

The “Magnificent Memo” Series: Trade-offs between Science and Engineering for the Square Kilometre Array

Submitted by Bryan Gaensler, SKA Project Scientist,
and Joseph Lazio,¹ Deputy Project Scientist,
on behalf of the SKA Science Working Group
2006 August 7

EXECUTIVE SUMMARY

In this document, we revisit the scientific rationale behind the instrumental specifications for the Square Kilometer Array (SKA), and explain the loss of science that would result from various design decisions or cutbacks. A summary of our findings is as follows.

- If a large number of pulsars need to be timed over long durations (e.g., years) to detect gravitational waves, multiple fields of view at full sensitivity will likely be required to achieve the needed throughput. Further simulations will be needed to quantify this requirement.
- The potential need for simultaneous availability of the SKA for deep surveys, short-term projects and outreach suggests that multiple, independently steerable fields of view would be valuable. A study needs to be made of how the SKA would be scheduled to achieve multiple science goals.
- Weak lensing experiments (requiring a very large number of pixels per image) and large-scale HI surveys (which require large images in a large number of spectral channels) both drive demanding requirements on combined angular resolution and field of view.
- Most Key Science will be adversely affected by a reduction in image dynamic range. Whenever it is necessary to limit the field of view, it will be important to use techniques that allow the highest dynamic range to be maintained.
- Weak lensing and the imaging of deep extragalactic continuum fields both require high angular resolution (baselines $\lesssim 1000$ km) at low frequencies ($\lesssim 2$ GHz). Without this capability, we will give up a crucial orthogonal measure of cosmological parameters, will be unable to study the faint active, star-forming and normal galaxy populations, and will lose the ability to identify the first black holes.
- The frequency range 200–500 MHz is essential for the SKA to capitalize on one of its greatest strengths: its unique ability to trace both the large-scale distribution of mass in the Universe and the transformation of HI gas into stars over a large fraction of cosmic time.
- The high surface brightness sensitivity provided by compact array configurations is needed for searches for pulsars and extraterrestrial intelligence, for studying HI in external galaxies, and for mapping magnetism in the intergalactic and interstellar medium. This capability is absolutely essential at low frequencies ($\lesssim 3$ GHz); at higher frequencies important science is lost, but the main Key Science goals remain achievable.
- Transient searching can be divided into two modes, slow and fast. With the current SKA design goals, a slow transient search that involves pixellating the entire field of view but no dedispersion will likely be possible with the computational resources available by the time the SKA is on-line. A fast transient search of a pixellated field of view will require algorithm improvements and considerable computational resources.

¹Basic research in radio astronomy at the Naval Research Laboratory is supported by the Office of Naval Research.

1 INTRODUCTION

The Square Kilometre Array (SKA)² is a next-generation radio telescope, currently being planned for the next decade. The SKA will be at least 50 times more sensitive than any other centimeter- to meter-wavelength telescope ever built, and will be able to answer fundamental questions about the origin and evolution of the Universe. In addition, the vast increase in sensitivity provided by the SKA will also almost certainly lead to the discovery of new and totally unexpected celestial phenomena.

The specifications for the SKA (Jones 2004, SKA Memo #45) require an angular resolution of $0''.02$ at 1.4 GHz, a frequency capability of 0.1–25 GHz, and a field of view of at least 1 deg^2 at 1.4 GHz. While the detailed design for the SKA is yet to be finalized, the SKA Reference Design (Schilizzi 2006, SKA Memo #69) currently consists of a central 5-km core of both steerable dishes and passive aperture tiles, with 50% of the total collecting area distributed on longer baselines, extending out to $> 3000 \text{ km}$. The dishes will be outfitted with phased arrays at low frequencies and wide-band feeds at higher frequencies. The aperture array will have a very wide field of view at low frequencies ($\sim 50 \text{ deg}^2$ at 700 MHz), to allow rapid surveys of redshifted H I. Operations for the SKA are expected to begin in 2015–2020.

One of the challenges for SKA design and planning is to make the appropriate trade-offs between scientific requirements and engineering reality. A compelling science case has been developed by the international community, as described in detail in *New Astronomy Reviews*, volume 48 (2004). The core of the science case is five Key Science Projects (KSPs):

- The Cradle of Life;
- Strong Field Tests of Gravity Using Pulsars and Black Holes;
- The Origin and Evolution of Cosmic Magnetism;
- Galaxy Evolution, Cosmology and Dark Energy;
- Probing the Dark Ages.

While these projects promise to answer fundamental questions in physics and astrophysics, they must be balanced against the estimated total cost for the instrument of approximately €1.5 billion. The combined KSP requirements entail a huge range in frequency, high sensitivity on all angular scales, and high angular resolution over a wide field of view. Whether all of this is feasible is a question that needs to be answered before a final design is implemented and construction begins.

1.1 The Magnificent Memos

In response to the above situation, the SKA Scientific Working Group (SWG) was asked to produce a series of memos, outlining the scientific rationale behind various SKA requirements, and explaining the loss of science that would result if certain design decisions were to be made. Originally seven questions were posed, leading to the “Magnificent Seven” memos. An eighth question was later added, resulting in the following set of questions:

1. What key science can be delivered by a Phase 1 SKA (10% collecting area), Phase 2 (50% collecting area), and a full SKA, in each case as a function of instantaneous field of view?
2. What is the science case for multiple independently-steerable fields of view?
3. What is the impact of limiting the field of view at high angular resolution?
4. What is the case for high angular resolution below a few GHz?
5. What is the case for frequencies between 200 and 500 MHz?

²See <http://www.skatelescope.org>.

6. What is the case for high filling factor at high frequency?
7. What are the options for transient detection?
8. What is the impact on key science of having just a high- or a low-frequency array?

The response to Question 1 has evolved into a comprehensive document on the science case for a 10% SKA, currently under preparation by the SWG. A report on Question 8 has been put on hold until further siting and engineering decisions have been made. This leaves responses to the remaining “Magnificent Six” questions, as detailed in the rest of this document.

2 WHAT IS THE CASE FOR MULTIPLE INDEPENDENTLY-STEERABLE FOVS?

Jim Cordes and the SWG

2.1 Preamble

The innovation of the SKA is the combination of a large field of view (FoV), an increase of up to two orders in sensitivity, and a rich palette of operating modes that can exploit these characteristics. Current SKA specifications³ call for a 1 deg² FoV at 1.4 GHz, increasing to ~ 200 deg² at 0.7 GHz.⁴ Specifications include one *separate* FoV with full sensitivity, ten sub-arrays and a goal of four separate FoVs with full sensitivity. Here we discuss the scientific need for multiple FoVs that are independently steerable. The KSPs do not provide a compelling need for multiple FoVs at full SKA sensitivity. A possible exception *may* be follow-up timing observations of pulsars, the number of which will be very large. However, further simulations of the process by which pulsars are discovered and then timed are needed in order to assess the throughput needed to make the pulsar KSP feasible. In considering a larger range of activities, accessibility and sociological issues come into play. Multiple, independent FoVs may be important for providing a multiplexing capability: (1) that would allow deep surveys to be conducted by one group while maintaining accessibility to other projects that are time constrained and to keep the project queue flowing; (2) that would ensure that the SKA is accessible to a wide range of users rather than only those involved with large surveys. Lastly, activities associated with the Exploration of the Unknown, such as transient studies and the search for extraterrestrial intelligence (SETI), might profit from multiple fields of view. Without knowing the luminosity functions for, e.g., transients or extraterrestrial intelligence (ETI) signals, we cannot yet say that multiple, independently steerable FoVs are required at full SKA sensitivity. Because independent FoVs at full sensitivity are a very constraining requirement, future work should aim at determining what science cannot be done by making use of sub-arrays to achieve sky coverage, either with contiguous or non-contiguous beams.

2.2 Definitions

Following the initial discussion on this topic at the Groningen SKA meeting in August 2002, we first define the large number of types of beams that enter the discussion:

Primary Beam: response of an individual antenna element. For a dual-antenna system (e.g., a paraboloid + feed antenna) this is the net angular response at a single focal point. For a single element of an aperture array, this would be the basic response of that element.

Station Beam = the synthesized beam of a *station* built up from primary elements.

SKA Beam = synthesized beam of the entire SKA.

Field of View (FoV) = the primary beam \times delay beam (or equivalent), i.e. the usable, instantaneous patch of solid angle that can be analyzed.

For aperture arrays, the FoV specification is much smaller than the intrinsic beam width of the tiles. Multi-beaming is therefore much more flexible in this design and it is possible to simultaneously sample regions of the sky separated by $\gg 1$ deg and using the entire collecting area of the SKA. Other designs may achieve large solid-angle coverage through use of multiple-beam systems at the focal point of a primary reflector (e.g. through use of focal-plane-sampling arrays or multiple feed antennas). In our definition of primary beam above, we would consider each pixel of a multiple beam system of this type to be a single FoV. While this differs from usage of the term in optical and other contexts, it is useful for the SKA because a single FoV defined this way indeed would form the boundaries of the analyzed field of view in, say, a VLBI observation.

³SKA Science Requirements: Version 2, SKA Memo 45, D. L. Jones

⁴Discussion at the Pune SKA 2005 meeting redefined this maximal FoV downward to 50 deg².

Core-Array Beam = the synthesized beam from elements within a central core array. The core array is simply a large station.

KSPs that need a core-array beam include epoch of reionization (EoR) studies, HI detection of galaxies, molecular line studies, and blind surveys for pulsars, transients and SETI.

It is useful to define the core array beam because elements within the core-array can be phased together or correlated individually rather than by first phasing them into station beams. Additionally, and most importantly, it is only for the core array that one can hope to sample the *entire* FoV of $\sim 1 \text{ deg}^2$ at 1.4 GHz at the high data rates needed for pulsar and transient applications (assuming that elements are coherently phased together).

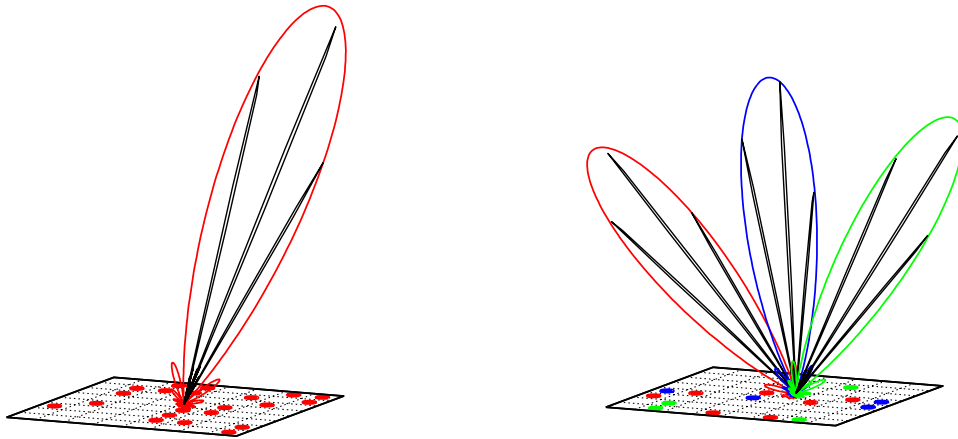


Figure 1. Illustrations of primary and phased-array beams. Left: primary beam (red) and three phased-array beams (black). The phased array could represent an outlier station for the SKA or it could represent the core elements of the overall SKA. Right: Illustration of individually steerable fields of view, which we identify with the primary beam. In practice, depending on the size of the array and on other factors, the FoV could be much smaller than the primary beam. In the example shown, the multiple FoVs are achieved through use of sub-arrays, as indicated by the color-coding of the beams and elements in the array.

2.3 Discussion of Multiple FoV Needs for Different Key Science Projects

At some level, all science areas would benefit from multiple, simultaneously available FoVs, because they would effectively provide larger solid angle coverage, and hence greater survey throughput. However, the FoVs specified in SKA Memos 45 & 69 already take into account throughput requirements for the KSPs. The key question for the present document is the simultaneity of multiple FoVs, especially if they are widely separated, for example, by tens of degrees on the sky. Such coverage can always be gotten through sub-array capability at less than full sensitivity. So the real issue is: what science areas require multiple, widely separated FoVs at full sensitivity?

2.3.1 The Cradle of Life

The three components of the Cradle of Life (imaging of protoplanetary disk evolution, inventories and chemistry of biomolecules, and SETI) do not require simultaneous, widely spaced FoVs at full sensitivity. Anti-coincidence

tests of radio-frequency interference (RFI) are crucial for SETI and thus SETI would benefit from the capability. However, multiple-site observations are at least as important for anti-coincidence tests and would most likely involve less-than-full sensitivity. The lack of knowledge of the luminosity function of ETI sources implies that we cannot estimate the sensitivity required to detect such sources.

2.3.2 Strong-Field Tests of Gravity Using Pulsars and Black Holes

The topic of this document was discussed at the August 2005 Pulsar SKA/KSP Project Workshop at the ATNF. Discussion of the pulsar KSP requires us to distinguish between the three observational components: searching, timing, and astrometry using long baselines. Briefly, the workshop discussion lead to the following assessment:

- *Searching*: Multiple FoVs are not essential but would be useful in terms of increasing survey throughput.
- *Timing*: Precision timing is needed at a frequency of about 2 GHz. FoVs separable by 10 to 45 deg are desirable or required, depending on the results of simulations that need to be done on simultaneous multiple pulsar timing. Note that this is an efficiency issue, but efficiency may be so low that the KSP is not feasible, keeping in mind that $\sim 10^4$ pulsars will be discovered in the search and these will all require follow-up timing of some sort in order to identify objects that will warrant long-term timing. For many objects, it may be possible to use sub-arrays for getting multiple FoVs. Simulations are required to really answer the question. Conducting the needed simulations is an action item.
- *VLBI astrometry*: Multiple FoVs are not necessary.

2.3.3 The Origin and Evolution of Cosmic Magnetism

There is no apparent need for simultaneous, multiple FoVs.

2.3.4 Galaxy Evolution, Cosmology and Dark Energy with the SKA

The massive surveys that comprise much of this KSP require large solid angle coverage and in fact drive the FoV requirement stated in SKA Memo 45 for frequencies below 1 GHz and down to about 0.4 GHz. However, there is no apparent requirement for multiple, widely spaced FoVs. The third component of this KSP consists of VLBI observations of water masers. The nominal high-frequency FoV will suffice.

2.3.5 Probing the Dark Ages with the SKA

The three components of this KSP are (1) high- z H I observations to detect the intergalactic medium (IGM) during the reionization stage; (2) detection of the first CO molecules in galaxies; and (3) surveying radio continuum emission from the first waves of star formation. The solid-angle requirements for each of these areas are already included in the SKA specifications of SKA Memo 45. Separate, simultaneous multiple FoVs are not needed to accomplish the science.

2.3.6 Exploration of the Unknown

The SKA will probe vast new regions of parameter space owing to its specified coverage of the frequency, time and spatial domains. Entirely new classes of objects are almost certain to be discovered. Transient sources represent one broad area in which the SKA will unveil new sources. A key factor for surveys of transients is instantaneous solid-angle coverage. As with the KSPs, any increase in the instantaneously sampled solid angle will improve survey throughput. Searches for fast transients require continuous dwelling on maximal solid angle, whereas slow transients (by definition) are those that can be surveyed through raster-scanning of the sky. In

either case, there is no compelling requirement for simultaneous, multiple FoVs. As for SETI (see above), anti-coincidence tests for RFI will be a necessary feature of transient searches and could benefit from multiple FoVs. But without further knowledge of the luminosity function for transients, it is not clear that such multiple FoVs would *require* full sensitivity and thus can be made using sub-arrays of the SKA.

