Radio Interferometry with Compressed Sensing

Jonas Schwammberger

Today

Abstract

Solar Flares Abstract

Contents

1 Interferometry and the Inverse Problem

Astronomy requires its instruments to have a high angular resolutions. This is an issue for radio wavelengths: The longer the wavelength, the bigger the diameter of a single dish antenna. Single dish antenna's are expensive to build and harder to steer accurately. Interferometers, where multiple smaller antennas act as a single large instrument, have had great success in Radio Astronomy. Instruments like VLA, ALMA and LOFAR have produced high-resolution images.

Interferometers do not observe the sky directly. Each antenna pair measure Fourier Components (Visibilities) of the sky brightness. The observed image has to be reconstructed from the measured Visibilites. Since the interferometer can only observe a limited number of Visibilities, the reconstruction is an ill-posed inverse problem.

For small Field of View imaging, the CLEAN class of Algorithms[**?**] have been developed and is the de-facto standard in Radio Astronomy. It is not guaranteed to reconstruct the true image in theory, it has produced remarkable results with expert tuning. New generation Interferometers like ASKAP, Pathfinder and SKA are built with wide Field of View imaging in mind. The CLEAN Algorithms have been extended for Wide Field of Views, but require even more tuning by experts.

The Theory of Compressed Sensing[**?**] has seen success in solving ill-posed inverse problems in other fields like image de-noising[**?**], in-painting[**?**] and super-resolution[**?**]. Applying the Theory of Compressed Sensing to Radio Interferometry is an active field of research.

It can work in Wide Field imaging, It has the potential to reduce the expert tuning, is guaranteed to reproduce the true image under the right conditions and has the potential to super-resolve.

A proof of concept Compressed Sensing Algorithm implemented in the Common Astronomy Software Application (CASA).

1.1 Inverse Problem for small Field of View Imaging

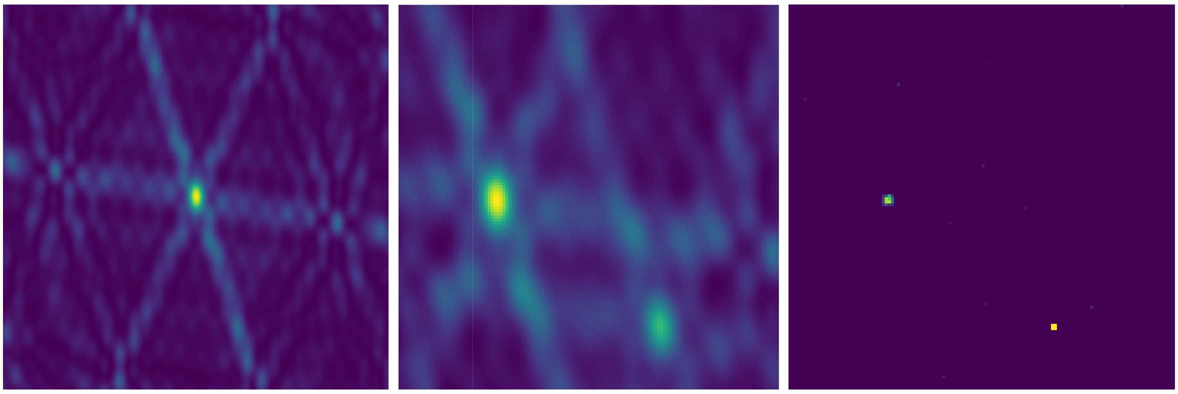
For small Field of Views, the inverse problem can be simplified(although it is still ill-posed). Each antenna pair measures a Visibility of the sky brightness. The distance between the antenna's

For the current class of interferometers, retrieving the observed image of the sky is calculating the Inverse Fourier Transform of the measured Visibilities. (This is a simplification of the actual problem, which does not hold true for future instruments like SKA. How the Inverse Problem changes for future instruments is handled in section 2). Since the Visibility Space is undersampled, the resulting image is 'dirty', as it contains effects from the instrument. The effects of the instrument can be modelled as a Point Spread Function (PSF). The Dirty Image is the Result of the True Image convolved with the PSF. Figure 1 shows an example of the PSF, Dirty Image and True Image.

The Dirty Image is the result of the Inverse Fourier Transform, the PSF on the other hand follows directly from the antenna configuration. Figure 2 shows an example of the VLA instrument.

