations.

What is nevertheless clear is that the IQN model is certainly not without merit. Indeed, when the probability distribution of the input signal obeys certain criteria, the quantisation noise is shown to be independent of the input signal (see for e.g. Widrow, 1956; Sripad and Snyder, 1977; Widrow and Kollar, 2008). The IQN model turns out to be a reasonable approximation to actual quantisation noise due to the exponential roll-off of the Gaussian probability distribution.

2.8.1 Statistical Properties of Astronomical Signals

The central issue in astronomical digital processing is that signals received by low-frequency radiometers do not satisfy conditions guaranteeing quantisation noise independence, nor do they satisfy even those conditions that predict approximate independence. Firstly, the received signal is a combination of the predominantly Gaussian sky signal and contaminant RFI of much greater amplitude with generally unknown probabilistic distribution. Furthermore, expecting RFI to be limited even to narrow bands may be naïve (see e.g. Lemmon, 1997). Gaussian noise analysis that predicts little deviation from the nominal error becomes invalid in this case. Second, in order to ensure against saturation from intermittent powerful RFI bursts, the astronomical signal must be confined to at most the bottom few bits of the digitiser range (see e.g. Parsons et al., 2010; Pober et al., 2011). This means that the astronomical signal standard deviation, σ_{sky}, may be required to be represented by approximately a single quantum of the digitiser and will thus not satisfy Widrow and Kollar's standard deviation criterion. Third, signal integration is the key technique to increase the SNR of buried signal components. Non-white spectral quantisation distortion which has sharp spectral components will also sum coherently at the same rate as the SNR (i.e. as $\sqrt{N_{\text{int}}}$). Such complicating factors are intensified when the quantisation occurs multiple times on each input sample, as it does for finite-bit-width signal processing after quantisation.

The independent quantisation noise model is an estimate of the quantisation error that tends toward an

(2008) present separate analysis for sinusoidal signals; both are forced to provide upper and lower bounds to quantised signal moment deviation from the IQN model due to the complexity of the approximate functions.

optimistic result: for a general input signal, quantisation noise can invariably be shown to deviate from it. This section will introduce a probabilistic analysis of the quantisation noise which was first formulated by Widrow (1956), and give some expressions for certain input signals.

2.8.2 Probability Density Function of the Quantisation Noise

The quantisation input variable is continuous and takes on the continuum of values. The output takes only discrete values. Quantisation noise is the difference between the continuous signal and the quantised signal $d = s - s'$. Its probability density function is an incoherent sum of quantum-sized sections of the original signals PDF:

$$f_d(x) = \begin{cases} \sum_{m=-\infty}^{\infty} f_s(-x + m\Delta) & \frac{-\Delta}{2} \leq x \leq \frac{\Delta}{2}, \\ 0, & \text{elsewhere.} \end{cases} \tag{2.18}$$

$$= \text{rect}(x) \cdot \sum_{m=-\infty}^{\infty} f_s(-x + m\Delta) \tag{2.19}$$

Essentially, the Probability Density Function (PDF) of the quantised signal is uniformly distributed over the interval $\frac{-\Delta}{2} \leq x \leq \frac{\Delta}{2}$. The characteristic function of a signal is the Fourier transform of its probability density function and allows calculation of actual quantised signal moments, and is able to show deviations from the IQN signal model.

2.8.3 Characteristic Function of the Quantisation Noise

The Characteristic Function (CF) of the quantisation noise is obtained by Fourier transformation of the probability density. Then the CF is composed of the Fourier Transform of the window function, $\mathcal{F}\{\text{rect}(\Delta)\} = \Delta\text{sinc}\left(\frac{\Delta\omega}{2}\right)$, convolved with the Fourier Transform of the original signal's PDF, which is

$$\begin{aligned} \mathcal{F}\left\{\sum_{m=-\infty}^{\infty} f_s(x - m\Delta)\right\} &= \frac{1}{2\pi}\Phi_s(\omega)\left(\cdots + e^{i2q\omega} + e^{iq\omega} + 1 + e^{-iq\omega} + e^{-i2q\omega} + \ldots\right) \\ &= \frac{1}{2\pi}\Phi_s(\omega)\sum_{l=-\infty}^{\infty} e^{il\Delta\omega} \end{aligned} \tag{2.20}$$

Since this function is periodic, it is equal to a Fourier series representation[6]

$$\Phi_s(\omega) \sum_{l=-\infty}^{\infty} e^{il\Delta\omega} = \Phi_s(\omega) \sum_{k=-\infty}^{\infty} T\delta\left(\omega - kT\right) \tag{2.21}$$

where

$$T = \frac{2\pi}{\Delta}. \tag{2.22}$$

Then the CF is the convolution

$$\begin{aligned}
\Phi_d(u) &= \frac{1}{2\pi}\Phi_s(\omega) \sum_{l=-\infty}^{\infty} e^{il\Delta\omega} * \Delta \cdot \mathrm{sinc}(\frac{\Delta\omega}{2}) \\
&= \frac{\Delta}{2\pi}\Phi_s(\omega) \sum_{k=-\infty}^{\infty} T\delta\left(\omega - kT\right) * \mathrm{sinc}(\frac{\omega}{2}) \qquad (2.23) \\
&= \sum_{k=-\infty}^{\infty} \Phi_s(-kT) \cdot \mathrm{sinc}\left(\frac{\Delta(\omega - kT)}{2}\right). \qquad (2.24)
\end{aligned}$$

This, the CF of the quantisation noise is a series representation in which delayed and advanced sinc functions are multiplied with the CF of the input signal. Widrow (1956) show that the properties of this characteristic function in terms of its distribution in sampling-radian-frequency space directly influences independence of the quantisation noise from the input signal.

Particularly, if the replicas of the CF of the input signal overlap in amplitude-transform space, 'aliasing' of probability distributions occurs, and the PDF of the original signal may not be recovered exactly from the quantised output. Heuristically, this condition is somewhat analogous to the sample-timing condition of the sampling theorem.

There are other, weaker conditions, proposed by Sripad and Snyder (1977), Gray (1990) and Widrow et al. (1996) which allow the recovery of all or some of the input signal moments. Knowledge of all signal moments allow reconstruction of the original signal and thus a way to measure the true quantisation noise deviation.

[6] A periodic train of exponentials may be represented by a Fourier series of Dirac sampling functions:

$$\frac{1}{T} \sum_{n=-\infty}^{\infty} e^{i2\pi n\omega/T} = \sum_{k=-\infty}^{\infty} \delta(\omega - kT)$$

Moments of the Quantisation Noise

The mean and higher moments of d may be obtained by the values of the nth derivatives of its Characteristic Function (CF) at the origin. The nth derivative of $\Phi_d(\omega)$ is

$$\begin{aligned}\frac{\mathrm{d}^n\Phi_d(\omega)}{\mathrm{d}\omega^n} &= \frac{\mathrm{d}^n}{\mathrm{d}\omega}\sum_{k=-\infty}^{\infty}\Phi_s(-kT)\cdot\mathrm{sinc}\left(\Delta(\omega-kT)/2\right)\\ &= \sum_{k=-\infty}^{\infty}\Phi_s(-kT)\frac{\mathrm{d}^n\mathrm{sinc}\left(\Delta(\omega-kT)/2\right)}{\mathrm{d}\omega^n}.\end{aligned} \tag{2.25}$$

The sinc function may be evaluated by making use of Euler's result,

$$\mathrm{sinc}(x) = \prod_{n=1}^{\infty}\cos\left(\frac{x}{2^n}\right)$$

as well as the nth derivative of the co-sinusoidal function, which is

$$\frac{\mathrm{d}^n\cos(x)}{\mathrm{d}x^n} = \begin{cases}\sin(x), & \{4n;\ \forall n\in\mathbb{Z}\},\\ -\cos(x), & \{4n+1;\ \forall n\in\mathbb{Z}\},\\ -\sin(x), & \{4n+2;\ \forall n\in\mathbb{Z}\},\\ \cos(x), & \{4n+3;\ \forall n\in\mathbb{Z}\}.\end{cases} \tag{2.26}$$

With these results, 2.25 becomes

$$\frac{\mathrm{d}^n\Phi_d(\omega)}{\mathrm{d}\omega^n} = \sum_{k=-\infty}^{\infty}\Phi_s(-kT)\prod_{k=1}^{\infty}\frac{\mathrm{d}^n}{\mathrm{d}\omega^n}\cos\left(\frac{\Delta(\omega-kT)}{2^{k+1}}\right). \tag{2.27}$$

Then the nth moment of d is the nth derivative evaluated at $\omega = 0$

$$\mathrm{E}\{d^n\} = \frac{1}{i^n}\left.\frac{\mathrm{d}^n\Phi_d(\omega)}{\mathrm{d}\omega^n}\right|_{\omega=0}$$

Dithering

Using the CF frequency-replica result, Gray (1990) show that when an independent process $r(n)$ with known, uniform PDF (and thus 'band-limited' CF) is used to modulate the input signal before quantisation, the quantisation error is able to be made independent of the input signal. This is known as 'dithering' and has

broad application inside and outside of quantisation analysis. While a real possibility for initial digitisation (e.g. Gammaitoni, 1995; Palonpon et al., 1998), it is not practical to digital processing since it requires the dither signal $r(n)$ to be removed post-quantisation, quintupling (at least) the processing requirement and vastly increasing the complexity of each digital operation. There is also a 'recursive' quantisation problem in ultimately digital processing, since a quantised approximation of the dither signal must be used to remove it from the quantised signal.

2.8.4 Quantisation Analysis for a Stationary Gaussian Input Signal

A stationary, zero-mean Gaussian-distributed input signal has a characteristic function that is not band-limited, nor does it meet the conditions in order for its associated quantisation noise to be uniformly distributed, nor independent of the quantiser input, nor even orthogonal to it. That is, there is a strong deterministic relationship between the quantisation noise and a Gaussian input signal.

The moments of the quantisation noise for a zero-mean Gaussian input are obtained from its CF. The expected value is the first moment of the quantisation noise,

$$\begin{aligned}
\mathrm{E}\{d\} &= \frac{1}{i}\left.\frac{\mathrm{d}\Phi_d(\omega)}{\mathrm{d}\omega}\right|_{\omega=0} \\
&= \sum_{k=-\infty}^{\infty} \frac{\Phi_s(kT)}{i}\left.\frac{\mathrm{d}\cdot\mathrm{sinc}\left(\frac{\Delta(\omega-kT)}{2}\right)}{\mathrm{d}\omega}\right|_{\omega=0} \\
&\qquad \left(\text{where } \frac{\mathrm{d}\cdot\mathrm{sinc}(z)}{\mathrm{d}z} = \frac{\cos(z)}{z} - \frac{\sin(z)}{z^2}\right) \\
&= \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \frac{\Phi_s(kT)}{i}\frac{\Delta}{2}\frac{(-1)^k}{k\pi} \\
&= \frac{2}{T}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \Phi_s(kT)\frac{(-1)^k}{ik}. \qquad (2.28)
\end{aligned}$$

This is the general moment formulation. The Gaussian Characteristic Function (CF)[7] is now substituted for

[7]The CF of the zero-mean Gaussian distribution is given by the integral

$$\Phi(\alpha) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha} e^{-t^2/2}\,dt = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\alpha}{\sqrt{2}}\right)\right], \quad \alpha \in \mathbb{R}.$$

where erf is the 'error function' and is defined by the preceding line. The Taylor expansion of the error function $\mathrm{erf}(x)$ is

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\left(x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - \cdots\right)$$

Φ_s in Equation 2.28 to give

$$\begin{aligned}
\mathrm{E}\{d_\mathrm{n}\} &= \frac{2}{T}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt\frac{(-1)^k}{ik}\\
&= \frac{2}{T}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\left[1+\mathrm{erf}\left(\frac{kT}{\sqrt{2}}\right)\right]\frac{(-1)^k}{ik}\\
&= \frac{4}{T}\sum_{k=1}^{\infty}\left[1+\mathrm{erf}\left(\frac{kT}{\sqrt{2}}\right)\right]\frac{(-1)^k}{ik}\\
&= 0. \qquad (2.29)
\end{aligned}$$

where the $-\infty$ summation index in the second line of Equation 2.29 collapses since $\Phi_s(-x) = -\Phi_s(x)$. Similarly to the first moment, the second moment (signal power) is calculated from the value of the second derivative of the CF at $\omega = 0$ and is

$$\begin{aligned}
\mathrm{E}\{d^2\} &= \sum_{k=-\infty}^{\infty}\frac{\Phi_s(kT)}{i}\left.\frac{\mathrm{d}^2\cdot\mathrm{sinc}\left(\frac{\Delta(\omega-kT)}{2}\right)}{\mathrm{d}\omega^2}\right|_{\omega=0}\\
&= \frac{1}{3}\Phi_s(0)\frac{\Delta^2}{4} - \sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\Phi_s(kT)\frac{\Delta^2}{4}\frac{2(-1)^{k+1}}{i^2k^2}\\
&= \frac{\Delta^2}{12} + \frac{2}{T^2}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\Phi_s(kT)\frac{(-1)^k}{i^2k^2}.
\end{aligned}$$

Substitution of the Gaussian CF gives for the second moment

$$\begin{aligned}
\mathrm{E}\{d_\mathrm{n}^2\} &= \frac{\Delta^2}{12} + \frac{2}{T^2}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\left[1+\mathrm{erf}\left(\frac{kT}{\sqrt{2}}\right)\right]\frac{(-1)^k}{i^2k^2}\\
&= \frac{\Delta^2}{12} + \frac{4}{T^2}\sum_{k=1}^{\infty}\left[1+\mathrm{erf}\left(\frac{kT}{\sqrt{2}}\right)\right]\frac{(-1)^k}{i^2k^2}. \qquad (2.30)
\end{aligned}$$

The higher moments are calculated similarly. These results may then be compared with the IQN model moments. Widrow and Kollar (2008) take the first few terms of the Taylor expansion and evaluate the first

80 terms of the infinite sum when $\Delta = 2\sigma_s$ to give the numerical result

$$\mathrm{E}\{d_\mathrm{n}^2\} \simeq \frac{\Delta^2}{12}(1 - 0.0876)$$

This is a deviation of order 10^{-1} from the IQN-predicted result per quantisation operation, when the quantisation step-size is particularly coarse. Moment deviation will also be present for each of the even higher-order moments, and also for the odd higher-order moments when the mean of the input signal is not zero.

This may seem like a positive result, in that to first order, the Gaussian signal will have approximately independent quantisation noise, even when the quantisation step size is relatively large. In fact, the deviation of the second moment is shown to be less than would be the independent noise contribution. However, there remains the complication that the noise is correlated with the input. Thus, it will not sum incoherently if successive samples of the input signal are not independent (i.e. if the input signal is not an independent random variable). Since integrating radiometers rely on instrumental noise summing incoherently in order to increase the effective SNR of weak, coherent signals, coherent quantisation noise will have an unwanted larger integrative effect.

2.8.5 Arbitrary Distribution Deviation

A signals with arbitrary distribution may be analysed in a similar manner provided it is known — particularly, if it has a known Probability Density Function (PDF) — and is a mathematical function (i.e. it is differentiable everywhere). Second moments are then calculable by

$$\begin{aligned}\mathrm{E}\{d^2\} &= \sum_{k=-\infty}^{\infty} \frac{\Phi_s(kT)}{i} \left.\frac{\mathrm{d}^2 \cdot \mathrm{sinc}\left(\frac{\Delta(\omega - kT)}{2}\right)}{\mathrm{d}\omega^2}\right|_{\omega=0} \\ &= \frac{\Phi_s(0)\Delta^2}{12} + \frac{2}{T^2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \Phi_s(kT)\frac{(-1)^k}{i^2 k^2}. \end{aligned} \tag{2.31}$$

For instance, the deviation from the IQN due to a sinusoidal function is calculated. The PDF of the sinusoid is

$$f_s(x) = \begin{cases} \frac{1}{\pi\sqrt{A^2 - x^2}}, & |x| < A, \\ 0, & elsewhere\}. \end{cases} \tag{2.32}$$

(Bennett, 1948) gives the CF as

$$\Phi_s(\omega) = J_0(A\omega) \tag{2.33}$$

where J_0 is the $0th$ order Bessel function of the first kind. Moment calculation requires an approximation to the Bessel function be used. Next, the deviation from the IQN due to band-limited (with band from $-v_0$ to v_0) white noise is calculated. The PDF of this input will be

$$f_s(x) = \begin{cases} \dfrac{1}{2A}, & |x| < v_0, \\ 0, & \text{elsewhere.} \end{cases} \tag{2.34}$$

and the CF will be the sinc function: $\Phi_s(\omega) = \mathrm{sinc}(A\omega)$. Its first moment will be

$$\begin{aligned} \mathrm{E}\{d\} &= \frac{2}{T} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \Phi_s(kT) \frac{(-1)^k}{ik} \\ &= \frac{2}{T} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \mathrm{sinc}(AkT) \frac{(-1)^k}{ik}, \end{aligned}$$

which has no closed-form solution. Numerical approximation is again required to be used to find the approximate moment deviations. This is problematic; even some of the simplest input signals have no closed-form analytic expression of the distortion under quantisation. More complex representative signals, such as the sinusoid plus noise, may not even have closed forms for their probability distributions (see, for e.g. Webster, 1983, who present an approximation to the probability density distribution of a sinusoid-plus-noise).

2.9 Deviation of N_q for a Non-Stationary Input Signal

The results of the analysis for arbitrary signals motivates two conclusions. The first is that the quantisation noise is correlated with the input signal, and will co-vary with the input. Purely Gaussian signals have rapidly decreasing correlation coefficients between quantisation noise and the input signal, and thus the noise is *practically* independent. However, arbitrary signals, particularly astronomical signal models, do not have this property and there is significant correlation. The second is that due to the lack of an analytic form of the quantisation noise, analytic modeling of the input signal is not generally possible. Quantifying quantisation noise power may be determined by simulation. However, simulation must be performed for each

input signal class, and will be different for each.

Quantisation noise power is calculated for three classes of representative input signals. The first is the white Gaussian noise (WGN) sequence. The spectrum of this signal is shown in the first panel of Figure 2.4; this signal has uniform spectral power over frequency. The probability density function of the quantisation noise, $f_d(x_1)$ is, as expected, approximately uniform, shown in Figure 2.7. A fit is made to the histogram of occupation densities for 10^6 quantisation operations, and also plotted in the figure.

The second example is a sinusoidal signal embedded in WGN. Its spectrum is shown in the second panel of Figure 2.4; the sinusoid has a frequency of $0.2f_s$. The signal time-series and its occupation density histogram is shown in Figure 2.6. The shape of the underlying probability density function traced by the histogram shows a sinusoidal probability density function modulated with a Gaussian noise kernel. The probability density function of its associated quantisation noise signal $f_d(x_2)$ is not uniform over the range $-\Delta/2 \leq x \leq \Delta/2$, and the fit to the PDF histogram in Figure 2.7 shows a periodic structure.

The third example signal is a set of random-amplitude band-limited WGN embedded in wide-band WGN to simulate radio interference sources. The spectrum of this signal is shown in the bottom panel of Figure 2.4. The powerful, band-limited WGN noise appears at $0.125f_s$, $0.285f_s$, and $0.485f_s$. The input signal and its probability distribution are shown in Figure 2.5.

The quantisation noise power is calculated for the input signals over a bit width range of $b \in \{4, 6, 8, 10, 12, 16\}$. The quantisation noise power varies with bit width as well as the input signal. This is shown in Figure 2.8. The figure shows that the variation of the quantisation noise power depends on the presence of non-Gaussian structure in the input signal. Thus, digital quantisation must be treated as a non-Gaussian error component which is measured for a particular input signal.

2.9.1 Concluding Remarks on Quantisation Analysis

The central conclusion of the statistical quantisation analysis is that although digital radio instruments rely on finite-bit-width signal processing, quantisation distortion inherent to such processing is not generally correctly modelled or removed. Furthermore, closed-form analytic expressions do not exist for quantisation distortion of most simple input signals, which means there is no simple replacement for the independent noise model. Simulations have been shown to yield different results for the quantisation noise power due to different input signals.

It may be argued that the error may be made arbitrarily small by using an arbitrarily large number of

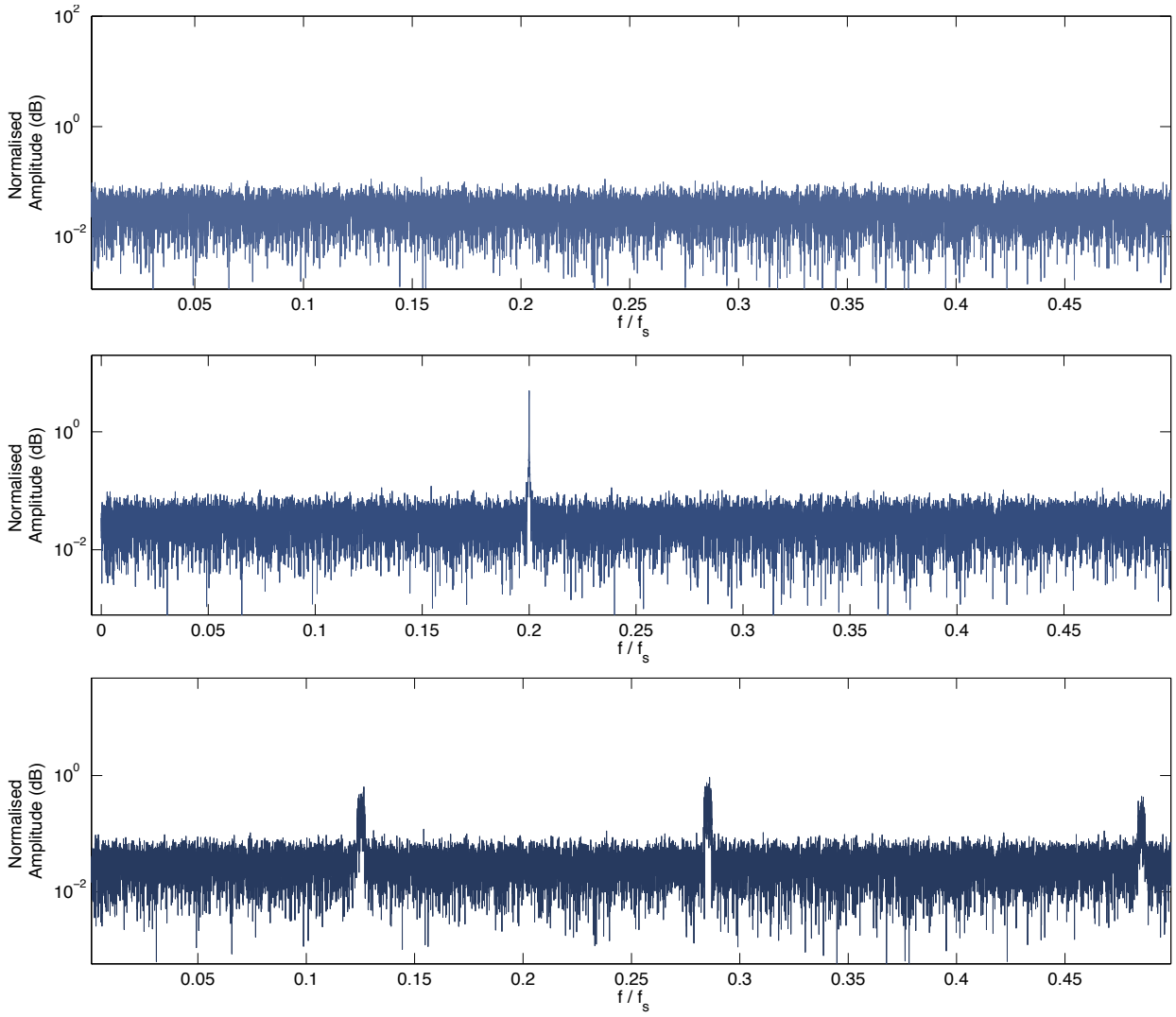


Figure 2.4: Spectra of the input signal models. The top panel is the spectrum of a white Gaussian noise sequence. The middle panel shows an embedded sinusoidal signal at $0.2f_s$. The bottom panel shows an expected signal model, which consists of Gaussian noise with embedded narrowband components at arbitrary positions in frequency space.

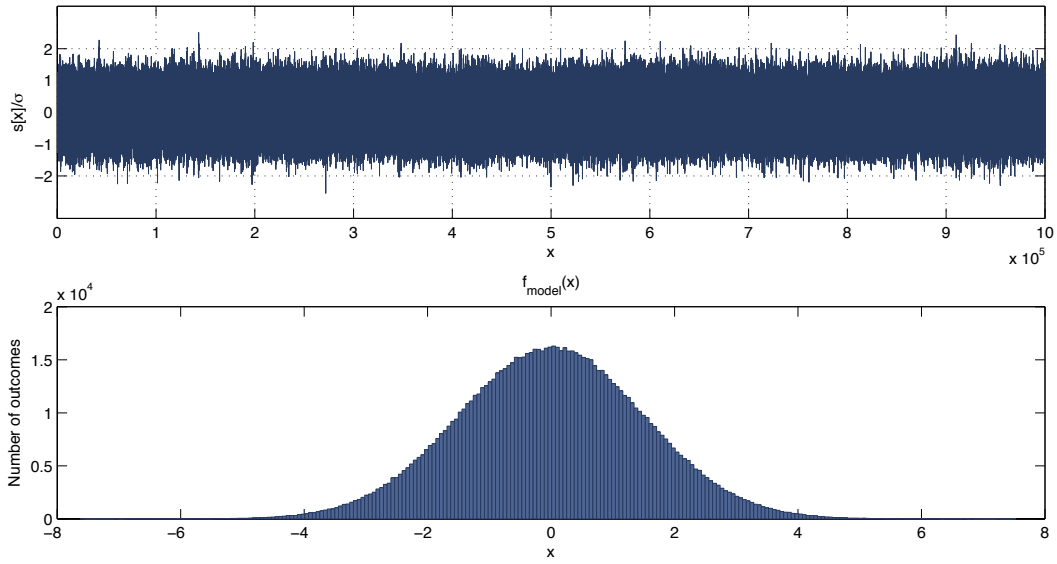


Figure 2.5: Time series plot and probability distribution of the interference signal model. The underlying probability function is clearly normally distributed. The log likelihood of a normal fit with a mean of 1.76×10^{-3} and variance 1.56 is is -1.64×10^{6}. Jarque-Bera test of composite normality: the null hypothesis — i.e. the data are normally distributed with zero mean — cannot be rejected at the $p \leq 10^{-3}$ level).

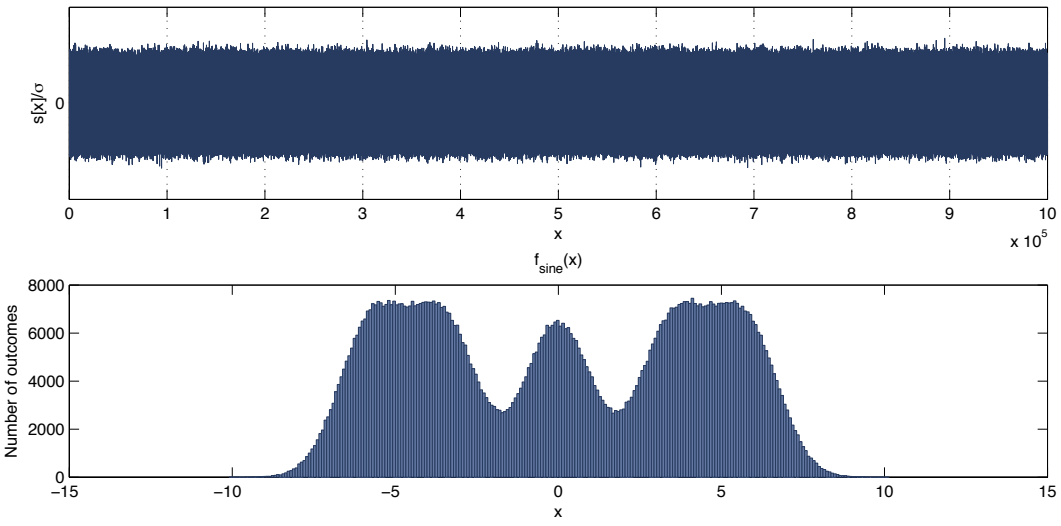


Figure 2.6: Time series plot and probability distribution of the embedded sinusoid signal model. The shape of the underlying probability density function shows a sinusoidal probability density function modulated with a Gaussian noise kernel.

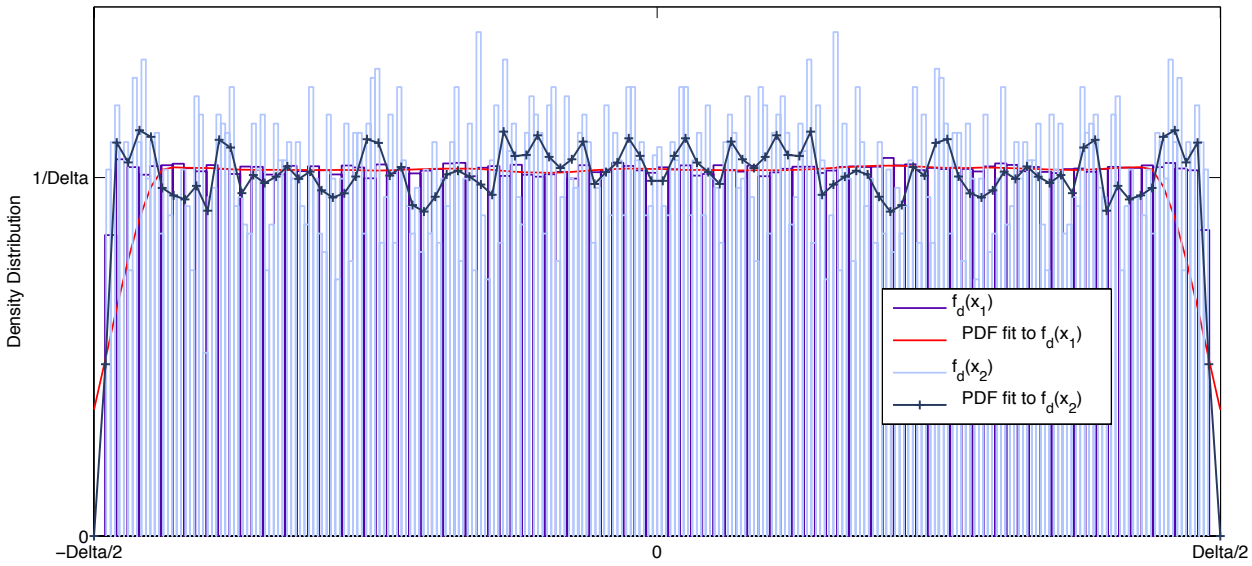


Figure 2.7: Probability density distributions of the quantisation noise for input signals of WGN, $f_d(x_1)$, and an embedded sinusoid, $f_d(x_2)$. The density distribution of the quantisation noise, $f_d(x_1)$ is, as expected, approximately uniform, as shown by the histogram plotted in red. A fit is made to the histogram of occupation densities for 10^6 quantisation operations, which is an approximation to the underlying density distribution, and is plotted in dark red. The probability density distribution of the embedded sinusoid's quantisation noise signal, $f_d(x_2)$, is not uniform over the range $-\Delta/2 \leq x \leq \Delta/2$; the fit (dark blue) to its histogram (light blue) shows a periodic structure.

representation levels. There are two issues with this argument. It may not be possible, from an energy perspective to make the number of bits arbitrarily large. It will be shown in Chapter 6 that an energy-cost argument requires that low-bit-width systems be used. Secondly, the problem is not removed by using more representation bits; it is simply reduced. Signal detection in radio interferometry relies on integration to improve the signal-to-noise ratio (SNR) to detectible levels. Non-Gaussian signal components will increase with the signal. Thus, even very small spectral components due to to quantisation may be comparable in magnitude to the astronomical signal, and are not able to be attenuated by incoherent summation. The predictability of the input data influences bit-depth decisions; the number of quantisation levels is dictated by the acceptable level of quantisation noise for a particular observation.

Modeling and Measurement Strategies

With an understanding of this inherent systematic problem, the distortion may be corrected approximately. This is done by either precise measurement or numerically approximate modeling of the distortion of a

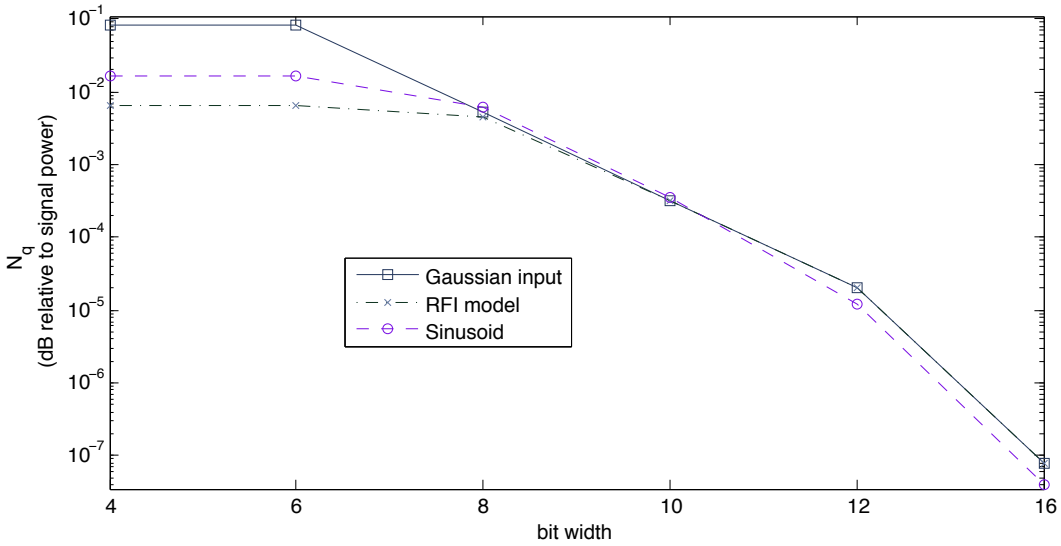


Figure 2.8: Quantisation noise power N_q relative to input signal power in a logarithmic scale for three input signal models as a function of bit width, $b \in \{4, 6, 8, 10, 12, 16\}$.

particular system for an expected input signal model. In this work, and in particular the digital system for 2-Polarisation All-Digital (2-PAD) in Chapter 4, the approach taken is to normalise the system response for an WGN noise input, since the input signal is likely to be modulated by a low-amplitude WGN signal. In general, the more accurately the input signal model is known or is able to be approximated, the more accurately the quantisation distortion may be removed by a correction matrix.

2.10 Summary

Since finite-bit-width digital systems will be a central part of most new digital radio instruments, full characterisation of systematic digital processing effects is merited. This chapter has introduced telescope receiver signal sensing techniques, including detection, spectroscopy, beamforming and interferometry. Quantisation noise is the only distortive process that occurs in digital systems.

A popular model of quantisation noise analysis, the Independent Quantisation Noise (IQN) model, is reviewed. A statistical quantisation analysis shows the assumption of quantisation noise independence is not generally correct. This finding is important since it shows that quantisation is a systematic distortion that covaries with the input signal.

The central conclusion of the statistical analysis is that quantisation distortion is not correctly modelled

by IQN; models which rely on this assumption must do so with care. However, statistical modeling does not either provide a convenient approximation, since closed-form analytic expressions do not exist for quantisation distortion of most input signals. Simulations have shown that different representative input signals give different results for the quantisation noise power.

Chapter 3

Digital Filtering of Astronomical Signals

On their journey toward the earth, the photons emitted at the moment of decoupling of energy and matter are joined by other photons on the same path, at first by those emitted by chance atomic transitions of hydrogen atoms in a largely uniform universe, and later as the gravitational force starts to have a cosmic effect by those emitted by luminescent ordinary matter from clustered structures, by *Bremsstrahlung* caused by deceleration of charged particles in dense gas regions of the galaxies encountered along the path, and later still by those photons originating from the stars, gas, dust and compact objects in our own galaxy before finally encountering photons from man-made sources within the earth's own atmosphere.

All these arrive at detectors on the telescopes of earth at the same time, producing a signal which contains a contribution from each of the astrophysical, cosmological, and terrestrial processes to which they owe existence. The challenge of astronomy is to disentangle the contributions from each of these sources, that the history of our universe — through the emission history of photons — becomes clear.

A radio telescope is essentially a signal filter; astronomical radiometers attempt to separate particular elements in the received signal through filtering, interferometry and integrative detection. This chapter expands on the concepts of discrete spatial filtering (beamforming) and discrete temporal filtering at radio-frequencies. It is a signal-theoretical introduction to the core concepts of digital beamforming that will form a base for digital design and beam-shape optimisation in the chapters that follow.

3.1 Digital Spatial Filtering

A wideband beamformer is both a spatial and a temporal filter. It separates signals of co-incident frequency and direction-of-arrival (DoA) through measurement of the electric field over its aperture. If this measurement is performed with a discrete set of sensors, the spatial filter is known as a beamformer. The spatial discrimination potential of the array is dependent on the size in wavelengths of the aperture, and is thus dependent on frequency.

Expressed in matrix-vector notation, the uniform linear array (ULA)[1] narrowband beamformer output is

$$\mathrm{b}_i(n) = \mathbf{w}^H \mathbf{s}(n) \quad (0 \leq i \leq N_b - 1) \tag{3.1}$$

where $b_i(n)$ is the beamformer output, $\mathbf{w}^H$ is the complex conjugate (Hermitian) transformation of the weight vector, N_b is the number of beams and $\mathbf{s}(n)$ is the signal vector. A wideband beamforming operation is here expressed as a set of vectors of each narrowband response. The wideband matrix equation is then

$$\mathbf{b} = \mathbf{W}^H \mathbf{s} \tag{3.2}$$

where the time index n has been dropped for clarity.

The 'beam' describes the spatial response of the aperture. An input signal is modulated with the beam and integrated over all θ to give a (complex) scalar quantity representing the array response in the nominal beam direction. In essence, the beam shows power contributions to the beamformer output from all spatial directions. Thus motivates the desired quantities of maximally-suppressed side-lobes (which are caused due to the finite extent of the aperture) and a narrow main-beam width.

The dynamic range of a beamformer instrument is the difference between the least-distinguishable power — which depends on the side-lobe level, system noise floor and integration time — and the largest representable power, which depends generally on input signal power and bandwidth.

3.1.1 Definitions Based on the Monotone Response

The response of the beamformer to a sinusoidal input allows several definitions to be conveniently derived. The output of the beamformer as a result of a complex sinusoid of frequency ω impinging on the array

[1] a uniform linear array (ULA) is an equi-spaced array configuration with antennas arranged along a line in a single dimension.

sensors is

$$y(n) = b_\omega(n) = e^{i\omega n} \sum_{k=1}^{N_a-1} w_k^* e^{-i\omega\tau_k(\theta)}$$

$$= e^{i\omega k} \mathbf{w}^H \mathbf{a}(\omega, \theta)$$

$$= e^{i\omega k} \mathrm{r}(\omega, \theta) \tag{3.3}$$

$\mathrm{r}(\omega, \theta)$ is a complex scalar function of time called, alternatively, the *beamformer response* or the *array factor* (Mailloux, 2005). The *array response*, $\mathbf{a}(\omega, \theta)$ is a vector of complex exponential terms that represent the phase and amplitude which the array presents to a sinusoidal input signal. The *beampattern* is the amplitude squared of $\mathrm{r}(\omega, \theta)$ (i.e. it is $|\mathrm{r}(\omega, \theta)|^2$). The white noise gain, G_{wn}, of the beamformer is the output power due to each input sensor receiving white noise with unit variance and is equal to the norm squared of the weight vector, $\mathbf{w}^H\mathbf{w}$.