2.4 Caveats

Accessibility considerations: Usage of the SKA will include deep integrations of particular fields lasting months or more. Accessibility of the SKA to other groups may thus require the existence of another FoV that can be pointed independently, particularly for projects that are time constrained, such as monitoring observations and target-of-opportunity cases. Sub-arraying capability may address these situations, and it remains to be seen whether full sensitivity is needed for more than one FoV.

Sociological considerations: an instrument as large as the SKA should serve and indeed requires a large user base. Also, there are numerous exciting ways to use the instrument for scientific outreach. Multiple FoVs would enable the availability of the SKA for these purposes. Depending on the demand for a FoV at full sensitivity for key science areas, the SKA could be made available at partial (and even very low) sensitivity. Sub-arrays may suffice but a thorough assessment should be made of situations that may require multiple FoVs at full sensitivity.

2.5 Summary

Overall, consideration of KSPs indicates that there is as yet no compelling need for a design that provides multiple fields of view at full sensitivity. However, such a need may emerge for the pulsar KSP. A decision for the pulsar KSP requires careful study of the process by which new pulsars discovered in a massive SKA survey will be timed efficiently. If a large fraction of the pulsars needs timing observations sustained over long durations (e.g., years), multiple FoVs are probably required to achieve the needed throughput. However, a winnowing protocol that allows us to rapidly select “important” pulsars for long-term timing may drastically reduce the array time needed for follow-up timing.

Other considerations, such as the general topic of Exploration of the Unknown, and the availability of the SKA for deep surveys, short-term projects, and outreach, suggest that the availability of multiple, independently steerable FoVs would be valuable. Careful study should be made of the following: (a) the sensitivity levels at which such FoVs need to be provided, and (b) the design and cost implications. In addition to the science requirements for the individual KSPs, study should be made of how the SKA would be scheduled to achieve multiple science goals.

3 WHAT IS THE IMPACT OF LIMITING THE FIELD OF VIEW AT HIGH ANGULAR RESOLUTION?

Dayton Jones, Bryan Gaensler, Chris Blake, Joseph Lazio, Kavilan Moodley, Ue-Li Pen

3.1 Introduction

It is generally recognized that imaging the maximum possible FoV at the highest possible angular resolution is likely to be impractical with the SKA. This is due to two separate concerns: (1) the data transport and processing rates implied by combining the full primary beam FoV of small antennas with pixel size determined by the maximum baseline lengths are currently unaffordable, and (2) the image dynamic range required increases as the field of view increases because the total range of source strengths within the field increases.

For wide-field survey observations, this difficulty can be avoided by using less than the maximum possible angular resolution. Full angular resolution observations are expected to concentrate on one (or a small number) of radio sources, and consequently the total area to be imaged can be a small fraction of the full primary beam area. Thus, limiting the FoV for the highest angular resolution observations does not impose a serious scientific penalty.

The most challenging key science project in this respect is the galaxy evolution, cosmology, and dark energy project. The large-scale HI survey aspect of this project requires an extremely wide field of view combined with sufficient angular resolution to avoid confusion, while the weak lensing/cosmic shear aspect requires a moderately wide field of view and angular resolution well under one arcsecond. Both of these combinations will be difficult to meet with current signal transport and processing concepts.

The ways in which data processing and imaging problems scale with FoV and angular resolution have been discussed in a number of recent memos (e.g., Carilli 2002; Perley and Clark 2003; Lonsdale 2003; Bunton 2004; Wright 2004; Cornwell 2004; Lonsdale, Doeleman & Oberoi 2004). This memo discusses the science requirements for wide FoVs at high angular resolution. In an appendix, we consider the pros and cons of several techniques for limiting the FoV, with emphasis on possible reductions in the imaging dynamic range that may occur within the reduced FoV.

3.2 Current Specifications

The science-driven specifications for the SKA are summarized in SKA Memos 45 & 69 (Jones 2004; Schilizzi 2006). The items most relevant to FoV issues are given below, in a slightly simplified form.

- Current FoV spec: 200 deg² at 0.1–0.3 GHz, 50 deg² at 0.3–1 GHz, 1–10 deg² at 1–3 GHz, and 0.33 $(3/\nu)^2$ deg² at 3–25 GHz.
- Current angular resolution spec: 0.02 ν^{-1} arcsec (maximum baselines at least 3000 km).
- Current correlator spec: 1 deg² at 1.4 GHz with 0.1 arcsecond resolution, and 200 deg² at 0.7 GHz with 0.2 arcsecond resolution.
- Current image dynamic range spec: 10^6 .

Note that the first two items listed here imply image sizes exceeding 10^{13} pixels (per spectral channel) at frequencies just below 1 GHz. However, it is not expected that the full FoV will ever be imaged with the full angular resolution at these frequencies; different science goals require one or the other, but not both simultaneously. The third item implies image sizes $\sim 10^{11}$ pixels. This, combined with the last item, represents the worst-case scenario for data processing.

3.3 Effect on Key Science Projects

For each of the SKA key science projects, we summarize the likely need for FoV limiting and possible detrimental effects.

3.3.1 The Cradle of Life

The imaging of protoplanetary disks demands milliarcsec resolution and a high observing frequency, but only requires a narrow, sub-arcminute, FoV. This is readily achievable. For SETI and astrochemistry observations a wide FoV is desirable, but high angular resolution is not needed. Use of the inner regions of the SKA configuration should be a good match for these observations.

3.3.2 Strong-Field Tests of Gravity Using Pulsars and Black Holes

The census of pulsars in the Galaxy requires a wide FoV, but only high enough angular resolution to avoid being limited by confusion. Pulsar timing requires neither very high angular resolution, very high imaging dynamic range, nor very wide FoVs, although an ability to have multiple (perhaps small) FoVs widely separated on the sky would be valuable (see discussion in §2). Full SKA sensitivity will not generally be needed for pulsar timing, so this is a good application of sub-arraying if full-sensitivity multi-beaming is not available. Pulsar astrometry requires the longest baselines, but only for very limited FoVs.

As was the case for the Cradle of Life, this key science project does not require very high angular resolution and a very wide field of view simultaneously. Consequently, it is not strongly affected by limiting the FoV at high angular resolution.

3.3.3 The Origin and Evolution of Cosmic Magnetism

Moderate angular resolution ($\approx 0''.1 - 0''.5$) and moderate FoV ($\gtrsim 1 \text{ deg}^2$) are needed at 1.4 GHz for the “all-sky Faraday rotation grid” and for deep polarization imaging of extragalactic fields. In these experiments the angular resolution is needed not just to prevent confusion, but to mitigate depolarization caused by structure both in the extragalactic sources themselves and in the foreground Faraday rotation screen. Imaging of polarization “silhouettes” requires $0''.1$ resolution, but only over a FoV $\approx 0.1 \text{ deg}^2$. Any FoV limiting for these experiments must permit high polarization accuracy across the imaging FoV (time-varying station beams may be particularly troubling here). These experiments also require a few hundred spectral channels, to allow accurate measurements of large rotation measures.

Most of the other experiments currently envisaged within this KSP involve either imaging of small discrete sources or studies of diffuse polarized emission at resolutions $\gg 1''$, and so do not imply demanding FoV requirements. Either the SKA core alone or imaging over a small FoV will be appropriate for these experiments.

3.3.4 Galaxy Evolution, Cosmology and Dark Energy

The science case for this project proposes large-scale structure surveys for redshifted HI, weak lensing (cosmic shear) continuum surveys, the cross-correlation of galaxies in different source and lens planes to probe cosmic magnification, and distance measurements and black hole mass determinations through H_2O maser emission from accretion disks in active galactic nuclei (AGN).

In order for SKA surveys to be competitive in studies of large-scale structure, they must cover essentially the entire observable sky to a redshift depth $z \sim 1.5$, which in turn demands a very wide FoV (at least 10 deg^2 , and preferably 10–20 times more). The angular resolution requirement for the large-scale structure survey is simply the avoidance of confusion. This should be possible even with angular resolution ~ 10 arcseconds because overlapping galaxies can be separated in frequency space. Even though relatively low resolution is sufficient for this experiment, data processing will be very challenging for FoV $\geq 100 \text{ deg}^2$ and a large number of spectral channels.

Weak lensing continuum surveys have stronger requirements in terms of pixels per image, because shapes can only be robustly measured for resolved galaxies. In this case, resolutions of $\approx 0''.3$ (and hence pixel sizes of $0''.1$) are probably needed. We note that this resolution needs to be achieved at the lowest possible frequency, to gain maximum benefit from the steep radio spectra of distant star-forming galaxies and reach the required source density of about $100 \text{ galaxies arcmin}^{-2}$. For a 10 deg^2 FoV, the required image sizes are a bit over 10^{10}

pixels. However, these will be continuum observations so far fewer spectral channels are needed than for the H I large-scale structure experiment.

Neither the cosmic magnification experiment nor observations of extragalactic H₂O masers require a very large FoV. In addition, cosmic magnification does not require high angular resolution. Astrometry of H₂O masers does require the very longest baselines.

The large-scale structure and weak lensing surveys appear to require images containing more than 10⁸ and 10¹⁰ pixels, respectively. They are closer in over-all data processing requirements when the larger number of spectral channels for the large-scale structure survey is included.

3.3.5 Probing the Dark Ages

Low angular resolution, wide FoV, observations are required to image the highly redshifted H I signal from the epoch of reionization. The dense core region of the SKA configuration will be needed, but not the longer baselines. It may be desirable to use long baselines to detect individual sources so their effects can be removed from the H I images, but this does not require very high resolution imaging of the confusing sources.

Surveys for CO emission from high-redshift star forming protogalaxies require angular resolution ≈ 50 milliarcsec over a FoV ≈ 10 arcmin². This implies about 10⁸ pixels/image, but also a large number of spectral channels. In this respect it is similar to the H I large-scale structure survey in terms of data processing. In practice redshifted CO observations are likely to be easier because the FoV per image is much smaller (even though the number of pixels is nearly the same) and both RFI and confusion are less of a problem at high frequencies than at low frequencies.

Continuum observations of the first AGN and starburst galaxies require angular resolution of a few milliarcsec and at least a moderately large FoV (~ 10 arcmin²). This will be challenging to achieve; it is likely that a more limited FoV will be necessary in order to obtain the required brightness temperature discrimination. One possibility is to combine an initial lower resolution survey with follow-up high resolution observations that imaged only isolated small regions within a large FoV. This approach requires a correlator capable of using multiple simultaneous phase centers, but it would greatly reduce the image processing requirements.

3.3.6 Exploration of the Unknown

The goal here is to expand observational parameter space as much as possible beyond what can be observed with EVLA and other near-future instruments. Both higher angular resolution and wider FoVs are valuable, but very wide (or multiple) FoVs are the more promising route to new discoveries. Sub-mas angular resolution has been explored, at least for high brightness temperature sources, but very wide FoV deep surveys have not yet been possible.

3.4 Conclusions

Table 1 provides a summary of the imaging and FoV requirements for the experiments covered by the five KSPs. The KSP experiments that place the most demanding requirements on combined angular resolution and FoV are the weak lensing experiment, which requires by far the most pixels per image, and the large-scale H I and EoR H I surveys, both of which require large images in a large number of spectral channels. The survey for high-redshift CO emission shares the requirement for large images over many channels, but benefits from having a smaller absolute FoV to deal with and a more benign sky at high frequencies.

Most, but not all, of the key science projects would be adversely affected by a reduction in image dynamic range. Consequently, whenever it is necessary to limit the FoV it will be important to use techniques that allow the highest dynamic range to be maintained.

Basic research in radio astronomy at the Naval Research Laboratory is supported by the Office of Naval Research. Part of this work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the U.S. National Aeronautics and Space Administration.

Table 1. Summary of Imaging Requirements

Key Science Project	Experiment	Angular Res. (arcsec)	Field of View (deg ²)	Pixels / Image / Freq. Channel
Cradle of Life	Planet formation	~ 0.001	< 0.001	~ 10 ⁷
Cradle of Life	SETI	> 1	≥ 1	Not imaging
Cradle of Life	Biochemistry	0.2	~ 0.1	~ 10 ⁷
Pulsars & Gravity	Survey	> 1	~ 1	~ 10 ⁶
Pulsars & Gravity	Timing	> 1	Multiple FoVs?	Not imaging
Pulsars & Gravity	Astrometry	< 0.005	< 0.1	Not imaging
Cosmic Magnetism	RM Grid	~ 0.5	~ 1	~ 10 ⁸
Cosmic Magnetism	Silhouettes	0.1	~ 0.1	~ 10 ⁸
Gal. & Cosmology	Large-scale H I	~ 10	≥ 50	≥ 10 ⁸
Gal. & Cosmology	Weak Lensing	~ 0.3	≥ 10	≥ 10 ¹⁰
Gal. & Cosmology	H ₂ O Masers	~ 0.001	<< 0.1	~ 10 ⁴
Dark Ages	H I during EoR	> 1	> 1	~ 10 ⁸
Dark Ages	High-z CO	~ 0.05	~ 0.1	~ 10 ⁸
Dark Ages	High-z AGN	~ 0.005	~ 0.01	~ 10 ⁸

REFERENCES

- Bunton, J., 2004, “SKA Correlator Input Data Rate”, SKA Memo 43.
- Carilli, C., 2002, “Number of Cross-Correlations Required to Synthesize a 1 sq. degree Field of View”, SKA Memo 24.
- Cornwell, T. J., 2004, “SKA and EVLA Computing Costs for Wide Field Imaging (Revised)”, SKA Memo 49.
- Cotton, W. D., 1994, “Wide-field Polarization Correction of VLA Snapshot Imaging at 1.4 GHz”, AIPS Memo 86.
- Jones, D. L., 2004, “SKA Science Requirements: version 2”, SKA Memo 45.
- Lonsdale, C. J., 2003, “Data Rate and Processing Load Considerations for the LNSD SKA Design”, SKA Memo 32.
- Lonsdale, C. J., Doleman, S. S., Oberoi, D., 2004, “Imaging Strategies and Post-processing Computing Costs for Large-N SKA Designs”, SKA Memo 54.
- Perley, R., and Clark, B., 2003, “Scaling Relations for Interferometric Post-Processing”, EVLA Memo 63.
- Walker, R. C., 1984, “Non-Closing Offsets on the VLA”, VLA Scientific Memo 152.
- Wright, M., 2004, “SKA Imaging”, SKA Memo 46.

3.5 APPENDIX: Techniques for Limiting the Field of View

There are at least six ways to limit the FoV of an aperture synthesis observation. They are listed here in approximate order of the angular scale on which they are effective:

1. Response of primary antenna beams
2. Station-level beam-forming
3. Averaging over time and frequency during correlation

4. Weighting of individual visibility points during correlation
5. Post-correlation averaging in time and frequency
6. Post-correlation weighting of visibilities

We will discuss each of these techniques in order.

First, and most fundamentally, the FoV is limited by the angular response of the primary antenna elements in the array. For high-gain, steerable antennas this can be extremely effective and may be the only technique required at high frequencies. For low-gain antennas, of course, this is much less effective (although even nearly isotropic antenna elements will have some FoV limits caused by interactions with the ground and with neighboring elements). The main difficulty with primary beam FoV limitation is that relatively small changes in beam direction can cause a time-varying response to radio sources near the edges of the beam, and small changes in pointing are difficult to avoid for large, mechanically steered structures. For small antennas the primary beam is large enough that it will not adequately limit the FoV for high resolution observations.

