Python implementation and validation of an Optimization Method for Interferometric Antenna Arrays

Luca Cabral
Argentine Institute
of Radio Astronomy (IAR)
UNLP - CONICET - CICPBA
Berazategui, Buenos Aires
Instituto de Investigaciones
en Electrónica, Control
y Procesamiento de Señales
(LEICI)
Facultad de Ingeniería
UNLP - CONICET
La Plata, Buenos Aires
Email: lucacabral41@gmail.com

Mariano Fernández Corazza
Instituto de Investigaciones
en Electrónica, Control
y Procesamiento de Señales
(LEICI)
Facultad de Ingeniería
UNLP - CONICET
La Plata, Buenos Aires

Guillermo Gancio
Argentine Institute
of Radio Astronomy (IAR)
UNLP - CONICET - CICPBA
Berazategui, Buenos Aires

Paula Benaglia
Argentine Institute
of Radio Astronomy (IAR)
UNLP - CONICET - CICPBA
Berazategui, Buenos Aires

***Abstract*—A main task to solve while designing a radio interferometer is the location of the antenna-elements, a problem that defines the interferometer response. The solution determines which points of the Fourier plane (or uv plane) will be sampled, together with their density. These characteristics are extremely important because the antenna locations, in turn, define the synthesized beam or point spread function (PSF) of the whole instrument. An inadequate array configuration implies an ill-constructed synthesized beam. Consequences of inadequate arrays are for example the need to delete measures over a range of spatial frequencies due to low signal-to-noise ratio (SNR), or, at the data reduction stage, to degrade spatial resolution in order to get a cleaner image. In this work, we implement a method to optimize the antenna locations, starting with an initial random configuration and a desired or objective sample density function, taking into account if there are terrain constraints.**

***Index Terms*—interferometric instrumentation, optimization, numerical methods.**

I. Introduction

A radio interferometer is an array of antennas which, via correlation between pairs of them, sample components of the Fourier transform plane (or uv plane) of an astronomical source. The more covered the uv plane is, the better the image quality will be. The aim is then, to cover as best, according to the circumstances, the uv plane, in order to get as much information as possible to build the image.

On designing an interferometer array there are mainly three problems to face, namely: determining the antenna locations, building an appropriate correlator, and synchronizating the data acquisition and data transfer systems. This work focuses on the antenna location problem. One approach to solve this problem is using optimization methods.

Regarding the optimization methods, as radio inteferometers were built around the world, different methods were also developed in order to optimize different characteristics of the instrument. One case of optimization is for linear array, radio interferometers where the antennas are located along a straight line. In this case, there is a configuration where the resolution is maximum and the number of times that a sample of the measured astronomical source is minimum, named as minimum-redundancy linear arrays [1]. Due to Earth rotation and changes in the interferometer response made by changes on the declination celestial coordinate, this strategy cannot be applied to bi-dimensional arrays. Keto [2] developed a method with neural networks, where the antennas are moved in order to minimize the distance between the points on the uv plane of the desired distribution and the one generated by the antennas. This method was developed for short time observations (snapshots), but it can be extended to optimize long track observations. Kogan [3] used a different approach, where the optimization minimizes the sidelobes of the PSF, which is an advantage in the steps on imaging but does not bring any information about how the points are sampled on the uv plane.

The Argentine Institute of Radio Astronomy (IAR) is working with a project to install the first radio interferometer in Argentina, which will be called the Multipurpose Interferometer Array (MIA). This interferometer will be placed in the Andes foothills region and will consist of 64 antennas. Given this particular problem with a certain number of antennas and geographic location, one aim is to find the optimal distribution for the antenna locations. The objective of this work is to reimplement an existing optimization method for

ISBN: 978-950-766-230-0

the later use on MIA [4], understanding in detail the algorithm, implementing it in Python, testing it and, if possible, even improve the algorithm. The decision to do the implementation in Python relies on the coding simplicity, and on being an open source environment that has packages like NumPy with high performance implementations for managing arrays and matrices.