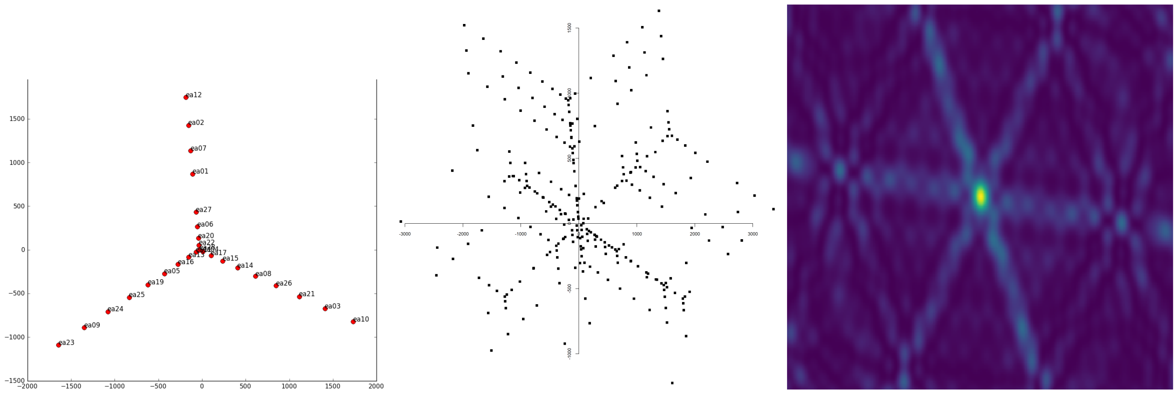
Inverse Fourier Transform

Each antenna pair samples a point in the UV Space. The distance between the antenna pair is called the baseline. Shorter baselines sample UV-Points closer to the center, while longer baselines sample points further away. Each point in the UV Space represents a frequency Visibilities of the observed image: High frequency components are located away from the center and low frequency components are towards the center. Since each antenna pair contributes a Visibility, an Interferometer of N antennas measures $(N-1)/2$



(a) Point Spread Function (PSF) (b) Dirty Image (c) True Image

Figure 1: The Inverse Problem: Finding the true image of the sky when only the PSF and the dirty image are known.



(a) Antenna Configuration (b) UV Space (c) Point Spread Function (PSF)

Figure 2: The Antenna Configuration sets up the UV Space. The UV Space dictates the PSF.

Visibilities. The configuration of the antenna sets up which points are sampled in the UV Space. The PSF can be calculated by using the Inverse Fourier Transform on the UV-Space. (Rotation of Earth)

Since the Dirty Image and the PSF are known, one needs to deconvolve the Dirty Image with the PSF and try to retrieve the True Image. More formally, it tries to find X from the equation (1.1) in a noiseless environment.

$$X \star PSF = D_{dirty} \tag{1.1}$$

Ill-Posed

1.2 Deconvolution with CLEAN

The CLEAN class of Algorithms[1][2][3][4] are widely used in Radio Astronomy. It does not solve the deconvolution problem (1.1) directly. Instead, it minimizes a similar optimization problem with the objective (1.2). In each iteration, CLEAN searches the highest peak in D_{dirty} and removes a fraction of the PSF at that point. It is a greedy optimizes the objective (1.2). It is easy to show that if CLEAN minimizes the objective to zero, then it has found a solution to the original problem (1.1).

$$\underset{X}{minimize}\ \|D_{dirty} - X \star PSF\|_2^2 \tag{1.2}$$

In a noiseless environment, the true image would be located where the objective of (1.2) is zero. In the real world however, noise corrupts the problem and the true image may not be at the minimum any more. To address this, CLEAN is stopped early by limiting the number of iterations or when the peak is below a threshold. The algorithm should find the minimum of the objective (1.2) where only a limited number of pixels are allowed to be non-zero; an implicit regularization term. With the right parameters, the regularization should stop the noise from corrupting the result and CLEAN should reproduce the true image.

So does CLEAN reproduce the image? It can, but it does not have to. The algorithm has no guarantees, we do not even know if the result is close to the true image.

1.3 From CLEAN to Compressed Sensing

An Algorithm in the Compressed Sensing Framework has three components:

- An objective function with a data and regularization term
- A Matrix P in which the signal can be sparsely represented.
- An optimization algorithm that is able to handle the objective function

CLEAN can be converted into the Compressed Sensing Framework. First, the regularization term has to be explicit, it gets added to the objective function. The resulting objective (1.3) has two terms: A data term and a regularization term. The data term forces the reconstruction to be close to the measurement, while the regularization term forces the reconstruction to be plausible. λ models the trade off between the terms. Note that the zero norm $\|PX\|_0$ acts as an indicator function and is not technically a norm.

$$\underset{X}{minimize}\ \|D_{dirty} - X \star PSF\|_2^2 \ + \ \lambda \|PX\|_0 \tag{1.3}$$

the new objective is the Compressed Sensing version of CLEAN. It assumes that the reconstruction is sparse in the image domain. When P is the identity matrix, This is true when there are only a few point sources located in the image. Extended sources are are not well represented and are harder to detect.