Second, station-level beam-forming has been proposed to reduce the long-distance data transmission bandwidth for arrays with large numbers of small antennas. This is essential for arrays with very large numbers of low-gain antenna elements (low frequency arrays and the phased-array tile concept for the SKA), and is also part of the Large-N Small-D (LNSD) SKA concept. The resulting station beam will have time-varying sidelobes, and for stations with only a few antennas (as in the LNSD concept) the instantaneous sidelobe level will be high. Consequently the response to distant sources will be significant, and one or more of the other FoV limiting techniques listed here will also be needed.

Third, time and frequency averaging of data during cross-correlation is a long-standing way to reduce the response to sources far from the phase center. At some level this type of FoV limiting is unavoidable because of finite sampling of the data. It can be effective on long baselines, but less so on short baselines. The attenuation is not uniform at a given distance from the FoV center, and varies with both time and baseline. As a result, it does not provide a constant, large attenuation of distant sources.

Fourth, the individual visibility points can be weighted by their location in the $u - v$ plane during correlation. This technique has not been applied to real data yet, but promises to provide a much cleaner and well-controlled limit to the imaging FoV. Unlike traditional time and frequency averaging, correlator field-of-view tailoring (Lonsdale et al. 2004) can provide an attenuation of more than 20 dB everywhere outside of a well-defined central region. However, it does decrease point-source sensitivity.

Fifth, additional averaging over time and frequency can be applied to data after correlation. Post-correlation averaging provides more flexibility than averaging during correlation, and in principle different averaging intervals could be applied to different baselines to improve the attenuation of distant sources. However, there will always be some variations in the attenuation provided by this technique, because any given baseline responds to sources at multiple locations that, for a given set of averaging parameters, will have different levels of decorrelation. Consequently the effectiveness of this technique will be limited.

Sixth, the visibility data can be weighted by $u - v$ location after correlation. This is often done to apply a smooth tapering function to reduce the weight of the outer region of $u - v$ coverage and thus reduce the effective angular resolution and increase surface brightness sensitivity (at the expense of point-source sensitivity). However, it is possible to apply a tapering function that reduces the weight of the inner $u - v$ region instead, effectively limiting the response to large-scale emission and (when combined with time and frequency averaging) reducing the response to distant compact sources as well.

3.5.1 Effect on Dynamic Range

The primary concern for each of these techniques is the extent to which they may limit the imaging dynamic range that can be achieved within the reduced FoV. This could occur if a technique introduced variations in the system response that change faster than the signal-to-noise-ratio based time scale for self-calibration corrections, or increased the number of calibration parameters that needed to be solved for, or increased the size of non-closing

errors (see Walker 1984), or increased sensitivity to RFI, or significantly reduced the number of independent visibility measurements.

Primary beam attenuation has always been with us, and has not prevented existing small- N array from achieving dynamic ranges exceeding 10^5 . For phased arrays with many small elements the primary beam (element FoV) is so large compared to the imaging FoV that “edge effects” are likely to be negligible. This technique has no effect on the number of visibility measurements or on non-closing errors.

Station-level beam-forming will clearly produce a time-varying response to sources outside of the FoV. The extent to which this affects dynamic range is not yet known, but can be empirically tested in the near future with the Allen Telescope Array (ATA) and large-N low frequency arrays. For large-N stations it is likely that the instantaneous station beams will have a stable main beam profile and low sidelobe levels, so station calibration will be much less of a concern than it is for small-N station beam-forming. Note that station beam-forming effectively produces one visibility measurement per station per unit time from the N independent (but nearly identical) measurements that are potentially available from the individual antennas. In this sense it reduces the number of measurements that can be used to constrain calibration. For small-N stations the response across the central part of the station beam will vary with time, potentially requiring additional position dependent self-calibration parameters. In all cases it will be important to minimize the time-varying polarization response across the station beams. Calibration of a time varying polarization response across the FoV can be applied to snapshot images (e.g., Cotton 1994), but this will be more difficult for long integrations or for station beams with very different polarization responses at a given time.

Time and frequency averaging during correlation is not a desirable technique from a dynamic range point of view, although it may be necessary at some level to reduce the correlator output data rate. The varying response to distant sources as a function of time and baseline will be difficult to correct for accurately. This technique also reduces the number of visibility samples that can be used for calibration and obscures possible low-level bandpass mismatches that could produce non-closing errors.

Correlator weighting by $u - v$ location is very promising because the smooth and deep attenuation that it can produce should minimize time-variable effects, and because we have precise control and knowledge of what is done to the data at a very fine sampling level. It maintains the number of independent visibility measurements, does not introduce either amplitude or phase errors, and should have no detrimental effect on imaging dynamic range. It does reduce point source sensitivity. This technique does not reduce the correlator data output rate, although traditional time and frequency averaging can probably be used with fewer ill effects after correlator FoV tailoring has been applied. This technique does require additional computation during correlation and will necessarily have some impact on detailed correlator design.

Post-correlation time and frequency averaging should have little effect on dynamic range if the resulting decorrelation occurs well outside of a FoV that is already limited by one or more of the other techniques listed here, and if the bandpass responses are well matched on the scales being averaged over. If too much averaging is applied, the same time and position dependent effects as described for data averaging during correlation will occur.

Post-correlation data weighting by $u - v$ location should not reduce dynamic range. Like all data weighting techniques, it will reduce point source sensitivity.

3.5.2 Conclusions

Techniques that limit the FoV in a highly time varying manner, like station beam-forming, need to be tested with real data to determine their effects on both image dynamic range and polarization response over the reduced FoV. The ATA should be useful for this in the LNSD context, where station sidelobe levels are particularly large. Station sidelobe levels will be much lower for large-N phased array tiles, but geometric projection will produce time-dependent response variations across the FoV. The large adaptive reflector / focal plane array approach offers potentially “clean” station beams, but projection will still introduce some changes over time and rapid small-scale changes in the station beams caused by reduced mechanical rigidity may be difficult to calibrate.

Correlator visibility weighting by $u - v$ location, as proposed by Lonsdale et al. (2004), appears to be the most

effective FoV limiting technique as well as the one least likely to reduce imaging dynamic range. It is not effective over an arbitrarily wide range of baseline lengths, but can be combined with primary beam attenuation and the response of the full array synthesized beam to provide very large attenuation of distant sources. We recommend that tests of this technique be carried out with real data (both connected-element and VLBI) as soon as practical, and that correlator designers consider how this technique could be incorporated into the SKA correlator.

4 WHAT IS THE CASE FOR HIGH ANGULAR RESOLUTION BELOW A FEW GHZ?

Raffaella Morganti, Melvin Hoare, Mike Garrett
(with additional input from H. Falcke, S. Rawlings and T. Oosterloo)

4.1 Introduction

The technical specifications for the SKA call for an angular resolution of $0''.1$ at 1.4 GHz. This value was proposed as the *minimum* acceptable angular resolution that would avoid the detrimental effects of source confusion in deep extra-galactic surveys. At 1.4 GHz, an angular resolution of $0''.1$ corresponds to a maximum baseline of 400 km.

This document explores the justification in the SKA science case for baselines of 1000 km or more, at frequencies of less than a few GHz. We discuss the various research areas that a high resolution/low frequency SKA configuration might address, considering each of the KSPs in turn.

4.2 KSP I: The Cradle of Life

The study of planet forming disks via thermal emission is exclusively confined to high frequencies. Most biomolecules of interest also emit at the high frequency end of the SKA range. The loss of high angular resolution at 1.4 GHz would preclude complementary studies of atomic hydrogen in proto-planetary disks, but that is not part of KSP I.

SETI searches will be carried out in the 1–10 GHz regime with an emphasis on frequencies near the “waterhole”, 1.4–1.7 GHz. The SKA should be capable of detecting “leakage radiation” from ETI transmitters out to ~ 300 pc (Tarter et al. 2004) and more powerful and directed ETI beacons out to 10 000 pc. If placed at a distance of 300 pc, the Solar System would subtend an angle of only $\sim 0''.2$. High resolution observations of low-frequency ETI beacons might play an important role in pin-pointing the location of planets harboring ETI, orbiting around relatively nearby stars. A comparison with observations made at other wavebands would be useful in understanding the general conditions under which intelligent life forms can develop and prosper.

4.3 KSP II: Strong-field Gravity Tests

Large-scale surveys and timing measurements associated with the proposed tests of general relativity via pulsars do not require high angular resolution. However, the astrometric measurements needed to determine their proper motion and parallax do require the highest angular resolution possible. These are necessary to provide kinematic corrections to the timing measurements in those sources found suitable for strong-field gravity tests. Proper motions and parallaxes will also allow tests of supernova collapse models and will map out the ionized medium distribution in the Galaxy. Even though Pulsars have a steep negative spectral index, the astrometric measurements are most likely to be carried out at frequencies higher than 2 GHz, in order to avoid errors introduced by ionospheric fluctuations. Hence, there is little impact on this KSP if the SKA does not include the provision of long baselines at low frequencies.

4.4 KSP III: Cosmic Magnetism

The construction of an all-sky rotation measure grid at 1.4 GHz only requires a resolution of about $1''$ and therefore does not need long baselines. This resolution is also sufficient for the Faraday tomography of diffuse emission in the Milky Way. The magnetic fields in intermediate redshift galaxies can be seen as silhouettes against polarized extended background sources, but these galaxies are still sufficiently large that the highest angular resolution is not vital in mapping out the global features.

Direct imaging of the polarization pattern in more distant spirals would clearly benefit from higher angular resolution in order to resolve the disk structure of these systems. At higher redshift still, the deeper integrations necessary to look for polarization in unresolved sources may require longer baselines to minimise confusion problems.

4.5 KSP IV: Galaxy Evolution, Cosmology and Dark Energy

Dark Energy, HI Surveys

The all-hemisphere HI survey of galaxies out to $z \sim 1.5$ planned for this KSP will permit the detailed investigation of dark energy, and only requires arcsecond resolution. However, “orthogonal” constraints determined via weak lensing observations (see, e.g., Figure 7 of Rawlings et al. 2004), may require much higher angular resolution to provide the necessary intrinsic shape information on the faint (μJy & nJy) background radio source population. The current nominal SKA configuration would permit $0''.1$ resolution, and it is unclear whether this is sufficient or not for weak lensing studies. The precision measurement of Hubble’s constant using water masers is performed at high frequencies, and so there is no impact on this aspect of the programme.

Deep Continuum Surveys & Confusion

In addition to detecting vast numbers of HI systems, the SKA can also perform (in parallel) very deep radio continuum surveys at ~ 1 GHz. A typical eight-hour observing run should reach $1\text{-}\sigma$ rms noise levels of ~ 20 nJy/beam (assuming a nominal total bandwidth of 700 MHz [dual polarization] and a system temperature of 25 K).

Reaching such depths in only a few hours of observing time raises the question of how quickly confusion noise will begin to limit the depth that continuum surveys can attain. The standard rule of thumb is that confusion becomes important when the source counts reach an angular density of one source per 30 beams (see Hogg 2000 for a detailed discussion), the exact number depending on the steepness of the $\log N - \log S$ function. We can estimate the depth at which confusion will become a limitation for the standard SKA specification by assuming a steep ($\alpha = -0.7$) spectral index for the bulk of the faint sub- μJy and nJy radio source population, and extrapolating the source counts determined via recent deep field studies (e.g., Richards et al. 2000). With these assumptions, the SKA sky at 1 GHz will reach an angular density of one source per 30 beams (assuming a standard baseline length of 400 km) at a flux density level of ~ 60 nJy , i.e., at the $3\text{-}\sigma$ rms noise level. Low-frequency radio continuum surveys (centered around 1 GHz) run the risk of being confusion limited on time scales that approach a typical 8 hour observing run.

If a low-frequency SKA instrument also includes a simultaneous multiple field-of-view capability, then much deeper (longer integration) radio surveys should be commonplace. Even without a multiple field-of-view capability, integration times of many days (possibly weeks) ought to be possible, in principle $1\text{-}\sigma$ rms noise levels of a few nJy per beam may be attained. At a flux density level of 10 nJy , extrapolation of the μJy source counts suggests a baseline length of 1400 km is required in order to maintain an angular density less than one source per thirty beams. Thus, for ultra-deep, low frequency surveys, a baseline length of ~ 1400 km (corresponding to an angular resolution of $0''.04$) may be required just in order to avoid confusion effects (see also Garrett 2002).

It should also be noted that deep surveys, conducted at even lower frequencies (e.g., 300 MHz), will reach an angular density of one source per 30 beams at a flux density of only 0.8 μJy for a baseline length of 400 km (see also Kellermann & Richards 2000). Such sensitivity levels should be attained by the SKA; indeed they are already within the grasp of other pathfinder instruments such as LOFAR.

The Scientific Justification for Long Baselines at 1 GHz

One of the motivations for conducting deep surveys at low frequencies is that it is expected that the dominant radio source population at these faint flux density levels will be optically thin, non-thermal emitters with steep spectral indices. The fact that many of these sources will also be located at high redshift ($z > 5$) makes low-frequency survey observations particularly advantageous.

Moderate ($0''.1$) resolution radio continuum observations can be used to determine the global characteristics of star formation over the history of the Universe, but another key goal of KSP IV includes the detailed study of the morphology of complex star-forming regions located within galaxies at high redshift. Here the SKA can bring to bear the highest resolution possible at any wavelength, especially for regions that are dust obscured.

The study of nearby star-forming systems such as Arp 220 suggests that linear resolutions better than 500 pc will be required to begin to resolve regions of intense star formation in such systems. At $z \sim 2$, 500 pc subtends an angle of ~ 60 milliarcseconds, requiring a baseline length of at least 1000 km at 1 GHz. Individual hypernovae (such as those observed in Arp 220) should be detectable out to cosmological distances with the

SKA — separating these hypernovae will typically require much better angular resolution, ~ 10 milliarcseconds, corresponding to a baseline length of 3000 km. A detailed knowledge of the radio morphology of the earliest forming galaxies will provide important information on the role of both mergers and AGN in galaxy formation. Identifying AGN within these systems will require milliarcsecond resolution, but for flat-spectrum sources this may be better achieved at higher frequencies.

Both the science case (and related confusion issues) strongly argue for a significant fraction of a low-frequency SKA to be distributed over baselines of up to several thousand km in extent. Longer baselines will inevitably lead to a sparse SKA array, but the $u - v$ coverage will still be good given the large fractional bandwidths employed. Indeed, the incorporation of longer baselines might alleviate image dynamic range limitations — the brightest and most troublesome confusing sources at the edge of the beam being largely resolved on baselines longer than a few thousand km.

4.5.1 Neutral hydrogen in nearby and high- z AGN

Observations of associated HI around AGN has helped in understanding their nuclear structure and the environment of the central black hole. The study of the neutral hydrogen in absorption against radio-loud AGN has shown that this gas can have a variety of structures distributed on scales ranging from a few, to a few hundred pc. A summary of the most recent results is given by Morganti et al. (2004).