The results of this first implementation were validated by optimizing the arrays with AntConfig [5], a software that uses the same base method that the one used in this paper but applying tomographic projection to reduce the problem dimension into a one-dimensional one. AntConfig was chosen to be the comparison software because it has the same base method, and also was the software used to optimize the distribution of interferometric arrays like ASKAP [6] and KAT-7 [7]. The decision to do our new implementation is to solve the optimization problem without reducing the dimension and also, for later improvement, to approach the problem of long observations in a way more related to the operation mode of a radio interferometer.

This paper is divided in different sections: section II explains how the samples are distributed over the uv plane from the place where the array is located, the antenna relative locations, and the desired astronomical source. In a subsection we explain the relationship between the uv density distribution and the PSF, outlining the array characteristics on this distribution. In section III, we introduce the gradient's optimization method [8]. The application of this method is described in section IV for a simplified case of MIA. Finally, conclusions and future work are outlined.

II. The Antenna Location Problem

To determine the optimal location of the antennas, first the desired characteristics of the radio interferometer have to be detailed, such as the angular resolution or the type of astronomical sources that the array will observe. Depending on the approach, these characteristics can be described in terms of the ground plane (North-East position of the antennas domain), or the uv plane (the transformed domain).

A. uv sample distribution

A simple radio interferometer consists of a pair of antenna-elements that make a baseline. This baseline, projected to the direction of the desired astronomical source, samples a pair of points on the uv plane. These points are symmetric respect to the plane's origin because two antenna-elements make two baselines when changing the reference antenna. Due to the Earth rotation, this projected baseline changes with time, and the sampled points describe two ellipsoidal branches in the uv plane. These long time observations are named as Earth synthesis rotation.

Eq. 1 shows the expression that relates the sampled points of the uv plane with the baselines [9]:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \sin(H) & \cos(H) & 0 \\ -sin(\delta)\cos(H) & \sin(\delta)sin(H) & \cos(\delta) \end{pmatrix} \mathbf{D}_\lambda, \quad (1)$$

where H is the hour angle, δ is the declination, and $\mathbf{D}_\lambda$ is the baseline vector, normalized to the observation wavelength λ, and expressed in an Earth-based Cartesian system (Fig 1), so the latitude where the array is located is involved.

It is important to highlight how an observation becomes more difficult to make when the declination module increases, because the v component becomes foreshortened.

Summarizing, the interferometer response is not only a function of the array geometry, but also of the source location and the observation time length. Therefore, in the antenna location problem, it is important to define the measurement conditions, the declination observation range, and the hour angle range.

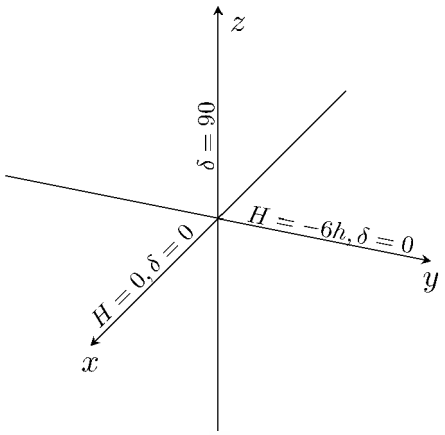

Fig. 1. Earth-based Cartesian system. In this system the $z = 0$ plane corresponds to the Earth's equator, the $y = 0$ plane is the meridian plane, the y axis points to the east and the z axis points to the North Pole. Because of the system definition, the geographical latitude ϕ measure matches the declination.

B. uv sample density

Due the amount of antenna-elements of an array, or the measurement time, there may be points of the uv plane which are sampled more than one time. Because of this, it can be defined a sample density function $D(u, v)$ whose value at a given (u, v) point is the amount of times which that point is sampled by all the baselines on the observation. Then, the spatial transfer function is defined in a similar way as $D(u, v)$, but instead of the times that a point is sampled, its value is 1 at each sampled point; and the Fourier transform of the spatial trasnfer function is the PSF [9].

