PERFORMANCE ANALYSIS OF CLUSTERED RADIO INTERFEROMETRIC CALIBRATION

S. Kazemi¹, S. Yatawatta², S. Zaroubi¹

¹ Kapteyn Astronomical Institute, Univ. of Groningen, The Netherlands ² ASTRON, The Netherlands

ABSTRACT

Subtraction of compact, bright sources is essential to produce high quality images in radio astronomy. It is recently proposed that 'clustered' calibration can perform better in subtracting fainter background sources. This is due to the fact that the effective power of a source cluster is greater than the power of an individual, weak source therefore enabling better calibration. In this paper, we present performance analysis of clustered calibration and show that, indeed, clustered calibration gives better results compared to un-clustered calibration, which is already seen in the previous work. The use of this analysis enables us to choose the optimal number of clusters for a given observation in an efficient way.

Index Terms— Calibration, Estimation: Cramer-Rao bounds, Interferometry: Radio interferometry

1. INTRODUCTION

Calibration of radio interferometers [1] is the procedure of estimating and correcting for the atmospheric and instrumental signal corruptions before imaging. It plays the key role in achieving the instruments desired precision and sensitivity [2]. Although calibration's performance is highly improved by various newly devised methods when the observed sources have a high enough Signal to Noise Ratio (SNR) to be distinguished from the background noises [3, 4], it is still a great challenge to calibrate at a very low SNR.

Radio interferometric images are produced using the data observed during the total observation (integration) time. However, calibration is done at shorter time intervals of that total time. This increases the noise level in data for which calibration should be executed compared to the one in the images. In other words, there are some weak sources appearing in the images whose signals are well below the calibration noise level. In calibrating for the case of having sources with a very low SNR (when SNR refers to the one obtained for a time interval during calibration is done) clustered calibration [5] exhibits a novel efficiency.

The clustered calibration method clusters sources in the sky model, considers the same corruptions for the signals of the sources at the same cluster, and solves for every cluster as a single source. The solutions are obtained for the coherency of every source cluster, which is the summation of its individual sources coherencies, containing an upgraded information level. Therefore, when the SNR is very low, their accuracy is highly improved compared with the solutions derived by un-clustered calibrations. However, since the true corruption of a given source might not be equal to the one obtained for the source cluster, this procedure might introduce an additional error which depends on the observation and the choice of source clusters. In this paper, we focus on analyzing the performance of clustered calibration, regardless of how the clusters are made. We relate this performance to the effective Signal to Interference plus Noise Ratio (SINR) obtained for each cluster. For this purpose, we use statistical estimation theory and the Cramer-Rao Lower Bounds (CRLB) [6]. Simulated observations illustrate the validity of the presented clustered calibration performance analysis.

The following notations are used in this paper: Bold, lower case letters refer to column vectors, e.g., **y**. Upper case bold letters refer to matrices, e.g., **C**. All parameters are complex numbers, unless stated otherwise. **I** is the identity matrix. The transpose, Hermitian transpose and conjugation of a matrix are presented by $(.)^T$, $(.)^H$ and $(.)^*$, respectively. The matrix Kronecker product is denoted by \otimes . E{} is the statistical expectation operator. CN and U represent the complex Gaussian and real uniform distributions, respectively.

2. CALIBRATION AND CLUSTERED CALIBRATION

In this section we represent the measurement equation of a polarimetric clustered calibration. The reader is referred to [5] for details on clustered calibration and to [7, 8] for some introduction to radio polarimetry and calibration.

Consider calibrating an observation of K radio sources provided by an N receiver synthesis array, with orthogonal polarized feeds. The induced voltage at receiver p, $\tilde{\mathbf{v}}_{pi}$, due to radiation of source *i* polarized waves, \mathbf{e}_i , is given by $\tilde{\mathbf{v}}_{pi} = \tilde{\mathbf{J}}_{pi}\mathbf{e}_i$ where $\tilde{\mathbf{J}}_{pi}$ is the 2×2 Jones matrix [7], representing some sky and instrumental corruptions in the signal.

