Compressed Sensing Image Reconstruction for CASA

Jonas Schwammberger

Today

Abstract

Reconstruction from under-sampled measurements, Theory of Compressed Sensing tells us how it is done. Theory was applied on measurements from radio interferometers which naturally produce under-sampled measurements.

Contents

1 Reconstruction from under-sampled Measurements

In Signal Processing, continuous signals are represented with discrete samples. A digital recording of music, an image of a tree, or the measured velocity of a particle, all are discrete samples of continuous signals. The Nyquist-Shannon sampling rate tells us how many samples are needed to fully represent a signal: A signal which contains at most frequency f should be sampled with a frequency higher than $2f$. For example if we record a piece of music where the highest tone is at 20kHz, sampling rate should be more than 40kHz. Then, the music piece gets recorded above the Nyquist-Shannon sampling rate. It is fully sampled and there is exactly one continuous signal(with maximum frequency f) which fits the measurement.

The Nyquist-Shannon sampling rate is not always achievable: Samples may be expensive to acquire, corrupted by noise, or incomplete by the nature of the measurement instrument. In this case we are dealing with under-sampled measurements. Many possible signals fit the measurement and from the measurement alone, we cannot distinguish the true signal from all possibilities.

With the Theory of Compressed Sensing[1][2] however, we can use prior information about the signal and find the most likely candidate from all possibilities. Under the right conditions, the most likely candidate is guaranteed to be the true signal. With the Theory of Compressed Sensing, we exploit prior information to reconstruct the true signal from under-sampled measurements.

In this project, the Theory of Compressed Sensing was applied to an image reconstruction problem of Radio Astronomy. Interferometers produce under-sampled measurements of the sky which have to be reconstructed by an algorithm. A Compressed Sensing approach was developed and implemented in the Common Astronomy Software Application (CASA). The reconstruction quality was compared to standard reconstruction algorithm in astronomy on VLA data of Supernova Remnant G55.

1.1 Image Reconstruction for Radio Interferometers

Radio Interferometers use several antennas spaced apart from each other. Each antenna pair measures the different arrival time of a radio wave. For small field of view imaging, each antenna pair measures approximately two dimensional Fourier components of the sky (called Visibility in Astronomy). An interferometer with 26 antennas measures 325 Visibilities. The distance between the antenna pair(called a baseline) dictates which Visibility of the image gets observed. The longer the baseline, the higher frequency of the observed Visibility. The longest baseline of an interferometer is therefore a rough estimate of its maximum resolution.

For wide field of view imaging, the two dimensional Fourier relation breaks apart as additional effects dominate the measurement. This project used small field of view imaging, and the two dimensional Fourier component is a good approximation. In this document, a measured Visibilitiy equals a two dimensional Fourier component.

The interferometer only samples as many Visibilities as it has antenna pairs. The task is to reconstruct the observed image from an under-sampled Fourier space. In Astronomy, the CLEAN class of algorithms[3][4][5][6] were developed for this task. New interferometers like MeerKAT also observe new phenomenons which, in an ideal world, can be included in the CLEAN prior and improve its reconstruction. This is not possible in CLEAN, the prior is a fixed part of the algorithm. It cannot be modified as our prior knowledge of radio images increases.

Compressed Sensing can be thought of as a generalization of the CLEAN algorithm. The prior can be exchanged without changing the rest of the reconstruction algorithm. Furthermore Compressed Sensing reconstructions may be able to super-resolve the image[7].

1.2 Deconvolution of the Dirty Image

The ideal interferometer would fully sample the Fourier Space and introduce no instrumental effects and the observed image can be calculated with the inverse Fourier Transform. A real world interferometer corrupts the observed image with the image with under-sampling and other instrumental effects. The inverse Fourier Transform of the Visibilities can still be calculated, but it results in a corrupted ”dirty” image. The observed image is convolved with a Point Spread Function (PSF) image with a Point Spread Function(PSF). The task is to reconstruct the observed image from the dirty image and the PSF, or more formally we try to solve for x in equation (1.1), where only the PSF and I_{dirty} are known($\star$ is used as the convolution operation).

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

A deconvolution algorithm should find the observed image x. However, there may be many possible solutions to the equation (1.1). Noise N further complicates the deconvolution. The algorithm has to decide what is the most likely x given PSF and I_{dirty}. The PSF is specific to the interferometer. It models the sampling pattern in Fourier space. A deconvolution algorithm is not limited to a single interferometer and can easily be used on other instruments[1].

CASA produces a the dirty image and the PSF for the VLA instrument.

Note that the deconvolution problem is just one way of formulating the reconstruction. It was used in this project since the CASA interface was built for deconvolution algorithms. The same problem can be formulated as in-painting the missing Visibilities: If we apply the Fourier Transform to equation (1.1), then the convolution turns to a multiplication and we arrive at a similar equation $X * M + N = V$. The Fourier Transform of I_{dirty} is just the measured Visibilities V, PSF turns into a masking matrix M. M is one for all measured Visibilities and zero everywhere else. If an algorithm can in-paint the missing Visibilities X, it has also found a solution x to the deconvolution problem of (1.1). In section 3.2 the different formulations get discussed in more detail.

