

# Radio Interferometry with Compressed Sensing

Jonas Schwammberger

Today

## **Abstract**

Solar Flares Abstract

## Contents

<b>1</b>	<b>Interferometry and the Inverse Problem</b>	<b>1</b>
1.1	Simplified Inverse Problem . . . . .	1
1.2	Deconvolution with CLEAN . . . . .	2
1.3	From CLEAN to Compressed Sensing . . . . .	3
1.3.1	Modelling Prior Knowledge and Compressible Signals . . . . .	3
1.3.2	Parts in a Compressed Sensing Framework . . . . .	3
<b>2</b>	<b>Interferometry and the actual Inverse Problem</b>	<b>5</b>
2.1	Calibration . . . . .	5
<b>3</b>	<b>The Theory of Compressed Sensing</b>	<b>7</b>
3.1	Modelling Signals . . . . .	7
3.2	Objective Function for Compressed Sensing . . . . .	7
3.3	Optimization Algorithms . . . . .	8
3.4	Compressed Sensing Algorithms in Astronomy . . . . .	8
3.4.1	SASIR . . . . .	8
3.4.2	PURIFY . . . . .	8
3.4.3	Vis-CS . . . . .	8
3.5	Implementation In Casa . . . . .	8
<b>4</b>	<b>Results</b>	<b>9</b>
4.1	Algorithms . . . . .	9
<b>5</b>	<b>Conclusion and Future Development</b>	<b>10</b>
<b>6</b>	<b>Ehrlichkeitserklärung</b>	<b>13</b>

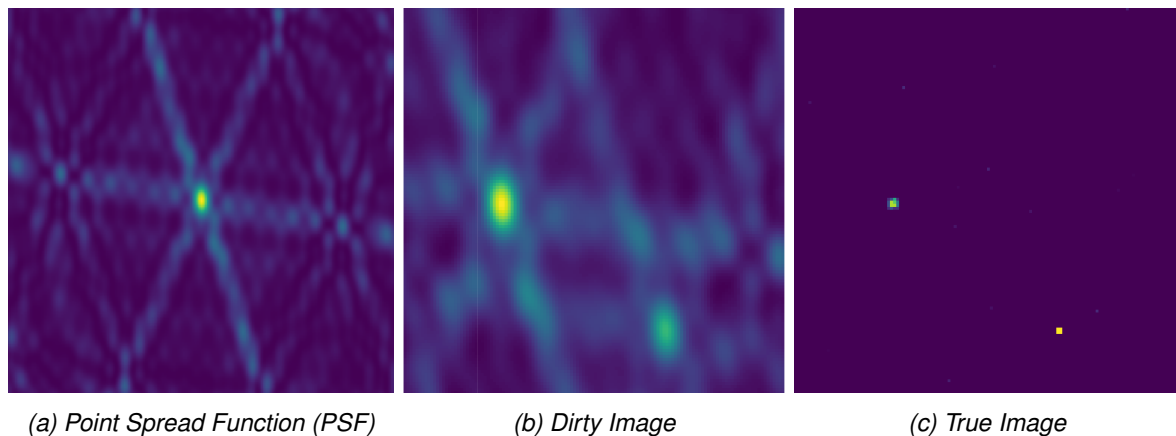
# 1 Interferometry and the Inverse Problem

Radio Antennas in Astronomy require a sub-arcsecond angular resolution. For a single dish antenna, the angular resolution is  $\theta \approx \lambda/D$  radians, which leads to impossibly large diameters for radio wavelengths. Interferometers, where several dishes act as a single instrument, can achieve a high angular resolution for an economical price. (Radio Astronomy has a lot of interest in Interferometers, building VLA, ALMA, LOFAR and planning SKA). However Interferometers do not observe the sky directly. Each antenna pair measures a Fourier component of the observed image. Furthermore the Fourier space is not fully sampled, it is incomplete. The task is to reconstruct an image in Pixel space from an incomplete set of Fourier measurements (Visibilities). From here on forward term Visibility is used. It is interchangeable with Fourier Component.

Each antenna pair, quadratic number of data points with the number of antennas: a lot of data)(An interferometer does not observe the image directly. Simplified, each antenna pair observes a Fourier component of the image.) (If all the fourier components are measured, the observed image could be retrieved by calculating the inverse fourier transform. However, only a limited set of visibilities can be measured.)(Inverting the problem results in an image with artefacts (Dirty Image).) (Clean can solve it for the current class of interferometers)

## 1.1 Simplified Inverse Problem

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.



