

Bayesian Reconstruction through Adaptive Image Notion [†]

Fabrizia Guglielmetti ^{1,*}, Eric Villard ² and Ed Fomalont ^{2,3}

¹ European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany

² Joint ALMA Observatory/ESO, Avenida Alonso de Córdova 3107, Vitacura, Santiago 763-0355, Chile; Eric.Villard@alma.cl (E.V.); efomalon@nrao.edu (E.F.)

³ National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA

* Correspondence: fgugliel@eso.org

† Presented at the 39th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Garching, Germany, 30 June–5 July 2019.

Published: 5 December 2019



Abstract: A stable and unique solution to the ill-posed inverse problem in radio synthesis image analysis is sought employing Bayesian probability theory combined with a probabilistic two-component mixture model. The solution of the ill-posed inverse problem is given by inferring the values of model parameters defined to describe completely the physical system arising from the data. The analysed data are calibrated visibilities, Fourier transformed from the (u, v) to image planes. Adaptive splines are explored to model the cumbersome background model corrupted by the largely varying dirty beam in the image plane. The de-convolution process of the dirty image from the dirty beam is tackled in probability space. Probability maps in source detection at several resolution values quantify the acquired knowledge on the celestial source distribution from a given state of information. The information available are data constraints, prior knowledge and uncertain information. The novel algorithm has the aim to provide an alternative imaging task for the use of the Atacama Large Millimeter/Submillimeter Array (ALMA) in support of the widely used Common Astronomy Software Applications (CASA) enhancing the capabilities in source detection.

Keywords: methods: data analysis; methods: statistical; techniques: image processing

1. Introduction

A software package investigates a method of Bayesian Reconstruction through Adaptive Image Notion (BRAIN) [1] to support and enhance the interferometric (IF) imaging procedure. Extending the works of [2,3], the software kit makes use of Gaussian statistics for a joint source detection and background estimation through a probabilistic mixture model technique. The background is defined as the small scale perturbation of the wanted signal, i.e., the celestial sources. As shown in [3], statistics is rigorously applied throughout the algorithm, so pixels with low intensity can be handled optimally and accurately, without binning and loss of resolution. Following the work of [2], a 2-D adaptive kernel deconvolution method is employed to prevent spurious signal arising from the dirty beam. Continuum, emission and absorption lines detection is foreseen to occur without an explicit subtraction, but propagating the information acquired on the continuum for line detection.

The novel algorithm aims at improving the CASA [4,5] imaging and deconvolution methods currently used, especially for the use of ALMA, but also applicable to next generation instruments as ngVLA [6] and SKA [7]. The developing software is going to be as compatible with CASA as possible by using CASA tasks, data formats, including python scripts. BRAIN has the aim to provide

an advancement in current issues as stopping thresholding, continuum subtraction, proper detection of extended emission, separation of point-like sources from diffuse emissions, weak signal, mosaics.

In the following sections, a brief description of the ALMA observatory, the acquired data and radio synthesis imaging is provided. The most used software package for ALMA data analysis, i.e., CASA, is introduced and the complex data transformation for the formation of interferometric images is shown. Last, the technique under development is outlined in the main features as differing from [3].

2. ALMA

Located at 5000 m altitude on the Chajnantor plateau (Chile), ALMA is an aperture synthesis telescope operating over a broad range of observing frequencies in the mm and submm regime of the electromagnetic spectrum, covering most of the wavelength range from 3.6 to 0.32 mm (84–950 GHz).