1.3 Deconvolution with CLEAN

CLEAN assumes the image consists of single pixel wide point sources. This is true when the image contains stars, the majority of which are too far away from earth to have any extension. CLEAN therefore takes a dirty image and a PSF as input and tries to find point sources in the image. In each iteration of CLEAN, it searches the highest peak of the dirty image and removes a fraction of the PSF. 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 tunable by the user.

CLEAN optimizes the objective (1.2) which is split in a data and a regularization term. The data term forces CLEAN to reconstruct close to the measurements, while the regularization term should account for noise. In each iteration, CLEAN searches the step which minimizes the objective the most, it uses a greedy optimization scheme. Due to the L0 ”norm”[2], the objective is non-convex and it may have local minima. Note that the L0 ”norm” acts as the sum of non-zero elements in the image.

$$\underset{x}{minimize} \ \|I_{dirty} - x \star PSF\|_2^2 + \ \|x\|_0 \tag{1.2}$$

[1] This is true for small field of view imaging. Wide field of view introduces additional effects that typically do not get modelled with a PSF

[2] The L0 ”norm” in this context is technically not a norm, hence the quotation marks. The L0 ”norm” is a common notation in Compressed Sensing literature, therefore it is used here.

In practice, CLEAN gets stopped before it reaches a local minima. A common stopping criteria is if the next peak in the dirty image is below 3 to 5 times the estimated noise level. In that case, CLEAN detects all peaks that are higher than 3 to 5 times the estimated noise.

CLEAN does a good approximation of the observed image, if it contains only point sources. Radio images may also contain extended emissions like hydrogen clouds, which are spread over several pixels. The maximum of extended emissions tends to be far lower than for point sources. Any pixel of the extended emission that is lower than 3 to 5 times the estimated noise gets ignored. If an extended emission is above the threshold, CLEAN approximates it with faint point sources. Instead of a hydrogen clouds, CLEAN may detect a cluster of faint stars.

To combat the problem, the reconstructed image of CLEAN gets convolved with the antenna beam-pattern. The beam-pattern is approximately a two dimensional Gaussian and represents the accuracy of the antennas. Clusters of faint stars get blurred into a more plausible cloud. This essentially lowers the resolution of the reconstructed image. The rationale is that any higher resolved structure is likely a reconstruction artefact. Over the years, the algorithm was extended, with the MS-MFS-CLEAN[6] as a recent example. All versions do blur the reconstructed image with the antenna beam-pattern.

The CLEAN prior works well for point sources, but it does not accurately reconstruct extended emissions. With the right prior super-resolved image reconstruction is possible. It was demonstrated in laboratory environment with the One Pixel Camera[8], or with the SASIR[7] reconstruction algorithm on the LOFAR radio interferometer. Since the prior of CLEAN is a fixed part of the algorithm it cannot be modified without creating a new algorithm.

1.4 CLEAN as Compressed Sensing Image Reconstruction

Compressed Sensing Image Reconstruction is a generalization of the CLEAN algorithm. A Compressed Sensing Reconstruction algorithm consists of three separate parts:

- An objective with a data and regularization term, and a parameter λ.
- A prior function $p()$.
- An optimization algorithm.

As the prior function $p()$, CLEAN uses the L0 ”norm”. The optimization algorithm is similar to Matching Pursuit. In each iteration, Matching Pursuit searches the element in the regularization term that minimizes the objective the most. The objective of CLEAN can also be generalized into (1.3). The objective still has a data and regularization term, but now the parameter λ represents the trade-off between reconstructing close to the measurement and reconstructing a plausible image.

$$\underset{x}{minimize} \ \|D_{dirty} - x \star PSF\|_2^2 \ + \ \lambda \, p(x) \tag{1.3}$$

CLEAN is similar to a Compressed Sensing Reconstruction consisting of the L0 ”norm”, matching pursuit and the objective (1.3). But now, individual parts of the Compressed Sensing Reconstruction can be replaced. For example instead of the L0 ”norm”, one could use the L2 norm, or apply the Haar Transform and use the L0 ”norm” on the Haar Wavelet coefficients of x. The prior function $p()$ can be chosen according to our prior knowledge and potentially super-resolve the image x.

Now we can minimize (1.3) with non-convex optimization techniques, we can analyse how calculate lower limits for the objective.

The guarantees of Compressed Sensing Reconstruction: Incoherent from the measurement space and sparse space is sparse.

Incoherence is easy. Interferometers measure in the Fourier space(This is an approximation for small field of view imaging. The approximation breaks apart in wide field of view). The image space is maximally incoherent from the Fourier space. Intuitively, A change in a single pixel will change all fourier components. A change in a single fourier component, changes all pixels.

maximize the information gained for each element in the sparse space.

The sparse space is here to distinguish true image from unlikely candidates. It models our prior knowledge.

Then, one can reconstruct the true image from undersampled measurements. How many measurements are needed? that depends on how sparse it is.

Taking again CLEAN as an example, if we know the image contains only one point source, we can locate it with only a few Visibilities. However if the image contains many point sources located closely together, we need more Visibilities.

The average case analysis is not trivial,

The prior and the optimization algorithm are disconnected and the prior $p()$ can be replaced for example with the L2 norm.

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 1 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 [9] algorithm approximates the effect of the w term restores the two dimensional Fourier relationship.

