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, can achieve high angular resolutions while being cheaper to build. In the past, interferometers like VLA, AlMA and LOFAR have made numerous discoveries.

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[1][2][3][4] have been developed and is the de-facto standard in Radio Astronomy. It is not guaranteed to reconstruct the true image in theory. In practice 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[5][6] has seen success in solving ill-posed inverse problems. It is flexible in its application and has produced remarkable results image de-noising[**?**], in-painting[**?**] and super-resolution[**?**]. Applying Compressed Sensing to wide Field of View imaging is an active field of research. In the last decade numerous approaches have been developed showing the potential of Compressed Sensing: Accurately modelling the effects of wide Field of View imaging, while reducing the tunable parameters and possibly super-resolved images[7]. Current research focuses on how the effects of wide Field of View can be accurately modelled while still being computationally efficient.

In this project, a proof of concept Compressed Sensing approach was developed and implemented in the Common Astronomy Software Application(CASA). The approach is focused on small Field of View imaging and the reduction of expert intervention.

1.1 Inverse Problem for small Field of View Imaging

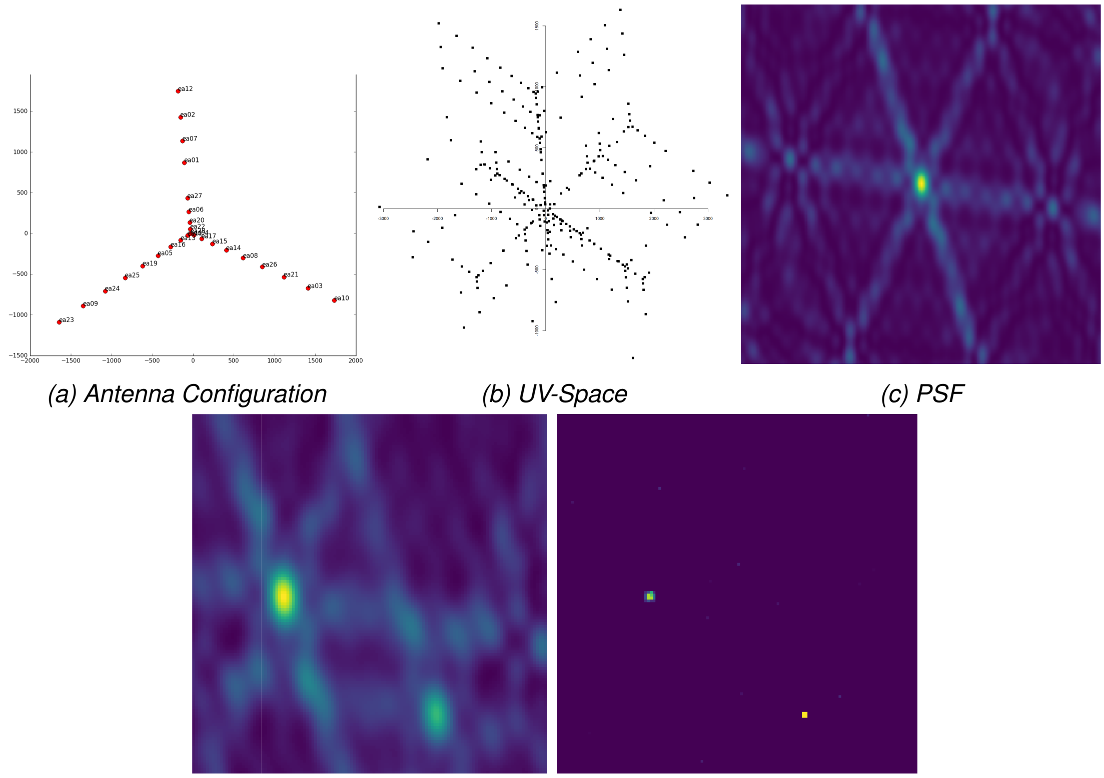
Each antenna pair measures a complex Visibility of the sky brightness. The distance between the antennas, the baseline, dictates the sample point in the Fourier Space (called UV-Space). Longer baselines sample higher frequency components, while shorter baselines sample lower frequency components.

