IMPACT OF STATION SIZE ON CALIBRATION OF SKA-LOW

Cathryn M. Trott

Curtin University, International Centre for Radio Astronomy Research (ICRAR)
ARC Centre of Excellence for All Sky Astrophysics
Bentley WA 6845 Australia

ABSTRACT

The size of SKA-Low stations affects calibration of the instrument, and the science that can be performed. We analyse the differences in instrumental calibration when varying station size, for a fixed total effective area and array layout. The Cramer-Rao Bound (CRB) is used as a metric to quantify the theoretical optimal performance for estimating direction-dependent station gains (amplitude and phase), in the presence of radiometric (thermal) noise and noise from unmodelled background point sources, and for multiple field calibrators. The multiplicity of the calibrators is approximated by the ratio of the station field-of-view (FOV) to the ionospheric isoplanatic patch size. Smaller stations have larger sky FOV, increasing both the noise from unmodelled background sources and number of calibrators required to determine the full gain characteristics, but are calibrated with a larger number of measurements (more baselines) compared with larger stations. We find that the number of calibrators has little impact on the estimation precision, and the results depend on the noise characteristics of the data. For thermal noise-dominated datasets, larger stations yield improved estimation performance. For datasets where unmodelled background point sources dominate the error budget, performance is comparable for large and small stations.

Index Terms— Calibration, signal processing, interferometers, estimation

1. INTRODUCTION

SKA-Low is a wide-field radio interferometer, and will contend with variations in calibration across the field-of-view. As such, it will make use of 'field-based' calibration, whereby calibration sources across the field are used to form a time-dependent model for the ionosphere [1, 2, 3]. [4] (hereafter, WV09) performed an analysis of multisource self-calibration, and described four ionospheric regimes for an instrument. They studied the third regime, where all stations observe a single source through the same ionospheric conditions, but sources sample different ionospheric conditions. In this scenario, which occurs for large fields-of-view and short baselines, a single calibration source cannot be used to determine station gain parameters, and independent calibration must be performed for multiple sources sampling the ionosphere. WV09 considered CRBs for the direction-independent gain amplitude and phase for each station, as well as the direction-dependent gain amplitude toward multiple sources in the FOV, embedded within radiometric noise. They did not consider the effect of varying station size, or of the additional signal noise contributed by background sources.

Increasing station size leads to smaller sky FOV. For the same ionosphere, fewer calibration sources are required to determine the full direction-dependent gain parameters for a station. In addition, smaller FOVs yield lower background source noise, due to the smaller sky area. Conversely, for a constant effective area of the instrument, larger stations means fewer baselines, and less information about the sky (fewer measurements). In this note we consider the effect of station size on calibration of direction-dependent station gain parameters, for a fixed total effective collecting area and baseline distribution. Point source calibrators do not reside alone in a field, but are accompanied by the other sources present in the field. This contributes unwanted signal to the measurements. Background source contamination can be modelled statistically and considered as an additional component of noise.

We start by describing the instrument and experiment design, and then present the CRB formalism, and the form of the signal. We then compute the additional noise in the visibility measurement due to unmodelled background sources, and conclude by computing the gain parameter estimation precision for each array.

2. INSTRUMENT AND EXPERIMENT DESIGN

Throughout we use SKA-Low parameters, as described in the SKA System Baseline Design [5]. We construct a proxy SKA-Low antenna layout based on the broad description within the document: 866 35 metre stations placed randomly within a 3000 m radius, and 45 stations placed on three spiral arms to a radius of $\sim$50 km. The total physical aperture is 8×10^5 m, with an effective area per dipole of $A_{\rm eff}$=1.6 m^2 at

cathryn.trott@curtin.edu.au

Array	Size (m)	Stations (core)	Stations (arms)
Baseline	35	866	45
Medium	50	410	45
Large	100	66	45

Table 1. Parameters for the three arrays considered.