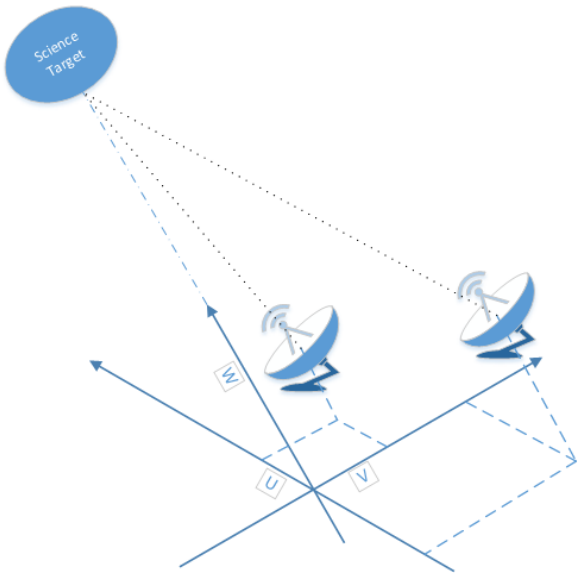


Figure 1: U V and W coordinate space

A-Projection [10]

2.1 Directionally Dependent Effects (DDE)

spread spectrum phenomenon

2.2 Calibration

2.2.1 Self-Calibration

3 Compressed Sensing Image Reconstruction

We want to reconstruct an image from under-sampled measurements. We can retrieve the true image if:

- we have prior information about the image
- measurement space and reconstruction space are incoherent

If the interferometer measures sky section which only contains point sources, the objective (1.3) with the L0 "norm" as prior $p()$ can reconstruct the true image from under-sampled Fourier components. The L0 "norm" is the sum of non-zero elements. The more elements of the image are zero, the lower the regularization term in the objective (1.3). This forces the reconstructed image to be sparse, it contains only few non zero elements, or in other words, it assumes the image contains few point sources.

Let us assume there are s point sources in the image. If we would know the number and location of the s point sources before, we could sample the s pixels to determine their magnitude. Sadly, we do not know the location before in general. The naive approach is therefore to sample at every pixel location. However, this is wasteful: Note that if we land at an empty pixel, the sample gives us almost no new information. If we would sample the whole image pixel for pixel, our average information gain per sample is low. In the under-sampled environment, we only have a limited number of samples available, we want to maximize the information gained for each sample. This is the case when the measurement space is incoherent from the reconstruction space.

Interferometers measure in the Fourier domain, which is maximally incoherent from the image domain. A change in a single pixel will modify all Visibilities, while a change in a single Visibility will modify all pixels. Intuitively, with each sample we now learn information about the s point sources. The average information gain per sample is maximized and we can reconstruct the even from under-sampled measurements. It works because we only need information about the s non-zero pixels.

The question remains, how many samples are needed to reconstruct the image? The answer to that question is an active field of research[11], but depends (among other factors) on the number of non-zero entries in the prior s. If the image also contains extended emissions, it increases the number of non-zero pixels s, which in turn increases the needed samples. However, if we can represent the image in a more sparse domain, we can reconstruct the image from fewer samples.

3.1 Sparseland Prior and over-complete Representations

An image containing both point sources and extended emissions is not sparse in the image domain. When using the L0 "norm" as the prior, we want to reconstruct in a domain where the image can be as sparsely represented as possible. From image compression, we know that natural images tend to be sparse in the Wavelet domain. We can use for example the Haar transform in the prior: $p(x) = \|Hx\|_0$ and potentially have a more sparse representation.

But why limit the prior to Wavelets? The prior function can be designed as we wish. Wavelets, sine functions or splines together can be used for the reconstruction. In practice many Compressed Sensing applications use a "sparseland" prior (3.1), a matrix D where each column contains a function potential part of the image. The dictionary matrix D is potentially a large, but has a finite number entries. Any image x we measure consists only of a few entries 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.

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

We can decide with what we fill the columns. Since D can be a mixture of functions and we know point sources are sparse in the image domain, we can use the identity matrix (also sometimes called Dirac basis), which represents single pixel values, and further add columns consisting of wavelets. The number of columns m can be much larger than the number of pixels n, which lends itself to over-complete representations.

An over-complete dictionary has more columns than rows (more entries in the dictionary than pixels) and can be used represent an image with even fewer non-zero entries s. Starlets[12] and Curvelets[13] have been developed as over-complete representations for astronomy.

In theory, there is no limitation on how many columns we add. In practice the computational power limits the number of columns.

Computational limits

Gurobi[14] was used.constrained optimization.

The prior $p()$ can be any function like the LP norm.

We could use L2, but what we mostly want in Compressed Sensing Reconstruction is sparsity. With that, L0 ”norm” is often used. 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.

3.2 Choosing the Objective Function

In general, there are three different reconstruction 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.

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

All three objective functions have the same global minimum. Retrieving x for the analysis objective is trivial. For 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 and prior, an objective might become more practical.

The analysis and in-painting objective require the inverse of the dictionary D^{-1}. It exists for orthogonal transformation like the Haar Wavelet transform, but may not be defined for over-complete dictionaries. In that case the synthesis objective gets selected. Similarly, the in-painting method is useful when the prior is defined as a convolution with the image, since it can be represented as a multiplication in the Fourier domain.During this project, no reconstruction algorithm was found which uses the in-painting method.