For small Field of View imaging, the measured Visibilities equal two dimensional Fourier components. The observed image can be calculated by the two dimensional Inverse Fourier Transform. However the interferometer cannot sample the whole UV-Space. The image calculated by the Inverse Fourier Transform is 'dirty', it contains artefacts introduced by undersampling.

The Inverse Problem is now to remove the artefacts of the interferometer and retrieve the true image. The effects of the undersampling can be modeled by a Point Spread Function (PSF). The interferometer sees the true image of the sky, but due to undersampling it gets convolved with a PSF, resulting in the dirty image. More formally, we try to find a solution x for equation (1.1), where only the PSF and I_{dirty} are known. This problem is ill-posed: it may have multiple solutions, and a small change in the I_{dirty} or the PSF may result in large changes in x. Furthermore, the whole problem gets corrupted by noise.

$$x \star PSF + N = I_{dirty} \tag{1.1}$$

The PSF is surprisingly easy to calculate. The Fourier Transformed PSF equals the sampling pattern in UV-Space. Remember that a convolution in image space is a multiplication in Fourier. The effects of under-



(a) Antenna Configuration *(b) UV-Space* *(c) PSF*

(d) dirty image *(e) true image*

Figure 1: Inverse Problem example for VLA: Retrieve the true image when only PSF and dirty image are known

sampling in image space are a convolution with the PSF. In the Fourier space it is masking all components other than the measured ones. From the Antenna Configuration we can infer the masking matrix M in UV-Space. Calculating the Inverse Fourier Transform of M results in the PSF.

1.2 Deconvolution with CLEAN

In each iteration of CLEAN, it searches the highest peak of the dirty image and removing a fraction of the PSF at that point. It stops until the next highest peak is below a threshold, or if the maximum number of iterations was reached. The fraction of the PSF, threshold and number of iterations are all tunable by the user. State of the art implementations expose even more parameters. The reconstruction quality depends on the chosen parameters and require extensive user input.

CLEAN does not solve the deconvolution problem (1.1) directly. Instead, it greedily minimizes the objective (1.2). It is easy to see that if CLEAN minimizes the objective to zero, it has found a solution to the original deconvolution problem in a noiseless environment.

$$\underset{x}{minimize} \; \|I_{dirty} - x \star PSF\|_2^2 \tag{1.2}$$

Since the original problem is ill-posed, the objective (1.2) may have several zero points. In practice, CLEAN is stopped before it reaches zero. The addition of noise can add spurious peaks in the dirty image. By stopping early, CLEAN regularizes the objective. It assumes only a limited number of point sources exist in the image. The larger the magnitude of the peak, the more likely it is to be a real point source.

In short, CLEAN does a greedy approximation of the deconvolution problem, and assumes the resulting image consists out of a few point sources. The question remains, how close the CLEAN approximation is to the true image? If the true image consists out of a few point sources, CLEAN produces a good approximation. Extended emissions however are harder for CLEAN to reproduce. The peak of extended sources is lower than that of point sources. It is harder for CLEAN to distinguish extended sources from noise.

The CLEAN regularization scheme is not ideal for extended sources. Ideally another way of regularization would be chosen for extended emissions, but the regularization is a fixed part of the CLEAN algorithm.

1.3 From CLEAN to Compressed Sensing

In the Compressed Sensing Framework, an approach is split into three separate parts:

- An objective with a data and regularization term.
- A prior P in which transforms the image into a sparse domain.
- An optimization algorithm that is suited for the objective.