160 MHz, and system temperature, $T_{\rm sys}$=370 K. We consider three arrays, with three station sizes, as outlined in Table 1. For each array, we compute CRBs on gain parameters for a contiguous bandwidth of 10 MHz and a single 10 s integration. The radiometric noise in each visibility is given by:

$$\sigma_{\rm thermal} = \frac{2k_B T_{\rm sys}}{A_{\rm eff}\sqrt{2.{\rm BW}.\Delta t}} = \frac{72.2}{A_{\rm eff}} \quad {\rm Jy}. \tag{1}$$

3. ESTIMATION PERFORMANCE: CRAMER-RAO LOWER BOUND

To determine the theoretical optimal estimation performance with a given dataset, we can calculate the CRB on the precision of parameter estimates [6, 7]. The CRB calculates the precision with which a minimum-variance unbiased estimator could estimate a parameter value, *using the information content of the dataset*. It has been used to assess the SKA calibration performance for redundant arrays [8] and calibration algorithms [9]. The CRB is computed as the square-root of the corresponding diagonal element of the inverse of the Fisher information matrix (FIM). The (ij)th entry of the FIM for a vector $\boldsymbol{\theta}$ of unknown parameters is given by:

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{ij} = -E\left[\frac{\partial^2 \log L(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right], \tag{2}$$

where E denotes the expectation value. For N independent samples in WGN and complex data, this expression simplifies to,

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{ij} = 2{\rm Re}\left[\frac{1}{\sigma^2}\sum_{n=1}^{N}\frac{\partial\tilde{s}^H[n;\boldsymbol{\theta}]}{\partial\theta_i}\frac{\partial\tilde{s}[n;\boldsymbol{\theta}]}{\partial\theta_j}\right]. \tag{3}$$

The CRB is a useful metric because it places a fundamental lower limit on the measurement precision of any parameter. In this work it will be used to gain an understanding of the calibration limits of an instrument.

The signal is expressed as the real and imaginary components, under which the noise can be modelled as white Gaussian. In the flat field approximation, the signal, $\tilde{s}[f,n]$, is the complex visibility for channel f and baseline n, and is given by:

$$\tilde{s}[f,n] = V(u_{fn},v_{fn}) = \iint B(l',m')I(l',m')\left(\frac{\nu(f)}{\nu_0}\right)^{\alpha} \exp\left[-2\pi i(u_{fn}l' + v_{fn}m')\right]dl'dm', \tag{4}$$

where $B(l',m')$ and $I(l',m')$ are the antenna response function and source intensity function at sky position (l',m'), and the spectral dependence is modelled as a power-law with index α and normalized by the base frequency, ν_0. Assuming a point source located at $(l'=l, m'=m)$, the visibility function becomes:

$$V(u_{fn},v_{fn}) = B(l,m)I(l,m)\left(\frac{\nu(f)}{\nu_0}\right)^{\alpha} \exp\left[-2\pi i(u_{fn}l + v_{fn}m)\right]. \tag{5}$$

For simplicity, we consider flat spectrum sources (α=0), and the frequency dependence (over a small fractional bandwidth) is removed from the problem. We form the likelihood function under WGN, assuming initially that the noise variance is known and identical for all baselines and channels. The likelihood function, which is the joint PDF for N baselines, is given by [10]:

$$\begin{aligned} L(\tilde{\mathbf{x}};H_1) &= \prod_{n=1}^{N}\frac{1}{\pi\sigma^2}\exp\left[-\frac{1}{\sigma^2}(\tilde{x}[n]-\tilde{s}[n])^*(\tilde{x}[n]-\tilde{s}[n])\right] \\ &= \frac{1}{\pi^N\sigma^{2N}}\exp\left[-\frac{1}{\sigma^2}(\tilde{\mathbf{x}}-\tilde{\mathbf{s}})^H(\tilde{\mathbf{x}}-\tilde{\mathbf{s}})\right] \end{aligned} \tag{6}$$

where H denotes Hermitian conjugate (complex conjugate transpose), σ is the noise per visibility (across the whole bandwidth) and the product has been collected in the matrix inner product.

For a small FOV, only one source is required to characterise the ionosphere and calibrate the station. In general, we require,

$$N_c \approx \frac{\rm FOV}{\rm FOV_{ion}} = \frac{\lambda^2}{A_{\rm eff}{\rm FOV_{ion}}}, \tag{7}$$

calibration point sources, where $A_{\rm eff}$ is the station effective area and $\rm FOV_{ion}$ is the characteristic scale of the ionosphere. In this note, we consider the patch size to have a diameter of 2 degrees ($\rm FOV_{ion} = 1\times10^{-3}$ sr), and calibrators with flux density, S = 1 Jy.

