Revisiting the spread spectrum effect in radio interferometric imaging: A sparse variant of the w-projection algorithm

5 authors, including:

Laura Wolz
The University of Manchester
23 PUBLICATIONS   **350** CITATIONS

SEE PROFILE

Jason D. McEwen
University College London
74 PUBLICATIONS   **3,188** CITATIONS

SEE PROFILE

Rafael E. Carrillo
Centre Suisse d'Electronique et de Microtechnique
73 PUBLICATIONS   **1,069** CITATIONS

SEE PROFILE

Yves Wiaux
Heriot-Watt University
161 PUBLICATIONS   **2,463** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   HyperSARA View project

Project   Faceted HyperSARA View project

Revisiting the spread spectrum effect in radio interferometric imaging: a sparse variant of the w-projection algorithm

L. Wolz[1,2]⋆, J. D. McEwen[1,3], F. B. Abdalla[1], R. E. Carrillo[4], Y. Wiaux[4,5,6,7]

[1]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*

[2] *Sub-Department of Astrophysics, Department of Physics, University of Oxford, The Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

[3] *Mullard Space Science Laboratory, University College London, Surrey RH5 6NT, UK*

[4] *Institute of Electrical Engineering, Ecole Polytechnique Federale de Lausanne, CH-1015 Lausanne, Switzerland*

[5] *Department of Radiology and Medical Informatics, University of Geneva, CH-1211 Geneva, Switzerland*

[6] *Department of Radiology, Lausanne University Hospital, CH-1011 Lausanne, Switzerland*

[7] *Institute of Sensors, Signals & Systems, Heriot Watt University, Edinburgh EH14 4AS, UK*

ABSTRACT

Next-generation radio interferometric telescopes will exhibit non-coplanar baseline configurations and wide field-of-views, inducing a w-modulation of the sky image, which induces the spread spectrum effect. We revisit the impact of this effect on imaging quality and study a new algorithmic strategy to deal with the associated operator. In previous studies it has been shown that image recovery in the framework of compressed sensing is improved due to this effect, where the w-modulation can increase the incoherence between measurement and sparsifying signal representations. For the purpose of computational efficiency, idealised experiments with a constant baseline component w were performed. We extend this analysis to the more realistic setting where the w-component varies for each visibility measurement. Firstly, incorporating varying w-components into imaging algorithms is a computational demanding task. We propose a variant of the w-projection algorithm, which is based on an adaptive sparsification procedure, and incorporate it in compressed sensing imaging methods. Secondly, we show that for varying w-components, reconstruction quality is significantly improved compared to no w-modulation, reaching levels comparable to a constant, maximal w-component. This finding confirms that one may seek to optimise future telescope configurations to promote large w-components, thus enhancing the fidelity of image reconstruction.

Key words: techniques: interferometric – methods: numerical.

1 INTRODUCTION

Future radio interferometric telescopes are planned to have increased sensitivity and wider field-of-views (FoV) than ongoing projects. Hence, imaging techniques must be able to handle more complex settings like wide FoVs and non-coplanar telescope alignments in order to ensure accurate image recovery. Assuming a small FoV and a coplanar baseline alignment, the imaging problem can be posed in two dimensions neglecting the spherical nature of the sky and Earth. For the next generation of radio interferometric surveys, these approximations are not reasonable and it is necessary to develop imaging techniques that model the radio interferometric measurement equation more accurately.

The generalisation from planar approximations to wide FoV settings has been considered previously (Cornwell 1992; Cornwell et al. 2008; McEwen & Scaife 2008; Carozzi & Woan 2009; Wiaux et al. 2009b; McEwen & Wiaux 2011; Cornwell et al. 2012). A

common approach that is finding widespread application is the incorporation of the w-projection algorithm (Cornwell et al. 2008). The w-projection algorithm reformulates the radio interferometric measurement equation in such a way that the two-dimensional Fourier transform remains applicable, thus leading to efficient computation; this approach has been incorporated effectively in the latest imagers (Bhatnagar et al. 2008; Tasse et al. 2013). In these imagers, a generalisation to the so-called A-projection algorithm is made to handle direction-dependent primary beams.

