P9 Distributed Image Reconstruction for the new Radio Interferometers

Jonas Schwammberger

May 3, 2019

Abstract

Contents

1 Introduction

In Astronomy, instruments with higher angular resolution allows us to measure ever smaller structures in the sky. For Radio frequencies, the angular resolution is bound to the antenna dish diameter, which puts practical and financial limitations on the highest possible angular resolution. Radio Interferometers get around this limitation by using several smaller antennas instead. Together, they act as a single large antenna with higher angular resolution at lower financial costs compared to single dish instruments.

However, Radio Interferometers do not measure the pixels of the sky image. Instead, Radio Interferometers measure an incomplete set of Fourier components. The sky image has to be reconstructed from the measurements. This forms an ill-posed inverse problem: There are potentially many different images that fit the measurements, and from the measurements alone we cannot decide which image was actually observed.

Algorithms that solve the ill-posed Image Reconstruction problem. Find most likely images. Extensive Research in the last decades.

SKA wants to create Radio Interferometers on a new scale. MeerKAT is the precursor of SKA-Mid. Create a new scale of Fourier measurements. Push towards distributing The Image Reconstruction Problem has to be solved with distributed computing.

Algorithms were developed before the advent of distributed computing. Distribution so far has been difficult. Only small number of nodes.

Target to distribute the image reconstruction First tests

1.1 Radio Interferometry System

This project is focused on distributing Image Reconstruction for Radio Interferometers, which is only one of three steps in the pipeline from measurements to the final image. We give a quick overview over the whole pipeline in figure 1 and how Radio Interferometers work in principle: The antennas observe the arriving electromagnetic wave, gets processed in three steps, Correlation, Calibration and Image Reconstruction.

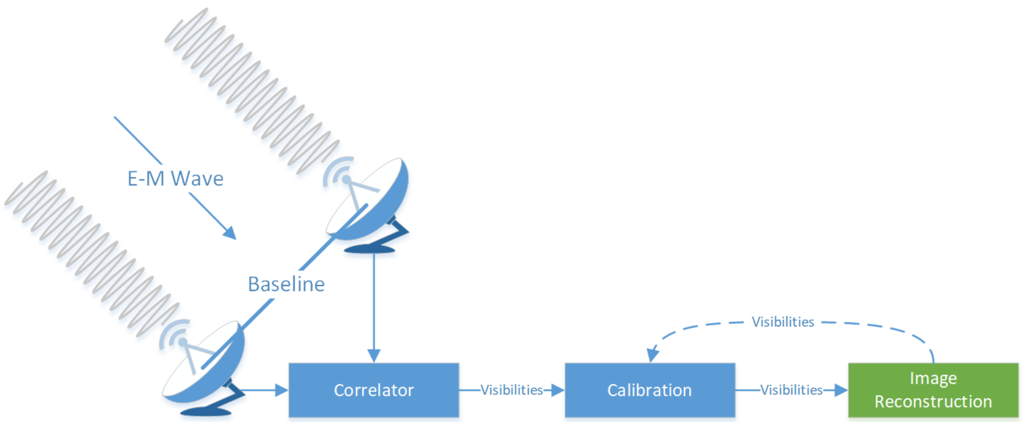


Figure 1: Radio Interferometer System

First, the electromagnetic wave gets measured by the different antennas of the interferometer. The measurements of each antenna pair get correlated into a complex-valued Fourier Component (called Visibility in Radio Astronomy). Each antenna pair measures a noisy amplitude and phase of a single Visibility (Fourier Component) of the sky image. The distance and orientation of the antenna pair relative to the incoming signal, called the baseline, dictates which Visibility gets measured. The longer the baseline the higher-order Visibility gets measured, resulting in a higher angular resolution. After correlation, the Visibility data is saved for later processing.

The Calibration step is done after all Visibility data has been recorded. This step corrects the amplitude and phase of the measurements for varying antenna sensitivities, pointing errors and other effects. Also, this step removes corrupted data from the measurements. After the Visibilities are calibrated, we can average the measurements and reduce the data volume by several factors. Typically, averaging is done to reduce the runtime costs of the Image Reconstruction step.

The Image Reconstruction step takes the calibrated, and potentially averaged down Visibility data and finds the most likely observed image of the sky. Although the Interferometer produces a potentially large number of Visibilities, they are incomplete: For example an Interferometer with dish-antennas is typically blind to the microwave background radiation. The largest structures in the image it can detect depends on the shortest baseline. Since the antennas have to be at least the dish-diameter placed apart, the Interferometer is simply blind to very large structures in the sky image, like the microwave background radiation. This property makes Image Reconstruction Problem ill-posed. In the Image Reconstruction step, we therefore have to find the most likely image which fits the measurements.

1.1.1 Earth's rotation and arbitrary large Number of Visibilities

To improve the final image, we want to measure as many different Visibilities as possible. Modern Radio Interferometers use the the earth's rotation to change the baselines. When the earth rotates, it modifies the length of the baseline and by extent, what Visibilities get measured. Modern Interferometer can produce an almost arbitrary large number of measurements by just increasing the observation time.

The data volume can be averaged down in the Calibration step. However, with self-calibration, the Image Reconstruction is tasked with solving both the most likely image and the calibration parameters at the same time. This further improves the reconstruction quality[**?**], but requires the Image Reconstruction step to handle all the Visibility data.