We begin by calculating the theoretical optimal precision with which the amplitude and phase calibration can be measured for a single antenna, for an $N_{\rm ant}$-antenna interferometer and N_c calibrators, with positions (l_{N_c}, m_{N_c}) and flux densities, S_{N_c}. *We determine independent gain parameters for each calibrator for each antenna.* We write the complex gain for baseline n comprising antennas β and γ, and calibrator i, as:

$$\tilde{G}_{in} = \tilde{G}_{i\beta}\tilde{G}_{i\gamma} = \frac{1}{b_{i\beta}b_{i\gamma}}\exp 2\pi i(\phi_{i\beta}-\phi_{i\gamma}), \tag{8}$$

where b and ϕ denote the amplitude and phase gain parameters. The signal is therefore given by,

$$\tilde{s}_{\beta\gamma} = \sum_{i=1}^{N_c} S_i b_{i\beta}b_{i\gamma}\exp -2\pi i(u_{f\beta\gamma}l_i + v_{f\beta\gamma}m_i + \phi_{i\beta} - \phi_{i\gamma}). \tag{9}$$

The FIM is a $2N_{\rm ant}N_c \times 2N_{\rm ant}N_c$ matrix to estimate all of the b and ϕ parameters for $N_{\rm ant}$ antennas and N_c sources. There are no covariances between the b and ϕ parameters, so the FIM is equivalent to two $N_{\rm ant}N_c \times N_{\rm ant}N_c$ matrices. Therefore, there are two FIMs to invert, FIM$_b$ and FIM$_\phi$. Constructing FIM$_\phi$ yields a singular matrix, due to the phases being relative quantities. Typically, the phase gain for one antenna is set to zero, and the others are defined relative to this, and FIM$_\phi$ has one less parameter.

4. FORM OF VISIBILITY NOISE

The noise in a visibility is ideally white Gaussian, and equal to the radiometric noise. In practice, there are other sources of uncertainty. In the context of this work, we consider the additional noise-like signal contained within a visibility due to unmodelled point sources present within the station beam. To compute the contribution of these sources, we perform the following calculation, where we consider a simplified Gaussian beam-shape, characterised by the Gaussian width, θ.

1. Calculate the Poisson noise due to the random number of sources within a small patch of sky and a small range of source flux density ($N(S, S+dS; l, l+dl; m, m+dm)$); We assume that the number of sources within a small area of the sky is Poisson-distributed. For a Poisson distribution, the variance is equal to the mean. The number of sources in a small patch of sky is given by,

$$N(S, S+dS; l, l+dl; m, m+dm) = \frac{dN}{dS} dS dl dm, \quad (10)$$

where dN/dS is the source density per unit flux density, and is given by [11],

$$\frac{dN}{dS} = 3600 S_{\rm Jy}^{-2.5} {\rm Jy}^{-1}{\rm sr}^{-1}. \quad (11)$$

2. Compute the variance on a visibility measurement due to the Poisson noise from a differential patch of sky at (l', m'); For N sources with flux density S located at sky position (l', m') within the beam $B(l, m)$, the mean visibility is given by;

$$\langle V(u,v)\rangle = NSB(l', m') \exp -2\pi i(ul' + vm'), \quad (12)$$

and the variance is[1];

$$\Sigma_{Vuv} = NS^2 B(l', m')^2. \quad (13)$$

3. Compute the total variance within a visibility from all sources (covariance matrices sum because sources are independent); The mean number of sources in the sky is given by;

$$N_{\rm Tot} = \int \frac{dN}{dS} dS dl dm, \quad (14)$$

and the total variance on a visibility is given by;

$$\begin{aligned} \Sigma_{Vuv} &= \int S^2 B(l,m)^2 \frac{dN}{dS} dS dl dm \quad (15)\\ &= 3600\pi\theta^2 \sqrt{S_{\rm max}} \ \ {\rm Jy}^2, \end{aligned}$$

where $S_{\rm max}$ is the brightest unmodelled source in the field (the peeling limit). As expected the brightest sources dominate the noise, and the variance scales linearly with the beam area.

