

# SparseRI: A Compressed Sensing Framework for Aperture Synthesis Imaging in Radio Astronomy

S. Wenger<sup>1,2</sup>, M. Magnor<sup>1</sup>, Y. Pihlström<sup>3</sup>, S. Bhatnagar<sup>4</sup>, and U. Rau<sup>4</sup>

## ABSTRACT

In radio interferometry, information about a small region of the sky is obtained in the form of samples in the Fourier transform domain of the desired image. Since this sampling is usually incomplete, the missing information has to be reconstructed using additional assumptions about the image. The emerging field of compressed sensing (CS) provides a promising new approach to this type of problem which is based on the supposed sparsity of natural images in some transform domain. We present a versatile CS-based image reconstruction framework called *SparseRI*, an interesting alternative to the CLEAN algorithm, that permits a wide choice of different regularizers for interferometric image reconstruction. The performance of our method is evaluated on simulated data as well as on actual radio interferometry measurements from the VLA, showing that our algorithm is able to reproduce the main features of the test sources. The proposed method is a first step towards an alternative reconstruction approach that may be able to avoid typical artifacts like negative flux regions, to work with large fields of view and non-coplanar baselines, to avoid the gridding process, and, in particular, to produce results not far from those achievable by human-assisted processing in CLEAN through an entirely automatic algorithm, making it especially well-suited for automated processing pipelines.

*Subject headings:* techniques: image processing — techniques: interferometric

## 1. Introduction

Since the middle of the twentieth century, interferometric techniques have been used to obtain images of the sky at radio wavelengths (Ryle & Vonberg 1946; Ryle & Hewish 1960; Swenson Jr. & Mathur 1968; Perley et al. 1989). The correlations between the signals from multiple antennae yield information about the *spatial* frequency content of the image, eventually allowing the image

---

<sup>1</sup>Computer Graphics Lab, TU Braunschweig, 38106 Braunschweig, Germany

<sup>2</sup>Institut für Geophysik und extraterrestrische Physik, TU Braunschweig, 38106 Braunschweig, Germany

<sup>3</sup>Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA

<sup>4</sup>National Radio Astronomy Observatory, Socorro, NM 87801, USA

itself to be reconstructed. Since the size of the synthesized beam of such a telescope is inversely proportional to the largest distance between any two antennae, very high spatial resolutions can be obtained this way. However, the process of reconstructing the image from incomplete frequency information is highly nontrivial, as missing information has to be appropriately reconstructed.

The reconstruction of missing information is only possible by specifying additional prior information about the image, such as its supposed smoothness, or the assumption that it contains a minimal amount of energy or a maximum amount of image ‘entropy’. Traditionally, the iterative deconvolution algorithm CLEAN (Högbom 1974) is used for the reconstruction of radio interferometric images. It implicitly assumes that the image is composed of a small number of point sources. Even though many extensions to the algorithm have been proposed (Sect. 2), extended intensity distributions are still not always well reconstructed by the algorithm, and the process may require considerable user guidance in order to yield satisfactory results.

Besides, new telescopes like the Long Wavelength Array LWA (Ellingson et al. 2009), the Low Frequency Array LOFAR (de Vos et al. 2009) or the Square Kilometre Array SKA (Ekers 2003; Schilizzi 2004) that image large parts of the sky at once require reconstruction algorithms that handle increasing amounts of data as well as non-coplanar telescope geometries (Cornwell et al. 2008; McEwen & Scaife 2008). Also, with a growing number of telescopes, a more automatic reconstruction pipeline is desirable.

A solution to many of these problems is proposed by the emerging theory of compressed sensing (Candès et al. 2006a,b; Candès 2006; Donoho 2006; Baraniuk 2007). It generalizes and formalizes the notion of *prior information* about a signal by relating it to its *compressibility*, a property that is present in most natural images (Clarke 1985; DeVore et al. 1992) and distinguishes them from random noise. Compressed sensing makes use of this fact in order to decide which of the possible images that explain a set of measurements is most probable, thereby turning the reconstruction problem into an optimization problem.

Subsequent to the advent of compressed sensing, existing reconstruction algorithms have been reinterpreted in the context of the new theory. This includes the mature CLEAN algorithm, which is an instance of a class of algorithms known as matching pursuit (MP) algorithms (Lannes et al. 1997). In addition, a multitude of new algorithms is now available to efficiently solve the numerical problems that occur in compressed sensing (e.g. Figueiredo et al. 2007; Tropp & Gilbert 2007), and some have successfully been applied to astronomy (Bobin et al. 2008), including radio interferometry (Wiaux et al. 2009a,b; Suksmono 2009). In this paper, we present SparseRI, a first step towards a general compressed sensing reconstruction framework for radio interferometry based on a recently developed accelerated *iterative shrinkage / thresholding* (IST) technique by Wright et al. (2009). We evaluate our algorithm using different sparsity priors on synthetic as well as real data. The evaluation shows that SparseRI is able to reproduce the main features of the test sources (Sect. 4.1), and our reconstruction of real observations (Sect. 4.2) proves its practical applicability.