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

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.

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

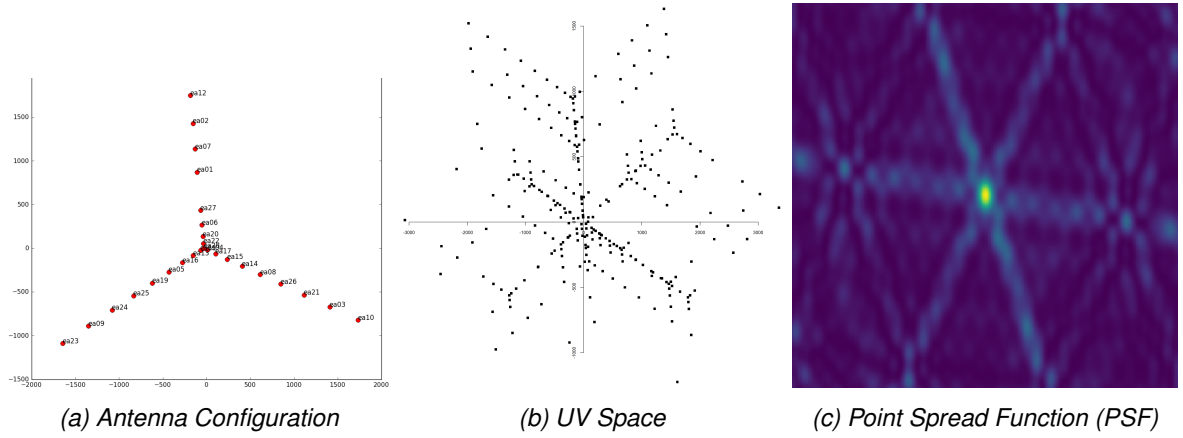


Figure 2: The Antenna Configuration sets up the UV Space. The UV Space dictates the PSF.

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  $(N - 1)/2$  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.

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.

$$X \star PSF = D_{dirty} \quad (1.1)$$

## 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).

$$\underset{X}{\text{minimize}} \quad \|D_{dirty} - X \star PSF\|_2^2 \quad (1.2)$$

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

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 (3.1) 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.  $\lambda$  models the tradeoff between the terms. Note that the zero norm  $\|PX\|_0$  acts as an indicator function and is not technically a norm.

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

When  $P$  is the identity matrix, the new objective is the Compressed Sensing version of CLEAN. It assumes that the reconstruction is sparse in the image domain. This is true when there are only a few point sources located in the image. Extended sources are not well represented and are harder to detect. The objective function is optimized by a greedy solver. An Algorithm in the Compressed Sensing Framework needs three parts:

- 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

The important questions are: In an undersampled, noisy environment, does the new objective (3.1) have a global minimum? What are 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

#### 1.3.1 Modelling Prior Knowledge and Compressible Signals

The objective (3.1) requires the product  $PX$  to be sparse (produces few non-zero entries) when  $X$  is plausible. This is the same requirement as a compression algorithm and indeed, if we know our signal is compressible in the wavelet domain, then the wavelet transform is a good  $P$  for Compressed Sensing. But  $P$  is not limited to a transformation. The only requirement is that the signal's representation is as sparse as possible. There is a case for overcomplete representations for compressed sensing (vis-cs).

$P$  should be independent from the measurement space.

How sparsely the signal can be represented dictates how many samples are needed Restricted Isometry Property, RIP

CLEAN uses the  $P$  as the identity matrix.

## 2 Interferometry and the actual Inverse Problem

The simplified inverse Problem says that each antenna pair measures a Fourier Component (Visibility) of the sky.

$$V(u, v) = \iint X(x, y) e^{2\pi i(ux+vy)} dx dy \quad (2.1)$$

But that is not true, the UV space actually has a third component W. image 3 shows it. W is a vector in the direction of the science target. V is what we measure, X is the image. calculating X is the inverse fourier transform.