- HI outflows: outflows are now recognized to represent a major source of feedback, and therefore play an important role in galaxy evolution as they have a major effect on the interstellar medium (ISM) in forming galaxies. There is a growing number of cases of fast and massive HI outflows detected (see Morganti et al. 2005). Only a few cases of HI outflows have been imaged so far. They are often off-nuclear (kpc scale), but some could also be connected to broad-line regions (pc scale). In general, it would be crucial to resolve the kinematics of the outflow (e.g., looking for gradients in velocity, etc.) and compare this with the structure of the ionized gas to understand their physical conditions.
- Tori and/or circumnuclear disk structures: these are seen on scales of a few to a few tens of pc. One of the best examples is the case of Cygnus A (Conway 1999), where a structure about 50 pc in extent has been detected against the counter-jet and a velocity gradient (indicating rotation) has been measured. Confirming the presence and studying the characteristics of these structures in more objects are crucial for our understanding of AGN, because these HI disks give a handle on the physics of the nuclear disks associated with AGN.
- Infalling gas: in a few cases the HI gas has been claimed to be the component feeding the AGN (as proposed by van Gorkom et al. 1989), but nothing is known about location and structure.
- Off-nuclear clouds: even in cases in which the HI absorption appears at the systemic velocity, the actual location of the gas can be off-nuclear and may not be at all related to disk structure, but may just trace the ISM around the source. A dense ISM surrounding an AGN represents an important phase in the evolution of AGN.

The SKA will permit us to dramatically expand our knowledge on the occurrence of neutral hydrogen around AGN; HI absorption with optical depths of the order of $\tau \sim 0.01$ is typical of the absorption that we see with the current instruments in powerful radio sources (where τ is related to the measured depth of the HI absorption, ΔS , and the covering factor c_f as: $\Delta S = S c_f [1 - e^{-\tau}]$ where S is the continuum flux density). With the SKA, this will be detectable in sources as faint as a few mJy, two orders of magnitude fainter than with current instruments. The number of sources available for study will increase most dramatically. For example, all the sources in the NRAO VLA Sky Survey (NVSS) with redshifts up to $z = 6$ can be searched for HI in absorption.

The implications of this huge step forward in understanding the nuclear regions of different type of AGNs are enormous, but they rely also on the possibility of spatially resolving the absorption and therefore imaging the location of the absorption and its characteristics.

Table 2. Parameters for 1000 and 5000 km baselines

z	pc/mas	Freq H I	1000 km baseline		5000 km baseline	
			resolution mas	linear size pc	resolution mas	linear size pc
0.05	0.965	1352.8	23	22	5	4
0.5	6.082	946.9	33	198	7	40
1.0	8.041	710.2	44	350	9	70
2.0	8.475	474.0	65	554	13	111

The complexity and scales of these structures justify the requirement for high-resolution capabilities for the SKA, to identify the location, as well as the structure of the associated absorption.

The information about the size of the structures is derived from a limited number of high-resolution studies at low redshift. In order to understand the origin and properties of the neutral gas around AGN, we need not only to locate it, but also to spatially resolve it, so that we can derive information about the kinematics of the gas. *This implies that a linear resolution of the order of at most a few tens of pc is needed.* This needs to be compared with the numbers in Table 2, where we summarize the spatial resolution expected for 1000 and 5000-km baselines. It is clear that baselines of 1000 km are not enough if we want to study the structure of H I for sources at $z > 0.5$.

Furthermore, a very important aspect is that the high-resolution capability will open up the study of the evolution of gaseous structures with redshift, something that is now almost impossible. So far, studying this outside the classical 20 cm has only been possible thanks to observations with a UHF ($\sim 0.5 - 1.2$ GHz) system. This, however, has brought mainly information about the occurrence of H I absorption (in particular in compact radio sources). Follow-up at high resolution to learn about the properties (morphology and kinematics) of gas in these sources is very difficult at the moment, given the limited number of telescopes with UHF systems (see, e.g., Vermeulen et al. 2006). Only a few attempts have been made and indeed detections (at frequencies below 1 GHz) on Europe to Greenbank (5000 km) baseline have already been made. However, a systematic study showing the evolution of the gas distribution with z is not currently possible.

4.6 KSP V: Dark Ages

The nature of the first sources of radiation that were responsible for reionizing the Universe and bringing an end to the so-called “dark ages” is not known. Possible candidates include the first generation of massive stars and black holes.

4.6.1 First Black Holes

The key science case for the dark ages contains the goal of identifying the first generation of black holes.

One of the clearest signs of black hole activity is the presence of a compact radio core in the nucleus of a galaxy. While in the past the focus had been on the few bright and relativistically beamed sources, new surveys now show that essentially all black holes produce compact radio emission that can be used effectively for large radio surveys. Radio has the advantage of not being affected as much by obscuration.

The SKA will be able to observe the faintest and possibly first AGN up to very high redshifts in the radio. It will also be able to see almost every supermassive black hole in the nearby Universe and allow us to study the cosmological evolution of the main AGN parameters.

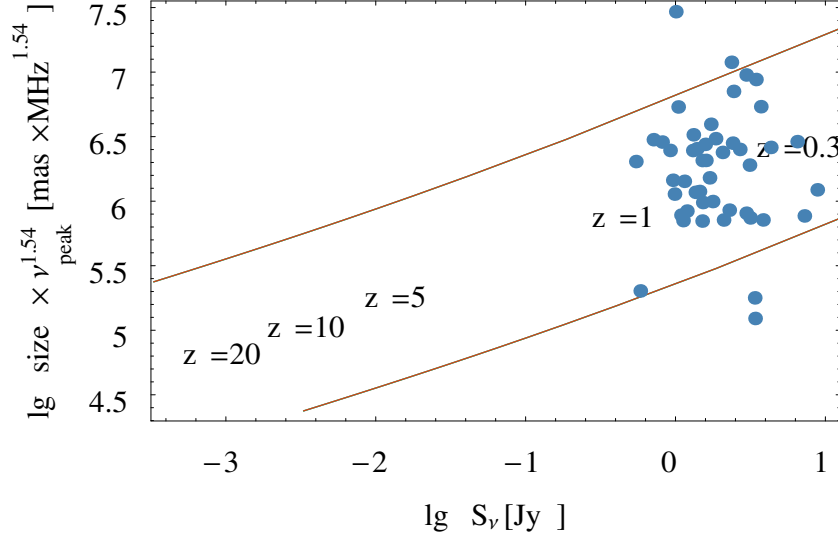


Figure 2. Size, frequency, and flux density roughly form a fundamental plane for GPS radio galaxies. Here we plot a combination of frequency and size ($\nu_{\text{peak}}^{1.54} \times r$) vs. the flux density for a range of empirical GPS models and actual GPS sources from O’Dea (1998). Size and frequency for the model are coupled according to the relation between size and turn-over frequency discussed by Falcke et al. (2004) for sources with different sizes of 0.1 (blue), 1 (green), and 10 kpc (red) and fall on top of each other. The two lines delimit the scatter in the empirical relation. The labels indicate the typical redshift of sources in that regime. Standard GPS sources are found in the redshift range around $z = 1$. The bottom left corner is not occupied and is the discovery space for young high-redshift black holes with the SKA.

One way to identify high redshift black holes has been discussed by Falcke et al. (2004). The first generation of black holes may have jets that are frustrated in their dense environment and thus appear as gigahertz-peaked spectrum (GPS) sources. Since their intrinsic size and peak frequency are related, and angular size and frequency scale differently with redshift, there is a unique region in parameter space that should be occupied by emerging black holes in the epoch of reionization. This can be well probed by radio-only methods with the SKA.

One way of doing this is to look for GPS-like sources at $z = 10 - 20$. These can, for example, be uniquely identified by a combination of turnover frequency and compactness (see Falcke et al. 2004) and this calls for a *broad bandwidth around 600 MHz and some tens of mas resolution. Therefore, 1000-km baselines are not enough.* This is described in more detail in the SKA science book.

In summary, a survey strategy for finding the first black holes would then be to:

- conduct a shallow all-sky multi-frequency survey in the range 100–600 MHz down to 0.1 mJy at arcsecond resolution;
- identify compact, highly peaked spectrum sources in that frequency range;
- identify empty fields in the optical;
- reobserve to exclude variable sources;
- observe with long baselines and resolutions of ~ 10 mas to determine sizes and to pick out the ultra-compact low-frequency peaked sources;
- spectroscopically confirm remaining candidates with HI observations or by other means.

In general, with the huge number of radio sources in our surveys, we will need the resolution to also pick out AGN populations and distinguish source types.

The First Generation of Massive Stars

Simulations of the collapse of metal-free clumps in the early Universe suggest that the first generation of stars must have been very massive, with masses of $\sim 100 - 1000 M_{\odot}$. Such stars will evolve rapidly, exploding catastrophically as hypernovae (an extreme form of supernova) and gamma-ray bursts after only a few million years. The SKA should be able to detect the death throes of such massive systems as the ejecta interacts with the dense, surrounding star-forming medium. This is expected to be a dynamic event in these over-dense regions; accurate astrometry will be required in order to determine the extent of these regions, again arguing for baselines of 3000 km in order to resolve clusters of these objects. The detection of HI absorption in these systems may be the only way of securely establishing the redshift of these systems.

4.7 Exploration of the Unknown

As a general consideration, it should be kept in mind that given the timescales expected for the SKA to be operational, it is important to consider the progress that will occur in astronomical instrumentation at other wavebands (Garrett 2002). A significant driver for future ground- and space-based next-generation telescopes (e.g., VLTI, NGST, ELTs, XEUS) is to provide high angular resolution capability at these wavebands for the first time. Given that the 1.4 GHz frequency range will be optimal for studying the high redshift Universe, there is a strong case for ensuring that the SKA can at least match these next generation instruments in terms of spatial resolution. Accurate astrometry, reliable identification and complementarity with these new instruments will be essential, in order to ensure that the SKA can also react to new discoveries made by other next generation telescopes. This again argues for baselines extending towards several thousand kilometres.

4.8 Summary

Table 3 summarizes the impact of the loss of a high angular resolution capability at frequencies below 2 GHz, i.e. in a “SKA-low” or even “SKA-mid” only implementation. It can be seen that the impact is limited to only sub-sections of two of the KSPs, and as such could be seen as fairly limited impact. However, the weak lensing results are a crucial orthogonal measure of cosmological parameters that would weaken the results from a stand-alone redshifted HI survey that does not require high spatial resolution. Studies of the normal, star-bursting and AGN sub μ Jy populations would be severely impaired by confusion and the inability to resolve structures and isolate counterparts without long baselines at low frequencies. Very long baselines are vital to identify the first black holes.

REFERENCES

- Conway J.E., 1999, in *Highly redshifted radio lines*, Carilli C.L. et al. eds., ASP Conf. Series 156, 259
Falke, H. et al. 2004, *NewAR* 48, 1157.
Garrett, M.A. 2002, 6th EVN Symposium Procs., p. 189 (astro-ph/0206270)
Hogg, D.W. 2000, *AJ* 121, 1207.
Kellermann, K. & Richards, E.A. 200, *Perspectives on Radio Astronomy Procs.*, p 157 (astro-ph/9909083)
Morganti, R. et al., 2004, *NewAR*, 48, 1195.
Morganti et al. 2005, *A&A*, 444, L9
O’Dea C.P. 1998 *PASP*, 110, 493
Rawlings, S. et al. 2004, *NewAR* 48, 1013
Richards, E.A. 2000, *ApJ* 533, 611.
Tarter, J.C. 2004, *NewAR*, 48, 1543
van Gorkom et al. 1989 *AJ*, 97,708
Vermeulen et al. 2006, *A&A*, 447, 489

Table 3. The impact on the SKA KSPs of not having long (~ 1000 km) baselines below 2 GHz. X means the science cannot be done at all, x means there is some impact.

KSP	Science	Impact
I	Planet formation	
	SETI	
	Biomolecules	
II	Pulsar searches	
	Pulsar timing & astrometry	
III	RM Grid	
	Deep fields	x
	Silhouettes	
	Polarization mapping	x
IV	Tomography	
	Redshifted H I	
	Weak lensing	X
V	Hubble Constant	
	HI from EoR	
	CO from proto-galaxies	
	Deep continuum fields	X

5 WHAT IS THE SCIENCE CASE FOR FREQUENCIES BETWEEN 200 AND 500 MHz

Chris Blake, Frank Briggs, Nissim Kanekar, Elaine Sadler, Thijs van der Hulst,
with additional input from Rainer Beck, Luigina Feretti, Bryan Gaensler & Michael Kramer

5.1 Summary

In this memo, we discuss the case for operating the SKA in the 200–500 MHz frequency range. Although the case for SKA to observe at 200–500 MHz is presented in more detail below, it rests so strongly on observations of redshifted neutral hydrogen that it can be summarized in very simple terms:

If we wish to understand how large-scale structure develops, and how gas is assembled into galaxies from the Epoch of Reionization to the present day, then the SKA must be able to observe H I in emission over the full redshift range permitted by its sensitivity ($0 < z < 4$), and in absorption over at least the redshift range $0 < z < 6$. In other words, the frequency range 200–500 MHz is essential for the SKA to capitalize on one of its greatest strengths: its unique ability to trace both the large-scale distribution of mass in the Universe and the transformation of H I gas into stars over a large fraction of cosmic time.

5.2 SKA KSPs and the 200–500 MHz frequency range

Two of the five SKA KSPs require the SKA to operate below 500 MHz. These are KSP IV (Galaxy Evolution, Cosmology and Dark Energy), for which observations at 200–500 MHz are identified as essential, and KSP V (Dark Ages), which requires observations below 200 MHz.

Two other KSPs (II. Gravity and III. Cosmic Magnetism) would also benefit from observations at 300–500 MHz, though in both cases (as discussed below) the main science goals could still be carried out with an SKA working only at 500 MHz and above. The remaining KSP (I. Cradle of Life) does not require the SKA to observe below 1 GHz.

The present document is therefore almost entirely concerned with science issues related to KSP IV (Galaxy Evolution, Cosmology and Dark Energy), though we also discuss briefly KSPs II and III.

5.3 KSP IV. Galaxy evolution, Dark Energy and Cosmology

5.3.1 H I emission-line surveys at $1.8 < z < 4$ (250–500 MHz)

Tracing galaxy evolution through observations of neutral hydrogen has always been a prime driver for a next-generation radio telescope. In fact, a square kilometer of collecting area is necessary to trace H I to high redshifts, and this requirement led to the SKA acronym.

To be effective at probing the epoch of galaxy assembly, the SKA must have frequency coverage at frequencies below 500 MHz. A number of tracers of galaxy evolution indicate a shift in the nature of galaxy formation from active and vigorous assembly to a more passive, settled evolution at redshifts around $z \sim 2$. To probe the regime where primordial clouds and primitive protogalaxies are assembling, the SKA must access redshifts of 2 to 5.

A sampling of the evidence is shown in Figures 3 & 4, which summarize the behaviour of four tracers of evolution. Of great interest to radio astronomers is the amount of neutral gas in the Universe as a function of cosmic time, since this is the building block from which stars, galaxies and the heavier elements are formed. Evidence from the study of damped Lyman- α (DLA) absorption lines in the spectra of high redshift quasars shows that there was more H I at early times than at present. It is thought that this decline in H I is at least partly responsible for the decline in the cosmic star formation rate with time (see Figure 4).