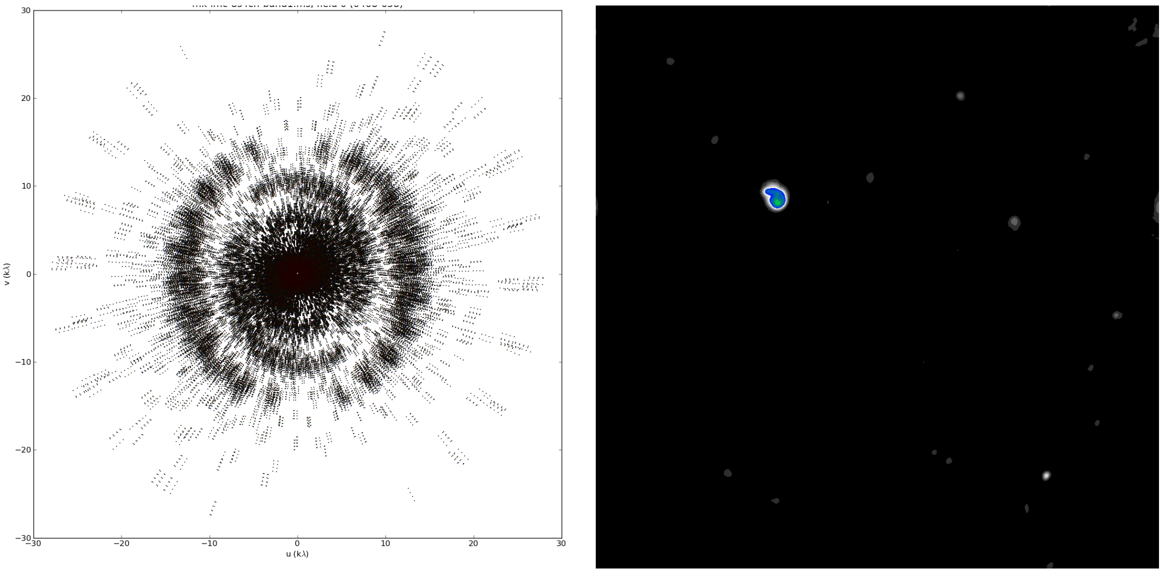
Modern Radio Interferometers can produce an almost arbitrary large number of measurements. The reconstruction quality benefits from a large number of different Visibility measurements. The two main limiting factors however, are the cost of data storage and the scalability of the Image Reconstruction algorithm.

1.2 The Image Reconstruction Problem

For Radio Interferometers, Image Reconstruction forms an ill-posed inverse problem. There are potentially many different images that fit the measurements. The Image Reconstruction is tasked with finding the most likely image $I()$ given the (calibrated) Visibility measurements $V()$. Figure 2 shows an example from a MeerKAT observation for $V()$ and $I()$. The image 2a shows the incomplete sampling in the Fourier space. The sample density decreases away from the center[1]. The observed image 2b contains two classes of structure, point sources and extended emissions. Point sources arise from stars and other distant objects, their emissions are concentrated around a single pixel. Extended emissions arise from hydrogen clouds and other sources which span an area of several arc-seconds of the sky.

Formally, we want to invert the measurement equation of the Radio Interferometer.

[1]The center has actually no samples, since the MeerKAT Radio Interferometer uses antenna's with dishes. Due to the resolution, it might not be visible in the image 2a.



(a) Measurements $V()$ in the uv-plane. *(b) The observed image $I()$.*

Figure 2: The Image Reconstruction Problem

Equation (1.1) shows a measurement equation for Radio Interferometers,

$$V(u,v,w) = \int\int \frac{I(x,y)}{c(x,y)} e^{2\pi i[ux+vy+w(c(x,y)-1)]}\,dx\,dy\ , \quad c(x,y) = \sqrt{1-x^2-y^2} \tag{1.1}$$

The measurement equation (1.1) contains three parts. The Visibility measurements $V(u,v,w)$, the observed image with a normalization factor $\frac{I(x,y)}{c(x,y)}$ and the Fourier Transform $e^{2\pi i[\ldots]}$. The x and y coordinates are angles from the image center. A pixel in $I(x,y)$ represents how much radio emission was measured from a the direction x, y. Note that the Visibilities $V(u,v,w)$ are three dimensional, while the image $I(x,y)$ only has two. However, also note that the third component w only depends on the directions x and y. In a sense, the Visibilities $V()$ and the Image $I()$ have a two dimensional Fourier relationship ($V(u,v,w) = \int\int I(x,y)e^{2\pi i[ux+vy]}\,dx\,dy$), but with a directionally dependent correction factor $e^{2\pi i[\ldots+w(c(x,y)-1)]}$.

The third component w is an example of a Directionally Dependent Effect (DDE) which have a tendency to increase the runtime costs of the Image Reconstruction. The w-component keeps us from using the Fast Fourier Transform (FFT) for the measurement equation (1.1). Research in this area tries to use approximations which lets us use faster algorithms like the FFT, and correct for DDE's accurately enough [1, 2, 3]. In this project, the w-correction is the only DDE we handle.

1.2.1 System of Linear Equations

Even though the Fourier Transform in equation (1.1) contains a w-correction factor, $V()$ and $I()$ have still a linear relationship. This means we can represent the Image Reconstruction Problem as a system of linear equations: (1.2). F is the Fourier Transform with w-correction, x is the image we are searching and V are the measured Visibilities[2].

$$\underset{x}{find} \quad V = Fx \tag{1.2}$$

[2] V in the equation (1.2) is a vector. We use the lower-case v to denote the axis in the Fourier space uvw, and the upper-case letter to denote the Visibility vector.