## 2. Interferometric image formation

In radio interferometry, correlations between multiple antennae are used to synthesize an aperture the size of the largest baseline. After applying a band-pass filter to select the desired frequency, the measured *visibility*  $V(\mathbf{b})$  for a baseline  $\mathbf{b}$  depends on the intensity  $I(\mathbf{s})$  of the sky in the direction  $\mathbf{s}$  according to

$$V(\mathbf{b}) = \int I(\mathbf{s}) e^{-2\pi i \mathbf{s} \cdot \mathbf{b}} d\Omega. \quad (1)$$

Under the assumption that the sources being imaged are confined to a small region of the sky, this corresponds to a two-dimensional Fourier transform multiplied by a phase (e.g. Perley et al. 1989). This approximation may be rendered obsolete by compressed sensing techniques in the future (Sect. 5).

Obtaining the sky image from the visibilities is an ill-posed inversion problem, and several approaches have been used to solve it. The CLEAN algorithm (Högbom 1974), the standard reconstruction method used in radio interferometry software like AIPS <sup>1</sup> and CASA <sup>2</sup>, starts from the dirty image and successively subtracts a user-defined fraction of the point spread function around the brightest spots of the image. This process works well for images of isolated point sources, but does not always produce satisfying results for extended sources. Also, the convergence of the algorithm as well as the uniqueness of its solutions are not always guaranteed (Schwarz 1978, 1979). Regions that are supposed to be dark in the image are often excluded from the reconstruction process by hand in order to avoid artifacts in these regions, although some modifications to the algorithm try to reduce this and other objectionable effects (Segalovitz & Frieden 1978; Cornwell 1983; Schwab 1984) and to achieve better reconstructions of extended sources using multi-scale approaches (e.g. Cornwell 2008).

Other reconstruction algorithms used in the context of radio interferometry include the *maximum entropy method* (Cornwell & Evans 1985), which attempts to maximize an entropy functional while satisfying the constraints imposed by the measured data, and more generally any algorithm that maximizes a plausibility function.

The compressed sensing method that we are going to present allows for a mathematically very well-founded formulation of an important subclass of such algorithms. In simple terms, our algorithm searches the space of all images that are consistent with the measurements for a solution that minimizes an appropriate metric, for example the magnitude of the coefficients of the image in a wavelet or gradient domain representation. Results similar to those of CLEAN can be obtained when the magnitude of the pixels is minimized instead.

---

<sup>1</sup><http://aips.nrao.edu/>

<sup>2</sup><http://casa.nrao.edu/>

### 3. Compressed sensing

The theory of compressed sensing, recently introduced by Emmanuel Candès and others (Candès et al. 2006a,b; Candès 2006; Donoho 2006; Baraniuk 2007), generalizes the way of thinking about sampling. While the usual Shannon–Nyquist sampling theory requires that a band-limited signal be sampled with at least twice its highest frequency, compressed sensing states that a signal that is not necessarily band-limited but sparse in some basis (i.e. the majority of its coefficients in that basis are zero) can be reconstructed from a small number of measurements in another basis (i.e. a number of linear combinations of these coefficients). As a trivial example, if a signal is known to consist of a single Gaussian, it can be reconstructed perfectly from three known points even though the signal is not band-limited. Similarly, a signal that is sparse in the spatial domain (i.e. a combination of a small number of point sources) can be reconstructed from a small number of Fourier coefficients, which is typically the case in radio interferometry.

The measurement of a signal  $\mathbf{x}$  using a linear measurement operator  $\mathbf{M}$  can be written as  $\mathbf{y} = \mathbf{M}\mathbf{x}$ , where  $\mathbf{y}$  is the resulting vector of measurements. In radio interferometry,  $\mathbf{y}$  contains the visibilities,  $\mathbf{M}$  is a sampled Fourier transform, and  $\mathbf{x}$  is the vector of image pixels. If  $\mathbf{x}$  has a sparse representation  $\mathbf{s} = \mathbf{B}\mathbf{x}$  in a sparsity basis  $\mathbf{B}$ , then  $\mathbf{x} = \mathbf{B}^T\mathbf{s}$  (because  $\mathbf{B}$ , representing a change of basis, is orthonormal by design), and  $\mathbf{y} = \mathbf{M}\mathbf{B}^T\mathbf{s}$ .

The possibility of perfect reconstruction of the signal  $\mathbf{x}$  depends on the properties of the combined measurement and sparsity matrix  $\mathbf{M}\mathbf{B}^T$ . Candès & Tao (2005) prove that perfect reconstruction occurs when (in simple terms) all sets of a sufficiently high number of columns of  $\mathbf{M}\mathbf{B}^T$  are approximately orthonormal. The required number of columns depends on the sparsity level of the signal. In particular, any *random* measurement – or a randomly sampled Fourier measurement – satisfies this criterion with high probability (Candès 2006). Stable reconstruction is possible even when  $\mathbf{y}$  is perturbed by measurement error and when  $\mathbf{s}$  is not perfectly sparse (Candès et al. 2006b), two properties that are important for real-world applications.