To demonstrate the flexibility of the Compressed Sensing Framework, we convert CLEAN into a Compressed Sensing approach. First we add a regularization term to (1.2) and arrive at the new objective (1.3). The objective contains the original CLEAN data term and a new 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 expected noise in the problem. Note that the $\left\|Px\right\|_0$ acts as an indicator function.

$$\underset{x}{minimize} \; \left\|D_{dirty} - x \star PSF\right\|_2^2 \; + \; \lambda \left\|Px\right\|_0 \tag{1.3}$$

The Prior P transforms the image in a sparse domain. CLEAN assumes the x contains a few point sources. In Compressed Sensing terminology, it assumes x is sparse in image space. Since x is already an image, the Prior P in Compressed Sensing CLEAN is the identity matrix.

The last step is choosing a similar optimization algorithm: In every iteration, CLEAN searches the highest peak in the dirty image. Matching Pursuit is a greedy optimization algorithm. In every iteration it searches the step which minimizes (1.3) the most. This Compressed Sensing approach is similar to CLEAN, but the new objective has a unique global minimum even with the presence of noise. The tunable parameters of CLEAN are replaced by a single parameter λ, which can be estimated.

The strength of Compressed Sensing Framework is its flexibility: The CLEAN prior works well on point sources, but is not ideal for extended emissions. In this Framework, the prior P can be replaced without changing the objective or the optimization algorithm. This has led to increased interest in Compressed Sensing for wide Field of View imaging.

2 Inverse Problem for wide Field of View Imaging

So far the small Field of View inverse problem has been introduced where each antenna pair measures a Visibility of the sky brightness distribution. This leads to the small Field of View measurement equation (2.1). It is identical to the two dimensional Fourier Transform. In practice the Fast Fourier Transform (FFT) is used, since it scales with $n\,log(n)$ instead of n^2 pixels.

$$V(u,v) = \int\int x(l,m)e^{2\pi i(ux+vy)}dldm \tag{2.1}$$

For wide Field of View imaging, two effects break the two dimensional Fourier Transform relationship: Non-coplanar Baselines and the celestial sphere which lead to the measurement equation (2.2). Note that for small Field of View $1 - x^2 - y^2 \ll 1$, and (2.2) reduces to the 2d measurement equation (2.1).

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

Non-coplanar Baselines lead to a third component w for each Visibility. Figure 2 shows the the u v and w coordinate system. w is essentially the pointing direction of the instrument. The UV-Plane is the projection of the antennas on a plane perpendicular to the pointing direction. Which point in the UV-Plane get sampled and what w component it has depends on the pointing direction. If the instrument points straight up, the UV-Plane is a tangent to earth's surface, and the w term compensates for earth's surface curvature. If however the instrument points at the horizon, the projected UV-Plane gets squashed and w compensates for antennas which lie far behind the UV-Plane. In essence, w is a phase delay that corrects antenna positions in three dimensions. The wide Field of View measurement equation (2.2) would account for the w phase delay, but it breaks the the two dimensional Fourier relationship and the FFT cannot be used. The W-Projection [8] algorithm approximates the effect of the w term restores the two dimensional Fourier relationship.

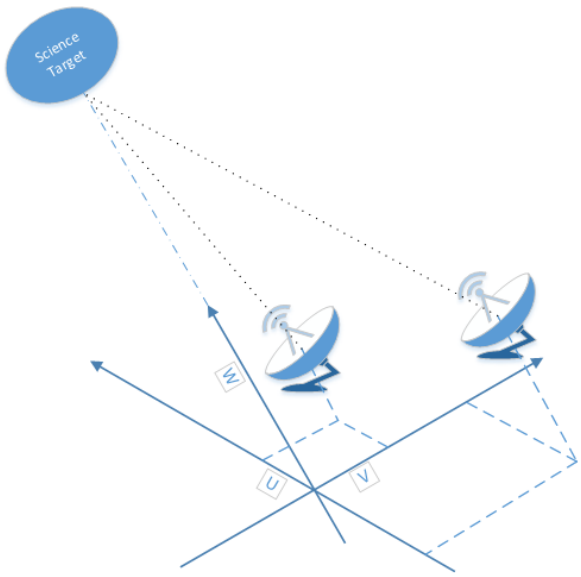


Figure 2: U V and W coordinate space

A-Projection [9]

2.1 Directionally Dependent Effects (DDE)

spread spectrum phenomenon

2.2 Calibration

2.2.1 Self-Calibration

3 Compressed Sensing for Radio Astronomy

The push to wide Field of View imaging has led to increased research into Compressed Sensing. The flexibility allows for a lot of freedom in design which can lead to different approaches for the same problem. So far, there are no 'best practices' for Astronomy: No prior, objective or optimization algorithm works strictly better than every other choice. Furthermore the choices for the individual parts influence each other. Compressed Sensing is flexible, but not every optimization algorithm works with every prior.