The array response $\mathbf{a}(\omega, \theta)$ and the weighting vector $\mathbf{w}$ are vectors in N_a-space (where N_a is the number of elements in the array). Thus, the cosine of the angle between the two vectors determines the magnitude of the beamformer response, $\mathrm{r}(\omega, \theta)$.

3.2 Spatial Response Shaping

The distinguishing feature of a beamformer is its spatial response. The response is primarily defined by the physical qualities of number of elements, the spatial aperture and observation frequency. However, it is also able to be further modified by controlling the effective current distribution in the beamformer. By modification of these current distributions, a response that is optimised toward a particular set of characteristics may be achieved. It is this ability to modify the response both before observation or even adaptively during observation that makes beamforming —and particularly digital beamforming — such a powerful technique.

This leads to a classification of two tasks in beamformer design. The first is the development of the physical beamformer architecture with a certain number of elements and bandwidth. The second, that of optimising the response of a given architecture, is a mathematical optimisation problem. The latter is introduced in this section.

A convenient equivalence to the established techniques of signal frequency filtering is applicable to beamforming: a narrowband spatial filter may be seen as mathematically equivalent to the FIR filter. Techniques

for the generation of frequency response curves with certain desirable properties for FIR filters are applicable to spatial response shaping in a narrowband beamformer (a discrete *spatial* filter). Consequently, and as an example, frequency-response shaping with desired responses — such as those defined by Dolph using Chebychev polynomials (Dolph, 1946), Blackman-Harris (Harris, 1978), and others — may be used to similar effect in beam-shaping. The response shaping function is also known as an apodisation or 'windowing' function. Completing this analogy, a uniformly-weighted narrowband beamformed array will have the same spatial response as the temporal-frequency response of the running-average filter (see e.g. Oppenheim and Schafer, 2009). However, weighting functions may only *reduce* the total physical beamformer forward gain; this is the price paid for improved spatial discrimination.

3.2.1 Data-Independent Response Optimisation

Approaches to beamformer response design may be generally classified as being 'active', in that the response is reactively-optimised during observation in response to received data, or as data-independent, in that the response is designed prior to the observation. Data-independent beamforming attempts to produce a desired beamformer response based on *a priori* knowledge of the desired beam shape or of the signal environment, such as the location and frequency distribution of interference sources, or the known statistics of a target signal. The array shading vector in a data-independent beamformer approximates the desired response, $\mathbf{a}_{\mathrm{d}}(\theta, \omega, n)$, as closely as possible.

In contrast to the data-independent approach, the signal spatial correlation characteristics may be dynamically measured in order to reactively optimise the response to reject previously unknown noise and interference. This 'data-optimum' approach (also termed 'statistically-optimum beamforming') is a significant area of research in the beamforming literature and includes, for example, the techniques of multiple sidelobe cancellation (MSC), reference signal methods, SNR maximisation and linearly-constrained minimum variance (LCMV) beamforming (see also Veen and Buckley, 1988). Optimisation during observation is, however, problematic for astronomical beamforming for two reasons. The first is that the signal of interest is inherently unknown — a fundamentally different situation to that of communications-based beamforming — and thus may itself be cancelled by interference-elimination algorithms (see Cox, 1973; Cox et al., 1987, for discussion of the difficulties with adaptive algorithms when source statistics are unknown). The second is that the signal and interferers may be partially correlated. In this case, interference minimisation would eliminate the correlated portion of the desired source, altering its properties in an undesired manner (see

Widrow, 1982; Tie-Jun, 1985).

In both cases, this desired response is in general a function of time, as signal and interferer statistics change in either a known or measured manner. In the absence of complete astronomical source statistics (which is the case in SKA surveys) automatic adaptive optimisation may have the unwanted effect of optimising away the desired signal. It is proposed here that adaptive beamforming is not generally applicable for astronomical beamformers since it is inherent to the very act of observation that some aspect of the target signal is unknown.

a priori Beam Response Design

Generation of the desired beam response is achieved by adjusting coefficients such that the error between the desired response, $\mathbf{a}_d(\theta, \omega, n)$, and actual response, $\mathbf{a}(\theta, \omega, n)$, is minimised according to a sensible error condition. The desired response may have any shape and be optimised to be statistically matched to the observation.

The generalised optimisation problem may be approached in at least two ways. One method, discussed in detail in Chapter 5 and not further in this Chapter, is through the use of stochastic algorithms governed by heuristic selection criteria[2] for $\Delta\mathbf{a}$ minimisation. Another method is through the use of norm discriminants. In this case, a set of points, $(\theta_i, \omega_i), \quad 1 \leq i \leq P$ is chosen (a reasonable choice for point location is a uniform distribution) about which the response is to be optimised. If the number of optimisation points is greater than the number of antenna inputs, an overdetermined set of equations is obtained, which may be solved with standard methods solvable. A linear least squares method is used by setting the array error as a matrix minimisation problem over $\mathbf{w}$:

$$|(\mathbf{AD})^H\mathbf{w} - \mathbf{a}_d|^2 = 0 \tag{3.4}$$

where

$$\mathbf{A} = [\mathbf{a}(\theta_1, \omega_1), \mathbf{a}(\theta_2, \omega_2), \ldots, \mathbf{a}(\theta_1, \omega_P)]$$

$$\mathbf{a}_d = [\mathbf{a}_d(\theta_1, \omega_1), \mathbf{a}_d(\theta_2, \omega_2), \ldots, \mathbf{a}_d(\theta_1, \omega_P)]$$

and

$$\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_P]$$

[2] A stochastic or heuristic algorithm is one that will quickly arrive at good, but not provably exact solutions in intractably-large parameter search spaces

is a weighting function for the optimisation points such that certain $\mathbf{a}_i$ may be prioritised. If $\mathbf{B} = \mathbf{AD}$, and is invertible[3], the solution to 3.4 is then

$$\mathbf{w} = (\mathbf{B}\mathbf{B}^H)^{-1}\mathbf{B}\ \mathbf{a}_d \tag{3.5}$$

Weighting Functions and General-Norm Discriminants

Particular areas of the response may be targeted for optimisation. To achieve this, a weight vector $\mathbf{w}$ may be chosen chosen to increase the importance of these optimisation points in desired areas such as to minimise the difference between actual and desired spatial responses (the 'error') with an optimised difference-weighting metric.

Furthermore, squared-error is not not necessarily guaranteed to be the best error-weighting function. A general norm (L_p-norm) can be used as the error function instead of a weighted-square norm (L_2-norm). As an example of both of these techniques, Kumar and Murthy (1976, 1977) use unequal error weighting with an iterative norm-adaptive algorithm (Lawson's L_∞ algorithm) to minimise the response error in the minimax[4] error-function sense.

3.3 Digital Downconversion

Downconversion is the process of shifting the observation band so that it is aligned to the receiver baseband[5] in order to reduce the performance requirements of a digitiser and downstream digital processing hardware. It is a standard process in radiometric devices and is traditionally performed with an analogue multi-stage heterodyne mixer. However, there are advantages to performing down-conversion digitally. If the bandwidth of the digitiser is wide enough, the entire signal range, from baseband to the highest frequency extent of the observation band may be digitised and subsequently down-converted more predictably and precisely in the digital domain. This enables an inherently variable analogue component of the receiver chain to be replaced with an arbitrarily precise in-phase/quadrature (I/Q) downconverter. Furthermore, the observation band may be dynamically (run-time) selected if the desired band is to change dynamically during observation.

[3] the non-invertibility of $\mathbf{B}$ implies that the desired response could be unimplementable since it may, for instance, require excitation currents which are singular.

[4] An estimator of a deterministic parameter $\mathbf{a}_d \in \Theta$ from a set of underdetermined measurements $x \in \mathcal{X}$ is termed 'minimax' if the error is minimised in the worst possible case of $\Delta\mathbf{a}$.

[5] the down-converted signal may be slightly offset from baseband to reduce the interference due to down-converted $\frac{1}{f}$ noise and residual DC offset.

3.3.1 Nyquist Subsampling

Another approach to the downconversion process is the sub-sampling mixer. If a digitiser has the dynamic performance to measure signals n times higher in frequency than its maximum baseband-bandwidth, it may be used with some caution to down-convert and sample the n-th Nyquist band *directly*.

The note of caution mentioned in the previous paragraph concerns two properties of the sub-sampling technique. The first of these is noise aliasing. In contrast to single-stage heterodyne downconversion, Nyquist subsampling aliases noise from *each* of the higher Nyquist zones. Figure 3.1 is a schematic showing the operation of a sub-sampling mixer with an emphasis on noise aliasing properties. In particular, noise is aliased into the output spectrum from *each* of the Nyquist zones. In a real, noisy radiometer, the sub-sampling mixer has a signal to noise gain ratio of considerably less than unity.

The second note of caution concerns the dynamic performance of the digitiser. Almost counterintuitively, a digitiser is an inherently *analogue* device. The settling time of the digitiser output code, and particularly the settling time of each physical bit in the binary representation of the output code dictates the minimum sampling interval of the device *per effective bit*. This discussion is continued in Section 4.1.6 for the specific digitiser used in this design.

3.4 Digital Spectral Decomposition

Fundamental to both spectroscopic observation and the frequency-domain matrix-vector beamforming architecture is the efficient spectral decomposition of a received broadband signal. A computationally efficient algorithm for this task is the Fourier transform-based filter bank, for which well-known fast algorithms exist (Bellanger and Daguet, 1974). However, the spectral isolation of the naïve Fast Fourier Transform (FFT) filter bank is insufficient for some purposes, such as RFI containment, or spectral line observation. The frequency response of the FFT may be improved with the use of one of the set of pre-conditioning windowing functions, in order to optimise a particular aspect of the Fourier transform frequency response (see, for e.g. Crochiere and Rabiner, 1983).

3.4.1 Fourier Transform Polyphase Decomposition

Filter banks which contain frequency-modulated replicas of a basic filter may be computed efficiently by using the polyphase FFT FIR decomposition first introduced by Bellanger and Daguet (1974). The filter

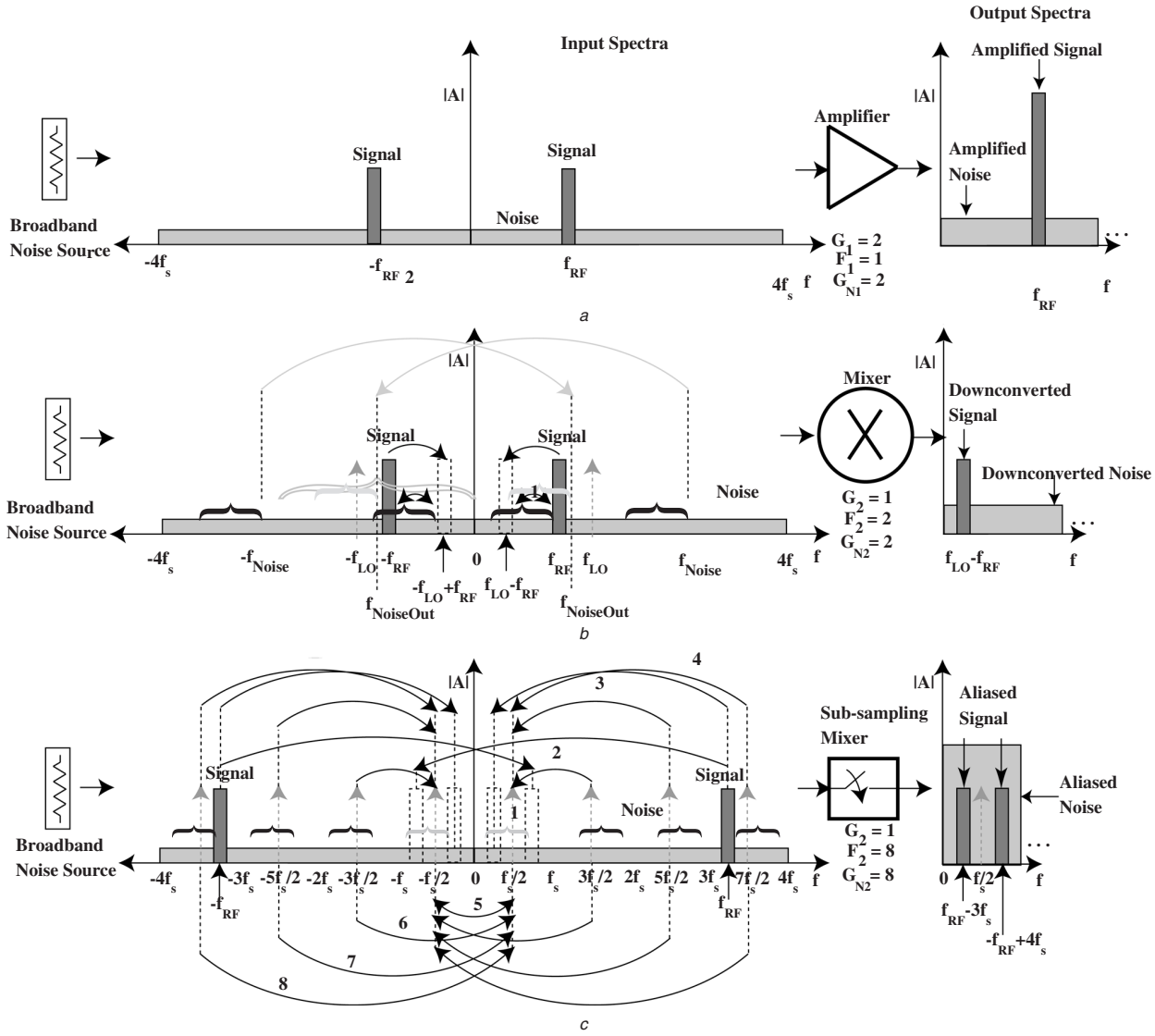


Figure 3.1: A schematic highlighting the concept of noise aliasing in a sub-sampling mixer. Figure 3.1a shows that noise is not aliased into the output spectrum by an ideal amplifier. The signal gain to noise gain ratio, $\frac{G_1}{G_{N1}}$, is unity. In the following two figures, arrows indicate direction of aliasing of 'mixed' signals. Figure 3.1b shows a single-stage mixer, where the down-converted noise is added to the baseband noise in the output spectrum, resulting in a signal gain to noise gain ratio $\frac{G_2}{G_{N2}} = 1/2$. Figure 3.1c shows the sub-sampler aliasing effect. In this case, down-converted noise from all Nyquist zones of the baseband signal is summed in the output spectrum. The signal gain to noise gain ratio in an n-th Nyquist-zone sub-sampling mixer is $\frac{G_3}{G_{N3}} = 1/n$ (i.e. inversely proportional to n). The sub-sampling mixer is sensitive to the instrumental noise floor as well as to the degree of sub-sampling. In order to clarify the contribution of the aliased noise, all other components are assumed to be ideal. From Pekau and Haslett (2006).

output for each channel is

$$Y_i = \frac{1}{N}\left[\sum_{k=0}^{N-1} H_i(W^k z)X(W^k z)\right], \tag{3.6}$$

where $W = e^{-j2\pi/N}$, H_0 is a prototype low-pass filter, and N is the length of the Discrete Fourier Transform (DFT). The analysis filter bank is composed of

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{N-1}(z) \end{bmatrix} = \begin{bmatrix} H_{p,0}(z^N) & z^{-1}H_{p,1}(z^N) & \dots & z^{-N+1}H_{p,N-1}(z^N) \\ H_{p,0}(z^N) & W^{-1}z^{-1}H_{p,1}(z^N) & & W^{-N+1}z^{-N+1}H_{p,N-1}(z^N) \\ \vdots & \vdots & & \vdots \\ H_{p,0}(z^N) & W^{-N+1}z^{-1}H_{p,1}(z^N) & \dots & W^{-N+1}z^{n+1}H_{p,N-1}(z^N) \end{bmatrix} \cdot \begin{bmatrix} X_0(z) \\ X_1(z) \\ \vdots \\ X_{N-1}(z) \end{bmatrix}$$

The filter coefficients may be factored from the analysis matrix to yield

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{N-1}(z) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & W & \dots & W^{N-1} \\ \vdots & \vdots & & \vdots \\ 1 & W^{N-1} & \dots & W^{N-1} \end{bmatrix} \cdot \begin{bmatrix} H_{p,0}(z^N)X_0(z) \\ z^{-1}H_{p,1}(z^N)X_1(z) \\ \vdots \\ z^{n+1}H_{p,N-1}(z^N)X_{N-1}(z) \end{bmatrix}$$

The weight matrix is then the Fourier matrix. Further, this matrix equation may be expressed in matrix-vector form as

$$\mathbf{y}(z) = \mathcal{F} \cdot \mathbf{h}(z^N) \cdot \mathbf{x}(z) \tag{3.7}$$

where the filter output is obtained by polyphase convolution (convolution of the input signal with modulated prototype filter set $H_i(z)$) followed by the Fourier transform matrix operation $\mathcal{F}$. The number of operations per set of output samples is then equivalent to the evaluation of N polyphase filters of length k, an $O(1)$ operation if $k \ll N$[6], as well as an FFT of size N, which is of order $O(N \log N)$.

3.4.2 Perfect Reconstruction

In such polyphase filter bank representations, is the exact original signal able to be re-composed from the sub-banded signals? Is the operation of spectral decomposition is reversible for a particular choice of filter coefficients?

[6] it is an $O(N)$ operation if $k \approx N$.

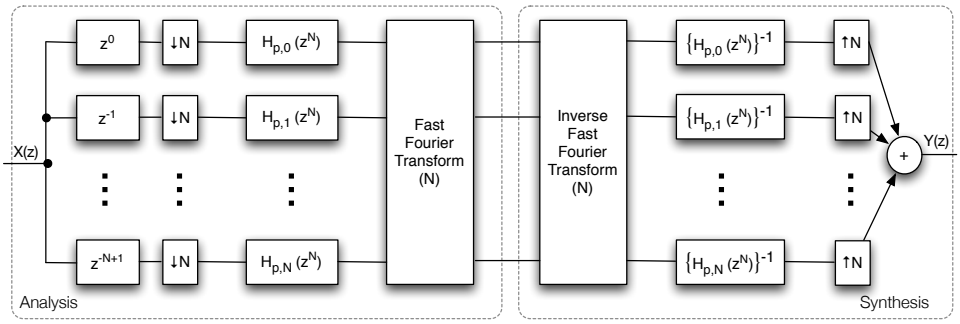


Figure 3.2: A modulated filterbank showing analysis and synthesis filters. The analysis filters are polyphase copies of the base filter H_p. Synthesis filters are the inverse of $H_{p,i}$, and will exist for all perfect reconstruction filter coefficients.

Perfect reconstruction is a desirable attribute for multi-purpose radio instruments, since different observations may require either further channelisation, such as in precision pulsar desispersion (Cordes, 1975), or precise reconstitution of the entire observation band for continuum studies.

Vetterli (1986) show that perfect reconstruction filters are possible for FIR-based filter banks. Particularly, if the synthesis (reconstruction) filters are composed from the first column elements from the cofactor matrix[7] $\mathbf{C}(z)$ of the analysis filter matrix $\mathbf{M}(z)$ and chosen such that the determinant of the analysis filter matrix, $\det(\mathbf{M}(z))$, is a pure delay, then the aliasing of filter elements is perfectly cancelled, and the original signal may be perfectly reconstructed.

The reconstructed filterbank output is a linear combination of $X(z)$ and its N-1 aliased components $X(W^k z)$

$$X_{\mathrm{r}} = \sum_{k=0}^{N-1} F_k(z) X(W^k z), \quad \text{where } F_i(z) = \frac{1}{N} \sum_{l=0}^{N-1} H_l(W^i z) G_l(z)$$

Attempting to get $F_0(z)$ to be equal to an ideal delay and $F_i(z) = 0,\ (i \neq 0)$ yields

$$\mathbf{F}(z) = \mathbf{M}(z)\mathbf{g}(z) = \mathbf{u}(z),$$

$$\mathbf{f}_i(z) = F_i(z), \quad \mathrm{M}_{ij}(z) = (1/N) H_j(W^i z), \quad \mathbf{g}_i(z) = G_i(z), \quad \mathbf{u}(z) = [z^{-k}\ 0 \ldots 0].$$

Perfect reconstruction may be achieved by choosing the sub-band filters $H_n(z)$ from a class of filters that

[7]The cofactor matrix of $\mathbf{M}(z)$ is equal to the inverse of $\mathbf{M}$ when it is multiplied by the determinant, i.e. $\mathbf{M}^{-1}(z) = \det(\mathbf{M}(z))\mathbf{C}(z)$.

will automatically satisfy the constraint of the determinant being a pure delay. One such class, proposed in Wackersreuther (1985), consists in choosing $H_0(z)$ such that its autocorrelation function has zero even-index samples. As an example, in the case of a two-filter bank with $H_0(z)$ of length M (that is, $H_n(z) = [H_0(z) \;\; H_1(z)]$), setting $H_1(z) = z^{-M+1}H_0(-z^{-1})$ will satisfy the determinant constraint.

3.5 Beamforming

3.5.1 Spatial Sub-sampling: Sparse and Dense Apertures

Aliasing in the spatial domain corresponds to an ambiguity in source origin. That is, there is no guaranteed ability to discriminate between signals resulting from sources at two different locations if the aperture is not critically sampled. Even within a narrow frequency band ω_0, a source signal originating at (θ_0, ω_0) may be indistinguishable from a separate signal $(\theta_1, \omega_0) = (k\theta_0, \omega_0)$ if the aperture is not fully-sampled and both will be included in a the beam response to (θ_0, ω_0). Wideband responses have added ambiguity due to the scalar projection of frequency-angle tuples. These 'ambiguity beams' are referred to as the *grating lobes* of the array.

Both grating lobes and side-lobes introduce an unwanted power response into the beam shape. Ideally, there would be no signal response in either of these regions of spatial ambiguity. For the SKA, large side- and grating lobe responses need to be combated if a large aperture array is to be able to observe signals that are much ($\lesssim -60$ dB) weaker than the signals which may appear in grating- or side- lobes.

There is an important qualification to be made about grating lobes regarding the eventual use of the beamformer. In a *total-power* beamformed radio instrument, sensitivity and dynamic range performance is dependent on the placement and spacing of array elements. It is desirable to space elements such that all grating lobes due to particular antenna pairs sum incoherently, minimising the effect of the ambiguity beam due to each antenna pair on the total summed beam.

However, the use of a beamformer as an input to an interferometric instrument provides additional discrimination through the cross-correlation function. A source appearing in the beamformer grating lobes and side-lobes is modulated across the accumulated interferometer output at a higher frequency and is thus able to be filtered. This means that discrimination potential is theoretically greater than in total-power instruments, and the filtering function depends now also on baseline separation as well as the quality of imaging techniques. Quantifying the precise interferometric discrimination potential is beyond the scope of

this work; much greater detail and further references are in Thompson et al. (2001).

3.5.2 Narrowband Beamforming Techniques

If a frequency band is considered to be negligible in width compared to its central frequency f_c, a single coefficient may be used to phase-delay the narrow band with range $f_c - \Delta$ to $f_c + \Delta$. However, a signal undergoes spectral smearing if a single phase term is applied to signals in a finite range of frequency space; a narrowband signal is defined as one with bandwidth of small enough extent that this error is negligible.

Phase-Shift Beamforming

In a phase-shift beamformer, the sub-banded beam vector, $\mathbf{b}(k)$, $(0 \leq i \leq N_b)$ is a linear combination of the signal vector $\mathbf{s}$, $(0 \leq j \leq N_a)$ and an N_b x N_a matrix $\mathbf{M}$ which is a combination of the beamforming, equalisation and calibration coefficients for each desired beam.

Fast Transform Beamforming

Calculation of the full set of spatial frequencies is equivalent to a Discrete Fourier Transform (DFT) of the spatial samples. Fast Transform beamforming exploits the redundancy in this calculation with use of the Fast Fourier Transform (FFT) in an almost identical fashion to Fourier transform spectral decomposition[8]. This technique requires a dense, uniform grid of sample points.

Fast transform arrays have been used for radio-astronomy. Otobe et al. (1994); Daishido et al. (2000) describe the *Waseda Digital Lens*: a 2^6-element, 2D-FFT Nyquist-rate digital imager operated from 1989-1995 at Waseda University, in which the resultant spatial beams are directly integrated. More recently, Tegmark and Zaldarriaga (2009) suggest a similar design with more elements and wider bandwidth in a purpose-built instrument for 21-cm tomography for Epoch of Reionisation (EoR) signature signal detection. Analogue full-beam beamformers such as the Rotman Lens and the Butler Matrix (e.g. Elliott, 2003) perform the equivalent operation. Such techniques are well-suited to approximately fully-sampled apertures (where the spatial occupation matrix is dense) since inputs must be fit to a uniform input grid.

The pointing directions of the set of independent beams produced by fast transform beamformers is arbitrary from the sky reference frame and in general will not coincide with desired beam pointing direction. In this case, beam interpolation may be used to generate secondary beams pointed at precisely desired

[8]Note that the full set of spatial frequencies are formed over a narrow temporal band.

locations. This second level of beamforming requires further computations, compromising the efficiency of the fast transform approach if a full set of secondary, *beam-space* beams is generated.

3.5.3 Wideband Space-Frequency Beamforming Techniques

Khlebnikov et al. (2010) present a novel approach to the interleaved space-frequency beamforming process. Essentially, the technique involves interleaving temporal and spatial filtering in a computationally-efficient manner. Initially, coarse beams are formed by frequency independent, high-precision digital time-alignment of antenna signals before channelisation and subsequent precise phase alignment. In effect, this is a composite time-frequency-interleaved beamforming process. It has been shown to reduce the complexity significantly, and shows a large total computational improvement over traditional narrowband phase-shift beamforming.

3.6 Summary

In this chapter, the role of the beamformer as a digital spatial and temporal filter has been highlighted. It has the task of extracting each component of the received signal through signal processing techniques. The two chapters that follow each describe an implementation of a digital beamformer architecture, based on the algorithmic framework of this chapter. Chapter 5 uses a single-processor beamformer to perform beam-shape optimisation, while Chapter 4 describes the design of a multi-processor beamforming architecture.

Chapter 4

A Wideband, Dual-Polarisation Digital Beamformer for 2-PAD

This chapter discusses the design of a wideband, multi-beam, dual-polarisation digital beamformer for the 2-Polarisation All-Digital (2-PAD) instrument. 2-PAD is an aperture array demonstrator that will confirm the feasibility of beamforming tightly-coupled, wide-bandwidth signals entirely in the digital domain. Its goals are to demonstrate design of novel antennas, low-noise RF instrumentation for signal amplification and transmission, and signal digitisation, dataflow and processing for aperture array radio instruments[1]. Due to 2PAD's modular design, different implementations of subsystems (i.e. antennas, digitisers or processing systems) are able to be tested within a working system.

In this chapter, the digital design and characterisation of the signal processing architecture is detailed. Some important elements of the analogue signal chain of the instrument, built in a collaboration of the Universities of Oxford, Cambridge, Manchester is also described. Deployment and testing of the complete instrument at Jodrell Bank Radio Observatory is detailed.

[1] more information is available at www.skads-eu.org

4.1 The 2-PAD Digital Beamformer

The essential goal of the 2-PAD project is to design and test a *digital* phased array over a critically-sampled aperture[2], with the telescope aperture composed of many antenna elements over a large ($\Delta v \simeq v_c$) bandwidth. Digital signal processing systems have the advantage of environmental stability over their analogue counterparts.

Even though the processing requirement is large, the goal for fully-digital operation is well-founded: there has been an exponential increase in essentially every type of digital processing capacity, with capacity doubling times of memory, processor, and communication subsystems of between 10^1 and $\sim 10^{1.6}$ months (Meuer, 2008), which strongly suggests that that the cost of a particular digital system will fall exponentially with time.

The digital system designed and built for 2-PAD is the subject of this chapter. The physical implementation of a digital system is decomposed into several signal processing subsystems. Each stage is responsible for one of the operations of sampling, downconversion, filtering, channelisation, beamsteering and, finally, beamforming. In this section, the design of each these signal processing systems is described.

4.1.1 Digital Frequency-Domain Beamforming

Some types of beamforming architectures were introduced in the previous chapter. The optimal algorithmic architecture of these, in terms of efficient computation[3], depends on the number of 'spatial frequency channels' or 'beams' required. If all possible independent beams are required, the redundancy in steering coefficient multiplication may be exploited by the use of fast transform algorithms. Alternatively, if only a sparse set of beams is required, computing each beam directly (i.e. without making use of computational redundancies) is more efficient.

There exists a number of beams at which the fast all-beam calculation becomes more efficient; $\log_2(N_B) \geq N_b/C$, where N_B is the total number of independent sky beams (and antennas), N_b is the number of *required* beams, and $15 \gtrsim C \geq 34/9$ (Lundy and Buskirk, 2007; Johnson and Frigo, 2007) is a constant related to the fast transform total arithmetic complexity. 2-PAD requires multiple, but few beams. Thus, the phase-shift

[2]critical sampling of the antenna aperture ensures a unique solution to signal DoA. Element spacing of greater than $\lambda/2$ corresponds spatial aliasing: an ambiguity in source location. An array with 'critical' ($\lambda/2$) or smaller antenna spacing is also known as a 'dense array' in the SKA memoranda series (e.g. Schilizzi et al., 2007; Dewdney et al., 2010).

[3]Again it is noted that evaluation of computational efficiency is incomplete to the point of inaccuracy in the absence of actual hardware and communication infrastructure implementation. The direct computational efficiency comparisons presented here are derived as simply as possible; an ideal monolithic processing architecture is assumed.

beamforming method was chosen as it is the most computationally efficient.

4.1.2 Equalisation

Equalisation compensates for systematic bandpass non-linearity caused by both digital and analogue components. Digital artefacts are non-linearities caused by quantisation in digitisation and digital processing. Analogue non-linearities are those introduced by the various RF components in the radiometer signal path. Analogue equalisation must be performed at least as fast as the rate at which the system changes. The analogue system may change significantly on a sub-minute timescale, $\Delta t_a^{-1} \simeq 10$ s (see, for example Price et al., 2009).

An independent quantisation noise model assumes that the digital system is time invariant ($\Delta t_d^{-1} = \infty$). However, it has been shown in Chapter 2 that quantisation noise is in general deterministically related to the input signal, and its exact value depends on the (variable) properties of the input signal. Furthermore, determination of exact relationships between the signal and quantisation noise becomes intractable in the case of many, repeated quantisations. Thus, digital modeling is required, and must be done for each expected input signal class if quantisation effects are to be removed correctly. The magnitude of quantisation artefacts are reduced by increasing processor bit-width, but this must be traded against the corresponding increase in power consumption and therefore monetary cost.

Another form of equalisation is also used after re-quantisation to shift the signal into a linear power gain region of the quantisation curve of signal correlation (see Thompson et al., 2001).

In this work, bandpass characteristics are corrected by first obtaining the equivalent filter transfer function of the complete system for a particular input signal, $H_S(z)$, then multiplying the bandpass with its own gain-normalised inverse $H_S^{-1}(z)$. The equalisation coefficients are applied during the beamsteering operation.

4.1.3 Synchronous Digital Design

Globally Asynchronous, Locally Synchronous (GALS) design allows the data transfer characteristics of a digital system to be statistically modelled; that is, data transfer need not be deterministic, relying on the *average* transfer and delay rates between processing nodes in order to model data flow. Globally synchronous design, on the other hand, requires that all data is deterministically assured to be at a certain point in the system at a certain time.

The design pattern[4] of *global asynchronicity* in digital correlator design has been used for digital radio

[4]Johnson et al. (1995); Gamma et al. (1993) introduce the concept of re-usable system design; the *design pattern*. DeHon

correlator design (e.g. Jonas, 2009; Parsons et al., 2010). This design pattern describes an architecture of loosely-coupled processing blocks, each of which manages its own data buffering in a streaming signal processing system. The distinction between global asynchronous and global synchronous design templates is essentially that of distribution of synchronisation complexity; the larger a digital system becomes, the lesser the overhead of local synchronisation dominates total control complexity.

The digital design in 2PAD is optimised for a deterministic data stream with static data transposition, as opposed to a statistical data stream; thus, globally synchronous data transfer is used in 2-PAD in preference to switch-based asynchronous data transfer of Jonas (2009) and Parsons et al. (2010).

4.1.4 System Dataflow Requirements for 2PAD

The analogue input signal range of 2PAD is from 0.5GHz to 0.7GHz (a bandwidth of 0.2GHz). It is digitised with $q = 8$ bits of precision, which results in an input data rate from each antenna polarisation of $q \cdot 0.4 = 3.2\text{Gbs}^{-1}$. This means that the first-stage beamformer processor will have a raw input data-rate of $N_a \cdot q \cdot 0.4 = 51.2$ Gbs^{-1} and an output data rate of $N_b \cdot q \cdot s \cdot 0.4$ $\text{Gbs}^{-1} = 6.4\text{Gbs}^{-1}\text{beam}^{-1}$, where s is the bitwidth scaling factor of the beamformer, $q \cdot s$ is the output bitwidth, and N_b is the number of digital beams.

4.1.5 Architecture

A narrowband phase-shift beamforming architecture is chosen. The sub-banded beam vector, $\mathbf{b}(k)$, $(0 \leq i \leq N_b)$ is a linear combination of the signal vector $\mathbf{s}$, $(0 \leq j \leq N_a)$and an $N_b \times N_a$ matrix $\mathbf{M}$ which is a combination of the beamforming, equalisation and calibration coefficients for each desired beam[5]

The reason for this decision is that the simplicity of the basic structure of the matrix-vector approach allows a simple addition of beam processors with available processing resources. Furthermore, a simpler structure provides a more accessible platform for application of the calibration coefficients. The alternative, fast-transform beamforming, requires either an extra equalisation coefficient multiplication step, or a modified FFT algorithm which is able to apply arbitrary and changeable equalisation coefficients. Figure 4.1 shows a schematic representation of the 2-PAD matrix-vector beamforming architecture. The figure shows how the

et al. (2004) expand on the use of design patterns in reconfigurable hardware design, particularly that of synchronous system design.

[5]There is a strong distinction between real-time beamforming processing, which must be performed at the input data-rate and the *calculation* of equalisation and beamshaping co-efficients, which may be performed at a much reduced rate.

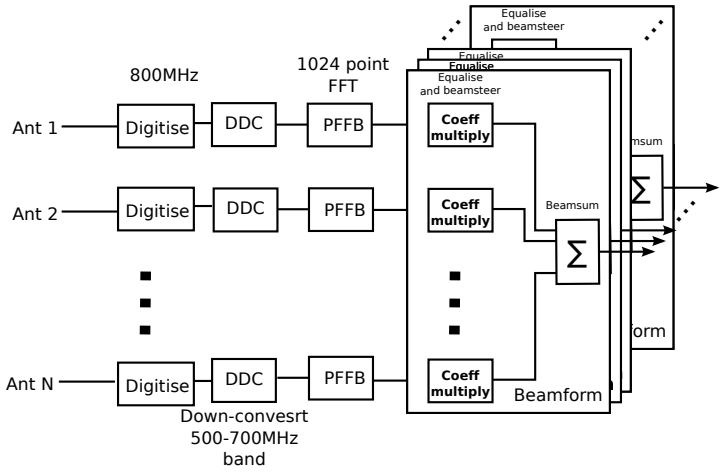


Figure 4.1: A schematic of the matrix-vector beamforming architecture. Each antenna signal is digitised, and the 0.5 - 0.7 GHz band is down-converted to baseband. Channelisation is performed with a polyphase FFT Filter bank in order to divide the broadband signal into narrowband channels. Each channel is multiplied with a phase coefficient which both shifts the output in phase and to equalises the channel to the correct input level. The aligned antenna channels are then summed to form each beam. The beamforming process is repeated for each additional (independent) beam which is required, as hardware resources (multipliers, adders) allow.

basic single-beam architecture is extendible to multiple-beam operation by repetition of the beamsteering coefficient multiplication step.

4.1.6 Hardware Platform

The digital beamformer for 2-PAD uses digital hardware from the Collaboration for Astronomy Signal Processing for Radio Astronomy (CASPER) group. The design is currently implemented on eight Interface Break-Out Board (IBOB)s and a single Berkeley Emulation Engine II (BEE2), which are both Field Programmable Gate Array (FPGA) processors. Figure 4.2 shows these processing boards and Figure 4.7 shows the processing boards racked in the 2-PAD bunker.

BEE2

The BEE2 board is a general-purpose reconfigurable processor. This multi-FPGA board forms the core of the Berkeley Emulation Engine II high-end reconfigurable computing (HERC) system, which is a reconfigurable supercomputer designed for System-on-Chip (SoC) and Application-Specific Integrated Circuit (ASIC) simulation. Figure 4.3 is a schematic of the board showing the layout and connection architecture.

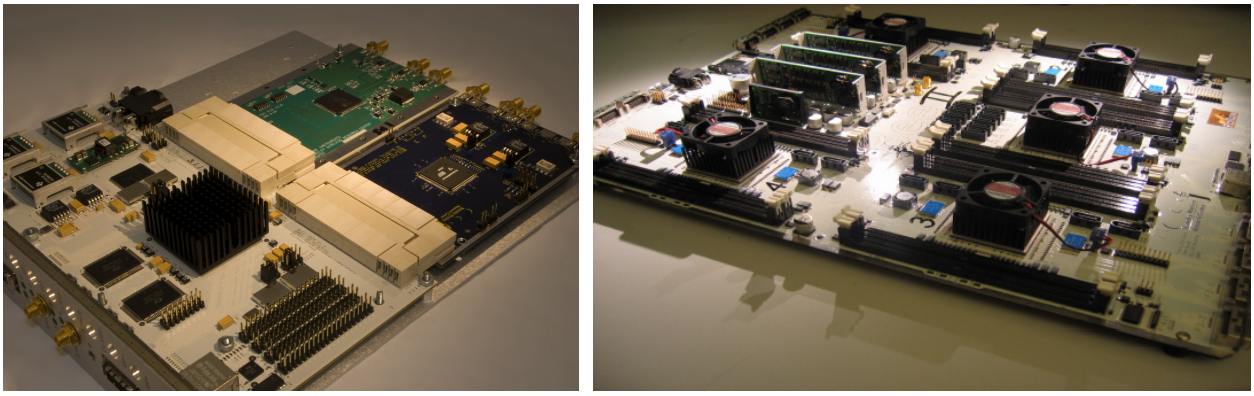


Figure 4.2: Field programmable gate array (FPGA) processing hardware. Figure 4.2(a) shows the IBOB board with a single Xilinx XC2VP50 Virtex-II Pro FPGA. Two different production versions of the iADC digitisation boards are connected to the IBOB. Figure 4.2(b) shows the BEE2 board containing 5 Xilinx XC2VP70 Virtex-II Pro FPGAs. The control FPGA is on the centre right of the image.

The processing power of the board is provided by five Xilinx XC2VP70 Virtex-II Pro FPGAs[6] laid out in a master-star topology; four 'user' FPGAs are connected in a ring architecture, and each of these is connected directly to the 'master' FPGA.

The master FPGA runs the BORPH (Berkeley OS for ReProgrammable Hardware) operating system (So and Brodersen, 2006). This is a modified version of the linux operating system which provides a software abstraction for reconfigurable computers. Particularly, FPGA gateware is exposed as a process abstraction, while FPGA resources, such as processor-visible registers and memory, are file abstractions. Programmable hardware abstraction through an operating system such as BORPH is a powerful way to interact with hardware, and is used in the 2-PAD digital beamformer for programmatic management of steering coefficients, as well as control of signal scaling and equalisation depending on input signal power.

The BEE2 has a total compute capability of $\sim 0.5 \times 10^{12}$ Operations.s^{-1}, provided by 328 multipliers (18×18 bit) and 74,448 logic cells, as well as 2 PowerPC CPUs on each of the five FPGAs.