A sparse signal  $\mathbf{s}$  can be reconstructed from its measurement vector  $\mathbf{y}$  by finding the sparsest vector  $\mathbf{s}$  that satisfies  $\mathbf{y} = \mathbf{M}\mathbf{B}^T\mathbf{s}$  (Candès et al. 2006a). Finding a sparsest vector is equivalent to minimizing the  $\ell_0$  norm which is defined as the number of nonzero coefficients of a vector. Unfortunately, the  $\ell_0$  norm minimization problem is computationally not tractable for large-scale problems. With increasing dimension of the vector  $\mathbf{s}$ , however, the solution of the  $\ell_0$  norm minimization approaches that of an  $\ell_1$  norm minimization, which can be solved efficiently. Algorithms also exist for the case of vector-valued entries of  $\mathbf{y}$  and  $\mathbf{s}$ , such as the different polarization components of a signal (Fornasier & Rauhut 2008).

As opposed to the procedural definition of CLEAN, the result of the optimization algorithm is mathematically uniquely defined. Therefore it does not depend on any specific implementation or parameter set, and theoretical analysis of the algorithm is greatly simplified. The only parameters that influence the optimal solution are the expected noise level of the measurement and the chosen sparsity prior.

### 3.1. Sparsity priors

The selection of a sparsity basis or *sparsity prior* represents our assumptions about the image to reconstruct and as such can strongly influence the reconstruction. For example, the pixel basis is obviously well suited to represent the assumption of isolated point sources, and is therefore implicitly used in the conventional CLEAN algorithm. In contrast, most terrestrial images are likely to contain large regions of homogeneous or slowly changing intensity, possibly with small-scale perturbations or sharp edges, and appropriate sparsity bases are known which can as well be used for astronomical imaging. For example, different wavelet representations like those proposed by Daubechies (1988) or Cohen et al. (1992) efficiently compress many natural images because they provide a scaling-independent but localized basis. Minimizing the total variation (which is however not a basis in the strict sense) will reliably localize extended emissive regions because in this case, sharp edges are not penalized more than smooth gradients. In Sect. 4, we provide sample reconstructions using pixel and wavelet sparsity priors.

### 3.2. Reconstruction algorithm

In previous work (Wiaux et al. 2009a,b), basis pursuit algorithms were used to recover simulated images of different radio sources (random Gaussians and string signals) from simulated measurements as a proof of concept. Suksmono (2009) applied similar methods to data from the Very Large Array (VLA), using total variation minimization for regularization. However, the exact solver employed there is not suitable for large-scale problems. We implement a different algorithm that can also be shown to work with data from real radio interferometers and that permits a wide choice of different regularizers in order to reconstruct a wider range of sources. It is also stable with respect to the errors introduced by the physical measurement process (Zhu 2008), and, as an approximate solver, is more efficient for high-dimensional problems.

Our algorithm is an adaptation of the very general compressed sensing framework *SpaRSA* by Wright et al. (2009). The algorithm minimizes a functional of the form

$$\|\mathbf{y} - \mathbf{M}\mathbf{B}^T\mathbf{s}\|_2^2 + \lambda f(\mathbf{s}), \quad (2)$$

where  $\mathbf{y}$  is the vector of measurements,  $\mathbf{M}$  describes the measurement operation – in this case, a Fourier transform – and  $\mathbf{s} = \mathbf{B}\mathbf{x}$  is the sparse representation of the vector  $\mathbf{x}$  of image pixels to be reconstructed. The first term (the *data term*) promotes consistency with the measurements, while the second term (the *regularizer*) demands plausibility of the reconstructed image.

The plausibility metric  $f(\mathbf{s})$  can be any (preferably convex) function for which the simpler term

$$\|\mathbf{u} - \mathbf{s}\|_2^2 + \beta f(\mathbf{s}) \quad (3)$$

can easily be minimized with respect to  $\mathbf{s}$  for any  $\mathbf{u}$  and  $\beta$ . When this minimization problem is solved as a sub-step of the SpaRSA algorithm,  $\mathbf{u}$  is replaced by a local linear approximation of

the data term  $\|\mathbf{y} - \mathbf{M}\mathbf{B}^T\mathbf{s}\|_2^2$ , forcing  $\mathbf{s}$  to move towards the subspace where  $\mathbf{y} = \mathbf{M}\mathbf{B}^T\mathbf{s}$ . The parameter  $\beta$ , on the other hand, is composed of the regularization parameter  $\lambda$  and a step size, which will become evident from the pseudo-code description of the algorithm.

A notable example for a plausibility metric is the  $\ell_1$  norm of  $\mathbf{s}$ , defined as  $\|\mathbf{s}\|_1 = \sum_i |(\mathbf{s})_i|$ , for which the minimum of (3) is analytically given as

$$(\mathbf{s})_i = \frac{\max(|(\mathbf{u})_i| - \beta, 0)}{\max(|(\mathbf{u})_i| - \beta, 0) + \beta} (\mathbf{u})_i. \quad (4)$$

Implementations for many different wavelet transforms that work well with  $\ell_1$  minimization are available in the WaveLab package (Buckheit et al. 2005). Another useful plausibility metric is the *total variation* – which can be regarded as the  $\ell_1$  norm of the gradient image of the image  $\mathbf{x}$  – for which numerical solvers are available (e.g. Chambolle & Darbon 2009). In this case,  $\mathbf{B}$  is conveniently chosen as the identity matrix, so that  $\mathbf{s} = \mathbf{x}$ .