Data term can also be in the Visibility domain

3.3 Compressed Sensing Reconstruction Algorithms in Astronomy

multiple

3.3.1 PURIFY

Prior: Mixture of Dirac functions and Daubechies Wavelet (DB1 - DB8)

Objective: analysis

Optimizer: SDMM

Dirac is a fancy way of saying ”it is sparse in pixel space”

3.3.2 Vis-CS

Prior: dictionary of gaussians

Objective: Synthesis

Optimizer: Coordinate descent

3.3.3 SASIR

Was chosen because it has an inverse. Multiscale effects included in prior.

Prior: Starlets

Multi scale prior, over complete representation but with a transformation from image space in starlet space.

Objective: synthesis

Optimizer: FISTA

3.4 Implementation In CASA

The process of reconstructing an image in CASA is split in two separate cycles: The major and the minor cycle. The major cycle transforms the Visibilities to image space and back using the Fourier Transform. The minor cycle is the deconvolution algorithm, which tries to find the true image from a dirty image and a PSF.

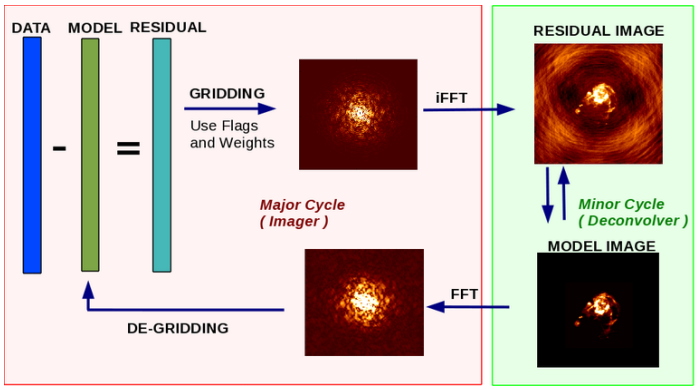


Figure 2: Casa Major Minor Cycle. Source [15]

The first major cycle iteration creates the PSF and the dirty image. Then, several minor cycle deconvolve the dirty image. The major cycle then continues and transforms the deconvolved image back to Visibilities. The major cycle ends by calculating the residual Visibilities from the measurements. The next major cycle continues by transforming the residual Visibilities. At the end of several major cycles (and with many, many minor cycle iterations) the model column should contain an approximation of the true visibilities while the residuals should be noise.

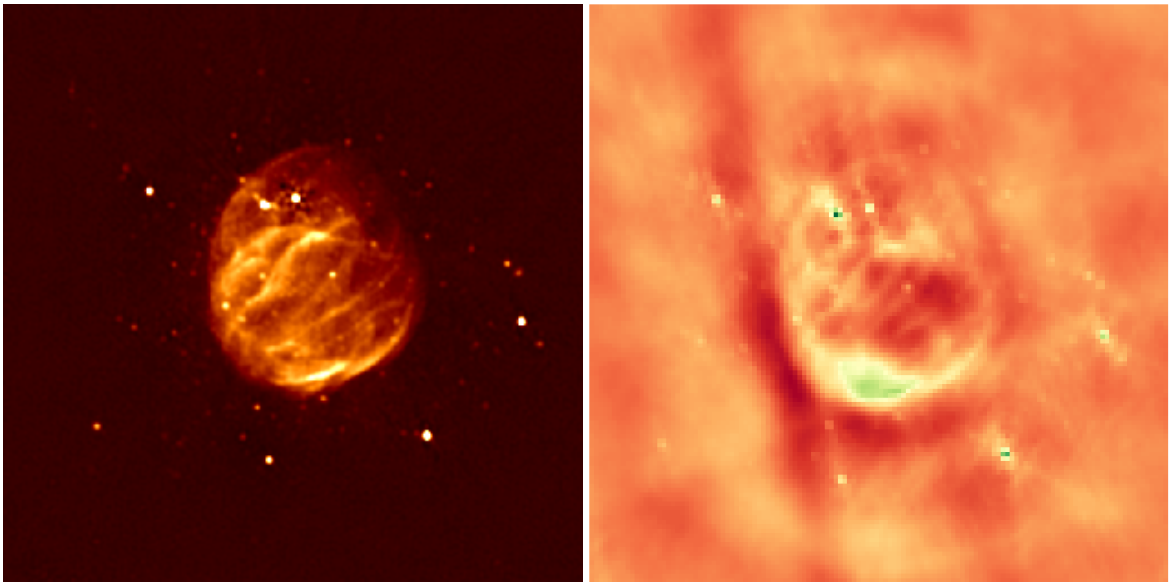
The major and minor cycle separation is built for CLEAN.

In CASA the major cycle is fixed. It was evaluated if it can be modified, but a modification was too time consuming in the context of the project. However CASA allows for the addition of new deconvolution algorithms.

Major cycle is more expensive to compute than a CLEAN minor cylce.

CLEAN needs potentially many major cycle iterations. A Compressed Sensing Reconstruction would converge to the optimum in one major cylce. Here lies a potential speedup for the Compressed Sensing Reconstruction.