As we discussed in section 1.1, Radio Interferometers can produce an almost arbitrary large number of Visibilities. In practice, we often reconstruct images with far fewer pixels than Visibility measurements, making the equation (1.2) an over-determined problem. Meaning equation (1.2) is either consistent and has a unique solution x, or is inconsistent with no solutions. Sadly, the Visibility measurements are subject to noise. The equation (1.2) is inconsistent does not have a solution. To solve the Image Reconstruction Problem, we need to account for noise in the measurements, which leads us to a new inequality (1.3).

$$\underset{x}{find} \; \|V - Fx\|_2^2 < \epsilon \tag{1.3}$$

We use the L2 norm and find the solution x which, when the Fourier Transform is applied, has the smallest distance to the measurements V. The L2 norm leads to a strictly-convex function, meaning (1.3) has only one minimum, and we can use convex optimization techniques to find it. So the question arises, is the observed image $I()$ located at the global minimum of (1.3)? Although we have more measurements than pixels in the reconstruction, the set of Visibilities are still incomplete. They shift the observed image some distance away from the global minimum of (1.3). From the measurements alone, we cannot locate the observed image. We know it is near the global minimum, but we do not know where it is. The equation (1.3) has many candidate solutions x, one of which is the observed image. From the measurements alone, we cannot decide which candidate was the observed image, and (1.3) is an ill-posed inverse problem.

Intuitively, the incomplete measurements allow the pixels x in the reconstruction to form physically implausible solutions.

In many Image Reconstruction Problems of Radio Interferometry, the pixels cannot be negative. We know it contains stars,

If we include prior information about the image, we can retrieve the most likely image from all the candidates of (1.3). The question is, what are the chances that the most likely image is the observed image? The theory of Compressed Sensing tells us this.

1.2.2 Theory of Compressed Sensing

Theory of Compressed Sensing [4, 5] is a recently developed sampling theorem. How we can include prior information about the image and how likely we are to retrieve the observed image. Prior.

sparsity Dictionary D. For example Wavelets. Can be larger than the number of pixels

For our considerations LASSO objective (1.4)

$$\underset{x}{minimize} \; \|V - Fx\|_2^2 + \lambda \left\|D^{-1}x\right\|_1 \tag{1.4}$$

Data term Regularization term

Problem of a result: we still don't know how close we are, or if certain structures

So we need a Regularization Dicitionary D, which captures the prior information we have about the image. Then, we can use convex optimization techniques to minimize the objective (1.4).

1.2.3 Different Representations for the LASSO Objective

There are different ways to represent the LASSO objective (1.4). Has the data term in the Fourier space, and reconstructs the image directly. We have different design decisions here. We can choose the Data term to be either in Fourier or image space, and we can set the Variables to be in either image, uniform Fourier or in the sparse space.

These possibilities arise from the Fourier Transform Matrix F, shown in equation (1.5). The first line shows the normal Fourier transform, from image x into Visibilities V. Since x is in a uniformly-sampled space, we can factorize F into the FFT and the masking matrix M and we arrive at the second line of equation (1.5). M is essentially a matrix contains values between zero and one. Represents the degradation due to incomplete measurements.

$$\begin{aligned} V &= Fx \\ V &= M\ FFT(x) \\ V &= M\ V_2 \\ I_{Dirty} &= x \star PSF \end{aligned} \tag{1.5}$$

We can use the matrix M directly, and use it to transform from uniformly sampled Visibilities V_2 to non-uniformly sampled Visibilities V, arriving at line three of equation (1.5).

Since a multiplication in Fourier space is a convolution in image space, we can also represent the masking matrix M as a Point Spread Function PSF in image space. Degradation is a convolution. I_{Dirty} calculation is a Fourier Transform of the measurements. So we need to solve a deconvolution problem.

All these lines are of (1.5) all represent the Image Reconstruction Problem in different spaces. For example, if we change the data term in our LASSO objective from $\|V - Fx\|_2^2$ to a deconvolution $\|I_{Dirty} - x \star PSF\|_2^2$ we arrive at the same global minimum. But, in Radio Astronomy, the images are magnitudes smaller than the calibrated Visibility measurements V. Reducing the problem size. In practice, we have an accuracy issue with the PSF, since it is not constant. The PSF changes over the image due to DDEs like w-correction. We will use deconvolution in the state-of-the-art.

$$\begin{aligned} \underset{x}{minimize}\ \|V - Fx\|_2^2 + \lambda \left\|D^{-1}x\right\|_1 \\ \underset{x}{minimize}\ \|I_{Dirty} - x \star PSF\|_2^2 + \lambda \left\|D^{-1}x\right\|_1 \end{aligned} \tag{1.6}$$

We can set the data term in our LASSO objective to either Visibility or image space. The reconstruction space has three possible design choices, which leads to three different LASSO objectives.

$$\text{analysis: } \underset{x}{minimize}\ \|V - Fx\|_2^2 + \lambda \left\|D^{-1}x\right\|_1 \tag{1.7}$$

$$\text{in-painting: } \underset{V_2}{minimize}\ \|V - MV_2\|_2^2 + \lambda \left\|D^{-1}F^{-1}V_2\right\|_1 \tag{1.8}$$

$$\text{synthesis: } \underset{\alpha}{minimize}\ \|V - FD\alpha\|_2^2 + \lambda \left\|\alpha\right\|_1 \tag{1.9}$$

Analysis, standard formulation.

To our knowledge, in-painting (1.8) is not used for the Image Reconstruction Problem. Difficulty in the Regularization term.

Synthesis reconstructs in the sparse space, increases the number of free variables, we have at least as many α than x. But, we do not need the Matrix D^{-1}, which for certain spaces may not even be defined. We can use over-complete representations. Why use over-complete? More freedom?

1.3 Solving the Image Reconstruction Problem: The Major/Minor Cycle Architecture

We want to solve ill-posed Image Reconstruction problem (1.3).

In this section we present the state-of-the-art in solving the Image Reconstruction Problem, the Major/Minor Cycle algorithm. Like the name implies, we split the problem into two parts, the Major Cycle, which is responsible for calculating the Fourier Transform, and the Minor Cycle, which is responsible for deconvolving the image. The figure 3 shows the Major/Minor Cycle architecture and their steps.