In (2), the parameter  $\lambda$  describes the balance between data fidelity and *a priori* plausibility of the image. In practice, it is most appropriately chosen so that  $\|\mathbf{y} - \mathbf{M}\mathbf{B}^T\mathbf{s}\|_2^2$  lies just within the expected noise level of the instrument when the optimization is finished. This can be implemented by starting with a relatively large  $\lambda$ , which is then successively decreased in a so-called *continuation scheme* (Wright et al. 2009) until the noise level is reached.

The algorithm for reconstructing an image  $\mathbf{x}$  from visibilities  $\mathbf{y} = \mathbf{M}\mathbf{x} = \mathbf{M}\mathbf{B}^T\mathbf{s}$ , given a sparsity basis  $\mathbf{B}$  and the rms noise level  $\sigma$ , can be written in pseudo-code as follows:

```

s ← BMTy (dirty image in sparsity domain)
α ← 1 (initial step size)
repeat (continuation scheme for λ)
  r ← y − MBTs (residual visibility data)
  λ ←  $\|\mathbf{B}\mathbf{M}^T\mathbf{r}\|_\infty$  (regularization parameter)
  repeat (minimize (2))
     $e \leftarrow \frac{1}{2} \|\mathbf{r}\|_2^2 + \lambda f(\mathbf{s})$  (error term)
     $\mathbf{s}' \leftarrow \arg \min_{\hat{\mathbf{s}}} \frac{1}{2} \|\mathbf{s} + \frac{\mathbf{B}\mathbf{M}^T\mathbf{r}}{\alpha} - \hat{\mathbf{s}}\|_2^2 + \frac{\lambda}{\alpha} f(\hat{\mathbf{s}})$  (minimize (3))
    r ← r − MBT(s' − s) (update residual visibilities)
    α ←  $\frac{\|\mathbf{M}\mathbf{B}^T(\mathbf{s}' - \mathbf{s})\|_2^2}{\|\mathbf{s}' - \mathbf{s}\|_2^2}$  (adapt step size)
    s ← s' (update dirty image in sparsity domain)
  until change in  $e \leq 10^{-6}$  (convergence of the inner loop)
until  $\|\mathbf{r}\|_2 \leq \sigma$  (residual visibilities below observational noise)
x ← BTs (transform solution from sparsity to image domain).

```

Here,  $\leftarrow$  assigns a value to a variable,  $\cdot^T$  denotes matrix adjoint (or transpose),  $\|\mathbf{x}\|_\infty = \max_i |(\mathbf{x})_i|$  is the maximum norm, and  $\|\cdot\|_2$  is the Euclidean norm. For efficiency reasons, matrix products with  $\mathbf{M}$  and  $\mathbf{M}^T$  are implemented as two-dimensional forward and backward Fast Fourier Transforms (FFTs), respectively. Since in practice the visibilities do not lie on a regular grid as required by the FFT, *gridding* is necessary, which inevitably introduces slight quantization errors. If computation time is no primary issue, these errors may be reduced by applying convolution-based gridding and degridting before and after the forward and backward transform, respectively. In addition, gridding may be avoided entirely by future extensions of our framework (Sect. 5).

## 4. Evaluation

In the following, we evaluate the performance of SparseRI using two different datasets, a synthetic one and one from the Very Large Array.

### 4.1. Simulated data

For a synthetic source, the ‘ground truth’ image is known, allowing us to objectively compare the performance of different reconstruction algorithms. We provide three different error metrics: the signal-to-noise ratio (SNR), which is defined as  $\text{SNR} = -20 \log_{10} \frac{\sigma_{\text{residual}}}{\sigma_{\text{original}}}$  (cf. Wiaux et al. 2009a), the RMS error normalized to the average of the true image, and the dynamic range, which is the ratio of the highest peak in the reconstruction to the standard deviation of the reconstruction noise that can be measured in empty background regions of the image. Since SparseRI and CLEAN are non-linear reconstruction algorithms, the dynamic range has to be interpreted with care.

In Fig. 1, we present two simulated sources (b) that are reasonably close to real extended radio sources while not being particularly easy or difficult to reconstruct for either algorithm: a uniform gradient with Gaussian decay and a series of Gaussians of increasing size. The measurements were simulated using the CASA `simdata` task. The large regions of extended emission, together with comparatively low sampling density, push the algorithms to their limits so that the differences in reconstruction quality become visible. The UV coverage – that is, the fraction of grid cells for which visibility data is known after gridding at the specified image resolution, in this case 256 by 256 pixels – is 4.2 per cent from a simulated 6000 s VLA observation in the ‘D’ configuration. Each pixel in the simulated image corresponds to one arcsecond, and the CLEAN beam has a full width at half maximum (FWHM) of 7.4 by 4.7 arcseconds.

As a reference, CLEAN reconstructions (c) were made using the Clark (1980) algorithm from the CASA software package. In order to make the results user-independent and comparable, both reconstruction algorithms have been run with their default parameters (for CLEAN, these were gain 0.1, threshold 0 Jy, natural weighting; however, the number of CLEAN iterations had to be increased from 500 to 1000 for the first and to 90000 for the second source in order to yield

satisfying results). For the CLEAN reconstructions, the SNRs are 6.1 dB and 15.9 dB, the dynamic ranges are 17.6 and 121.5, and the RMS errors per pixel are 0.12 and 0.05 per cent of the true mean intensity for the first and second source, respectively. It is noticeable that in the second source, the Gaussians are not clearly separated from each other.