Intra-board communication in the user-FPGA ring is through 138 LVCMOS direct traces operated at 300MHz, for a maximum communication bandwidth of 41.8Gbs^{-1}. User FPGAs connect to the control FPGA with 64 LVCMOS direct traces with a maximum bandwidth of 19.2Gbs^{-1}.

External communication capacity totals 108Gbs^{-1}. User FPGAs have four independently-controlled, dual-way 10Gbs^{-1} transceivers, which are accessed through a copper CX4 cable assembly. The control

[6] see www.xilinx.com

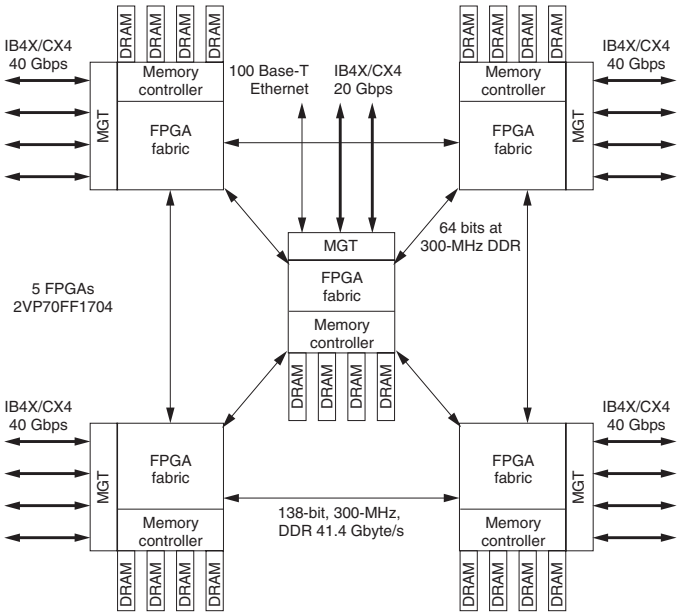


Figure 4.3: Schematic showing the BEE2 FPGA processing board. FPGAs are inter-connected in a master-star topology. User FPGAs on the outside of the ring have mono-directional connection to their two nearest neighbours totalling 41.8Gbs^{-1}, as well as an additional 19.2Gbs^{-1} connection capacity to the master FPGA. Each user FPGA has 4 10Gbs^{-1} external connections that are able to interface to XAUI, Xilinx Aurora©, 10GbE, and other high-speed-serial (HSS) protocols. FPGAs have access to 4GB of DDR2 SDRAM. The master FPGA runs a version of linux, BORPH (see So and Brodersen, 2006, for more information about the BORPH operating system), that provides, among other useful attributes, process abstractions of FPGA gateware and file access to FPGA resources.

FPGA has two such transceiver/CX4 pairs, as well as connections to common interfaces of 10/100 ethernet, USB 1.1, universal serial (through the RS232 protocol), DVI, and to General Purpose I/O (GPIO). The low-speed interfaces have been used for control and software interaction, while high-speed interfaces are used for signal transport.

The BEE2 has a maximum capacity of 20GB of DDR2 DRAM memory (4GB per FPGA) and high memory bandwidth provided by the four independent DDR2 access channels. The BEE2's large-storage capacity is not used. Instead, all used storage is local to the FPGA fabric.

The combination of high communication and computation capability, as well as the high ratio of the former to the latter make the BEE2 particularly suited to high-performance signal processing of broadband radio signals. The BEE2 forms the central computational engine of the beamformer.

IBOB

The first stage of the beamformer hierarchy is implemented on an IBOB board. The IBOB's central processing resource is of the same family as the BEE2, with slightly lesser specifications than the BEE2 FPGAs; a Xilinx XC2VP50 FPGA, which has 232 DSP 'slices' — each able to perform a clock-rate 18x18-bit multiplication — and 53,448 logic cells provides processing power. An approximate maximum compute capacity is 60×10^9 Operations.s^{-1} Similarly to the BEE2 FPGA, the XC2VP50 contains a PowerPC© core which provides telnet access to the fabric resources including registers and block random access memories (BRAMs).

The IBOB connects to digitisation hardware through a 40-pin ZDOK connector. It connects to further 'downstream' processing devices (in this case, the BEE2) through two 10Gbs^{-1} CX4 interfaces connected to RocketIO© FPGA interfaces.

Digitisation

Signals from array elements are digitised with the Atmel BiCMOS AT84AD001B dual 8-bit Analogue-to-Digital Converter (ADC), each of which is able to digitise two streams at 1 GSa.s^{-1} on each of two iADC boards connected to each IBOB. Each channel of the ADC may receive a 500 mVpp differential RF input. Two iADC boards are connected to each of the IBOB processing boards. Table 4.1 lists the characteristic parameters of the device.

The ADCs are run at 800MHz, which results in a sampling bandwidth of 400 MHz. 200 MHz of this band is digitally down-converted to baseband. In other words, the signal is super-Nyquist sampled as a single, real voltage before the required 0.5 - 0.7GHz band is digitally mixed to baseband.

Table 4.1: Characteristic Performance of the Atmel BiCMOS AT84AD001B Dual 8-bit Digitiser from Atmel (2006).

Power Consumption	0.7W.channel^{-1}
Full Power Input Bandwidth	1.5 GHz (-3 dB)
Worst-case Signal-to-Noise Ratio	42 dB (6.8 ENOB)
Total Harmonic Distortion	-51 dBc
Spurious Free Dynamic Range	-54 dBc (Fs = 1GSa.s^{-1})
Differential non-linearity	0.25 LSB
Integral non-linearity	0.5 LSB
Bit Error Rate	10^{-13} at 1GSa.s^{-1}
Channel Crosstalk	< -40 dBc ($F_s = 0.5$GSa.s^{-1}, $F_{\text{in}} < 0.7$GSa.s^{-1})

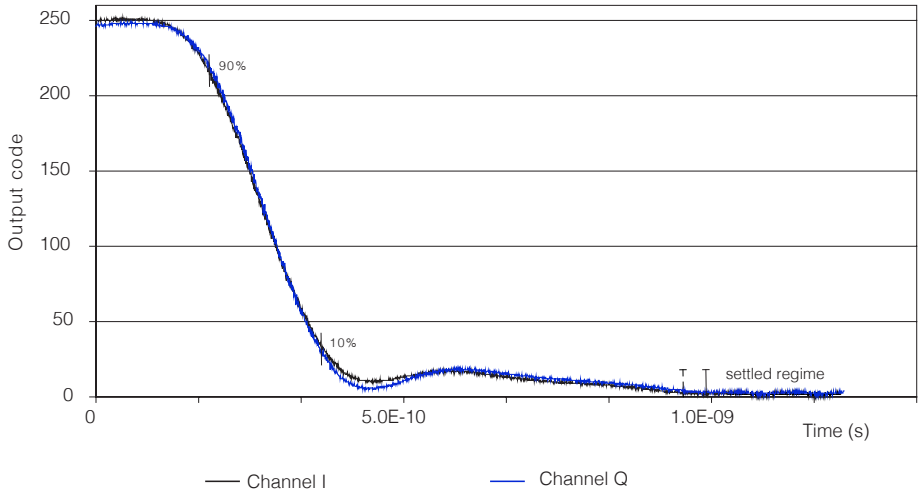


Figure 4.4: A figure highlighting the undersampling capability of the Atmel AT84AD001B digitiser by showing the internal settling time. The figure shows the output code response when a 0.5GHz, -1dBFS-power, sinusoidal input signal is stepped in value (changed from maximum to minimum signal power) at $t = 0$s. The ordinate is the integer value of the output code. The response is plotted for both channels, which are labelled IA and QA respectively. The full-specification output code settling time is 997 ns (i.e. the time required for the output code to settle from its maximum to its minimum value). This dictates that the maximum signal frequency of 1.003GHz (i.e. this is the maximum frequency possible if the device is to operate at the specified effective number-of-bits (ENOB)). Error-free operation is possible in only those Nyquist zones which are less than this frequency (i.e. 1st and 2nd Nyquist zones). The ENOB of the digitiser will be degraded if operated with an input signal in the higher Nyquist zones. As an example, the AT84AD001B has a settling time of 170 ns for the output code value to change from 10% to 90% of its value. Disregarding differential and integral nonlinearities, this enables this device to be used with > 3 effective bits in under-sampled applications with a signal frequency up to 5.88GHz. The measured internal settling time of the digitiser was extracted from data in Atmel (2006).

Harmonic Sampling

Sampling signals above the first Nyquist zone is often desirable. The advantage is that high frequency signals may be sampled directly and then digital techniques used to process the signal, thereby obviating the requirement for analogue down-conversion. There are issues of noise aliasing from the other Nyquist zones (discussed in Section 3.3.1) and dynamic performance at super-Nyquist frequencies. Clearly, as signal frequencies become higher, the *dynamic* performance of the ADC becomes critical. Distortion performance is important at the *signal* frequency, not just at the bandwidth frequency or baseband. A digitiser suitable for under-sampling applications must possess sufficient dynamic performance in the required higher-order Nyquist zone.

Of particular importance for super-Nyquist operation is the internal settling time of the ADC, manifest

by the measured *internal* step response. Figure 4.4 shows the output code as a function of time after an input step is applied at $t = 0$ s. At higher frequencies, fewer bits of the ADC output code are able to be used reliably. Operation in higher Nyquist zones decreases the effective number of bits of a particular digitiser. The full-swing (FS) output code is the value of the output code of reliable bits. The figure shows a measure of the maximum operating frequency of the ADC.

Polyphase Filter Bank

In Chapter 3, the concept of frequency decomposition was introduced. Narrowband beamforming requires signals to be decomposed into frequency components which have range small enough that application of a single coefficient does not introduce large smearing at the band edges. The FFT is a computationally efficient algorithm for spectral decomposition. However, frequency isolation of the FFT may not be sufficient for some applications; of particular concern is that spectral isolation in adjacent channels is approximately -13 dB. Spectral isolation is improved by the use of a poly-phase windowing function. The implementation of the FFT and FIR filter of the poly-phase channelisation architecture is described below.

Parsons et al. (2008) report substantial hardware savings of the CASPER library FFT over commercial implementations, due to the use of the radix-2 biplex pipelined FFT architecture (see, for example Emerson, 1976; Crochiere and Rabiner, 1983). Real FFTs may be performed with a single complex FFT core by making use of the Hermitian conjugation operation. Two real signal streams are encapsulated within a pseudo-complex data stream. Post-transforming, the first spectrum is extracted as the real part of the complex spectrum and the second spectrum is the bin-reversed conjugate of the imaginary part of the complex spectrum. Thus, four real data streams may be fast-transformed inside a single such core. The biplex, real-FFT-encapsulation architecture is implemented by Parsons et al. (2008) and available for implementation as a CASPER library. This biplex radix-2 FFT forms the core of the polyphase FFT filter bank.

Poly-phase decomposition of the Hann-windowed apodisation function (see for example Blackman and Tukey, 1958) improves the frequency isolation of each channel. Figure 4.5 shows the frequency response of both the FFT and the Hann-windowed poly-phase FFT. The filter coefficients have 18-bit quantisation, which is the native bit-width of FPGA multipliers. Frequency isolation is improved to ~ 30 dB with an 8-tap poly-phase filter bank compared to the FFT.

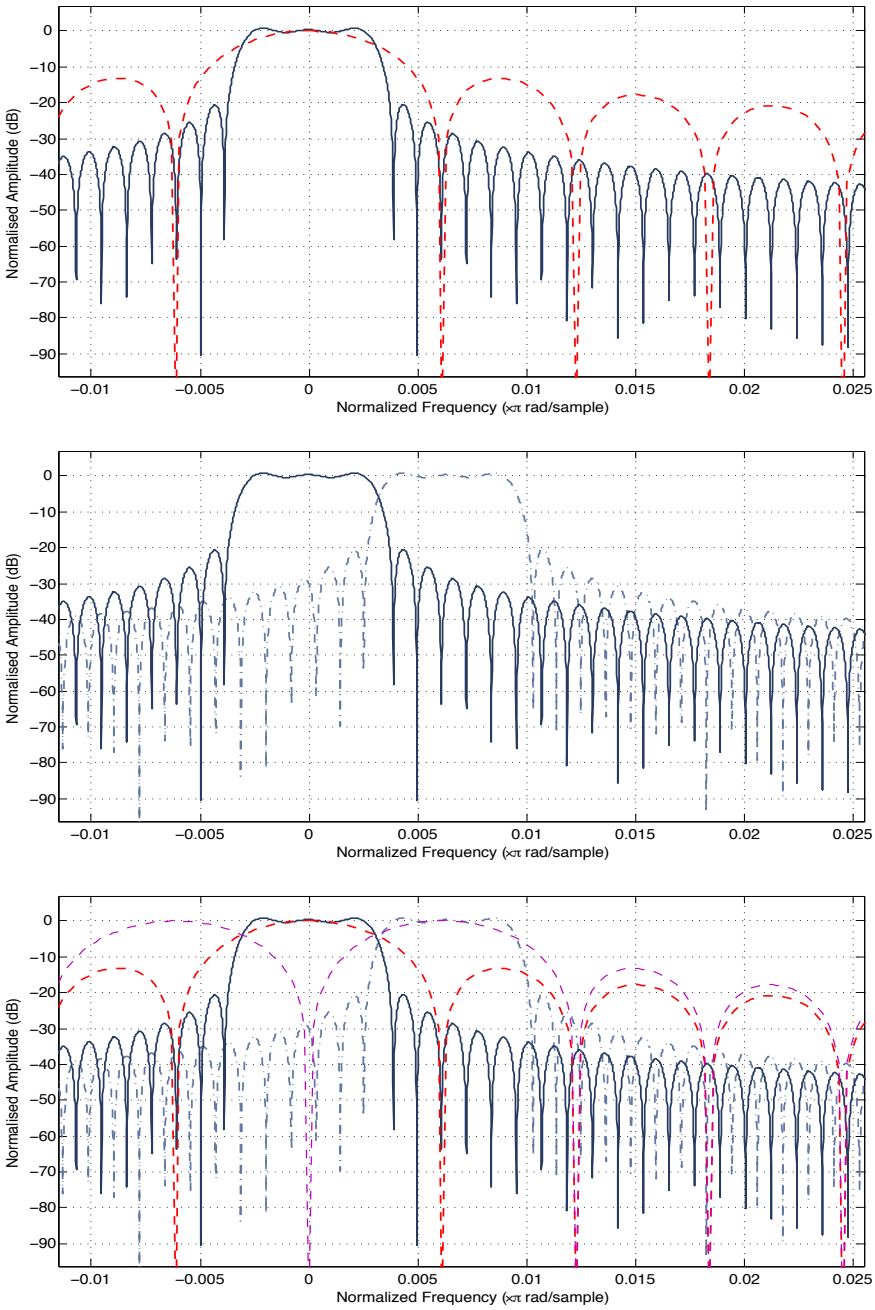


Figure 4.5: Response of the 2^3-tap Hann-windowed finite impulse response (FIR) filter bank with 18-bit quantised coefficients (solid, dark blue line) compared with the frequency response of the nominal discrete Fourier transform (DFT) (dashed red line) shown in 4.5(a) demonstrating the improvement in spectral isolation through the use of the pre-conditioning filter bank. 4.5(b) shows adjacent channels of the filter bank, with the second bin response the dashed light blue line. Spectral suppression in adjacent channels is greatly improved from -13.7dB for the FFT and better than -30dB at the equivalent point in the filter bank. 4.5(c) is a combined plot incorporating also the second channel of the FFT.

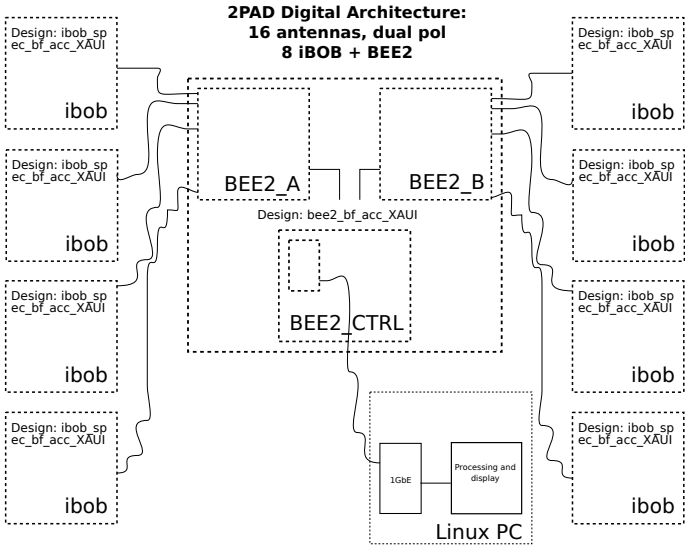


Figure 4.6: Schematic representation of the 4x4x2 digital beamformer showing communication and synchronisation between the iBOB boards and the BEE-2 boards.

Beamformer

The approach taken to beamforming is that of an hierarchical summation network, in order to reduce inter-board communication, as shown in Figure 4.6. Beamsteering is performed for each antenna element on the IBOB board. Beam summation is performed after this coefficient multiplication, as well as after each data transfer and agglomeration; once each per polarisation on the first and second BEE2 FPGA, and once again on the third FPGA.

The FPGA-based design uses embedded multipliers in a Xilinx Virtex II Pro device. Since these fixed hardware blocks are natively 18-bit, 18-bit multiplications and datapaths are used throughout the design.

XAUI Synchronisation

The Ten gigabit Attachment Unit Interface (XAUI) protocol operating over Xilinx RocketIO© transceivers is used to stream data between digital processing hardware. The lower-order XAUI protocol is used in preference to 10Gb Ethernet since it has a simpler structure (particularly, XAUI has no guaranteed data delivery and has no inherent packet re-transmission and is thus has a theoretically deterministic data transmission rate) and requires less bandwidth overhead to implement.

Parsons et al. (2008) estimate XAUI bit-error rates for the IBOB as $2.2{\cdot}10^{-16}$ Hz, based on the observation of no transmission errors in a test bit-error-counting system over $\sim 10^5$ s. These results (i.e. no transmission

errors) have been confirmed in a test system over a similar time period. However, much higher bit-error rates were observed in the operation 2PAD system: a frame error every $6.55 \cdot 10^{-6}$ s over 7.5 hours, which gives instead a bit-error-rate of $4.20 \cdot 10^{-4}$ Hz.

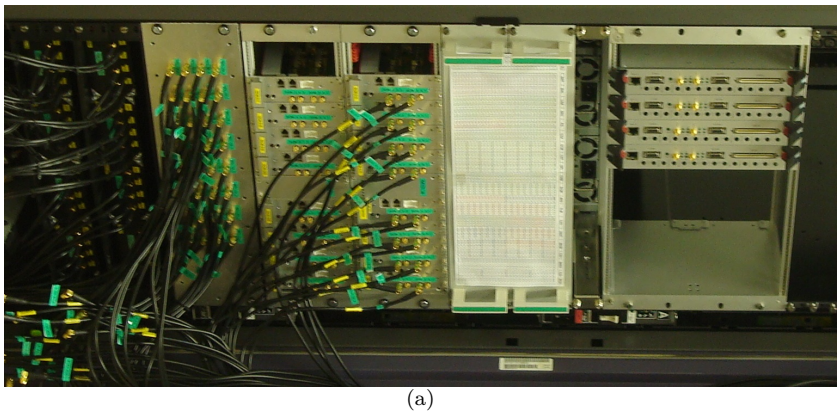
The likely cause of this inconsistency is the difference in FPGA temperature when executing a 'full' design (i.e. a design with high resource utilisation, close to 100%) compared to executing a test system, where resource utilisation is low. The full design execution causes higher silicon temperature. During full-system tests, the room temperature was $\sim 27^{o}$C and FPGA chip packaging temperature was measured to be 68^{o}C, which suggests that the RocketIO© circuit junction temperature would have been significantly higher. Overly high temperature may have been the cause of faulty transmission and forced re-transmission, even though all FPGAs were individually air-cooled and the entire system separately air-cooled by a rack-mounted fan module, as visible in Figure 4.7(a).

Under such conditions, the XAUI data transmission medium was found to have far higher bit-error rates than either predicted by Xilinx (2007), or measured in test systems. Several attempts at synchronisation of the separate streams were impeded by unpredictable errors caused by this effect. The synchronisation was eventually solved with a robust XAUI fault-detection/correction and synchronisation protocol developed by Hickish (2010), which allows the system to function (with some lost, flagged data) in the presence of transmission bit errors.

4.2 2PAD Analogue Subsystem and Calibration

The 2-PAD RF signal system is a wideband radiometer. The analogue subsystem is the work of the other members of the 2-PAD collaboration, at the University of Oxford and the University of Manchester. Extensive details of low-noise amplification, gain signal conditioning, and signal transport are described in Ikin et al. (2009), and only a broad summary relevant to this work is presented here.

Figure 4.8 is a system diagram for the signal transport before digitisation. After reception by the antenna, the signal is converted to coaxial by a balun transformer. This signal is then amplified by a Mini-Circuits ZX60-33LN low-noise amplifier (LNA) before being passing through an optional filter module, which selects a 40 MHz band centred at 875 MHz for RFI-contained testing. Three gain stages in the gain-chain module (GCM) amplify the signal and a transformer drives one twisted pair of the category-7 (CAT-7) ethernet standard cable. Primary analogue equalisation and filtering occurs in the signal conditioning module (SCM).



(a)

(b)

Figure 4.7: Signal (analogue and digital) processing rack. 4.7(a) is a replica setup of the system at Jodrell Bank Observatory showing IBOB boards for a single polarisation racked above the analogue conditioning hardware. 4.7(b) is the 2-PAD RFI containment bunker which provides a secure, portable and temperature controlled ($\sim$ 20°C) environment for the 2-PAD digital systems and signal conditioning equipment. The Lovell Telescope is visible in the background.

Front-end components draw power from the SCM over the CAT-7 cable. It is noted that a balun transformer is a severely sub-optimal solution since it is a lossy RF component which enters into the radiometer *before* low-noise amplification. It was designed for integration with the commercially-available ZX60-33LN LNA. A future iteration of the analogue chain will contain a custom differential LNA, obviating the use of a balun (see Bhaumik and George, 2009, for details of LNA development).

4.2.1 Analogue System Calibration

Calibration is the process of measurement of and compensation for bandpass characteristics which are not constant after characterisation. Digital system characteristics are invariant with environmental conditions.

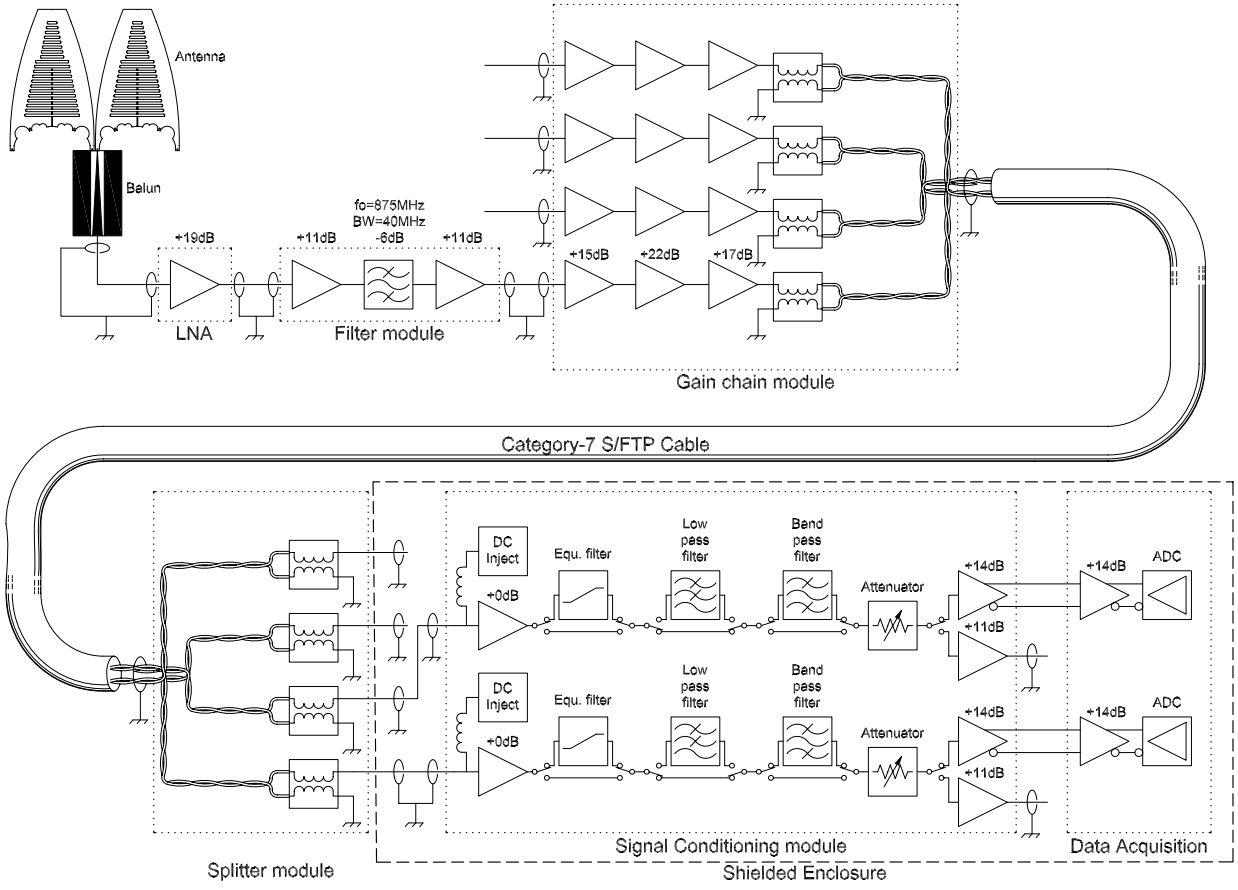


Figure 4.8: A sketch of the 2-PAD analogue system showing a single analogue gain chain. After reception by the antenna, the signal is converted to a a (single-ended) coaxial-compatible form with a balun transformer. It is then amplified by a Mini-Circuits ZX60-33LN low-noise amplifier (LNA) before being passed through an optional, custom filter module (FIL). Three gain stages amplify the signal before being sent over one twisted pair of a category-7 ethernet cable. Primary analogue equalisation and filtering occurs in the signal conditioning module (SCM) before digitisation. The components that are 'downstream' of the category-7 ethernet cable are housed in an electromagnetically-shielded container to reduce self-generated RFI. Figure from Ikin et al. (2009).

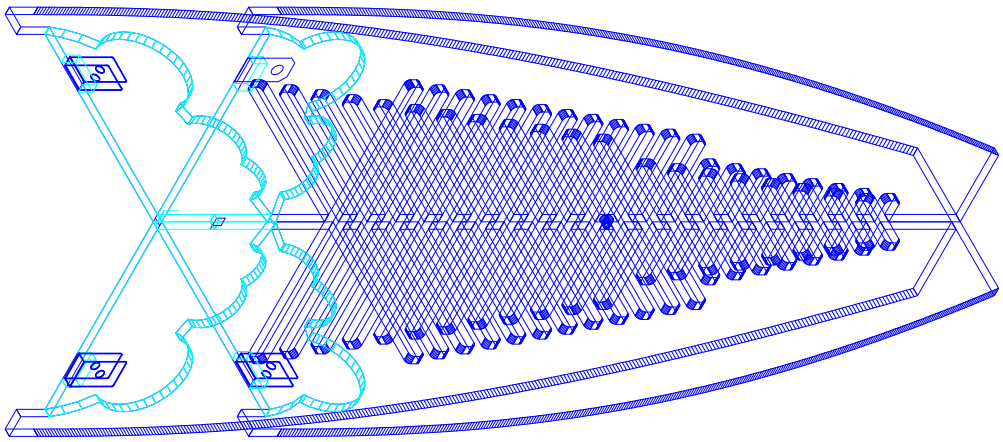


Figure 4.9: A schematic of the antenna element used in the 2-PAD instrument, the Bunny-Ear Combline Antenna (BECA). It is a projected schematic of an individual BECA element with both polarisations. This is the same antenna used in the beam pattern optimisation of Chapter 5. Figure courtesy Tim Ikin, Jodrell Bank Observatory.

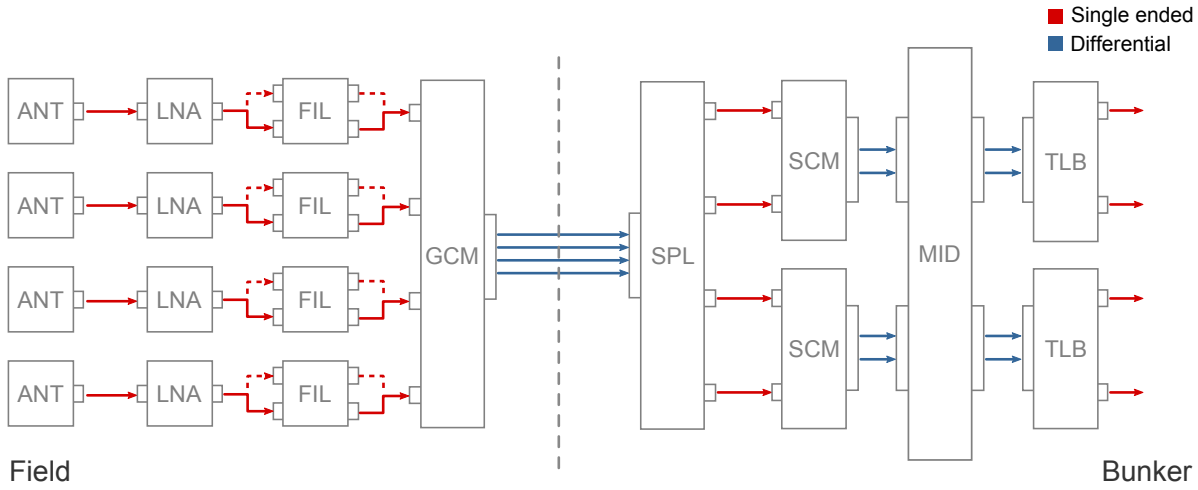
Analogue components, on the other hand, are expected to change with environmental conditions[7]. Differences in materials and construction of components result in non-identical RF performance, especially over wide bands. A typical analogue system is dependent on the observation frequency, the ambient temperature, the characteristics of specific components and the environmental RFI, (for e.g., see Tyler et al., 2004; Price et al., 2009) and needs to be continuously *calibrated* on timescales faster than the fastest varying of the set of all these.

Thus, each antenna signal channel will include a randomly distributed gain (δ_n) and phase (Φ_n) with a generally correlated non-zero mean. Calibration intends to remove the mean of random gain and phase (μ_δ and μ_Φ) errors generated in the analogue signal chain, so that all that remain are Gaussian residuals.

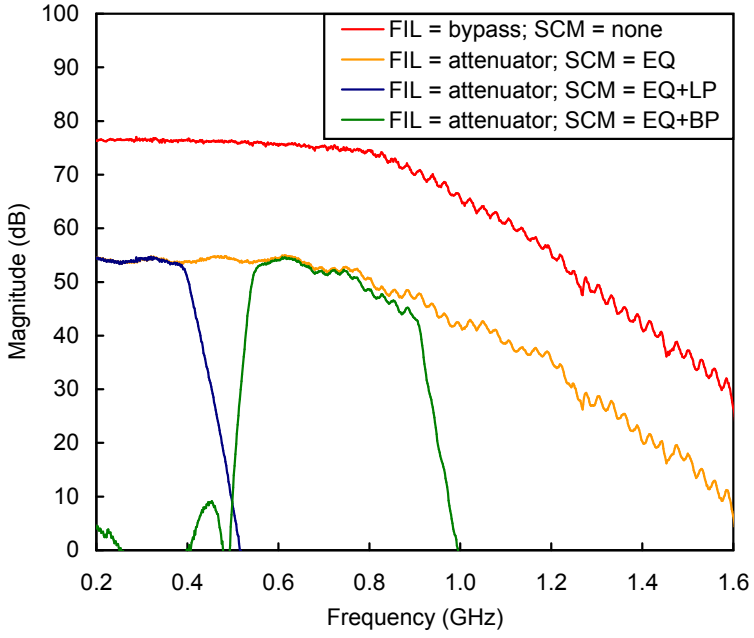
There are various methods for determination the decrease in beamformer gain due to random phase errors. Residual (Gaussian) channel errors decrease the array power pattern and increase sidelobe level. For example, Skolnik (1969) derive the average power pattern due to random amplitude (δ) and phase (Φ) errors. The gain reduction due to calibration residuals is

$$\langle |\mathbf{b}_\omega(\theta)|^2 \rangle = |\mathbf{b}_\omega(\theta)|^2 + (\delta^2 + \Phi^2) \frac{\sum_m x_{m,\omega}^2}{(\sum_m x_{m,\omega})^2} \tag{4.1}$$

[7]Price et al. (2009) discuss the 2PAD analogue signal chain, including modeling with S-parameter propagation, response variation due to temperature, component anisotropy and radio-frequency interference (RFI) effects.



(a)



(b)

Figure 4.10: Figure 4.10(a) is a schematic of the 2PAD analogue system from Price et al. (2009). The left half of the figure depicts the components that are on the physical antenna structure. Signals are received by the antenna, proceed through a low-noise amplifier (LNA) module, filter module and gain-chain module (GCM). The GCM modulates 4 signal channels over category 7 ethernet cable (CAT-7) for transmission to the shielded processing bunker represented on the right half of the figure. Signal demodulation is performed by the splitter (SPL) module. Next, the signal conditioning module corrects for bandpass nonlinearity (i.e. it performs analogue equalisation) and filtering before being sent through the interface boards (MID and TLB). Figure 4.10(b) is a plot of the gain magnitude of the 2-PAD analogue system in broadband mode, taken as a single measurement. The blue curve is the response with the attenuator activated on the filter module and the equalisation and low-pass anti-aliasing filters activated on the signal condition module, while the green curve is the same except that it shows the response with the band-pass anti-aliasing filter.

4.2.2 Bandpass Equalisation

The system bandpass is equalised digitally, for a Gaussian input signal. A calibrated broadband noise source signal[8] is amplified, filtered to the range 0.5-0.7 GHz and then split with a 16-way power divider[9]. After the power divider, the input level is -22dBm, which is used as the input to the 2PAD instrument. Figure 4.11(a) shows the measured bandpass due to this band-limited noise source that is passed through the analogue signal chain and connected to each of the beamformer inputs.

The calibration procedure proceeds to find the inverse of the bandpass at the central frequency for each frequency channel (Figure 4.11(b)). This is applied as an 16-bit amplitude coefficient in the beamformer to yield the amplitude calibrated bandpass (4.11(d)).

4.3 Deployment and Testing at Jodrell Bank Observatory

The digital beamformer was installed at the 2-PAD site inside the RF-shielded bunker at Jodrell Bank Radio Observatory. Figure 4.15 shows the antenna array and processing bunker at the site.

The test source was provided by a narrowband, polarised transmitter positioned atop a 35m mast at a distance of ∼0.2km (in the far-field of the receive array) and pointed at the array. Test frequencies that were permitted at the radio-quiet site were Continuous Wave (CW) sources at 0.675GHz and 0.530GHz. Phase calibration was performed at these frequencies.

4.3.1 Element-Phase Calibration

Elemental phase calibration corrects for signal path length differences between each of the array elements due to, for instance, phase mismatch between digitisation channels and array pointing offset. Even minute offsets in pointing direction, such as structural deformations severely affect the ability to coherently sum signals from multiple elements.

Calibration proceeds as follows: a sinusoidal input signal is amplified and transmitted from a (far-field) radio mast toward the array. The algorithm iteratively shifts the digital phase applied to each channel signal in order to maximise the ensemble average of N measurements of beamformer output power, $\langle |b_k|^2 \rangle$, $(0 \leq k \leq N)$ of the beamformer for each incremental frequency. The phase of a reference element is fixed,

[8] the Atlantec AS4803 calibrated noise source (www.atlantec.com).
[9] Minicircuits 16-way power splitter ZC16PD-252+ (www.minicircuits.com).

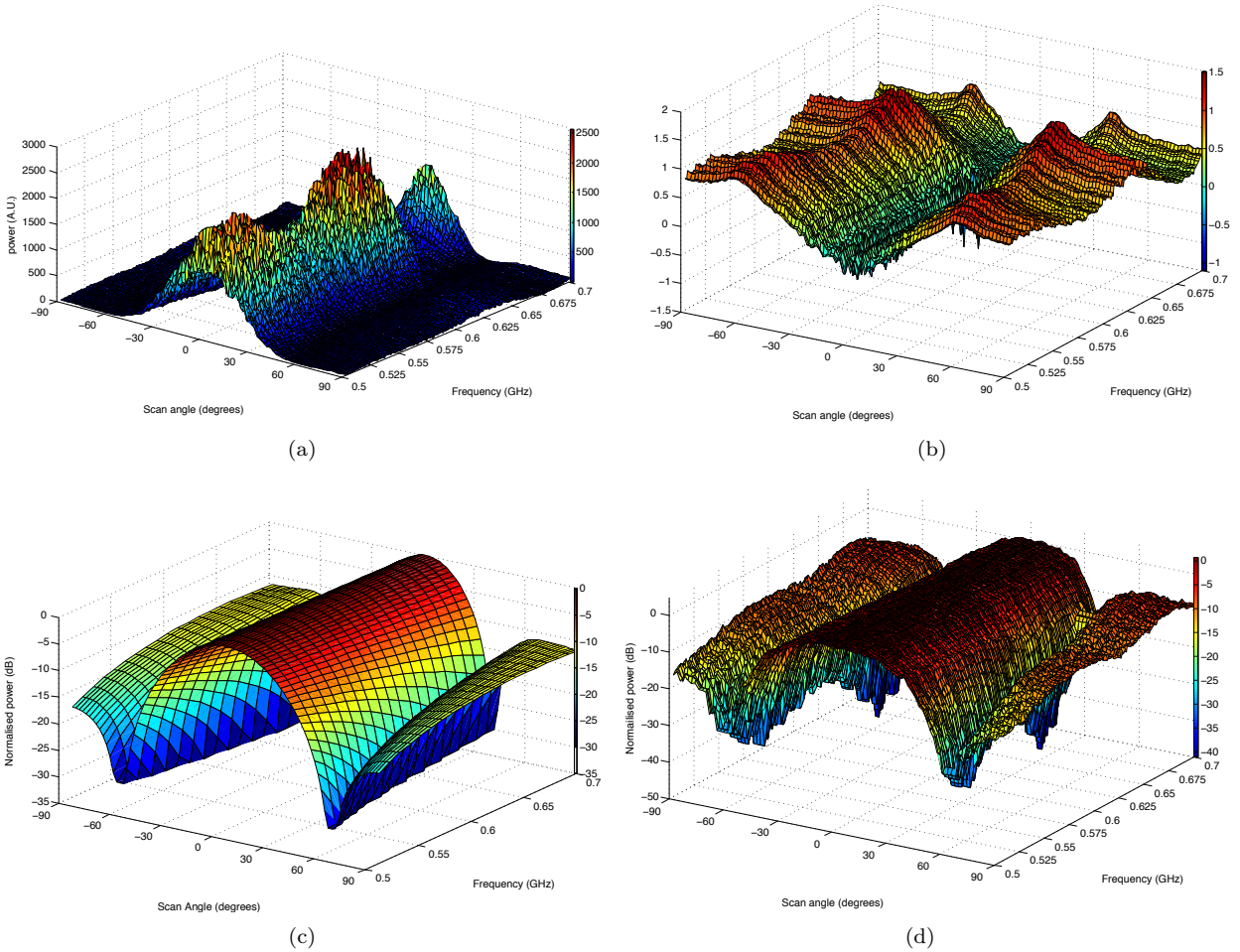


Figure 4.11: Broadband bandpass equalisation. This set of plots show how the power response — which initially varies greatly with frequency — is corrected with a measured correction matrix. (a) is a linear-scale surface plot showing the pre-equalisation beam bandpass plotted over the frequency range 0.5 to 0.7 GHz, and over the full range of scan angles. The bandpass has a characteristic shape due to both the analogue and the digital system contributions. (b) shows the (unit-less) bandpass correction factor surface, which is the inverse system transfer function $H^{-1}(z)$. Application of this correction matrix should correct the system bandpass to within numerical accuracy. (c) shows the ideal broadband beam surface (where the channel width $\Delta f \rightarrow 0$ and has no frequency smearing) plotted on a a normalised logarithmic scale. The main lobe of the beam broadens with decreasing frequency. (d) is the corrected instrumental bandpass after application of the correction surface. A 16-bit quantised representation of the equalisation factor is used for each frequency bin.

while the phase of each other elements are swept through 2π in the horizontal plane. The set of correction factors, $\mathbf{p}_\omega$, is that which results in maximum beam power. Tyler et al. (2004) describe a similar approach to adaptive antenna calibration, but do not take the ensemble average of the output power, producing 'snapshot' calibrations. This process is repeated for each (narrowband) frequency channel. However, RFI requirements at the Jodrell Bank radio-quiet site meant that only 0.675GHz and 0.530GHz were tested. If continuous calibration is required during operation, a strong reference signal, such as a periodically-injected sinusoidal reference or bright sky source is required.