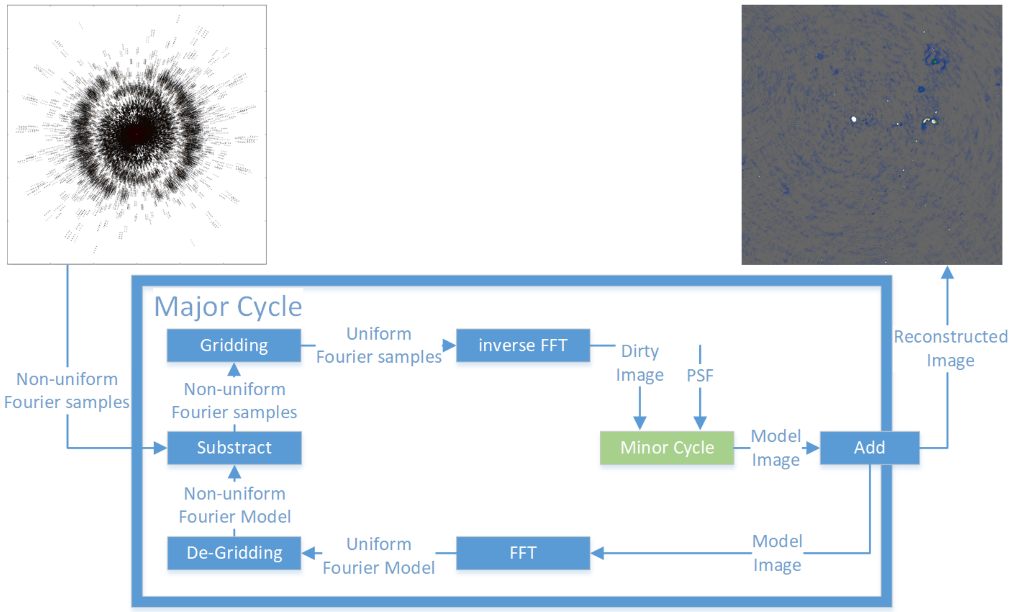


Figure 3: The Major/Minor Cycle Architecture

Let us first look at the Major Cycle in more detail. The Gridder first takes the non-uniformly sampled Visibilities and interpolates them on a uniform grid. This allows us to use the inverse FFT to calculate the acutal Fourier Transform. All DDEs, like w-correction fall into the Gridder in this architecture. Gridder is not perfect. Remember that we can factorize the Fourier Transform Matrix into two parts, $F = M\ FFT(x)$. In the Gridding step, we essentially applies M^{-1} on the calibrated Visibilities.

After the inverse FFT, we arrive at the Minor Cycle deconvolution. Deconvolution problem is a lot smaller and (mostly) free from DDEs. The PSF is just an approximation Here in the Minor Cycle, we include prior knowledge about the image to approximate the ill-posed inverse problem. We have different choices of algorithms here. Our choice defines the reconstruction quality.

Then De-gridding.

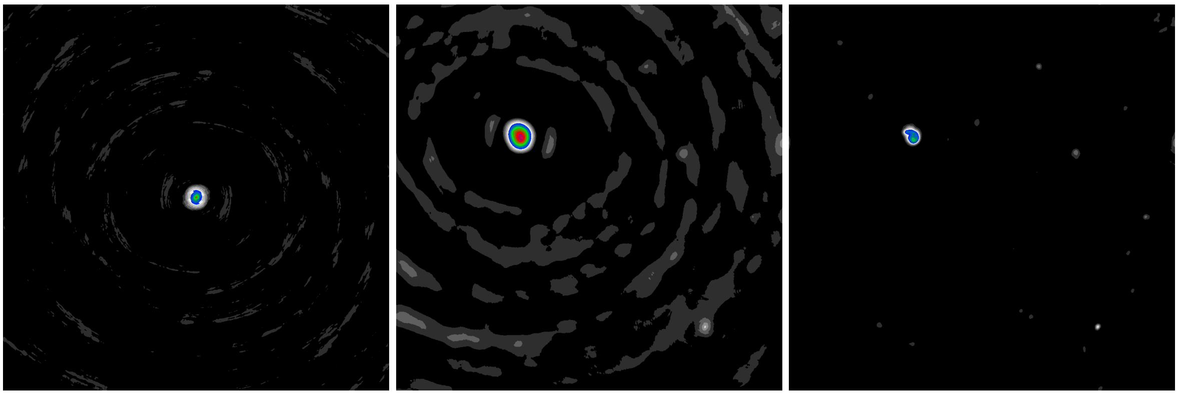
What the major cycle does, Minor Cycle can approximate the deconvolution.

Runtime: Gridder problematic. It has to deal with the calibrated Visibilities, which are LARGE. The unifrom grid is typically smaller. Gridder is the first [**?**]

1.3.1 Minor Cycle: CLEAN Deconvolutions

The Minor Cycle was conceived with CLEAN in mind. CLEAN is state of the art algorithm

Contains two classes of objects: Point sources, which are essentially stars, and extended emissions, which span over several pixels.



(a) Point Spread Function. (b) Dirty image (c) Deconvolved image.

Figure 4: Image reconstruction of two simulated point sources.

Find the fainter sources in later iterations Because we can only estimate the psf.

1.4 Distributing the Major and Minor Cycles

Difficulty in separation Each Pixel depends on all Visibilities, and a single Visibility depends on all pixels. Doing work several times

1.4.1 Compressed Sensing and the evolution of Priors

Positivity constraint. Contains two classes of objects: Point sources, which are essentially stars, and extended emissions, which span over several pixels.