4 Reconstruction of Supernova Remnant G55



(a) Reconstruction by NRAO. Source:[16] *(b) Dirty Image*

Figure 3: SNR G55 source observed by VLA.

The supernova remnant (SNR) G55 was observed by VLA. 10 seconds of the 8 hour observation is publicly available through the CASA imaging tutorial[17]. 3b is the dirty image calculated from the 10 second observation. The full 8 hours are not readily available. The image 3a is a reconstruction from an unknown VLA observation. The deconvolution algorithm is also unknown. For this project, the reconstructed image is assumed to show the true image of the sky.

3 shows G55 to be a slightly ”egg shaped” extended emission with six strong point sources. Several fainter point sources are inside and around the egg shaped extended emission. The dirty image 3b shows a corrupted version of G55. The six strong point sources are clearly visible as are the brighter parts of the extended emission. The dirty image also shows a negative ”trench” striking through the image as well as brighter regions around the remnant.

The size of the images 3 is about twice the size of the primary beam (the primary beam is approximately the size of the extended emission). In the real world, wide field imaging would be used. In this project, small field of view imaging was used because it is quicker to compute. It limits the dynamic range of the dirty image, the whole task gets harder.

The CLEAN algorithm gets compared to Compressed Sensing Reconstructions. The parameters of CLEAN were taken from the CASA imaging tutorial[17]. The reconstructed images of Compressed Sensing are constrained to have no negative pixels. Negative pixels are not physically plausible and was shown to improve Compressed Sensing reconstructions for synthetic data[18]. In total six different priors were tested with the analysis objective:

1. No Regularization
2. L1
3. L2
4. L1+L2
5. Total Variation

6. Starlet Transform

The regularization parameter λ needs to be estimated for each prior. The Miller[19] λ estimation was used and is shown in equation (4.1). [Estimation of the noise level e, divided and E]. An approximate solution is needed to define the noise level and the amount of regularization. In this project, the result with no regularization was used for the λ estimation.

This is not an ideal estimation, the image effectively gets reconstructed twice. Other Compressed Sensing Reconstructions approximate x of equation (4.1) by running their optimizer a couple of iterations without regularization, which reduces the computational costs.

$$\lambda \approx e/E \qquad \|I_{dirty} - x \star PSF\|_2^2 \leq e \qquad p(x) \leq E \tag{4.1}$$

Two figures compare the Dirty Image, CLEAN and the Compressed Sensing Reconstrutions. Figure 4 shows the reconstructed images on the same intensity scale. Figure 5 shows the flux profile of a cut through the reconstructed images. The Cut through the remnant is located in the center of the profile image.

CLEAN: CLEAN detects the brightest point sources, but finds only part of the extended emission. The top half of the ”egg” emission is missing. The structures of the remnant are blurred compared to the Compressed Sensing reconstructions. With the parameters of the imaging tutorial, CLEAN models the trench as a region with negative emissions. This is not physically plausible, although the behaviour can be changed with additional parameters. The profile 5 shows that CLEAN accurately reconstructs the flux of the large peaks, although the peaks are wide in comparison.

No Regularization: Detects the Egghead of the remnant and the non-negativity constraint keeps it from modelling the negative trench. It detects the smaller emissions around the remnant, but also detects ”fake” extended emissions. The profile 5 shows two smaller peaks in the center which also appear in 3a as do the two small point sources at the edge of the remnant. However, the dirty image has a rise in flux at the borders of the image, which is likely an artefact of the measurement, considering it does not exist in the NRAO reconstruction 3a.

L1: There is almost no visible difference between no regularization and L1. This is possibly an interaction with the Miller λ estimation, since the result of no regularization was used to estimate the λ of L1. The L1 regularization removes part of the ”fake” extended emission, particularly in the top region, but also a few structures in the center. The peaks in the profile 5 of L1 and no regularization are narrower than CLEAN, although they do not reach the same peak flux. L1 is prone to produce unlikely extended emissions: L1 also tries to approximate extended emissions with a number of faint point sources. This can introduce artefacts like pixel wide holes in extended emissions.

L2: Forces the extended emissions to be more smooth. It also considerably lowers the flux of bright point sources. The profile 5 shows L2 forces the bright peak to widen and lower. The peak is almost as wide as the CLEAN reconstruction. In the remnant center, it smoothers the smaller peaks.

L1+L2: Since L1 does a good job with point sources, but needs to be more continuous for extended emissions, why does one not combine both regularizations? The flexibility of Compressed Sensing Reconstructions allows for it. Sadly, the result is indistinguishable from the L1 regularization. In the dirty image, all pixels are very close to zero (Maximum: 0.0076 in Dirty Image). If the L1 and L2 regularization receive the same λ, the L1 term dominates. If the dirty image would contain pixels much larger than 1, then the L2 term would dominate.