As the array perfomance may be analized from the ground plane, the same can be done with the density function. One way to see this is thinking the (u, v) points of the plane as a linear transformation of the baselines. The angular resolution is given by the points of the uv plane that are farthest from the origin, that correspond to the longest baselines. Sub-arrays with lower angular resolution (in order to observe extended sources, i.e. with larger angular size, or sources with extended emission) correspond to regions of the uv plane near the origin.

There is an important relationship between the shape of the PSF and the shape of the density function. One example

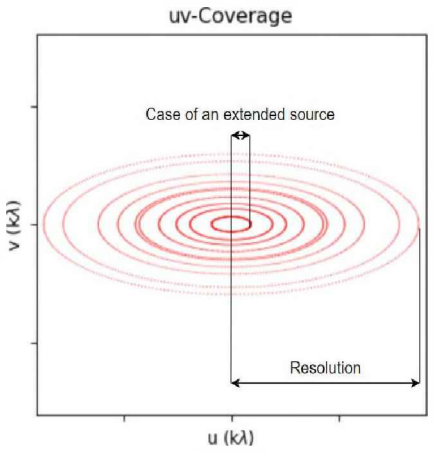

Fig. 2. Example of how some parameters of the instrument can be interpreted by means of the uv coverage.

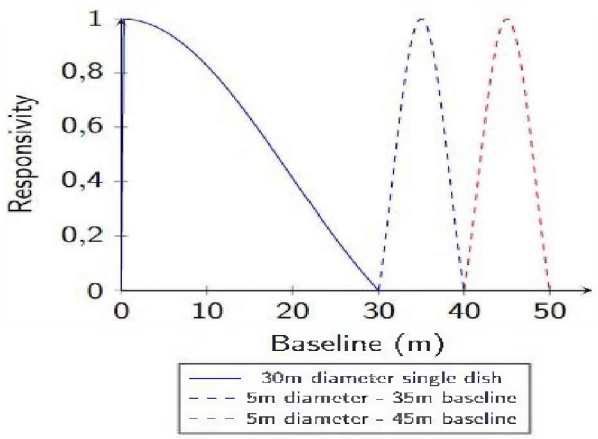

Fig. 3. Example of the relation between single dish and interferometric observations with MIA parameters in terms of responsivity. The antennas from MIA have a 5m diameter dish and a single dish with a diameter of 30m.

of this is an exponential sample density function, which corresponds to an exponential PSF shape. The exponential is a remarkable shape because it has no sidelobes but, for a given angular resolution, it requires longer baselines than other shapes. Another important characteristic is the signal-to-noise ratio (SNR) of the sampled points, which for the exponential case is highest at the origin and lowest at the farthest regions. The second example is the uniform coverage or the disk density function, with constant SNR everywhere and with a PSF that corresponds to a jinc function [9]. The jinc function is like the sinc function but instead of a sine in the numerator, it is the zero species and first order Bessel function. This transformation can be achieved via the Hankel transform, because of the radial symmetry [10]. Intermediate densities between these examples, starting with the uniform case, lead to PSFs with lower sidelobes, lower SNR in the outer regions and higher angular resolutions.

The maximum baseline defines the angular resolution of a radio interferometer, and the minimum baseline defines the range of the extended emission that the instrument can measure (Fig. 2). There is also a minimum distance between antennas to avoid problems like shadowing, which also depend on the declination at a given observation. One possibility is complementing an observation with single dish observations [11] [12]. In that case, a single dish antenna can sample points from the origin to the diameter of the dish (Fig 3), covering the short baselines that can not be covered by the interferometer. A circular mask can be added to the model density to impose the minimum baseline condition and/or single dish observations can be used as a complement.

In conclusion, at the stage of defining the desired characteristics of the radio interferometer, these can be defined both on the ground plane domain and on the uv plane. Related to the antenna location problem, the first approach represents the direct method, and the second represents the indirect method. In this work we focus on the indirect method.