Recently, the radio interferometric imaging problem has been addressed in the context of compressive sensing (Wiaux et al. 2009a,b; Suksmono 2009; Wiaux et al. 2010; Wenger et al. 2010; McEwen & Wiaux 2011; Li et al. 2011a,b; Carrillo et al. 2012), where wide FoVs have sometimes been considered (Wiaux et al. 2009b; McEwen & Wiaux 2011). In particular, the spread spectrum effect that arises in the wide FoV setting was first studied by Wiaux et al. (2009b), where it was shown both theoretically and empirically to enhance reconstruction quality in the compressive sensing framework. A non-negligible w-component (the component of

⋆ lwolz@star.ucl.ac.uk

the baseline in the pointing direction of the telescope) produces a w-modulation of the sky image, which acts to increase the incoherence between sensing and sparsity bases, thereby enhancing reconstruction quality. In the first study of the spread spectrum effect performed by Wiaux et al. (2009b), a constant w-component was examined in order to reduce significantly the computational cost of image recovery.

In the current article we revisit the spread spectrum effect in a more realistic setting, where the w-component varies for each visibility measurement. We support varying w-components by incorporating a variant of the w-projection algorithm (Cornwell et al. 2008) into the compressed sensing radio interferometric imaging framework. We propose an adaptive sparsification procedure to recover a sparse matrix representation of the w-modulation kernel. By applying the w-modulation via a sparse matrix multiplication, the computational burden and memory requirements are reduced substantially even when no approximation is made. The computational cost may be further reduced, for a small loss in reconstruction fidelity, by approximating the w-modulation kernel in a controlled manner. This procedure is generic and applicable for all types of direction dependent effects. We also extend the analysis to more realistic images and to different sparse dictionaries, while also considering different levels of visibility coverage.

The remainder of the article is organised as follows. In Sec. 2, we give a brief description of the radio interferometric measurement equation, recast the description in a discrete setting, and briefly review how to solve the imaging problem with compressed sensing methods. In Sec. 3, we describe a variant of the w-projection algorithm, incorporate it into the imaging framework and study its fast application. In Sec. 4, we describe the experimental set-up of our simulations and present our results regarding the impact of the spread spectrum effect in the setting of varying w-components. In Sec. 5, we present the main conclusions of this work.

2   RADIO INTERFEROMETRIC IMAGING

In this section we describe the radio interferometric measurement equation in the most general form and then proceed to recast the equation in operator formalism. Solving the interferometric imaging problem by compressed sensing techniques is outlined in a general form in the final part of this section.

2.1   Radio interferometric measurement equation

A radio interferometer is comprised of an array of telescopes and measures the Fourier coefficients, the so-called visibilities $\mathcal{V}$, of an image on the sky, where the visibility coordinates are given by the relative position between the antenna pairs. The baseline components (u, v, w) are measured in units of the wavelength λ of the incoming signal. The components u and v specify the planar baseline coordinates, while the third component w is associated with the basis vector of the coordinate frame pointing towards the direction of the center of the FoV of the telescope. The brightness of the sky I can be described in the same coordinate frame as the baseline, with components $\boldsymbol{l} = (l, m, n)$. Representing the celestial sphere by the unit sphere, the component n can be expressed in terms of (l, m) by $n(l, m) = \sqrt{1 - l^2 - m^2}$. The radio interferometric measurement equation for monochromatic, unpolarized signals reads

$$\mathcal{V}(u, v, w) = \int \frac{I(l, m)}{\sqrt{1 - l^2 - m^2}} \times e^{-2\pi i(ul + vm + w(\sqrt{1 - l^2 - m^2} - 1))} \mathrm{d}l\mathrm{d}m \quad (1)$$

(without loss of generality we neglect the shape of the primary beam pattern; one may consider it essentially set to unity or absorbed into the sky image). See Thompson et al. (2004) for further detail.

The $(n(l, m) - 1)$ term in the complex exponential of Eq. (1) can be neglected if the condition

$$(l^2 + m^2)w \ll 1 \quad (2)$$

is fulfilled (Thompson et al. 2004). This is in general true for either coplanar baseline configurations which lead to $w = 0$ (as realised by, *e.g.*, the VLA[1]; Condon et al. 1998) or a small FoV such that l and m are small. This approximation leads to the van Cittert-Zernike theorem (see *e.g.* Thompson et al. 2004), which states that the visibilities measured by a radio interferometric telescope are related to the sky brightness by a two-dimensional Fourier transform under the small FoV approximation. Some of the next-generation radio interferometric telescopes (*e.g.* the SKA[2]; Carilli & Rawlings 2004) will have considerably larger field of views and longer baselines than current telescopes, such that the condition specified by Eq. (2) is no longer valid.

