Computing Cost of Sensitivity and Survey Speed for Aperture Array and Phased Array Feed Systems

Stefan J. Wijnholds
Netherlands Institute for Radio Astronomy (ASTRON)
Oude Hoogeveensedijk 4
7991 PD Dwingeloo
The Netherlands

Rik Jongerius
IBM Netherlands
Postbus 9999
1006 CE Amsterdam
The Netherlands

***Abstract*—Aperture array (AA) and phased array feed (PAF) systems are envisaged to play an important role in the Square Kilometre Array (SKA), because their multi-beaming capability can be used to improve survey speed and enhance observing flexibility. In this paper, we demonstrate that computing costs and I/O rates should be considered as an integral part of the overall system design. We do so by comparing the correlator and imaging computing requirements for the current SKA phase 1 (SKA_1) AA-low baseline design with those for an alternative design with the same survey speed and sensitivity. We also compare the correlator and imaging computing demands and survey speed for the proposed SKA_1 survey array (dishes with PAFs), the envisaged SKA_1 dish array (dishes with single pixel feeds (SPFs)) and an AA-mid alternative design for the 300 – 1000 MHz range. We conclude that the current SKA_1 baseline design may not be the optimal solution in view of computing requirements (hence operating costs) for given sensitivity and survey speed.**

I. Introduction

The Square Kilometre Array (SKA) is a future radio telescope envisaged to provide at least an order of magnitude more sensitivity and survey speed than current instruments. Phased array technology plays a key role in the SKA design [1], either in aperture array (AA) or in phased array feed (PAF) systems. This technology allows multi-beaming, which significantly enhances survey speed and observing flexibility compared to dishes with single pixel feeds. In the SKA phase I (SKA_1) baseline design, AA technology is used to cover the 50 – 300 MHz range and is part of the advanced instrumentation program towards SKA phase II for frequencies up to 1.4 GHz, while PAF systems are a crucial part of the SKA_1 survey telescope covering the 650 – 1670 MHz range.

The performance of AA and PAF systems will, probably, initially be limited by the available data transport capacity (I/O rates) and computing resources. Since those are expected to improve over time, this provides a natural future upgrade path. Since they will also be significant cost drivers for the operating costs of the SKA, it is interesting to look into the required resources for the subsystems in the SKA_1 baseline design. We also include an AA-mid alternative for the 300 – 1000 MHz range and a previously proposed alternative for the AA-low system in our analysis. In total, five different subsystems for the SKA are considered in our analysis, which are described in the next section.

II. Top-level description of SKA subsystems

A. AA-low

The AA-low system described in [1] consists of 911 stations with 35-m diameter spread over a circular area with a diameter of 50 km and 289 receiving elements each. This system covers the 50 – 300 MHz frequency range. Since the low noise amplifiers (LNA) are not cryogenically cooled, the receiver temperature will be $\sim$ 50 K. Each station forms a single beam. In our analysis, we will assume that the station beamformer output samples will be encoded in 8 bit per sample.

B. AA-low, alternative option

We will see that the AA-low system described in the previous subsection leads to considerable computing requirements. This issue was raised earlier by Paul Alexander, who proposed a system with 200 75-m stations in the core area and 85 70-m stations outside that area [2]. In our analysis we will take 280 75-m stations with 940 antennas per station, keeping the total number of antennas fixed. We will also assume the same maximum baseline (50 km), frequency range (50 – 300 MHz), receiver temperature (50 K) and sample word size (8 bit) as before. Since the station diameter is about twice as large as for the previous configuration, we form four beams covering the full frequency range to retain survey performance.

C. Dishes with single pixel feeds (SPFs)

Dishes with SPFs are part of SKA_1 and will be built in South Africa. The array is envisaged to consist of 64 13.5-m dishes from the MeerKAT project and 190 15-m SKA dishes [1]. The maximum baseline of the SKA_1 dish array will be 200 km. In our analysis, we will simplify this to 254 15-m dishes and will consider the frequency range from 300 to 1000 MHz. The SPFs will be cryogenically cooled to bring the receiver temperature down to 20 K. We will further take a constant aperture efficiency of 70% over the aforementioned frequency range and assume 8 bit sample size.