1.5 The Major Cycle Architecture

Major cycle how to reconstruct the image with deconvolution In an efficient manner

$V()$ problems of non-uniform sampling, and the 3 dimensions. keep us from using the Fast Fourier Transform. We first interpolate on a regularly spaced grid, in the "Gridder". Use the FFT. For large numbers of Visibilities, this is faster. than inverting equation directly (1.1). And now we do a deconvolution in image space

So we have three basic components, Gridder, FFT and Deconvolution algorithms. The Major Cycle Architecture makes this a cycle. Shown in figure 3.

Why the cycle is necessary

1.5.1 Minor Cycle

2 State of the art Image Reconstruction

2.1 Gridder

2.1.1 -stacking

2.1.2 IDG

2.2 Reconstruction

2.2.1 CLEAN

2.2.2 SARA

3 Distributing the Image Reconstruction

Distributing the whole image reconstruction. So far, only parts were distributed. First time that end to end, everything gets distributed.

OpenMPI

Gridding and Devonvolution

3.1 Distributed Gridder: The IDG algorithm

Veeneboer et al[1] developed the Image Domain Gridder. It uses Subgrids and solves each subgrid seperately. It is in the image domain, because it can do Radio Interferometer specific corrections efficiently. Furthermore, it leads to a structure which is primed for GPU processing. We use this algorithm to distribute the gridding.

W-Projection, Spheroidal are convolutions in the Fourier space.

The figure 5 shows the different parts of the image domain algorithm.

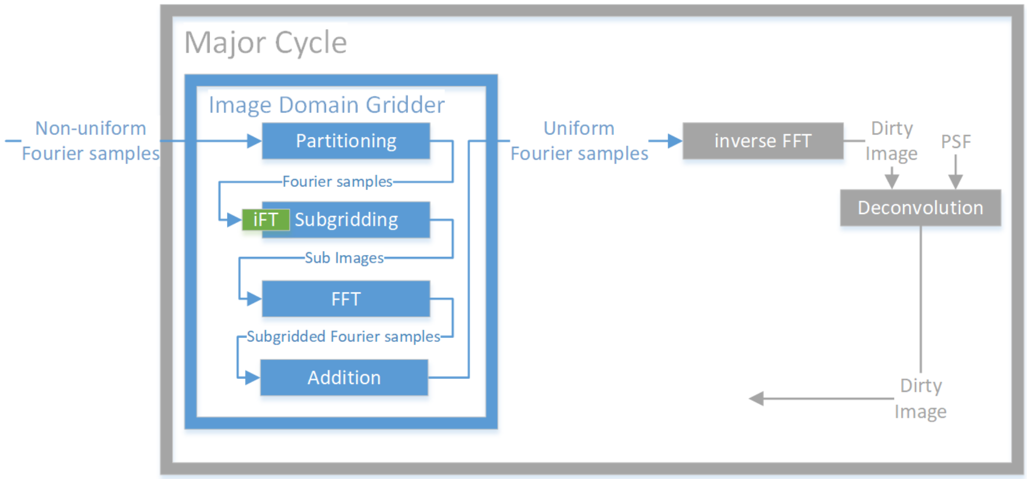


Figure 5: Image Domain Gridder in the Major Cycle Architecture

Algorithm

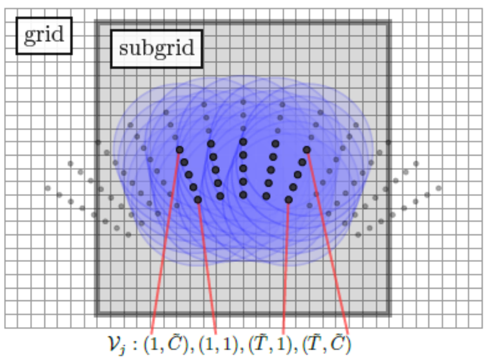


Figure 6: Subgrid

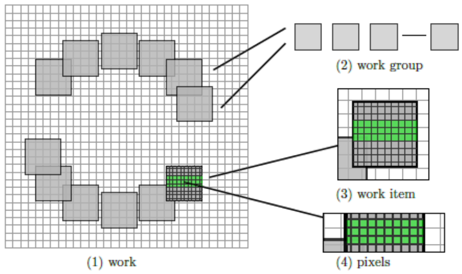


Figure 7: parallel

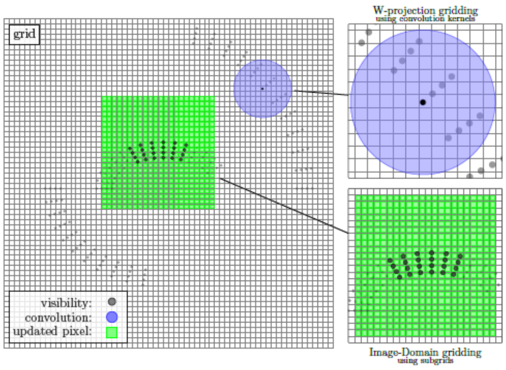


Figure 8: Image Domain Gridder in the Major Cycle Architecture

3.2 Distributed Deconvolution: Coordinate Descent

4 Conclusion

References

[1] Bram Veenboer, Matthias Petschow, and John W Romein. Image-domain gridding on graphics processors. In 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS), pages 545–554. IEEE, 2017.

[2] AR Offringa, Benjamin McKinley, Natasha Hurley-Walker, FH Briggs, RB Wayth, DL Kaplan, ME Bell, Lu Feng, AR Neben, JD Hughes, et al. Wsclean: an implementation of a fast, generic wide-field imager for radio astronomy. Monthly Notices of the Royal Astronomical Society, 444(1):606–619, 2014.

[3] 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.