III. Gradient based optimization

Given an unconstrained function $f(\bar{\mathbf{x}})$, where $f(\bar{\mathbf{x}}) : \mathbb{R}^n \rightarrow \mathbb{R}$, a way to minimize it is changing $\bar{\mathbf{x}}$ in the direction $-\nabla f(\bar{\mathbf{x}})$ until reaching the minimum, known as the gradient descent method. If the function is convex and differentiable, there will be a minimum when $\nabla f(\bar{\mathbf{x}}) = 0$, and $\bar{\mathbf{x}}$ will be the optimal point because there is only one global minimum [8]. This optimization method consists of iterations over the calculation of Eq. 2

$$\bar{\mathbf{x}}_{k+1} = \bar{\mathbf{x}}_k - g\nabla f(\bar{\mathbf{x}}_k), \tag{2}$$

until the gradient reaches either zero or some stop condition, where it is considered that the minimum was found. The constant g, with $g \in \mathbb{R}$, is a step gain which can be defined by the user or calculated as the optimal step towards the minimum. When giving a value to g, it has to be observed that for high values, $f(\bar{\mathbf{x}})$ may oscillate around the problem solution. But for low values of g, it will require more iterations for the optimization to reach the optimal $\bar{\mathbf{x}}$.

For the antenna location problem, this method was previously implemented [4], defining $f(\mathbf{x})$, as

$$f(\bar{x}) = G(u,v) = D(u,v) - D_m(u,v), \tag{3}$$

where $D(u,v)$ is the sample density and $D_m(u,v)$ is the desired or model density. For a given array with n antenna-elements, there are $n-1$ samples related to one antenna in an instant of the observation, situation that has to be considered when displacing the antennas. Therefore, the optimization is rebuilt as

$$\mathbf{x}_{j,k+1} = \mathbf{x}_{j,k} - g \sum_{i=1, i\neq j}^{n} M\nabla G(u_{ij}, v_{ij}), \tag{4}$$

where (u_{ij}, v_{ij}) is the point of the uv plane which correspond to the baseline made by the antenna i and the antenna j. M

ISBN: 978-950-766-230-0

is the transformation matrix from the uv plane to the ground plane, defined as [9]:

$$M = \begin{pmatrix} \frac{\sin(\delta)\sin(\phi)cos(H) + cos(\delta)\cos(\phi)}{\cos(\delta-\phi)} & \frac{\sin(H)\sin(\phi)}{\cos(\delta-\phi)} \\ \frac{\sin(H)\sin(\delta)}{\cos(\delta-\phi)} & \frac{\cos(H)}{\cos(\delta-\phi)} \end{pmatrix}, \tag{5}$$

where ϕ is the latitude. M is a constant for snapshot cases and a function of H for long track observations. The sum in Eq. 4 considers the direction of the minimum for all the baselines for each antenna. Then, in each iteration, the antennas are displaced in the average direction of the minimum of $G(u, v)$. An intuitive analysis of the convergence of the method and avoidance of local minima is found in [4].

Another important point is the grid used in the numerical methods. If the grid is too small for a particular case, the sample density accumulates a lower number of (u, v) points, so there will be densities with null value and then, the numerical derivates will be also null. Therefore the antennas will not move and the method will not converge to the optimal distribution. A larger size of the grid will delete the smoothness of the model density function, and therefore the final distribution will produce a different density to the expected one. For example, when using a coarser grid for a uniform density coverage, it can be seen that the edges of the disk turn into straight lines, so the final distribution will describe straight lines in the circle instead the expected curve for this model. In conclusion, the grid also has to be taken into account, because it affects the convergence of the method. This trade-off of the grid size is an example of an implementation detail addressed in this work that was not described in the literature.

IV. Implementation

Our implementation was coded on a Jupyter Notebook, where we included the feature of adding terrain constraints. At each iteration, the algorithm checks if any antenna is inside a forbidden region and if that happens, that antenna is pulled out to the nearest border.

After the implementation, we simulated a simple version of the MIA case, with these characteristics:

- Geographical latitude of -34.5 deg (it is the value where the array will be placed).
- Equatorial declination of -34.5 deg.
- Angular resolution of 1 arcsec with a center wavelength of 21 cm.
- Snapshot observation.
- Uniform disk uv plane coverage.