The number density of luminous quasars (and bright radio galaxies) shows a peak near redshift 2–3 (see Figure 3), implying that the redistribution of mass that feeds the nuclei of these monsters was most effective at these times. Also of interest is the onset of metal-rich halos (detected in large equivalent width, C IV quasar

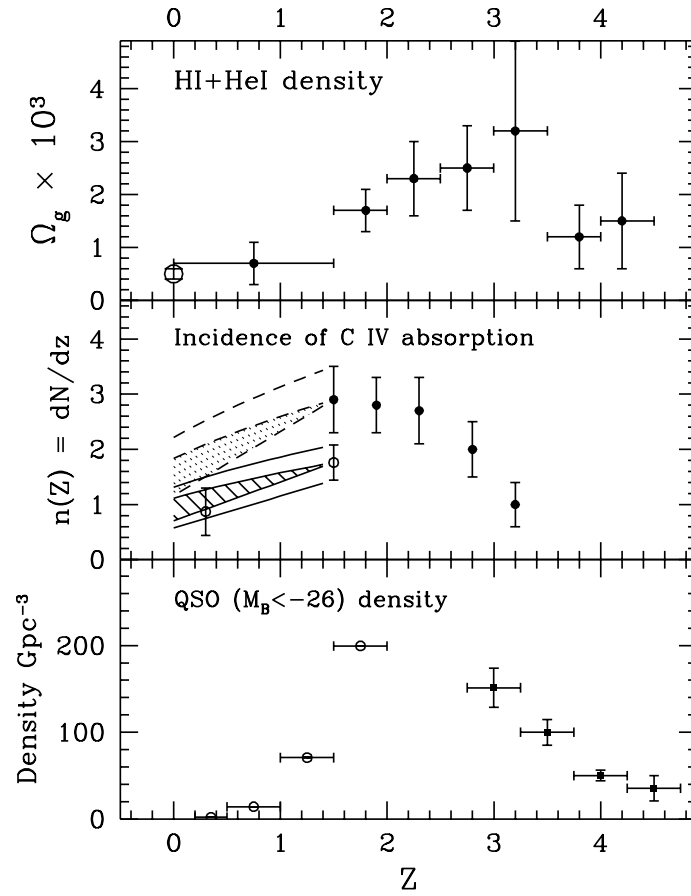


Figure 3. Cosmological density of neutral gas, incidence of C IV absorption, and co-moving density of luminous quasars as a function of redshift, from Briggs (1997).

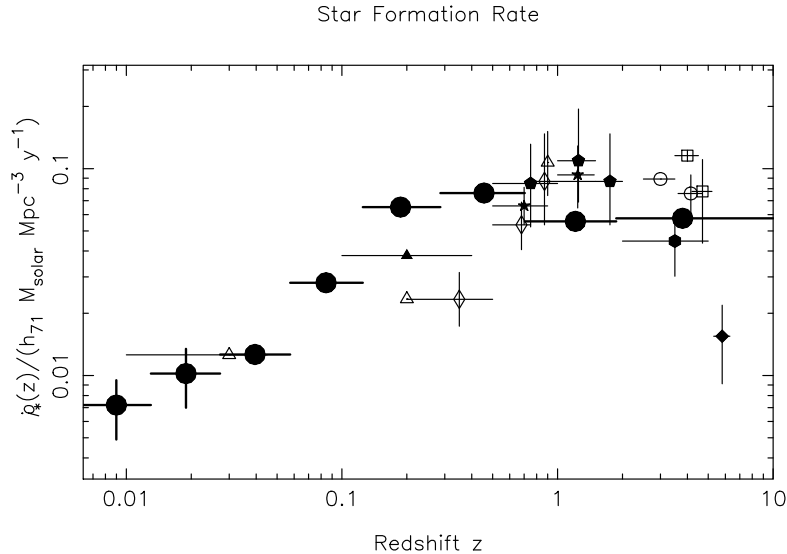


Figure 4. The star formation history of the Universe, from Heavens et al. (2004). The star formation rate recovered from the fossil record of galaxies in the Sloan Digital Sky Survey is shown by the eight large filled circles. The other symbols correspond to independent determinations using instantaneous measurements of the star formation rate.

absorption lines, for example); these appear in the redshift range 3.5 to 2, but their subsequent behaviour is consistent with no evolution in cross section since that time. The implication is that the vigorous gravitational interactions and strong stellar winds of violent accretion have subsided. Evidence on the evolution of the galaxy luminosity function since $z = 1.5$ (e.g., Cohen 2002) favors passive evolution of the galaxy population in the period since redshift 1.5.

The overall evidence points to a gathering crescendo in the period prior to redshift 2, with a subsequent relaxation and settling since then as galaxies take their current form and the degree of interaction and mass transfer subsides. Figure 5 shows that deep integrations with SKA should detect the H I emission line in “normal” galaxies to redshift $z \sim 3$ and beyond. *It is imperative that the SKA cover the frequency range below 500 MHz (i.e., redshifts greater than 2) in order to study this critical period in the evolution of galaxies.*

5.3.2 H I absorption-line surveys at $1.8 < z < 6$ (200–500 MHz)

Absorption-line studies toward bright background sources provide a powerful observational probe of galaxy formation and evolution. The highest H I column density absorbers, the damped Lyman- α systems (DLAs, with $N_{\text{HI}} \gtrsim 10^{20} \text{ cm}^{-2}$) are of particular interest as these are the largest repository of neutral gas at high redshifts and hence believed to be the precursors of present-day galaxies (e.g., Wolfe et al. 1986).

Understanding the evolution of a “typical” DLA as a function of redshift has long been recognized as critical to understanding normal galaxy evolution, but despite 25 years of detailed study, the nature of high- z DLAs remains controversial, with models ranging from large, rapidly-rotating massive disks to small, merging sub-galactic systems. One reason for the uncertainty is that although present high- z DLA samples are quite large (e.g., ~ 500 DLAs in the SDSS-3 data release; Prochaska et al. 2005), they are almost entirely drawn from optically-selected surveys and so are likely to be biased against systems with a high metallicity (i.e., a high dust content). For example, current DLA samples contain no absorbers with both high H I column density and high metallicity.

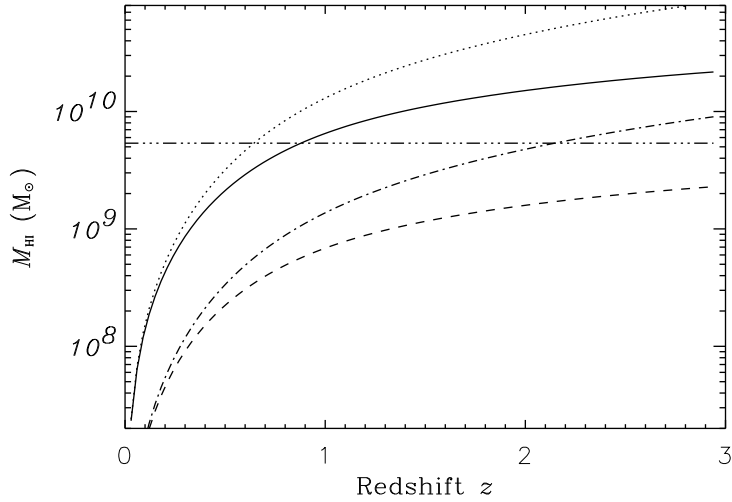


Figure 5. Limiting HI mass M_{HI} (in units of M_{\odot}) for HI emission-line surveys with the SKA, for a signal-to-noise (S_N) ratio of 10 in a 4-hour integration time (upper solid and dotted lines) and a 360-hour integration time (lower dot-dashed and dashed lines), from Abdalla & Rawlings (2005). The dotted and dot-dashed lines assume pointed observations, and the solid and dashed lines assume tiled surveys, as discussed by Abdalla & Rawlings. The horizontal line corresponds to the break of the HI mass function M_{HI}^* at low redshifts (Zwaan et al. 2003).

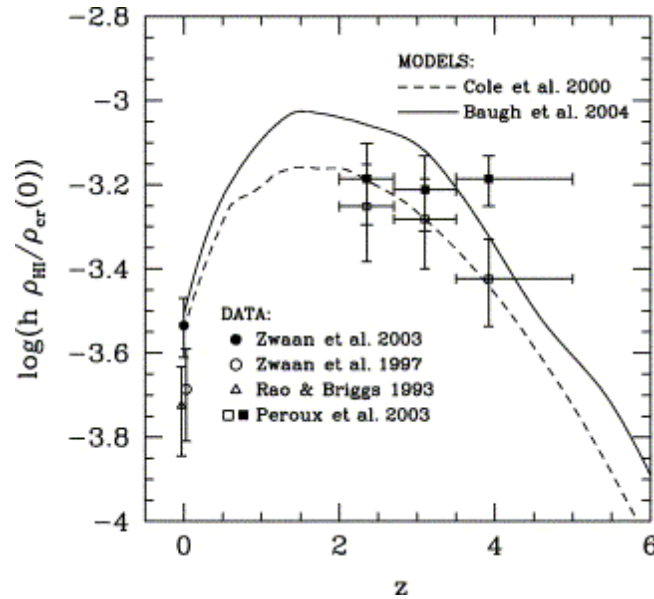


Figure 6. Theoretical predictions for the cosmic HI mass density at different redshifts, scaled to the present-day critical density for a flat Universe (Baugh et al. 2004). Solid and dashed lines show the predictions from two hierarchical galaxy-formation models with different assumptions about the star-formation timescales in galaxies. The HI mass density is predicted to peak near $z \sim 2$.

This “dust-bias” complicates the use of optical DLA samples in studying the evolution of the average galaxy population, which critically requires an unbiased sample of DLAs with a uniform sampling in redshift. Furthermore, direct optical imaging of high- z DLAs is difficult because of the need to subtract out the point spread function of the background quasar. Similarly, it is not possible to determine the transverse size of the absorbers using optical or ultraviolet absorption lines because quasars are unresolved at optical wavelengths.

H I absorption is only detectable in high column density gas, $N_{\text{HI}} \gtrsim 10^{20} \text{ cm}^{-2}$, precisely the type of absorber that would give rise to damping wings in the Lyman- α line. 21 cm absorption surveys towards radio-loud background sources are thus well-matched toward the construction of DLA samples. Radio surveys contain no “intrinsic” redshift bias, such as that arising from the ultraviolet cut-off in the atmosphere and, just as importantly, flux-limited radio samples are unaffected by dust. *Blind H I absorption surveys with SKA will therefore provide DLA samples that are entirely unbiased by dust extinction and redshift coverage.*

Quasar DLAs found by SKA could be followed up with optical spectroscopy to measure their H I column density and thus trace the global evolution of the cosmological H I mass density. Combining the 21-cm absorption spectra with the H I column density enables one to track the evolution of the gas temperature distribution (e.g., Kanekar & Chengalur 2003).

H I absorption-line surveys with SKA would target both compact quasars and extended radio galaxies, whereas optical surveys are limited to (unresolved) quasars. High-redshift H I detections against extended radio sources would allow estimates of the transverse size, kinematics and H I mass of the intervening galaxy and, in the case of disk absorbers, even their dynamical mass (Kanekar & Briggs 2004). Such studies can show whether the velocity field is quiescent and well-ordered, or whether disruptive events such as mergers have occurred in the recent past.

Kanekar & Briggs (2004) estimate that the SKA will be able to carry out 21-cm absorption measurements against thousands of distant radio sources out to $z \gtrsim 4$, with a survey sensitivity function ~ 20 times larger than that obtained by combining all present optical surveys. H I absorption studies with the SKA will thus provide critical inputs to understanding the evolution of galaxies at redshifts above those accessible to 21cm emission-line studies.

Frequency coverage below 500 MHz (i.e., at $z > 1.8$ for the H I line) is essential for this purpose, as both the cosmological H I mass density and the DLA number density show strong evolution in the redshift range $2 < z < 5.5$ (Prochaska et al. 2005). Coverage over the entire frequency range 200-1400 MHz would allow SKA 21-cm absorption studies to trace the evolution of H I in galaxies right from the build-up of large H I masses and column densities at the highest redshifts ($z \sim 4 - 6$), through the conversion of neutral hydrogen into stars at intermediate redshifts, all the way to the present epoch.

5.3.3 Dark energy surveys above $z = 1.8$?

The relevant SKA experiments for dark-energy science are:

1. Large-scale structure surveys for redshifted H I, yielding very precise measurements of the galaxy clustering pattern including features such as the baryon oscillations.
2. Weak lensing (cosmic shear) continuum surveys.
3. The cross-correlation of galaxies in different source and lens planes to probe cosmic magnification effects.

The fundamental scientific goal is to discriminate between different models for the dark energy using probes such as cosmic distance relations and the growth of structure with redshift.

The frequency range 200-500 MHz is not essential for this science. For redshifted H I, frequencies below 500 MHz correspond to $z > 1.8$. In our default picture where a Lambda-CDM Universe is the “null hypothesis”, $z > 1.8$ is not the critical regime for measuring the dark energy parameters because the cosmic expansion is matter-dominated.

However, there are many alternative non-lambda-CDM dark energy models compatible with current data for which dark energy has a significant effect at $z > 1.8$. Current data very weakly constrain the evolution of the dark

energy equation of state with redshift, the key measurement which discriminates between dark energy models. Depending on the results of intervening surveys, $z > 1.8$ may turn out to be a critical regime. Moreover, this range is very difficult for ground-based optical surveys (and even space missions) to cover in large (thousands of deg^2) areas. Including this capability in the dark energy surveys therefore opens up important “discovery space”.

In the Lambda-CDM model, the most significant gains from the $z > 1.8$ capability are:

- There is a large amount of cosmic volume at $z > 1.8$ which can be used to measure large-scale clustering modes accurately. Baryon oscillations at a given redshift measure the angular diameter distance to that redshift, which still depends on dark energy at lower redshifts via an integral.
- High-redshift observations can extend the background source redshift slices for cosmic magnification experiments.

We note that for any of these experiments to be competitive on the timescale of 2020, the SKA must be designed with a sufficient instantaneous field-of-view that these surveys can cover a hemisphere in ~ 1 year.

5.4 KSP II. Gravity/Pulsars

The frequency range 200–500 MHz is not essential, but would be used if available (Cordes & Kramer 2005). Since pulsars have steep radio spectra, frequencies below 500 MHz would be useful for detecting faint pulsars at Galactic latitudes where scattering is not an issue.

5.5 KSP III. Cosmic Magnetism

The ability to observe at frequencies below 500 MHz, while not essential to this KSP, would be of considerable interest. Interesting and potentially important investigations of the magnetic fields in the local ISM and in galaxy clusters would benefit from observations at 300–500 MHz, but this would be a specific subset of experiments.

Due to the steep spectrum of most galaxies and clusters, low frequencies give higher flux densities. On the other hand, Faraday depolarization increases even stronger with wavelength. Hence there is an optimum wavelength range to detect polarized emission, which is different for different sources. Galactic disks can be best observed around 1 GHz. Star-forming regions and central regions need higher frequencies, while halos and clusters are best observed at lower frequencies. The main KSP task, the all-sky RM survey, is best observed at 1.4 GHz.

It is hard to predict exactly what we can expect with the SKA until we have realistic models for the magnetized ISM. At the moment the optimal range for the Cosmic Magnetism KSP is thought to be 0.4–10 GHz. However, important information about intergalactic magnetic fields can still potentially be obtained in the range 200–400 MHz. Current observations only provide upper limits to the field strength of the IGM. The detection of synchrotron radiation at the lowest possible frequencies will be crucial for tracing the properties of magnetic fields in regions of the IGM more rarefied than in clusters. Due to the diffuse nature of this emission, this is best observed at lower frequencies. Finally, it would also be desirable to have the SKA frequency range extend down to at least 240 MHz, to provide continuity with the polarimetry to be carried out by LOFAR (which covers 30–240 MHz) and other low-frequency experiments now being planned.