[4] 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.

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

[6] Arwa Dabbech, Chiara Ferrari, David Mary, Eric Slezak, Oleg Smirnov, and Jonathan S Kenyon. More-sane: Model reconstruction by synthesis-analysis estimators-a sparse deconvolution algorithm for radio interferometric imaging. Astronomy & Astrophysics, 576:A7, 2015.

List of Figures

List of Tables

5 attachment

6 Larger runtime costs for Compressed Sensing Reconstructions

The MeerKAT instrument produces a new magnitude of data volume. An image with several million pixels gets reconstructed from billions of Visibility measurements. Although MeerKAT measures a large set of Visibilities, the measurements are still incomplete. We do not have all the information available to reconstruct an image. Essentially, this introduces "fake" structures in the image, which a reconstruction algorithm has to remove. Additionally, the measurements are noisy.

We require an image reconstruction algorithm which removes the "fake" structures from the image, and removes the noise from the measurements. The large data volume of MeerKAT requires the algorithm to be both scalable and distributable. Over the years, several reconstruction algorithms were developed, which can be separated into two classes: Algorithms based on CLEAN, which are cheaper to compute and algorithms based on Compressed Sensing, which create higher quality reconstructions.

CLEAN based algorithms represent the reconstruction problem as a deconvolution. First, they calculate the "dirty" image, which is corrupted by noise and fake image structures. The incomplete measurements essentially convolve the image with a Point Spread Function (PSF). CLEAN estimates the PSF and searches for a deconvolved version of the dirty image. In each CLEAN iteration, it searches for the highest pixel in the dirty image, subtracts a fraction PSF at the location. It adds the fraction to the same pixel location of a the "cleaned" image. After several iterations, the cleaned image contains the deconvolved version of the dirty image. CLEAN accounts for noise by stopping early. It stops when the highest pixel value is smaller than a certain threshold. This results in a light-weight and robust reconstruction algorithm. CLEAN is comparatively cheap to compute, but does not produce the best reconstructions and is difficult to distribute on a large scale.

Compressed Sensing based algorithms represent the reconstruction as an optimization problem. They search for the optimal image which is as close to the Visibility measurements as possible, but also has the smallest regularization penalty. The regularization encodes our prior knowledge about the image. Image structures which were likely measured by the instrument result in a low regularization penalty. Image structures which were likely introduced by noise or the measurement instrument itself result in high penalty. Compressed Sensing based algorithms explicitly handle noise and create higher quality reconstructions than CLEAN. State-of-the-art Compressed Sensing algorithms show potential for distributed computing. However, they currently do not scale on MeerKATs data volume. They require too many computing resources compared to CLEAN based algorithms.

This project searches for a way to reduce the runtime costs of Compressed Sensing based algorithms. One reason for the higher costs is due to the non-uniform FFT Cycle. State-of-the-art CLEAN and Compressed Sensing based algorithms both use the non-uniform FFT approximation in a cycle during reconstruction. The interferometer measures the Visibilities in a continuous space in a non-uniform pattern. The image is divided in a regularly spaced, discrete pixels. The non-uniform FFT creates an approximate, uniformly sampled image from the non-uniform measurements. Both, CLEAN and Compressed Sensing based algorithms use the non-uniform FFT to cycle between non-uniform Visibilities and uniform image. However, a Compressed Sensing algorithm requires more non-uniform FFT cycles for reconstruction.

CLEAN and Compressed Sensing based algorithms use the non-uniform FFT in a similar manner. However, there are slight differences in the architecture. This project hypothesises that The previous project searched for an alternative to the non-uniform FFT cycle. Although there are alternatives, there is currently no replacement which leads to lower runtime costs for Compressed Sensing. Current research is focused on reducing the number of non-uniform FFT cycles for Compressed Sensing algorithms.

CLEAN based algorithms use the Major Cycle Architecture for reconstruction. Compressed Sensing based algorithms use a similar architecture, but with slight modifications. Our hypothesis is that we may reduce the number of non-uniform FFT cycles for Compressed Sensing by using CLEAN's Major Cycle Architecture.

6.1 CLEAN: The Major Cycle Architecture

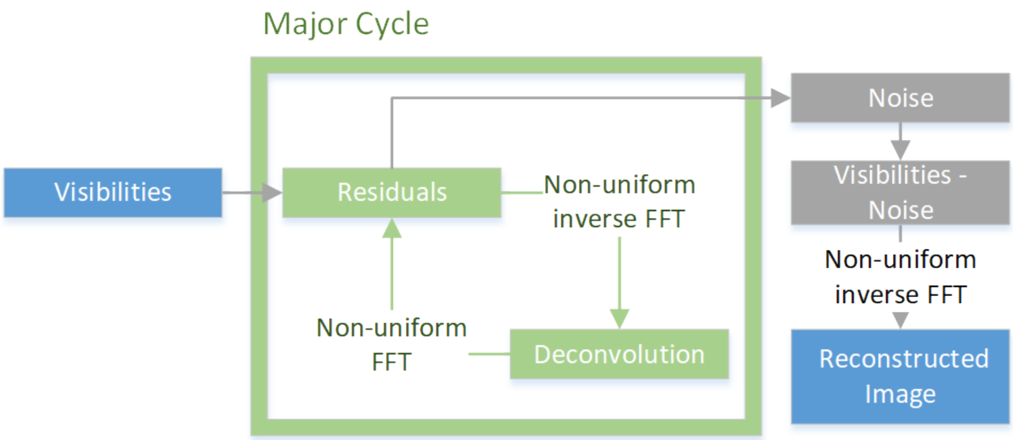


Figure 9: The Major Cycle Architecture

Figure 9 depicts the Major Cycle Architecture used by CLEAN algorithms. First, the Visibilities get transformed into an image with the non-uniform FFT. The resulting dirty image contains the corruptions of the measurement instrument and noise. A deconvolution algorithm, typically CLEAN, removes the corruption of the instrument with a deconvolution. When the deconvolution stops, it should have removed most of the observed structures from the dirty image. The rest, mostly noisy part of the dirty image gets transformed back into residual Visibilities and the cycle starts over.