Correlation-Based Calibration

Correlation-based methods allow calibration off the actual data signal rather than a specific reference signal. At a minimum, a correlation-based calibration technique must form an Nx1 correlation matrix, which gives phases relative to the reference element, similarly to the reference signal method. A full correlation (NxN) of all baselines would allow the correlation matrix of the array, $\mathbf{R}$, to be formed and a linear least-squares solution to the gain and phase of each element. However, a full real-time correlator may require an even larger digital system than the digital beamformer, since the computational complexity of direct correlation is $O(N^2)$. Signal correlations are not an immediate product of a beamformer, and a separate signal path must be provided if these are required. Calibration approaches which require the correlation matrix of signal inputs to be known (such as Wijnholds and Boonstra, 2006; Warnick et al., 2009; Wijnholds and van Cappellen, 2011) require a correlation matrix that is obtained at a rate less than the data rate, such as a time multiplexed scheme.

4.3.2 2PAD Beam Response Measurement

The beam response in the field is the essential test of a complete digital beamforming system. Figure 4.12 is a plot of the beam response of the 2-PAD digital beamformer. It shows the beam pattern at 0° (*cf.* the laboratory response of Figure 4.11(d)) at 0.675GHz. Plotted on the same axes is an ideal response for comparison. The response deviates a maximum of 0.22 dB in the main lobe and the first side lobes are measured as low as -23 dB compared to the maximum power.

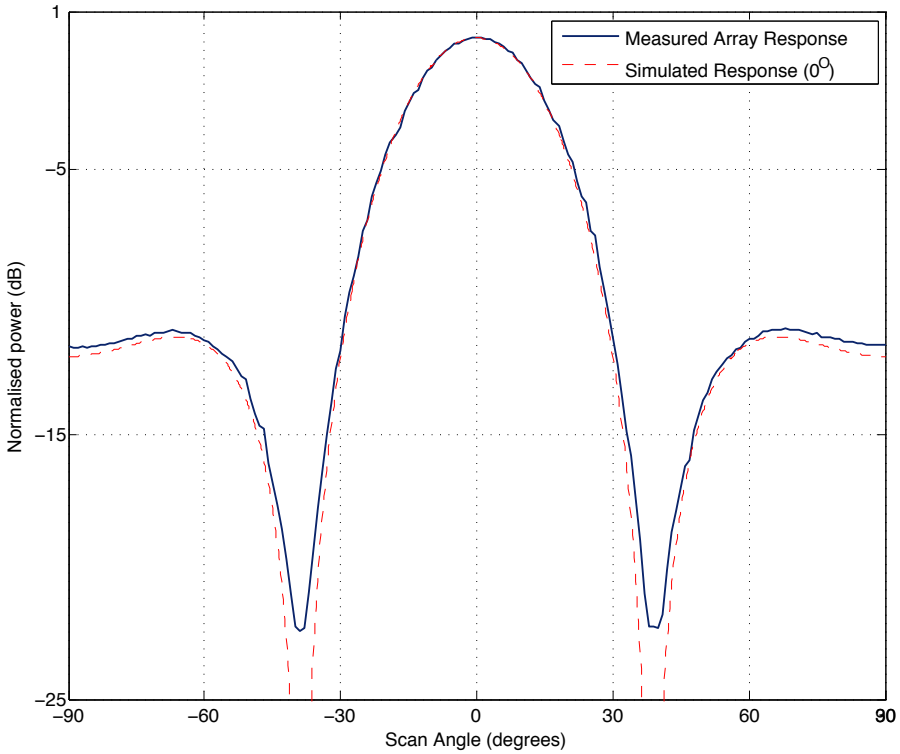


Figure 4.12: A plot of the sixteen-element (4x4) beam pattern of the array at 0.675 GHz as measured in the field at Jodrell Bank plotted alongside a simulated 4x4-element beam. The elements were calibrated while pointed toward the transmitter. The simulated beam was calculated for the aperture spanned by a 0.675 GHz signal for the 2-PAD array, which has Nyquist element spacing at 0.882 GHz. The abscissa is the scan angle in degrees from broadside and the ordinate is an arbitrary power scale relative to maximum signal power at 0° pointing.

Electromagnetic Coupling

Figure 4.13 shows the effect of environmental electromagnetic coupling. The plot shows the beam response of both beams as the array is mechanically rotated by 80 degrees. This is compared to a digital beam response for both test frequencies. This measurement shows the effect of coupling as a function of scan angle. The response deviates significantly compared with the ideal response within the main lobe of the beam. Deviation is more severe near the predicted nulls.

Application of a phase-gradient with the array still pointed at broadside will not have the same effect. A phase-gradient should shift the signal to match the simulated response more closely. This was not possible to confirm, since the position of the testing transmitter was fixed.

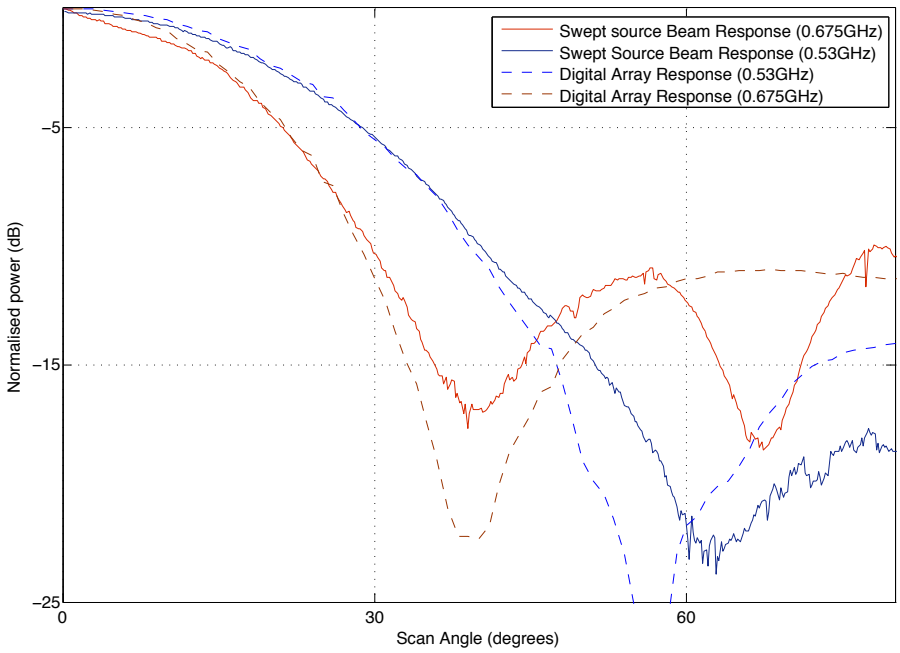


Figure 4.13: A plot of the digital beamformer response compared to the response as the array is mechanically rotated through 80°. This is done for both independent beams, at 0.675GHz as well as at 0.53GHz. The mechanical rotation exposes the electromagnetic connection of the array to its surroundings. The abscissa is the scan angle in degrees from broadside and the ordinate is an logarithmic power scale relative to maximum signal power at 0° pointing.

4.3.3 Polarisation Response

2-PAD is a dual-polarisation instrument. Isolation of the separate polarisation channels is a strong requirement of key SKA science. In particular, Rotation Measure (RM) surveys of Galactic magnetic sources require effective isolation of polarisation channels.

Polarisation crosstalk is an electromagnetic attribute of the array. It is able to be measured with the digital beamformer, but is not caused by it. The transmit antenna is linearly polarised, and aligned with the X-polarisation of the array, as closely as possible. Figure 4.14 is a plot of the *cross*-polarisation response of the 2PAD digital beamformer at 0.675GHz. The scan shows signal power leakage into the Y-polarisation channel from the X-polarisation channel.

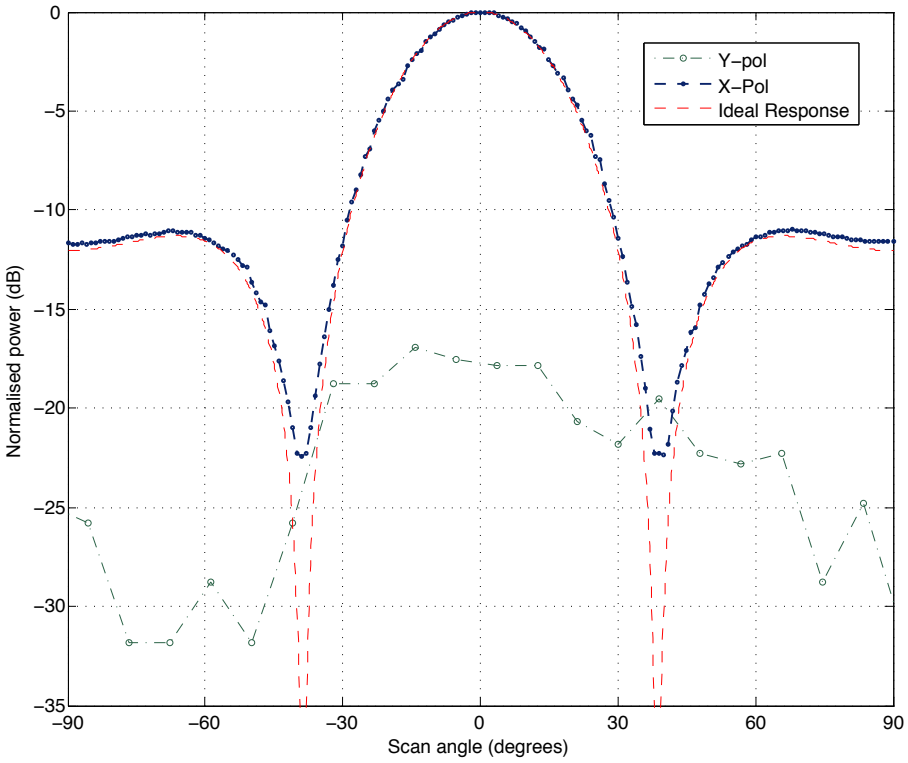


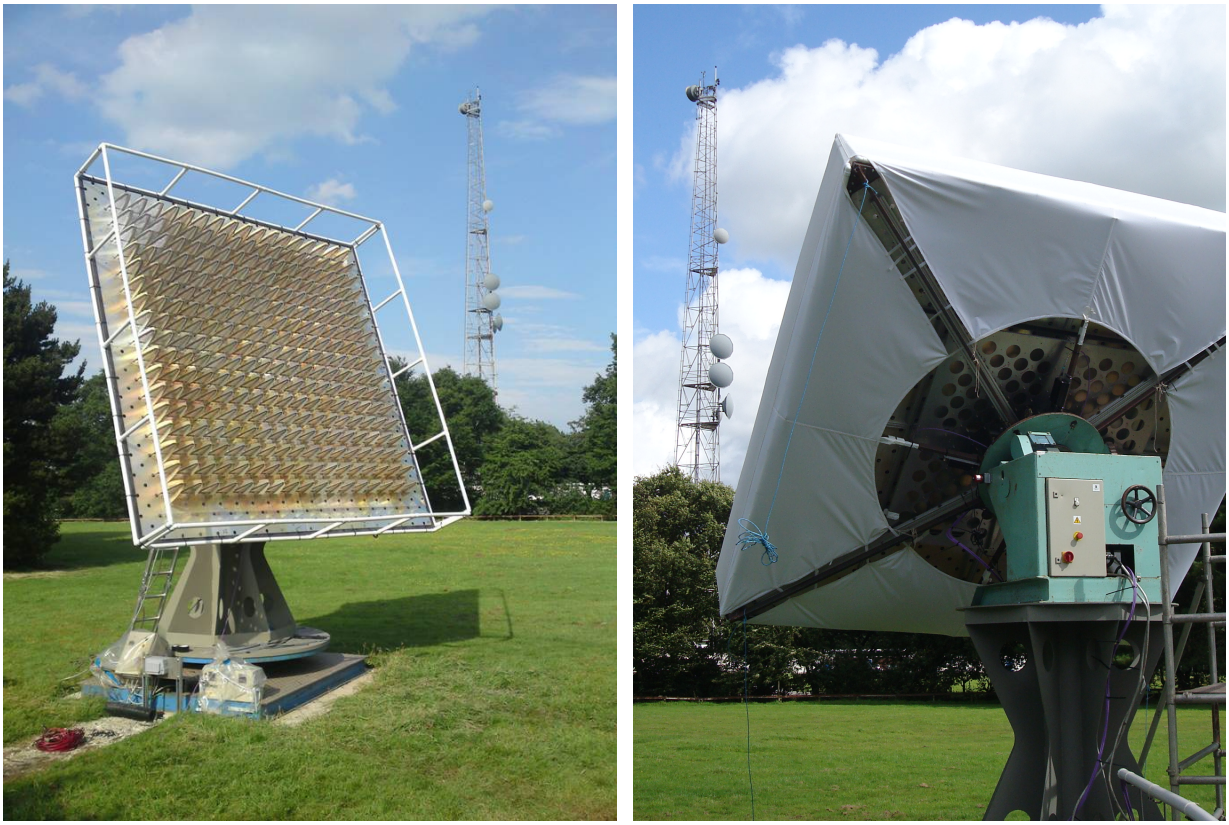
Figure 4.14: Polarisation response of the 2PAD sixteen-element (4x4) digital beamformer at 0.675GHz. The polarisation response is an electromagnetic attribute of the array. Digital computation itself does not add to cross-polarisation leakage. The plot shows the the calibrated, unweighted digital beamformer response in each of the two (X- and Y-) polarisations to a linearly polarised source transmitter. The abscissa is the scan angle in degrees from broadside and the ordinate power in dB.

4.3.4 Summary

A digital tile beamformer is the principal component of an adaptive aperture array. Signal processing systems for large arrays are constructed from many replicas of this computational structure. The basic computational structure is repeated many times over and combined hierarchically.

The 2-PAD project is a phased aperture array demonstrator that will confirm the feasibility of highly-coupled, wide-bandwidth beamforming entirely in the digital domain. Its primary aim is to test individual components of the system, including antenna design, RF amplification and transmission, digitisation, dataflow and signal processing for future aperture array radio instruments.

Central to 2-PAD is the wide-band digital beamforming system. This chapter has described the design, characterisation, and deployment and testing at Jodrell Bank Observatory of the 2-PAD digital beamformer. Future-generations of all-digital beamformers, of which 2-PAD is an early prototype, will form the core of



(a) (b)

Figure 4.15: The 2-Polarisation All-Digital (2-PAD) aperture array demonstrator at Jodrell Bank Radio Observatory. 4.15(a) shows the full antenna array configuration including terminated dummy elements installed on a 3-axis mount, which allows arbitrary directional pointing. In the background to the right is a radio mast with transmitting antenna used for field beam-shape measurement and phase-alignment calibration. 4.15(b) shows the array pointed toward the transmit antenna with the weather covering attached.

the SKA dense aperture array.

Chapter 5

Stochastic Optimisation of the Beam Response in Small Arrays

The problem of determining current distributions for optimal array beam patterns is compounded by having both a small number of discrete elements and non-isotropic element patterns. Classical analytic solutions, such as those determined by Dolph (1946) and Taylor (1955), do not produce optimal patterns in this case. In this regime, heuristic algorithms are able to outperform the analytic methods, and even correct for severely non-ideal individual element patterns. The primary purpose of this or, indeed, any array pattern synthesis optimiser is to maintain the desired pattern profile under such conditions.

In this chapter, the design of a four-element digital beamforming architecture is detailed, as is its use in the measurement of individual immersed element patterns of a electromagnetically-coupled four-element array. Hierarchical processing structures for SKA aperture array processing favours the use of a small array as the basic unit (a 'tile') of a large beamformer system; large arrays will likely be constructed in a modular way, composed of many small arrays, and thus the small-array beam shape is of importance. An heuristic algorithm, particle swarm optimisation, is used to optimise the digital elemental current distributions in order to improve the full array response. Stochastically-optimised current distributions obtained with the measured immersed element patterns are eventually applied to a physical digital beamforming system, measured, and compared with the predicted results.

5.1 Current Distribution in Beamformed Arrays

Element weighting effectively modifies the current distribution of electromagnetic radiation in the effective aperture defined by a beamformed array with the aim of improving particular aspects of the array response. The majority of early work in this area is analytical; results are provably optimal when there are many, ideal elements. This work concerns current distribution in small arrays when the element response is not ideal.

5.1.1 Dolph-Chebyshev and Taylor-Chebyshev Synthesis

Dolph (1946) develop a current distribution technique for a uniformly spaced linear array (ULA) that produce either a theoretically minimal beamwidth for a given sidelobe level, or, alternatively, a theoretically minimal sidelobe level for a given beamwidth. The technique is developed from the orthogonal polynomial sequence of Chebychev polynomials (Chebyshev, 1854).

Taylor (1955) bring to notice that the requirement of Dolph-Chebychev synthesis that all side-lobes be exactly equal causes deficiencies in the synthesised pattern and possibly unrealisable excitation values due to the concentration of energy at the end elements (i.e. elemental excitations tend to singular values in large arrays). To counter this issue, and in a similar manner to Dolph, Taylor derived a *continuous* aperture pattern from the Chebyshev polynomial sequence without the strong sidelobe equality requirement. The pattern is constructed such that only the first $\bar{n}$ side-lobes are equal, with those that are further from the phase reference[1] decaying as $\sin x/x$. This ensures that the aperture distribution at the edges of the array is not singular.

In discrete phased arrays, patterns are determined by sampling the continuous Taylor distribution at the location of the individual elements. In large arrays, such sampling results in beam shapes approximating the Taylor pattern. However, in small[2] arrays, and arrays with any or all of the properties of anisotropy, heterogeneity, and strong mutual coupling, the application of discrete Taylor weights may not produce optimal results. This will be shown below.

5.1.2 Schelkunoff's Unit Circle Form and Z-plane Transform

Schelkunoff (1943) express the array factor in terms of the polynomial form. The central concept is that every uniform linear array can be represented as a polynomial, as every polynomial can be interpreted as

[1] The phase reference of the array is defined to be the geometrical centre of the aperture for a Uniform Linear Array (ULA)
[2] A 'small' array is defined here, as in standard references like Elliott (2003), as one with less than 8 nodes per dimension.

a uniform linear array. Pattern synthesis, as regards positioning of nulls, side-lobes and grating lobes is performed in polynomial space and accomplished by positioning of polynomial roots of the array equation along the unit circle. The array polynomial is

$$F(u) = \sum_{n=0}^{N-1} a_n z^n$$

where z^n are the (complex) roots of the polynomial expansion and $a_n = A_n e^{ikd_x}$ are the elemental current excitations of an array with N elements. Central to Schelkunoff's theorem is that the roots of this polynomial all lie on the unit circle in the complex plane. This observation may be used to manipulate positions of the polynomial roots (which relate to the distribution of array factor nulls, side-lobes and grating lobes in the array factor) instead of manipulating the array factor directly. This leads to algorithmically simpler forms than the direct approach (see e.g. Elliott, 2003).

5.1.3 Heuristic Algorithms in Pattern Optimisation

The traditional analytic solutions for current distributions in phased arrays do not perform optimally when there is asymmetry between individual element patterns, there are not enough degrees of freedom in the array polynomial (i.e. there are only a small number of elements), or when there is a complex electromagnetic coupling relationship between elements. In these cases, an unbiased optimisation algorithm may return a set of current distributions which both correct for elemental anisotropies and enable synthesis of a more desirable array pattern.

A large number of authors have worked on general numerical array optimisation, as reviewed extensively by Bucci et al. (1994). Genetic Algorithms (GA) and, more recently, the Particle Swarm Optimisation (PSO) algorithm, are increasingly popular choices for array optimisation. Such heuristic optimisation techniques have been used together with the Schelkunoff's unit circle representation of the array polynomial for large uniform linear array optimisation. Yu et al. (2004) and Monorchio et al. (2007) use Genetic Algorithm (GA)s in Schelkunoff-optimisation schema to achieve an optimal main beam and side-lobe combination.

Previous work has also performed comparison of these heuristic algorithms. Boeringer and Werner (2004) compare the PSO and general GAs by application to the problem of Notch synthesis in large (10^2-element) arrays. Lopez and Basterrechea (2007) compare a large set of classical and hybrid PSO algorithms for array synthesis particularly with regard to accelerating convergence. Bevelacqua and Balanis (2007) use PSO for

non-uniform linear arrays.

Particle Swarm Optimisation

Introduced by Kennedy and Eberhart (1995), Particle Swarm Optimisation (PSO) is an iterative optimisation algorithm which attempts to obtain a near-optimal solution within a given problem domain. Like other metaheuristic computational methods, the algorithm evaluates and ranks potential solutions and selects further candidate solutions without strong priors or assumptions about the solution space. Furthermore, PSO does not require that the problem space be differentiable. Metaheuristic algorithms, including PSO and the class of genetic algorithms, do not guarantee that an optimal solution be found, nor do they guarantee convergence.

The metaphor upon which the algorithm relies is as follows: solution candidates in each iteration are termed 'particles', which each have velocities that govern their 'trajectories' (see Trelea, 2003). Selection of the next generation of candidates is determined by the Newtonian particle-like vector motions of the candidates within the problem hyperspace. The motion of each candidate particle is influenced by the 'swarm' of contemporary candidates (especially those ranked highest), the trajectory history of the particle, as well as its own current velocity. This generates a dynamical system which is essentially chaotic (Boeringer and Werner, 2004). Since the particles act classically and atomically, overshoot and undershoot of the target points stochastically probe regions around high-scoring subspaces.

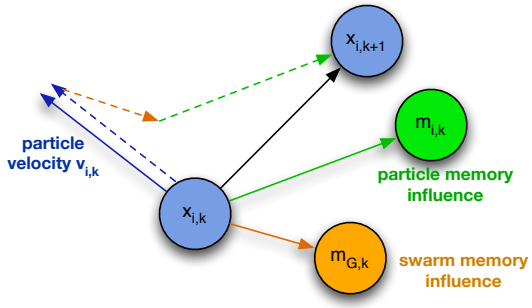


Figure 5.1: A schematic of the vector forces acting on each particle in the PSO algorithm. The particle's position at time $k+1$, $x_{i,k+1}$ is a combination of the individual particle velocity $v_{i,k}$, the particle memory influence $m_{i,k}$ and the swarm memory influence $m_{G,k}$.

The algorithm does not make any strong problem-space suppositions, nor requires differentiability, unlike traditional gradient-based GAs (see Trelea, 2003). PSO has advantages over other algorithms in its class; for

instance, Hsu and Shyr (2005) show that PSO outperforms genetic algorithms in a variety of unconstrained nonlinear array problems with continuous design variables. It is quite similar in effect to the class of genetic or evolutionary algorithms, but requires less computational administration, which partly explains the outperformances observed by Hsu and Shyr (2005). Since it is not desirable here to make any problem-space assumptions, PSO is well-suited to the small-array immersed-element pattern optimisation task.

5.2 Digital Beamforming System

The digital beamforming system is built on FPGA-based signal processing hardware, to implement a frequency-sub-band phase-shift beamforming architecture. A set of FPGA gateware modules perform signal digitisation, down-conversion and frequency-splitting of each antenna-element signal, and digital beam-shaping and summation for all the elemental signals.

The digital system for the four-element beamformer allows application of frequency-dependent element excitation coefficients, linear phase gradients (for beam scanning), and a frequency-dependent bandpass compensation. These attributes are not, in general, achievable with simple digital time-delay beamforming nor with analogue approaches.

5.2.1 Digitisation and Nyquist Sampling

Signals from array elements are digitised by two Atmel BiCMOS AT84AD001B dual 8-bit ADC, each of which is able to digitise two streams at $1\text{GSa}.s^{-1}$. Each channel of the ADC may receive a 500 mVpp differential RF input.

The ADCs are run at 800MHz. The entire required 0.5 - 0.7GHz band is super-Nyquist sampled[3] as a single, real voltage and then digitally mixed to baseband with precise in-phase and quadrature-phase components. Two iADC boards are connected to each of the IBOB processing boards.

On the IBOB, a Xilinx Virtex II Pro V2P50V FPGA performs the digital down-conversion. Spectral decomposition is again accomplished with a Poly-phase FFT Filter Bank (PFFB). The design uses embedded multipliers in a Xilinx Virtex II Pro device. Since these fixed silicon blocks are natively 18-bit, we

[3]Operation in this 2nd Nyquist zone is possible since the AT84AD001B ADC has a output code settling time of 997ns, which dictates a maximum frequency of 1.003GHz, and satisfactory operation in those Nyquist zones which are less than this frequency (i.e. 1st and 2nd Nyquist zones). Accurate full-swing performance will be degrade for the higher Nyquist zones. See Chapter 4 Section 4.1.6 for further discussion of super-Nyquist sampling issues.

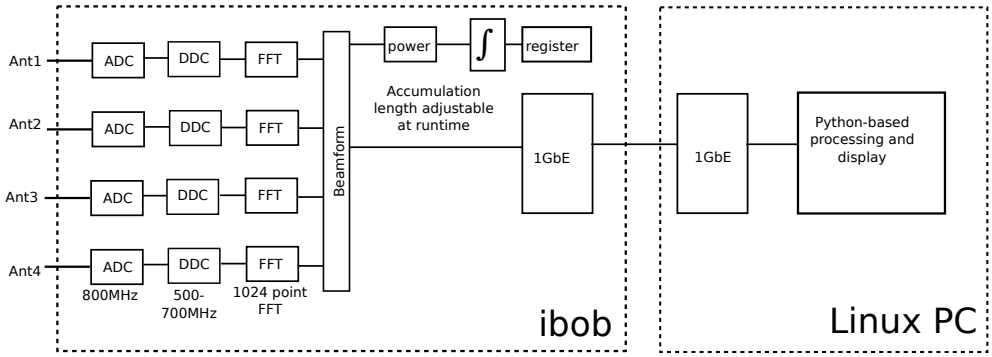


Figure 5.2: A schematic of the digital beamformer and digital signal processing architecture. The ADC samples the incoming antenna signal at 800MSas^{-1} with 8 bits of precision. This signal is immediately Digitally Down-Converted (DDC) by mixing with an complex sinusoid at 3/4 of the ADC clock rate to yield a base-banded 500MHz to 700MHz input signal. Beam spectra are then sent to a control computer over the UDP protocol.

preferentially use 18-bit multiplications and datapaths throughout the design. Figure 5.2 is an architectural schematic of the 4-element beamformer digital system.

The beamformer applies a complex coefficient to each of the spectrally decomposed, narrowband channels of the input signal, then coherently sums each of the four digital streams. The resulting beam is sent via the UDP ethernet protocol to a control machine.

5.3 Optimised Array Factor Measurement

The principal aim of this work is to improve the array response for small arrays with non-ideal element patterns. In order to achieve this, the immersed element patterns of each of the array elements is measured, then used as inputs to the PSO array pattern algorithm. Finally, the optimised element current excitations are applied, and the optimised array response is measured to compare with the algorithmically predicted response.

5.3.1 Algorithm Framework and Implementation

Khan and Brown (2009) show that a bounded, hybrid PSO algorithm improves the convergence time significantly for small arrays by using the mutual coupling impedance matrix and active element patterns derived from the commercial simulation tool, CST®, together with a MATLAB® implementation of the bounded hy-

brid PSO. It is particularly noted that the software framework used in this experiment is entirely attributable to Khan and Brown.

The framework has been modified for this task so that the actual (measured) immersed element patterns are used as the optimisation inputs instead of simulated element patterns. The digital beamformer allows the algorithm to be tested on a functional array.

5.3.2 Active Element Pattern Measurement

A phased array is a unified electromagnetic entity. Electromagnetic coupling between elements is a critical and significant determinant of the entire response of the array, and may not be ignored for densely-packed arrays.

In order to model such coupling in the simplest case of an infinite array, a single *active element pattern* may be assumed to exactly represent every element; in this scheme there is consequently no requirement to measure each element individually in order to correctly model the entire array.

However, finite arrays exhibit edge-element behaviour that is not uniform across elements. In order to account for this, many authors have used a hybrid solution for edge and central elements (see for e.g. the review article by Pozar, 1994). Edge element patterns in *large* arrays may be determined as the product of the infinite array element pattern and the finite array correction factor.

Mutual coupling behaviour in small arrays resists the hybrid correction factor approach, since each element is effectively an 'edge-element' from all sides, and is unique in the array. In addition to this, edge-effect deviation is compounded by anisotropic manufacturing of individual elements which becomes non-Gaussian in small arrays.

The coupling profile of the small array used in this work is complex; to illustrate the electromagnetic coupling behaviour, Figure 5.3 shows the measured coupling coefficients relative to the first element[4]. The measurement uses scattering-matrix parameter (S-parameter) characterisation (see e.g. Pozar, 1997), which show the reflected and transmitted power waves between two ports of a device.

Due to the small-array effect, complex coupling, and suspected anisotropic element performance, it was judged necessary to measure each of the elemental patterns instead of using electromagnetically-simulated

[4]Other options are available to obtain coupling coefficients; Steyskal and Herd (1990) present a matrix-inversion method (a derivation of the Steyskal-Herd method is detailed in Elliott (2003)) using isolated element patterns for performing coupling compensation. The optimisation method used in this work does not require such *isolated* element patterns, since *immersed* patterns may be used directly.

patterns.

Phase-Coherence in Element Measurement

Phase-referenced measurements are required in order to determine the element (complex voltage) patterns; therefore a phase-coherent measurement tool is necessary. Since the digital beamformer does not itself produce the transmitted signal, nor have an extra input for receiving the transmitted signal, there is no way of recovering the transmitter phase reference. Thus, a Vector Network Analyser (VNA) is used for the task active element measurement task. The VNA also provides external validation that the digital beamforming system operates correctly. This is shown in Figure 5.11, where the VNA response conforms to the digital beamformer response to within tolerances of the analogue power-combiner device.

5.3.3 Experimental Environment

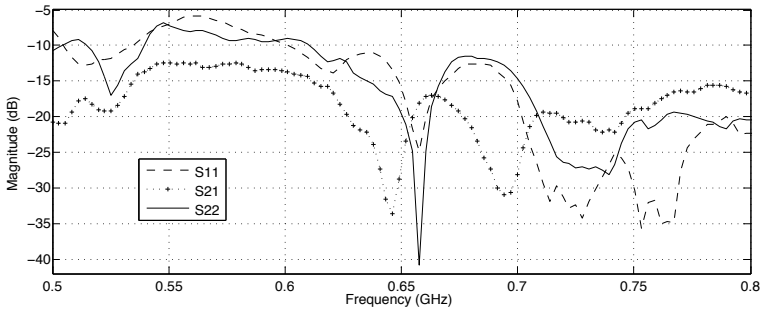
Antenna Array

Zhang and Brown (2010) present a dual-polarised antenna — called a Bunny-Ear Combline Antenna (BECA) — for wide-band signal applications, derived from the original Vivaldi concept by Lewis (1974). While conventional Vivaldi designs (such as those of Smolders 2000) perform acceptably in the principal planes (E- and H-planes), polarisation performance in the diagonal (D-) plane degrades significantly at high scan angles. The design goal of the BECA antenna is to combine Fermi tapering[5] (as introduced in Sugawara et al. 1997) with compact, Vivaldi-style geometry, in order to improve the radiation pattern with a lower side-lobe level, which is apposite to dual-polarisation, broadband electronically scanned arrays.

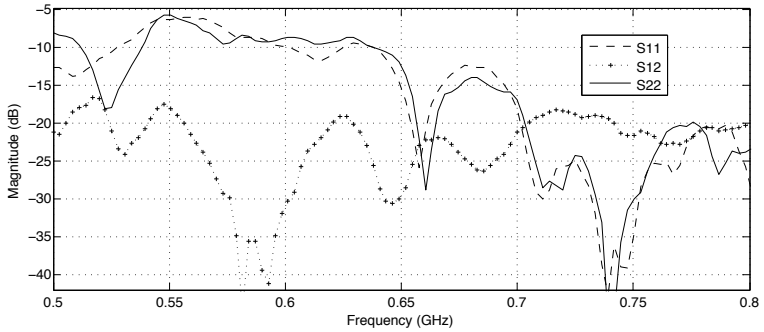
The antenna is designed for operation in the 0.3 to 1.0GHz frequency range. Half-wavelength spacing is at 882MHz, corresponding to ~ 170 mm, which is approximately one sixth of the free-space wavelength at the lowest frequency. This implies that the array element spacing is shorter than one wavelength at the top of the frequency range.

A 3.3:1 frequency bandwidth is achieved for a $\pm 45^\circ$ volume, and the characteristic impedance of the antenna feed is 169Ω. Zhang and Brown's antenna element is used in an 8-feed-element linear array with 4 active elements. The inactive or 'dummy' elements are terminated across 170Ω, approximately the char-

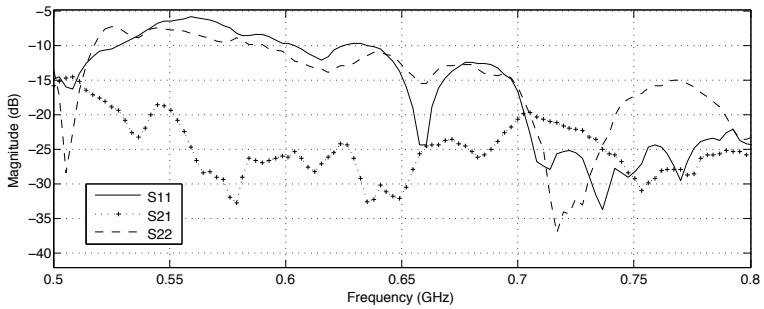
[5]The taper of the array uses a modified version of the Fermi-Dirac fermion occupation probability distribution, $f(E) = \frac{1}{1-e^{(E-E_F)/kT}}$, defined to be $f(x) = \frac{a}{1-e^{bx+c}}$, where a, b, and c are constants that define the particular attributes of the array. This tapering function has been shown to improve side-lobe response in scanned arrays by Sugawara et al. (1997)



(a) E(1,2)

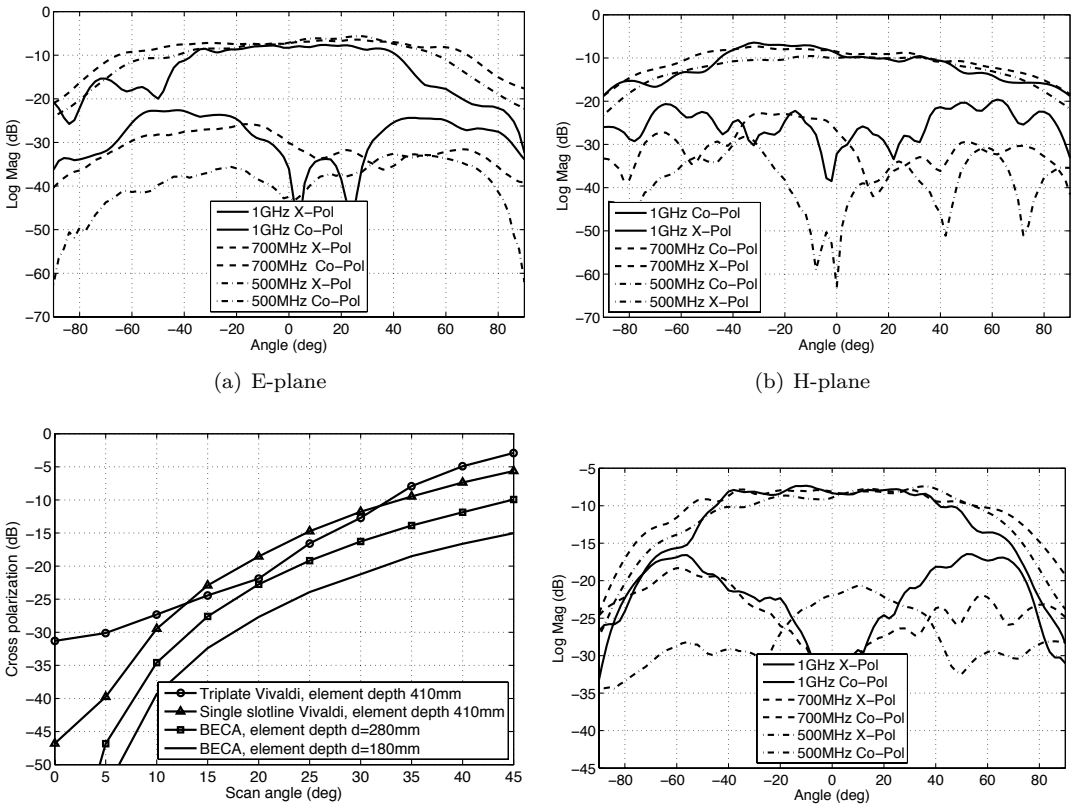


(b) E(1,3)



(c) E(1,4)

Figure 5.3: A plot showing coupling coefficients of the first element in the array over the operational frequency band. scattering-parameter measurements are taken with a vector network analyser (VNA). The coupling coefficients of the first element with each of the other elements is shown. Figure 5.3(a), labelled *E(1,2)*, shows the forward reflected (S11, dashed line), reverse-reflected (S22, solid line) and transmitted (S21, dotted line) power relationships between the inputs of elements 1 and 2. The other two figures (Figure 5.3(b) and 5.3(c)) show the same relationships for the inputs of elements 1 and 3, labelled *E(1,3)*, and for the inputs of elements 1 and 4, labelled *E(1,4)*. The coupling characteristics are the transmitted power relationships, which are a strong function of frequency. Coupling measurements for the other elements are omitted for space reasons.



(a) E-plane

(b) H-plane

(c) Cross-polarisation comparison

(d) D-plane

Figure 5.4: Simulated central-element radiation patterns for the BECA antenna design in the (a) E-plane and (b) H-plane. (c) shows the cross polarisation performance comparison of two Vivaldi-style antennas scans in the D-plane at 1 GHz, based on infinite array simulation for the dual-polarised array. Element spacing for the BECA antenna is 170mm, and element depth for the two Vivaldi antennas is 410mm and the element depth for BECA is 280mm and 180mm. From Zhang and Brown (2010).

acteristic impedance of the antenna. The active elements (the centre 4) are converted to 50Ω single-ended outputs (for connection to the digitisation circuitry) with a purpose-designed (lossy) balun. Figure 5.3.3 is a schematic showing the antenna array composed of BECA elements. The measured response at the design frequency is shown in Figure 5.7. Comparison of the simulated and measured responses show that an implementation of the array deviates from the expected pattern. Thus, an optimisation method based on predicted results is likely to fail.