D. Dishes with PAFs

The SKA_1 survey telescope is envisaged as an extension of the 36-dish Australian SKA Pathfinder with 60 15-m dishes placed in a 50-km area [1]. For the analysis in this paper, we will assume 96 15-m dishes with 64-element PAF for the

978-1-4673-5943-6/13/$31.00

300 – 1000 MHz range. Because PAFs are too large to cool cryogenically, their receiver temperature will be about 50 K. For simplicity, we take 70% aperture efficiency for all beams over the entire frequency range. The SKA_1 survey instrument will form 30 beams at the highest frequency. The number of beams will be reduced towards lower frequencies to provide a FoV that is more or less constant over the covered frequency range. Again, we will assume a sample size of 8 bit.

E. AA-mid

To show the potential of an AA-mid system, we present an alternative with similar sensitivity as the dish array with SPFs and similar correlator load as the SKA_1 survey telescope. We demonstrate that its larger field-of-view provides a factor 3 – 10 times higher survey speed than with the other SKA_1-mid systems and that this can be achieved with far less computational cost for imaging. This system consists of 64 35-m stations covering the 300 – 1000 MHz frequency range. Each station consists of 376 tiles consisting of 8×8 antennas with 0.2 m pitch. Since this system is uncooled, we assume a receiver temperature of 50 K. We further assume a maximum baseline of 50 km (as with SKA_1 survey) and 8 bit sample size. To match the correlator compute load of the SKA_1 survey system, we form 70 beams at the highest frequency trying to keep the FoV as constant with frequency as possible.

III. PERFORMANCE ASSESSMENT

In this section, we describe the various performance metrics and data transport and computing requirements that we calculate for each of the subsystems described in the previous section. We will use the word "station" for both dishes (with PAFs or SPFs) and AA stations.

A. System temperature

The system temperature T_{sys} is calculated by adding the receiver temperature T_{rec} (specified in the previous section) to the sky noise temperature T_{sky} given by

$$T_{sky} = (\lambda/\lambda_0)^{2.55} + (f/f_0)^{1.8} + T_{bg}, \tag{1}$$

where λ is the observing wavelength, $\lambda_0 = 0.2008$ m, f is the observing frequency, $f_0 = 10$ GHz and T_{bg} is the temperature of the cosmic background (assumed to have a constant value of 2.7 K in our analysis).

B. Effective area

For the AA systems, we assume elements with a low directivity to ensure a large scan range. Such elements have a typical effective area per element of $\lambda^2/3$. At the low frequency end, the physical area available for each element, calculated by dividing the total aperture size of the station by the number of elements, may constrain the effective area. We therefore use

$$A_{eff,e} = \min\left\{\frac{1}{4}\pi D_{stat}^2/N_e \;,\; \lambda^2/3\right\} \tag{2}$$

to calculate the effective area per element. Here, D_{stat} is the diameter of the station and N_e is the number of elements in a station. The total effective area then becomes

$$A_{eff} = N_s N_e A_{eff,e}, \tag{3}$$

where N_s is the number of stations in the array.

For dish systems (either equipped with PAFs or SPFs) with aperture efficiency η, the effective area is determined by

$$A_{eff} = \eta N_s \left(\frac{1}{4}\pi D_{stat}^2\right). \tag{4}$$

C. Sensitivity

The sensitivity is simply calculated as A_{eff}/T_{sys}. With A_{eff} as described in the previous section, this value corresponds to bore sight or zenith sensitivity for AAs.

D. Survey speed

The survey speed is given by

$$SS = \left(\frac{A_{eff}}{T_{sys}}\right)^2 \times FoV, \tag{5}$$

where *FoV* stands for the total field of view. For a single beam, the FoV is taken to be $\pi/4$ times the square of the half power beam width. For SKA-mid systems with multi-beaming capability, we have defined the number of beams at the highest frequency. This allows us to compute the total FoV at the highest frequency by multiplication of the single beam FoV with the number of beams. Towards lower frequencies, we decrease the number of beams by one each time we reach a frequency at which one beam less can deliver the same FoV. In other words, we try to keep the FoV of the multi-beam systems as constant over frequency as possible.

