P7 Compressed Sensing Image Reconstruction for CASA

Jonas Schwammberger

January 9, 2019

Abstract

abstract plcaceholder

Contents

1 Compressed Sensing Image Reconstruction for MeerKAT

An instrument in the real world measures noisy data. Measurements are corrupted by noise, interference sources or the measurement instrument itself. Image reconstruction problems appear when one tries to find the observed image of the instrument, removing the corruption from the data. This forms an ill-posed inverse problem: A small change in the measurements may create very different reconstructions, and many possible images match the measurements. An image reconstruction algorithm therefore has to find the observed image from a potentially large set of possible images.

In the past, image reconstructions applied simple heuristics and approximated a likely image. How close the approximation was to the observed image was in general not known. The theory of compressed sensing[3][4] introduced a new theoretical framework under which image reconstructions can be analysed. This has lead to new algorithms which apply the theory of compressed sensing. Under the right conditions, the compressed sensing reconstructions are guaranteed to find the observed image. Furthermore they have the potential to super-resolve the observed image, creating a reconstruction above the accuracy limit of the instrument.

Image reconstruction problems appear in different fields, from MRI scans in medicine to X-Ray reconstructions in Astronomy.

CLEAN, and its own architecture. Lots of approximations.

Compressed Sensing reconstructions were able to super-resolve images, but were more complex to solve

A different architecture if CS reconstructions could scale better. In this project, a proof of concept reconstruction algorithm was developed. It gets rid of the approximations of CLEAN. The simple structure lends itself to distribution. However, the overall complexity of the new algorithm mainly depends on the number of Visibilities. For MeerKAT reconstructions, the Visibilities is the largest number. Even for the theoretical ideal algorithm, it cannot improve on the current Compressed Sensing Reconstructions in terms of complexity.

In the real world, measurements are corrupted by noise. Noise is introduced by the measurement instrument itself, but additional interference sources may be present. For example in Audio Recording, the microphone measures a noisy signal from the real world tone, and a passing car introduces additional interference. In a controlled environment like a recording studio, sound proofing handles the interference and only the noisy measurements have to be dealt with. Controlled environments are not always possible, that's why in Signal Processing the fields of de-noising and reconstruction exists. With those, we want to reconstruct the truly observed signal without the noise or external interference.

Image de-noising and reconstruction problems appear in different fields. In Astronomy, Radio Interferometers pose a challenging image reconstruction problem. The Interferometer measures a noisy image of the sky, corrupted by various sources like the ionosphere, passing satellites and the Interferometer itself. In the past, reconstruction algorithms used simple approximations to correct the image. For new interferometers like MeerKAT the approximations of the past do not hold. Furthermore, the larger MeerKAT instrument poses the reconstruction problem on a new data volume.

CLEAN algorithm[1][2]

Reconstruction algorithms for MeerKAT should handle the more complex effects of ionosphere and the like, while scaling even more on larger problem instances. Reconstruction algorithms using the theory of compressed sensing[3][4] showed higher reconstruction accuracy than state of the art algorithms. However, they currently do require more computing hardware. This project searches for a way to make CS Algorithm scale on MeerKAT data size.

SASIR[5] or more recently reconstructions using SARA [6] [7]

A different reconstruction algorithm

1.1 The basic Measurement Equation of a radio interferometer

Real world radio interferometers have complicated measurement equations. They become even more complicated for large interferometers like MeerKAT. These problems get addressed in section 2. This section looks at the basic measurement equation of a radio interferometer (1.1) and discusses the two fundamental challenges for image reconstruction.

$$V(u,v) = \int\int I(x,y)e^{2\pi i(ux+vy)}\,dx\,dy \tag{1.1}$$

An interferometer measures Fourier Components V (called Visibilities in Radio Astronomy) from the sky image I at position x and y. The term $e^{2\pi i(ux+vy)}$ represents the two dimensional Fourier Transform. The task is to reconstruct the observed image I from the measured Visibilities V. In theory this task is trivial: Since the inverse Fourier Transform exists, we can reconstruct the image I by calculating the inverse Fourier Transform of V. However, two properties of the Visibilities make this task challenging in practice:

1. Non-uniform sampling pattern in Visibility space
2. Incomplete Visibility coverage.

Property 1: We want to reconstruct an image with uniformly spaced pixels. The instrument defines the sampling pattern in Visibility space and does not correspond to the exact pixels of the reconstructed image. This property keeps us from using the Fast Fourier Transform. The naive inverse Fourier Transform can still be calculated, but it has a quadratic runtime and does not scale to the data volume of interferometers. Current reconstruction algorithms use the non-uniform Fast Fourier Transform (nuFFT). The nuFFT approximates the non-uniform Fourier Transform.

Property 2: Interferometers sample only a limited set of Visibilities. It does not have the whole information for reconstruction. This has the effect that the instrument introduces fake structures into the image. With only knowing the incomplete set of Visibilities, a reconstruction algorithm has to decide which image structures were truly measured, and which are due to the instrument. This forms an ill-posed inverse problem. There are many images that fit the measurements, and a small change in the Visibilities can lead to a very different reconstruction.