The SparseRI reconstructions using a Daubechies wavelet basis – cf. Fig. 2 – yield SNRs of 5.0 dB and 6.6 dB, dynamic ranges of 19.2 and 24.6, and RMS errors per pixel of 0.13 and 0.15 per cent of the true mean intensity, Fig. 1(d). These metrics indicate that the performance of SparseRI is comparable with CLEAN, although the emerging algorithm is not yet on par with its mature counterpart. The largest contribution to the lower SNR values is due to the occurrence of negative flux regions which are not penalized in the current implementation (cf. Sect. 5). However, visual comparison shows that SparseRI is able to resolve the series of Gaussians, even though some larger scale stripes are present in the background. The computation time of the non-optimized, single-threaded SparseRI algorithm was about 4.5 seconds on conventional PC hardware. The automatic CLEAN reconstruction took three seconds for the 1000 iterations of the first source, and 160 seconds for the 90000 iterations of the second source.

## 4.2. Real data

The applicability of our algorithm to actual noise-affected measurements is demonstrated using snapshot observations from the VLA in the ‘D’ configuration at 14.965 GHz and a UV coverage of 2.4 per cent. Fig. 3 shows a series of reconstructions of this dataset containing the Sgr A West region, including the central ‘Minispiral’. Fig. 3(b) shows the reconstruction using CLEAN in CASA with default parameters. This CLEAN result can be considerably improved by manually optimizing the several parameters and constraining the intensity to specified regions (commonly referred to as *boxing*). Fig. 3(d) shows a radio map that was produced in AIPS by co-author YP in such a user-guided CLEAN session (100000 iterations, gain 0.1, manual boxing). Finally, Fig. 3(c) shows the results from SparseRI using the pixel basis. All images share a cell size of 1.2 arcseconds and are convolved with the CLEAN beam with a FWHM of 10.2 by 4.3 arcseconds. SparseRI appears to better reconstruct the image than the automated CLEAN, but still shows some systematic imaging effects at lower intensity levels. The dynamic range of the reconstructions (assuming the right quarter of the image is background) is 43.5 for CLEAN with the default parameters, 1367 for user-guided CLEAN and 610 for SparseRI. The computation time of the non-optimized, single-threaded SparseRI algorithm was 9 seconds on conventional PC hardware, the automatic CLEAN reconstruction took about two seconds, while approximately 15 minutes were needed for the user-guided reconstruction, including self-calibration steps.

## 5. Conclusion

The evaluation of our algorithm on simulated and real measurements shows that SparseRI is able to provide interferometric image reconstructions that reproduce the main features of complex sources without any manual parameter tweaking or boxing, at comparable computation times as traditional reconstruction algorithms. While SparseRI is a non-linear algorithm, the method has been proven to converge towards the optimal solution as well as to be stable with respect to noise (Zhu 2008). The reconstruction results of this first implementation are still above the observational noise level. However, a number of beneficial constraints and other potential improvements exist that have not yet been explored. We are confident that the compressed sensing approach has the potential to become on par with and possibly surpass traditional reconstruction algorithms and their modern variants, such as multi-scale techniques, with respect to automated performance and achievable resolution. A quantitative comparison to several up-to-date approaches will be the subject of future work.

A future systematic evaluation of different sparsity bases on a large database of images may be a first step towards a sparsity basis that is explicitly designed for typical radio images. However, one would have to make sure that the database does not suffer from systematic errors (such as reconstruction artifacts in the training database), and an efficient representation for such a non-systematic basis has yet to be found. A combination of the previously discussed wavelet, pixel and total variation representations would also be able to increase the reconstruction quality on a wider range of sources. In general, the choice of sparsity basis will reflect which features or spatial frequencies the user wants to emphasize.

Incorporating further constraints into SparseRI – such as suppressing negative flux regions, which cause most of the reconstruction error in the current implementation – will further increase reconstruction quality (cf. Wiaux et al. 2009a). Additionally, we aim to include calibration with respect to an absolute flux scale, which is not yet implemented in the present version.

We also consider implementing a fast, parallelized three-dimensional non-uniform Fourier transform, for example on inexpensive graphics hardware. Such a general transform could prevent the hitherto unavoidable gridding error, possibly further decrease computation time, and enable fully three-dimensional aperture synthesis for recent interferometry arrays with large fields of view or non-coplanar baselines, such as EVLA, ALMA, EVN, VLBA, or the recent LWA, LOFAR and SKA systems.

The authors would like to thank Soheil Darabi, Karl-Heinz Glassmeier, Dave Martin and Pradeep Sen for fruitful discussions as well as their support and advice. Marcus Magnor gratefully acknowledges support by the Feodor Lynen Alumni Program of the Alexander von Humboldt Foundation as well as by the Fulbright Visiting Scholars Program.