In the Major Cycle Architecture, we need several deconvolution attempts before it has distinguished the noise from the measurements. Both the non-uniform FFT and the deconvolution are approximations. By using the non-uniform FFT in a cycle, it can reconstruct an image at a higher quality. For MeerKAT reconstruction with CLEAN, we need approximately 4-6 non-uniform FFT cycles for a reconstruction.

6.2 Compressed Sensing Architecture

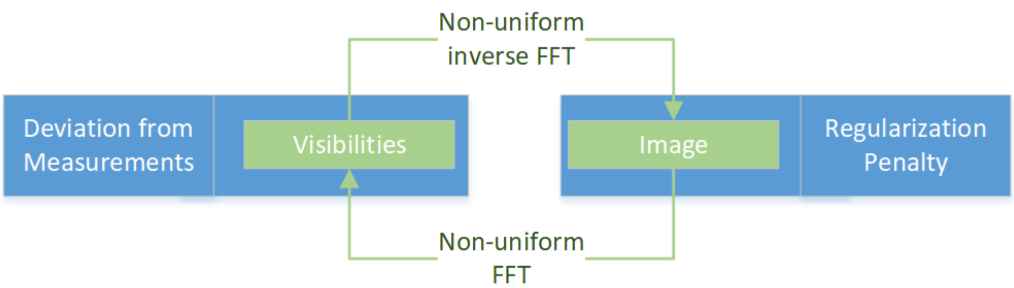


Figure 10: State-of-the-art Compressed Sensing Reconstruction Architecture

Figure 10 depicts the architecture used by Compressed Sensing reconstructions. The Visibilities get transformed into an image with the non-uniform FFT approximation. The algorithm then modifies the image so it reduces the regularization penalty. The modified image gets transformed back to Visibilities and the algorithm then minimizes the difference between measured and reconstructed Visibilities. This is repeated until the algorithm converges to an optimum.

In this architecture, state-of-the-art Compressed Sensing algorithms need approximately 10 or more non-uniform FFT cycles to converge. It is one source for the higher runtime costs. For MeerKAT reconstructions the non-uniform FFT tends to dominate the runtime costs. A CLEAN reconstruction with the Major Cycle Architecture already spends a large part of its time in the non-uniform FFT. Compressed Sensing algorithms need even more non-uniform FFT cycle on top of the "Image Regularization" step being generally more expensive than CLEAN deconvolution. There is one upside in this architecture: State-of-the-art algorithms managed to distribute the "Image Regularization" operation.

6.3 Hypothesis for reducing costs of Compressed Sensing Algorithms

Compressed Sensing Algorithms are not bound to the Architecture presented in section 6.2. For example, we can design a Compressed Sensing based deconvolution algorithm and use the Major Cycle Architecture instead.

Our hypothesis is: We can create a Compressed Sensing based deconvolution algorithm which is both distributable and creates higher quality reconstructions than CLEAN. Because it also uses the Major Cycle architecture, we reckon that the Compressed Sensing deconvolution requires a comparable number of non-uniform FFT cycles to CLEAN. This would result in a Compressed Sensing based reconstruction algorithm with similar runtime costs to CLEAN, but higher reconstruction quality and higher potential for distributed computing.

6.4 State of the art: WSCLEAN Software Package

6.4.1 W-Stacking Major Cycle

6.4.2 Deconvolution Algorithms

CLEAN MORESANE

6.5 Distributing the Image Reconstruction

6.5.1 Distributing the Non-uniform FFT

6.5.2 Distributing the Deconvolution

7 Handling the Data Volume

The new data volume is a challenge to process for both algorithms and computing infrastructure. Push for parallel and distributed algorithms. For Radio Interferometer imaging, we require specialized algorithms. The two distinct operations, non-uniform FFT and Deconvolution, were difficult algorithms for parallel or distributed computing.

The non-uniform FFT was historically what dominated the runtime []. Performing an efficient non-uniform FFT for Radio Interferometers is an active field of research[2, 3], continually reducing the runtime costs of the operation. Recently, Veeneboer et al[1] developed a non-uniform FFT which can be fully executed on the GPU. It speeds up the most expensive operation.

In Radio Astronomy, CLEAN is the go-to deconvolution algorithm. It is light-weight and compared to the non-uniform FFT, a cheap algorithm. It is also highly iterative, which makes it difficult for effective parallel or distributed implementations. However, compressed sensing based deconvolution algorithms can be developed with distribution in mind.

7.1 Fully distributed imaging algorithm

Current imaging algorithms push towards parallel computing with GPU acceleration. But with Veeneboer et al's non-uniform FFT and a compressed sensing based deconvolution, we can go a step further and create a distributed imaging algorithm.

8 Image Reconstruction for Radio Interferometers

In Astronomy, instruments with higher angular resolution allows us to measure ever smaller structures in the sky. For Radio frequencies, the angular resolution is bound to the antenna dish diameter, which puts practical and financial limitations on the highest possible angular resolution. Radio Interferometers get around this limitation by using several smaller antennas instead. Together, they act as a single large antenna with higher angular resolution at lower financial costs compared to single dish instruments.