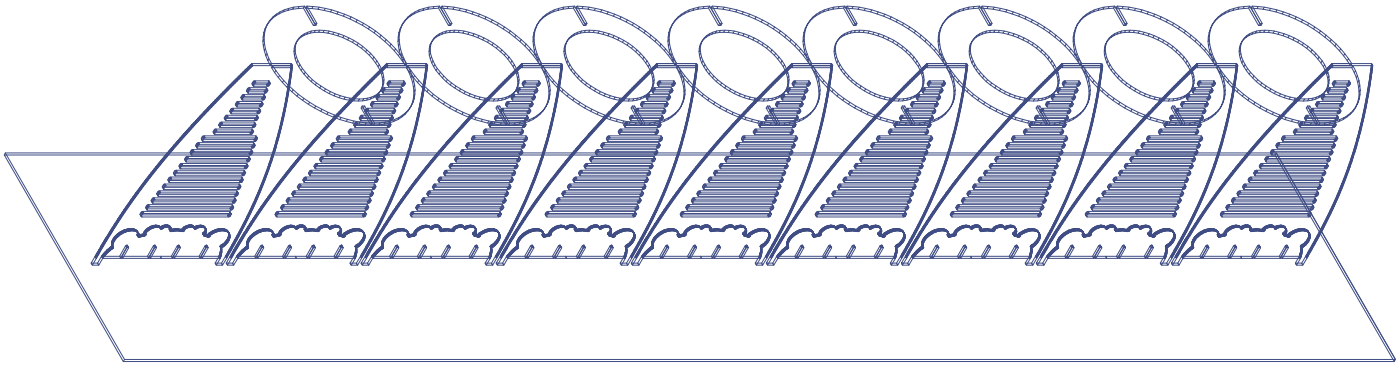


Figure 5.5: A schematic of the antenna array used in the experiments in this Chapter. The figure shows the linear array composed of Bunny-Ear Combline Antennas (BECAs). The array consists of eight single-polarisation feed elements. The four feeds used in the experiment are taken from the central four elements. The outer two elements on each side are terminated across 170Ω. The circular braces in the schematic were removed for these experiments. Schematic templates courtesy of Tim Ikin at Jodrell Bank Observatory.

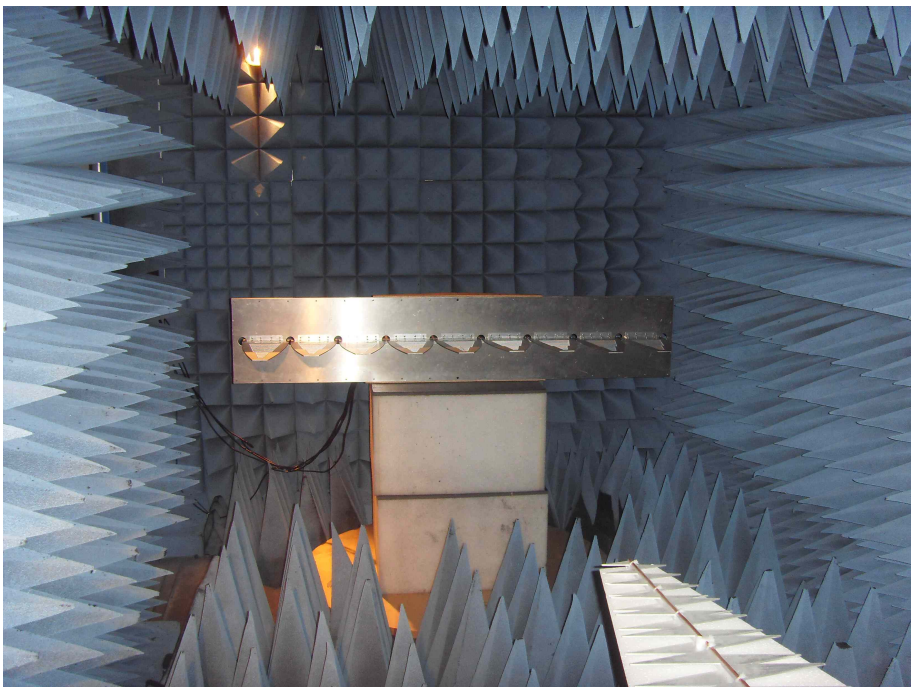


Figure 5.6: An image showing the measurement environment of the antenna array in the anechoic chamber. The BECA linear array is in the centre at the back of the chamber. A linearly-polarised transmit antenna is positioned at the other end of the chamber. This measurement environment is shown schematically in Figure 5.8. Figure 5.3.3 is a schematic of the linear array.

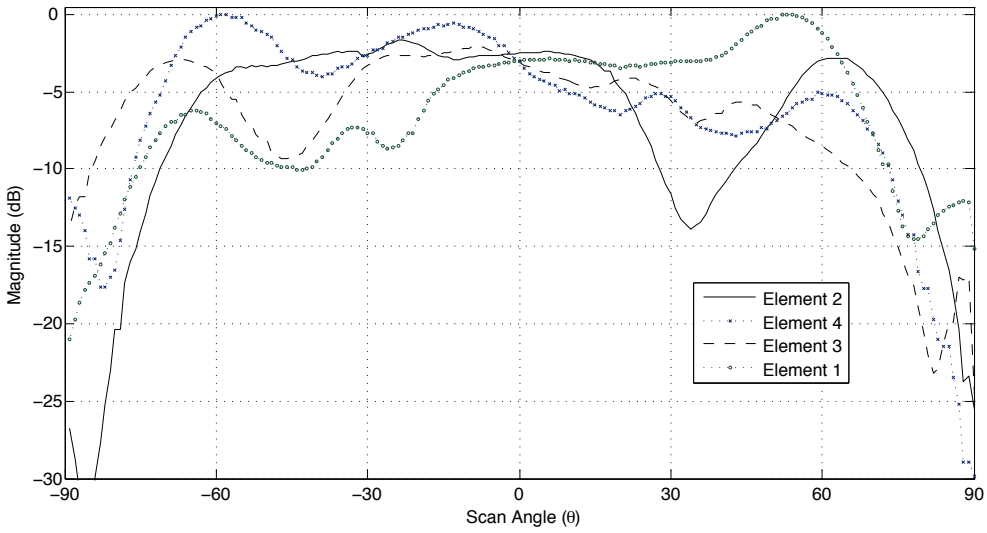


Figure 5.7: Magnitude of the immersed element patterns of the array elements at its design (half-wavelength spacing) frequency, 0.882 GHz, measured with the digital beamformer. Elemental responses are to first order approximately symmetrical. This is not the case at the testing frequency of 650 MHz (element responses are shown in Figure 5.9).

Anechoic Chamber and Fraunhofer Criterion Satisfaction

An anechoic chamber at the University of Oxford Department of Engineering is used for these measurements. The chamber is supplied by Rainford EMC. An image of the array in the anechoic chamber is shown in Figure 5.6.

The Fresnel number, $F = a^2/L\lambda$, where a is the aperture dimension and L the distance between aperture and receiver, is a dimensionless number which defines the division between the distance at which Fraunhofer or Fresnel[6] diffraction of an electromagnetic wave occurs. In antenna measurements, the generally-accepted (see for e.g. Elliott, 2003; Balanis, 1982) Fraunhofer criterion (i.e. the criterion for planar wave or far-field measurement) is that a receive antenna must be placed at a distance greater than $2D^2/\lambda$, where D is the physical dimension of the receive aperture. Operation in the 0.5 to 0.7GHz frequency range gives a wavelength range of 0.5995m to 0.4283m, while the total aperture length of the test array in the anechoic chamber is 0.684m. Operation in the far-field is valid for all frequencies in this range for transmitter separation greater than 2.1847m. Antenna separation in the anechoic chamber is approximately 3.5m, satisfying the Fraunhofer

[6]Fraunhofer diffraction is wave diffraction that occurs when field waves are passed through an aperture and observed as planar waves by a given receiver. It occurs when the Fresnel number, $F \gg 1$, where $F = a^2/L\lambda$. More complex diffraction, called Fresnel diffraction, occurs when $F \lesssim 1$, or when the distance from aperture to receiver is less than that required for planar reception. At Fraunhofer distances, the diffraction pattern of an observed aperture image is independent of the distance of the receiver from the aperture.

far-field criterion.

5.3.4 VNA Measurement Configuration

The VNA measurement configuration is shown in Figure 5.8(a) The transmit antenna was connected to Port 1 of the VNA. Port 2 was connected, in turn, to each of the individual array elements while its surrounding elements were terminated at the characteristic impedance. It was then connected to the Mini-Circuits ZFSC-4-1+ power combiner , which combines all active inputs.

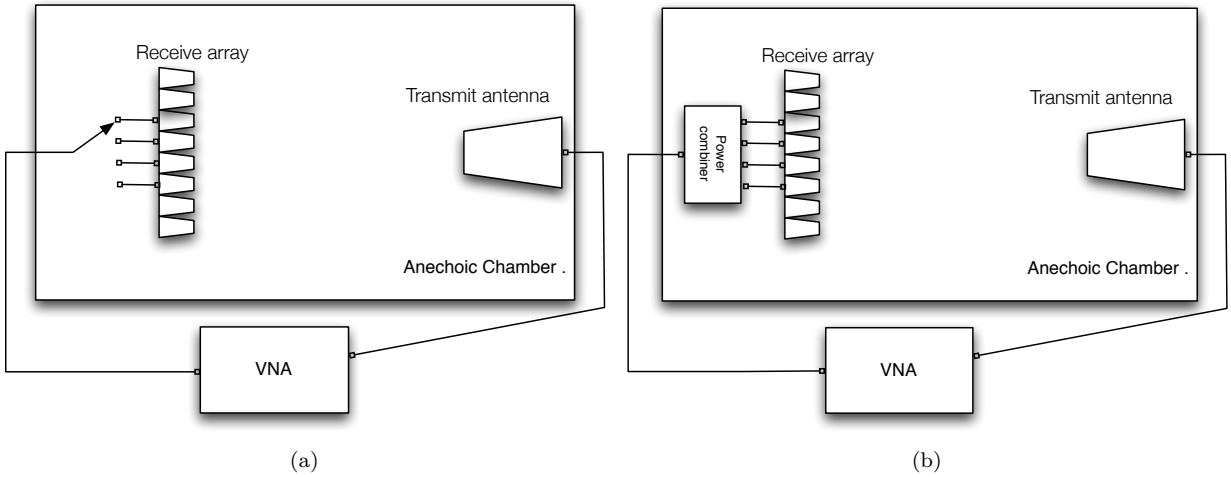


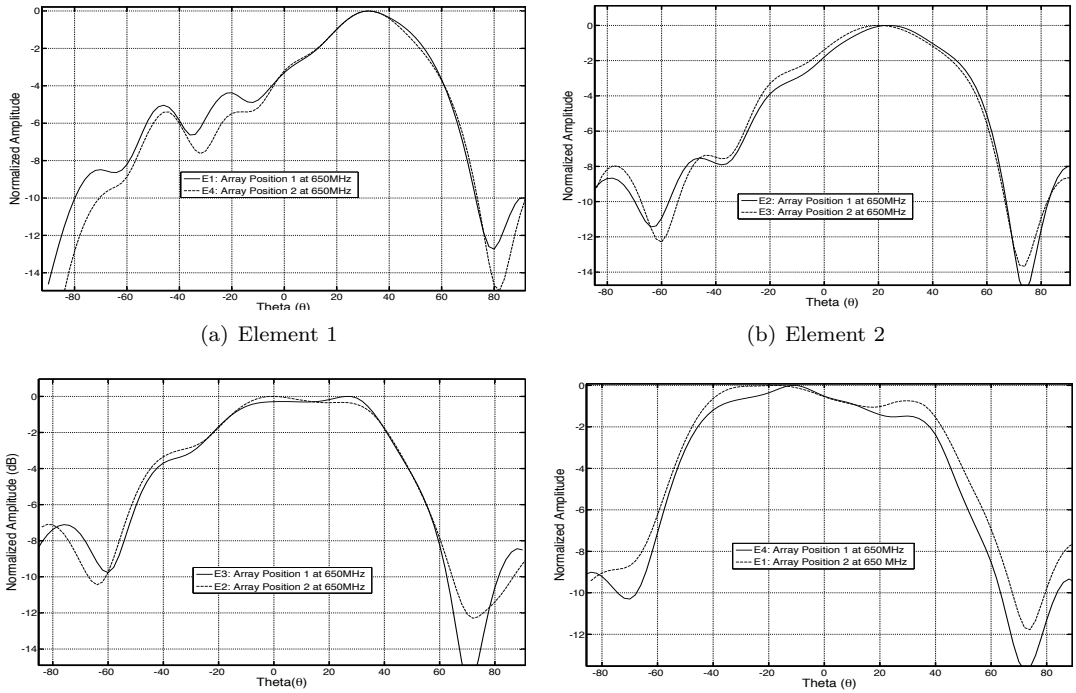
Figure 5.8: A schematic showing the vector network analyser (VNA) measurement setup. Figure 5.8(a) shows the individual pattern measurement. The VNA was required since a phase-coherent measurement is required for complex-voltage signals. Figure 5.8(b) shows the analogue power-combiner beam-summation setup. Element signals are not able to be phase-delayed in this setup, which is primarily used for verification of the digital system.

Element Pattern Verification

It is expected that while the element patterns may not match the predicted (simulated) ideal immersed element pattern of Zhang and Brown, they should, for such a physically-symmetrical array, be approximately symmetrical. Figure 5.9 shows that this is not the case at the observation frequencies (i.e. not the design frequency); neither of the element pairs, $E(2,3)$ nor $E(1,4)$ are mirror images of each other in the general broadband case. Element patterns match the response-symmetry criteria more closely at the design frequency (see Figure 5.7).

Therefore, in order to verify that the element patterns do actually belong to the array itself, rather

than the greater electromagnetic environment, measurements were performed first with the array in a given broadside orientation, and then with the array rotated by half a revolution in a plane normal to the transmit antenna. Element patterns are expected to remain with the specific elements, rather than with their 'positions.' Figure 5.9 shows each of the element patterns measured in the two positions of the array. Table 5.4.2 summarises the maximum deviation from perfect pattern match. It is evident that by far the most significant determinant of the element pattern is the element itself, rather than the (expected anechoic) electromagnetic environment. This test is a confirmation of the anechoic nature of the test environment.



(a) Element 1

(b) Element 2

(c) Element 3

(d) Element 4

Figure 5.9: Flipped and fiducial element pattern comparison. This measurement confirms that the chamber is electromagnetically uniform to good approximation, since element patterns followed closely the flip direction. These patterns deviate strongly from the simulated design-frequency (882 MHz) central-element pattern of Figure 5.4, hence an aspect of the requirement for immersed element optimisation. Table 5.4.2 lists the maximum elemental deviations over a 45° scan angle.

5.3.5 Digital Beamformer Configuration

The digital beamformer measurement configuration is shown in Figure 5.10. Each 50Ω antenna balun is connected to an input port of the ADC. Python-based control software on a central PC is used to control the

Element	Maximum elemental pattern deviation in ±45° scan volume	Scan angle (degrees)
1	1.286dB	-25
2	0.582dB	-18
3	0.337dB	26
4	1.382dB	32

Table 5.1: Maximum deviation of elemental patterns for a ±45° scan volume when the array is rotated on a plane normal to the transmit antenna.

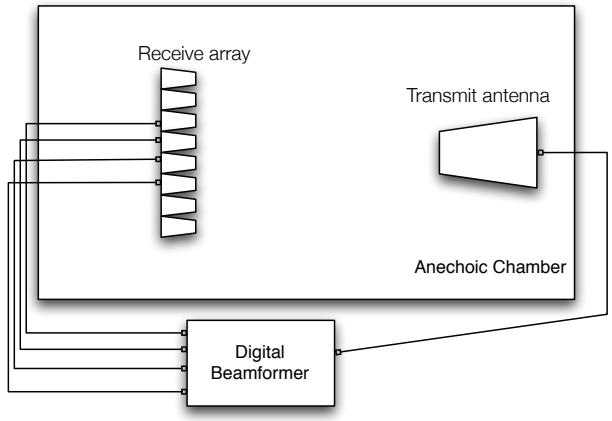


Figure 5.10: A schematic of the digital beamformer measurement configuration

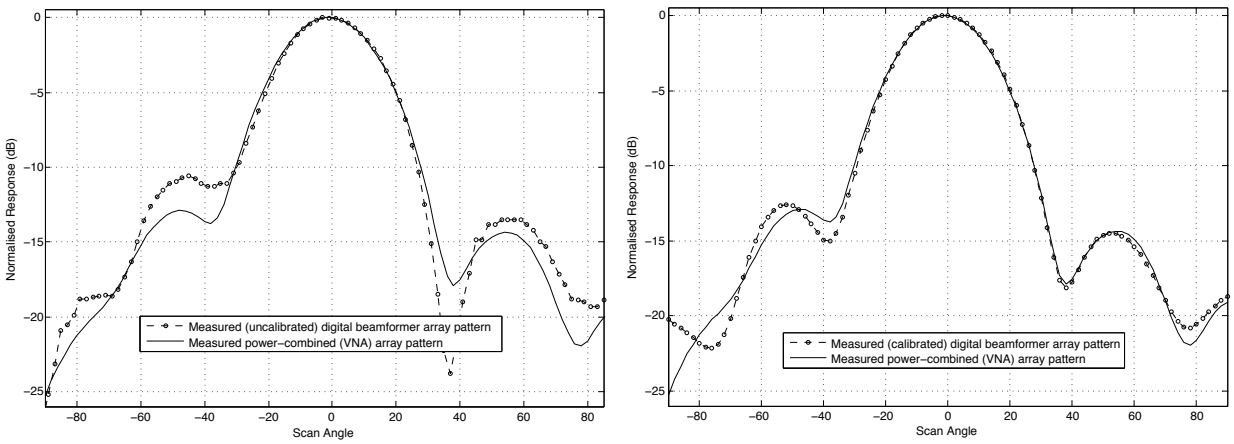
anechoic chamber stage, as well as read snapshots of the digital beam spectra as they are output from the IBOB. The transmit antenna was connected to a signal generator, which outputs a CW 0.65GHz sinusoidal waveform.

Current excitations are applied to each of the four signal paths by the digital beamformer. Excitations are applied as a tuple composed of a 16-bit real and 16-bit imaginary pair representing the complex excitation current determined by the PSO algorithm.

5.4 Application of Optimal Current Excitations

Phase Calibration

There remain small phase differences between the radio frequency (RF) path of each of the elements. In order to achieve correct gain and phase between the elements of the digital beamformer, a Gaussian mean calibration is performed. This procedure compensates for the ensemble average the phase difference. It



(a) (b)

Figure 5.11: Phase calibration of the digital beamformer. These graphs show the comparison between a Vector Network Analyser (VNA) system (configuration is shown in Figure 5.8(b)) and the digital beamformer system (configuration is shown in Figure 5.10) before (fig. 5.11(a)) and after (fig. 5.11(a)) phase calibration.

includes effects due to non-ideal match between baluns and ADC, environmental mismatch (non-ideal anechoic environment), and cable length and permittivity mismatch. The calibration procedure is similar to that described in Section 4.3.1.

This procedure produces a sinusoidal power curve as a function of phase angle for each antenna pair. The phase angle at maximum ensemble average power (the peak of the sinusoid) is then used as the calibration mean. Figure 5.11 shows the effect on the array pattern of the calibration procedure compared of the power-combined VNA measurement.

5.4.1 Obtaining Current Excitations

Excitation coefficients were determined for four beam-shaping methods; Chebychev and Taylor current distributions, as well as two current distributions from of the hybrid PSO method with different optimisation targets. This was done to be able to compare the PSO method. The two analytic current distributions were obtained with standard methods (see e.g. Elliott, 2003) in MATLAB®.

Heuristic current excitations were obtained by delivering the (complex) Schelkunoff form of the element responses to the MATLAB-based algorithm with a particular set of constraints for each case. The first set of coefficients attempt to conserve beam width (this is not a feature of the analytic weighting functions) while converging on a response with side-lobes less than -20dB. The second set relax slightly the beam-width

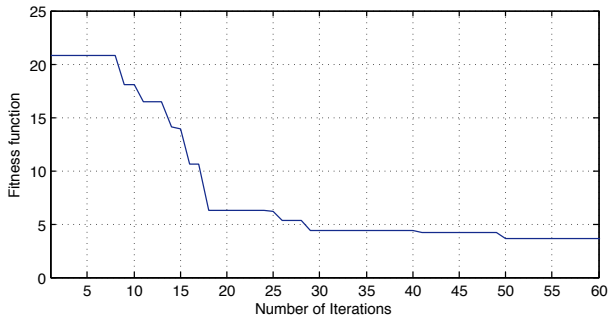


Figure 5.12: Algorithm convergence curve. The plot shows the best beam-shape fitness (the best result of beam divergence compared to the template beam shape) of the PSO algorithm as a function of algorithm iteration. This particular run of the algorithm, which attempts to achieve -20 dB side-lobes and a narrower main beam, converges at 50 iterations.

requirement while aiming for side-lobes less than -25dB. Figure 5.12 demonstrates convergence of the PSO algorithm. The plot shows the fitness function evaluation, which evaluates the 'nearness' of the best PSO solution to the template as a function of iteration.

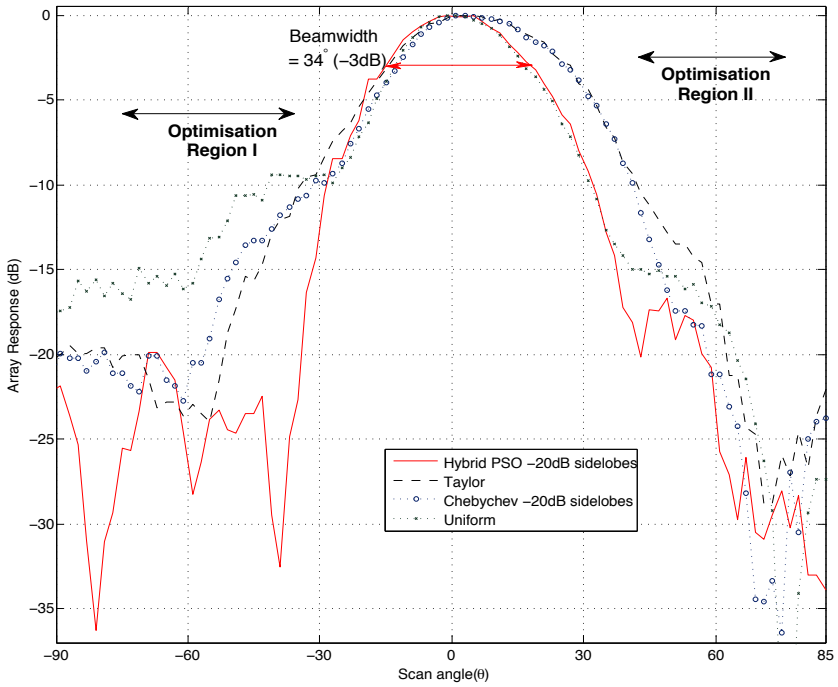
5.4.2 Application of Current Excitations

This optimisation procedure is not a closed loop. That is, measured antenna patterns are used to *predict* an optimised beam response. Once predicted excitations are obtained, they are verified by uploading to the beamformer and measuring the response.

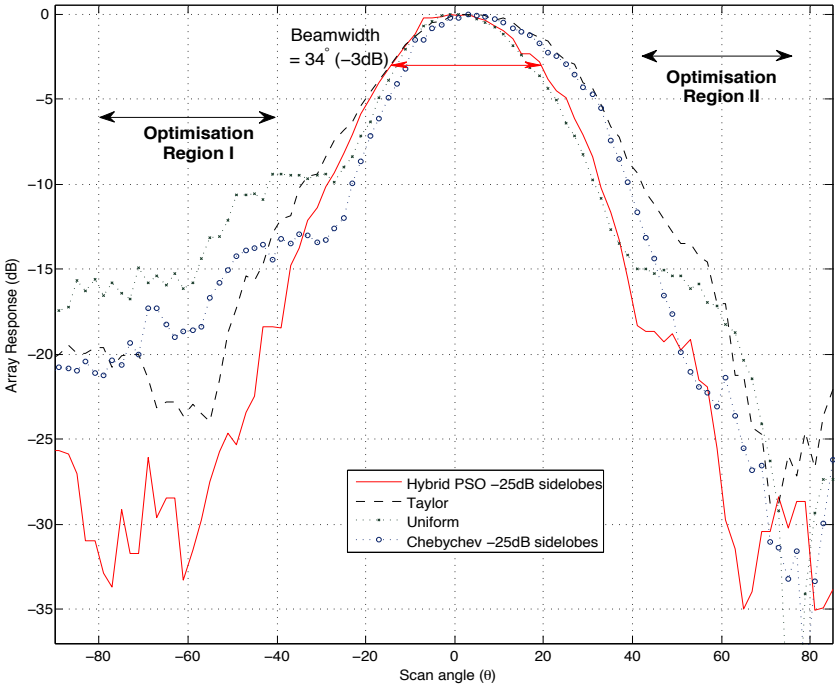
A comparison of resultant beam responses is shown in Figure 5.13. In the first subfigure, 5.13(a), the hybrid PSO beam shape is compared to the unweighted (uniform) case, as well as to responses due to Chebychev and Taylor weightings. The -3 dB beamwidth is maintained compared to the unweighted case, and improved for even greater scan angles. The PSO solution is not smooth, likely due to the greater numerical instability of the excitation set.

The second subfigure, 5.13(b), shows the optimisation attempt to obtain the smallest *side-lobes* possible, again compared to the other weighting methods. Beam narrowness is sacrificed for the sidelobe constraint.

The central result shown by the figures is that heuristic current distribution optimisation with the particle swarm optimisation method produces narrower beamshapes with lower side-lobes than both Taylor and Chebychev weightings.



(a)



(b)

Figure 5.13: A comparative plot of the resultant beam responses. Subfigure 5.13(a) shows the hybrid PSO beam shape compared to the unweighted case, as well as to responses due to Chebychev and Taylor weightings. The -3 dB beamwidth is maintained compared to the unweighted case, and improved at scan angles outside of the half-power response. The second subfigure, 5.13(b), shows the result of the optimisation scheme to obtain the smallest *side-lobes* possible, again compared to the other weighting methods. Beam narrowness in the second optimisation is sacrificed for the minimal sidelobe constraint.

Current Excitation	-3dB Beamwidth	Region I Minimum Sidelobe Level (dB)	Region II Minimum Sidelobe Level (dB)
-20dB Chebychev	40	-13.54	-17.41
-20dB Hybrid PSO	34	-19.87	-16.69
-25dB Hybrid PSO	34	-26	-19.15

Table 5.2: Minimum sidelobe level (SLL) and beamwidth are compared for the three sets of tapering weights. The corresponding patterns are shown in Figure 5.13

5.4.3 Concluding Remarks on Small-Array Optimisation

A digital beamformer was designed and used in conjunction with the Particle Swarm Optimisation (PSO) algorithm to improve the array response of small, uniform linear arrays (ULAs). Traditional current excitation approaches, such as Dolph-Chebyshev and Taylor synthesis do not perform optimally for small arrays, especially in case of anisotropic array elements.

Aperture arrays for the SKA will built hierarchically, from many small processing groups of array elements. It is worthwhile to both make uniform and improve the responses of the component small arrays. Heuristic optimisation provides a solution; non-isotropic element patterns of individual antennas may be combined in such a way as to produce a uniform beam shape from each small array.

Chapter 6

Signal Processing for Digital Radio Astronomy Instruments

One of the outstanding challenges for digital aperture arrays for the Square Kilometre Array is the provision of sufficient digital signal processing capacity. This task — of processing tens of thousands of data streams each with bandwidths of the order of 10^9 Sa.s^{-1}, as would be required of a large aperture array beamformer — is significant. The empirically-observed exponential increase in the size and efficiency and concurrent decrease in cost of complementary metal-oxide-semiconductor (CMOS) technology, known as Moore's Law, will not obviate the solution to the processing problem.

There are several central considerations in the design of a digital processing system for a station phased array. The first — and often the sole attribute which is considered — is the provision of sufficient computational resource. Evidently a processor should have adequate *total* processing capacity. Yet questions of the division of processing into some *hierarchical* structure are also important; it will not be possible to perform all processing operations in a single processor, and thus division of the system into logical processing groups is important. Second is the issue of data logistics: efficient data scheduling is required for high-capacity use of available computational resource. Furthermore, the *type* and *length* of data links are decisive factors in both cost and power consumption. The third consideration is that of adequate system flexibility. The SKA is a multi-purpose instrument with many and possibly conflicting observational modes. Thus, some degree of plasticity in the processing system is desirable and possibly necessary. Last are questions on energy; is it

possible to estimate accurately the expected (and minimum) energy consumption of a central processor and what are the implications for feasibility of the digital aperture array technology if these energy efficiency requirements are not met?

A qualitative discussion of these central issues forms the first half of this chapter. In the second half, a quantitative framework is developed and is used to assess the processing and energy issues for the SKA and array-based digital radio instruments in general. Particularly, an economic model of energy consumption for processing and data transport places limits on the the choice among the full spectrum of processing architectures to only the most energy-efficient.

6.1 Processing Capacity

The signal processing requirements for a digital aperture array system are large. In simple terms, the processing of some hundreds of thousands of elements per station, each with digital bandwidths of order 10^9 Hz is required. In order to illustrate the approximate extent of this requirement, a fiducial processing model is adopted. The model consists of a Fast Fourier Transform (FFT)-based filterbank (of $N_{\text{chans}} = 2^{10}$ channels) and beamformer (of $N_{\text{beams}} = 2^6$ spatial inputs/outputs), data rate[1] DR $= 10^9$ Sa.s^{-1}, spectral frame rate SFR[2], and FFT efficiency coefficient twice the maximum FFT efficiency achieved by Frigo and Johnson, $C = 2 \times \frac{34}{9}$ (Frigo and Johnson, 2005) set. Other models are possible (for e.g., see Chapter 4), and the $O(N \log N)$ algorithmic efficiency of this model (compared to the algorithmic efficiency of $O(N^2)$ of other models) is handled with care. Then the minimum total processing requirement per element O_{el} in complex multiply-accumulate operations per second (MACs) is estimated as

$$\begin{aligned}
\text{O}_{\text{el}} &= \text{O}_{\text{fil}} + \text{O}_{\text{chan}} + \text{O}_{\text{bf}} \\
&= 8 \cdot (\text{DR}) + C \cdot N_{\text{ch}} \cdot \log_2 N_{\text{ch}} \cdot (\text{SFR}) + C \cdot \log_2 N_{\text{beams}} \cdot (\text{DR}) \\
&= \text{DR} \cdot (8 + C(10 + 6)) \\
&= 1.29 \times 10^{11} \text{ complex MACs}
\end{aligned} \tag{6.1}$$

[1]Input data rate will be 10^9 Sa.s^{-1} per element, based on an RF bandwidth of ~ 0.5 GHz.

[2]The spectral frame rate is a measure of the degree of overlap in the spectral windows. If the SFR = DR/N_{chans}, then spectral windows are independent.

where the operation count for each of the processing stages of filtering, channelisation and spatial beamforming are O_{fil}, O_{chan}, and O_{bf}, respectively. An SKA station contains many tens of thousands of gigahertz-bandwidth receiver chains[3]. The station processing requirement, $O_{station}$ is derived from this result.

$$\begin{aligned} O_{station} &= O_{el} \cdot N_{el} \cdot N_{pols} \\ &\simeq 2.20 \times 10^{16} \text{ complex MACs} \end{aligned} \tag{6.2}$$

There are different predictions for this number. Faulkner et al. (2008) estimate the requirement for an SKA phased array station to be of order $\sim$ 10 PMACs (10^{16} MACs)[4]. All known previous estimates are within an order of magnitude of the model estimate in Equation 6.2. The difficulty of achieving such performance is highlighted by converting the estimated fixed-point signal processing metric of multiply-accumulate operations per second (MACs) into the metric of floating point operations per second (FLOPs)[5] and comparing it to the current most-powerful computers. The organisation which records the 500 most-powerful public domain supercomputers (www.top500.org) lists the most powerful machine as achieving a maximum performance of just less than 10^{16} FLOPs. This large processing requirement is the central reason that all-digital phased arrays have not previously been implemented on a large scale for radio astronomy.

6.1.1 The Exponential Increase in Processing Capacity

However, achieving this level of processing capacity appears at least possible if the expected capacity increases of future CMOS systems is considered. Following the predictions of the empirical trend known as 'Moore's law' (Mollick, 2006), the increase in transistor density as the technological era of transistor density known as the very-large-scale integration (VLSI) era transitions to the giga-scale integration (GSI)[6] era should make possible machines with adequate processing performance. This trend is by no means a guarantee, although there is convincing evidence to its continuation for at least the next 10 years (see Meuer, 2008, for

[3] An estimate for the number of receiver chains, based on a 56m×56m station array that is spatially Nyquist sampled to 0.882GHz, is $N_{elements} \sim 85,000$ dual polarisation elements (34.5744 elements per m^2, $\pi \cdot 28^2$ m^2 = 85155 elements)

[4] However, it is noted that Faulkner et al. (2010) revise this number to be 1.25×10^{15} complex MACs without divulging details of the calculation.

[5] It is expressly not claimed that fixed-point multiply-accumulate operations are equivalent to floating point operations. The differences in datapath design, architecture and numerical dynamic-range of the two processors mean that they are not comparable. The justification for the conversion is that signal processors of this scale do not exist; the SKA aperture array will push the boundaries of signal processors. The comparison is done in a strictly limited way and only to illustrate the order of the processing problem.

[6] The term giga-scale integration (GSI) refers to the next generation of silicon mega-circuits, which will be a *qualitative* advancement in size, energy-density and transistor scaling than the previous generation of very-large-scale integration (VLSI).

a document of the convincing empirical trend of exponential increase in digital processing capacity, as well as network, storage and reconfigurable logic capacity). Reliance on Moore's Law is not unique to aperture array signal processing. Indeed, the SKA relies on this increase in processing capacity. Machines with total processing capacity of 10 petaflops (10^{16} floating-point operations per second) are predicted by the beginning of 2012, and exaflop (10^{18} flops) machines by 2016 (see www.top500.org). Furthermore, the risk of reliance on a processor performance increase is not taken by the SKA alone; entire industries and scientific projects significantly larger than the SKA rely on the processing increases of Moore's law.

Thus, even public-domain supercomputers are expected to be able to provide the requisite processing capacity in time for the SKA. Slowing factors to Moore's Law, such as approaching physical limits of transistor feature miniaturisation, the unwieldy complexity of monolithic processors and manufacturing chip size limitations are compensated, respectively, by advanced lithography techniques enabling feature sizes down to 11 nm (see SIA, 2011, and www.itrs.net/), the introduction and broad industry acceptance of moderately and massively parallel on-chip processors and accelerators, as well as the emergence of multi-billion-dollar fabrication facilities.

In summary, the expected exponential processor performance increases with time should mean that processing capacity for an all-digital phased array is possible in some digital form in the near future.

6.1.2 Hierarchical Processing Groups

Beyond provision of total processing capacity, there is also the question of processing hierarchy. Given even the most optimistic projections of Moore's law, it will not be possible to process all the elements in a station in a *single* processor. There is no technology path for a future monolithic system which combines an I/O bandwidth of $\sim 8.5 \times 10^{14}$ bps with processing capacity of $\sim 10^{17}$ MACs. That is, it is not possible to process all elements on a single processor; some form of system hierarchy is required. A 'tile' is defined as a processing subsystem consisting of the number of elements that may be processed by a single monolithic processor at the first level of this hierarchy.

The tile as a cohesive processing unit is an important distinction since it will be the atomic unit of the aperture array instrument. A station is at the next processing level and will contain physically-adjacent tile arrays, combined hierarchically with 'tile-like' processing structures.

6.1.3 Taxonomy of Signal Processing Architectures

There are many processing architectures which are potentially usable for digital signal processing. For instance, radio astronomy signal processing has been performed on general-purpose CPU clusters (Deller et al., 2007) , commercial supercomputers (Mol and Romein, 2011), custom single-purpose hardware (Dewdney and Carlson, 2000), and FPGA hardware (Parsons et al., 2008; McMahon et al., 2007) . These architectures may be classified by the size of the smallest atomic processing unit; their *processing unit granularity*. At one end of the scale are custom or semi-custom logic designs. At the other end are large clusters of fully-programmable central processing unit (CPU)s, such as x86-series processors. The set of architectures may be visualised on a 'spectrum' of granularity. Figure 6.1 shows some common signal processing architectures organised on a atomic processor complexity scale. Viewed as such, there is also a trend toward programmability and programming-language-abstraction as the grain-size is increased.

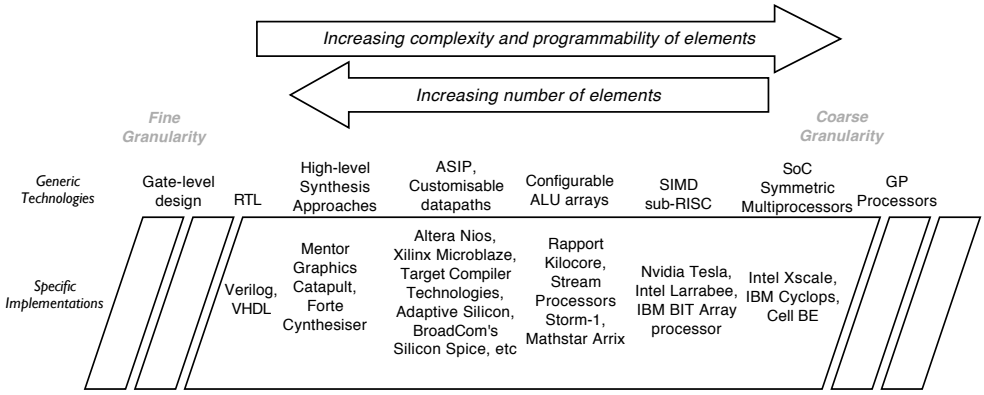


Figure 6.1: Schematic showing types of signal processing systems organised by the smallest atomic unit of computation. Within the architectural space of signal processing systems, there is scope for great variation in the parameters of the basic building blocks. Indeed, all architectures, from gate-level and ASIC designs to complex scalar General-Purpose Processors, may be seen fit within a spectrum of processing building blocks. The figure shows the general technological classes above the spectrum and examples of specific technologies that fall into these categories. Actual implementations fall onto discrete points on what can be viewed as a continuous space. Discontinuities exist in the space due to technological limitations. The spectrum begins on left side of the figure with designs at the gate-level (these are the most basic of building blocks in VLSI chips). The right side of the figure represents the opposite extreme of processing granularity where General Purpose Processors (GPP) are used as the building blocks of larger systems. Between these two extremes is a spectrum of system building blocks of varying size and complexity. With this unified view in mind, potential systems for signal processing are examined.

6.2 Data Logistics

An important concern in addition to processing performance is the dataflow and Input/Output (I/O) communication requirements of signal processing operations. It is essential that the inter-chip and on-chip bandwidth requirements of the processing are met. The first-stage signal processing requires a high ratio of I/O to processing capacity to be able to cope with gigahertz-bandwidth input signals.

6.2.1 Computation-to-I/O Ratio

One of the distinctions that can be made between signal processing tasks is by taking the ratio of the number of processing operations that must be performed on each data sample to the sample rate. This is the compute-to-I/O ratio of the processing task which is a decisive factor in the architectural structure of such processors. A processing task is considered compute bound if arbitrary data are transferred rarely compared to the number of computations performed on them (Stimler, 1969).