The total signal obtained at receiver p, \mathbf{v}_p , is a linear superposition of K such signals plus the receiver noise. After correcting for geometric delays and the instrumental effects, the p-th receiver voltage is correlated to the other N - 1 receivers voltages. The correlated voltages $E\{\mathbf{v}_p \otimes \mathbf{v}_q^H\}$, referred as *visibility* [7] of baseline p - q is given by

$$\mathbf{V}_{pq} = \sum_{i=1}^{K} \mathbf{J}_{pi}(\boldsymbol{\theta}) \mathbf{C}_{i\{pq\}} \mathbf{J}_{qi}^{H}(\boldsymbol{\theta}) + \mathbf{N}_{pq}, \qquad (1)$$

where \mathbf{N}_{pq} is the baseline's additive noise and $\mathbf{C}_{i\{pq\}}$ is the Fourier transform of the *coherency* matrix [9, 7] of the *i*-th source radiation. Note that in Eq. (1), the Jones matrix $\mathbf{J}(\boldsymbol{\theta})$ corresponds to the corruptions in signals which are identified by some parameters collected in vector $\boldsymbol{\theta}$. Calibration is essentially obtaining the maximum likelihood estimation of the unknown parameter vector $\boldsymbol{\theta}$.

Our current clustered calibration makes the underlying assumption that if the angular separation between two sources i and j is small enough, the sky corruptions in their signals are identical, $\mathbf{J}_{pi} = \mathbf{J}_{pj}$. Thus, it initially designs source clusters, L_i for $i \in \{1, \ldots, Q\}$ where $Q \ll K$, on which the sky variation is

The authors acknowledge the Netherlands Foundation for Scientific research (NWO) grant 436040 for support.

considered to be uniform, and applies calibration on

$$\mathbf{V}_{pq} = \sum_{i=1}^{Q} \widetilde{\mathbf{J}}_{pi}(\widetilde{\boldsymbol{\theta}}) \{ \sum_{l \in L_i} \mathbf{C}_{l\{pq\}} \} \widetilde{\mathbf{J}}_{qi}^{H}(\widetilde{\boldsymbol{\theta}}) + \widetilde{\mathbf{N}}_{pq}.$$
(2)

In Eq. (2), \mathbf{J}_{pi} is the clustered calibration Jones which is shared by sources of cluster *i* at receiver *p*, and $\widetilde{\mathbf{N}}$ is its effective noise which is explicitly discussed at section 3.

3. PERFORMANCE EVALUATION

In this section, we explain the reason of clustered calibration's superior performance, compared to un-clustered calibration, at a low SNR. For simplicity, we first consider the case of observing two point sources in the sky using an interferometric array.

3.1. Estimations of Cramer-Rao Lower Bounds

We are concerned with calculating CRLB to limit the performances of clustered and un-clustered calibrations.

Consider observing two point sources at baseline p - q. Based on Eq. (1), the visibilities are given by

$$\mathbf{V}_{pq} = \mathbf{J}_{p1} \mathbf{C}_{1\{pq\}} \mathbf{J}_{q1}^{H} + \mathbf{J}_{p2} \mathbf{C}_{2\{pq\}} \mathbf{J}_{q2}^{H} + \mathbf{N}_{pq}, \qquad (3)$$

in the un-clustered calibration strategy. Vectorizing \mathbf{V}_{pq} , the visibility vector of the baseline p - q is derived as

$$\mathbf{y} = \mathbf{J}_{q1}^* \otimes \mathbf{J}_{p1} \operatorname{vec}(\mathbf{C}_{1\{pq\}}) + \mathbf{J}_{q2}^* \otimes \mathbf{J}_{p2} \operatorname{vec}(\mathbf{C}_{2\{pq\}}) + \mathbf{n}_{pq},$$

where the noise vector $\mathbf{n}_{pq} \equiv \operatorname{vec}(\mathbf{N}_{pq})$ is considered to have a $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$ distribution. Therefore, we have

$$\mathbf{y} \sim \mathcal{CN}(\mathbf{s}, \sigma^2 \mathbf{I}), \quad \mathbf{s} \equiv \sum_{i=1,2} \mathbf{J}_{qi}^* \otimes \mathbf{J}_{pi} \operatorname{vec}(\mathbf{C}_{i\{pq\}}).$$
 (4)

Using Eq. (4), the log-likelihood function of the visibility vector **y** is given by