5.6 KSP V. Probing the Dark Ages

The frequency range required by this KSP for H I studies of the EoR is 100–200 MHz, which places it outside the scope of this discussion. As noted in §5.3 above, the SKA will be the first telescope powerful enough to detect neutral hydrogen in emission and/or absorption over the entire redshift range from the EoR to the present day. We believe that there are compelling arguments for providing the SKA with continuous frequency coverage over the range 0.1 to 1.4 GHz, so that we can observe and study neutral hydrogen over the entire span of cosmic time — one of the main reasons for conceiving a SKA in the first place.

REFERENCES

- F. Abdalla, S. Rawlings, 2005, MNRAS, 360, 27
Baugh, C.M. et al., 2004, New Astr. Rev., 48, 1239
P. Boissé et al. 1998, A&A, 333, 841
F.H. Briggs, 1997, PASA 14, 31
J.G. Cohen 2002, ApJ, 567, 67
J. Cordes, M. Kramer, 2005, Summary of SKA KSP meeting on Gravity, Aug. 2005.
A. Heavens, B. Panter, R. Jimenez, J. Dunlop, 2004, Nature, 428, 625
N. Kanekar and J.N. Chengalur 2003, A&A, 399, 857
N. Kanekar and F.H. Briggs 2004, New Astr. Rev., 48, 1259
J. X. Prochaska et al. 2005, ApJ, 635, 123
A.M. Wolfe et al. 1986, ApJS, 61, 249
M.A. Zwaan et al., 2003, AJ, 125, 2842

6 WHAT IS THE CASE FOR HIGH FILLING FACTOR AT HIGH FREQUENCY?

John Dickey, Rainer Beck, Luigina Feretti, Eric Wilcots

6.1 Summary

The filling factor or aperture covering factor of the SKA is a critical parameter for determining the sensitivity of the array to low surface brightness emission. This is of primary importance for observations of emission by the 21-cm line, but it is equally important at shorter wavelengths, as brightness sensitivity is the limiting factor in mapping any spectral line that is thermally excited. Thus brightness sensitivity is central for galaxy structure and dynamics studies using the H I line and highly redshifted CO lines, and for star and planet formation studies using molecular lines such as NH₃, CH₃OH, H₂CO, CH, OH, and other molecules which have transitions with wavelengths in the range 18 cm to 1.2 cm. These are relevant to the Cradle of Life and Galaxy Evolution KSPs.

The same surface brightness consideration applies to linearly and circularly polarized emission (Stokes Q , U and V) from the Milky Way ISM and from the intra-cluster medium in clusters of galaxies. Mapping this diffuse synchrotron emission is fundamental to the Cosmic Magnetism key science project. Sensitivity is particularly critical at frequencies above about 2 GHz where the synchrotron spectrum makes this emission very faint. Mapping the linear polarization at high frequencies shows the intrinsic position angle of the emission region; this is useful to isolate the effects of propagation in data taken at lower frequencies, where Faraday rotation in the intervening medium has a much stronger effect.

In addition to the diffuse synchrotron background, at the highest frequencies the SKA will be sensitive to structure in two thermal backgrounds, the 2.7 K cosmic blackbody, and emission from dust in the Milky Way and nearby spiral galaxies. The latter may show linear polarization due to alignment of dust grains in the Galactic magnetic field, and thus it is another valuable tracer for the Cosmic Magnetism project.

For computational reasons, searches for pulsars or SETI signals are much easier to carry out with an array configuration with a high filling factor. This is because a centrally condensed array provides maximal sensitivity on a relatively small number of synthesized beams per primary beam area. In this way a high filling factor is critical for the key science projects on Gravity (pulsars) and the Cradle of Life.

This report derives the relationship between the aperture filling factor and the brightness sensitivity, uses this relationship to derive the mapping speed, and then computes some rough numbers using a simple straw-man model of the distribution of the inner half of the collecting area of the array. The results are illustrated by several calculations determining the observing time needed for some representative applications. In an appendix, we emphasize the case for high filling factor in studies of the 21-cm H I line.

6.2 Background

The brightness sensitivity of an aperture synthesis array is set primarily by its aperture filling factor. This brightness sensitivity determines what kinds of studies are possible with the array, and the speed with which these observations can be done. There are several KSPs for the SKA that will be limited by the speed of the telescope for mapping emission with very low surface brightness. In the following discussion, we first review the connection between brightness sensitivity and filling factor, and then we describe several hypothetical observations that will be possible *only* if the filling factor is high in the central core of the telescope.

6.2.1 Sensitivity and Mapping Speed for a Flux Limited Survey

The flux sensitivity, i.e., the rms noise in the flux density, σ_s , of a conventional aperture synthesis array with movable antennas is given by the number of antennas, n , the integration time, τ_i , the system temperature of the receivers, T_{sys} , the bandwidth, BW , and the gain, G (in K/Jy), of the individual antenna elements:

$$\sigma_s = \frac{T_{sys}}{G\eta} \sqrt{\frac{2}{n(n-1)}} \frac{1}{\sqrt{BW \tau_i}}, \quad (1)$$

with η an efficiency set by the aperture efficiency of the antennas and by the tapering functions used on the spectral and spatial correlation functions ($\eta < 1$). This is independent of the placement of the elements, so for example the formula is the same for the VLA A, B, C, and D configurations. To make a mosaic survey of a large solid angle, Ω , down to a given *flux* sensitivity takes total integration time, τ , given by:

$$\tau = \frac{4 \Omega}{\Omega_b} \tau_i = \frac{16 k}{\lambda^2} \frac{T_{sys}^2}{\eta^2 BW} \frac{\Omega}{\sigma_s^2} \frac{1}{G n^2}, \quad (2)$$

with k Boltzmann's constant, λ the wavelength, Ω_b the primary beam solid angle (of the individual antennas) and we approximate $n(n-1) \simeq n^2$. The factor of four in the upper line comes from the need to observe points separated by roughly half the primary beam full width, and we have used the relationship between the gain and the primary beam solid angle of the individual antennas, $2 \Omega_b G k = \lambda^2$. This is just a factor of two different from the speed to make a flux limited survey with the same number of receivers mounted on a single dish telescope of the same total area:

$$\tau = \frac{4 \Omega}{\Omega_b} \tau_i = \frac{8 k}{\lambda^2} \frac{T_{sys}^2}{\eta^2 BW} \frac{\Omega}{\sigma_s^2} \frac{1}{G_t n}, \quad (3)$$

where now G_t is the gain of the single dish telescope and n is the number of independent receivers that illuminate its aperture, as in a multi-beam receiver. If the collecting area of an aperture synthesis telescope were all gathered together into one dish it would have gain $G_t = n G$; with this substitution the two mapping speed formulae become nearly identical. But this is for a flux-limited survey; the situation is quite different for a brightness-limited survey.

6.2.2 Mapping Speed for a Brightness-Limited Survey

The best possible brightness sensitivity is that of a single dish telescope, for which the rms noise in brightness temperature, σ_T , is given by the radiometer equation,

$$\sigma_T = \frac{T_{sys}}{\eta \sqrt{BW} \tau_i} \quad (\text{single dish}), \quad (4)$$

where η is now the beam efficiency rather than the aperture efficiency, which may include other effects such as the need to position switch or frequency switch for calibration. For an aperture synthesis telescope we must compute the brightness sensitivity from the flux sensitivity using the synthesized beam-width, θ_s (FWHM), by

$$\begin{aligned} \sigma_T &= \sigma_s G_e = \sigma_s \frac{\lambda^2}{2.26 k \theta_s^2} \quad (\text{interferometer}) \\ &= 6.1 \times 10^2 K \frac{\sigma_s}{mJy} \left(\frac{\lambda}{21.1cm} \right)^2 \left(\frac{\theta_s}{arcsec} \right)^{-2}, \end{aligned} \quad (5)$$

where G_e is here the gain of the equivalent single dish with resolution θ_s . Equation (5) expresses the familiar conversion from units of Jy/beam to K of brightness temperature:

$$\frac{T}{\left(\frac{Jy}{beam} \right)} = 0.73 \frac{T}{K} \left(\frac{\lambda}{cm} \right)^{-2} \left(\frac{\theta_s}{arcsec} \right)^2.$$

Substituting Equation (1) into Equation (5) we see that the brightness sensitivity of the aperture synthesis telescope is given by the radiometer equation with an extra factor, f :

$$\sigma_T = \frac{T_{sys}}{\eta \sqrt{BW} \tau_i} f^{-1}, \quad (6)$$

where f is the filling factor,

$$f = \frac{\theta_b^2}{n \theta_s^2} \simeq \frac{A_e}{A_{<}} \simeq \frac{A_e}{\lambda^2} \theta_s^2, \quad (7)$$

where θ_b is the FWHM of the primary beam of the individual antennas of the array, A_e is the effective total collecting area including all antennas, and $A_{<}$ is the geometric area enclosed by a circle whose diameter equals the longest baseline, i.e. $A_{<} = \pi d^2/4$ where d is the longest baseline. For a single dish, $f \rightarrow 1$, depending on the taper in the illumination pattern, but for most aperture synthesis telescopes $f \ll 1$, e.g., $f \sim 0.014$ for the VLA D configuration, $f \sim 0.015$ for the 375-m configuration of the ATCA, and in the extreme, $f \sim 10^{-10}$ for the VLBA. This means that an integration time and bandwidth combination that would give rms noise in antenna temperature of 1 K for a single VLBA antenna gives rms noise of $\sigma_T = 10^{10}$ K for the array.

The mapping speed for a survey sensitive to extended emission with a given brightness sensitivity follows from Equations (6) and (7), and is a strong function of f . To survey total solid angle Ω with 1σ brightness sensitivity σ_T takes total integration time

$$\tau = \frac{4 \Omega}{\Omega_b} \tau_i = \frac{4 \Omega A_e T_{sys}^2}{n \lambda^2 \eta^2 \sigma_T^2 f^2 BW}, \quad (8)$$

where the primary beam solid angle, $\Omega_b = \lambda^2 n/A_e$. For a movable-element array like the VLA this survey time goes as the fourth power of the longest baseline; so, for example, a brightness-limited survey with resolution $10''$ takes 16 times longer than one with resolution $20''$.

6.3 Filling Factors for the SKA

Even a square kilometer of collecting area has a low filling factor for baselines longer than a few kilometers. To maintain even moderate surface brightness sensitivity, a significant fraction of the collecting area must be concentrated near the center. If the array is configured to have half a square kilometer of collecting area inside a circle with radius 2.5 km, as described in SKA Memo 45, Table 1 (Jones, February 2004, ‘‘SKA Science Requirements: Version 2’’), this gives filling factor of only 5×10^{-3} . A simple configuration that is scale invariant (i.e., has no preferred particular baseline length) that gives this nominal specification can be described by a logarithmic variation of collecting area A_{eff} inside radius r as given by:

$$a = \frac{A_{eff}}{A_0} = 4 + 43 \log b, \quad (9)$$

where $b = \frac{r}{100m}$ and A_0 is the geometric area of a 100m diameter dish, i.e.,

$$A_0 = \pi(50m)^2 = 7.85 \cdot 10^3 \text{ m}^2 = \frac{1}{127} \text{ sq. km.} \quad (10)$$

This gives filling factor $f = a/4b^2$ decreasing from $f = 1$ (out to $b = 2.1$, i.e., $r \leq 210$ m, 1.4 times the radius of the Arecibo primary) to $f = 0.7$ at $r = 300$ m, $f = 0.34$ at $r = 500$ m, $f = 0.12$ at $r = 1$ km, $f = 0.375$ at $r = 2$ km. If the logarithmic slope continues out to 10 km radius this circle would contain 71% of the total area, with filling factor 2.3×10^{-3} . The optimistic $f = 1$ for $r < 210$ m approximation might have to be compromised to prevent shadowing or cross-talk between very close antennas. An innovative mounting structure (e.g., wedding-cake design) might alleviate the shadowing problems, but still a maximum of $f \simeq 0.7$ is probably more realistic. But for baselines longer than about 300 m the simple model of Equation (9) for the distribution of collecting area in the inner 10 km radius should be feasible (illustrated in Fig. 7). Note that the precise shape of the solid curve (Eq. [9]) may be determined by other design criteria, but the slope overall should not be much shallower than that given by Equation (9) or else the filling factor will get too low for the telescope to provide adequate brightness sensitivity.

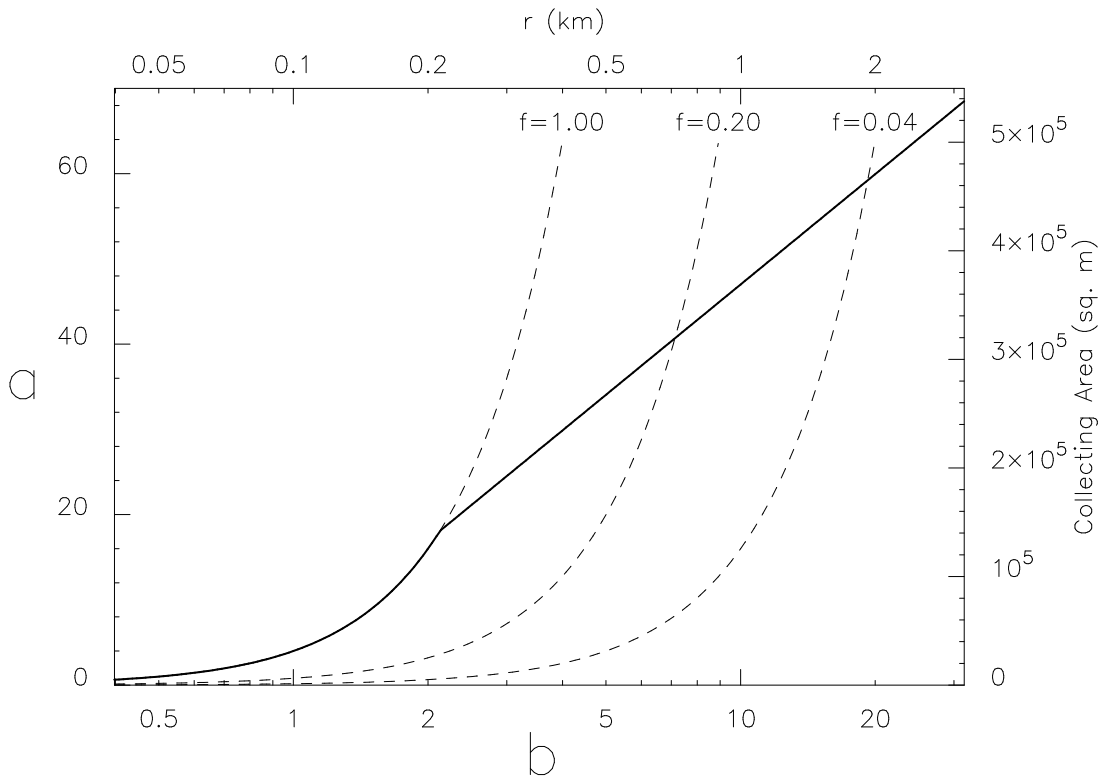


Figure 7. Filling factors for a simple, logarithmic distribution of collecting area that gives half of the total within a radius of 2.5 km from the center. The solid curve shows the aperture area, a , as a function of maximum baseline length, b , with a and b in units of a 100-m filled aperture; the straight line is given by Equation (9). The top scale gives the radius of the corresponding circle, in meters, and the right axis the total collecting area, in m^2 . Note that the horizontal axis is logarithmic. The dashed curves show filling factors of 1.0, 0.2 and 0.04.