## REFERENCES

- Baraniuk, R. 2007, *IEEE Signal Proc. Magazine*, 24, 118
- Bobin, J., Starck, J.-L., & Ottensamer, R. 2008, *IEEE J. Selected Topics in Signal Proc.*, 2, 718
- Buckheit, J., Chen, S., Donoho, D. L., Johnstone, I., & Scargle, J. 2005, *About WaveLab*, Stanford University
- Candès, E. J. 2006, *Int. Congress Math.*, 3, 1433
- Candès, E. J., Romberg, J., & Tao, T. 2006a, *IEEE Trans. Inform. Theory*, 52, 489
- . 2006b, *Comm. Pure and Appl. Math.*, 59, 1207
- Candès, E. J., & Tao, T. 2005, *IEEE Trans. Inform. Theory*, 51, 4203
- Chambolle, A., & Darbon, J. 2009, *Int. J. Computer Vision*, 84, 288
- Clark, B. G. 1980, *A&A*, 89, 377
- Clarke, R. J. 1985, *Transform coding of images* (Academic Press)
- Cohen, A., Daubechies, I., & Feauveau, J. C. 1992, *Comm. Pure and Appl. Math.*, 45, 485
- Cornwell, T. J. 1983, *A&A*, 121, 281
- . 2008, *IEEE J. Selected Topics in Signal Proc.*, 2, 793
- Cornwell, T. J., & Evans, K. F. 1985, *A&A*, 143, 77
- Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, *IEEE J. Selected Topics in Signal Proc.*, 2, 647
- Daubechies, I. 1988, *Comm. Pure and Appl. Math.*, 41, 909
- DeVore, R. A., Jawerth, B., & Lucier, B. J. 1992, *IEEE Trans. Inform. Theory*, 38, 719
- Donoho, D. L. 2006, *IEEE Trans. Inform. Theory*, 52, 1289
- Ekers, R. D. 2003, in *Proc. of the IAU 8th Asian-Pacific Regional Meeting*, Vol. 289, 21–28
- Ellingson, S. W., Clarke, T. E., Cohen, A., Craig, J., Kassim, N. E., Pihlstrom, Y., Rickard, L. J., & Taylor, G. B. 2009, *Proc. IEEE*, 97, 1421
- Figueiredo, M. A. T., Nowak, R. D., & Wright, S. J. 2007, *IEEE J. Selected Topics in Signal Proc.*, 1, 586
- Fornasier, M., & Rauhut, H. 2008, *SIAM J. Num. Anal.*, 46, 577
- Högbom, J. A. 1974, *A&AS*, 15, 417

- Lannes, A., Anterrieu, E., & Maréchal, P. 1997, *A&AS*, 123, 183
- McEwen, J. D., & Scaife, A. M. M. 2008, *MNRAS*, 389, 1163
- Perley, R. A., Schwab, F. R., & Bridle, A. H., eds. 1989, *Conference Series*, Vol. 6, *Synthesis Imaging in Radio Astronomy* (Astronomical Society of the Pacific)
- Ryle, M., & Hewish, A. 1960, *MNRAS*, 120, 220
- Ryle, M., & Vonberg, D. D. 1946, *Nat*, 158, 339
- Schilizzi, R. T. 2004, in *SPIE Conference Series*, Vol. 5489, 62–71
- Schwab, F. R. 1984, *AJ*, 89, 1076
- Schwarz, U. J. 1978, *A&A*, 65, 345
- Schwarz, U. J. 1979, in *IAU Colloq. 49: Image Formation from Coherence Functions in Astronomy*, Vol. 76, 261–275
- Segalovitz, A., & Frieden, B. R. 1978, *A&A*, 70, 335
- Suksmono, A. B. 2009, in *Proc. International Conference on Electrical Engineering and Informatics 2009 (ICEEI '09)*, Vol. 1, 110–116
- Swenson Jr., G. W., & Mathur, N. C. 1968, *Proc. IEEE*, 56, 2114
- Tropp, J., & Gilbert, A. 2007, *IEEE Trans. Inform. Theory*, 53, 4655
- de Vos, M., Gunst, A. W., & Nijboer, R. 2009, *Proc. IEEE*, 97, 1431
- Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., & Vanderghynst, P. 2009a, *MNRAS*, 395, 1733
- Wiaux, Y., Puy, G., & Vanderghynst, P. 2009b, *MNRAS*, in press
- Wright, S. J., Nowak, R. D., & Figueiredo, M. A. T. 2009, *IEEE Trans. Signal Proc.*, 57, 2479
- Zhu, C. 2008, *IEEE Trans. Inform. Theory*, 54, 3364