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{y}) = -4\ln\{\frac{\pi}{\sigma^2}\} - \sigma^{-2}(\mathbf{y} - \mathbf{s}(\boldsymbol{\theta}))^H(\mathbf{y} - \mathbf{s}(\boldsymbol{\theta})).$$
(5)

Consequently, the Fisher information matrix is obtained as

$$\mathcal{I}(\boldsymbol{\theta}) = -\operatorname{E}_{\mathbf{y}}\left[\frac{\partial^{2}\mathcal{L}(\boldsymbol{\theta}|\mathbf{y})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{T}}\right] = 2\sigma^{-2}\,\mathfrak{Re}(J_{\mathbf{s}}^{H}J_{\mathbf{s}}),\tag{6}$$

where J_{s} is the Jacobian matrix of s with respect to $\boldsymbol{\theta}$

$$J_{\mathbf{s}}(\boldsymbol{\theta}) = \sum_{i=1}^{2} \frac{\partial}{\partial \boldsymbol{\theta}} \{ \mathbf{J}_{qi}^{*} \otimes \mathbf{J}_{pi} \} [\mathbf{I} \otimes \operatorname{vec}(\mathbf{C}_{i\{pq\}})].$$
(7)

Thus, variations of any unbiased estimator of parameter vector $\boldsymbol{\theta}$, lets say $\hat{\boldsymbol{\theta}}$, is bounded from below by the CRLB as

$$\operatorname{Var}(\widehat{\boldsymbol{\theta}}) \ge [2\sigma^{-2} \operatorname{\mathfrak{Re}}(J_{\mathbf{s}}^{H}J_{\mathbf{s}})]^{-1}.$$
(8)

Lets try to bound the error variations of the clustered calibration parameters assuming that the two sources construct a single cluster, called cluster number 1. We reform Eq. (3) as

$$\mathbf{V}_{pq} = \widetilde{\mathbf{J}}_{p1} (\mathbf{C}_{1\{pq\}} + \mathbf{C}_{2\{pq\}}) \widetilde{\mathbf{J}}_{q1}^{H} + \Gamma_{1\{pq\}} + \Gamma_{2\{pq\}} + \mathbf{N}_{pq}, \quad (9)$$

where $\Gamma_{i\{pq\}}$, referred to as "clustering error" matrices, are given by

$$\boldsymbol{\Gamma}_{i\{pq\}} = \mathbf{J}_{pi} \mathbf{C}_{i\{pq\}} \mathbf{J}_{qi}^{H} - \widetilde{\mathbf{J}}_{p1} \mathbf{C}_{i\{pq\}} \widetilde{\mathbf{J}}_{q1}^{H},$$
(10)

and $\widetilde{\mathbf{J}}_{p1}(\hat{\boldsymbol{\theta}})$ is the clustered calibration solution at receiver *p*. Eq. (9) implies that what is considered as the noise matrix $\widetilde{\mathbf{N}}_{pq}$ in clustered calibration data model, Eq. (2), is in fact

$$\widetilde{\mathbf{N}}_{pq} \equiv \Gamma_{1\{pq\}} + \Gamma_{2\{pq\}} + \mathbf{N}_{pq}.$$
(11)

Vectorizing Eq. (9), clustered calibration visibility vector is obtained by

$$\mathbf{y} = \widetilde{\mathbf{J}}_{q}^{*} \otimes \widetilde{\mathbf{J}}_{p} \operatorname{vec}(\mathbf{C}_{1\{pq\}} + \mathbf{C}_{2\{pq\}}) + \widetilde{\mathbf{n}}_{pq}, \qquad (12)$$

where $\widetilde{\mathbf{n}}_{pq} = \operatorname{vec}(\mathbf{N}_{pq})$.

We point out that depending on the observation as well as the positions of the two sources on the sky, the clustering error $\Gamma_{i\{pq\}}$ will have different properties. However, in order to study the performance of clustering in a statistical sense, and to simplify our analysis, we make the following assumptions.

1. Consider statistical expectation over different observations and over different sky realizations where the sources are randomly distributed on the sky. In that case, almost surely $E{\{\tilde{J}\}} \rightarrow E{\{J\}}$ and consequently

$$\mathrm{E}\{\Gamma_{i\{pq\}}\} \to \mathbf{0} \tag{13}$$

In other words, we assume the clustering error to have zero mean over many observations of different parts of the sky.