The total Gaussian-distributed noise considered for the CRB calculations is therefore the quadrature sum of the radiometric noise and the background source noise, and is given by:

$$\begin{aligned} \sigma &= \sqrt{\sigma^2_{\rm thermal} + 3600\pi\theta^2\sqrt{S_{\rm max}}} \ \ {\rm Jy} \quad (16)\\ &\approx \frac{1}{A_{\rm eff}} \sqrt{\frac{4k_B^2 T^2_{\rm sys}}{2.{\rm BW}.\Delta t} + 3600\pi\lambda^2 A_{\rm eff}\sqrt{S_{\rm max}}}. \quad (17) \end{aligned}$$

Therefore, the background source noise variance scales with the inverse effective area, while the thermal noise scales with its square. This means that the relative contribution of the background noise is *greater* for larger stations. For fiducial values of the peeling limit and the observational setup described here, Table 2 lists the noise in an individual visibility for each array we consider. It also lists the number of calibrator sources required. Note that for highly peeled datasets, the thermal noise becomes more dominant, and the smaller stations have considerably worse noise (relative to the others).

Array	$A_{\rm eff}$ (m^2)	$S_{\rm max}$ (Jy)	σ (Jy)	N_c
Baseline	461	0.1	1.78	4
Baseline	460	0.01	1.00	4
Baseline	460	0.001	0.56	4
Baseline	460	0.0001	0.32	4
Baseline	460	0.0	1.4×10^{-2}	4
Medium	940	1	1.25	2
Medium	940	0.1	0.70	2
Medium	940	0.01	0.39	2
Medium	940	0.0001	0.22	2
Medium	940	0.0	6.9×10^{-3}	2
Large	3800	1	0.62	1
Large	3800	0.1	0.35	1
Large	3800	0.01	0.20	1
Large	3800	0.0001	0.11	1
Large	3800	0.0	1.7×10^{-3}	1

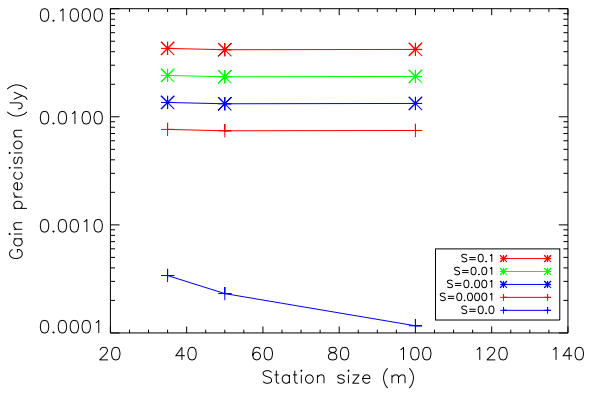
Table 2. Total noise on an individual visibility for the three arrays considered.

[1]We ignore covariances between baselines; in practice, baselines sampling closely-spaced regions of the uv plane will have correlated noise

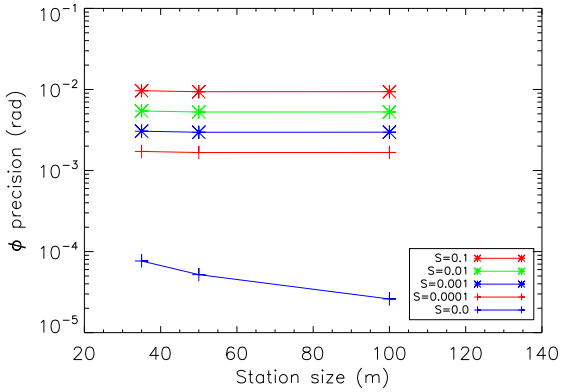
However, the background source noise dominates the error budget in most cases, alleviating some of the noise degradation of the smaller stations. This has the counter-intuitive effect of improving the estimation precision for small stations compared with large stations (although the overall noise per visibility is worse for the small stations, the noise uncertainty is only scaling as $1/\sqrt{A_{\rm eff}}$ when background sources remain).

5. RESULTS AND INTERPRETATION

Figure 1 presents the gain and phase precision for each array and each point source peeling limit. When there are no background sources, and the noise is purely thermal, the larger stations yield improved precision compared with the smaller stations. In this regime, the smaller FOV requires fewer calibrator sources. When background sources are included, this term dominates the error budget, and larger stations perform relatively worse. In this regime, there is little difference between the arrays; the additional uncertainty due to the larger number of calibrators required (larger FOV) is effectively countered by the additional measurements used to calibrate the instrument (larger number of baselines). We performed the same calculation for a single calibrator for each antenna, and found comparable results. Multiple calibrator sources do contribute additional uncertainty, but we found that this had minimal impact for fewer than $\sim$10 calibrators.



(a) Gain precision (Jy).