The CLEAN algorithms approximate the observed image with a deconvolution: The inverse Fourier Transform produces a corrupted image. The observed image was convolved with a known Point Spread Function (PSF), which represents the instrument. Finding a deconvolution reconstructs the observed image. The deconvolution is still an ill-posed problem, there are potentially many possible deconvolutions, and a small change in the input can lead to a very different output. Furthermore the CLEAN algorithms produce a greedy approximation of the deconvolution.

A CLEAN image reconstruction uses two different approximation algorithms: The nuFFT, which approximates the inverse Fourier Transform, and the CLEAN deconvolution, which approximately removes instrumental effects. In real world reconstructions, these two approximations are used in the major cycle architecture to further increase its accuracies.

1.2 The Major Cycle Architecture

Major cycle was created with CLEAN in mind, but the CS approaches use a similar architecture

An image reconstruction for radio interferometers consists of two different steps: A nuFFT, which approximates the inverse Fourier Transform efficiently and an image constraint/deconvolution algorithm, which approximates the instrumental effects on the image (for example CLEAN). These two approximations are an error source of the reconstruction. The major cycle architecture 1 therefore tries to iteratively minimize the errors of the nuFFT and the deconvolution.

The first major cycle uses the non-uniform inverse FFT to approximate the 'dirty' image from the Visibilities. The image constraint/Deconvolution algorithm decides parts which parts belong to the observed image and which are due to instrumental effects and other noise sources. It then returns the noise part, which get transformed back into residual Visibilities. The next major cycle iteration continues with the residual Visibilities. The residual Visibilities get minimized until they contain only the instrumental effects and noise.

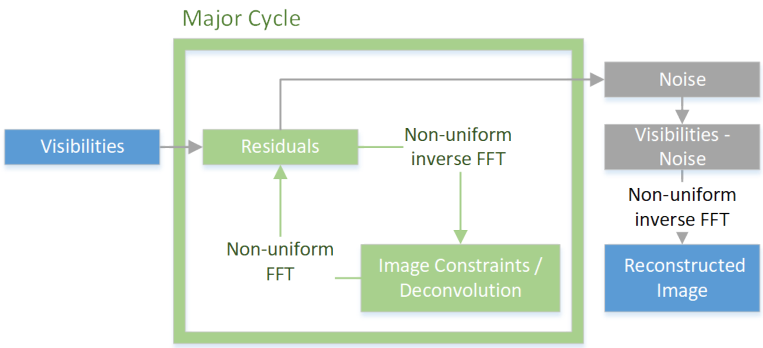


Figure 1: The Major Cycle Framework

After several major cycles the residual on a regularly spaced image which has a small error from non-uniform samples, and a small error from incomplete measurements.

A single major cycle is expensive. MeerKAT needs several cycles to have a good image result with a CLEAN algorithm.

Compressed Sensing reconstructions essentially use the same architecture. They do not use an explicit deconvolution anymore, but

Slightly different implementation in Compressed Sensing approaches, does not neatly fit in the figure 1, but still gains the same advantages and disadvantages as the standard major cycle.

Compressed Sensing algorithms use essentially the same architecture. But in turn uses more major cylces, on top of being more expensive to compute inside a cycle. Even though it was shown to produce better reconstructions on different datasets, the added runtime is a big reason what keeps it from wide-spread adoption.

The question thefore is, can the overall complexity of the comporessed sensing reconstruction be reduced, if a different architecture is used.

1.3 Compressed Sensing Reconstructions

Compressed sening reconstructions use the However, compressed Sensing algorithms come with the drawback of requiring more major cycles.

MeerKAT due to wide field of view introduces even more troubles

Current Compressed Sensing reconstructions reduce the number of major cycles. However, the question is if Compressed Sensing can use a different architecture, and scale better to problems of the size of MeerKAT.

Furthermore on the new MeerKAT instruments, we have a big data problem. We want to create a large image from a large amount of Visibilities. 32k*32k pixels and terabytes of raw Visibility data.

Scalability is a big problem.

There are ways to get rid of the major cycle, but overall the complexity could not be reduced.

2 Challenges for imaging MeerKAT data

New interferometers like MeerKAT put additional challenges on the image reconstruction problem. First, the new instrument produce a new magnitude of data, forcing reconstruction algorithms to be highly scalable and distributable. Second, the more sensitive instruments with large field of view amplify effects which were negligible in older instruments, like non-coplanar baselines or the ionosphere.

Compressed sensing has to be able to work with these issues somehow. Solutions all in the context of the major cycle, a new architecture may need to handle effects differently. Hopefully in a manner that is at least as efficient.

In this work, the effects of non-coplanar baselines gets handled in more detail. The effect has a neat mathematical notation, it adds a third fourier component to the measurement equation, but breaks the two dimensional fourier relationship introduced in section 1.1.

Calibration used to be a task before image reconstruction. Calibration gets not explicitly handled here, but it is important to keep in mind. For older interferometers, the data was calibrated before image reconstruction. With the advent of self-calibration, the image reconstruction is also used to improve calibration. MeerKAT requires more calibration parameters, so an image reconstruction algorithm has to be able to calibrate, or it is not interesting.