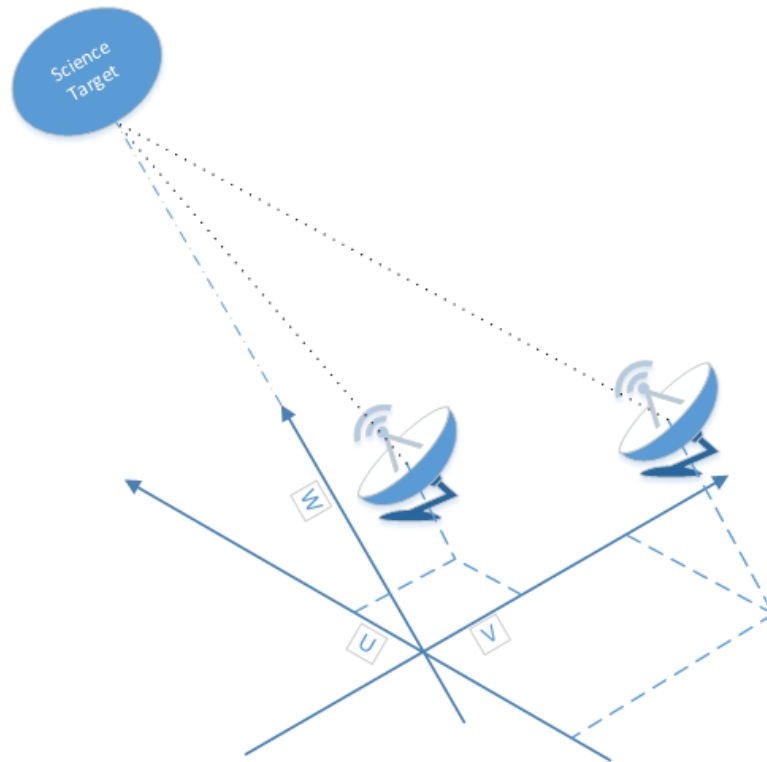


Figure 3: U V and W coordinate space

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

for small field of view the term  $x^2 + y^2 \ll 1$  and therefore can be neglected. so we arrive at the 2d fourier transform. for wide field of view, the term.

"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]

### 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.

Self-cal?



### 3 The Theory of Compressed Sensing

#### 3.1 Modelling Signals

$P(X)$  compressible signals like wavelets. It works if the sparse

Idea of a dictionary,  $X = D\alpha$  Dictionary, over complete representations. What is possible depends on the objective function

Finding a sparse space, can even be over complete and pretty much anything one wants.

#### 3.2 Objective Function for Compressed Sensing

$$\underset{X}{\text{minimize}} \|D_{\text{dirty}} - X \star PSF\|_2^2 + \lambda \|P(X)\|_1 \quad (3.1)$$

L1 Relaxation of the Objective function

$$\underset{X}{\text{minimize}} \|D_{\text{dirty}} - x \star PSF\|_2^2 + \lambda \text{indicator}(P(X)) \quad (3.2)$$

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 (3.2) is not Convex anymore. There are specialized optimization algorithms that minimize the objective (3.2) like Matching Pursuit.

Another solution is to use the L1 Relaxation instead of the indicator function. Convex Problem. And all the problems were solved.

Reconstruction in different spaces

$$\begin{aligned} D_{\text{dirty}} &= F^{-1}MV \\ X &= D\alpha \end{aligned} \quad (3.3)$$

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

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 well-defined. Coordinate descent works well on minimizing in the sparse space  $\alpha$  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.

#### 3.3 Optimization Algorithms

Not all are equal. A good algorithm depends on what.

GuRoBi was used.

### **3.4 Compressed Sensing Algorithms in Astronomy**

#### **3.4.1 SASIR**

#### **3.4.2 PURIFY**

#### **3.4.3 Vis-CS**

### **3.5 Implementation In Casa**

Because of major and minor cycle

## **4 Results**

### **4.1 Algorithms**

## **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] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on information theory, 52(2):489–509, 2006.
- [6] David L Donoho. Compressed sensing. IEEE Transactions on information theory, 52(4):1289–1306, 2006.
- [7] Institut de RadioAstronomie Millimétrique. Proceedings from imiss2 “iram millimeter interferometry summer school 2”, 2000.

## List of Figures

1	The Inverse Problem: Finding the true image of the sky when only the PSF and the observed image are known. . . . .	1
2	The Antenna Configuration sets up the UV Space. The UV Space dictates the PSF. . . . .	2
3	U V and W coordinate space . . . . .	5

## 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 15, 2018

Jonas Schwammberger