Total Variation: A simple prior that has its origins in image de-noising. The objective is to reduce the gradients over the image. With an infinite λ, the Total Variation forces all pixels of the reconstruction to have the same value. The idea of the prior is to work well for both extended emissions and point sources. It has trouble with

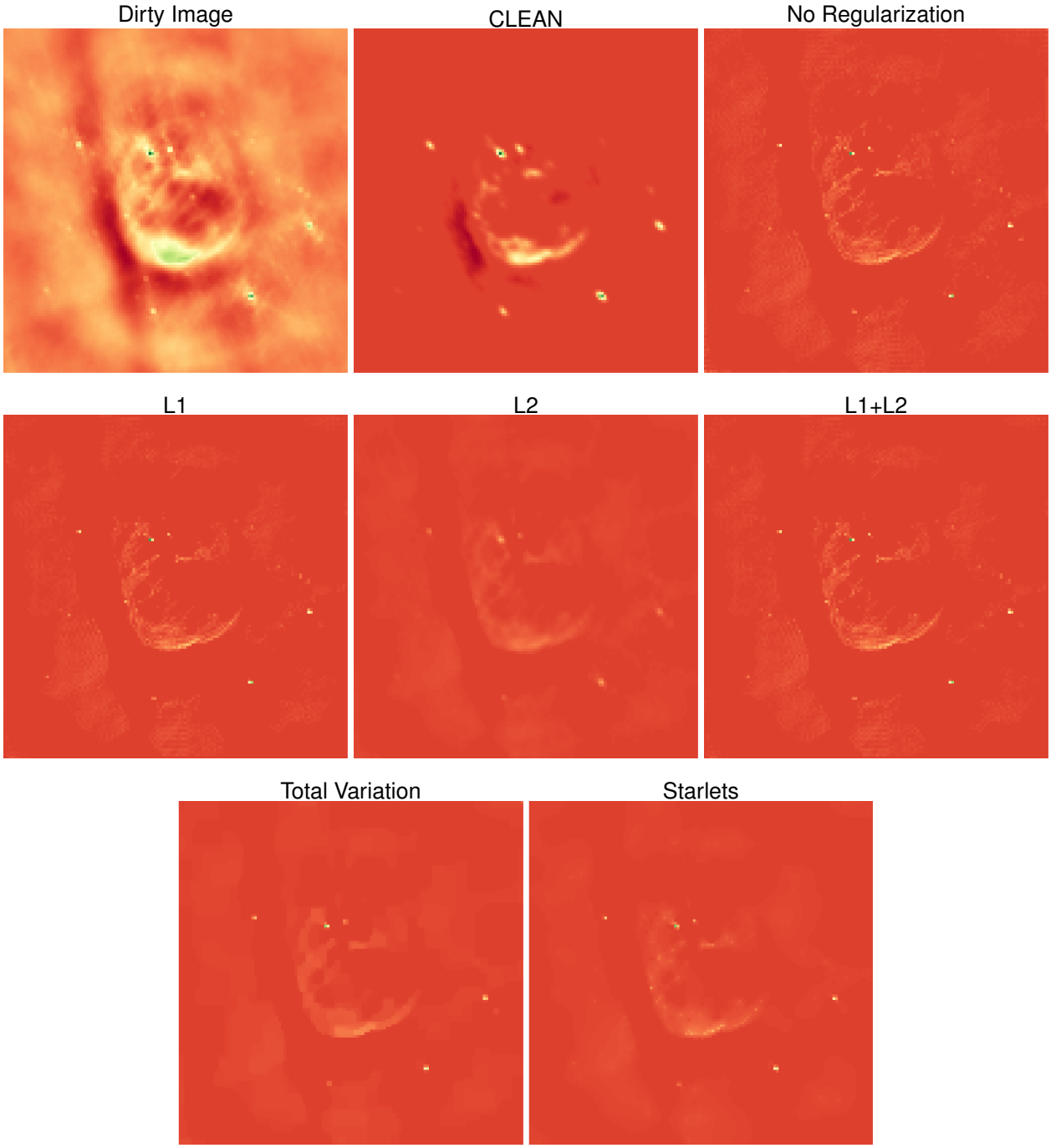
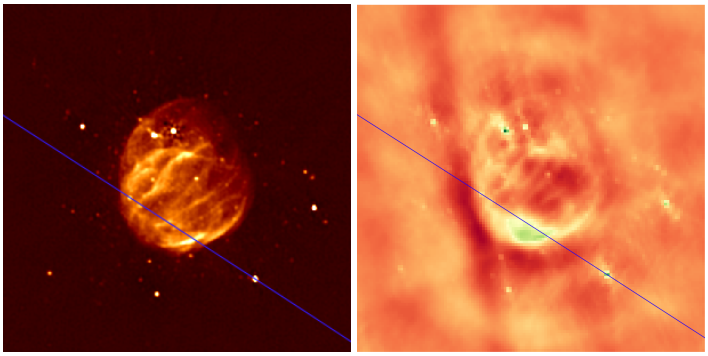
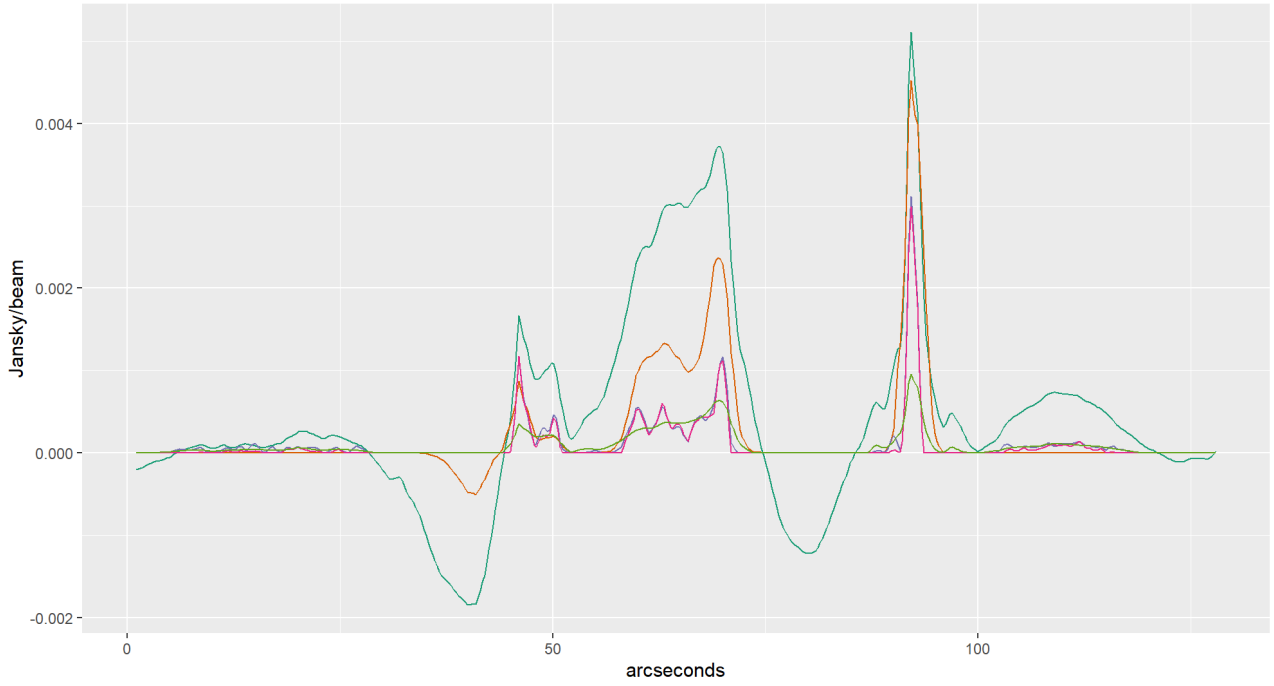