For instance, there are tasks that require many millions of operations to be performed on each data sample, that are well suited to systems where inter-processor connectivity is relatively low. An example of such a processing task and a distributed solution in the field of radio astronomy is the SETI@home project (Werthimer et al., 2001), where small amounts of astronomical data are transmitted to volunteer machines and a large number of operations is performed on each data fragment.

However, in the domain of spectrometer and beamformer processing, the number of operations that are required to be performed per sample is relatively low. This requires tight coupling between data transmission and processing systems in order to make efficient use of processing resources. It also requires a much larger I/O bandwidth than typical of the largest supercomputers. The type of machine desired is essentially a 'super-signal processor'.

6.2.2 Reducing Intra-Chip Interconnect

A further consideration is that on-chip communication is inherently more efficient than inter-chip communication. If the number of elements to process in one group (i.e. the tile granularity) is required to be larger than a single chip could process, a large amount of inter-chip communication would be required, consuming more power. Inter-chip communication requirements reduce with the processing 'tile' granularity. A model of inter-processor connections is developed below.

Interconnection Model

An hierarchical matrix-vector beamforming scheme is taken as the fiducial model. The required processor interconnect model has a large, deterministic communication load, but requires relatively few uni-directional interconnections; an appropriate interconnection architecture is a minimum spanning tree. The interconnection architecture is shown in Figure 6.2.

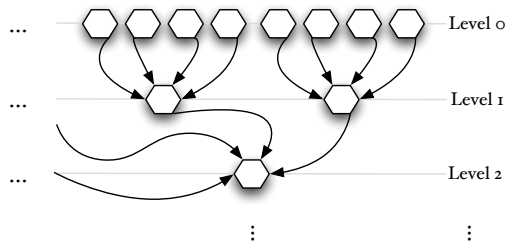


Figure 6.2: Schematic showing the hierarchical processing interconnect model. Nodes are connected in a minimum-spanning-tree. The number of physical interconnections between nodal levels, N_h (arbitrarily designated as 4 in the figure) is determined by processor bandwidth constraints.

The total number of inter-processor interconnections is

$$N_i = N_P \cdot N_b \cdot \alpha_p \cdot \log_{N_h}(N_P) \tag{6.3}$$

where N_h is the number of interconnections between nodal levels, and N_P is the total number of processors in the first level of the hierarchy. N_b is the number of beams, and α_p is the processing bit growth ratio. The product of N_i and the single-link bandwidth B_s determines the total interconnect bandwidth of a hierarchical beamformer array. The single-link bandwidth is the width of the band transferred to the first-stage processor, even if this is smaller than the elemental bandwidth. This model will be returned to later in order to perform energy consumption analysis.

6.3 Flexibility of Operation

The scientific goals of the SKA dictate that a number of different surveys be performed, each requiring different parameters to be optimised. For instance, a galaxy survey requires high-resolution beams over a limited FoV, while transient detection surveys may preferentially require low resolution, all-sky monitoring.

Flexibility is required in the operation of the instrument to be able to cater for as many scientific surveys as possible. Furthermore, *all* possible operational modes of the SKA are not yet known and may not be known for some time, perhaps long after the instrument has been built. To give a concrete example, commensality of transient-monitoring programs with survey science has become a requirement, which was not explicitly anticipated when the original SKA specifications were developed. It is desirable that a signal processing system will be able to adapt to new and evolving specification requirements. In other words, it is desirable that the processing system of an aperture array be programmable, and that the system have some plasticity.

6.4 Energy Considerations

It is difficult to imagine it ever being contentious to state that energy-efficiency is a desirable attribute of *any* processing system. Energy reduction carries with it a reduction in both running costs and fixed costs to provide the power generation capacity. Energy efficiency is especially important for the remote digital signal processing centres of SKA stations precisely because of their remoteness yet also because there is a limit on the running and maintenance cost of the instrument to be of not significantly greater than $\sim 10\%$ of capital cost per annum.

6.4.1 Feasibility of Digital Aperture Arrays Due to Cost of Energy

Is the cost of energy of a digital aperture array prohibitive? A simple economic model can provide some constraints on the processing power consumption. Running costs for the entire 250-station, $6.16 \times 10^5 m^2$ aperture array must be $\lesssim$ 37.5 M€ per year[7]; this requirement is based on running costs per year not exceeding $\sim 10\%$ of the instrument capital costs (e.g. Bolton et al., 2009). This means that the total station running costs are 0.15 M€ per year, that the running cost per m^2 of aperture array is ~ 61 €.yr^{-1}, and that a single-polarisation element is allocated just 0.88 €.yr^{-1} in running costs.

In the following, feasibility is considered from the perspective of energy consumed per processing element chain per year. Energy supplied at the average of the site rates[8] costs $\xi = 0.049$ €.kWh^{-1}. A single-polarisation element consumes $\chi_{\rm el} = 8.76$ kWh.yr^{-1}.W^{-1} of energy per year per unit of power. This would

[7] All costs are derived in 2008 €s

[8] Industry electrical energy cost in South Africa is 0.0248 USD kW^{-1}h^{-1} in 2007. Energy cost in Australian is 0.0827 USD kW^{-1}h^{-1} in 2006. No data are available for dates beyond this. The average, projected to 2008 is 0.074 USD kW^{-1}h^{-1} (see IEA, 2011, for cost data and projection methodologies)

contribute a cost C_{station} of

$$\begin{aligned} C_{\text{station}} &= \xi \cdot \chi_{\text{el}} \cdot N_{\text{el}} \cdot N_{\text{pol}} \\ &= 7.31 \times 10^4 \text{ €.yr}^{-1}.\text{W}^{-1} \end{aligned} \tag{6.4}$$

to the running costs of the entire array, or approximately 49% of the entire maintenance budget per Watt of power consumed per element. This calculation includes the cost of data-centre infrastructure, particularly that of cooling. World-averaged data-centre analysis shows that every € spent on computing resource is at least matched by that spent on cooling and infrastructure, resulting in a *more-than-doubling* of the total energy compared to processing alone (see Koomey, 2008, and references therein). Let it be assumed for this argument that a generous 50% of the maintenance budget may be spent on energy. This means that each element is allocated approximately 1W for all processing including infrastructure, data transport[9] and cooling.

Table 6.1: A list of the 10 most energy-efficient supercomputers as at June 2011. At the top of the list is an updated version of the IBM Blue Gene/Q, which is able to achieve an MFLOPS/W rating of 2,097. This makes the IBM Blue Gene/Q the first *supercomputer* to surpass 2,000 MFLOPS/W. The fastest supercomputer in the world, the K supercomputer from RIKEN in Japan, also happens to be one of the greenest supercomputers in the world; it has a ranking of 6th on the Green500. Interestingly, 70% of the 20 greenest supercomputers are accelerator-based. However, average performance efficiency for non-accelerated supercomputers is 67.7%, while GPU-accelerated machines have an average performance efficiency of 47%. All data are from www.green500.org/

Rank	MFLOPS/Watt	Power (kW)	Computer Description
1	2097.191697	40.95	NNSA/SC Blue Gene/Q Prototype 2
2	1684.201031	38.8	NNSA/SC Blue Gene/Q Prototype 1
3	1375.876168	34.24	DEGIMA Cluster, Intel i5, ATI GPU, Infiniband QDR
4	958.353433	1243.8	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU
5	891.875	160	iDataPlex DX360M3, Xeon 2.4, nVidia GPU, Infiniband
6	824.5643811	9898.56	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect
7	773.3750434	57.54	QPACE Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus
8	773.3750434	57.54	QPACE Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus
9	773.3750434	57.54	QPACE Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus
10	718.1286331	416.7777	Supermicro Cluster, Opteron 2.1 GHz, ATI GPU, Infiniband

[9]It is noted that Koomey (2008) find that total communication and networking costs are not a dominant contributor to total energy cost of datacentres worldwide. However, data transport costs become significant in those systems where relatively few operations are performed per sample. This is also the case in systems where processing and cooling are particularly efficient. This is shown later in this Chapter where dsata transport energy costs are modeled in Section 6.7.

6.4.2 A Top-Down Approach: General-Purpose Supercomputers for Aperture Array Signal Processing

Summarising briefly to the economic analysis, each element requires 1.29×10^{11} complex multiply-accumulate operations to be performed per second (see Equation 6.1) all for 1W. When converted to the metric of IEEE single-precision FLOPS per unit of energy[10] amounts to 516 GFLOPS.W^{-1}.

Current Supercomputer Energy Efficiency

A list is maintained of the most energy-efficient supercomputers (detailed in Sharma et al. 2006) at www.green500.org/. The 10 most-efficient supercomputers from this list as at June 2011 are shown in Table 6.1. As of June 2011, the most efficient of the world's supercomputers achieve 2.1×10^9 FLOPS.W^{-1} (see www.green500.org/). This means that even the most energy-efficient supercomputer, even at maximum efficiency, is currently approximately 250 times too energy-inefficient for the needs of SKA aperture array processing. Moore's law certainly predicts energy-efficiency increases in each CMOS feature-size iteration, but these are of the order of 30% per generation, and at this rate will not come close to the computation efficiency required.

The essential conclusion is that the most efficient general-purpose supercomputers are not feasible for aperture-array signal processing; energy-efficient architectures must be sought, or developed. One solution is an equivalently powerful signal processing-focussed machine, which would combine the required capacity with energy-reduction techniques such as fixed-point processing, optimised instruction paths or even static data-paths in custom hardware.

It is worth commenting that similar arguments hold for the other processing operations required of the SKA. Both correlation and interferometric imaging are particularly problematic, since they are both expected to require even greater amounts of total computing capacity than beamforming processing, and are believed to be required to perform their operations in floating-point arithmetic (see Dewdney et al., 2010). The energy-cost limit derived above must drive fundamental algorithmic change in current interferometric imaging practice if the SKA is to be a reality.

[10] It is again noted that it is not claimed here that fixed-point multiply-accumulate operations are equivalent to floating point operations. The justification for the 'comparison' is that this section is attempting to show that while both powerful and fully able to perform signal processing operations, general-purpose supercomputers do not currently and will not in the future possess the required efficiency for SKA signal processing.

6.4.3 Alternative Technologies

Other architectures on the processing spectrum may not fare as badly as general purpose processors (GPPs). Theoretically, such architectures may outperform GPP supercomputers but often little empirical data are available. Furthermore, it is already evident that there are complicating factors to direct comparisons; for instance, non-accelerated supercomputers are able to achieve 44% better average utilisation fraction than accelerated supercomputers[11]. There are other issues. Some architectures must be connected to GPP processors to function, and thus their own efficiency is dependent also on the host machine architecture.

There is recent interest in the massive processing power of modern scientific-processing Graphics Processing Unit (GPU) systems. An example of the best of the new generation of systems, the Nvidia Tesla S2050, starts to challenge the most efficient general-purpose machines by achieving 4.5 GFLOPS.W^{-1} in the 28 nm feature-size technology. Nvidia map a roadmap down to to 2013 which will achieve 16 GFLOPS.W^{-1} with 22 nm GPUs. This figure is slightly misleading, since GPUs (at least currently) must be tethered to a traditional CPU system and involve energy-hungry memory transfers over custom copper interfaces. GPU-based supercomputers have comparable energy-efficiencies to CPU-only supercomputers (Sharma et al., 2006).

FPGA technology fares better. The Xilinx Virtex 7 XT 28 nm FPGA consumes a maximum of $\sim$ 100 W for a maximum of 6,737 GMACS. These are next-generation devices – current generations provide closer to 1000 GMACS consuming $\sim$ 45W. Full-system FPGA efficiency is difficult to predict, since few (if any) large-scale high-performance processing installations exist. Nevertheless, due to the advantage of both programmability and relatively low power consumption, FPGAs are a viable candidate for signal processing. Detailed SKA-scaled evaluations are important when actual performance and energy figures for large FPGA data-centres become available.

Custom-designed hardware is inherently more energy-efficient than the general-purpose solutions. The use of custom hardware is tempered by economies of scale; it will incur a large initial design cost, known as non-recurrent expenditure (NRE), which must be offset over production of a large number of chips. For example, Borkar (2009) estimate mask fabrication NRE per chip to be in excess of 1 M€ for each 22 nm generation chip. Furthermore, there is a risk of design failure (i.e. if the device fails to work after NRE is invested). Custom hardware is also fixed, and allows no algorithmic changes or general use once the design

[11] According to www.green500.org/, more than two-thirds of the 20 most energy-efficient supercomputers are accelerator-based. Average performance efficiency for non-accelerated supercomputers is 67.7%, while GPU-accelerated machines have an average performance efficiency of 47% (an algorithmic efficiency $\eta_{\rm alg}$ of 1.5 and 2, respectively). This confirms the heuristic observation that massively parallel architectures face significant programming and software infrastructure issues.

is finalised. However, an efficiency argument may require its use despite these limitations.

A Supercomputer Standard?

It may be seen as a controversial choice to take the supercomputer list as fiducial, especially since large speed-ups are often reported for particular codes on systems that use accelerators of various kinds, recently quite often with GPU-based systems (e.g. Barsdell et al., 2012; Kruse and Ransom, 2012; Clark et al., 2011). While undeniably impressive, these architectures have inherent scaling limitations. Importantly, they are not guaranteed to be able to be scaled to the level required of SKA processing. As a hypothetical example, it is unlikely to be accepted as a large scale processing solution if it is found that a low-power mobile accelerator achieves a phenomenal rate of operations per joule precisely because it cannot be operated independently nor scaled. Furthermore, accelerator performance measures are incomparable if they do not include hosts and host interface power consumption and scaling limitations.

For this reason, supercomputer statistics are used as a standard in this work since these represent systems that *are* able to be scaled. Indeed, many supercomputers *do* make use of accelerators, yet few claim the large speed-ups and power economies reported for smaller systems. It is argued here that supercomputer performance figures are less likely to be skewed by misleading results due to, for example, intense hand-optimisation or poor definition of system energy boundaries if the systems are both enterprise-sized and subject to standard industry performance tests. Thus, supercomputers, including those that contain accelerators, are taken as the standard for processing capacity and energy predictions.

6.5 A Quantitative Model of Processing Energy Consumption

The cost analysis leads to the question of theoretical feasibility. Is it *technologically* possible with *any* CMOS-based architecture to meet the processing and infrastructure power budget of 1 W per element within a decade, even given the exponential efficiency predictions? In this section, a quantitative model of energy consumption in CMOS circuits is developed using the thermodynamic minimum as the basis and a composite device efficiency factor, η_d, to parameterise the circuit model. The model includes physical, material and architectural limits of the minimum possible power consumption in CMOS digital systems and is used in this and the remaining sections of this chapter to answer energy feasibility questions.

6.5.1 Physical Minimum Energy of Classical Binary Computation

Any logically irreversible manipulation of binary information is associated with a physical irreversibility, and thus requires a minimum amount of heat generation. A physical implementation of a binary operation may always be decomposed to a reversible operation which can be performed at no energy cost, followed by an erasure of data. Information erasure is bound by the laws of thermodynamics and will result in an increase in thermal energy. The statement of the fundamental connection between mathematical and physical computation is attributed to Landauer (1961).

Shannon-von Neumann-Landauer Minimum

Classical mathematical computation is a formal system composed of manipulations on logical states of an initial set of data. It is a map C from a finite set A of input states to a finite set B of output states. However, realisable computation, in contrast to the mathematical form, is bound by physical laws. An irreversible logical operation may be decomposed into a reversible operation followed by an erasure of data. If data erasure is accomplished by a classical physical system, it has an associated thermodynamic irreversibility. Any logically irreversible operation can not be realised in a thermodynamically reversible way. This means that classical binary computation requires a minimum amount of energy.

The minimum limit of energy required for a single binary transition in a physical system, E_{bit}, is always greater than the theoretical minimum energy of information erasure, E_{min}, and is given by Landauer as

$$E_{\text{bin}} \geq E_{\text{min}} = k_B T \ln 2 = 0.020 \ e\text{V} \tag{6.5}$$

where T = 333K is a typical operational temperature of a CMOS transistor and k_B is the Boltzmann constant. This is the physical minimum energy required of computation, and connects the abstract mathematical concept of classical computation[12] with its physical implementation.

[12]It is noted that this statement is true of *classical* computation. Work on quantum computation shows that the erasure of quantum information by an entangled observer may actually even decrease the thermal energy in the environment, dependent on the conditional entropy (del Rio et al., 2011).

Then, the *physical* minimum power for SKA station processing is

$$\begin{aligned} P_{\min} &= E_{\text{bit}} \cdot e \cdot 2^{w_\kappa(n)} \cdot \mathrm{O}_{\text{station}} \\ &= 0.020 \cdot 1.602 \times 10^{-19} \cdot 2^{w_\kappa(n)} \cdot 1.19 \times 10^{17} \\ &= 2^{w_\kappa(n)} \cdot 3.81276 \times 10^{-4} \text{ W} \end{aligned} \tag{6.6}$$

$w_\kappa(n)$ is a function of the word length n, in bits, of the processing pipeline. There are approximately 2^{n-2} binary transitions in an optimised, efficient n-bit Multiply and Accumulate Operation (MAC) operation (for the 8-bit case, see Costa et al., 2001)[13]. Then there are 2^n binary transitions in a *complex* MAC operation. For fixed-point (integer) processing, $w_\kappa(n) = w_{\text{fi}}(n) = n$ (Costa et al., 2001).

As a concrete example, $P_{\min} \simeq 98$ mW for a station beamformer processor which performs all the operations of Equation 6.2 with 8-bit fixed-point datapaths.

The complexity of floating-point pipelines — in particular modern general-purpose processors with out-of-order execution, vector (Single Instruction stream Multiple Data stream (SIMD)) instruction pipelines, and branch predication — requires a significantly more complex relationship for $w_\kappa(n) = w_{\text{fl}}(n)$, likely best determined empirically and not derived here. However, it is expected in general that energy consumption of floating-point arithmetic units will be greater than fixed-point pipelines.

Thus, a figure has been derived for minimum computation energy of an SKA station. This is a mildly interesting theoretical result, but so what of it? The real interesting result emerges when the minimum physical energy of *realisable* devices is calculated.

The theoretical limit is not the final word in minimum energy computation. It is clear that a real (classical) system may not violate the physical minimum energy, but it is not either guaranteed nor expected to operate at this minimum. It is no surprise that the implementation details of CMOS logic circuits mean that the energy consumed is greater than the theoretical lower limit. In particular, circuits require immunity against logical *indistinguishability*, which is not encompassed by the physical energy limit. There are also other unavoidable circuit issues which affect efficiency in realisable circuits, which are elaborated upon in the following paragraphs.

[13] As a simple check to this number, the n-bit multiply operation is composed of $n \times n$-bit addition, followed by a $2n$-bit addition with with 3 transistor gates per bit. On average, half of the bits are toggled in an arbitrary multiplication.

6.5.2 The Connection to Realisable Systems

Quite evidently, there will be sources of inefficiency beyond the theoretical minimum in realisable systems. Transistor devices which accomplish binary computation have physical minimum energy requirements demanded by structural and noise immunity factors. These insurances against error define η_d, the transistor *device efficiency* relative to the Shannon-von Neumann-Landauer (SNL) minimum in terms of energy. η_d quantifies the physical circuit energy limit for *any* 10^9-transistor CMOS chip, including custom hardware, GPUs, CPUs, and others. It is the maximum efficiency for a particular transistor process generation.

Secondly, a particular processing architecture does not devote all active transistors to computation, and will have an efficiency inversely related to the average proportion of transistors involved in non-computational tasks.The relative *efficiency of the architecture*, η_{arch}, encapsulates this property. This efficiency term differentiates architectural efficiency from transistor efficiency and should have a minimal dependency on the process generation.

Programmable devices have a further determining factor; the code or set of instructions which implements processing operations will include non-computational instructions that will introduce further inefficiencies. This is parameterised as the *algorithmic efficiency*, η_{alg}. It is equal to 1 in devices that do not execute programs.

So there are thus three classes of processor efficiency. Including these parameterisations, the final power consumption may be expressed as

$$P_{\text{actual}} = \frac{P_{\min} \cdot 2^{w(n)}}{\eta_d \cdot \eta_{arch} \cdot \eta_{alg}} \tag{6.7}$$

In order to meet the cost-of-power requirement developed above, P_{actual} must be less than 1.25×10^6 W. This in turn implies that the product of the efficiency factors $\eta_d \cdot \eta_{arch} \cdot \eta_{alg}$ must be greater than $\sim 3.125 \times 10^{-8}$. In the next section, the physical contributors to each of these factors will be developed.

6.6 Physical Circuits: a Transistor Model

The thermodynamic limit is not physically realisable in CMOS devices for a number of reasons, in three main categories. The three categories of inefficiency are static transistor effects, dynamic transistor effects and circuit implementation issues. The model is developed for the 45 nm transistor node size, which is likely

to be the most cost-effective technological node size by 2016 (SIA, 2011). This node size is able to be scaled up or down several generations within the Giga-Scale Integration (GSI) era, as will be detailed below.

6.6.1 Static Transistor Effects

In order to ensure unambiguous distinguishability between the two logical levels at the output of a CMOS transistor requires a minimum energy of twice E_{min}. It is required since the probability of detecting the correct digital level must be greater than 0.5. This is the minimum energy barrier for *distinguishable* computation, which is not often clearly defined.

Device limits are independent of the particular circuit configuration in which a transistor or an interconnect is applied. To account for particular CMOS device noise immunity — due to the quantum behavior of the field effect transistor (FET) circuit structure — requires a further multiplier of at least 2 on E_{min} (see Meindl, 1995).

Thirdly, the probability of a logical error must be $< 10^{-14}$ in a CMOS circuit in order to ensure a system failure probability of $< 10^{-1}$ in a billion-transistor circuit – as is the case in an average-sized chip. Larger switching energies result in a decrease in the probability of internal thermal noise error. This probability may be described by a Boltzmann probability distribution function

$$P_{\text{error}} = \exp(E_{\text{bit}}/kT) \tag{6.8}$$

Thus, the probability of error decreases exponentially as E_{bit}/kT increases. To meet the chip-wide requirement for error immunity, the transition energy must be increased approximately fourteen-fold to $\sim 56 k_B T \ln 2$ (see also Meindl, 1995; Mukhopadhyay et al., 2005). This is a fundamental limit for any type of CMOS FET-based architecture that requires system-wide noise immunity; i.e. essentially all modern transistor-based chips.

6.6.2 Dynamic Transistor Effects

Charge-based CMOS logic also has associated *dynamic* effects that influence device efficiency. The word dynamic is used to emphasise the finite time available for devices to switch their logical state through a *change* in charge. In order to charge a capacitor to the turn-on voltage required to drive the next gate in the device chain (i.e. to switch the next inverter) requires a certain amount of charge governed by

$Q_{min} = C \cdot V_{min}$. Parasitic capacitance, which adds to the device charge C, increases the amount of current required to drive the next transistor gate.

A minimum size transistor with load capacitance, C_t, of $\sim 4 \times 10^{-17} F$ (a typical value for the 45 nm node – SIA, 2011) requires Q_{min} of $260 \cdot e$ (where e is the electron charge) to ensure that the normalised voltage is sufficient to force a binary transition.

Thirdly, an inverter circuit consists of 3 transistors each with capacitance of at least $Q_{\min}$. This requires a total charge of $780 \cdot e$ for the inverter gate. This amount of charge results in a total device switching energy of $43680 k_B T \ln 2$.

An Independent Confirmation

As a confirmation of this model thus far, the International Technology Roadmap for Semiconductors (ITRS) at http://www.itrs.net/ predicts a switching energy of a single inverter in the 45 nm node based on empirical trend extrapolated from existing technologies to be $5 \times 10^4 (k_B T \ln 2)$, which compares closely to the $4.37 \times 10^4 (k_B T \ln 2)$ of switching energy predicted by the above transistor model. The empirical switching energy prediction is comparable to the minimum predicted by a complete model.

6.6.3 Circuit Implementation and Global Interconnect Limits

The third category of non-ideal behaviour are circuit and global interconnect issues. These arise due to transistors being components of highly-integrated large systems with a mix of global and local interconnects. Local circuit behaviour is due to the direct connection of individual gate outputs to downstream gate inputs, known as 'fan-out'. A CMOS circuit in a digital system requires enough power to drive 4 further devices. This requirement means that $\sim 2 \times 10^5 k_B T \ln 2$ of switching energy is required.

Global Chip Interconnect Limits

Global interconnect behaviour is subtler. At modern switching frequencies, global interconnect must take into account transmission line models. It has been found to be most-accurately modelled by empirical functions of average interconnect depth and length (e.g. Davis et al., 1998, 2001).

Physically-distant devices are required to be driven in digital systems; thus, interconnect capacitance becomes an issue. Davis et al. (1998) derive a stochastic wire-length distribution for the large, specific chip architectures of the 90-22 nm feature-size generation (the giga-scale integration era). This distribution

relates the number of signal I/O terminals T_{IO} to the number of gates N in a random logic circuit, based on an extension of the well-established empirical relationship (e.g. Landman and Russo, 1971; Christie, 1993) known as Rent's rule ($T_{\mathrm{IO}} = kN^p$). It has been shown by Davis et al. (1998, 2001) to be a good model for the average interconnect length in real GSI CMOS circuits, with appropriately-selected parameters for each technology generation. Furthermore, the ITRS[14] predicts that CMOS circuit interconnect will have a capacitance per metre for the 45 nm node of $2.5 \times 10^{-10} F.\mathrm{m}^{-1}$. Combining these models gives a value for the switching energy which is approximately $10^6 k_B T \ln 2$ per binary transition for a large CMOS chip.

6.7 Processing Hierarchy and Interconnect Energy

Since the processing requirements of digital aperture arrays demand large interconnect bandwidth, a full-system energy model must include energy expenditure due to data-transfer between processing units. The model would then incorporate the energy component due to the *hierarchy* of processing, and completes a minimum-energy system processing model.

In this section, a deeper analysis of the effect on total interconnect bandwidth due to processor hierarchy is undertaken. It will be shown that inter-processor bandwidth can be made a small fraction of total system energy as long as the processor node I/O is large enough. If not, processor interconnect starts to dominate total system energy. This is true even when optical interfaces are used due to the dominant energy cost of logistical transmission overhead, which is predominantly performed in CMOS logic.

There is very little research directly targeting system-level modeling of inter-processing-unit interconnect energy consumption either empirical or theoretical. Some work peripherally concerns interconnect power. For example, Patel et al. (1997) and Dally (1990) use power as a constraint for *latency* minimisation. This is in contrast to processor-specific power modeling tools, which are more abundant[15]. Intra-chip communication on billion-transistor chips is widely researched (e.g. Landman and Russo, 1971; Christie, 1993; Davis et al., 1998, 2001), and is modeled well by the wire-length distribution model (based on Rent's rule) that is incorporated into the device efficiency, η_{d} of Chapter 6.

Inter-processing-unit communication is qualitatively different to intra-chip communication, due to the substantive increase in length of interconnects and the high-frequency of data transfer. High-speed copper

[14] http://www.itrs.net/

[15] Architecture power simulation systems include, for e.g. SimplePower (Ye et al., 2000), XTREM Power and Performance Simulator for the Intel XScale Core (Contreras et al., 2007), SpecPower http://www.spec.org/power/, and others.

interconnects, for example, require that transmission line effects be considered as the transmission wavelength becomes comparable to the interconnect length.

6.7.1 Interconnection Model

This data-transfer model attempts to determine the minimum possible energy consumption of an interconnection network with known communication load. This is achieved in two ways. First, a fundamental energy model is developed. It is compared to an empirical model based on existing hardware. These will be shown to differ substantially, showing that while processing energy efficiency is close to the theoretical minimum, actual interconnect energy consumption is many orders of magnitude from the minimum.

Operational Modes of Transmission and a Unifying Model

In general, Maxwell's equations provide the basis for transmission line analysis. At very low frequencies, however, the line may be characterised only by its line impedance, composed of a line capacitance, source impedance and load impedance. As frequency increases, the line resistance, inductance and capacitance become important. The line must be modeled as a distributed *RLC* circuit, with effects of a general transmission line included.

Figure 6.3: Schematic showing the interconnect sequence in a single bidirectional interconnection line. The energy consumption model treats each of these stages as a separate component.

In section 6.2.2, a number count of processor interconnections was developed. The interconnect energy consumption is based on a model which separates components of the transfer circuit. The sequence of interconnect energy consumers is shown in Figure 6.3. The transmission line energy expenditure is just one component of the total interconnect architecture, and also one component of the energy budget.

Since buffering and arbitration circuitry is entirely CMOS-based, and much of the modulation, transmission and detection circuitry is also centrally CMOS-based, it can be treated with the minimum processing energy model. Transmission buffering, logistical, and arbitration circuitry is modeled as an additional *processing* energy cost.

6.7.2 Interconnect Energy per Bit

The interconnect energy per bit is parameterised by

$$E_i = E_r + E_t \tag{6.9}$$

where E_r and E_t are the routing and transmission-line energies respectively. Bit-wise routing energy has both a buffering cost, which concerns data movement from memory to a transmission buffer, and a logistical transmission cost, which concerns error detection/correction, encoding and transmission readiness operations.

The bit-wise transmission-line energy cost, E_t, is governed by the physics of the transmission media. Entropy arguments dictate the minimum physical energy irrespective of the physical information-carrying medium (i.e. electrons or photons), and distributed transmission line equations derive the interactions of the medium through which the information-carriers travel.

6.7.3 Physical Limits to Data Transfer Energy

Once again, the starting point of a minimum-energy model is the physical limit. This time the limit is for information transfer. Shannon's communication theorem states that the maximum capacity of a communication channel is dependent on the ratio of signal power to the Johnson thermal noise power delivered to a matched load (Shannon, 1949):

$$\begin{aligned} C &\leq B \log_2 (1 + \mathrm{SNR}) \\ &\leq B \log_2 \left(1 + \frac{P_s}{kTB}\right) \end{aligned} \tag{6.10}$$

where C is the maximum channel capacity measured in bits.s^{-1}. The snrSNR is composed of P_s, the average signal power of the input, and kTB, the Johnson-Nyquist thermal noise power delivered to a matched load (Johnson, 1927; Nyquist, 1928b; Johnson, 1928). B is the receiver bandwidth, k is Boltzmann's constant, and T is the temperature in Kelvin.

The energy of a single binary transition, E_{bit}, is the ratio of signal power to channel capacity, P_s/C. Substituting into 6.10 and solving for the binary transition energy gives

$$E_{\mathrm{bit}} = \frac{kTB}{C}\left(2^{C/B} - 1\right). \tag{6.11}$$

Davis et al. (2001) derive the minimum binary transition energy $E_{\min}$ by letting $C/B \to 0$ (i.e. the channel capacity is much larger than the bandwidth) and using the l'Hôpital-Bernoulli rule to give

$$E_{\min} = E_{\text{bit}}(C/B \to 0) = kT \ln 2. \tag{6.12}$$

The quantisation of charge sets a limit on the minimum potential difference (voltage) required for information transfer. The minimum switching potential is

$$V_{\min} = \frac{2E}{q} = \frac{2kTln2}{q} = \simeq 0.036 \text{ V}. \tag{6.13}$$

where q is the fundamental charge. Again, a theoretical minimum energy for an operation, this time the transfer of information over a finite channel, is available. It is not very useful until rooted in realisable systems. This is done below.

6.7.4 Transmission Lines

A transmission line is a general model of a current-carrying wire which treats the wire as an infinitesimal array of resistors, capacitors and inductors. Differential equations give attenuation (α) and phase (β) constants per unit length of line based on a incremental RLC circuit model.

Matched Transmission Line

Matched transmission lines ensure that all driven power is delivered to the load circuit, and that the generator (the source signal driver) will not resonate with the line. A generator (signal source) is designed to drive the characteristic impedance of a particular transmission line. If the load is matched with the line, the generator will see no reactive part and the length of the line will be independent of the power consumption. In a matched transmission line, the input impedance is equal to the load impedance and the characteristic impedance of the transmission line (i.e. $Z_{in} = Z_L = Z_0$). Thus, the transmission bit energy will be proportional to the bit time τ_{bit} and the square of the threshold voltage V_{thr} and inversely proportional to the characteristic impedance of the transmission line only,

$$E_t = \frac{\tau_{\text{bit}}}{Z_0} \cdot V_{\text{thr}}^2 \tag{6.14}$$

As an example, a typical value for a common high-speed copper interconnect is $100\Omega \pm 5\Omega$ (e.g. ITO, 2007), where each of 4 differential pairs is operated at 3.125 Gbps. If the threshold voltage is in the range $1 \lesssim V_{\text{thr}} \lesssim 2$ and assumed to be approximately $\sqrt{2}$ V, then E_t is 6.4×10^{-12} J.bit^{-1}. Now, returning to Equation 6.3, and assuming a fan-in hierarchy $N_h = 16$ and total number of leaf-node processors $N_P = N_{els}/N_h = 85155/16 \approx 5323$, the total minimum power required due only to matched physical transmission per station beam is

$$\begin{aligned} P_{I\text{T}} &= E_t \cdot \{\text{data rate}\} \cdot N_i \\ &= E_t \cdot 0.5 \times 10^9 \cdot N_P \log_{N_h}(N_P) \\ &= 52.7 \text{ W}. \end{aligned} \tag{6.15}$$

and the total interconnect bandwidth for this example is 8.23×10^{12} bps. Total interconnect bandwidth is an exponential function of the fan-in hierarchy. Figure 6.4 shows the total internal system bandwidth as a function of the number of node connections ('fan-in').

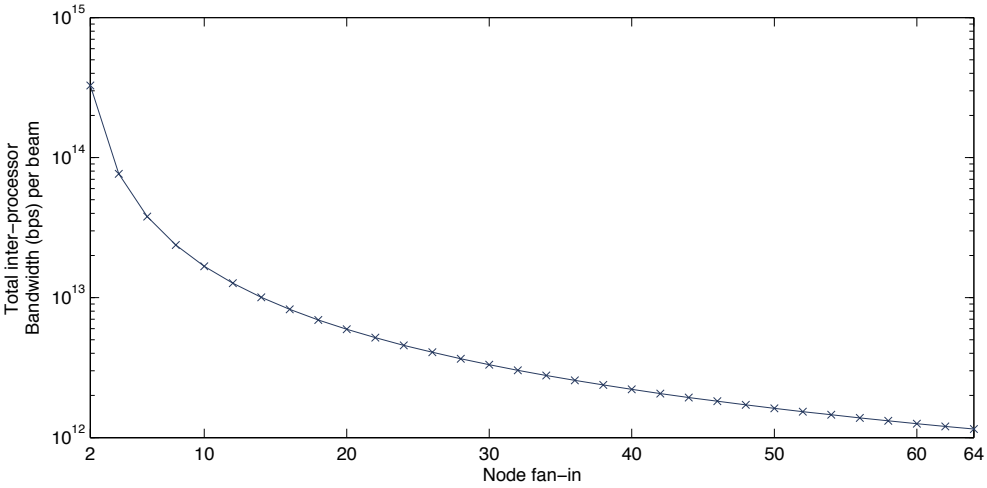


Figure 6.4: Total system bandwidth as a function of the number of connections to each node processor. Since the architecture requires conglomeration of all streams, the minimum number of such connections is 2. The difference between N_h of 2 and 32 is approximately 2 orders of magnitude in internal bandwidth requirement.

Real interconnects require immunity against attenuative loss. A typical value for the attenuation constant

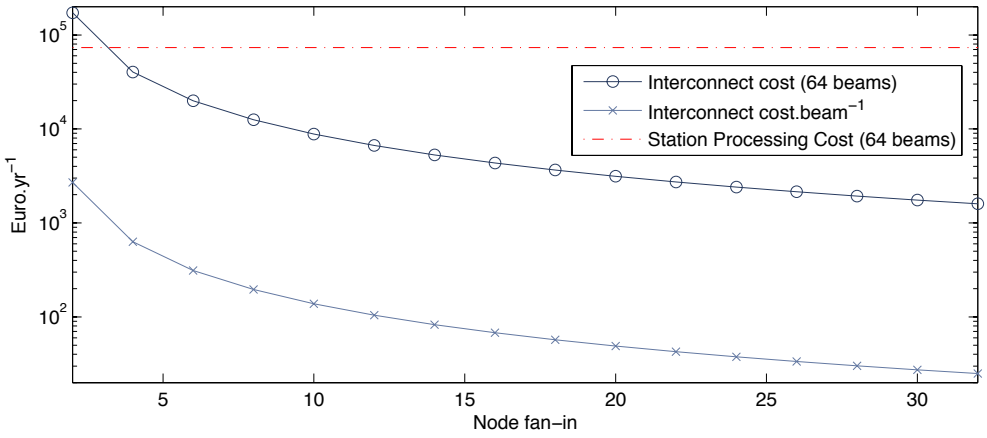


Figure 6.5: Power cost per year for interconnect as a function of the number of connections to each node processor. The power cost is that from Chapter 6. Also shown is the station cost, and the interconnect cost per beam. Interconnect cost becomes becomes a greater contributor to total cost for $N_h < 3$ and significant for $N_h \lesssim 6$

per metre α_m is $\sim 0.57\text{dB.m}^{-1}$ (ITO, 2007)[16]. Assuming an average of 5 m interconnect length for processors housed in the same central facility means that the power loss over the transmission line is ~ 3 dB. Thus, approximately twice the matched, lossless power is required for distinguishable detection in matched, lossy transmission lines.

However, this derivation includes only the interconnect energy. Transmission buffering, control, and logistical management, which are digital CMOS tasks, are not included in this energy budget. When they are, the interconnect energy, even for a lossy transmission line, may only be a fraction of the total energy cost.

6.7.5 Empirical Derivation

A complementary approach to the interconnect problem is to empirically predict the interconnect energy based on existing implementations. It was shown in the previous section how the bottom-up and top-down approaches converged for processing energy consumption (particularly, that ITRS energy predictions for 45 nm node-size transistors matched the device energy model developed above).

Single-link copper communication interface cards consume more power than that predicted by the lossy transmission line model. An industry-based figure for the copper interconnect-based Infiniband and ten-

[16] Also see www.infinibandta.org/home.

gigabit ethernet (10GbE) is E_t of 0.3 nJ.bit^{-1}(0.3 W.Gbps^{-1}) (Paduit, 2007)[17]. Current devices which, for example, include subsystems for data buffering and transfer, pre/post-emphasis, error correction strategies and other functions are $\eta_R \approx 10^5$ less efficient than the transmission-line model.

A cost analysis is developed using the figures presented in Paduit (2007). Using this model, the cost of data transmission in the hierarchical processor per beam with node fan-in of 16 is 1.03×10^3 €.yr^{-1}. This can also be parameterised by the node fan-in, as shown in Figure 6.5. It is shown that the power cost per year for interconnect becomes a significant proportion of total processing power when the node fan-in is less than approximately 6.

The analysis prompts two conclusions. First, processing node I/O bandwidth is a key component determining total energy cost. It is imperative to have a processing node with sufficient processing power and bandwidth if the signal processing is to be tractable. Second, that power spent on transmission-line portion of the interconnect is a small fraction of the total power consumed by modern interfaces, and that most of the energy consumed is in the logistical details of high-speed serial data transfer and signal buffering.

It strongly suggests that processing systems with node bandwidth less than ~ 6 the input signal bandwidth are ruled out, and the interconnect energy will become a minor contributor to total energy for node fan-in of ~ 16. Potential processing systems must match high-performance with sufficient I/O bandwidth.

Photon-Based Transmission

Optical interconnects are well-known to have lower energy consumption over long distances than copper equivalents. The use of photons rather than electrons means that the speed of propagation of a signal using optical interconnects is related only to the speed of light in the medium and its refractive index. Furthermore, capacitative effects which dominate electron-based interconnects (particularly crosstalk and capacitative loading) do not effect optical interconnects. For example, physically-demonstrated modern vertical-cavity surface-emitting lasers for optical interconnects consume of order 100 fJ/bit or < 0.1 mW.Gbps^{-1} (Moser et al., 2011), which is approximately 2 orders of magnitude lower than the theoretical minimum for copper-electron transmission.