Further issues that do not get handled here

- (Beam Pattern, A Projection)
- Full polarization
- Wide band imaging

There are several challenger for imaging meerkat data. One problem is the new amount of data.

terabytes of measurements. Large image size 32k squared are the obvious problems to solve. Distributing the problem is not part of this work.

In this work, it is focused on Wide field of view issue.

2.1 Wide Field of View Imaging and the third Fourier Component

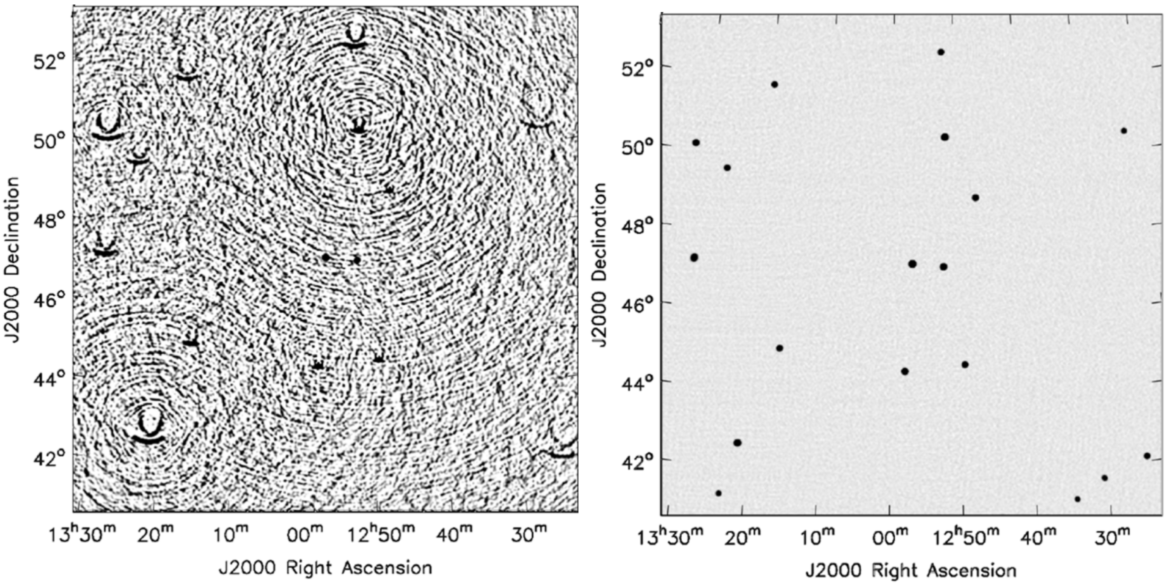
The measurement equation of radio interferometers contains a third fourier component, the w component. This leads to the following measurement equation (2.1).

$$V(u,v,w) = \int\int \frac{I(x,y)}{\sqrt{1-x^2-y^2}} e^{2\pi i [ux+vy+w(\sqrt{1-x^2-y^2}-1)]} \, dx \, dy \tag{2.1}$$

For small field of view observations, the term $\sqrt{1-x^2-y^2} \approx 1$, and we arrive at the basic measurement equation (1.1). For small field of view, the w term can be ignored, and the image can be calculated with the inverse two dimensional non-uniform FFT. For wide field of view measurements, the w term breaks the two dimensional Fourier relationship. The measurement equation (2.1) can still be inverted, but the resulting algorithm has a quadratic runtime and is not feasible in practice.

Note that the image $I(x,y)$ of the wide field of view the measurement equation (2.1) is still two dimensional, even though the Visibilities $V(u,v,w)$ have three. The relationship between Visibilities and image is neither the two, nor the three dimensional Fourier transform. The Visibilities represents an image on the celestial sphere. This means the image is naturally curved, and the w component represents the distortion introduced by the curvature.

One can simply ignore the w component, ignore the curvature and use the two dimensional Fourier Transform. The image then does not lie on a sphere, but on a tangent plane. At the tangent point (where $\sqrt{1-x^2-y^2} \approx 1$, typically the image center in radio astronomy) the two reconstructions are identical. But the distortion gets more severe away from the tangent point. The image 2a shows a tangent plane reconstruction over a large field of view, while 2b shows the same reconstruction but w corrected. Close to the tangent point, the image center, the reconstructions both locate the point sources. At the image edges, the distortion gets severe enough to decorrelate the Visibilities and the point sources get 'torn' apart.



(a) 2D Fourier Transform. *(b) Fourier Transform with w-correction.*

Figure 2: Celestial sphere distortion on simulated data. Source: [8]

Image reconstructions for MeerKAT need to account for the w component efficiently. In the Major Cycle framework, this is part of the non-uniform FFT. The first approach was to image facets, creating several tangent planes for the celestial sphere. A more efficient approach came in the form of the w-projection algorithm[8]. It uses convolution on the Visibilities during the non-uniform FFT. MeerKAT currently uses the w-stacking algorithm, which more efficient than w-projection on MeerKAT's data volume and is simple to distribute.

This is still an active field of research. The latest work as of time of writing is a hybrid approach[9] with both w-stacking and w-projection. As of time of writing, it was not yet adapted.

2.1.1 State of the art: w-stacking algorithm

Here a short explanation of the w-stacking algorithm is provided. It is a surprisingly simple algorithm that seems to be here to stay. It is a modification of the nuFFT operation displayed in figure 1.