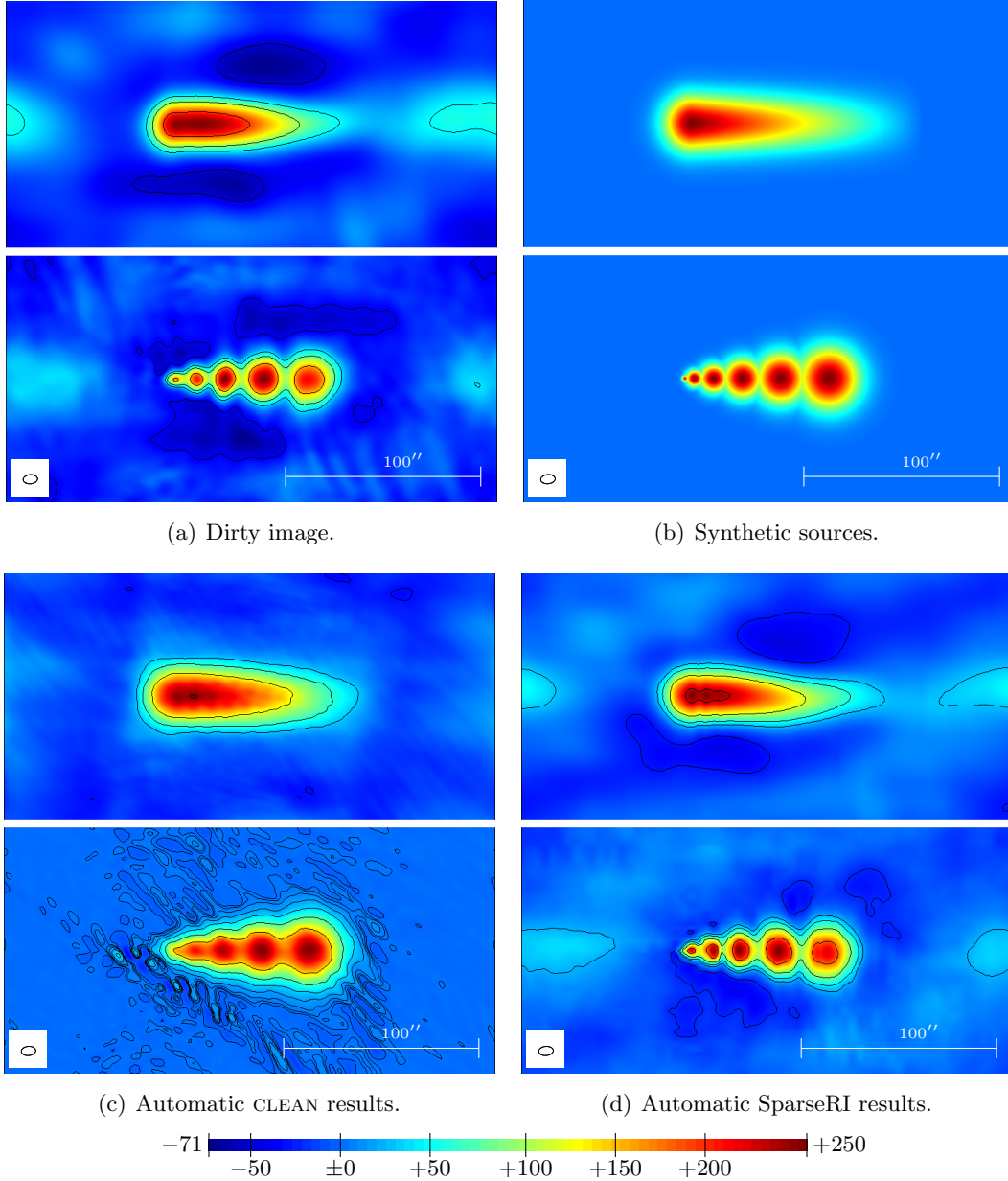


Fig. 1.— Synthetic sources (b) with reconstructions from simulated measurements using CLEAN with the default parameters (c) and SparseRI’s Daubechies wavelet minimization (d), respectively. Contour levels are plotted – except for the noise-free originals (b) – at  $-1, 1, 2, 4, \dots, 2^n$  times the respective  $3\sigma$  RMS noise. All images are 256 by 256 pixels (256 by 256 seconds of arc), and the UV coverage is 4.2 per cent. The lower left corner of each image shows the FWHM of the CLEAN beam.

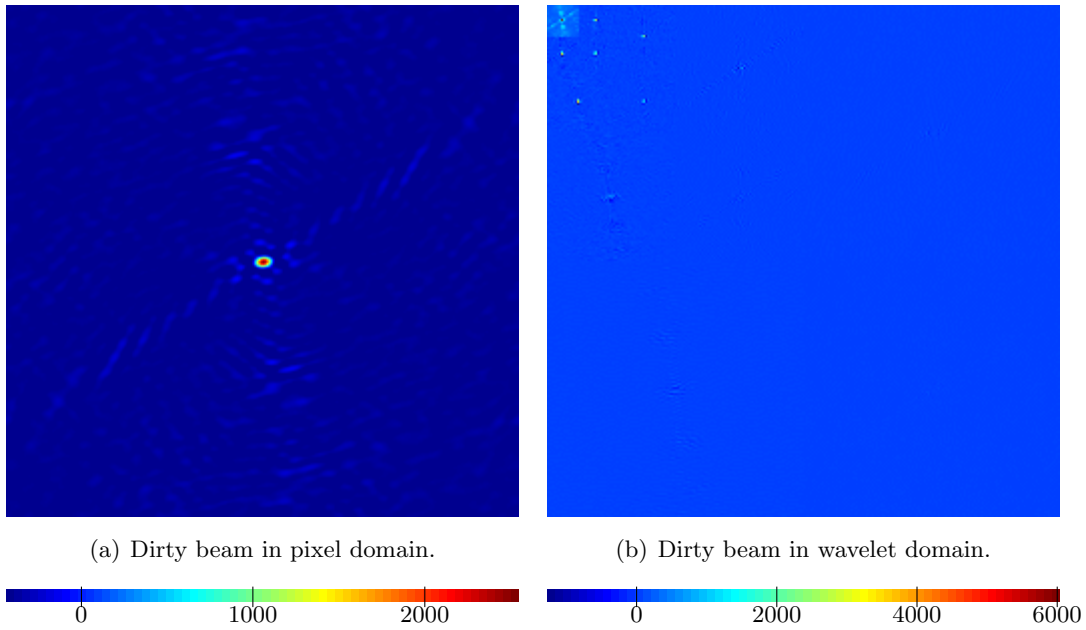


Fig. 2.— Dirty beam for the simulated data set, displayed in the pixel domain (a) and in the Daubechies-8 wavelet domain using four scale levels (b), which was used for the reconstruction in Fig. 1(d). Both plots are in arbitrary units.