2. We assume the error in clustering sources is less if the sources are closer together in the sky rather than far apart. Therefore, given a set of sources, the clustering error will reduce as the number of clusters increase. In fact, this error becomes **0** when we have clusters equal to the number of source (each cluster contains only one source). Therefore, given a set of sources, the variance of $\Gamma_{i\{pq\}}$ will decrease as the number of clusters increase.

Using Eq. (13) and bearing in mind that $E\{N_{pq}\} = 0$, we can consider $\tilde{n}_{pq} \sim C\mathcal{N}(0, \tilde{\sigma}^2 \mathbf{I})$ where $E\{\tilde{n}_{pq}\tilde{n}_{pq}^H\} = \tilde{\sigma}^2 \mathbf{I}$. Therefore,

$$\mathbf{y} \sim \mathcal{CN}(\widetilde{\mathbf{s}}, \widetilde{\sigma}^2 \mathbf{I}), \quad \widetilde{\mathbf{s}} \equiv \widetilde{\mathbf{J}}_q^* \otimes \widetilde{\mathbf{J}}_p \operatorname{vec}(\mathbf{C}_{1\{pq\}} + \mathbf{C}_{2\{pq\}}).$$

and similar to Eq. (8), we have

$$\operatorname{Var}(\widetilde{\widetilde{\boldsymbol{\theta}}}) \ge [2\widetilde{\sigma}^{-2} \,\mathfrak{Re}(J_{\widetilde{\mathbf{s}}}^H J_{\widetilde{\mathbf{s}}})]^{-1}.$$
(14)

3.1.1. Simulation 1: Two sources and one cluster

We simulate an observation of two point sources with intensities $\delta_1 = 11.25$ and $\delta_2 = 2.01$ Jansky (Jy) at sky coordinates [l, m] equal to [-0.014, -0.005] and [-0.011, -0.010], respectively. Baselines coordinates of Westerbork Radio Synthesis Telescope (WSRT) with 14 receivers are used in this simulation.

Fourier transform of the *i*-th source coherency matrix at baseline coordinates [u, v, w] is calculated as [1]

$$\mathbf{C}_{i} = \delta_{i} e^{\{j(ul_{i} + vm_{i} + w_{i}[\sqrt{1 - l^{2} - m^{2}} - 1])\}} \begin{bmatrix} .5 & 0\\ 0 & .5 \end{bmatrix}$$
(15)

where $j^2 = -1$. We Consider the J Jones matrices of Eq. (3) to be diagonal. Their amplitude and phase elements follow $\mathcal{U}(0.75, 0.95)$ and $\mathcal{U}(0.003, 0.004)$ distributions, respectively. The background



Fig. 1. Clustered and un-clustered calibrations CRLB. When the effective noise power of the clustered calibration, $||\tilde{\mathbf{N}}||^2$, is small enough, then its CRLB is lower than of the un-clustered calibration's and it reveals a superior performance.

noise is $\mathbf{N} \sim C\mathcal{N}(\mathbf{0}, 10\mathbf{I})$. Jones matrices of the clustered calibration, \mathbf{J}_{p1} for p = 1, 2, are obtained as $\mathbf{J}_{p1} + \mathcal{U}(0.02, 0.40)e^{j\mathcal{U}(0.5,2)}$. For twenty realizations of \mathbf{J} matrices, we calculated CRLB of the unclustered and clustered calibrations using Eq. (8) and Eq. (14), respectively. The results are presented by Fig. 1. As we can see in Fig. 1, for small enough errors matrices $\mathbf{\Gamma}$ of Eq. (10), the clustered calibration's CRLB stands below the un-clustered calibration's. Therefore, the clustered calibration's performance is superior compared to the un-clustered ones. Increasing the power of error matrices, or the power of effective noise $\mathbf{\tilde{N}}$, the result becomes the opposite.

3.2. Analysis of CRLB

We claim that if source 1 is considerably brighter than source 2, $||\mathbf{C}_{1\{pq\}}|| \gg ||\mathbf{C}_{2\{pq\}}||$, and if the weak source power is much lower than the noise level, $||\mathbf{C}_{2\{pq\}}|| \ll ||\mathbf{N}_{\{pq\}}||$, then clustered calibration's performance is superior than the un-clustered calibration. Note that the worst performance of both the calibrations is for the weakest source and we are more concerned to compare the CRLBs for this source.