The objective function is optimized by a greedy solver.

in the compressed Sensing Framework CLEAN is a specific objective function with an identity matrix as prior and a specific optimization algorithm

In this part, mostly the prior is

The important questions are: In an undersampled, noisy environment, does the new objective (1.3) have a global minimum? What is the chances that the minimum is equal to the true image? It turns out that even though we have fewer samples than the Nyquist-Shannon Theorem requires, we can guarantee a global minimum and that the minimum is equal to the true image, if we have enough prior knowledge P about the signal. How we model the signal and by extend, what P we choose is essential for Compressed Sensing

Incoherence

1.3.1 Modelling Prior Knowledge and Compressible Signals

2 Interferometry and the actual Inverse Problem

So far the simplified inverse problem has been introduced. Each antenna pair measures a Fourier Component of the sky brightness distribution. The distance between antenna pairs dictates what point is sampled in the UV plane. This leads to the measurement equation (2.1). Calculating the true image X is simply inverting the two dimensional Fourier Transform.

$$V(u,v) = \int\int X(x,y)e^{2\pi i(ux+vy)}dxdx \tag{2.1}$$

In reality, each visibility has a third component w. It comes from the fact that the antennas are not on a flat plane but on the curved surface of the earth. Image 3 shows the three dimensional space. w is the vector points from the phase center to the science target. For small Field of Views, the effect is negligible and the measurement equation (2.1) is a good approximation. The field of view is limited by the primary beam of the antennas. Primary beam widens with wavelength. So far, wide Field of View problems were encountered in low frequencies like LOFAR.

new instruments like SKA, ASKAP and Pathfinder are constructed with a wide Field of View in mind. The simple two dimensional Fourier Transform does not hold true anymore and we arrive at the wide Field of View measurement equation (2.2).

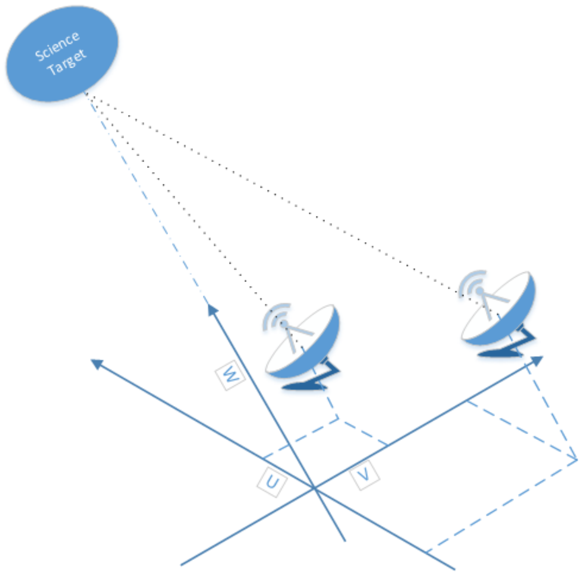


Figure 3: U V and W coordinate space

$$V(u,v,w) = \int\int \frac{X(x,y)}{\sqrt{1-x^2-y^2}}e^{2\pi i(ux+vy+w\sqrt{1-x^2-y^2})}dxdy \tag{2.2}$$

Note that for small Field of View $1 - x^2 - y^2 \ll 1$, and equation (2.2) can be approximated by the 2d Fourier Transform (2.1). Two separate effects: w Component for non coplanar baselines.

W-Projection [5]

A-Projection [6]

All here to try to get back to the 2d fourier transform.

"The field of view of a telescope is limited by the primary beams of the antennas. To map a region of sky where the emission is at a scale larger than the angular width of the primary beams, mosaicing needs to be done. This is discussed in the second part of this lecture." Phase

source [7]

Strength of compressed sensing is modelling these effects.

2.1 Calibration

A lot of effects, weather, noise, antenna temperature, drift.

Antenna based calibration, holds true for current interferometers but is not true for SKA. Possible switch to baseline based calibration.

2.1.1 Self-Calibration

?

3 Compressed Sensing for Radio Astronomy

A compressed sensing algorithm consists of three components: A Prior, an Objective and an Optimization Algorithm.

To Include wide field imaging in the objective function directly, which has the potential to improve reconstruction with Model the signal to so the reconstruction is plausible Use an optimization algorithm that is able to handle the expected amount of data. Interferometers produce a large amount of data. In this project, the optimization algorithm was not further investigated. The GuRoBi simplex solver was used.