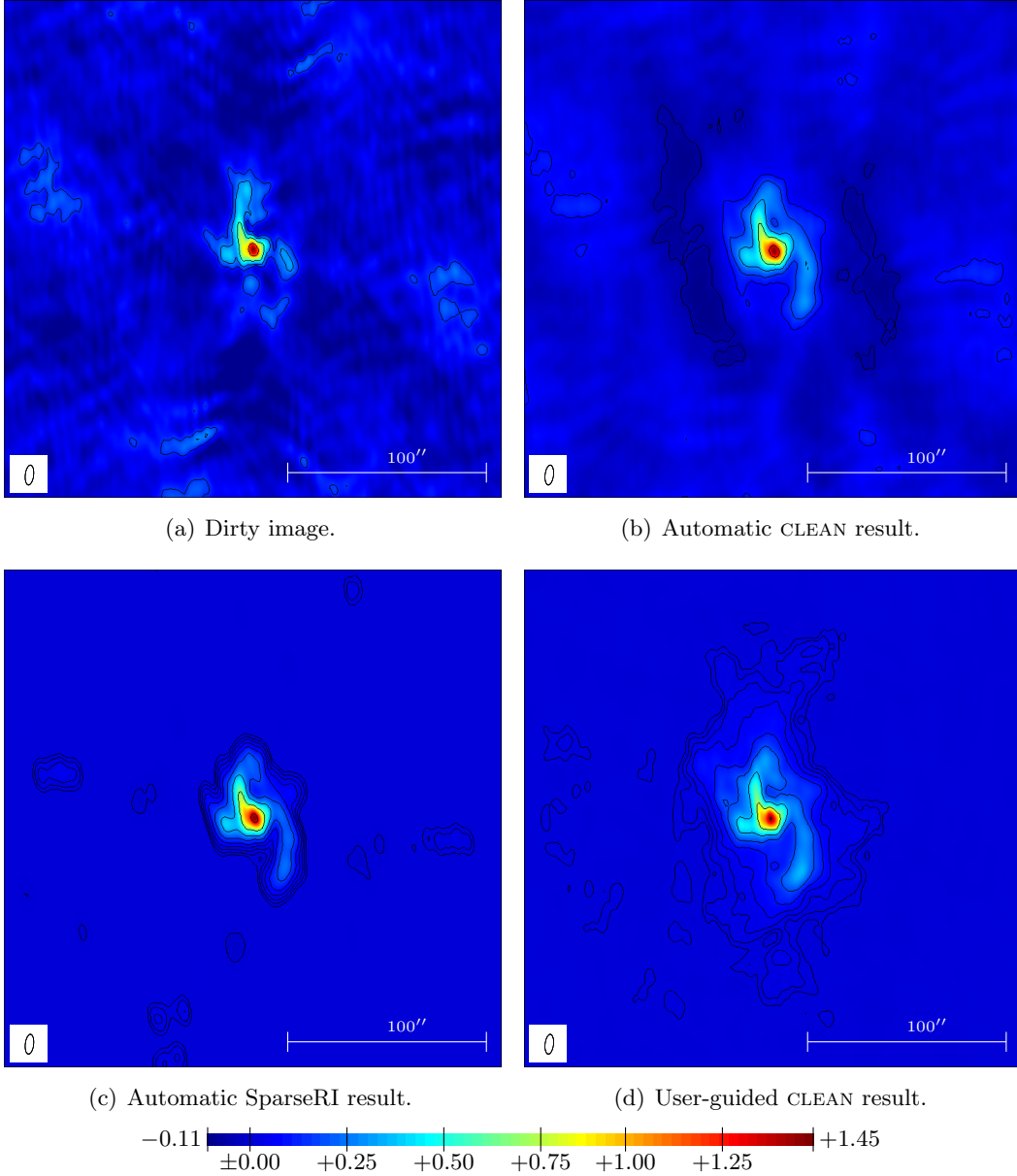


Fig. 3.— Sgr A West reconstructed from VLA data using CLEAN with the default parameters (b), SparseRI’s pixel magnitude minimization (c), and CLEAN with user guidance (d). Contour levels are plotted at  $-1, 1, 2, 4, \dots, 2^n$  times the respective  $3\sigma$  RMS noise. The color scale is in arbitrary units, with the peak flux of all images normalized to the same level. All images are cropped to the inner 210 by 210 pixels (about 250 by 250 seconds of arc), and the UV coverage is 2.4 per cent. The lower left corner of each image shows the FWHM of the CLEAN beam.