The simple model of Equation (9) gives rms brightness temperature, σ_T , as a function of synthesized beam width, θ_s , using Equation (6) given one hour of integration time, $T_{sys} = 50$ K, and $\eta = 0.65$ as shown in Table 4 for different wavelengths and bandwidths.

The first block in Table 4 is appropriate for spectroscopy of the 21-cm line with velocity resolution 10 km s^{-1} , the middle block is appropriate for continuum mapping (Stokes I , Q , U or V), and the lower block is appropriate for mapping NH_3 or any other thermally excited spectral line at roughly 1.3 cm wavelength with velocity resolution 1 km s^{-1} . This will be particularly significant for mapping CO $J=(1-0)$ emission at redshifts higher than $z \sim 3$ ($\nu < 29$ GHz). A wavelength of $\lambda = 1.25$ cm as shown in Table 4 corresponds to $z = 3.8$ for CO ($J=1-0$).

If the total SKA collecting area is scaled down to match a lower system temperature, holding the ratio A_e/T_{sys} constant, then Equations (6) and (7) show that the brightness sensitivity is unchanged, since σ_T is proportional to T_{sys}/A_e . The slope of the linear portion of the curve on the Figure is reduced to match the new total collecting area (right axis). Alternatively the smaller collecting area can be arranged so as to satisfy Equation (9) with the same coefficient, 43, on the log term. In this case the surface brightness sensitivity is improved, but only out to some smaller radius than 10 km, since a significant fraction of the collecting area must be committed to long and very long baselines.

Table 4. Brightness sensitivity after 1 hour integration. Here r is the longest baseline used in the imaging, which determines the synthesized beam width, θ_s .

λ cm	BW MHz	r m	θ_s arc sec	σ_T mK
21	0.05	210	105	5.7
21	0.05	300	75	8.5
21	0.05	1000	22	47
21	0.05	3000	7.5	310
21	0.05	10000	2.2	2600
21	100	210	105	0.13
21	100	300	75	0.19
21	100	1000	22	1.1
21	100	3000	7.5	6.9
21	100	10000	2.2	57
1.25	0.08	210	6.3	4.5
1.25	0.08	300	4.5	6.8
1.25	0.08	1000	1.3	37
1.25	0.08	3000	0.44	240
1.25	0.08	10000	0.13	2020

6.3.1 Minimum Baseline

In addition to the design requirements set by brightness sensitivity, there is the question of the shortest baselines sampled by the SKA, which set the largest angular scale to which the array is sensitive. To image large scale emission on scales of $\sim 30'$ at 20 cm, the core of the SKA should have a minimum baseline of ~ 20 m. For some projects single dish maps will have to be made to supplement the interferometer data to supply the true “zero-spacing” flux, i.e., the constant offset of the brightness scale for the map. This is particularly important for Stokes Q and U , which as vector quantities become very distorted if a constant term is missing due to unsampled, smoothly spread emission.

6.4 Low Surface Brightness Science at High Frequencies

The science goals of the SKA involve many kinds of observations that require high brightness sensitivity. Particular examples of this science include studies of diffuse linear polarization in continuum (KSP3, The Magnetic Universe), surveys of hydrogen inside and outside galaxies (KSP4, Evolution and Cosmology), and emission and absorption by the intra-cluster medium in clusters of galaxies at high and low redshifts (KSP4). In this section we detail the surface brightness sensitivity and resolution needed for some representative observations in these areas.

6.4.1 Magnetic Fields and Polarized Emission

Several different kinds of observations will be used to address this key science project. Here we list considerations associated with these observations that require relatively high filling factors.

- Study of the small-scale magnetic structures (filaments, loops, reconnection regions) in the interstellar medium, supernova remnants, molecular clouds, or H II regions of the Milky Way by mapping their polarized emission.

This requires high angular resolution (hence high frequencies) *and* high surface brightness sensitivity. For example, the 5-GHz synchrotron emission from cosmic-ray electrons in a 15 μG magnetic field along a 1 pc path length is ~ 3 nJy per 1'' FWHM beam (assuming energy equipartition between total cosmic rays and the magnetic field, and a ratio of 100 between cosmic-ray protons and electrons).

- Study of the distribution of magnetic fields and thermal electrons in the dense interstellar medium of nearby galaxies by mapping diffuse polarized emission and Faraday rotation measures.

In the Galactic plane and in the Galactic central region, Faraday rotation measures reach several ± 100 rad m^{-2} . Even larger values are expected for massive galaxies with stronger fields than in the Milky Way. To avoid Faraday depolarization and ambiguous polarization angles, observations at high frequencies are required where synchrotron emission is weak. High surface brightness sensitivity is essential. Furthermore, all angular scales of the object must be imaged with similar sensitivity in order to avoid distortions in polarization angle due to missing large-scale emission in the maps of the Stokes parameters Q and U .

- Detection of intervening magneto-ionic objects which depolarize and/or Faraday-rotate a diffuse polarized background signal (“polarization silhouettes”), e.g., by intervening galaxies, planetary nebulae, supernova remnants, pulsar nebulae, molecular clouds or H II regions.

The lobes of distant radio galaxies or the extended Galactic emission can be used as polarized background signals. In the latter case, all angular scales must be imaged with similar sensitivity to avoid distortions of the measured polarization angle due to missing large-scale emission.

6.4.2 Clusters of Galaxies at Low and High Redshift

Diffuse Synchrotron Emission: The diffuse synchrotron emission from clusters (see, e.g., Feretti, Burigana, and Ensslin, 2004, *New Astr. Rev.*, 48, 1137) is presently detected down to values of the brightness, B , of 100 nJy/arcsec² (at 1.4 GHz), for sources of total flux of 10–100 mJy (depending on the redshift). The brightness of faint diffuse sources of total flux down to 1 mJy can be computed assuming a source size of about 1 Mpc. In this case most of the flux is expected to be within the central region of 0.5 Mpc diameter (concordance cosmology is used for the computation). This gives the values in Table 5.

For a detection at the 3- σ level, a sensitivity of 6 nJy arcsec⁻² is needed for nearby clusters. The same sensitivity would allow detection of a 10% polarized flux in clusters at $z = 0.5$, whereas for nearby clusters the sensitivity needs to be 0.6 nJy/arcsec² for this.

Table 5. Expected 1.4 GHz brightness from diffuse synchrotron emission in clusters of galaxies at different redshifts.

Redshift	$B_{1.4}$ (nJy arcsec $^{-2}$)	T_B (mK)
0.1	17	10
0.5	188	116
1.0	329	203

For example, at 21 cm, 1 K of brightness temperature corresponds to $1.6 \mu\text{Jy arcsec}^{-2}$, so $1 \text{ nJy arcsec}^{-2}$ corresponds to $T_B = 0.6 \text{ mK}$. Thus an rms noise level of $6 \text{ nJy arcsec}^{-2}$ corresponds to 3.6 mK . The second block in Table 4 shows that this is achieved in about four hours of integration for a synthesized beam width of $7''.5$ or in about seven minutes of integration for a synthesized beam width of $22''$, using a 100 MHz bandwidth at 21 cm.

Due to their steep spectra, cluster radio halos and relics have been detected mostly at low frequencies (1.4 GHz or lower). However, information on their spectra at higher frequencies (in particular up to 10 GHz) is crucial. This range is totally unexplored with interferometers. Knowledge of high frequency spectra gives information on the properties of the radio emitting particles, i.e., diffusion, ageing and any reacceleration process. The expected brightnesses at higher frequencies can be obtained from Table 5 by scaling $B_{1.4}$ with an average spectral index of ~ 1.2 (where $S_\nu \propto \nu^\alpha$).

Studies of the SZ Effect: The detection of the Sunyaev-Zel'dovich (SZ) effect both from clusters and from proto-galactic gas on galaxy scales has been discussed by Burigana et al. (2004, *New Astr. Rev.*, 48, 1107). Many thousands of galaxy clusters will be identified by future missions like Planck. Their typical angular sizes range from $\sim 1 \text{ arcmin}$ to tens of arcmin. This effect will be detectable with the SKA down to arcsec scales, provided that the interferometer contains baselines that sample scales $> 10''$ as well as the small angular scales. At increasingly higher resolutions, there is some possibility of detecting kinematic effects associated with fast motions. Although the scale of these is quite unclear, they would represent some prospect of measuring the cluster dynamical state.

In addition, what is of large importance is the detection of small-scale structure due to pressure anisotropies in the centers of clusters, associated with AGN-driven shocks and blast waves (Cavaliere & Lapi, 2006, *ApJL*, in press, astro-ph/0607021), or radio jets or lobes, or some other mechanism transferring energy to the cluster gas. These effects would be of the order of $\sim 1 \mu\text{Jy beam}^{-1}$ with $1''$ resolution, at frequencies $> 20 \text{ GHz}$. Since this is extended low-surface brightness emission, its detection requires high surface brightness sensitivity.

6.4.3 Pulsar and SETI Searches

The problem for searches for modulated or transient signals from point sources is that every synthesized beam area in the primary beam of the telescope must be analysed separately. Pulsar searches are a simple example. To search all reasonable dispersion measures over a broad bandwidth for a wide range of pulsar periods is an immense computational job, which increases as the square of the ratio of the synthesized beam width to the primary beam width, θ_b^2/θ_s^2 as in Equation (7). Thus a low filling factor array is much more difficult to use for such searches. The most computationally efficient situation is a high filling factor instrument, such as a multi-beam receiver on a single dish telescope. Future advances in computing power may mitigate this problem to some extent, but these advances may be offset by the push to broader bandwidths and longer integrations. Thus high filling factors are strongly preferred for all key science goals related to pulsar searches and SETI.

6.4.4 Thermal Backgrounds at High Frequencies

Studies of the spatial structure of the CMB can be done best from space, however there may be niche areas in which the SKA can contribute, particularly to studies of the polarization of the CMB. Subtraction of foreground confusion, both from compact sources and from diffuse Galactic emission, is a critical area in which the broad frequency range of the SKA can be very useful. There is a strong possibility that at frequencies above 15 to 25 GHz the diffuse polarized emission is dominated by dust emission, which can be polarized by alignment of the grains with the magnetic field (Lazarian & Finkbeiner, 2004, *New Astr. Rev.*, 47, 1107). This is a much stronger effect at higher frequencies, but it may be detectable with the SKA, if the filling factor is kept sufficiently high.

6.5 APPENDIX: Diffuse H I Around Nearby Galaxies

The case for high filling factor at 21 cm is clear and uncontroversial. As an illustration, this appendix discusses several important issues in the study of galaxies (Galaxy Evolution KSP), and the requirements that they place on the filling factor of the inner part of the array.

Mapping the extended H I environment of nearby galaxies is extremely important to testing cosmological models of galaxy formation, directly measuring the intergalactic radiation field, and connecting the observed physical conditions in quasar absorption line systems with the properties of the extended disks of galaxies at $z \sim 0$.

Where Do Galaxies End? This is a question that has relevance to both studies of galaxy formation, as well as to our understanding of the conditions of the IGM. Corbelli & Salpeter (1993) used the observed sharp edge (at an H I column density of $2 \times 10^{19} \text{ cm}^{-2}$) to the H I disk of M33 to estimate that the intergalactic ultraviolet background radiation at the Lyman edge is $6 \times 10^{-23} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$. A number of more recent interferometric studies suggest that the H I disks of galaxies extend to column densities lower than 10^{19} cm^{-2} . Using the VLA, Wilcots & Hunter (2002) reached a sensitivity of $2 \times 10^{19} \text{ cm}^{-2}$ in a survey of Sextans A. Van Gorkom (1993) carried out the deepest synthesis observations of an individual isolated galaxy with a 3σ detection limit of $4 \times 10^{18} \text{ cm}^{-2}$ over a 40 km s^{-1} channel width. The “magic number” is an H I column density of $1.6 \times 10^{17} \text{ cm}^{-2}$, at which point the optical depth for continuum radiation at the Lyman edge is equal to one (Corbelli, Salpeter & Bandiera 2001).

Connecting Nearby Galaxies to Lyman Limit Systems: We have recognized for the better part of a decade that Lyman Limit systems (LLS), and the Mg II quasar absorption line systems with which they are usually associated, are related to the outer disks of galaxies (Rao & Turnshek 2000). Churchill and collaborators (Charlton & Churchill 1998, 2000; Churchill & Vogt 2001; Churchill et al. 2003) have shown that Mg II absorbers at intermediate redshifts arise from system of clouds that are optically thick at the Lyman limit and, therefore, have H I column densities of $N_{HI} > 1.6 \times 10^{17} \text{ cm}^{-2}$. Nestor, Turnshek & Rao (2005) show that the Mg II absorption line systems have cross sections of 50–100 kpc. Thus a significant fraction of the gaseous component of the nearby Universe may lie in the outskirts of galaxies at column densities that are an order of magnitude lower than what we can currently achieve with available instrumentation.

Structure in the Outer Disks of Galaxies: In thinking about the need for sensitivity to H I column densities of $\sim 10^{17} \text{ cm}^{-2}$, one cannot ignore the complementary need to maintain good angular resolution. Current deep mosaics of the H I distribution around nearby galaxies reveal a wealth of structure in both the distribution and kinematics of the neutral gas. Hunter et al. (1998) uncovered a complex distribution of H I streamers and filaments around NGC 4449, Wilcots & Miller (1998) detected clouds of infalling gas around IC 10, and Westmeier, Braun, & Thilker (2005) detected a population of low mass clouds around M31 that may be tracing the galaxy formation process. In many nearby galaxies the precipitous fall-off in H I column density at the edge of the disk

is unresolved with $1'$ synthesized beams.

Where Are We Now? The most sensitive 21-cm line single dish surveys to date achieve column density sensitivities of $N_{HI} \sim (1 - 5) \times 10^{18} \text{ cm}^{-2}$. Minchin et al. (2003) achieved an HI column density sensitivity of $4.2 \times 10^{18} \text{ cm}^{-2}$ over a 32 deg^2 region of Centaurus with the Parkes multi-beam system. A more recent Westerbork survey of the environment of M31 achieved a 1σ column density sensitivity of $9 \times 10^{17} \text{ cm}^{-2}$, equivalent to a brightness temperature rms of 40 mK (Westmeier, Braun & Thilker 2005). These observations revealed a collection of low mass HI clouds that could represent a high velocity cloud population associated with M 31.

The SKA HI Survey of the Edges of Nearby Galaxies: In addition to the lack of sensitivity to column densities of $\sim 10^{17} \text{ cm}^{-2}$, our efforts to understand the extended environments of nearby galaxies remain hindered by the “short-spacing” problem. The full extent of the HI disk associated with a Mg II absorption line system of a galaxy at 10 Mpc is likely to be $\sim 16' - 32'$. A $3\text{-}\sigma$ detection of HI at a column density of $1.6 \times 10^{17} \text{ cm}^{-2}$ corresponds to a brightness temperature of $\sim 3 \text{ mK}$.

A column density of 10^{18} cm^{-2} of HI gives a brightness temperature of 10 mK over a line-width of 50 km s^{-1} (10 km s^{-1} corresponds to 47.4 KHz bandwidth at $z = 0$). From the first block of Table 4, 10 mK would give a 5:1 signal-to-noise ratio after about 18 hours of integration for a beam width of $75''$. To get a resolution of $22''$ would require a much longer integration, 550 hours to get an rms brightness temperature of 2 mK.