Inverse nuFFT with w-stacking: Visibilities get grouped into stacks by their w-term. Every visibility inside a stack has a similar w-term. The more stacks, the more accurately you can handle the w component. The figure displays the relation

The nuFFT gets called on each stack separately, and a constant w correction factor applied. Since w-term is constant, we can use the two dimensional nuFFT again. At the end, all the stacks are summed together to form the dirty image.

nuFFT with w-stacking: The inverse is done in a similar fashion. The sum cannot be re-separated, so the image gets copied into each stack. The constant w factor gets divided out. The nuFFT gets applied.

the runtime

(talking about more recent hybrid approach, limiting the total number of ?w-stacks?)

2.2 Self-Calibration

Complex gain term. Corrects amplitude and phase.

Traditional Calibration

A calibration source close by, with a known brightness. Phase and amplitude calibration was done before image reconstruction. For older interferometers, there was a very limited number of calibration terms to solve.

but again effects of wide field of view also increase the number of necessary calibration terms. The advent of self calibration, in which the image reconstruction was used to solve for both, the observed image and the calibration of the instrument.

Self calibration with the major cycle algorithm and a CLEAN deconvolution.

Initial phase calibration Shallow clean phase calibration deep clean phase and amplitude calibration deep clean reconstruction

3 Eliminating the Major Cycle

There is little research in exploring a different architecture for Compressed Sensing. The closest to this work was done by Pratley et al[10], which optimizes the non-uniform FFT in the context of compressed sensing. The Major cycle is still used. On the upside, progress in w-term accuracy and runtime.

There are ways to form a different architecture. However, the advent of self-calibration pushes a forward-backward kind of architecture on the reconstruction.

3.1 Separating the Major Cycle into two Optimization Problems

The major cycle reduces two errors simultaneously. The error introduced by the non-uniform FFT approximation, and the error introduced by the deconvolution approximation. In a sense, the major cycle is a simple optimization algorithm: It iteratively calls the non-uniform FFT and deconvolution approximations with the error from the last iteration, until the error is below a certain threshold.

$$FT : \underset{I_{dirty}}{minimize} \left\|V - A^{-1} I_{dirty}\right\|_2^2 \tag{3.1}$$

$$deconvolution : \underset{x}{minimize} \left\|I_{dirty} - x \star PSF\right\|_2^2 + \lambda \left\|P(x)\right\|_1 \tag{3.2}$$

The major cycle can also be separated into two objectives (3.1) and (3.2), and solved in sequence. The first objective (3.1) tries to find the best uniformly spaced image I_{dirty} matching the observation. A^{-1} represents the non-uniform inverse FFT approximation ($A^{-1}I_{dirty} \approx F^{-1}I_{dirty}$). When this is solved, the original Visibility measurements are not needed any more. Then, the second objective (3.2) searches for the deconvolved image x regularized by some prior $P()$.

By separating the two objectives, the hope is we can use specialized optimization algorithms that overall converge faster than several major cycles. Also, after the first objective was minimized, we can drop the Visibilities, which can be magnitudes larger than the image.

At first glance, this is a sensible idea, but the devil is in the objective (3.2). Namely in the PSF is not constant, it varies over the image. In the Major Cycle Architecture, a constant PSF[1] is sufficient: It only needs to approximate the deconvolution. The error introduced by a constant PSF can be reduced in the next Major Cycle. By eliminating the Major Cycle, we also eliminated the chance to approximate the deconvolution with a constant PSF. In the worst case, we would need to estimate the PSF for each pixel in the image.

In the worst case, these many PSFs may stop us from using this approach. It is plausible that can use fewer PSFs and be close enough. No known work done in the radio astronomy. The Major Cycle is ubiquitous. But it is plausible.

However, the problem is that the imaging algorithm should also be able to calibrate. With this approach, there is a strict one way flow from measurements to image. For calibration, a backwards flow from image to measurements is necessary.

For calibrated data, this approach may be potentially faster. but since MeerKAT will be difficult to calibrate correctly, this is unlikely.

[1]CLEAN implementations in Radio Astronomy typically use a constant PSF. This is not strictly necessary, one can also implement CLEAN with a varying PSF.

3.2 Reconstruction on the Sphere

The signal naturally lives on the celestial sphere.

Not that much used in Radio Astronomy reconstructions. Work that uses to project on the sphere and looks at the reconstruction quality [11].

The wide field of view measurement equation (2.1) can be expanded into spherical harmonics.

[12] measurement equation with spherical harmonics. Essentially replaces the non-uniform FFT from the Major cycle with a new approach based on the fact that it has a solution to the Helmholtz equation. , see [12] for further detail. It can be used as a different way to arrive at a dirty image from wide field of view interferometers. It is useful in full sky image, because it fulfills the (property from paper) and the full sky dirty image does not have artifacts. It does not gain any complexity advantage, quite the contrary.

Even if the spherical harmonics transform would gain a speed advantage, used in the way of [12] it would still be the same major cycle architecture, just with spherical harmonics transform.

However there is the possibility to just do everything on the sphere.