E. Station beamformer output data rates

For all systems, we assume that the beamformer covers the full specified frequency range. For a single beam, this means that the total beamforming bandwidth Δf_{BF} is simply the difference between the maximum and minimum frequency. For multi-beaming, we have to split the frequency range into intervals with a constant number of beams, following the procedure outlined in the previous subsection, and sum the total beamforming bandwidth for each interval over all intervals. Assuming that the beamformer output data are Nyquist sampled and that the beamformer produces full polarimetric output, the output data rate per station will be

$$ODR_{BF} = 4\Delta f_{BF} N_{bit}, \tag{6}$$

where N_{bit} is the number of bits per sample.

F. Correlator processing power

The correlator input data rate is simply the output data rate per station multiplied by the number of stations in the array. The required correlator processing power is

$$P_{cor} = 4\Delta f_{BF}\,(2N_s)^2 \tag{7}$$

real valued multiply or add (floating point) operations per second.

G. Correlator output data rate

To minimize the correlator output data rate ODR_{cor}, we want to make our frequency channels as wide as possible and integrate as long as possible unless there are strong scientific reasons to opt for higher time or frequency resolution. The integration time of the correlator, or correlator dump time, and correlator channel width are limited by the time it takes for the longest baseline to produce a new, fully independent visibility sample. Limiting the signal degradation on the longest baseline to 2%, the correlator dump time should be at most [3]

$$\tau = 1200 \frac{D_{stat}}{B_{max}} \tag{8}$$

and the channel width should be at most [3]

$$\Delta f = \frac{1}{10} \frac{D_{stat} f}{B_{max}}. \tag{9}$$

For f we will assume the lowest frequency in the range of operating frequencies. The correlator will therefore have to produce $N_{ch} = \lceil \Delta f_{BF} / \Delta f \rceil$ frequency channels. This leads to a correlator output data rate of

$$ODR_{cor} = 32 N_{ch} \left(2 N_s\right)^2 / \tau, \tag{10}$$

assuming single precision floating point samples (32 bit) and a visibility matrix with the real parts stored in the upper triangle and the imaginary parts in the lower triangle.

H. Computing power for imaging

The scaling laws for imaging depend strongly on the algorithm. In our analysis, we will use the scaling law for the w-projection algorithm, since it is a well established algorithm, that is easily extended to include direction dependent effects by transforming it to the AW-projection algorithm. For simplicity, we focus on the w-projection imaging only and ignore other common operations like source subtraction and deconvolution. The main goal here is to make a relative comparison of different system designs in terms of computing costs of imaging. Since gridding dominates the total computing costs, we ignore the FFT-stage. The computing cost of w-projection for a 1 MHz continuum image can then be estimated using [4], [5]

$$P_{im} = N_{op} \underbrace{\frac{10^6}{\Delta f} \left(2 N_s\right)^2 \frac{T_{obs}}{\tau}}_{N_{vis}} \left(\left(\frac{\lambda_{max} B_{max}}{D_{stat}^2} \right)^2 + N_{kernel}^2 \right) \tag{11}$$

where N_{op} is the number of operations required to grid one visibility to one grid point, T_{obs} is the duration of the observation, N_{kernel} is the size of the convolution kernel and N_{vis} is the total number of visibilities that need to be gridded. The number of operations required to grid one visibility to one grid point depends strongly on the complexity of the convolution kernel and whether it has been pre-computed or not. For our analysis, we assume the minimal 7×7 kernel has been pre-computed, such that $N_{op} = 10$ suffices.

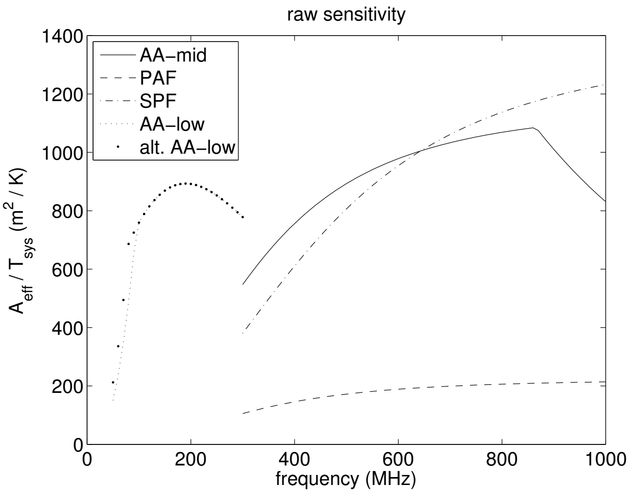

Fig. 1. Sensitivity as function of frequency for the SKA-low and SKA-mid subsystems described in Sec. II.