ALMA is characterized by 66 high-precision antennas and a flexible design: Several pairs of antennas (baselines) build a single filled-aperture telescope whose spatial resolution depends on the used configuration and observing frequency. Two main antennas arrangements are identifiable: the 12-m Array and the Atacama Compact Array (ACA). Firstly, with fifty relocatable antennas on 192 stations, the 12-m Array allows for baselines ranging from 15 m to 16 km. Secondly, ACA is composed by twelve closely spaced 7-m antennas (7-m Array) and four 12-m antennas for single dish observations, named Total Power (TP) Array. The 7-m and the TP arrays sample baselines in the (9–30) m and (0–12) m ranges, respectively. The data re-combination from the 7-m and TP arrays overcome the well-known “zero spacing” problem [8] being crucial during the observation of extended sources. The final resolution reached by these arrays is given by the ratio between the observational wavelength and the maximum baseline for a given configuration. Therefore, the 12-m Array is used for sensitive and high resolution imaging, while ACA is preferable for wide-field imaging of extended structures. The 12-m Array and ACA can reach a fine resolution up to 20 mas at 230 GHz with the most extended configuration and 1.44 arcsec at 870 GHz, respectively: These numbers refer to the point spread function (PSF) or Full Width at Half Maximum (FWHM) of the synthesized beam, which is the inverse Fourier transform of a (weighted) (u, v) sampling distribution.

ALMA delivers continuum images and spectral line cubes with frequency characterizing the third axis. The data cubes are provided with up to 7680 frequency channels, with the channel width indicating the spectral resolution (from 3.8 kHz to 15.6 MHz).

2.1. ALMA Images

The sky brightness distribution, collected by each antenna, is correlated for given baselines. The measurements obtained from the correlators in an array allow one to reconstruct the complex visibility function of the celestial source. The Van Cittert–Zernike theorem [9] provides the Fourier transform relation between the sky brightness distribution and the array response (or primary beam) [10].

In simplified form, the complex visibility can be stated as follows:

$$\mathcal{V}(u, v) = \iint A(l, m) \cdot I(l, m) e^{2\pi i(ul+vm)} dl dm \approx A e^{i\phi}. \quad (1)$$

In aperture synthesis analysis, the goal is to solve Equation (1) for $I(l, m)$ in the image plane by measuring the complex-valued and calibrated $\mathcal{V}(u, v)$. The coordinates u, v , measured in units of wavelength, characterize the spatial frequency domain. The (u, v) plane is effectively composed by the projections of baselines onto a plane perpendicular to the source direction. The calibrated $\mathcal{V}(u, v)$ and $I(l, m)$ are given in units of flux density ($1 \text{ Jy} = 10^{-23} \text{ ergs cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1}$) and of surface brightness (Jy/beam area), respectively. The antenna reception pattern $A(l, m)$ is relevant for the primary beam correction. The amplitude (A) and the phase (ϕ) inform us about source brightness and position relative to the phase center at spatial frequencies u and v .

2.2. The Data

A direct application of the inverse Fourier transform of $\mathcal{V}(u, v)$ (Equation (1)) is desirable but suitable only in case of a complete sampling of the spatial frequency domain $B(u, v)$. For a given configuration, the array provides a discrete number of baselines, being the number of baselines provided by $N \cdot (N - 1)/2$, where N is the number of antennas. The sampling of the (u, v) plane is consequently limited. Hence, the measured complex visibility function is corrupted by the sampling function: $B(u, v) \cdot \mathcal{V}(u, v)$.

The sampling function $B(u, v)$, composed by sinusoids of various amplitude and phase, is characterized as an ensemble of delta functions accounting for Hermitian symmetry. The inverse Fourier transform of this ensemble of visibilities provides the dirty image $I^D(l, m)$:

$$I^D(l, m) = FT^{-1}\{B(u, v)\mathcal{V}(u, v)\}. \quad (2)$$

Following the convolution theorem, Equation (2) can be written as follows:

$$I^D(l, m) = b(l, m) \times I(l, m)A(l, m). \quad (3)$$

The dirty image is the convolution of the sky brightness modified by the antenna primary beam $A(l, m)$ with the dirty beam, given by $b(l, m) = FT^{-1}\{B(u, v)\}$ (FWHM). The true sky brightness reconstruction $I(l, m)$ depends on the image fidelity, since the coverage of the (u, v) plane is by definition incomplete. For this reason the reconstruction of $I(l, m)$ is challenging due to the artefacts introduced by the dirty beam $b(l, m)$, the sparsity in the data and the big data volume.