Spherical haar wavelets used for simulating visibilities from a known image, essentially the opposite operation from imaging. [13]

based on the spherical harmonic transform. Doing a Spherical harmonic transform efficiently is still an active field of research [14]

transform . The wide field of view measurement equation (2.1) can be factorized into spherical harmonics. This means spherical harmonics are equivalent, calculating an image from the measurements with the three dimensional fourier transform is the same as using the spherical harmonics transform. [15]

[16]sampling theorem on a sphere

Spherical harmonics were researched in the context of Compressed Sensing image reconstructions (cite wiaux). They improved the quality of the image reconstruction, but no speed advantage.

The push on solving on the sphere directly. Using spherical haar wavelets, (cite) was able to speed up the simulation problem. Simulation tries to solve to problem of finding the measurements corresponding to a given image x. However, this has so far not been used for imaging.

3.3 Explicit Fourier Transform Matrix

Hardy et al[17] handled the whole inverse Fourier Transform as an explicit matrix F^{-1} of size $M * N$. They dropped the non-uniform FFT and w-stacking approximations. This is trivially to distribute later. The only problem is of Course that for MeerKAT data sizes $M * N$ is too large (4gb of visibilities times 1 Billion Pixels).

However there is potential for improvement: Only a limited number of pixels in a reconstruction are not zero. We spent a lot of processing power on pixels which do not matter in the reconstruction. If we could predict which pixels are not zero, we would only need to calculate a subset of F^{-1} columns to reconstruct the image. The starlet transform[18] has a way to predict its non-zero components. We represent the image as a combination of starlets, and estimate which starlets are likely non-zero from the Visibilities directly.

To my knowledge, such an algorithm has never been tried out.

A proof-of-concept algorithm was developed that shows the prediction actually works in practice. It uses Coordinate Descent with the starlet transform, and only needs to calculate a subset of the matrix F^{-1} for a reconstruction. The question remains how many columns of F^{-1} are needed, and is it more efficient than an algorithm using the major cycle architecture.

4 Compressed Sensing with Coordinate Descent

The idea behind compressed sensing is we use prior information about the image, and find the most likely image x from all possible solutions. In practice, we formulate a minimization problem with a data and regularization term. The data term forces the image to be as close to the measured Visibilities as possible, and the regularization term tells us how likely the image is to be true.

The regularization is responsible for an accurate reconstruction. In this work, two regularizations are used. The L1 pixel regularization is used to demonstrate a simple version of the coordinate descent algorithm. For the actual algorithm starlet regularization is used. Starlets have been used as regularization for the LOFAR interferometer[5], which has been shown to work.

First, let us look at the coordinate descent algorithm in its simplest form, and discuss its properties. we want to minimize the objective (4.1), where the data term $\left\|V - F^{-1}x\right\|_2^2$ forces the image to be close to the Visibilities V, and the regularization term $\|x\|_1$ forces the image to have as few non-zero pixels as possible. The parameter λ represents our trade-off between reconstruction accuracy and regularization.

$$\underset{x}{minimize} \ \left\|V - F^{-1}x\right\|_2^2 + \lambda \|x\|_1 \tag{4.1}$$

The objective function (4.1) has a property, which we can exploit with Coordinate Descent: The optimum for a single pixel, if we keep all others fixed, turns out to be a parabola. The optimum for a single pixel of (4.1) is simply finding the apex of a parabola, followed by a shrink operation[2]. Sadly, the pixels are not independent of each other, and we need multiple iterations over all pixels to converge. This leads to the following reconstruction algorithm, written in python code:

```python
def coordinate_descent(V_residual, x, lambda, max_iter):
  for k in range(0, max_iter):
    for i in pixels_row:
      for j in pixels_column:
        x_old = x[i, j]
        fourier_column = calculate_fourier_transform(i, j)
        fr = real(fourier_column)
        fi = imag(fourier_column)
        rr = real(V_residual)
        ri = imag(V_residual)

        #find apex
        a = sum(fr**2 + 2*fr*fi + fi**2)
        b = sum(fr*rr + fr*ri + fi*rr + fi*ri)
        x_new = b / a + x_old

        x_new = shrink(x_new, lambda)
        x[i, j] = x_new
        V_residual = V_residual - fourier_column * (x_new - x_old)
```

The simple version of coordinate descent converges to a solution which is similar to CLEAN. The convergence rate of coordinate descent is not well understood. Our image reconstruction problem falls in the class of quadratic programming, in coordinate descent converges at least linearly[19]. In practice, the convergence rate can be sped up with heuristics like active set[**?**] and λ[**?**]. Furthermore, coordinate descent is robust, even when the residual vector is only approximated, which is useful for distributing the algorithm.

[2] The shrink operation reduces the magnitude of the pixel by λ. For example: Pixel $x = -12.4$, $\lambda = 0.6$. The new pixel value after shrinkage follows as $shrink(-12.4, 0.6) = -11.9$

The simple approach calculates the whole Fourier Transform Matrix F^{-1} explicitly. The upside is that we do an exact w-term correction. The downside is F^{-1} is too large to keep in memory, and too expensive to calculate on the fly for the scale of MeerKAT reconstructions.