7 WHAT ARE THE OPTIONS FOR TRANSIENT DETECTION WITH THE SQUARE KILOMETER ARRAY?

Joseph Lazio (NRL), James Cordes (Cornell & NAIC), & Dan Werthimer (UCB/RAL)

7.1 Summary

We consider radio transient detection possibilities with the SKA. Transient searching can be divided into two modes, *slow* in which the sky can be raster scanned to search for known and new classes of transients and *fast* in which a given field of view must be observed for sustained periods of time. With the current SKA design goals, a slow transient search that involves pixellating the entire field of view but no dedispersion will likely be possible with the computational resources available by the time the SKA is on-line. A fast transient search of a pixelated field of view will require algorithm improvements and considerable computational resources. However, fast transient searching also may ultimately offer the possibility of searching for time-frequency behaviors that are more complex than broadband, dispersed, and periodic pulsar-like signals. We also discuss two strawman implementations of a transient buffer, one involving an archive of relatively low-resolution images, the other involving either a disk- or RAM-based near-real-time buffer. The hardware costs for either kind of transient buffer appear achievable for a small fraction of the SKA's total cost. However, implementation details for transient searching will require additional careful study of the SKA design as well as possible impact on other KSPs.

7.2 Introduction

Cordes, Lazio & McLaughlin (2005) and Wilkinson et al. (2005) describe the scientific potential associated with searching for radio transients, particularly with respect to the detection of new classes of objects, new phenomena, or both. Those authors make some general statements about the requirements on the SKA to optimize it for searching for radio transients. Cordes (2006, in preparation; hereinafter C06) considers detection aspects of pulsars, radio transients, and extraterrestrial transmitters in more detail. We shall summarize salient details when appropriate, but focus more on the instrumental design considerations arising from optimizing the SKA for transient detection.

Our intent is to consider the detection and identification of radio-selected transients, with an emphasis on *blind* searching. Radio telescopes have long been used for monitoring and follow-up observations of transients selected at other wavelengths, particularly X- and γ -ray, and we expect that the SKA will continue that tradition. From the standpoint of follow-up observations, the requirements on the instrument are not significantly different than those required for high-resolution, high-dynamic range imaging.

Following C06, we divide transients into two broad classes, “slow” and “fast,” based on the technique most suitable for searching for them. We discuss the search techniques for these two different classes, then conclude with a discussion of certain hardware-related issues.

7.3 Slow Transient Searching

Slow transients are those that can be found by repeated imaging of a given (“large”) section of sky or “tiling” the sky. As such, at least from the standpoint of *detection*, the time scale over which they evolve must be comparable to or longer than the time it takes to revisit the region of sky surveyed.

As a fiducial slow transient search, we consider an all-hemisphere imaging survey. We take the design goal of a FoV of 1 deg^2 at 1 GHz. Allowing for a 3-s integration per FoV (including survey overhead), the entire hemisphere ($20\,000 \text{ deg}^2$) could be surveyed in less than 1 day. The nominal point-source sensitivity of such a survey would be $50 \mu\text{Jy}$, for a 100 MHz bandwidth and assuming only the inner 25% of the SKA is used. For comparison, Amy, Large & Vaughan (1989) placed a limit of $0.017 \text{ events s}^{-1} \text{ sr}^{-1}$ on radio transients at 0.84 GHz with flux densities greater than 10 Jy (0.4 Jy) having durations of 1 ms (25 ms), and Katz et al. (2003) placed a limit of $1.5 \times 10^{-8} \text{ events s}^{-1} \text{ sr}^{-1}$ on *non-solar* radio transients at 0.61 GHz with flux densities greater than about 27 kJy and having durations shorter than a few minutes.

At a minimum, our fiducial slow-transient survey would be sensitive to transients having durations of order hours to days. Examples of such transients include γ -ray burst afterglows and extreme scattering events.

Transients having time scales shorter than the observation duration (~ 1 s) might be able to be detected, though probably at the cost of compensating for dispersion effects in the imaging (see below).

Obviously, there are a number of variants on such a survey. One could consider restricting the survey to the Galactic plane, in which case more time could be spent on each pointing or the pointings could be observed multiple times in a day. At lower frequencies, the FoV becomes larger so that the sky could be surveyed either deeper or more often. Also, as C06 discusses, X- and γ -ray transient surveys have been quite successful by sacrificing sensitivity for large FoVs. Imaging with sub-arrays of the SKA would yield a speed-up factor of N_{sub} (where N_{sub} is the number of sub-arrays) for surveying at an equivalent cost in sensitivity.

The primary requirement on such a survey is the development of an imaging pipeline. Imaging pipelines have been developed for comparable surveys, so we do not regard this as requiring substantial methodological or technological development. An example of such an imaging pipeline was one used for the 1400 MHz NVSS (Condon et al. 1998), which imaged the full sky accessible to the VLA with an FoV of 0.2 deg^2 .

C06 discusses the processing rate required for a slow transient search. For a survey aimed at detecting only slowly varying sources, for which dedispersion is not important, roughly only 10^{10} operations s^{-1} per FoV would be required. Moreover, slow transient searching is potentially an embarrassingly parallel problem. To first order, each FoV can be processed independently of each other. Indeed, existing sky surveys (such as the 74 MHz VLA Low-frequency Sky Survey) already utilize simple parallelization, and parallelization could offer a means for more rapid processing, particularly if sub-arrays are used to increase the FoV.

Finally, slow transient searching allows for a natural growth path for the SKA. When the SKA first comes on-line, the initial searches could be for sources not requiring dedispersion. As computational power, the sophistication of the imaging pipeline, and experience with the telescope increases, the slow transient search could transition to a search that also looked for dispersed signals.

7.4 Fast Transient Searching

Fast transients are those that evolve on a time scale sufficiently rapid that it is not useful to consider re-pointing the telescope. Examples of fast transients include a deep search of a globular cluster for pulsars, a search of the central region of the Virgo cluster for giant pulses, or searching for radio emission from extrasolar planets. As such, fast transient searching may often be more productive when considered as a *non-imaging* use of the telescope.

In this regard, the processing associated with pulsar observations with the Parkes multi-beam system and the Arecibo L-band focal-plane array (ALFA) can be considered prototypes for the SKA. As a fiducial survey, we consider searching a single field for 10 hr. Then a deep pulsar search could approach a sensitivity of 500 nJy; for comparison, Ransom et al. (2005) were able to detect pulsars of order $10 \mu\text{Jy}$ from a deep (6.5 hr) observation with the GBT. However, if the entire FoV can be pixellated, the initial positions of the transients detected will be considerably more precise than can be obtained from current single-dish observations. This is perhaps less of an issue for targeted searches (e.g., pulsars in globular clusters or radio emission from extrasolar planets), though it may be an issue for follow-up of other kinds of transients (e.g., localizing the giant pulses from a nearby galaxy). Somewhat lower sensitivities will be obtained for objects not having the highly periodic signals of pulsars, nonetheless the SKA will improve dramatically upon current limits for fast transient searching.

C06 discusses the processing rate required for a transient search in which one dedisperses the signals received from a pixellated FoV. For nominal parameters, he estimates $10^{14.4}$ operations s^{-1} per FoV. Moreover, with improvements both from Moore's law and from algorithm development, it may be possible to expand to more complicated signal descriptions rather than simply periodic emissions, such as pulsars. Searching for more complicated time-frequency signal behavior may be useful for projects such as SETI or searches for radio emission from extrasolar giant planets.

The primary requirement on fast transient searching is the development of search algorithms and accurate RFI excision. As noted above, in some respects, the existing (single-dish) multi-beam systems provide a means for useful initial developments in these aspects.

7.5 Hardware Aspects of Transient Searching

While there will clearly be a need for considerable processing hardware associated with transient searching, here we highlight two other aspects of transient searching.

7.5.1 Data Archive

An SKA data archive offers a possible means for searching for transients or perhaps even a means of implementing a slow-transient survey in a commensal manner. Consider imaging a 100 deg^2 FoV with $1'$ resolution, which requires forming an approximately 1000×1000 pixel image. For instance, such images could be produced at frequencies near 1 GHz from a core array composed of small dishes outfitted with focal plane arrays.

Even if such images were produced every second, the data rate would be less than 10 MB s^{-1} , which is much less than that produced by some extant radio telescopes. Over the course of a year, such an image archive would require about 250 TB. Given that petabyte data storage systems exist today, an image archive should be able to be implemented by the time the SKA becomes operational.

Clearly, various trade-offs exist for an image archive. For instance, higher resolution images (allowing for more precise localization) could be formed on longer time scales (e.g., 10 s). Nonetheless, an image archive would be a valuable, long-term means of searching for transients and monitoring more slowly varying sources.

7.5.2 A Transient Buffer

A possibility that has been discussed regarding transients is to outfit the SKA with additional hardware to provide a “look-back” buffer. If the individual elements or stations in the SKA were to be equipped with a RAM or disk buffer, a trigger from another instrument, or the SKA itself, could be used to re-phase the telescope in another direction, effectively allowing *ex post facto* observing.

Providing this capability imposes a number of requirements on the telescope. In addition to the obvious requirement of a buffer for each element or station, there would be additional monitor & control (M&C) software as well as a data communication channel.

The most simple operations concept would be one in which M&C software would be notified of the existence of a trigger and allowed to override the regular data acquisition process. At each element or station, the signals would be written continuously into a rolling buffer until notified by the M&C system that the buffer would need to be dumped. In principle, data communications in this operations concept need not be any more complex than required by regular observations. The M&C communication to the elements or stations might require only a few bits to request that the buffer be dumped rather than the regular element or station signal, while the communication of the buffer from the element or station to the central processing facility would be no more demanding than the regular data acquisition.

Below we shall illustrate how such a transient buffer might be implemented, based on current technology. More importantly, however, we have the following recommendations:

1. The implementation of “look-back” buffers on telescopes under or soon to be under construction should be evaluated after experience has been developed from these buffers. Instruments that could or will be outfitted with buffers include a radio instrument designed to search for low-frequency prompt emission associated with γ -ray bursts (GASE), the Low Frequency Array (LOFAR), and the Long Wavelength Array (LWA).
2. Careful examination of the observing programs of the KSPs must be conducted. For instance, while a brief interruption of a deep H I survey, as envisioned by the “Galaxy Evolution & Cosmology” Project, may be tolerable, a brief interruption in a pulsar timing observation, as required by the “Strong Field Tests of Gravity Using Pulsars and Black Holes” Project, may not be tolerable.
3. The size of the transient buffer at each element or station should be as large as possible while keeping the cost of the additional hardware and software for the buffer subsystem to be a fixed and small percentage of

Table 6. SKA Transient Buffer Options

Type	Cost	Size (TB)	Maximum Data Rate (TB s ⁻¹)
Disk	USD\$10M	10000	2
RAM	USD\$10M	20	80

Table 7. SKA Transient Buffer: Disk vs. RAM

Number of Beams (stations or antennas)	Bandwidth (per beam, GHz)	Disk Storage Time (s)	RAM Storage Time (s)	Data Rate (TB s ⁻¹)
1	10	10 ⁶	2000	0.010
1000	1	10 ⁴	20	1
1000	10	1000	2	10 ^a
10 ⁴	5	200	0.4	50 ^a
10 ⁵	0.5	200	0.4	50 ^a
10 ⁶	0.05	200	0.4	50 ^a

^a This data rate cannot be obtained with a disk based system (with 2005 disk technology and a cost limit of US\$10M).

the total cost of the SKA. As a nominal value for the cost of the buffer subsystem, we suggest 1% of the total SKA cost.

The final recommendation follows from the first. While a transient buffer could provide a novel observing capability, acquiring such a capability will impact other capabilities of the telescope. Thus, we recommend that the cost of a transient buffer be kept to a modest level.

As an illustration that an effective transient buffer could still be obtained for a modest amount, we consider two types of transient buffers, each costing US\$10M (about 1% of the SKA cost), each using 2005 technology and prices. Of course, 2010 technology will store more, provide higher bandwidth, and/or cost less, so these illustrations are largely in the vein of an “existence proof.”

Table 6 summarizes the characteristics of the two buffers types, disk and RAM. A disk based buffer has 500 times more storage but has 40 times less bandwidth than a RAM based buffer.

Table 7 shows the storage times and data rates for disk and RAM buffers for a variety of SKA configurations (assuming a single polarization per beam, storing 4 bits real and 4 bits imaginary).

Both disk and RAM buffers have considerable computational capability. For instance, a RAM buffer can perform 320 teraops/second. Because of this super-computer ability, the transient buffer should probably not be built as a separate dedicated instrument that can only be used as a buffer, but the buffer should be part of a general purpose compute cluster.

Disk Drive Based Data Buffer: One could assemble a disk based transient storage buffer of 10 PB. Such a buffer would utilize AMD Opterons running Linux, a software RAID system, octal SATA controllers, and contain 20 000 SATA disk drives (0.5 TB per drive), and provide a maximum data transfer rate of 2 TB s⁻¹, which corresponds to a maximum (beam × bandwidth) product of 2 × 10¹² beams Hz.

Such a system will fill 60 equipment racks. SATA disk drives fail at 7% per year, so the annual replacement cost would be US\$0.5M (using 2005 drive prices). SCSI drives fail at a rate of 5% per year, but SATA drives are considerably cheaper.

The limitation of a 2005-technology, disk-based transient buffer is the data rate. Each SATA drive can record 100 MB s⁻¹, so the maximum data rate of this 20 000 drive system is 2 TB s⁻¹.

RAM Based Data Buffer: A RAM-based buffer would likely be built using a bank of identical PC boards. In our strawman design, which is similar to UC Berkeley's BEE2 board, the basic storage module would consist of a Xilinx XC4VP20 FPGA chip (US\$335) and two DDR2 GByte DRAMs (US\$80 each).

Each DDR2 DRAM can store data at 4 GB s^{-1} using Xilinx's DDR2 controller. Each FPGA has eight 10-Gb s^{-1} ports for high-speed I/O, so a single module could store 2 GB at a rate of 8 GB s^{-1} , with an estimated cost of US\$1000 per module including PCB, power supplies, racks, cooling and connectors.

Each module's FPGA can calculate 32 Gops/sec, so the transient buffer could also be used as a 320 teraop/sec supercomputer with 80 TB s^{-1} I/O.

ACKNOWLEDGEMENTS

Basic research in radio astronomy at the NRL is supported by the Office of Naval Research.

REFERENCES

- Amy, S. W., Large, M. I., & Vaughan, A. E. 1989, *Pub. Astron. Soc. Au.*, 8, 172
Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. 1998, *AJ*, 115, 1693
Cordes, J. M., Lazio, T. J. W., & McLaughlin, M. A. 2005, *New Astron. Rev.*, 48, 1459 (SKA Science Book)
Katz, C. A., Hewitt, J. N., Corey, B. E., & Moore, C. B. 2003, *PASP*, 115, 675
Ransom, S. M., Hessels, J. W. T., Stairs, I. H., Freire, P. C. C., Camilo, F., Kaspi, V. M., & Kaplan, D. L. 2005, *Science*, 307, 892
Wilkinson, P. N., Kellermann, K. I., Ekers, R. D., Cordes, J. M., & Lazio, T. J. W. 2005, *New Astron. Rev.*, 48, 1551