3. Common Astronomy Software Applications

CASA is a suite of tools allowing for calibration and image analyses for radio astronomy data [11]. It is especially designed for the investigation of data observed with ALMA and the VLA [12].

The calibration process determines the net complex correction factors to be applied to each visibility. Corrections are applied to the data due to, e.g., temperature effects, atmospheric effects, antenna gain-elevation dependencies. It includes solvers for basic gain (A, ϕ), bandpass, polarization effects, antenna-based and baseline-based solutions. Between calibration and imaging procedures, image formation follows in three steps: weighting, convolutional resampling, and a Fourier transform. This is a general procedure employed by most systems over the last 30 years. In order to produce images with improved thermal noise and to customize the resolution, the calibrated visibilities undergo a variety of weighting schemes. The commonly used robust weighting [13] has the property to vary smoothly from natural (provides equal weight to all samples) to uniform (gives equal weight to each measured spatial frequency irrespective of sample density) based on the signal-to-noise ratio of the measurements and a pre-defined noise threshold parameter. The robust weighting scheme has the advantages to make the effective u - v coverage as smooth as possible while allowing for sparse sampling, to modify the resolution and theoretical signal-to-noise of the image. Nonetheless, care has to be taken on the choice of weighting scheme due to sidelobes effects. The weighted visibilities are resampled onto a regular uv -grid (convolutional resampling) employing a gridding convolution function. A fast Fourier transform is applied on the resampled data and corrections are introduced to transform the image into units of sky brightness. The resulting dirty image $I^D(l, m)$ is deconvolved from the well-known dirty beam $b(l, m)$. A model of the sky brightness distribution is obtained, allowing for the creation of the dirty/residual image, as the result of a convolution of the true sky brightness and the PSF of the instrument. The dirty image arises or from single pointing or from mosaics. Last, CASA offers facilities for simulating observations.

3.1. Deconvolution

Most interferometric imaging in CASA is done using the *tclean* task. This task is composed by several operating modes, allowing for the generation of images from visibilities and the reconstruction

of a sky model. Continuum images and spectral line cubes are handled. Image reconstruction occurs employing an outer loop of major cycles and an inner loop of minor cycles, following the Cotton-Schwab CLEAN style [14]. The major cycle accounts for the transformation between data and image domains. The minor cycle is designed to operate in the image domain. An iterative weighted χ^2 minimization is implemented to solve for the measurement equations.

Several algorithms for image reconstruction (deconvolvers) in the minor cycle are available, e.g., Hogbom [15], Clark [16], Multi-Scale [17], Multi-Term [18]. Each minor cycle algorithm can be characterized by their own optimization scheme, framework and task interface with the pre-requisite to produce a model image as output. BRAIN has the potentials to become a deconvolution algorithm within the *tclean* task.

3.2. Simulated ALMA Data

The CASA simulator takes on input a sky model to create a customized ALMA interferometric or total power observations, including multiple configurations of the 12-m array. The CASA simulator is characterized by two main tasks: *simobserve*, *simanalyze*. The task *simobserve* is used to create a model image (or component list), i.e., a representation of the simulated sky brightness distribution, and the (u, v) data. The dirty image is created with the task *simanalyze*.

In Figure 1, 13 point-like and extended sources are simulated at 97 GHz in the 12-m array for an integration time of 30 s and a total observing time of 1 h. The simulated sources are characterized by a Gaussian flux distribution in the range 0.2–5.0 Jy. Low thermal noise is introduced. The image size is composed by [1024, 1024] pixels, covering 10 arcsec on the side. These and more sophisticated simulated data are planned to be used for the feasibility study.