The CRLBs obtained for un-clustered and clustered calibrations in Eq. (8) and Eq. (14), respectively, are both almost equal to SINR⁻¹. In the un-clustered calibration, the effective signal for the weakest source is $C_{2{pq}}$ where the noise is $N_{{pq}}$. Therefore,

$$SINR_2 = \frac{||\mathbf{C}_{2\{pq\}}||^2}{||\mathbf{N}_{\{pq\}}||^2}.$$
 (16)

But, in clustered calibration, the effective signal and noise are $\widetilde{\mathbf{C}}_{\{pq\}} \equiv \mathbf{C}_{1\{pq\}} + \mathbf{C}_{2\{pq\}}$ and $\widetilde{\mathbf{N}}_{\{pq\}}$, respectively. Thus,

$$\operatorname{SINR}_{c} = \frac{||\widetilde{\mathbf{C}}_{\{pq\}}||^{2}}{||\widetilde{\mathbf{N}}_{\{pq\}}||^{2}}.$$
(17)

Clustered calibration has a superior performance when

$$SINR_c \gg SINR_2.$$
 (18)

Consider the two possible extremes in a clustered calibration procedure:

First, the case of clustering many sources of a large field of view to a very small number of clusters. In this case, the angular diameter of a cluster could be so large that the characteristics of the sky are considerably changed for different source directions of the cluster. Subsequently, dedicating a single solution to all the sources of every cluster by clustered calibration introduces clustering error matrices Γ with large variance of Eq. (10). Having high interference power, the clustered calibration effective noise \widetilde{N} of Eq. (11) becomes very large. Therefore, clustered calibration SINR will be very low and it does not produce high quality results.

Second, the case of clustering sources of a small field of view to a very large number of clusters. In this case, the variance of Γ matrices are almost zero while the signal powers of sources clusters are almost as low as of the individual sources. Therefore, the clustered calibration SINR is almost equal to the un-clustered calibration one and the calibration performances are expected to be almost the same as well.

Thus, clustered calibration best efficiency is obtained at the smallest number of clusters for which Eq. (13) and Eq. (18) satisfy simultaneously. The SINR of Eq. (18) could be used as an efficient criterion for detecting the optimum number of clusters.

3.3. Generalizations to many sources and many clusters

Vectorizing the visibilities of Eq. (1) and staking them in vector y, the general data model of un-clustered calibration is

$$\mathbf{y} = \sum_{i=1}^{K} \mathbf{s}_i(\boldsymbol{\theta}) + \mathbf{n},\tag{19}$$

where $\mathbf{n}\sim\mathcal{CN}(\mathbf{0},\boldsymbol{\Pi})$ and

$$\mathbf{s}_{i} \equiv \begin{bmatrix} \mathbf{J}_{2i}^{*} \otimes \mathbf{J}_{1i} \operatorname{vec}(\mathbf{C}_{\{12\}i}) \\ \vdots \\ \mathbf{J}_{Ni}^{*} \otimes \mathbf{J}_{(N-1)i} \operatorname{vec}(\mathbf{C}_{\{(N-1)N\}i}) \end{bmatrix}.$$
(20)

Therefore, the un-clustered calibration CRLB is resulted as

$$\operatorname{var}(\boldsymbol{\theta}) \geq \left[2 \operatorname{\mathfrak{Re}}\left\{ (\sum_{i=1}^{K} J_{\mathbf{s}_{i}}(\boldsymbol{\theta}))^{H} \boldsymbol{\Pi}^{-1} (\sum_{i=1}^{K} J_{\mathbf{s}_{i}}(\boldsymbol{\theta})) \right\} \right]^{-1},$$

where $J_{\mathbf{s}_i}$ is the Jacobian matrix of \mathbf{s}_i with respect to $\boldsymbol{\theta}$.