Each antenna pair of an Interferometer measures a single Fourier component of the observed image. We can retrieve the image by calculating the Fourier Transform of the measurements. However, since the Interferometer only measures an incomplete set of Fourier components, the resulting image is "dirty", convolved with a Point Spread Function (PSF). Calculating the Fourier Transform is not enough. To reconstruct the from an Interferometer image, an algorithm has to find the observed image with only the dirty image and the PSF as input. It has to perform a deconvolution. The difficulty lies in the fact that there are potentially many valid deconvolutions for a single measurement, and the algorithm has to decide for the most likely one. How similar the truly observed image and the reconstructed images are depends largely on the deconvolution algorithm.

State-of-the-art image reconstructions use the Major Cycle architecture (shown in Figure 11), which contains three operations: Gridding, FFT and Deconvolution.

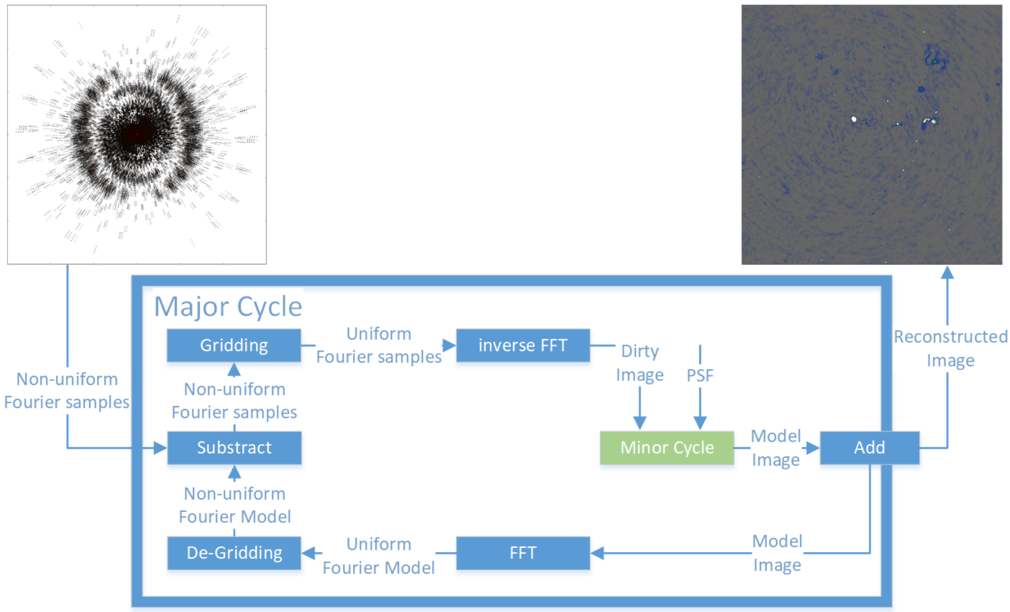


Figure 11: The Major Cycle Architecture of image reconstruction algorithms

The first operation in the Major Cycle, Gridding, takes the non-uniformly sampled Fourier measurements from the Interferometer and interpolates them on a uniformly spaced grid. The uniform grid lets us use FFT to calculate the inverse Fourier Transform and we arrive at the dirty image. A deconvolution algorithm takes the dirty image plus the PSF as input, producing the deconvolved "model image", and the residual image as output. At this point, the reverse operations get applied to the residual image. First the FFT and then De-gridding, arriving at the non-uniform Residuals. The next Major Cycle begins with the non-uniform Residuals as input. The cycles are necessary, because the Gridding and Deconvolution operations are only approximations. Over several cycles, we reduce the errors introduced by the approximate Gridding and Deconvolution. The final, reconstructed image is the addition of all the model images of each Major Cycle.

8.1 Distributed Image Reconstruction

New Interferometer produce an ever increasing number of measurements, creating ever larger reconstruction problems. A single image can contain several terabytes of Fourier measurements. Handling reconstruction problems of this size forces us to use distributed computing. However, state-of-the-art Gridding and Deconvolution algorithms only allow for limited distribution. How to scale the Gridding and Deconvolution algorithms to large problem sizes is still an open question.

Recent developments make a distributed Gridder and a distributed Deconvolution algorithm possible. Veeneboer et al[1] found an input partitioning scheme, which allowed them to perform the Gridding on the GPU. The same partitioning scheme can potentially be used to distribute the Gridding onto multiple machines. For Deconvolution, there exist parallel implementations for certain algorithms like MORESANE[6]. These can be used as a basis for a fully distributed image reconstruction.

In this project, we want to make the first steps towards an image reconstruction algorithm, which is distributed from end-to-end, from Gridding up to and including deconvolution. We create our own distributed Gridding and Deconvolution algorithms, and analyse the bottlenecks that arise.

8.2 First steps towards a distributed Algorithm

In this project, we make the first steps towards a distributed Major Cycle architecture (shown in figure 11) implemented C#. We port Veeneboer et al's Gridder, which is written in C++, to C# and modify it for distributed computing. We implement a simple deconvolution algorithm based on the previous project and create a first, non-optimal distributed version of it.

In the next step, we create a more sophisticated deconvolution algorithm based on the shortcomings of the first implementation. We use simulated and real-world observations of the MeerKAT Radio Interferometer and measure its speed up. We identify the bottlenecks of the current implementation and explore further steps.

From the first lessons, we continually modify the distributed algorithm and focus on decreasing the need for communication between the nodes, and increase the overall speed up compared to single-machine implementations. Possible Further steps:

- Distributed FFT
- Replacing the Major Cycle Architecture
- GPU-accelerated Deconvolution algorithm.

A state-of-the-art reconstruction algorithm has to correct large number of measurement effects arising from the Radio Interferometer. Accounting for all effects is out of the scope for this project. We make simplifying assumptions, resulting in a proof-of-concept algorithm.

9 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, May 3, 2019

Jonas Schwammberger