3.1 The Sparseland Prior and Incoherence: How Compressed Sensing works

For Compressed Sensing we need 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) which is at the core of Compressed Sensing: We assume for our signal x there exists a dictionary D. Each entry represents a signal part which can be present. D is potentially a large, but has a finite number entries. We assume that any x can only consist of a few signal parts of D. This means the coefficients for the signal parts in the dictionary α are all zero except for s entries for all valid x.

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

In image compression this phenomenon was can already be observed: image depicting nature scenes tend to be sparse in the wavelet domain. If x in (3.1) are nature scenes, we can create a Dictionary D of wavelets. A single image x can be represented with a few wavelets, meaning the number of non-zero entries s in α is far lower than the number of pixels n. All that is left to do for compression is save the non-zero entries of α. Two effects are of note: When x is noisy, or when x is not a nature image, the resulting α is not sparse. In Compressed Sensing this fact is exploited to reconstruct the true image from under-sampled measurements.

Back to the ill-posed inverse problem of interferometry: We measure the complex Visibilities of a band-limited signal. The Nyquist-Shannon rate states that if our band limited signal has at most frequency f, our sample frequency needs to be higher than $2f$. For n pixels, this is the case when we measure all n complex Visibilites (two samples per Visibility). If we have fewer samples, the Nyquist Shannon theorem states we cannot reconstruct the true image.

But what if we know our image is a Sparseland Signal and we happen to know the dictionary? Let us assume our image consists of $n = 20 * 20$ pixels and our dictionary of $m = 1000$ wavelets. Further assume at most $s = 10$ of the wavelets are non-zero for a given image. Could one just measure the 10 non-zero components of α and reconstruct the image? If we have prior knowledge about the location of the non-zero components, we would need 10 samples to reconstruct the image. Sadly, this is not the case in general. With the first sample we have about a $1/100$ chance of measuring a non-zero component. Note that if we measure a non-zero component, we learn $1/10$ of the information about the image. If we hit a zero component, we learn practically nothing. The question is, is there a way we can maximize our information gain of the non-zero components for each sample? In fact, there is: By having the measurement space be as incoherent as possible from the dictionary space, we maximize the information gained per sample.

[How many samples are needed]

Constructing an incoherent sampling space is surprisingly easy. Random projections are likely to produce a incoherent sampling space. Since we use an interferometer, we are bound to measure in Fourier Space. Depending on the prior, the Fourier Space might be coherent with the dictionary space and we do gain the maximum amount of information. As discussed in section 2, wide Field of View imaging breaks the two dimensional Fourier relationship. McEwen et al[10] showed that the wide Field of View measurement equation can help with incoherence, and demonstrated higher image reconstruction quality on simulated data.

The Sparseland prior is the basis of which Compressed Sensing builds. The dictionary can contain any function and is not limited to wavelets. It can even contain a mixture of for example wavelets, gaussians and cosine functions. Sparseland priors lend themselves to over-complete representations, where the number of dictionary entries is multiple times higher than the number of pixels($m \gg n$).

An appropriate prior for Radio Astronomy is still under research, currently Starlets[11] and Curvelets[**?**] are on top of the foodchain

3.2 Objective Function

The Compressed Sensing CLEAN objective function (1.3) uses the L0 norm for it's regularization term, which means the Objective Function is not convex. There are specialized solvers for the L0 compressed sensing. The L1 relaxation however is practically guaranteed to have the same minimum as the L0 norm and results in a convex objective function. Since GuRoBi works better on the L1 relaxation it was chosen for this project.