(b) Phase precision (radians).

Fig. 1. Gain and phase precision for each antenna and calibrator source, for the three arrays considered and different values for the point source peeling limit. Each calibrator had S = 1 Jy.

The results can be interpreted using the following arguments. The thermal noise uncertainty scales as the inverse effective station area, while the background source noise scales as the square-root of the inverse effective station area. For each antenna, there are $(N_{\rm ant}-1)$ independent measurements of the gain parameters. Therefore, the CRBs for the gain of a single antenna, with a thermally-noise dominated, and a background source noise dominated array are:

$$\Delta G_{thermal} \sim \frac{\sigma_{\rm thermal}}{\sqrt{N_{\rm ant}}} \sim \frac{1}{A_{\rm eff}\sqrt{N_{\rm ant}}} \quad (18)$$

$$\Delta G_{back} \sim \frac{\sigma_{\rm back}}{\sqrt{N_{\rm ant}}} \sim \frac{1}{\sqrt{A_{\rm eff}N_{\rm ant}}}. \quad (19)$$

Because we are comparing arrays for which $A_{\rm eff}N_{\rm ant}$ = constant, the gain parameter precision is the same for each array *for an error budget dominated by background sources*. For pure thermal noise, smaller station sizes yield larger uncertainty in gain parameters (there are more parameters to estimate).

The balance between noise terms depends on many factors. For this work, we chose a small bandwidth (10 MHz) to calibrate the instrument. Use of a larger fraction of SKA-Low's bandwidth will further reduce the thermal noise component, as well as providing additional information with which to model and peel sources. The use of the point source number counts is an approximation in this work. SKA-Low's large baseline lengths will resolve most sources in the sky, and the level of peeling is uncertain.

6. CONCLUSIONS

In this work we considered the impact of station size on gain calibration parameter precision, for a fixed total collecting area and array layout. Although we considered the additional point sources required to calibrate large FOV, due to ionospheric differences towards different lines-of-sight, we found this had minimal impact for a well-behaved ionosphere. For arrays that are thermal-noise dominant, smaller station sizes yield poorer estimation performance. For arrays for which the dominant noise term is due to unmodelled background sources, performance is comparable for each array. Given the expected high sensitivity of SKA-Low, contribution of signal from background sources may be a significant contributor to the overall error budget.

7. REFERENCES

[1] D. A. Mitchell, L. J. Greenhill, R. B. Wayth, R. J. Sault, C. J. Lonsdale, R. J. Cappallo, M. F. Morales, and S. M. Ord, "Real-Time Calibration of the Murchison Wide-field Array," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, pp. 707–717, Nov. 2008.

[2] S. Wijnholds, S. van der Tol, R. Nijboer, and A.-J. van der Veen, "Calibration challenges for future radio telescopes," *IEEE Signal Processing Magazine*, vol. 27, pp. 30–42, Jan. 2010.

[3] 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, pp. 0, Nov. 2011.

[4] S. J. Wijnholds and A.-J. van der Veen, "Multisource Self-Calibration for Sensor Arrays," *IEEE Transactions on Signal Processing*, vol. 57, pp. 3512–3522, Sept. 2009.

[5] P. Dewdney, W. Turner, R. Millenaar, R. McCool, J. Lazio, and T. J. Cornwell, "SKA1 System Baseline Design," *SKA Memo Series*, 2013.

[6] H. L. van Trees, *Detection, estimation and modulation theory*, John Wiley & Sons, New York, 2001.

[7] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall, 1993.

[8] P. Noorishad, S. J. Wijnholds, A. van Ardenne, and J. M. van der Hulst, "Redundancy calibration of phased-array stations," *A&A*, vol. 545, pp. A108, Sept. 2012.

[9] S. Salvini and S. J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *A&A*, vol. 571, pp. A97, Nov. 2014.

[10] C. M. Trott, R. B. Wayth, J.-P. R. Macquart, and S. J. Tingay, "Source Detection in Interferometric Visibility Data. I. Fundamental Estimation Limits," *ApJ*, vol. 731, pp. 81–+, Apr. 2011.

[11] S. E. G. Hales, J. E. Baldwin, and P. J. Warner, "The 6C survey of radio sources. II - The zone delta = 30-51 deg, alpha = 08h30m-17h30m," *MNRAS*, vol. 234, pp. 919–936, Oct. 1988.