Compressed Sensing for Radio Astronomy is an active research field. The new instrument's push to wide Field of View imaging has led to increased interest.

The guarantees in Compressed Sensing depend on the signal, in what space the signal is measured and how well our Prior models the signal.

3.1 Sparseland Prior and Incoherence

From compression algorithm, we assume there is a Prior P in which our signal can be sparsely represented. It is not guaranteed that such a space exists, but for natural signals there always seem to be. This has led to the idea of the Sparseland Prior (3.1): We assume for our signal there exists a Dictionary Dic of signal parts. There is potentially a large, but finite number entries of parts. When we measure our signal x, we see a combination of only a few entries non-zero entries α of the Dictionary.

$$\begin{aligned} x = Dic\,\alpha \qquad x \in \mathbb{R}^n, \alpha \in \mathbb{R}^m, Dic \in \mathbb{R}^{n*m}, \qquad n \leq m \\ \|\alpha\|_0 = s \qquad s \ll n \leq m \end{aligned} \tag{3.1}$$

Pictures of nature scenes for example tend to be sparse in the wavelet domain. If x in (3.1) are nature scenes, we can create a Dictionary Dic of wavelets. A single image x then consists out of a combination of a few wavelets, meaning the number of non-zero entries is far lower than the number of pixels n.

Noise tends to affect all entries of the Dictionary. Sparseland Prior has had success in image denoising.

The signal parts is not restricted to be in one domain. It can consists of wavelets, cosine functions, a combination of both, or any other function.

Work with over-complete dictionaries, where the number of pixels n is far smaller than the number of signal parts m. There are redundant entries which is OK, as long is it is not too redundant. There is a tradeoff in practice in how sparse a signal can be represented and how redundant the dictionary is.

Back to the ill-posed inverse problem of interferometry. Nyquist Shannon Sampling Theorem requires that if we want an image of n pixels, we need at least $2n$ measurements. The small Field of View Interferometer measures complex visibilities, so for n fully sampled pixels we need n complex Visibility measurements.

But when we know it is a Sparseland signal and we know the Dictionary, we could try to measure in the Dictionary space, so we could try to retrieve the non-zero components and only measure s samples. Sadly, this is only possible if we know which entry of the dictionary are non-zero beforehand, which is in general not possible. So we would measure different α and are more likely to hit a zero component. Note however that if we measure a non-zero component of α, we learn a lot more about our signal than when we hit a zero component. So the next question is, can we measure in a space where we learn more about the non-zero components of α?

Surprisingly, the answer is yes, there is a way. The space in which we measure our signal x should be incoherent from our Dictionary. With that we maximize the amount we learn about the non-zero components of $alpha$ with each measurement. Interestingly enough, constructing an incoherent measurement space is easy: Random projections are very good at being incoherent from the dictionary.

Random Projections are not possible, the Interferometer measures complex visibilities. Antenna configuration could help. But luckily what also can help is the Wide Field imaging. McEwen [8] showed the theoretical improvement on synthetic data.

P in our original objective function. $Dic = P^{-1}$. So if the conversion from image space to Sparseland is only defined when the dictionary is a square matrix. When it has as many signal parts as pixels. The Objective Function can be modified to work with over-complete dictionaries.

3.2 Objective Function

L0 norm is not technically a norm. The problem is not convex.

L1 Relaxation of the Objective function

$$\underset{X}{minimize} \; \|D_{dirty} - X \star PSF\|_2^2 \; + \; \lambda \left\|PX\right\|_1 \tag{3.2}$$

The actual CS formulation uses the indicator function for regularization. The indicator function is Zero for each element of $P(X)$ that is also zero, and 1 for each element that is not zero. This objective function (**??**) is not Convex anymore. There are specialized optimization algorithms that minimize the objective (**??**) like Matching Pursuit.

$$\begin{aligned} PSF = F^{-1}MD_{dirty} = F^{-1}MV \\ X = Dic\,\alpha \end{aligned} \tag{3.3}$$

$$\begin{aligned}
&\underset{X}{minimize} \; \|D_{dirty} - X \star PSF\|_2^2 &&+ \lambda \left\|P^{-1}X\right\|_1 \\
&\underset{\alpha}{minimize} \; \|D_{dirty} - P\alpha \star PSF\|_2^2 + \lambda \left\|\alpha\right\|_1 && \\
&\underset{V_2}{minimize} \; \left\|D_{dirty} - F^{-1}MV_2\right\|_2^2 &&+ \lambda \left\|P^{-1}F^{-1}V_2\right\|_1 \\
&\underset{V_2}{minimize} \; \|V - MV_2\|_2^2 &&+ \lambda \left\|P^{-1}F^{-1}V_2\right\|_1
\end{aligned}$$

All these equations should produce the same result in theory. In practice these different approaches have implications on the model and the chosen optimization algorithm. Not every D^{-1} is defined. Coordinate descent works well on minimizing in the sparse space α directly. Also the convolution has to be explicitly handled when reconstructing in the image space.