Figure 4: Reconstructed images of CLEAN and the different Compressed Sensing priors.

point sources inside extended emissions. In the profile, it is clearly visible how Total Variations cuts the peaks inside the remnant.

Starlets: Is a more sophisticated prior which also tries to model both point sources and extended emissions. It locates the point sources accurately. In the profile 5 it also finds the faint point sources in the extended emission. However, it also smoothed out the structure inside the remnant. For the LOFAR instrument, the starlet regularization was able to find smaller structures than the antenna beam-width[7]. The smallest starlet has a $5 * 5$ pixels dimension. For this reconstruction, the antenna beam-width is about two pixels wide. The resolution might be too coarse for the starlet regularization.





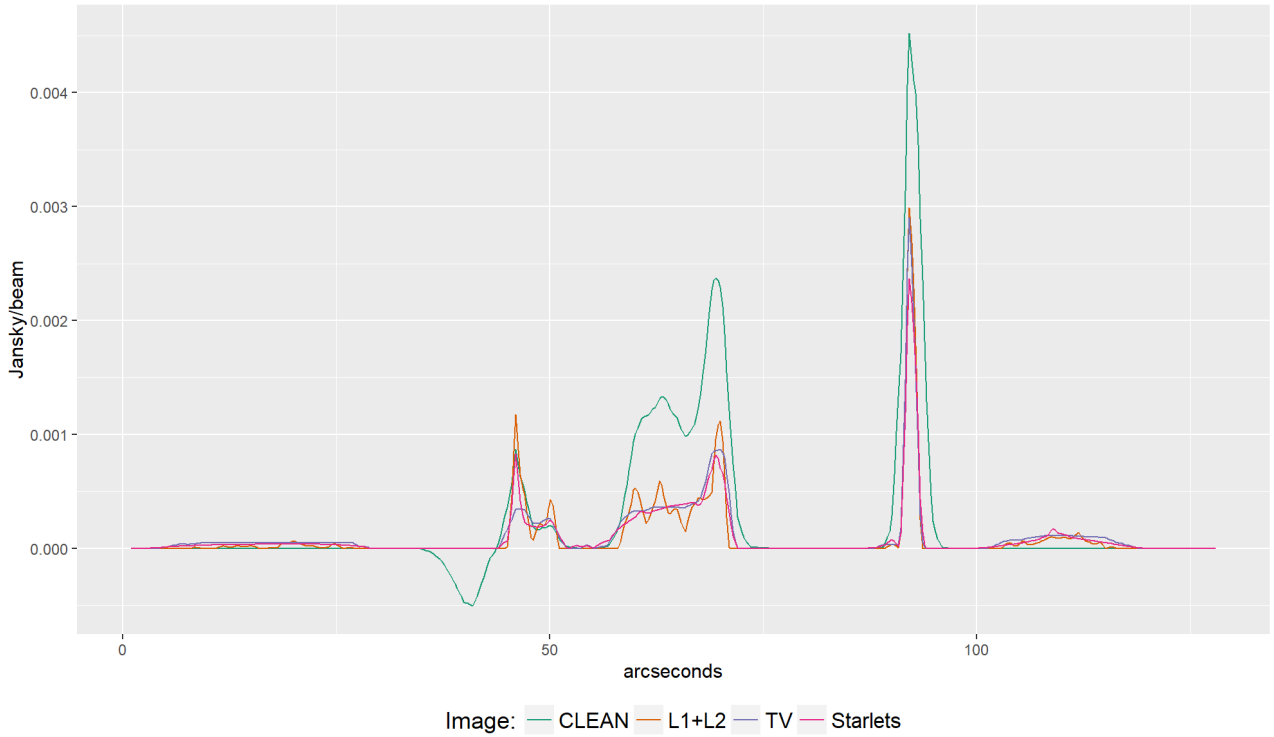


Figure 5: Intensity profile of CLEAN and the different Compressed Sensing Regularizations.

CLEAN produces the most accurate flux for strong point sources, even though in other reports Total Variation and Starlets reconstructed comparable peak flux to CLEAN [20][18].This is due to the implementation in CASA: The PSF CASA produces does not sum up to one. Convolving an image with the PSF increases the total flux, and naturally a deconvolution decreases the flux. CLEAN is the only algorithm that comes close to the measured flux, because in the last iteration, the reconstructed image gets convolved with the beam pattern (usually a 2d gaussian function), which is also not normalized. If we also convolve the Compressed Sensing reconstructions with the beam pattern (and smear away details), the fluxes become similar.

In this example, the L1 normalization was able to find smaller, plausible structures than CLEAN inside the remnant. Outside the remnant, all Compressed Sensing Reconstructions found ”fake” extended emissions. The L2 and starlet were also expected to find smaller structures. One possible explanation is that the antenna beam-width is about two pixels wide. Any structure smaller than the beam-width is one pixel wide.

For higher resolutions it is expected that L1 introduces more artefacts. The current implementation cannot increase the resolution since it needs a quadratic amount of memory per pixel: The image size of about 128^2 was the maximum which was feasible to solve on a laptop. The convolution operation $x \star PSF$ gets modelled as a matrix-vector product. The matrix therefore has 128^4 entries.

5 Future in Compressed Sensing Reconstruction

The flexibility of Compressed Sensing Reconstructions allows us to change the prior according to our previous knowledge, all while keeping the same objective and optimization algorithm. The prior can be updated as our knowledge changes. It was able to find plausible structures smaller than the antenna beam-width. Using Compressed Sensing Reconstructions gives us theoretical guarantees that, if our prior models the observation well enough, we reconstruct the true image from under-sampled measurements.