MIA, as the name says, has to be a multipurpose interferometer, so it must be optimal to carry out observations in both long tracks in the case of Earth synthesis rotation, and in short observations like in the case of snapshots. For these first optimization cases, only simulations for snapshots were carried out. Three model cases were simulated for testing: the uniform coverage, a gaussian coverage, and a uniform coverage with terrain constraints. Starting with the first case, in order to include the angular resolution θ_{res} in the model, it was used the relationship between this parameter and the longest baseline B_{max} given by

$$\theta_{res} = \frac{\lambda}{B_{max}}. \tag{6}$$

Then, it can be found the radius of the disk using the transformation between the baseline vector to the uv plane as:

$$u^2 + v^2 = \left(\frac{B_{max}}{\lambda}\right)^2, \tag{7}$$

which is around $262 \times 10^6 \lambda$.

Fig. 4 depicts the initial distribution of the antenna array with blue dots (each dot represents an antenna), which is given by a random distribution, and the optimized distribution with red dots. The light blue circumference is given by the longest baseline, calculated from the farthest (u, v) point to the origin. It can be seen that most of the antennas lie around this circumference, which diameter corresponds to the angular resolution of the instrument (around 55.54 km); while the rest of the antennas are inside the circle to cover the intermediate baselines. These results have been previously observed [9] [4] [2], so this simulation was also a way to check if our implementation was working properly. In the first row of Fig. 5 we show the final density distribution and the horizontal PSF for the uniform case, where only one profile is plotted because of the radial symmetry. Regarding the density, it can be seen that it has uniform coverage except at the borders of the circle, where the density is a little bit lower than at the origin of the plane. About the PSF, it shows some weak side lobes, a characteristic of this distribution as it was explained in section II. Additionally, the achieved angular resolution was 0.8 arcsec, lower than the desired characteristic for the interferometer, because the largest (u, v) sample in the calculation with Eq. 7 was rounded to the next integer value.

In the second simulation, we used a Gaussian density model with zero mean and variance σ given by

$$\sigma = \frac{2\sqrt{2\ln(2)}\frac{180 * 60 * 60}{\pi}\frac{\lambda}{2\pi}}{FWHM}, \tag{8}$$

expression used in the simulator AntConfig, software developed by De Villiers [5]. The optimized distribution for this distribution can be seen in the second row of Fig. 5. Note that the antenna distribution also has an exponential baseline distribution on the ground plane. This is because the density is larger near the uv plane origin (shorter baselines). Then, the density decreases when moving away from the origin, so the antenna distribution must have less density of longer baselines. In the second row of Fig. 5 it can be seen that the final distribution approaches the model density. Regarding the PSF, the 1 arcsec of angular resolution is achieved, at the expenses of adding longer baselines than in the first case of an uniform density coverage, being the longest baseline of 65.17 km. As an advantage of the Gaussian distribution, the PSF has

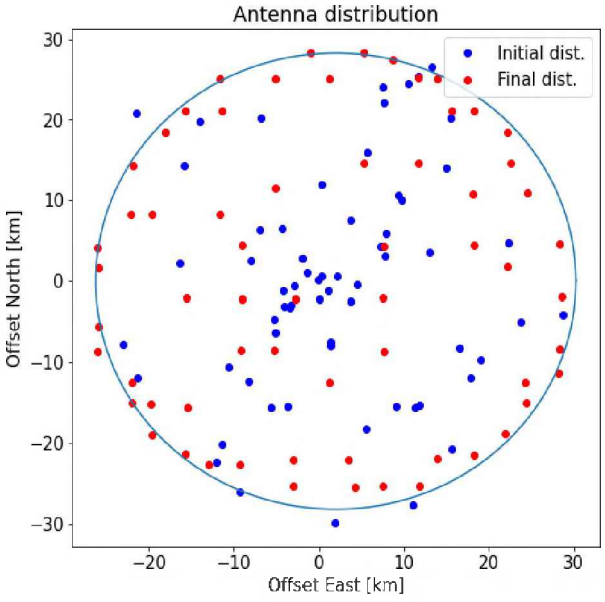

Fig. 4. Initial and final antenna distributions for a uniform coverage density model.

no sidelobes, which impacts later in the processing time of the images.