Computing the exact CRLB is more complicated when we have clustered calibration. Based on Eq. (2), clustered calibration measurement equation is

$$\mathbf{y} = \sum_{i=1}^{Q} \widetilde{\mathbf{s}}_i(\widetilde{\boldsymbol{\theta}}) + \widetilde{\mathbf{n}}.$$
 (21)

In Eq. (21), $\tilde{\mathbf{s}}_i$ is defined similar to Eq. (20) where **J** and **C** are replaced by $\tilde{\mathbf{J}}$ and $\tilde{\mathbf{C}}$, $\tilde{\mathbf{n}} \equiv \sum_{i=1}^{K} \Gamma_i + \mathbf{n}$,

$$\Gamma_i = [\operatorname{vec}(\Gamma_{i\{12\}})^T \dots \operatorname{vec}(\Gamma_{i\{(N-1)N\}})^T]^T, \qquad (22)$$

and $\Gamma_{i\{pq\}}$ is given by Eq. (10).

Calculation of the conventional CRLB for the clustered calibration's data model of Eq. (21) is impractical due to the existence of the nuisance parameters Γ_i . This leads us to the use of Cramer-Rao like bounds devised in the presence of the nuisance parameters [10]. We apply the Modified CRLB (MCRLB) [11] to the clustered calibration's case.

The MCRLB for estimation errors of $\tilde{\theta}$ in the presence of the nuisance parameters Γ error matrices is defined as [11]

$$\operatorname{var}(\widehat{\widetilde{\boldsymbol{\theta}}}) \geq \left[\operatorname{E}_{\mathbf{y}, \mathbf{\Gamma}} \left\{ -\operatorname{E}_{\mathbf{y} \mid \mathbf{\Gamma}} \left\{ \frac{\partial}{\partial \widetilde{\boldsymbol{\theta}}} \frac{\partial}{\partial \widetilde{\boldsymbol{\theta}}^{T}} \ln\{P(\mathbf{y} \mid \mathbf{\Gamma}; \widetilde{\boldsymbol{\theta}})\} \right\} \right\}^{-1}, (23)$$

where $P(\mathbf{y}|\boldsymbol{\Gamma};\boldsymbol{\theta})$ is the PDF of the visibility vector \mathbf{y} assuming that the $\boldsymbol{\Gamma}$ matrices of Eq. (22) are priori known. Since $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Pi})$, from Eq. (21) we have

$$\mathbf{y}|\mathbf{\Gamma} \sim \mathcal{CN}(\sum_{i=1}^{Q} \tilde{\mathbf{s}}_i + \sum_{i=1}^{K} \mathbf{\Gamma}_i], \mathbf{\Pi}),$$
 (24)

and therefore in Eq. (23), $- \mathbb{E}_{\mathbf{y}|\mathbf{\Gamma}} \left[\frac{\partial}{\partial \tilde{\boldsymbol{\theta}}} \frac{\partial}{\partial \tilde{\boldsymbol{\theta}}^T} \ln\{P(\mathbf{y}|\mathbf{\Gamma}; \tilde{\boldsymbol{\theta}})\} \right]$, which is called the modified Fisher information, is equal to

$$2\mathfrak{Re}\{[\sum_{i=1}^{Q}J_{\tilde{\mathbf{s}}_{i}}(\tilde{\boldsymbol{\theta}})+\sum_{i=1}^{K}J_{\Gamma_{i}}(\tilde{\boldsymbol{\theta}})]^{H}\boldsymbol{\Pi}^{-1}[\sum_{i=1}^{Q}J_{\tilde{\mathbf{s}}_{i}}(\tilde{\boldsymbol{\theta}})+\sum_{i=1}^{K}J_{\Gamma_{i}}(\tilde{\boldsymbol{\theta}})]\}.$$

 $E_{\mathbf{y},\Gamma}$ in Eq. (23) could be estimated by Monte-Carlo method.

As a rule of thumb, skipping heavy computational cost of MCRLB, one can interpret the SINR test of Eq. (18) as follows: If in average the effective SINR of clustered calibration, $SINR_c$, gets higher than the effective SINR of un-clustered calibration obtained for the weakest observed signal, $SINR_w$,

$$E\{SINR_c\} \gg E\{SINR_w\},\tag{25}$$

then clustered calibration can achieve a better quality results. In Eq. (25), the expectation is taken with respect to the thermal noise \mathbf{N} , error matrices Γ , and all the baselines.