There are three different Compressed Sensing objectives: The analysis method, where the image x is minimized directly, the synthesis method where the sparse vector α is minimized, or by in-painting the missing Visibilities V_2.

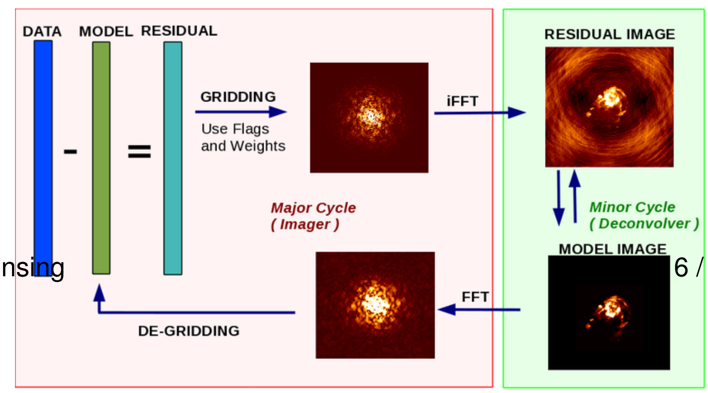
$$\begin{aligned}
analysis: \quad & \underset{x}{minimize}\ \|D_{dirty} - x \star PSF\|_2^2 \quad + \lambda\, \|Px\|_1 \\
synthesis: \quad & \underset{\alpha}{minimize}\ \|D_{dirty} - D\alpha \star PSF\|_2^2 + \lambda\, \|\alpha\|_1 \\
in-painting: \quad & \underset{V_2}{minimize}\ \left\|D_{dirty} - F^{-1}MV_2\right\|_2^2 \quad + \lambda\, \left\|PF^{-1}V_2\right\|_1
\end{aligned}$$

All three objective functions have the same global minimum. Retrieving x for the analysis objective is trivial, or the second and third objective x can be retrieved by $x = D\alpha$ and by $x = F^{-1}V_2$ respectively. [Empirical and theoretical studies have shown an advantage of the analysis objective over the other two [**?**]]. However, depending on the measurement space, prior and optimization algorithm, one objective may be easier to solve than others. The analysis objective is not useful when there is no transformation from x into the sparse space. This is the case for most over-complete priors: In that case, P is a $m * n$ matrix and $Px \neq \alpha$. The synthesis method just requires a transformation from the sparse space to image $D\alpha = x$. Similarly one might chose the in-painting method when the prior is a convolution in image space: Convolutions in image space are equivalent to a multiplication in Fourier Space.

$$\underset{x}{minimize}\ \|V - MF_{wFOV}x\|_2^2 + \lambda\, \|Px\|_1 \tag{3.2}$$

3.2.1 Implementation In Casa

Major minor cycle



3.3 Compressed Sensing Algorithms in Astronomy

3.3.1 SASIR

Prior

Objective

Optimizer

3.3.2 PURIFY

3.3.3 Vis-CS

4 Results

Physical plausible and shown to produce
better results on synthetic data [10]

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] 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] Hugh Garsden, JN Girard, Jean-Luc Starck, Stéphane Corbel, C Tasse, A Woiselle, JP McKean, Alexander S Van Amesfoort, J Anderson, IM Avruch, et al. Lofar sparse image reconstruction. Astronomy & astrophysics, 575:A90, 2015.

[8] 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.

[9] S Bhatnagar, TJ Cornwell, K Golap, and Juan M Uson. Correcting direction-dependent gains in the deconvolution of radio interferometric images. Astronomy & Astrophysics, 487(1):419–429, 2008.

[10] 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.

[11] Jean-Luc Starck, Fionn Murtagh, and Mario Bertero. Starlet transform in astronomical data processing. Handbook of Mathematical Methods in Imaging, pages 2053–2098, 2015.

[12] 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.

[13] 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, July 6, 2018

Jonas Schwammberger