In the last simulation, the same model $D_m(u, v)$ of the first simulation was used, with the addition of two rectangular areas as terrain constraints. In Fig. 7 the final antenna distribution is depicted together with the rectangles. At the borders of these rectangles it can be seen a few antennas which felt inside, and were pulled out to the nearest border. This case shows that the algorithm to contemplate terrain constraints works properly.

With these three previous simulations we checked that our implementation of the gradient-based method works correctly. It was tested for different models and in one case with terrain constraints. All the optimizations were carried out with 500 iterations and a fixed gain of $g = 0.1$. In this first implementation of the method, we used a uniform rectangular grid with size $10\lambda \times 10\lambda$. The computation time was between 1 min and 1 min 30 s for all the cases. Therefore, with this first grid and a large amount of antennas like the MIA case with 64 antennas, the implementation requires a low time to optimize the distribution. It has to be outlined that this kind of grid is not always the optimal one, and because the sample density usually has radial symmetry, polar grids will be a future improvement for the method, because it will not delete the smoothness and may be increase the grid size range for the problem.

In the previous simulation, M was a constant because the optimizations were planned for snapshot observations. In case of designing an optimization for long track observations, M becomes a function of the hour angle H. The calculation of $\bar{\mathbf{x}}_{k+1}$ will require more time, because there will be one gradient to compute and one $n-1$ sum for each value of H.

The optimization of the uniform distribution case was made also in AntConfig, where the final distribution is depicted in Fig. 8. It can be seen that the AntConfig distribution fits a little bit more with the circumference than our implementation. The resulting optimization in both cases were uniform but, making the comparison with the standard deviation, of the difference between the model density and the density of the optimized array for both cases, the result was 0.42 for the AntConfig case and 0.38 for our implementation. This result leads to conclude that our implementation has a similar performance than AntConfig.

Summarizing, our implementation of the method can handle not only cases like the previous ones, but this can be improved in order to handle situations with more requirements.

V. Conclusion and Future Work

We successfully implemented the gradient descent optimization method in Python, applied it to optimize the antenna distribution of a radio interferometric array for three different distributions, and carried out a simulation for a real case such as the MIA. Our future work will include different features to the implementation such as:

- Define the gain g to be the optimal step gain in the calculation of $\bar{\mathbf{x}}_{k+1}$ [8].
- Include the polar grid type which size may increase depending on the uv plane region under consideration, as suggested in [4].
- In the calculation of $\bar{\mathbf{x}}_{k+1}$, redefine the multiplication $MG(u, v)$ in order to account M as a function of the angle hour H, and then include optimizations for longer time observations.
- Including a decomposition of the model density $D_m(u, v)$ in a sum of many model densities, in order to include sub-array optimizations.
- Another point to treat is to include optimizations, either for snapshots or long tracks, for different declinations, because it is desired that the final antenna distribution performs well for a declination range, and not only for one specific declination [13].

References

[1] Moffet, A. *Minimum-redundancy linear arrays*. IEEE Transactions on antennas and propagation, 16, 2, 172, 1968

[2] Keto, E. *The shapes of cross-correlation interferometers*. The Astrophysical Journal, 475, 2, 843, 1997

[3] Kogan, L. *Optimization of an array configuration minimizing side lobes*. MMA Memo, 171, 1997

[4] Boone, F. *Interferometric array design: Optimizing the locations of the antenna pads*. Astronomy & Astrophysics, 377, 1, 368, 2001

[5] De Villiers, M. *Interferometric array layout design by tomographic projection*. Astronomy & Astrophysics, 469, 2, 793, 2007

[6] Gupta, N., Johnston, S. & Feain, I. *ASKAP Array Configurations: Technical Studies*. ATNF SKA memo series, 16, 2008