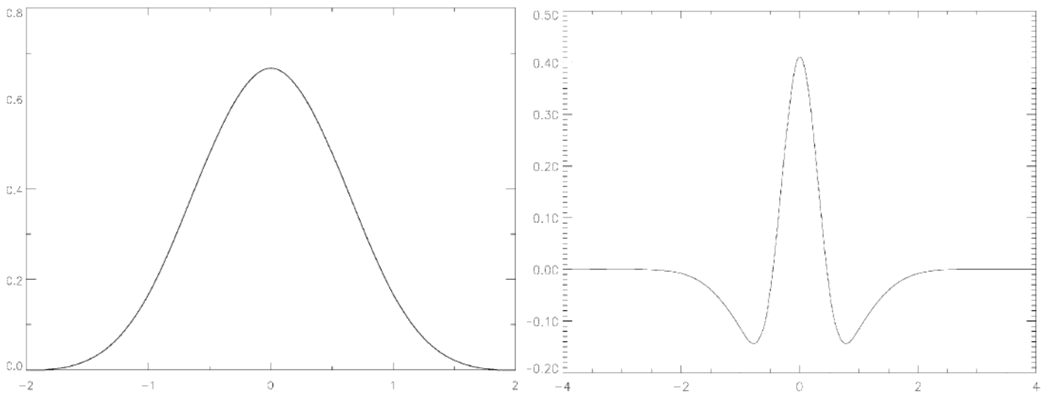
The starlet transform is used from here on. Instead of reconstructing the image directly, we represent the image as a combination of starlets $x = D\alpha$. Where D is an over-complete dictionary of starlets is the starlet component vector. Since the dictionary D is over-complete, it is a matrix with more columns than rows, and does generally not have an inverse $D^{-1}x \neq \alpha$. However, the starlet transform does have an pseudo-inverse. In this work, the pseudo inverse is used to only look at a subset of starlets, which is likely to be non-zero.

The new objective (4.2) tries to find th

$$\underset{\alpha}{minimize} \ \left\|V - F^{-1}D\alpha\right\|_2^2 + \lambda \left\|\alpha\right\|_1 \tag{4.2}$$

Starlet transform can be used to estimate the component vector parabola for each starlet.

4.1 Starlet Transform



(a) Scaling function *(b) Wavelet function*

Figure 3: Scaling and wvelet function of the starlet wavelet. Source: [18]

The starlet transform is based on the starlet wavelet, shown in figure 3.

Starlet is a multi-scale wavelet representation which were specifically developed for astronomy.

Over-complete representation. More starlets than there are pixels. Sparse representation, the number of non-zero starlets is smaller than the number of pixels

Starlets as a series of convolutions.

Forward transform, from image to starlets

From Starlets to image

4.2 Active set heuristic with Starlets

Even though the starlets are an over-complete dictionary, they have an approximate transform from image to starlet space. For Coordinate Descent, this can be used as a active set heuristic: We try to find the coefficients which are likely to be non-zero. This helps us so we do not need to calculate the whole matrix product $F^{-1}D$. We only use columns that are likely to be not zero.

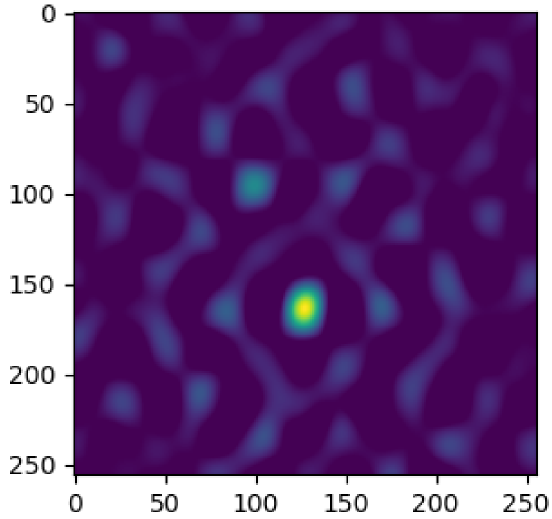


Figure 4: Starlet Level 0

The higher the number, the more likely this component is to be non-zero. It is essentially a probability distribution for which starlet components are non-zero.

Stupid approach with line search. Could be done more efficiently by using the histogram of the starlet level.

4.3 Implementation

$$\underset{\alpha}{minimize} \ \left\| V - F^{-1}D\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_1 \tag{4.3}$$

coordinate descent with active set heuristic

5 results and scalability of Coordinate Descent

MeerKAT simulation with Simkat64 Simulation suite. Not a lot of noise, fairly simple data to test the algorithm.

Two simulations, one of point sources and one of mixed sources.

Super resolution. The reconstruction should be able to locate the point sources below the accuracy of the instrument.