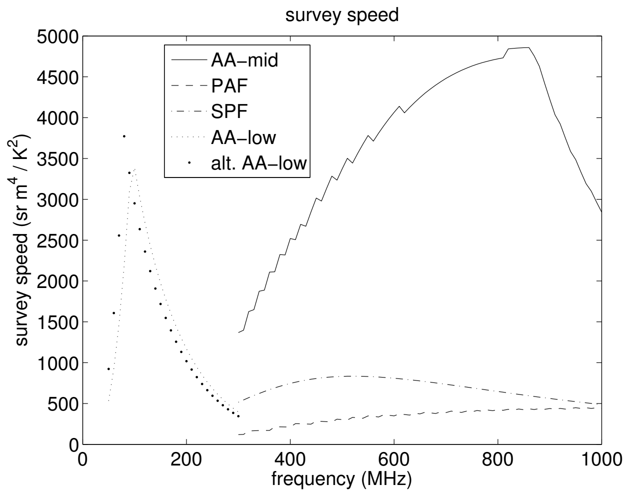

Fig. 2. Survey speed as function of frequency for the SKA-low and SKA-mid subsystems described in Sec. II.

IV. Results

Figures 1 and 2 show the sensitivity and survey speed for each of the five subsystems described in Sec. II. The sensitivity plot shows that the two AA-low options have the same raw sensitivity over a wide range of frequencies. The antenna arrangement in the 75-m stations is slightly sparser, pushing the turnover caused by complete filling of the station aperture towards a slightly lower frequency. The plot also shows that the AA-mid system has a sensitivity comparable to the dish array with SPFs over the 300 – 1000 MHz frequency range. In this range, the sensitivity of SKA_1-dish is significantly higher than SKA_1-survey, because it has over 2.5 times more collecting area and cryogenically cooled receivers.

As can be seen from Table I, the number of beams for the AA-mid system has been chosen to match the correlator compute load of the dish array with PAFs. Figure 2 shows that

TABLE I

COMPARISON OF I/O AND PROCESSING LOAD FOR DIFFERENT SYSTEM CONCEPTS.

	dish + SPF	dish + PAF	AA-mid	AA-low	alt. AA-low
Station beamformer output data rate (Gbps)	22.4	322	734	8	32
Correlator input data rate (Tbps)	5.69	31.0	47.2	7.29	8.96
Correlator processing power (Tflops)	722.6	1486	1511	3320	1254
Max. correlator dump time (s)	0.09	0.36	0.84	0.84	1.8
Max. correlator channel width (kHz)	2.25	9.0	21.0	3.5	7.5
Correlator output data rate (Tbps)	28.5	3.70	0.69	9.03	0.74
1D image size (pixels)	13333	3333	1429	1429	667
Image processing per MHz (Tflops)	10056	5.61	0.016	677	0.67
Total image processing (all beams, Tflops)	7.0e6	56.6e3	367	1.69e5	671

with this number of beams, the survey speed of the proposed AA-mid system is a factor 3 – 10 higher than that of both the SKA_1-dish and the SKA_1-survey system proposed in [1]. It also shows that the array with PAFs cannot make up for its reduced sensitivity compared to the dish array with SPFs in the 300 – 1000 MHz range, but, extrapolating, this will be different in L-band. Figure 2 also shows that the alternative AA-low system can compensate for its reduced station FoV compared to the baseline design by forming four beams over the full frequency range. The two AA-low systems described in Sec. II thus have very similar sensitivity and survey speed.