However, it is clear from the above derivation that interconnect energy consumption is not the main contributor. The energy cost of modulation, buffering, error detection/correction, and transmission circuitry means that optical interconnects will have equivalent energy costs to copper for the relatively short distances

[17]It is noted that this is the *average* bit energy, since both 10GbE and Infiniband use several lower-frequency transmission lines. The bit time implied by τ_{bit} is not necessarily physically possible on copper interconnect.

required of inter-chip communication.[18]

6.8 Conclusions

6.8.1 Summary of Real-Circuit Minimum Energy Model

CMOS circuits may not be operated at the theoretical energy limit of $2k_BT\ln 2$ due to static issues of circuit reliability, noise immunity and distinguishability, and dynamic issues in charge-based digital logic of capacitance delay. Circuits cannot be operated at the minimum charge, Q_{min} (i.e. the *physical* quantisation of charge: that of a single electron, q) due to the requirement for drivability of downstream circuits and circuit-level parasitic capacitance. Average global interconnect capacitance is modelled with the use of a stochastic wire distribution model. These effects result in devices that require $\sim 10^6 k_BT$ per binary transition in a real CMOS circuit.

Furthermore, interconnect energy consumption is not guaranteed to be a negligible contributor to total system energy; in particular, processors need large I/O capacity ($\gtrsim 6 \cdot \{\text{instrument bandwidth}\}$) to ensure that interconnect energy consumption does not dominate.

In order to meet the power requirement of equation 6.6 requires that $\eta_{arch} \cdot \eta_{alg} \lesssim 32$. This means that it *is* theoretically possible to build a digital system of this scale in a circuit CMOS technology likely to be readily available in 2016. It also strongly suggests that highly-efficient architectures are required to ensure that the energy usage is within the available budget.

The current best figure for $\eta_{arch} \cdot \eta_{alg}$ from www.green500.org/ as at June 2011 is 1.18×10^4. General-purpose computers and supercomputers are not suited to the signal processing task.

6.8.2 Concluding Remarks

Signal processing for the Square Kilometre Array (SKA) introduces unique challenges. Total processing requirements are large. Furthermore, a decision of a potential system which may provide this capacity must take account of energy consumption, have appropriate flexibility, and manage the logistics of communication amongst hierarchies of processing subsystems.

[18] Optical detectors fabricated from silicon have a cut-off at ~ 1.1 μm which is a fundamental limiting property of the silicon material. For smaller transmission wavelengths, other compounds must be used, such as gallium arsenide (GaAs).

It has been shown through the use of a detailed device model and interconnect model that it *is* theoretically possible to build a digital system of the required scale in a circuit CMOS technology likely to be readily available in 2016 within the energy and cost budget. The analysis strongly suggests that highly-efficient architectures are required to ensure that the energy usage is within the available budget.

There is also an evident conclusion to the energy-consumption analysis that general-purpose computers and supercomputers are not suited to the signal processing task. If accepted processing estimates (e.g. Dewdney et al., 2010) are used, a similar problem exists for imaging processing and correlation processing.

Each architecture will have complex algorithmic and architectural inefficiencies. These may only be fully characterised through detailed exploration to the level of implementation. Constraining η_{arch} and η_{alg} for a novel architecture is an important contribution of this thesis, and is the subject of the following chapter.

Chapter 7

Cycle Accurate Architectural Simulation of the IBM BIT Integer Processor

The significant processing and power requirements of future digital radio telescopes demand energy-efficient yet powerful processing architectures. General purpose supercomputers will not be sufficient for this task; non-standard alternative architectures need to be considered, with particular attention being paid to power consumption per processing operation. A class of supercomputers achieve energy-efficiency through massive parallelism, shared instruction scheduling and software-controlled data-memory management. Novel programming and compilation frameworks are required to fully exploit the power of these array processor architectures. This need is addressed in this chapter.

This chapter discusses a programming language and cycle-accurate simulator for the IBM BIT array processor — a novel, multi-core array processor supercomputer that is particularly suited to streaming signal processing applications. A new cycle-accurate simulation environment, consisting of a functional language, compiler and full-architectural simulator is described. Using these tools, the architectural efficiency of the array processor is characterised. These results have been fed into the processor design; a rare example of true hardware/software co-design. The algorithmic efficiency of radio signal processing operations is characterised by implementation of a fiducial digital beamformer processing code.

7.1 The IBM BIT Integer Processor

The IBM BIT Integer Array Processor (IBM BIT) is an energy-optimised research supercomputer which combines a simplified execution pipeline, a low-order interconnection architecture and massive parallelism to achieve high computational efficiency per joule. The IBM BIT, to be built at the IBM T.J Watson Research Laboratory will be a SIMD[1] array processor consisting of 2^{10} processing elements per chip, $\sim 2^6$ chips per 'blade' (processing board) and $\sim 2^{10}$ blades in the rack array; a total of $\sim 2^{26}$ processing elements in the supercomputer.

Each individual processing element consists of two 64-bit execution units which include a multiplier, shifters, an adder/subtractor and a logical unit. Arithmetic and logical operations can be performed on 'pockets' of size 2^k, $(k \in \{6, 5, 4, 3\})$ within the 64-bit word[2].

Processing elements are supplied with microcode each cycle by an IBM Cyclops C64-based[3] sequencing and control unit. The inter-processor communication (IPC) network is similarly microcoded and operated in lock-step with the processor cores. The combined instructions for tile operations and mesh data movement forms a single long microcode word.

The IBM BIT array supercomputer is designed to process parallel data streams. Each 2^{10}-element processor array (the 'BIT array processor') is a single independent node in a systolic array. This means that the basic computational unit of the IBM BIT supercomputer is a massively parallel array processor itself. The smallest data-independent execution object is a single processor element, called a 'tile'. However, the smallest instruction-independent object is the BIT array processor. Since each BIT array processor is independent and operates on parallel data streams, simulation of the operation the 2^{10}-element processor array is the critical objective. This array processor is the object of the simulation environment.

7.1.1 Processing Elements

The architecture of the smallest complete computational element in the array is shown in Figure 7.1. Each of these is a full processor itself, but share instructions with the entire array.

[1] A classification of parallelism in computers, introduced by Flynn (1972) is based on the multiplicity of instruction and data streams (sequential flows of a particular type of object). A SIMD computer allows one set of instructions to operate on a multiplicity of data streams, which for certain codes results in high efficiency.

[2] an exception to the arithmetic 'pocketisation' is the multiplication operation, which can only be performed on byte pockets (i.e. 8-bit) and double-byte (i.e. 16-bit) pockets.

[3] The IBM Cyclops C64 is a Shared Memory multi-Processor (SMP) SoC multicore processor (see for e.g. Zhang et al., 2006; del Cuvillo et al., 2006; Chen et al., 2007). The Cyclops C64 has been used for beamforming processing previously by AhmedSaid (2009). The architecture of the IBM BIT includes a single processing core of the Cyclops C64 SMP used for instruction sequencing and communication control.

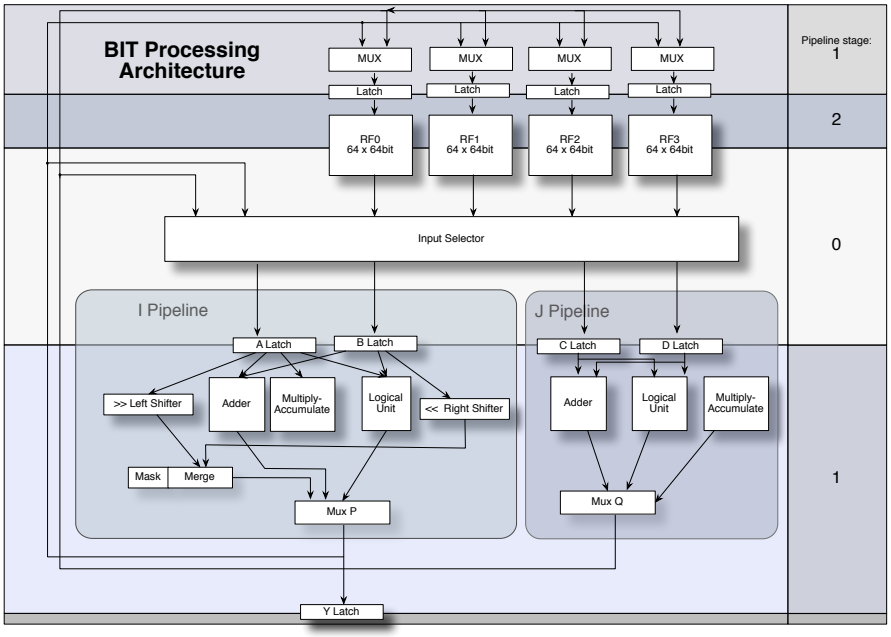


Figure 7.1: Architecture of a single Processing Element or *tile* of the BIT array processor. Processor memory is distributed in 16kb blocks, arranged as four 64×64-bit register arrays which are independently connected to the communication mesh. Total chip software-controlled cache memory is 2MB. Dual, pipelined 64-bit execution units perform arithmetical and logical operations; one of these units has the ability to also perform shifting, data mask/merge and multiply-accumulate operations. Processor pipeline stages are underlaid. A three-stage pipeline with multi-cycle processor stalling is implemented. Pipeline stages are data fetch, execute and store. Multiplication operations cause multi-cycle pipeline stalls dependent on execution unit pocketisation.

A processing element consists of four 64-bit register files that feed two execution units, the 'I' and 'J' processing pipelines. The I pipeline is fed by the first two latches (A and B) which may access any of the register arrays. The I pipeline is an integer, byte-sliced Arithmetic Logic Unit (ALU) with adder, logic unit, multiply-accumulate, left-shift, right-shift, comparator and merge units.

The J pipeline is fed instead by the second pair of latches (C and D). It is also integer, byte-sliced ALU pipeline with adder, logic, comparator and multiply-accumulate units but without shifting functionality. Any of the latches (A, B, C and D) may source data from any of the register files. However, only one register can be read from any register file concurrently.

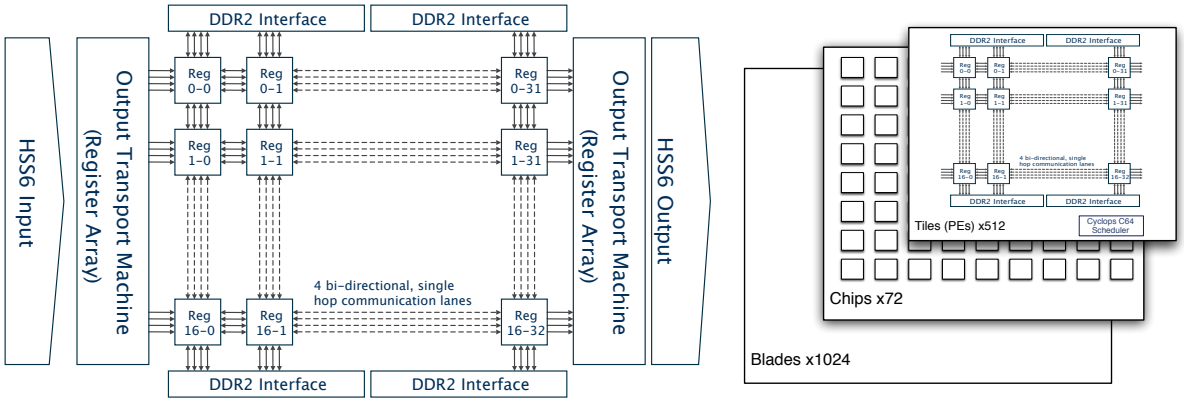


Figure 7.2: BIT Array processor arrangement at the tile (a) and system (b) level. In 7.2(a), the register-file interconnection interface is shown. Processing elements communicate through this shared memory. High-speed copper interconnects are shown at each end of the processor. These interfaces transmit 11 Gbps. Data are buffered into input and output transport machines (TM), which are register arrays of the same memory capacity as the processor array (i.e. 16Mb). 7.2(b) is a schematic showing the eventual supercomputer architecture. Each 'blade' is a processing board containing 72 IBM BIT chips, and 1024 blades compose the supercomputer.

7.1.2 Processor Interconnection Network: Low-Dimensional k-ary n-cube

In the context of processor communication, the k-ary n-cube is an interconnection model where processors are modelled as the k nodes on the edges of a hypercube and inter-processor connection occurs along edges which span the n dimensions of the hypercube. The *dimensionality* of the interconnection network is the number of connections that each node will have to all other nodes. In a k-ary n-cube, the number of processing elements, N, nodes, k, and dimension, n, are related by: $N = k^n$ (see e.g. Dally, 1990).

Communication Latency Optimisation

In order to reduce maximum communication latency, interconnection networks designed with performance as the primary constraint typically have high interconnectivity; that is, dimensionality is made as large as practical. Patel et al. (1997) find that the network dimension which achieves minimal latency for a fixed node frequency is a slowly growing function of system size. In other words, large-N systems ($> 2^{14}$ nodes) require higher dimensionality to achieve their theoretical minimum latency. This observation generally results in low-latency, highly-redundant interconnection networks. Exemplifying this philosophy is the architecture of one of the earliest massively parallel array processor families, the Thinking Machines Connection Machine (CM) -

series processors, where the CM-2 2^{14} single-bit-node processor is connected with a binary (nearest neighbour) 16-cube architecture (Tucker and Robertson, 1988). Similarly high-dimensional in its connection architecture is the Cyclops C-64, which has approximately the same number ($13\,824 \simeq 2^{14}$) of 'nodes' (although in this case the nodes are themselves SMPs with 160 tightly-coupled thread units), may be connected with dimensionality up to a binary 12-cube (or a k-ary 6-cube) architecture (see e.g. del Cuvillo et al., 2006; Chen et al., 2007). High-dimensionality of interconnection networks is characteristic of designs where latency performance is the primary optimisation objective.

Energy and Size Optimisation

In contrast to low-latency communication networks, low-dimensional k-ary n-cube connection architectures attempt to minimise the *total number of processor interconnections*. Such connection schema are shown to be *wire efficient* (which occurs at the knee of the latency-interconnect density curve), and hence physically smaller by Dally (1990).

There are advantages to low-dimensionality in terms of energy efficiency as well as size. In reducing the chip size, such connection architectures alleviate some of the energy cost of the interconnection network of high-wire-count routing systems at the expense of an increase in routing path length and an increase in the theoretical maximum communication latency. It is noted that theoretical maximum latency is calculated for maximum communication channel capacity. However, if the required signal processing codes entail inter-processor communication which is deterministic, the maximum communication channel capacity can be shown to never exceed a certain level and worst-case latency might never become problematic.

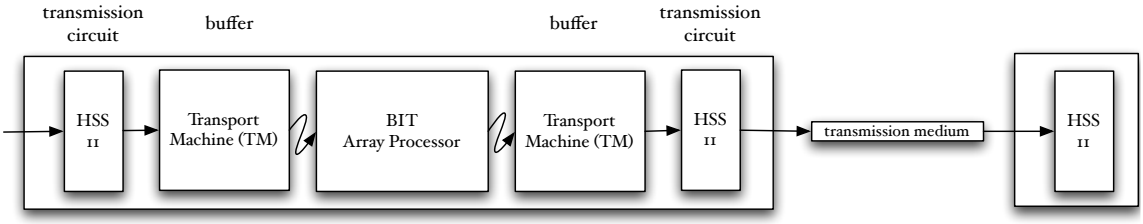
Simulation is critical to the processor connection decision; there is a clear energy advantage to the use of low-dimensional processor interconnection if it can be shown that this will not impact significantly the execution efficiency for the required codes. Specifically, if processing tasks are not communication-bound and are known not to exceed a certain communication utilisation threshold, interconnection dimensionality is able to be decreased at a small total performance cost.

Communication Infrastructure

One of the primary advantages of the IBM BIT is that communication architectures are built in to the chip itself. The schematic view of Figure 7.3 shows the standard interconnect energy consumption sequence from Chapter 6. The IBM BIT includes circuitry for modulation, buffering, emphasis, transmission and detection

on-chip, and so interconnect power is much reduced.

(a)



(b)

Figure 7.3: Schematic showing the interconnect sequence of a single bidirectional interconnection line and the integration of these components into the IBM BIT architecture. Since the IBM BIT includes circuitry for modulation, buffering, emphasis, transmission and detection on-chip, extra interconnect energy is used only by the transmission medium itself. It will be shown below that most interconnect power is consumed by components other than the transmission medium itself. A processor architecture which includes these components will have much lower communication energy costs.

BIT Array Processor Communication Infrastructure

IBM BIT processing elements are connected in a 2-dimensional 'nearest-neighbour' scheme within the chip, a variation of the 32-ary 2-cube torus with the distinction that end-elements are unconnected (or end-processing-elements are connected with a route cost of N) and the array is composed of instead 32×32-elements. At higher levels of the supercomputer, the array is structured as a systolic array of linear, independent pipes; that is, there is minimal interaction between nearest-neighbour array meshes. Figure 7.2 shows the interconnection of processors at the tile level (7.2(a)), as well as processor organisation at the system level (7.2(b)).

7.2 Cycle-Accurate Processor Simulation

Custom architectural simulators have necessarily evolved in parallel to the development of computing. Their applications include hardware development, performance profiling and architectural evaluation. Effective and representative simulation systems allow processor architectures to be evaluated before they are available in silicon. This has two principal advantages. First, it allows software co-design. This means that algorithms may be implemented in parallel to the hardware design, which reduces total system design time, but also allows evaluation of the merits of a yet-to-exist architecture. Secondly, it is a feedback path into the hardware design process, which enables the hardware to be optimised for precise tasks based on the software design and implementation analysis.

To be more concrete, IBM are currently in the process of building the BIT integer processor, but the design is still able to include small architectural changes that will increase the efficiency of codes which will eventually run on it. However, there are two issues with this scenario. The first is that there exists no final programming language. The second is that there does not yet exist the hardware upon which codes may be run. Thus, if the hardware design of the processor is to be influenced, both programming language and a system simulator must be developed.

The objectives of the architectural simulator designed for this task and described in this chapter are, firstly, to profile a suite of typical signal processing algorithms for accurate performance estimates of the IBM BIT processor, and secondly, to use these results to influence the hardware design of the processor before it is finalised. Because simulators are not subject to the same restrictions and constraints as the actual system, they are more easily modified and observed.

7.2.1 Simulator Taxonomy

Processor simulators are of two main types, differentiated by the detail and accuracy with which they model a system. There is an inherent trade-off between the level of detail that a hardware simulator models and the speed at which it runs.

On the one hand, functional simulators model a system as a ‘black-box’; that is, they produce functionally correct results for a given program (or possibly for its compiled machine-code equivalent). This simplification means that micro-architectural detail in the processor pipeline is not modelled; in some cases nor is such detailed timing information is not necessary, such as when model-based timing estimates are sufficient.

Functional simulators must, however, present all system elements that would be visible to a programmer. They are characterised by relative speed and simplicity, but rely on simplifying models for processor operation at the cost of accuracy in exact timing due to the model abstraction level.

On the other hand, cycle-accurate simulators attempt to additionally model timing on a simulated system. This requires the modeling of at least some parts of the system down to the hardware level, often to the level of a hardware description language (HDL) or register transfer language (RTL). Consequently, they are more complicated in design and implementation than functional simulators and thus take a relatively long physical time to compute a relatively short period of simulated time. The complexity of this type of simulator means that only recently and only by one group (see Yourst, 2007, and http://www.ptlsim.org/) has a modern (superscalar, out-of order execution-cycle) x86-64 processor been modelled closely enough to simulate accurate execution of an unmodified operating system.

7.3 Building the CABSim Environment

CABSim is a cycle accurate simulation environment for the IBM BIT array processor, built for the analysis in this thesis. It has three essential elements: a pipelined execution datapath, a functional language and compiler, and a cycle-accurate simulator down to the micro-architectural level. CABSim is able to compile functional microcode and then simulate the resulting low-level IBM BIT instruction set (the microassembler or symbolic microcode) as well as generating vertical microcode.

7.3.1 Pipelined Execution Cycle

It was decided to operate the pipeline as a three-stage, stalled pipe composed of *data-fetch* (0), *execute* (1) and *register write-back* (2) phases. In contrast to many other shallow pipeline architectures, instruction fetching is intrinsic to the array processor architecture and thus does not require its own pipeline stage; the Cyclops C-64 sequencer decodes microcode, which is identical for each element and unified for the communications mesh, and guarantees delivery of instruction data to each tile each cycle. Data fetching is entirely programmatically controlled, which is a task often allocated to the compiler.

Other pipelined structures are possible; this structure was chosen since it encompasses instructions required for this signal processing task. Figure 7.1 shows the pipeline stages underlaid on the tile architecture.

7.3.2 Language Development

The IBM BIT programming language is architecture-specific and *functional*. A functional language treats computation as iterative mathematical function evaluation and specifically avoids use of state and mutable data (see, for e.g. Hudak, 1989)[4].

Functional languages suit signal processing abstraction, and have been used in this field previously. Sander and Jantsch (2004) use a functional language, Haskell[5], as a base for formal system design and modeling in signal processing systems. Both the LAVA[6] (Singh, 2003) and Feldspar[7] (Axelsson et al., 2010) languages also use a Haskell base for structural signal processing abstraction, the latter through embedding a domain-specific language (DSL) within Haskell, and the former directly.

Functional Microcode Language

At the top-level of the IBM BIT programming model is the functional microcode language (FML), composed of nested sets of pure functions. FML is a functional, architecture-specific language since it is designed with a fixed set of base operations for a particular architecture, and thus occupies a lower-level of abstraction in model space than the Haskell-embedded languages. FML allows algorithmic blocks to be defined as single instruction calls having no state-based *side-effects*.

7.3.3 Cycle-Accurate Simulator

The cycle-accurate simulator composes the final element of CABSim. The simulation kernel is extensible and object-oriented and schematically illustrated in the Unified Modelling Language (UML)[8] as a class diagram in Figure 7.4(b). In the figure, *HMSB Proc* is the highest-level processor object which contains processor tiles, communication interfaces, and it communicates with the the instruction scheduler. The simulator's task is twofold. Simulation generates processor timing statistics as well as executable vertical microcode (uncompressed machine register transfer instructions). Figure 7.4(a) schematically represents the structure of the language environment. FML is compiled into IBM BIT microcode in a two-pass process, generating

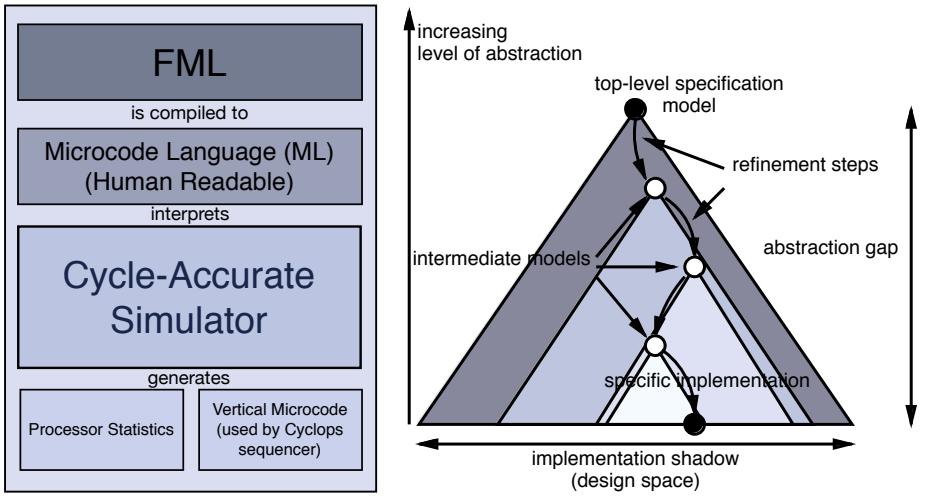
[4]The modification of state variables on independent calls to the same function in an *imperative* language may have a different, unpredictable result each time. Non-predictable effects or code clauses dependent on state are called side-effects. In contrast, functional languages obviate state and mutable data and thus do not have side effects.

[5]http://www.haskell.org/

[6]http://www.raintown.org/lava/

[7]http://dsl4dsp.inf.elte.hu/feldspar/index.html

[8]www.uml.org/



(a) (b)

Figure 7.4: A schematic representation of the CABSim environment. 7.4(a) is a relational model between language, simulator and compilation products. The Functional Modeling Language (FML) represents algorithms as pure functions. This language is compiled to human readable microcode. The cycle-true simulator reads (compressed) microcode, and simulates array processor execution. Processor cycle information as well as vertical microcode (which is stored in the scheduler and used directly by tile processors) are both produced by the simulation system. 7.4(b) is a visualisation of the language design space, modified from the schematic of Jantsch and Sander 2005. Higher levels of abstraction have a wider implementation shadow. Intermediate models define iteratively narrower areas of the implementation space. Since FML is architecture-specific, it it situated lower down the implementation tree and at a lower level of abstraction.

a human-readable microassembly intermediately. These sets of instructions and control loops are compiled to a human-readable microcode language, where each line of microcode represents a single processor cycle.

7.4 Algorithm Implementation in Functional Microcode Language

One of the central motivations behind the CABSim environment is to determine the efficiency of the massively-parallel array processors — in which the IBM BIT is typical of its class — for radio-frequency processing. Efficiency is best determined by actually implementing beamformer algorithms. This section describes algorithm design and implementation in FML for the fiducial tile processing model.

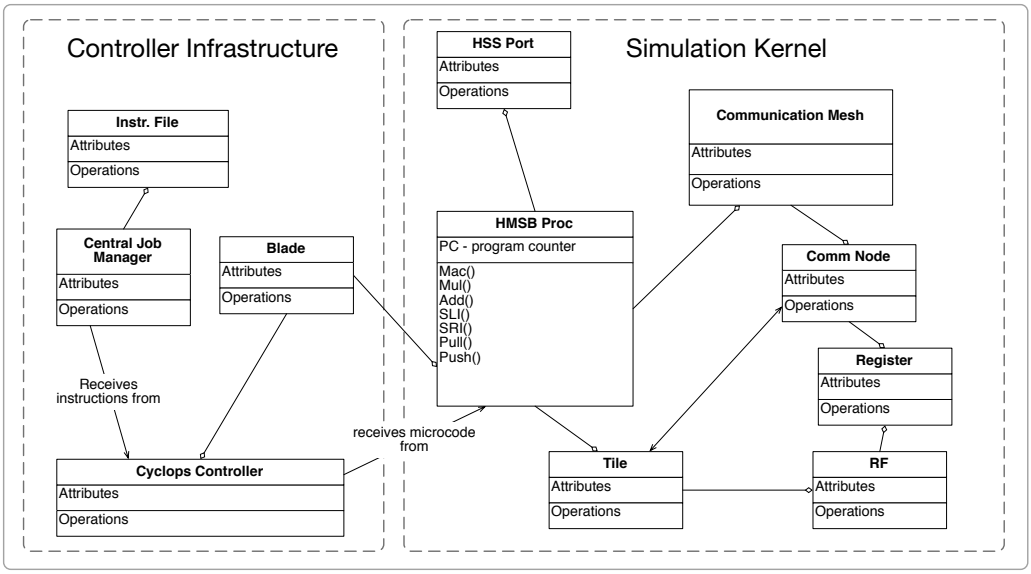


Figure 7.5: A Unified Modelling Language (UML) class-diagram (abridged for clarity) of the simulation environment. The simulator is extensible and designed with object-oriented principles. Particularly, the simulation kernel is parameterisable for different array processor configurations and numbers of elements.

7.4.1 Filter Architecture

The signal is decomposed into its frequency components with a Poly-phase FFT Filter Bank (PFFB), which is constructed from a bank of time-shifted groups of the prototype low-pass FIR filter[9]. Architectural filter design is not concerned with the actual coefficients that are to be implemented; certainly these will define the performance of the filter, but the architectural design is agnostic to the actual coefficients.

Each filter in the array maintains the same basic structure, requiring 8 complex multiply-accumulate operations per sample. Physical implementation requires at least 8 multipliers per channel and storage memory for 8 filter coefficients. For IBM BIT implementation, a circular buffer of sample values is maintained and stored on each tile's register file. These are then multiplied with the (complex) filter coefficients, and a sum formed over the products. Each complex multiply takes four real multiply-accumulate operations.

[9] An FIR filter is a system in which the response to the unit impulse, $\delta[n] = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0. \end{cases}$ will stabilise within a finite interval.

7.4.2 Fast Fourier Transform

The Discrete Fourier Transform is here implemented with one of the well-known class of fast algorithms, the radix-2 Decimation-In-Time (DIT) FFT algorithm (Cooley and Tukey, 1965), due to its symmetric structure and algorithmic simplicity compared to the highly-optimised generated set of algorithms detailed in, for e.g. Frigo and Johnson (2005). Indeed, the symmetry of this base-2 algorithm naturally complements the 2-dimensional grid arrangement of processor cores. This will be shown below as the algorithm is further partitioned over processor cores.

Splitting the Fourier Transform

A 2^N-point fast Fourier transform may be constructed by performing two sets of $2^{N/2}$ separate $2^{N/2}$-point Fourier transforms on input data, interspersed by a intermediate transpose operation. This decimation is precisely equivalent to one step of the iterative decomposition of the of the discrete Fourier transform employed by the radix-2 DIT fast Fourier transform. The discrete Fourier transform is

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} nk}, \quad (0 < k < N-1) \tag{7.1}$$

A radix-2 DIT decomposition (Cooley and Tukey, 1965) is such that the transform is iteratively split into its even and odd-numbered components,

$$X_k = \sum_{l=0}^{N/2-1} x_{2l} e^{-\frac{2\pi i}{N} 2lk} + \sum_{l=0}^{N/2-1} x_{2l+1} e^{-\frac{2\pi i}{N}(2l+1)k} \tag{7.2}$$

Defining E_k as the even-indexed transform, and O_k as the odd-indexed transform,

$$X_k = E_k + O_k e^{-\frac{2\pi i}{N} k} \tag{7.3}$$

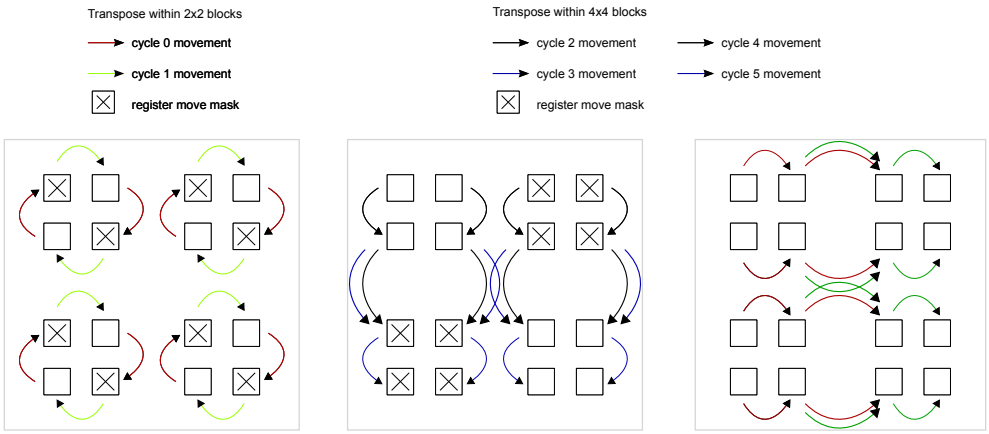
Which re-expresses X_k as two transforms of half its length, with one of these modulated by a complex exponential. This process is then repeated recursively until each transform is exactly 1 sample wide. This re-formulation reduces the complexity of the DFT from an $O(n^2)$ algorithm to the well-known complexity of $O(n \log_2 n)$.

A generalised binary factorisation re-interprets the discrete Fourier transform as two sets of smaller transforms, interspersed with a Fourier transpose operation. Making use of binary factorisation, the algorithm is

now mapped onto the array layout of the IBM BIT architecture. Given a square grid of $2^{A/2}$ x $2^{A/2} = 2^A$ processor tiles, a transform of size $2^{A/2}$ is performed on each tile. These intermediate data are transposed, then a second transform of size $2^{A/2}$ is performed; this process is equivalent to a 2^A-point transform.

Fourier Transpose

The transform developed above relies on an efficient and realisable intermediate data transpose between processor tiles. An $O(\sqrt{N}\log_2 N)$ transpose is designed on the processor mesh. The central approach is again a 'divide-and-conquer' strategy: a large transpose is built up by smaller transposes of logarithmically increasing size. There need be $k = \log_2 N$ stages of transpose, each consisting of $\log_2 k$ operations. Figure 7.6(a) shows the steps involved in transposing an 4×4 block of data registers.



(a)

Figure 7.6: Schematic showing representative component steps of an iterative $O(\sqrt{N}\log_2 N)$ transpose. A 4×4 transpose is composed from 4 sets of 2×2 intermediate transpose steps.

This Fast Fourier Transform (FFT) is used as the core of the PFFB. A spatial Fourier transform computes all sky beams of the incident wavefront over the array aperture. Again, the symmetric, radix-2 DIT algorithm structure of the PFFB is used.

7.4.3 Generator Program

At the highest level of the specification tree of Figure 7.4(b) lies the signal processing algorithm itself, which is typically a mathematical description of the algorithm. Specification at this level is implementation-agnostic. Specification refinement limits the implementation space to that spanned by specific architectures

or even by a single architecture. FML is a domain-specific language, and thus at a lower abstraction level. FML algorithms are composed of *generator* functions. The top-level FML program listing for beamformer processing is:

Listing 7.1: FML top-level beamformer generator

```
trans_mach2fir_gen(f,NUM_COLS,0)                    1
fir_complex_gen(f, NUM_FIR_TAPS, 3)                 2
fir2fft_gen(f)                                      3
fft_gen(f,int(sqrt(NUM_CHANNELS)))                  4
transpose_gen(f)                                    5
fft_gen(f, int(sqrt(NUM_CHANNELS)))                 6
fft2trans_mac_gen(f,NUM_COLS,0)                     7
```

Each of the above generators maps to a signal processing function. Generators are parsed and interpreted by the compiler, which generates expanded microcode language (ML) statements.

7.5 Performance Analysis

Table 7.1 lists the final contributions of each algorithmic stage of the beamformer to the total processing time for a window of 2^{10} time samples. The table lists the the number of cycles used by each coherent block (each block corresponds to an FML functional generator) within the algorithm, when the processor runs at 533MHz. This relatively low frequency is chosen to reduce energy consumption; 533MHz is approximately the optimal energy consumption knee frequency[10] for the 90nm CMOS node size (Denneau, 2008; Pullini et al., 2007). The FFT operation consumes the majority of the processor time, which totals 2333 processor cycles in 4305.82 ns.

7.5.1 Architectural Efficiency

In essence, architectural efficiency for a particular machine is a simple relation between operation rate and energy consumption. The question of the system boundary is important in this case. Some authors use the total system power to determine operations-per-Watt (see Feng and Scogland 2009 and www.green500.org/). This is useful when the infrastructure power is known. In a specific architecture comparison, the processor

[10]The energy consumption knee frequency is the frequency at which the exponential energy consumption curve intersects with the static power dissipation curve and defines an optimal energy consumption frequency for a particular processor .

Algorithm Block	Cycles	time (ns)	% time
Buffer to FIR	31	58.16	1.4
FIR Filter Bank	513	962.48	22.4
FIR to FFT	66	123.83	2.9
4 sets of 2^5-point FFT	1280	2401.50	55.8
3 sets of Inter-FFT Transpose	405	759.85	17.6
Total	2333	4305.82	-

Table 7.1: Processor utilisation in the beamformer by algorithmic block. The Fast Fourier Transform is the dominant processing operation

boundary may define the system boundary. It is this definition that is used in the calculations which follow. Total system power is derived from the final results.

The chip power consumption for the processor in 90nm CMOS technology P_{BIT} is $\sim$ 45 W (Denneau, 2008). Maximum processing capacity with 8-bit data pipelines for the IBM BIT processor is $\gamma_{\text{BIT8}} = 8.73 \times 10^{12}$ MACs. This information is used to derive η_{arch} for the IBM BIT processor.

First, the energy efficiency of the IBM BIT in 8-bit pocketisation mode, E_{BIT8} is

$$\begin{aligned} E_{\text{BIT8}} &= \frac{\gamma_{\text{BIT8}}}{P_{\text{BIT}}} \\ &= \frac{8.73 \times 10^{12}}{45W} \\ &= 1.94 \times 10^{11} \text{ MAC.J}^{-1} \end{aligned}$$

and the ideal energy efficiency E_{ideal} derived from the device model in Chapter 6 is

$$\begin{aligned} E_{\text{ideal}} &= \frac{\gamma_{\text{proc8}}}{P_{\text{ideal8}}} \\ &= \frac{\gamma_{\text{proc8}}}{E_{\text{min}} \cdot e \cdot 2^{n-1} \cdot \eta_{\text{dev}} \cdot \gamma_{\text{proc8}}} \\ &= 2.44 \times 10^{12} \text{ MAC.J}^{-1} \end{aligned}$$

Now, the ideal energy model was derived for 45 nm CMOS technology, so requires a scaling factor for direct comparison to the IBM BITs technology generation. The generation-energy-scaling factor is $\varepsilon_{\text{BIT}} \simeq 0.7g$ where g is an integer representing the generation difference (i.e. $g = 2$ and $\varepsilon_{\text{BIT}} = 0.49$). Then the architectural efficiency η_{arch} is the difference between the ideal device energy consumption E_{ideal} and the

actual energy consumption of the processor:

$$\begin{aligned}\eta_{\text{arch}} &= \frac{E_{\text{proc8}} \cdot \varepsilon_{\text{BIT}}}{E_{\text{dev}}} \\ &= 12.57 \end{aligned} \tag{7.4}$$

Algorithmic Efficiency

Algorithmic efficiency is derived for the IBM BIT. Theoretical maximum 8-bit-pocketised processing capacity is $\gamma_{\text{BIT8}} = 8.73 \times 10^{12}$ MACs. Algorithmic efficiency is a measure of the ratio of actual to theoretical calculations for a particular code. Now, following the convention of Section 6.1, the total number of complex multiply-accumulate operations required of an this beamformer implementation is

$$\begin{aligned} N_{\text{ops}} &= N_{\text{els}} \cdot (N_{\text{fil}} + N_{\text{chan}} + N_{\text{bf}}) \\ &= N_{\text{els}} \cdot (8 \cdot N_{\text{ch}} + C \cdot N_{\text{ch}} \cdot \log_2 N_{\text{ch}} + C \cdot N_{\text{ch}} \cdot \log_2 N_{\text{beams}}) \\ &= N_{\text{els}} \cdot N_{\text{ch}} \cdot (8 + 16C) \\ &= 2^{15} \cdot 168 \\ &= 5.51 \times 10^6 \end{aligned}$$

where N_{chan} is here 2^{10}, N_{beams} is 2^6 and the number of filter taps is 2^8. These operations are performed in $t_s = 4.38$ μs.

$$\begin{aligned}\gamma_{\text{alg8}} &= \frac{N_{\text{ops}}}{t_s} \\ &= 1.26 \times 10^{12} \text{ complex MACs} \\ &\simeq 5.03 \times 10^{12} \text{ MACs.}\end{aligned}$$

then the algorithmic efficiency is calculated from the theoretical maximum processor output $\eta_{\text{alg}} = \frac{\gamma_{\text{BIT8}}}{\gamma_{\text{alg8}}} = 1.27$ (alternatively, this is a *utilisation fraction* of 78%). This is an above-average utilisation fraction for a massively parallel processor. For example, the LOFAR correlator is a highly-optimised implementation which achieves a 44%, 86% and 96% of the FPU peak performance for the assembly-optimised FFT, FIR filter and correlator respectively (Romein et al., 2010). Average utilisation fractions for efficient supercomputers

are 47% and 67% for accelerated and non-accelerated supercomputers respectively (see www.green500.org/).