5.0.1 Point Sources

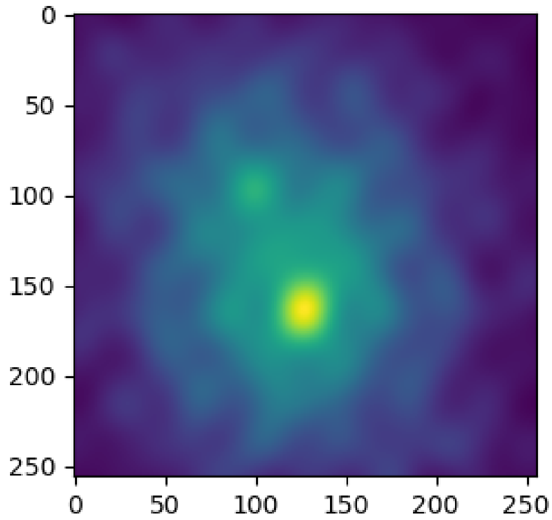


Figure 5: Dirty Image of two point sources

faint source is spread more widely. Not good, The paper [18] used different λ values for different starlet levels. Maybe with this it can be forced to do more precise results, but it was not tried in this work.

5.0.2 Gaussian and Point Sources

5.1 Scalability estimates with ideal heuristics

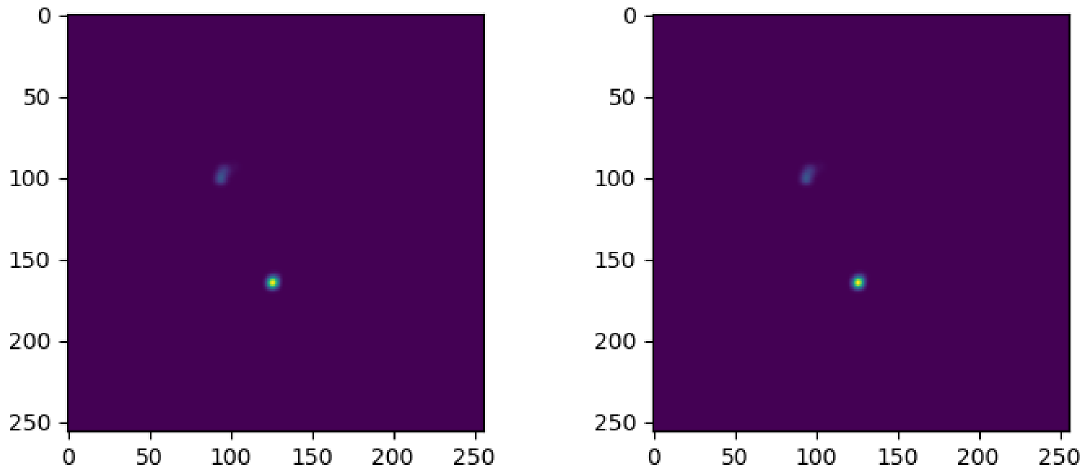
Bunch of heuristics. There are a lot of little ways to optimize this algorithm. The question is, is it worth going further and try to improve the algorithm further. So we want to estimate the lower bound of the algorithm.

Convergence is hard to determine in general. Following simplified assumptions were made: We have a heuristic with oracle performance. It returns only the locations of α in constant time. Furthermore, we assume all axes are independent from each other, only one descent per non-zero axis is necessary.

S

$res * starlet = M$ descent: $genfcol = 3 * M$ $a = 3 * M$ $b = 4 * M$ $residuals, fcol * diff = M$

$total_m it_g encol = 11M$



(a) Reconstruction after one full iteration *(b) Reconstruction after 4 Full iterations*

Figure 6

total CD = $S * 11M + 2M * starlets$

S depends on the image content directly. For example if the image contains 15 point sources and five Gaussian extended emissions, then S equals 20 non-zero components (if we assume the Gaussian sources require only one starlet for representation). Coordinate Descent therefore is independent of the image size N. It solely depends on the size of the measurements M, and the number of non-zero components in the dictionary S.

It does not use any approximation for the fourier transform.

(Cite New W-stacking approach) Nufft: $M + 2Nlog2N$ W-stacking = $M + W * (2Nlog2N + 2N) + NlogN$ Deconvolution = ??

The major cycle algorithm depends on more parameters. Assumptions were made in favor of Coordinate Descent.