A proof of Compressed Sensing Reconstruction was implemented in CASA as a minor cycle deconvolver. It takes the dirty image and point spread function as input and calculates the optimal deconvolution according to the objective. The current implementation has a quadratic memory requirement. It does not scale to any practical image size for VLA observations. New interferometers like MeerKAT will require even larger images. The current implementation is not suited for large scale reconstructions.

Self-calibration is a task which was not covered in this project. Self-calibration aims to reconstruct both the image and the the antenna gain calibration from the same measurements. Here with Compressed Sensing Reconstructions one may be able to apply complex optimization techniques to self-calibration, and find the global optimum of both calibration and reconstructed image.

New interferometers are built with wide field of view imaging in mind, which introduces new measurement effects that a reconstruction algorithm should account for. State of the art Compressed Sensing Reconstructions take the measurements as input and also correct wide field of view effects. Due to the CASA interface, the current implementation is restricted to be in the image space, the effects of wide field of view are currently handled by CASA. The next step is a Compressed Sensing Reconstruction which accounts for the effects of wide field imaging and scales to astronomical image sizes.

References

[1] 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.

[2] David L Donoho. Compressed sensing. IEEE Transactions on information theory, 52(4):1289–1306, 2006.

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

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

[5] 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.

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

[7] Julien N Girard, Hugh Garsden, Jean Luc Starck, Stéphane Corbel, Arnaud Woiselle, Cyril Tasse, John P McKean, and Jérôme Bobin. Sparse representations and convex optimization as tools for lofar radio interferometric imaging. Journal of Instrumentation, 10(08):C08013, 2015.

[8] ams.org. Compressed sensing makes every pixel count, 2007.

[9] 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.

[10] 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.

[11] Emmanuel J Candes and Yaniv Plan. A probabilistic and ripless theory of compressed sensing. IEEE transactions on information theory, 57(11):7235–7254, 2011.

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

[13] Jean-Luc Starck, David L Donoho, and Emmanuel J Candès. Astronomical image representation by the curvelet transform. Astronomy & Astrophysics, 398(2):785–800, 2003.

[14] Gurobi Optimization. Gurobi optimizer, 2018.

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

[16] National Radio Astronomy Observations. Glowing bubble of an exploded star, 2016.

[17] National Radio Astronomy Observations. Vla casa imaging-casa5.0.0, 2017.

[18] 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.

[19] Keith Miller. Least squares methods for ill-posed problems with a prescribed bound. SIAM Journal on Mathematical Analysis, 1(1):52–74, 1970.

[20] 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.

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, August 1, 2018

Jonas Schwammberger