[7] Foley, A., Alberts, T., Armstrong, R., Barta, A. et al. *Engineering and science highlights of the KAT-7 radio telescope*. Monthly Notices of the Royal Astronomical Society, 460, 2, 1664, 2016

[8] Boyd, S. P. & Vandenberghe, L. *Convex optimization*. Cambridge university press, 2004

[9] Thompson, A. R., Moran, J. M. & Swenson, G. W. *Interferometry and synthesis in radio astronomy*. Springer Nature, 2017

[10] Bracewell, R. N. & Bracewell, R. N. *The Fourier transform and its applications*, tomo 31999. McGraw-Hill New York, 1986

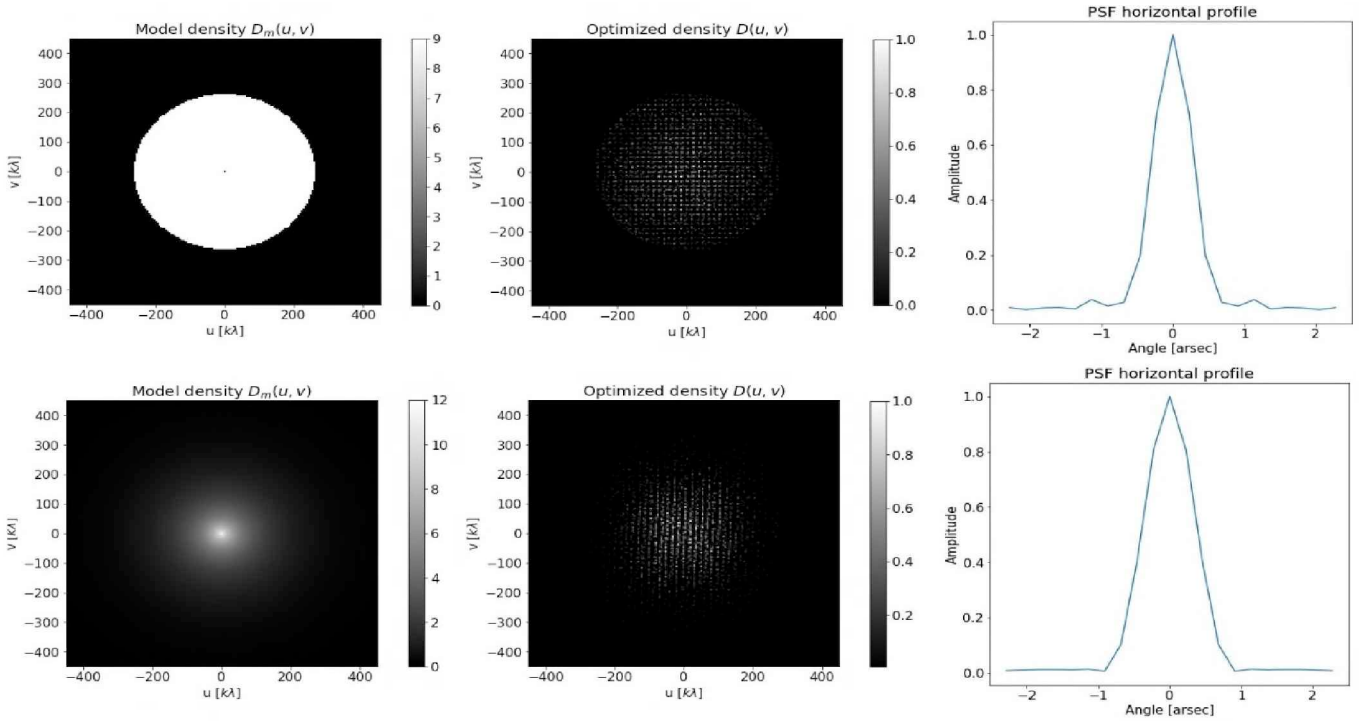

Fig. 5. Plots of the simulation results with the models for each case. The first row corresponds to uniform coverage while the second row corresponds to the gaussian samples density. In the first column are the density models, in the second column the final densities of the optimized distributions, and in the third column the horizontal PSF profiles.