Three improvement contribute to this high utilisation fraction: the optimised $O(\sqrt{N} \log N)$ Fourier transpose, a hand-optimised butterfly inner-loop, and compiler instruction pipelining optimisation. Together, these improvements contribute a total improvement of 59.18%.

7.6 Summary

The IBM BIT Integer Array Processor (IBM BIT) represents a new class of energy-optimised Single Instruction stream Multiple Data stream (SIMD) supercomputer. It combines a simplified execution pipeline, a low-order interconnection architecture and massive parallelism to achieve high computational efficiency per joule. The processor is currently being built in the 90 nm CMOS technology node. Further iterations into 65nm and beyond promise even greater processing performance and energy efficiency.

The advantage of programmability makes heterogenous use possible. Fixed-point (integer) processing has advantages for energy-efficiency compared to floating point pipelines of general-purpose processing architectures. Such classes of architectures are absolutely required if the energy-efficiency target for signal processing as well as general data processing is to be met for the SKA aperture arrays.

This chapter has focussed on several aspects of the IBM BIT. Since the hardware is not yet available, the design and construction of a cycle-true architectural simulator for the IBM BIT array processor is described. A functional programming language has been designed in conjunction with a compiler specifically for the array processor. All of these elements are the author's own work designed for this thesis.

Together, the language and simulator enable both the architecture itself and potential signal processing codes to be characterised in terms of performance and energy-efficiency. The cycle-accurate simulator allows the architectural efficiency of the IBM BIT to be determined; it is 12.57 times the theoretical device model minimum. Secondly, a digital software beamformer is designed and implemented on the IBM BIT. Hand-optimisation of the FFT as well as compiler optimisation of the resulting code allows a utilisation ratio of 78% to be achieved ($\eta_{\text{alg}} = 1.27$). In particular, the FML compiler improves execution efficiency by 34%. It is able to exploit the reduction in control complexity provided by the explicitly software-controlled memory. It achieves this by overlapping data transport operations as well as adjacent instructions up to depth of 2.

What is clear from this analysis is that energy-efficient architectures are well-suited to beamformer signal processing.

Chapter 8

Conclusion

This thesis has focused on instrumentation research into digital aperture arrays, which are a key technology and thus a central area of development leading toward the Square Kilometre Array. In this brief chapter, I summarise the three strands of work constituting this thesis and make some suggestions for future avenues of digital instrumentation research as the two large Southern hemisphere radio telescopes, MeerKAT[1] and ASKAP[2] are being built, and as the Square Kilometre Array itself becomes a reality.

A low-frequency radio facility with collecting area of a square kilometre will further our understanding of the Universe. The shift toward aperture arrays as the major collector technology at low radio frequencies — a important range for future cosmological and astrophysical experiments on such facilities — is both well-justified and possibly inevitable, since parabolic dish-based instruments may not be economically scaled to the required physical size at low radio frequencies. Furthermore, the era of increasing construction commodity prices and exponentially-decreasing computing costs only makes aperture arrays more attractive with time.

Yet several questions face the development of digital aperture arrays as a usable technology. Since it is a new technique in astronomy, is wideband digital spatial filtering able to be demonstrated for astronomical applications across the required frequency band? Is the potential of digital aperture arrays for adaptive electronic optics able to be exploited? A demonstration of equivalent or better beam quality than traditional parabolic receivers is a critical first step toward truly adaptive control. Finally, how will the extreme amount

[1] www.ska.ac.za/
[2] www.atnf.csiro.au/projects/askap/

of digital signal processing that is required to enable the use of large digital arrays able to be provided in an energy-efficient manner?

This thesis addresses elements of each of the central concerns facing the digital aperture array and contends that despite these challenges it is indeed a feasible technology for future instruments. The exact contributions of this thesis in each of these three areas is discussed in more detail in the paragraphs that follow.

The LOFAR telescope[3] is already showing the broad scientific merit of fully digital aperture arrays at low radio frequencies. Yet it is important to demonstrate the feasibility of this technology at mid-frequencies, where the sky noise contribution begins to diminish and radiometer noise becomes increasingly important. The essential goal of the 2-Polarisation All-Digital (2-PAD)[4] project is exactly thus: to build and demonstrate a modular, fully-digital array operating over a frequency range where sky noise is a minor contributor to total system noise. The design, characterisation and deployment of the digital processing system for 2-PAD has been a major contribution of this thesis. This work has been described in Chapter 4.

Another aspect of concern for aperture arrays is provision of sufficient quality of beam response. Precise control over beam response unlocks the potential of adaptive optics with digital beamformers. In the work done in this thesis toward this broader goal, it was shown that a heuristic optimisation framework results in beam response that are improved over the analytical apodisation solutions, as well as correcting for non-ideal element responses. A smaller digital beamforming system is designed for this work. The results of digital beamforming with stochastic optimisation for small arrays show an improvement in both main beam width and side-lobe level compared to both Taylor and Dolph-Chebychev synthesis solutions.

The third central issue is that of processing infrastructures for large digital instruments. The magnitude of processing operations that must be performed for aperture array stations is dauntingly large. At the same time, cost requirements dictate the use of highly-energy-optimised architectures for the task. The first part of Chapter 6 is a qualitative exposure of some of the conflicting requirements placed on high-performance signal processing architectures. It is then shown through a quantitative model that the cost of energy limits the choice of processor architectures. Chapter 7 is a comprehensive investigation into one such energy-optimised processor, the IBM BIT Integer Array Processor (IBM BIT), a research supercomputer built at the IBM T.J. Research Centre. Before the architecture became available, I designed and built a cycle-accurate simulator and functional programming language for the array processor architecture. This enabled

[3] www.lofar.org/
[4] again, see 2-pad.physics.ox.ac.uk/

beamforming algorithms to be developed for the array processor. Using these results, the architectural and algorithmic efficiency of the IBM BIT array integer processor and its accompanying software beamformer respectively have been determined. Furthermore, this work has also afforded the opportunity for true co-design; software bottlenecks identified in code executed in the simulator informed architectural modifications in the processor design.

8.1 Further Work: Next-Generation Radio Astronomy Instrumentation

The work described in this thesis is a set of early results and findings in what will surely develop further as a large area of astronomy instrumentation research. As aperture array technology for the SKA becomes closer to becoming reality, there are several areas in which further research on digital instrumentation is merited, some of the most pressing of which are identified below.

8.1.1 The Processing Issue

In Chapter 6, a two-fold quantitative energy model of digital computation has been developed. The top-down model, based on industry predictions for future system energy costs, is in broad agreement with a detailed bottom-up energy model consisting of device, system, architecture and algorithmic efficiency contributions.

In the area of signal processing for the SKA, the two most important considerations are as follows. The processing requirement is large, and is not feasible with current technology. Yet, this level of processing capacity will be available in future if the trend of exponential increase in processing capacity continues, which it seems likely to do. However, the cost of energy is also significant, and is also an important limitation on potential architectures. In particular, based on an economic model, supplying more energy than approximately 8.76 kWh.yr^{-1} per element for processing (or equivalently, an average power of approximately 1 W) is not possible. Several conclusions emerge from this analysis, and are discussed below.

Energy-Optimised Supercomputers (and 'Super-Signal-Processors')

It was found that even the most efficient *general-purpose* supercomputers are not and will not be feasible for aperture-array signal processing by the time of phased construction of the SKA . Energy-efficient, highly-interconnected architectures must be sought, or developed.

Similar economic arguments of energy cost hold for the other large processing tasks required of the SKA. Both correlation and interferometric imaging are particularly problematic. Each is expected to require even greater amounts of computing capacity than beamforming processing (see Dewdney et al., 2010), albeit with lower communication requirements.[5] Furthermore, imaging processing is expected to be required to be performed in floating-point arithmetic. Energy-cost restrictions have often been neglected in past analyses. However, the energy-cost limit will likely motivate fundamental algorithmic change in current interferometric imaging techniques, particularly in their efficiency, if the SKA is to be a reality.

Future Processing Architectures

The issues of energy and capacity stand out as central determinants on the eventual choice of processing architecture. Yet it is often difficult to compare directly the broad spectrum of possible architectures on these metrics. In order to facilitate a fair comparison, a goal for digital architecture research in this area in future would be to compile a set of standard performance metrics (such as the device, architecture and algorithmic efficiency, as has been done in this thesis for a particular architecture) for the set of architectures that can be scaled to the requisite size. The set of possible architectures is likely to be smaller than may be expected, due to the threefold requirement for a large amount of *processing capacity*, large *inter-processor bandwidth* and communication capacity, and large *I/O-to-compute* ratios.

8.1.2 Quantisation

There is a fundamental systematic error facing digital computation, particularly important for narrow-bit-width architectures; in Chapter 2 it has been shown that physically-implementable finite-bit-width processing introduces an unavoidable error term. Quantisation inherent to initial digitisation as well as finite-bit-width arithmetic introduces non-linear distortion into the signal path. Furthermore, this distortion is *not* independent of the signal, and is imperfectly modelled by an independent, Gaussian additive noise process — a ubiquitous model in digital signal processing analysis. It is also shown that generally no exact expression for the quantisation noise exists, and that both the analytic and numerically-approximate noise expressions which do exist for certain input signals co-vary with the input signal. Initial quantisation is not the only noise contributor; the repeated re-quantisation of finite-bit-width arithmetic and integrative detection introduces

[5] These two processing tasks are differentiated by their I/O-to-compute ratios: correlation requires many fewer operations per sample than does imaging (where the latter is estimated to require of order $\sim 10^5$ OPs/sample (Dewdney et al., 2010)

further complexity to the analytic modeling approach.

Nevertheless, quantisation error is generally small, with quantisation noise power being some fraction of the smallest representable quantum Δ, and *approximately* equal to $\Delta^2/12$. If this is the case, then why is it such a concern? Even such a small non-Gaussian error is concerning in the context of detection of faint signals with instantaneous power much smaller than the instrumental noise floor. Such detection relies on many digital processing operations as well as deep integration of the detected signal to expose the buried signal. Non-Gaussian quantisation spectra are not reduced as expected through integration and digital processing. Many experiments planned for the SKA, and other radio telescopes that use digital systems, require data product precision that is much finer than the quantisation error.

Furthermore, while the quantisation distortion problem is reduced by using wide data-paths (since it is always related to the smallest representable quantum), it is not removed by doing so. And as has also been shown, wide data paths cost energy that may not be available. What is clear is that since digital computation is so central to modern radio telescopes, systematic quantisation effects, some of which were shown in this thesis, merit deeper analysis to be able to better constrain this systematic, non-linear contributor.

8.2 Outlook and Concluding Remarks

The essential scientific goal of the SKA is to determine the large-scale properties of the Universe, particularly the amount, distribution, and nature of its matter and energy, its age, and its evolutionary history. Neutral hydrogen tomography spanning significant redshift will be a strong observational constraint on our cosmology. An HI survey instrument able to provide three-dimensional matter distribution measurements possesses much further scientific potential. Pulsar surveys of the Galaxy and beyond are positioned to discover exceptional systems which test relativistic gravity in the ultra-strong-field limit. Measurement of the minute variations in the time-of-arrival of pulses from thousands of millisecond-period pulsars should enable the detection of the stochastic background signature of gravitational radiation, which is an as-yet-unobserved but key prediction of general relativity. The evolution and formation of magnetic fields and their relation to structure formation is promised by all-sky rotation measures of $\sim 10^8$ extragalactic sources. Light will be shed on the astrophysical conditions necessary for formation of life and earth-size ($\gtrsim 4M_\oplus$) planets. What is also clear is that an instrument that will map unexplored portions of the frequency- and time- variable sky will likely yield unexpected or even unimagined discoveries. The SKA is uniquely positioned to provide a fundamental

scientific contribution.

Anticipating the transformational astronomical tool provided by spatial matter tomography, as well as the tools provided by the other SKA surveys, digital instrumentation research acquires an urgency in its rôle as an enabler of these measurements. Low-frequency radio astronomy is entering a phase of extensive worldwide development in the instruments that are designed in preparation for and directly toward the Square Kilometre Array. At the same time, the ubiquity of digital processing is becoming ever more extensive; there are few future radio telescopes which do not include a significant amount of processing that will be performed digitally. This motivates renewed theoretical work on digital processing with extremely low signal-to-noise ratios. Furthermore, while the SKA instrumentation community is certainly positioned to benefit from the exponential processing increases that seem inevitable, it must too take ownership of the development of the class of processing architectures that meet its particular, energy-constrained requirements.

Yet over all, the outlook for the Square Kilometre Array is extremely promising. What is to all of our great benefit is the vast community worldwide of exceptional minds being applied to the immediate and long-term aspects of the problem. I submit this before the site decision has been made, although it is believed to be imminent. Whichever of the excellent sites in Australia or Africa eventually becomes the permanent home of the SKA will not diminish that it will then be the fruition of the long plans of the *world-wide* radio astronomy community. The SKA project is well-positioned to achieve its great potential.

Bibliography

Abragam A., Proctor W. G. (1958), Phys. Rev., 109, 1441.

Ade P., Aghanim N., Arnaud M., Ashdown M., Aumont J., Baccigalupi C. *et al.* (2011), arXiv.

Aghanim N., Bartelmann M., Bersanelli M., Bouchet F., Davis R., Efstathiou G. *et al.* (2006), arXiv.

AhmedSaid A. (2009), Widefield Science and Technology for the SKA, 1, 289–293.

Amanullah R., Lidman C., Rubin D., Aldering G., Astier P., Barbary K. *et al.* (2010), ApJ, 716, 712.

Armstrong R. P., Hickish J., ZarbAdami K., Jones M. E. (2009), Widefield Science and Technology for the SKA, 1, 273–278.

Armstrong R. P., Hickish J., ZarbAdami K., Jones M. E. (2010), Phased Array Systems and Technology, IEEE Int. Symp. on, pages 686 – 689.

Armstrong R. P., Jones M. E. (2010), Phased Array Systems and Technology, IEEE Int. Symp. on, pages 680 – 685.

Atmel (2006), AT84AD001B Datasheet.

Axelsson E., Claessen K., Dévai G., Horváth Z., Keijzer K., Lyckegård B. *et al.* (2010), Formal Methods and Models for Codesign, 8th IEEE/ACM Int. Conf. on, pages 169 – 178.

Balanis C. A. (1982), Antenna theory: Analysis and design, New York, Harper and Row(4th edn.).

Barsdell B. R., Bailes M., Barnes D. G., Fluke C. J. (2012), eprint arXiv, 1201, 5380.

Beck R., Gaensler B. M. (2004), New Astronomy Reviews, 48, 1289.

Becker R. H., Fan X., White R. L., Strauss M. A., Narayanan V. K., Lupton R. H. *et al.* (2001), 122, 2850.

Bell J., Ekers R. (2003), SKA Memorandum, 1(30).

Bellanger M., Daguet J. (1974), Communications, IEEE Trans. on, 22(9), 1199 – 1205.

Bennett W. (1948), Bell Syst. Tech. J., 27(4), 446–472.

Bevelacqua P., Balanis C. (2007), Antennas and Propagation, IEEE Trans. on, 55(12), 3442 – 3449.

Bhaumik S., George D. (2009), Widefield Science and Technology for the SKA, 1, 279–282.

Blackman R. B., Tukey J. W. (1958), Bell Syst. Tech. J., 37(1), 185–282.

Blake C. A., Abdalla F. B., Bridle S. L., Rawlings S. (2004), New Ast. Rev., 48, 1063–1077.

Boeringer D., Werner D. (2004), Antennas and Propagation, IEEE Trans. on, 52(3), 771 – 779.

Bolton R., Harris G., Faulkner A., Ikin T., Alexander P., Jones M. *et al.* (2009), SKA Memorandum, 1(111).

Borkar S. (2009), Proceedings of the 46th IEEE Annual Design Automation Conference, pages 93–94.

Bucci O., D'Elia G., Mazzarella G., Panariello G. (1994), Proc. of the IEEE, 82(3), 358 – 371.

Burbidge G. R. (1956), ApJ, 124, 416.

Burns W. R., Yao S. S. (1969), Radio Sci., 4(5), 431–436.

Candes E. J., Romberg J., Tao T. (2005), Communications of Pure and Applied Mathematics, 59(8), 1207–1223.

Chakrabarti S. K., Rosner R., Vainshtein S. I. (1994), Nature, 368(6470), 434–436.

Chang T.-Y., Bibyk S. (1999), Circuits and Systems, IEEE Int. Symp. on, 2, 372 – 375.

Chebyshev P. (1854), Mémoires des Savants étrangers présentés à l'Académie de Saint-Pétersbourg, 7, 539 – 586.

Chen L., Hu Z., Lin J., Gao G. (2007), Parallel and Distributed Processing, IEEE Int. Symp. on, pages 1 – 8.

Chengalur J. N., Braun R., Wieringa M. (2001), A&A, 372, 768.

Christie P. (1993), Proc. of the IEEE, 81(10), 1492 – 1499.

Clark M. A., Plante P. C. L., Greenhill L. J. (2011), eprint arXiv, 1107, 4264.

Clavier A., Panter P., Grieg D. (1947), American Institute of Electrical Engineers, Trans. of the, 66(1), 989 – 1005.

Cline T. L., Desai U. D., Teegarden B. J., Evans W. D., Klebesadel R. W., Laros J. G. *et al.* (1982), ApJ, 255, 45–48.

Contreras G., Martonosi M., Peng J., Lueh G.-Y., Ju R. (2007), ACM Trans. Embedd. Comput. Syst., 6(1).

Cooley J. W., Tukey J. W. (1965), Math Comput, 19(90), 297–301.

Cordes J. (1975), Ph.D. Thesis, California Univ., San Diego.

Cordes J. (2007), *The Square Kilometer Array as a Radio Synoptic Survey Telescope: Widefield Surveys for Transients, Pulsars and ETI.*

Cordes J. M., Lazio T. J. W., McLaughlin M. (2004), New Ast. Rev., 48(11-12), 1459 – 1472.

Costa E., Bampi S., Monteiro J. (2001). Power Efficient Arithmetic Operand Encoding.

Cox H. (1973), The Journal of the Acoustical Society of America, 54(3), 771–785.

Cox H., Zeskind R. M., Owen M. M. (1987), Acoustics, Speech and Signal Processing, IEEE Trans. on, 35(10), 1365 – 1376.

Crochiere R., Rabiner L. R. (1983), Multirate Digital Signal Processing, Prentice-Hall (Englewood Cliffs, N.J.).

Daishido T., Tanaka N., Takeuchi H., Akamine Y., Fujii F., Kuniyoshi M. *et al.* (2000), Proc. of the SPIE, 4015, 73–85.

Dally W. (1990), Computers, IEEE Trans. on, 39(6), 775 – 785.

Davis J., De V., Meindl J. (1998), Electron Devices, IEEE Transactions on, 45(3), 580–589.

Davis J., Venkatesan R., Kaloyeros A., Beylansky M., Souri S., Banerjee K. *et al.* (2001), Proc. of the IEEE, 89(3), 305 – 324.

DeHon A., Adams J., DeLorimier M., Kapre N., Matsuda Y., Naeimi H. *et al.* (2004), Field-Programmable Custom Computing Machines, 2004. 12th Annual IEEE Symposium on, pages 13 – 23.

del Cuvillo J., Zhu W., Hu Z., Gao; G. R. (2006), High-Performance Computing in an Advanced Collaborative Environment, 20th IEEE Computer Society Symposium on, page 9.

del Rio L., Aberg J., Renner R., Dahlsten O., Vedral V. (2011), Nature, 474(7349), 61–63.

Deller A. T., Tingay S. J., Bailes M., West C. (2007), The Publications of the Astronomical Society of the Pacific, 119, 318.

Denneau M. (2008), Priv. Communication.

Dewdney P., bij de Vaate J.-G., Cloete K., Gunst A., Hall D., McCool R. *et al.* (2010), SKA Memorandum Series, (130).

Dewdney P. E., Carlson B. R. (2000), American Astronomical Society, 196, 719.

Dewdney P. E., Landecker T. L. (1991), Radio interferometry: Theory Techniques and Applications, 19, 415–419.

Dicke R. H. (1946), The Review of Scientific Instruments, 17(7), 268.

Dolag K., Stasyszyn F., Donnert J., Pakmor R. (2009), Proc. Int. Ast. U., 4(259), 519–528.

Dolph C. (1946), Proceedings of the IRE, 34(6), 335 – 348.

Dunkley J., Komatsu E., Nolta M. R., Spergel D. N., Larson D., Hinshaw G. *et al.* (2009), ApJS, 180, 306.

Efstathiou G. (2003), MNRAS, 346, 26.

Eisenstein D. J., Zehavi I., Hogg D. W., Scoccimarro R., Blanton M. R., Nichol R. C. *et al.* (2005), ApJ, 633, 560.

Elliott R. S. (2003), Antenna Theory and Design, Wiley-IEEE Press.

Emerson R. F. (1976), The Deep Space Network Progress Report, 34, 54.

Ewen H. I., Purcell E. M. (1951), Nature, 168, 356.

Fan X., Strauss M. A., Schneider D. P., Becker R. H., White R. L., Haiman Z. *et al.* (2003), 125, 1649–1659.

Faulkner A., Alexander P., Jones M., Bolton R., van Ardenne A., Torchinsky S. (2008), *Design of an Aperture Phased Array System for the SKA*.

Faulkner A., Kant D., Alexander P., Montebugnoli S., van Ardenne A., Picard P. *et al.* (2010), SKA Memorandum Series.

Feng W.-C., Scogland T. (2009). The Green500 List: Year One.

Field (1958), Proc. of the IRE, 46(1), 240 – 250.

Flynn M. J. (1972), IEEE Trans. Comput., C-21(9), 948–960.

Freudling W., Staveley-Smith L., Catinella B., Minchin R., Calabretta M., Momjian E. *et al.* (2011), ApJ, 727, 40.

Frigo M., Johnson S. (2005), Proc. of the IEEE, 93(2), 216 – 231.

Furlanetto S., Briggs F. H. (2004), New Ast. Rev., 48(11-12), 1039–1052.

Furlanetto S. R., Oh S. P., Briggs F. H. (2006), Phys. Rep., 433, 181–301.

Gaensler B. M., Beck R., Feretti L. (2004), New Ast. Rev., 48, 1003.

Gamma E., Helm R., Johnson R., Vlissides J. (1993), European Conference on Object-Oriented Programming, pages 406–431.

Gammaitoni L. (1995), Phys. Rev. E., 52(5), 4691–4698.

Gray R. (1990), Information Theory, IEEE Trans. on, 36(6), 1220 – 1244.

Gray R., Chou W., Wong P. (1989), Communications, IEEE Trans. on, 37(9), 956 – 968.

Gunn J. E., Peterson B. A. (1965), ApJ, 142, 1633.

Hagen J., Farley D. T. (1973), Radio Sci., 8(9), 775–784.

Harris F. (1978), Proc. of the IEEE, 66(1), 51–83.

Hayashida N., Honda K., Honda M., Inoue N., Kadota K., Kakimoto F. *et al.* (1996), Phys. Rev. Lett., 77(6), 1000.

Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F., Collins R. A. (1969), Nature, 224, 472.

Heywood I., Armstrong R. P., Booth R., Bunker A. J., Deane R. P., Jarvis M. J. *et al.* (2011), Astronomy with Megastructures: Joint Science with the E-ELT and SKA.

Hicken M., Wood-Vasey W., Blondin S., Challis P., Jha S., Kelly P. *et al.* (2009), ApJ, 700, 1097–1097–1140–1140.

Hickish J. (2010), Priv. Communications.

Hobbs G., Archibald A., Arzoumanian Z., Backer D., Bailes M., Bhat N. D. R. *et al.* (2010), Class. Quantum Grav., 27.

Hsu C.-H., Shyr W.-J. (2005), Circuits, Systems, and Signal Processing, 24(4), 327–341.

Hudak P. (1989), ACM Comp. Surveys, 21(3).

IEA (2011), International Energy Agency: Energy Prices and Taxes, 2011.

IEEE (1994), IEEE Std 1057-1994.

Ikin T., Wilkinson P., Faulkner A., Jones M., Baird A., Brown A. *et al.* (2009), Widefield Science and Technology for the SKA, 1, 267–272.

ITO (2007), InfiniBand Architecture Specification Release 1.2.1.

Jantsch A., Sander I. (2005), Computers and Digital Techniques, IEEE Proc. on, 152(2), 114 – 129.

Jenet F. A., Anderson S. B. (1998), Publ. Astron. Soc. Pacific, 110(754), 1467–1478.

Johnson J. B. (1927), Nature, 119, 50.

Johnson J. B. (1928), Phys. Rev., 32, 97.

Johnson R., Gamma E., Vlissides J., Helm R. (1995), Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley.

Johnson S. G., Frigo M. (2007), IEEE Transactions on Signal Processing, 55, 111.

Jonas J. L. (2009), Proc. of the IEEE, 97, 1522.

Kedar S., Hajj G. A., Wilson B. D., Heflin M. B. (2003), Geophys. Res. Lett., 30(16).

Kennedy J., Eberhart R. (1995). Particle swarm optimization. volume 4, pages 1942 –1948.

Khan A., Brown A. (2009), Antennas. & Propagation, 3rd Eur. Conf. on, pages 3011 – 3013.

Khlebnikov V. A., Zarb-Adami K., Armstrong R. P., Jones M. E. (2010), Phased Array Systems and Technology, 2010 IEEE Int. Symp. on, pages 911 – 916.

Kogut A., Spergel D. N., Barnes C., Bennett C. L., Halpern M., Hinshaw G. *et al.* (2003), ApJS, 148, 161.

Komatsu E., Smith K. M., Dunkley J., Bennett C. L., Gold B., Hinshaw G. *et al.* (2011), ApJS, 192, 18.

Koomey J. G. (2008), Environ. Res. Lett., 3, 034008.

Kotelnikov V. A. (1933). On the transmission capacity of the ether and of cables in electrical communications.

Kowalski M., Rubin D., Aldering G., Agostinho R. J., Amadon A., Amanullah R. *et al.* (2008), ApJ, 686, 749.

Kramer M. (2007), Exploring the Cosmic Frontier: Astrophysical Instruments for the 21st Century, ESO Astrophys. Symp. on., page 87.

Kramer M., Backer D. C., Cordes J. M., Lazio T. J. W., Stappers B. W., Johnston S. (2004), New Ast. Rev., 48, 993.

Kraus J. D. (1966), Radio Sstronomy, New York: McGraw-Hill.

Kruse E., Ransom S. (2012), American Astronomical Society, 219. (c) 2012: American Astronomical Society.

Kumar A., Murthy P. K. (1976), Antennas and Propagation, IEEE Trans. on, 24(6), 865 – 870.

Kumar A., Murthy P. K. (1977), Antennas and Propagation, IEEE Trans. on, 25(3), 425 – 428.

Lah P., Chengalur J. N., Briggs F. H., Colless M., de Propris R., Pracy M. B. *et al.* (2007), MNRAS, 376, 1357.

Lah P., Pracy M. B., Chengalur J. N., Briggs F. H., Colless M., de Propris R. *et al.* (2009), MNRAS, 399, 1447.

Landauer R. (1961), IBM J. of Research and Development, 5(3), 183 – 191.

Landman B., Russo R. (1971), Computers, IEEE Trans. on, C-20(12), 1469 – 1479.

Larson D., Dunkley J., Hinshaw G., Komatsu E., Nolta M. R., Bennett C. L. *et al.* (2011), ApJS, 192, 16.

Lemmon J. J. (1997), Radio Science, 32(2), 525–539.

Lewis (1974), Antennas and Propagation, IEEE Int. Symp. on, 12, 335 – 337.

Linsley J. (1963), Phys. Rev. Lett., 10, 146.

Lloyd S. (1982), Information Theory, IEEE Trans. on, 28(2), 129 – 137.

Lopez J., Basterrechea J. (2007), Antennas and Propagation, 2nd IEEE Eur. Conf. on, pages 1 – 6.

Lorimer D. R., Bailes M., McLaughlin M. A., Narkevic D. J., Crawford F. (2007), Science, 318(777).

Lundy T., Buskirk J. V. (2007), Computing, 80, 23–45.

Mailloux R. (2005), Phased Array Antenna Handbook, Boston: Artech House.

Max J. (1960), Information Theory, IRE Trans. on, 6(1), 7 – 12.

McLaughlin M. A., Lyne A. G., Lorimer D. R., Kramer M., Faulkner A. J., Manchester R. N. *et al.* (2006), Nature, 439, 817.

McMahon P., Langman A., Werthimer D., Backer... D. (2007), casper.berkeley.edu.

Meindl J. (1995), Proc. of the IEEE, 83(4), 619–635.

Meuer H. W. (2008), *The Top 500 Supercomputer Project Report: Looking Back Over 15 Years of Supercomputing Experience*, `http://www.top500.org/`.

Miniati F., Bell A. R. (2011), ApJ, 729, 73.

Mol J. D., Romein J. W. (2011), eprint arXiv, 1105, 661.

Mollick E. (2006), Annals of the History of Computing, IEEE, 28(3), 62 – 75.

Monorchio A., Genovesi S., Bertini S., Brizzi A. (2007), Antennas and Wireless Propagation Letters, IEEE, 6, 484 – 487.

Moser P., Hofmann W., Wolf P., Lott J. A., Larisch G., Payusov A. *et al.* (2011), App. Phys. Lett., 98, 1106. (c) 2011: American Institute of Physics.

Mukhopadhyay S., Raychowdhury A., Roy K. (2005), *Switching Energy in CMOS Logic: How far are we from the physical limit?*

Noordam J., Braun R., de Bruyn A. (1991), NFRA Note, 1(585).

Nyquist (1928a), American Institute of Electrical Engineers, Trans. of the, 47(2), 617 – 644.

Nyquist H. (1928b), Phys. Rev., 32, 110.

Obreschkow D., Rawlings S. (2009a), APJ Lett., 696, L129.

Obreschkow D., Rawlings S. (2009b), MNRAS, 394, 1857.

Oppenheim A. V., Schafer R. W. (2009), *Discrete-Time Signal Processing, 2nd Edition*, volume Prentice Hall.

Otobe E., Nakajima J., Nishibori K., Saito T., Kobayashi H., Tanaka N. *et al.* (1994), Publ. Astron. Soc. Japan, 46, 503–510.

Paduit C. (2007), An Introduction to InfiniBand.

Palonpon A., Amistoso J., Holdsworth J., Garcia W., Saloma C. (1998), Opt. Lett., 23(18), 1480–1482.

Parsons A., Backer D., Siemion A., Chen H., Droz P., Filiba T. *et al.* (2008), Publ. Astron. Soc. Pacific, 120(873), 1207–1221.

Parsons A. R., Backer D. C., Foster G. S., Wright M. C. H., Bradley R. F., Gugliucci N. E. *et al.* (2010), Astron. J., 139, 1468.

Patel C., Chai S., Yalamanchili S., Schimmel D. (1997), Computer Design: VLSI in Computers and Processors, IEEE Int. Conf. on, pages 408 – 416.

Pei Y. C., Fall S. M. (1995), ApJ, 454, 69.

Pekau H., Haslett J. (2006), Circuits, Devices and Systems, IEEE Proc. on, 153(6), 517 – 524.

Penzias A. A., Wilson R. W. (1965), ApJ, 142, 419.

Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C. *et al.* (2007), MNRAS, 381, 1053.

Pober J., Parsons A., Backer D., Bradley R., Parashare C., Gugliucci N. *et al.* (2011). The Precision Array for Probing the Epoch of Reionization. volume 43.

Pozar D. (1994), Antennas and Propagation, IEEE Trans. on, 42(8), 1176 – 1178.

Pozar D. (1997), Microwave Engineering, 2nd edition(Wiley - New York).

Press W. H., Schechter P. (1974), ApJ, 187, 425. A&AA ID. AAA011.162.012.

Price D., Schediwy S., Jones M. E. (2009), *The 2-PAD Analogue System: Characterisation, Commissioning and Implications for Calibration and Beamforming.*

Proakis J. G., Manolakis D. G. (1996), *Digital signal processing: principles, algorithms, and applications*, volume Prentice Hall.

Pullini A., Angiolini F., Meloni P., Atienza D., Murali S., Raffo L. *et al.* (2007), Networks-on-Chip, IEEE Comp. Soc. Int. Symp. on, pages 273 – 282.

Rawlings S., Abdalla F. B., Bridle S. L., Blake C. A., Baugh C. M., Greenhill L. J. *et al.* (2004), New Ast. Rev., 48, 1013.

Rickett B. (1975), ApJ, 197, 185–191.

Robertson P. (1992), Beyond southern skies. Radio astronomy and the Parkes telescope. Price Âč 40.00.

Rohlfs K., Wilson T. L. (1986), Tools of Radio Astronomy, Springer-Verlag(4th edn.), 332.

Romein J., Broekema P., Mol J., vanNiewpoort R. (2010), Principles and Practice of Parallel Programming, ACM Symp. on.

Sander I., Jantsch A. (2004), Computer-Aided Design of Integrated Circuits and Systems, IEEE Trans. on, 23(1), 17 – 32.

Schelkunoff S. A. (1943), Bell Sys. Tech. Jour., 22, 80–107.

Schilizzi R. T., Alexander P., Cordes J. M., Dewdney P. E., Ekers R. D., Faulkner A. J. *et al.* (2007), SKA Memorandum, 1(100).

Shannon C. E. (1949), Proc. of the IRE, 37(1), 10 – 21.

Sharma S., Hsu C.-H., chun Feng; W. (2006), Parallel and Distributed Processing, 20th IEEE Int. Symp. on.

Sheppard W. F. (1898), Proc. London Mathematical Soc., 29, 353–380.

SIA (2011), Semiconductor Industry Association: International Technology Roadmap for Semiconductors.

Silk J. (1968), ApJ, 151, 459.

Singh S. (2003), Design, Automation and Test, IEEE Comp. Soc. Conf. on, 1.

Skolnik M. (1969). *Antenna Theory*, volume McGraw-Hill, pages 207–234.

Smolders (2000), Antennas and Propagation Society, 1st International Symposium, pages 162 – 165.

So H. K.-H., Brodersen R. (2006), Field Programmable Logic and Applications, IEEE Int. Conf. on, pages 1 – 6.

Solomon P. M., VandenBout P. A. (2005), Annu. Rev. Astron. & Astrophys., 43, 677.

Spergel D. N., Bean R., Doré O., Nolta M. R., Bennett C. L., Dunkley J. *et al.* (2007), ApJS, 170, 377.

Spergel D. N., Verde L., Peiris H. V., Komatsu E., Nolta M. R., Bennett C. L. *et al.* (2003), ApJS, 148, 175.

Springel V., White S. D. M., Jenkins A., Frenk C. S., Yoshida N., Gao L. *et al.* (2005), Nature, 435, 629.

Sripad A., Snyder D. (1977), Acoustics, Speech and Signal Processing, IEEE Trans. on, 25(5), 442 – 448.

Steyskal H., Herd J. (1990), Antennas and Propagation, IEEE Trans. on, 38(12), 1971 –1975.

Stimler S. (1969), Commun. of the ACM.

Sugawara S., Maita Y., Adachi K., Mori K., Mizuno K. (1997), Microwave Symposium Digest, IEEE MTT-S, 2, 959 – 962.

Swarup G. (1991), Current Science, 60, 106.

Taylor T. T. (1955), Antennas and Propagation, IEEE Trans. on, 3, 16–23.

Tegmark M., Zaldarriaga M. (2009), Phys. Rev. D, 79(8), 083530.

Theuns T., Schaye J., Zaroubi S., Kim T.-S., Tzanavaris P., Carswell B. (2002), ApJ, 567, L103.

Thompson A. R., Moran J. M., Swenson G. W. (2001), Interferometry and Synthesis in Radio Astronomy, New York : Wiley(2nd edn.).

Tie-Jun (1985), Acoustics, Speech and Signal Processing, IEEE Trans. on, 33(3), 527 – 536.

Trelea I. C. (2003), Information Processing Lett., 85, 317 – 325.

Tucker L., Robertson G. (1988), Computer, 21(8), 26 – 38.

Tyler N., Allen B., Aghvami H. (2004), Communications Magazine, IEEE, 42(12), 114 – 122.

Vardeman (2005), Instrumentation and Measurement, IEEE Trans. on, 54(5), 2117 – 2119.

Veen B. V., Buckley K. (1988), IEEE assp magazine, 5(2), 4–24.

Vetterli M. (1986), Signal Processing, 10(3), 219–244.

Vleck J. H. V. (1943), *The Spectrum of Clipped Noise*.

Wackersreuther (1985), Acoustics, Speech, and Signal Processing, IEEE Int. Conf. on, 10, 73 – 76.

Walter F., Carilli C., Daddi E. (2011), Rev. Mod. Astron., 22, 167.

Warnick K., Jeffs B., Landon J., Waldron J., Jones D., Fisher J. *et al.* (2009), Antenna Technology, IEEE Int. Workshop on, pages 1 – 4.

Webster R. (1983), Information Theory, IEEE Trans. on, 29(5), 765 – 767.

Weinreb S. (1963), MIT Research Laboratory of Electronics.

Weinreb S., D'Addario L. (2001), SKA Memorandum, 1.

Werthimer D., Cobb J., Lebofsky M., Anderson D., Korpela E. (2001), Computing in Science and Engineering, 3(1), 78 – 83.

Whittaker E. T. (1915), Proc. of the Royal Society of Edinburgh, 35, 181–194.

Whittaker J. (1935), Cambridge University Press.

Widrow (1956), Circuit Theory, IRE Trans. on, 3(4), 266 – 276.

Widrow (1982), Antennas and Propagation, IEEE Trans. on, 30(3), 469 – 478.

Widrow B., Kollar I. (2008), *Quantization Noise: Round Off Error in Digital Computation, Signal Processing, Control and Communications*, volume Cambridge University Press.

Widrow B., Kollar I., Liu; M.-C. (1996), Instrumentation and Measurement, IEEE Trans. on, 45(2), 353 – 361.

Wijnholds S., Boonstra A. (2006), Sensor Array and Multichannel Processing, Fourth IEEE Workshop on, pages 200 – 204.

Wijnholds S., van Cappellen W. (2011), Antennas and Propagation, IEEE Trans. on, PP(99), 1.

Wilkinson P. N. (1991). The Hydrogen Array. volume 19, pages 428–432.

Wilkinson P. N., Kellermann K. I., Ekers R. D., Cordes J. M., Lazio T. J. W. (2004), New Ast. Rev., 48, 1551–1563.

Wolfe A. M., Turnshek D. A., Smith H. E., Cohen R. D. (1986), ApJS, 61, 249.

Wouthuysen S. A. (1952), 57, 31.

Wright M., Bregman J., Braun R., Bunton R., Cornwell T., Ekers R. *et al.* (2006), *SKA Survey Optimization*.

Xilinx (2007), Virtex-II Pro and Virtex-II Pro X FPGA User Guide and Datasheet.

Ye W., Vijaykrishnan N., Kandemir M., Irwin M. (2000), Design Automation, 37th IEEE Conf. on.

Yourst M. (2007), Performance Analysis of Systems and Software, IEEE Int. Symp. on, pages 23–34.

Yu F., Ronghong J., Zhengyi W., Bo L., Junping; G. (2004), Signal Processing, 7th IEEE Int. Conf. on, 1, 428 – 430.

Zhang Y., Brown A. (2010), for review in Antennas and Propagation, IEEE Trans. on.

Zhang Y. P., Jeong T., Chen F., Wu H., Nitzsche R., Gao G. (2006), Parallel and Distributed Processing, 20th IEEE Int. Symp. on, pages 1–10.

Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L. (2005), MNRAS Lett., 359, L30.