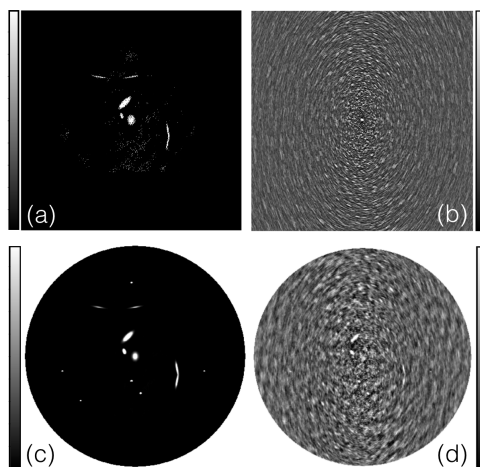


Figure 1. Simulated ALMA interferometric data: (a) the simulated model image $(2.435\text{--}10.685) \times 10^{-5}$ [Jy/pixel], (b) the synthesized (dirty) beam $(-0.02\text{--}0.04)$ [Brighness pixel unit], (c) the simulated sky $(-0.25\text{--}3.76) \times 10^{-3}$ [Jy/beam], (d) the dirty image $(-0.05\text{--}0.08)$ [Jy/beam].

4. BRAIN

An ALMA dirty image (Equation (3)) is characterizable or by continuum or by line detection for a given bandwidth. The detected celestial source intensities can be in absorption or emission with negative or positive values, respectively. We assume that the data went through a robust calibration process. Gaussian statistics describes the data distribution. Nonetheless, efforts are also sought to account for glitches in the calibration process. For those cases, BRAIN is planned to implement an improvement in the data modelling accounting for a Gaussian distribution with a longer tail. The data set in image space is $D = \{d_{ij}\} \in \mathbf{R}$ in Jansky per pixel cell $\{i, j\}$. Following the work of [3], an astronomical image consists of a composition of background and celestial signals. The background is

defined as the small scale perturbation of the wanted celestial signal. Two complementary hypotheses are introduced:

$$\begin{cases} B_{ij} : & d_{ij} = b_{ij} + \epsilon_{ij} \\ \bar{B}_{ij} : & d_{ij} = b_{ij} + s_{ij} + \epsilon_{ij}. \end{cases} \quad (4)$$

Hypothesis B_{ij} specifies that the data d_{ij} consists only of background intensity b_{ij} spoiled with noise ϵ_{ij} , i.e., the (statistical) uncertainty associated with the measurement process. Hypothesis \bar{B}_{ij} specifies the case where additional source intensity s_{ij} contributes to the background. Additional assumptions are that negative and positive values for source and background amplitudes are allowed and that the background in average is smoother than the source signal.

The background signal is propagated from the visibilities noise, taking into account that the Fourier Transform is a linear combination of the visibilities with some rotation (phase factor) applied. The noise on the visibilities is mainly introduced, e.g., by cosmic, sky and instrumental (due to receiver, single baseline or antenna, total collecting area, autocorrelations) signals. It is expected that real and imaginary part in the visibility noise is uncorrelated, with the phase factor not affecting the noise. Moreover, the individual visibilities are combined at the phase center and weighted, gridding correction and primary beam correction are increasing noise in the image. The noise in ALMA images is not uniform, with the noise increasing towards the edge. Mosaics of images are particularly cumbersome.

The background amplitude is modelled with a thin-plate-spline, with the support points chosen sparsely in order not to fit the sources [3]. The background model takes into account the dirty beam information, in order to deconvolve the dirty beam from the dirty image. The spline model is under further development to account for a more flexible design. The supporting points are analysed to account for a dynamic setting, where the number and positions of the supporting points are chosen on the basis of the data. This work follows [2] successfully developed in experimental spectra.