Table I summarizes the I/O rates of the correlator and processing power required to correlate and image the data at the same pace as the data are observed. It should be noted that the I/O rates do not include meta-data or encoding and that computing costs for channel separation are not taken into account. Despite the fact that sensitivity and survey performance of the two AA-low systems are comparable, the required correlator processing power and in particular the imaging processing power differ significantly. This can be explained by the scaling of the image processing given by (11). The number of visibilities depends on the correlator dump time and the channel width, which both depend on D_{stat} and B_{max}. These values also enter via the number of w-planes leading to a scaling proportional to B_{max}^4 and D_{stat}^{-6}. This also gives the AA-stations an advantage over dishes, because they can make up for the loss in FoV per beam by forming more beams for which the image processing scales linearly. The dish array with PAFs mainly requires less image processing power than the dish array with SPFs, because its longest baseline is a factor four shorter. Despite its much higher survey speed, the compute load for image processing with the suggested AA-mid system is more than two orders of magnitude lower than the compute load required for the SKA_1-survey system.

V. CONCLUSIONS

Imaging using w-projection scales with B_{max}^4 and D_{stat}^{-6}. This may be reduced to scaling with B_{max}^3 and D_{stat}^{-4} by w-snapshots [6], but even then the scaling with the maximum baseline and station diameter remains very strong. Reducing the maximum baseline of the proposed dish array with SPF to 50 km would reduce the total image processing to $27.5 \cdot 10^3$ Tflops, i.e. less than for the dish array with PAFs, while providing much more raw sensitivity and better survey performance. In terms of operating cost, a dish array with SPF may therefore be more suitable for the 300 – 1000 MHz range than a dish array with PAFs. An AA-mid array may be an even better option given its larger multi-beaming flexibility and upgradability if computing resources improve. For similar reasons, the alternative AA-low system with 280 75-m stations may be more attractive than the current baseline design with 911 35-m stations. A key conclusion is, that the SKA subsystems should not only be designed for calibratability to achieve the desired imaging performance [7], but also for feasibility of the imaging operation itself, since an unfortunate choice of front-end design may have significant consequences for computing requirements and hence operating costs.

In our analysis, the correlator dump time and correlator channel width are driven down by the longest baseline. Most baselines will be significantly shorter allowing for larger correlator dump times and channel widths. As an aside, we note that using different correlator dump times and channel widths for short and long baselines may significantly reduce the correlator output data rate and number of visibilities to be gridded.

REFERENCES

[1] P. E. Dewdney *et al.*, "SKA1 System Baseline Design," SKA Project Office, Manchester (UK), Tech. Rep. SKA-TEL-SKO-DD-001, 12 Mar. 2013.

[2] P. Alexander, "AA Processing Requirements for SKA1," in *Aperture Array Verification Program Workshop*, Dwingeloo (The Netherlands), 12-16 Dec. 2011.

[3] P. Alexander, A. J. Faulkner, K. Grainge, and J. Lazio, "Software and Computing CoDR - Analysis of Requirements Derived from the DRM," SKA Project Office, Manchester (UK), Tech. Rep. WP2-050.020.010-RR-001, 17 Jan 2012.

[4] T. J. Cornwell, K. Golap, and S. Bhatnagar, "The Noncoplanar Baselines Effect in Radio Interferometry: The W-Projection Algorithm," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 647–657, Oct. 2008.

[5] M. Yashar and A. Kemball, "Computational Costs of Radio Imaging Algorithms Dealing with the Non-coplanar Baselines Effect," SKA Project Office, Manchester (UK), Tech. Rep., 6 Nov. 2009, tDP CPG Memo 17.

[6] T. J. Cornwell, M. A. Voronkov, and B. Humphreys, "Wide field imaging for the Square Kilometre Array," in *Image Reconstruction from Incomplete Data VII*, ser. SPIE Conference Series, P. J. Bones, M. A. Fiddy, and R. P. Millane, Eds., vol. 8500, San Diego, California (USA), Aug. 2012.

[7] S. J. Wijnholds, J. D. Bregman, and A. van Ardenne, "Calibratability and its impact on configuration design for the LOFAR and SKA phased array radio telescopes," *Radio Science*, vol. 46, no. RS0F07, 8 Nov. 2011.