3.3.1. Simulation 2: SINR

We simulate WSRT including 14 receivers which observes 50 sources with intensities below 15 Jy. The source positions and their brightness are following uniform and Rayleigh distributions, respectively. The background noise is $\mathbf{N} \sim \mathcal{CN}(\mathbf{0}, 15\mathbf{I})$. We cluster sources, using weighted Hierarchical clustering [5], into Q number of clusters where $Q \in \{3, \ldots, 50\}$. Since for smaller number of clusters, we expect larger interference (errors) in clustered calibration's solutions, for every Q, we consider $\sum_{i=1}^{50} \mathbf{\Gamma}_i \sim \mathcal{CN}(\mathbf{0}, \frac{150}{Q}\mathbf{I})$. The choice of the complex Gaussian distribution for the errors matrices $\mathbf{\Gamma}$ is due to the central limit theorem and the assumptions made in section 3.1.

We proceed to calculate $E\{SINR_c\}$ by Monte-Carlo method. Signal power at each cluster is obtained using the cluster's brightest and weakest sources. As the result in Fig. 2 shows, $E\{SINR_c\}$ is low for very small Q, when the effect of interference is large. By increasing the number of clusters it increases and gets its highest peak for which the best performance of the clustered calibration is expected. After that, it decreases by the dominant effect of the background noise, and approachs to the un-clustered calibration $E\{SINR_w\}$.

4. CONCLUSIONS

Below the calibration background noise level, clustered calibration achieves a higher quality of solutions, compared to the un-clustered calibration. This is due to the fact that it promotes the level of SINR in the calibration procedure. We proved this superiority for a general sky model utilizing MCRLB and SINR analysis. Using SINR criterion, we could find the optimum number of clusters, for which the clustered calibration accomplishes its best performance. A future challenge shall be utilizing estimations of SINR or MCRLB in detection of the optimum number of clusters for real observations and for specific sky models.



Fig. 2. Clustered calibration's SINR. SINR is low for small Q, when the interference is large. By increasing the number of clusters it increases and gets its highest level for which the best performance of the calibration is expected. After that, it decreases by the dominant effect of the background noise.

5. REFERENCES

- A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., Interferometry and Synthesis in Radio Astronomy, John Wiley & Sons, May 2001.
- [2] T. J. Pearson and A. C. S. Readhead, "Image Formation by Self-Calibration in Radio Astronomy," ARA&A, vol. 22, pp. 97–130, September 1984.
- [3] G. Bernardi, D. A. Mitchell, S. M. Ord, L. J. Greenhill, B. Pindor, R. B. Wayth, and J. S. B. Wyithe, "Subtraction of point sources from interferometric radio images through an algebraic forward modeling scheme," *ArXiv e-prints*, December 2010.
- [4] B. Pindor, J. S. B. Wyithe, D. A. Mitchell, S. M. Ord, R. B. Wayth, and L. J. Greenhill, "Subtraction of Bright Point Sources from Synthesis Images of the Epoch of Reionization," *PASA*, vol. 28, pp. 46–57, July 2010.
- [5] S. Kazemi, S. Yatawatta, and S. Zaroubi, "Clustered Radio Interferometric Calibration," *IEEE Workshop on Statistical Signal Processing*, SSP2011, vol. 414, pp. 597 – 600, June 2011.
- [6] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation theory, Prentice Hall signal processing series. Prentice-Hall PTR, 1993.
- [7] J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding radio polarimetry. I. Mathematical foundations.," A&AS, vol. 117, pp. 137–147, May 1996.
- [8] J. P. Hamaker, "Understanding radio polarimetry. V. Making matrix self-calibration work: processing of a simulated observation," A&A, vol. 456, pp. 395–404, September 2006.
- [9] M. Born and E. Wolf, *Principles of Optics*, Cambridge University Press, October 1999.
- [10] F. Gini and R. Reggiannini, "On the use of cramer-rao-like bounds in the presence of random nuisance parameters," *IEEE Transactions on Communications*, vol. 48, pp. 2120–2126, December 2000.
- [11] F. Gini, R. Reggiannini, and U. Mengali, "The modified cramer-rao bound in vector parameter estimation," *IEEE*, vol. 46, pp. 52–60, January 1998.