Estimates of the hypotheses B_{ij} and \bar{B}_{ij} are the direct effort of this analysis. The likelihood probability for the hypothesis B_{ij} within Gaussian statistics is:

$$p(d_{ij} | B_{ij}, b_{ij}) = (2\pi\sigma_{ij}^2)^{-\frac{1}{2}} \exp[-(d_{ij} - b_{ij})^2 / 2\sigma_{ij}^2]. \quad (5)$$

For the alternative hypothesis, a similar equation is applied with included the signal contribution. Similar to what was done in [2,3], the signal is considered as a nuisance parameter. Following the Maximum Entropy distribution, a two-sided exponential function is chosen to describe the mean source intensity in the field ($\lambda_{+/-} = \langle s_{ij} \rangle$):

$$p(s_{ij} | \lambda_+, \lambda_-) = \begin{cases} \lambda_+^{-1} \exp(-s_{ij} / \lambda_+), & s_{ij} \geq 0 \\ \lambda_-^{-1} \exp(+s_{ij} / \lambda_-), & s_{ij} < 0. \end{cases} \quad (6)$$

The parameters λ_+ and λ_- are positive values, allowing one to introduce two different scales for the signal dependently on its sign. The likelihood for the hypothesis \bar{B}_{ij} is given following the marginalization rule:

$$\begin{aligned} p(d_{ij} | \bar{B}_{ij}, b_{ij}, \lambda) &= \int_{-\infty}^{+\infty} ds_{ij} p(d_{ij} | s_{ij}, \bar{B}_{ij}, b_{ij}) p(s_{ij} | \lambda_+, \lambda_-) \\ &= \frac{1}{2|\lambda_+|} \left\{ 1 + \operatorname{erf} \left[\frac{\lambda_+(d_{ij} - b_{ij}) - \sigma_{ij}^2}{|\lambda_+| \sqrt{2\sigma_{ij}^2}} \right] \right\} \exp \left[\frac{\lambda_+(d_{ij} - b_{ij}) + \sigma_{ij}^2/2}{\lambda_+^2} \right] + \\ &+ \frac{1}{2|\lambda_-|} \left\{ 1 + \operatorname{erf} \left[\frac{-\lambda_-(d_{ij} - b_{ij}) - \sigma_{ij}^2}{|\lambda_-| \sqrt{2\sigma_{ij}^2}} \right] \right\} \exp \left[\frac{-\lambda_-(d_{ij} - b_{ij}) + \sigma_{ij}^2/2}{\lambda_-^2} \right]. \end{aligned} \quad (7)$$

The data are modelled by a two-component mixture distribution in the parameter space within the Bayesian framework. In this way, background and sources with their respective uncertainties are jointly detected. The likelihood for the mixture model combines the probability distribution for the two hypotheses, B_{ij} and \bar{B}_{ij} :

$$p(D|\mathbf{b}, \mathbf{s}, \boldsymbol{\lambda}) = \prod_{ij} \{p(B_{ij}) \cdot p(d_{ij}|B_{ij}, b_{ij}) + p(\bar{B}_{ij}) \cdot p(d_{ij}|\bar{B}_{ij}, b_{ij}, \lambda_+, \lambda_-)\}, \quad (8)$$

where $\mathbf{b} = \{b_{ij}\}$, $\mathbf{s} = \{s_{ij}\}$, $\boldsymbol{\lambda} = \{\lambda_{+/-}\}$ and $\{ij\}$ corresponds to the pixels of the complete field. The prior pdfs $p(B_{ij})$ and $p(\bar{B}_{ij})$ for the two complementary hypotheses describe the prior knowledge of having background only or additional signal contribution, respectively, in a pixel. These prior pdfs are chosen to be constant, independent of i and j : $p(B_{ij}) = \beta$ and $p(\bar{B}_{ij}) = 1 - \beta$. The likelihood for the mixture model allows us to estimate the parameters entering the models from the data. The ultimate goal is to detect sources independently to their shape and intensity and to provide a robust uncertainty quantification. Therefore, Bayes' theorem is used to estimate the probability of the hypothesis $p(\bar{B}_{ij} | d_{ij}, b_{ij}, \boldsymbol{\lambda})$ for detecting celestial sources.