L2 norm in the Image space weights more heavily on the lower frequency components.

3.2.1 Implementation In Casa

Casa major and minor cycle. Major cycle calculates visibilities in image space. Minor Cycle Deconvolves the Problem, often with a CLEAN class Algorithm. This constrains the algorithm to use the data term in image space.

This forces the objective function to either minimize in the image domain or in the sparsity domain.

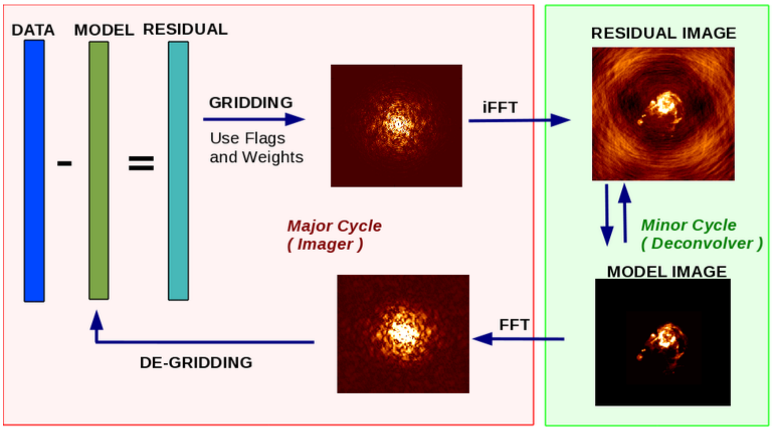


Figure 4: Casa Major Minor Cycle, source [9]

3.3 Compressed Sensing Algorithms in Astronomy

3.3.1 SASIR

3.3.2 PURIFY

3.3.3 Vis-CS

4 Results

image plane should be greater than zero. Physical plausability and shown to produce better results on synthetic data [8]

Results

Clean images:

Gurobi Clean

4.1 Simple Priors

pixels l1 norm

pixels l2 norm

Total Variation

Haar

4.2 Starlet Transform as Prior

starlet decomposition

the cJ map as a smart thresholder

Runtime issues

last comparison, CLean, TV, starlet

5 Conclusion and Future Development

References

[1] JA Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. Astronomy and Astrophysics Supplement Series, 15:417, 1974.

[2] FR Schwab. Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry. The Astronomical Journal, 89:1076–1081, 1984.

[3] JW Rich, WJG De Blok, TJ Cornwell, Elias Brinks, Fabian Walter, Ioannis Bagetakos, and RC Kennicutt Jr. Multi-scale clean: A comparison of its performance against classical clean on galaxies using things. The Astronomical Journal, 136(6):2897, 2008.

[4] Urvashi Rau and Tim J Cornwell. A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry. Astronomy & Astrophysics, 532:A71, 2011.

[5] Tim J Cornwell, Kumar Golap, and Sanjay Bhatnagar. The noncoplanar baselines effect in radio interferometry: The w-projection algorithm. IEEE Journal of Selected Topics in Signal Processing, 2(5):647–657, 2008.

[6] Cyril Tasse, S van der Tol, J van Zwieten, Ger van Diepen, and S Bhatnagar. Applying full polarization a-projection to very wide field of view instruments: An imager for lofar. Astronomy & Astrophysics, 553:A105, 2013.

[7] Institut de RadioAstronomie Millimétrique. Proceedings from imiss2 “iram millimeter interferometry summer school 2”, 2000.

[8] Jason D McEwen and Yves Wiaux. Compressed sensing for radio interferometric imaging: Review and future direction. In Image Processing (ICIP), 2011 18th IEEE International Conference on, pages 1313–1316. IEEE, 2011.

[9] National Radio Astronomy Observations. tclean overview, 2016.

List of Figures

List of Tables

6 Ehrlichkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende schriftliche Arbeit selbstständig und nur unter Zuhilfenahme der in den Verzeichnissen oder in den Anmerkungen genannten Quellen angefertigt habe. Ich versichere zudem, diese Arbeit nicht bereits anderweitig als Leistungsnachweis verwendet zu haben. Eine Überprüfung der Arbeit auf Plagiate unter Einsatz entsprechender Software darf vorgenommen werden.
Windisch, June 30, 2018

Jonas Schwammberger