the number of w-stacks W. If this is the case

[9] fast w-stacking w-projection

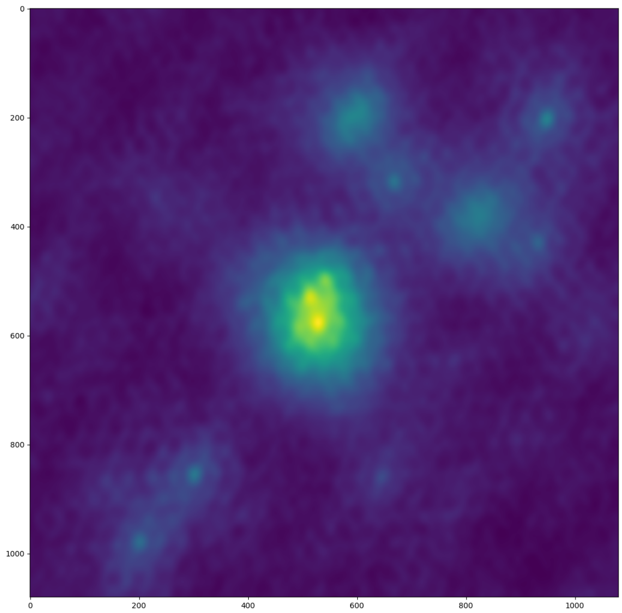
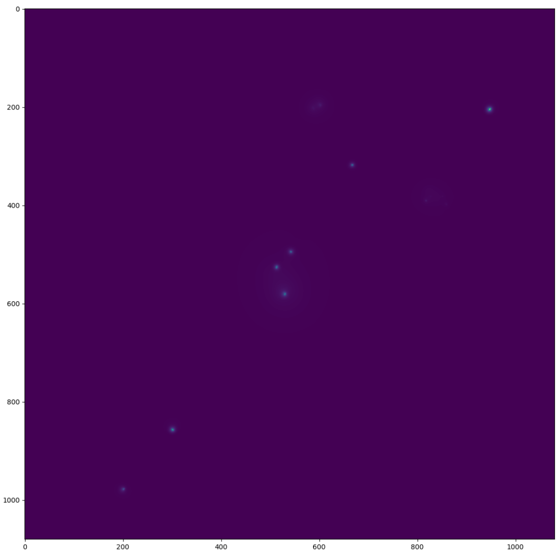
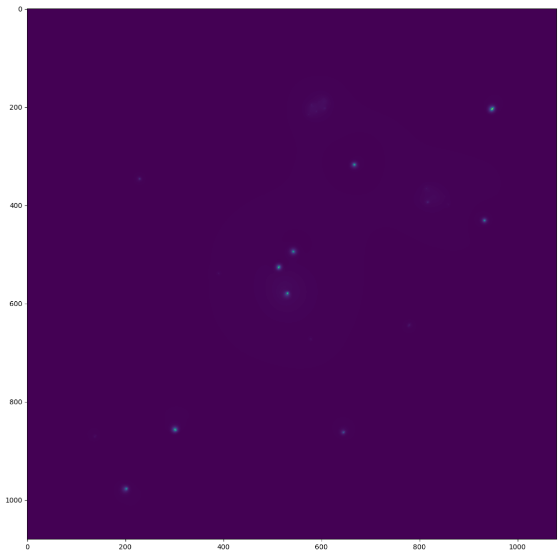


Figure 7: Dirty Image

6 Conclusion



(a) Reconstruction after one full iteration



(b) Reconstruction after 4 Full iterations

Figure 8

References

[1] 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.

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

[3] 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.

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

[5] 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.

[6] Arwa Dabbech, Alexandru Onose, Abdullah Abdulaziz, Richard A Perley, Oleg M Smirnov, and Yves Wiaux. Cygnus a super-resolved via convex optimization from vla data. Monthly Notices of the Royal Astronomical Society, 476(3):2853–2866, 2018.

[7] Jasleen Birdi, Audrey Repetti, and Yves Wiaux. Sparse interferometric stokes imaging under polarization constraint (polarized sara). arXiv preprint arXiv:1801.02417, 2018.

[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] Luke Pratley, Melanie Johnston-Hollitt, and Jason D McEwen. A fast and exact w-stacking and w-projection hybrid algorithm for wide-field interferometric imaging. arXiv preprint arXiv:1807.09239, 2018.

[10] Luke Pratley, Jason D McEwen, Mayeul d'Avezac, Rafael E Carrillo, Alexandru Onose, and Yves Wiaux. Robust sparse image reconstruction of radio interferometric observations with purify. Monthly Notices of the Royal Astronomical Society, 473(1):1038–1058, 2017.

[11] Jason D McEwen and Yves Wiaux. Compressed sensing for wide-field radio interferometric imaging. Monthly Notices of the Royal Astronomical Society, 413(2):1318–1332, 2011.

[12] TD Carozzi. Imaging on a sphere with interferometers: the spherical wave harmonic transform. Monthly Notices of the Royal Astronomical Society: Letters, 451(1):L6–L10, 2015.

[13] JD McEwen and AMM Scaife. Simulating full-sky interferometric observations. Monthly Notices of the Royal Astronomical Society, 389(3):1163–1178, 2008.

[14] Nathanaël Schaeffer. Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations. Geochemistry, Geophysics, Geosystems, 14(3):751–758, 2013.

[15] J Richard Shaw, Kris Sigurdson, Ue-Li Pen, Albert Stebbins, and Michael Sitwell. All-sky interferometry with spherical harmonic transit telescopes. The Astrophysical Journal, 781(2):57, 2014.

[16] Jason D McEwen, Gilles Puy, Jean-Philippe Thiran, Pierre Vandergheynst, Dimitri Van De Ville, and Yves Wiaux. Sparse image reconstruction on the sphere: implications of a new sampling theorem. IEEE Transactions on image processing, 22(6):2275–2285, 2013.

[17] Stephen J Hardy. Direct deconvolution of radio synthesis images using l1 minimisation. Astronomy & Astrophysics, 557:A134, 2013.

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

[19] Zhi-Quan Luo and Paul Tseng. On the convergence of the coordinate descent method for convex differentiable minimization. Journal of Optimization Theory and Applications, 72(1):7–35, 1992.

[20] Simon Felix, Roman Bolzern, and Marina Battaglia. A compressed sensing-based image reconstruction algorithm for solar flare x-ray observations. The Astrophysical Journal, 849(1):10, 2017.

List of Figures

List of Tables

7 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, January 9, 2019

Jonas Schwammberger