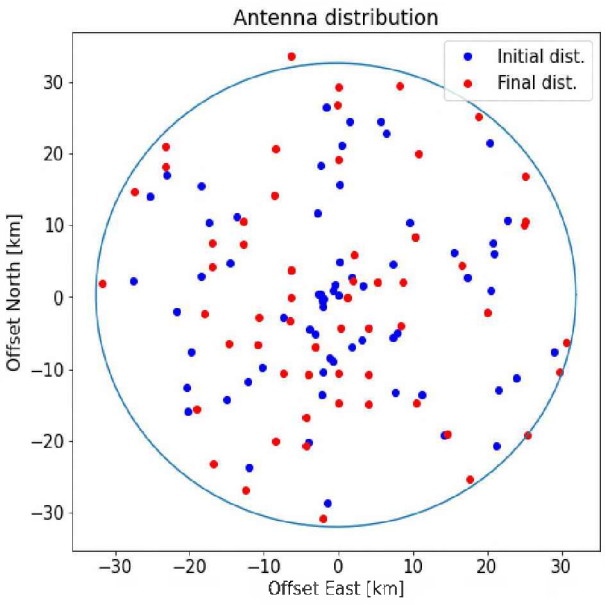

Fig. 6. Initial and final antenna distributions for a Gaussian coverage density model.

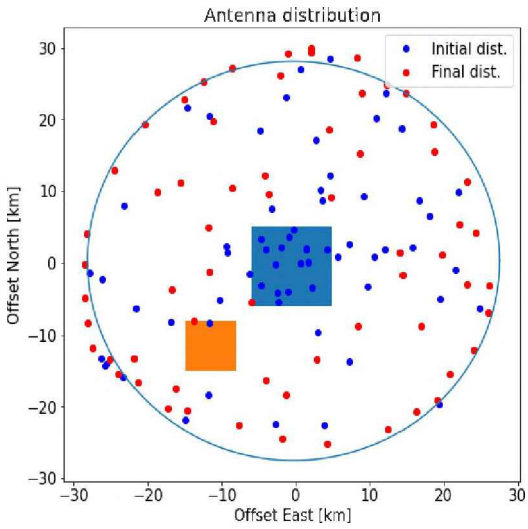

Fig. 7. Final antenna distribution for a uniform coverage density model and terrain constraints.

MMA Memo201, 1998

[11] Bajaja, E. & Van Albada, G. *Complementing aperture synthesis radio data by short spacing components from single dish observations*. Astronomy and Astrophysics, 75, 251, 1979

[12] Emerson, D. *Why single-dish?* En *Single-Dish Radio Astronomy: Techniques and Applications*, tomo 278, 27–43. 2002

[13] Holdaway, M. *Hour Angle Ranges for Configuration Optimization.*

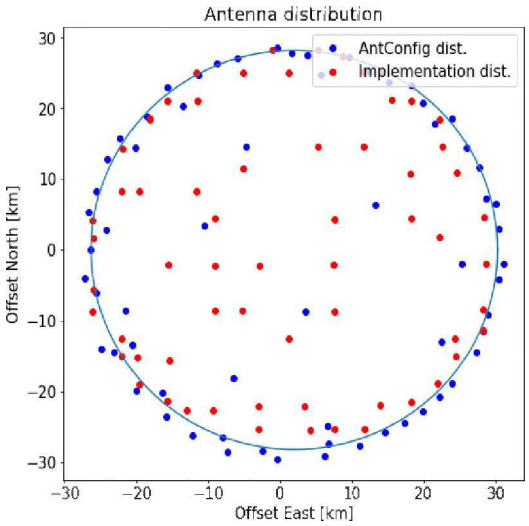

Fig. 8. Optimized antenna distributions for a uv plane uniform density coverage, made with AntConfig and with our implementation.

ISBN: 978-950-766-230-0