5. Concluding Remarks

The work is ongoing for this project. The novel technique, applied to ALMA interferometric data, is designed to create probability maps of source detection, allowing for a joint estimate of background and sources. A 2-D adaptive kernel deconvolution method will strengthen the deconvolution process of the dirty beam from the dirty image, reducing contaminations. BRAIN has the aim to provide an alternative technique in CASA image analysis. It is foreseen that a robust solution as model image is produced. The next effort is applied on the simulated data. Nonetheless, due to the statistical approach employed, BRAIN is also foreseen to be applied on data re-combination. In fact, advanced methods employing Bayesian probability theory have the potentials to address at best the analysis of combined short-spacing single-dish data with those from an interferometer, as provided by ALMA design.

Acknowledgments: This work was supported partly by the Joint ALMA Observatory Visitor Program.

References

- ALMA Dev. Workshop 2019. Available online: <https://zenodo.org/record/3240347#.XVq6PVBS9XQ> (accessed on 1 December 2019).
- Fischer, R.; Hanson, K.M.; Dose, V.; von der Linden, W. Background estimation in experimental spectra. *Phys. Rev.* **2000**, *E61*, 1151–1160.
- Guglielmetti, F.; Fischer, R.; Dose, V. Background-source separation in astronomical images with Bayesian probability theory—I. The method. *MNRAS* **2009**, *396*, 165–190.
- CASA Documentation. Available online: <https://casa.nrao.edu/casadocs-devel> (accessed on 1 December 2019).
- McMullin, J.P.; Waters, B.; Schiebel, D.; Young, W.; Golap, K. CASA Architecture and Applications. In *Astronomical Data Analysis Software and Systems XVI ASP Conference Series*; Shaw, R.A., Hill, F., Bell, D.J., Eds.; Astronomical Society of the Pacific: San Francisco, CA, USA; Volume 376, pp. 127–130.
- Next Generation Very Large Array. Available online: <https://ngvla.nrao.edu/> (accessed on 1 December 2019).
- Square Kilometre Array. Available online: <https://www.skatelescope.org/> (accessed on 1 December 2019).
- Braun, R.; Walterbos, R.A.M. A solution to the short spacing problem in radio interferometry. *Astron. Astrophys.* **1985**, *143*, 307–312.
- Zernike, F. The concept of degree of coherence and its application to optical problems. *Physica* **1938**, *5*, 785–795.
- Stanimirovic, S.; Altschuler, D.; Goldsmith, P.; Salter, C. Single-Dish Radio Astronomy: Techniques and Applications. In *ASP Conference Proceedings*; Stanimirovic, S., Altschuler, D., Goldsmith, P., Salter, C., Eds.; Astronomical Society of the Pacific: San Francisco, CA, USA; Volume 278, pp. 375–396.

11. Jaeger, S. The Common Astronomy Software Application (CASA). In *Astronomical Data Analysis Software and Systems ASP Conference Series*; Argyle, R.W., Bunclark, P.S., Lewis, J.R., Eds.; Astronomical Society of the Pacific: San Francisco, CA, USA; Volume 394, pp. 623–626.
12. Very Large Array. Available online: <http://www.vla.nrao.edu/> (accessed on 1 December 2019).
13. Briggs, D.S. High Fidelity Interferometric Imaging: Robust Weighting and NNLS Deconvolution. *Bull. Am. Astron. Soc.* **1995**, *27*, 1444.
14. Schwab, F.R. Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry. *Astron. J.* **1984**, *89*, 1076–1081.
15. Hogbom, J.A. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astron. Astrophys. Suppl. Ser.* **1974**, *15*, 417–426.
16. Clark, B.G. An efficient implementation of the algorithm ‘CLEAN’. *Astron. Astrophys.* **1980**, *89*, 377–378.
17. Cornwell, T.J. Multiscale CLEAN Deconvolution of Radio Synthesis Images. *IEEE J. Sel. Top. Signal Process.* **2008**, *2*, 793–801.
18. Rau, U.; Cornwell, T.J. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. *Astron. Astrophys.* **2011**, *532*, A71.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).