It is convenient to rewrite Eq. (1) as

$$\mathcal{V}(u, v, w) = \int \frac{I(l, m)C(l, m; w)}{\sqrt{1 - l^2 - m^2}} e^{-2\pi i(ul + vm)} \mathrm{d}l\mathrm{d}m, \quad (3)$$

where the w-modulation kernel is defined by

$$C(l, m; w) \equiv e^{-2\pi iw(\sqrt{1 - l^2 - m^2} - 1)}. \quad (4)$$

In the scope of this paper, we follow Wiaux et al. (2009b) and use the first order approximation of $C(l, m; w)$ given by the linear chirp

$$C_1(l, m; w) = e^{-\pi iw(l^2 + m^2)}, \quad (5)$$

which is valid if $(l^2 + m^2)^2 w \ll 1$. Note that for constant w, Eq. (3) reduces to a two-dimensional Fourier transform since $C(l, m; w) \equiv C(l, m)$, leading to an efficient computational implementation. The w-modulation kernel modulates the sky brightness image, giving rise to the spread spectrum effect (Wiaux et al. 2009b). This was studied by Wiaux et al. (2009b) for a constant w-component, which reduces computational costs significantly. To study the spread spectrum effect for varying w-component is computationally demanding; this is the issue addressed and studied in the current article.

2.2   Operator description

We recast the radio interferometric measurement equation of Eq. (3) in discrete operator formalism (although the measurement equation can also be expressed in the Mueller matrix formalism; see, *e.g.*, Mueller 1948; Smirnov 2011; Tasse et al. 2013). We follow the same operator description as Wiaux et al. (2009b), where a w-modulation with constant w-component is considered (this restriction will be relaxed in Sec. 3). The visibilities $\mathcal{V}$ are assumed to be measured on a uniform grid in (u, v), yielding the vector of measured visibilities $\boldsymbol{y}$. Similarly, the sky brightness I is discretised on a uniform grid, yielding the image pixels (concatenated into the vector) $\boldsymbol{x}$. The radio interferometric measurement equation can then be written as

$$\boldsymbol{y} = \mathbf{MFCU}\boldsymbol{x} + \boldsymbol{n} = \boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{n}, \quad (6)$$

where the masking operator **M** specifies the visibility sampling pattern of the telescope, **F** denotes the two-dimensional Fourier transform, which is implemented by a fast Fourier transform (FFT), **C** denotes the diagonal w-modulation operator that implements modulation by Eq. (5) (for constant w) and **U** denotes an upsampling operator. Multiplication of the image by the w-modulation term spreads the Fourier modes of the image. To avoid aliasing, we follow the approach taken by Wiaux et al. (2009b) and McEwen & Wiaux (2011) and introduce an upsampling operator **U**, which increases the supported band-limit by zero-padding the image in Fourier space. The noise n models the instrumental noise of the telescope. In this article, we do not consider visibility measurements at continuous coordinates, thus we neglect the gridding operator that would be required to model this setting; we leave a further investigation of the spread spectrum effect in the setting of continuous visibilities to future work.

2.3   Solving the interferometric inverse problem

Eq. (6) can be solved for the image x using techniques from compressive sensing. The theory of compressive sensing provides a flexible framework for sparse signal recovery (Candes & Tao 2006; Candes et al. 2006; Donoho 2006; Candes & Wakin 2008). Compressive sensing techniques have been applied effectively to radio interferometric imaging in simulated settings (*e.g.* Wiaux et al. 2009a,b; McEwen & Wiaux 2011; Carrillo et al. 2012), where reconstruction quality has been shown to be superior to conventional imaging algorithms.

Compressive sensing techniques are applicable under the presumption that the signal x has a sparse or compressible representation in a dictionary Ψ, for instance wavelets. The columns of Ψ are given by the atoms making up the dictionary. The inverse problem formulated in Eq. (6) can be solved by minimising the sparsity of the signal representation in the dictionary Ψ, subject to a data fidelity constraint. That is, by solving the optimisation problem

$$\min_{\alpha} \|\alpha\|_1 \text{ such that } \|y - \Phi\Psi\alpha\|_2 \leqslant \epsilon. \qquad (7)$$

Recall that the ℓ_1 norm is defined as $\|\alpha\|_1 \equiv \sum_i |\alpha_i|$ and the ℓ_2 norm as $\|\alpha\|_2 \equiv (\sum_i |\alpha_i|^2)^{1/2}$. The constraint ϵ may be related to a residual noise level estimator (see, *e.g.*, Wiaux et al. 2009a; Carrillo et al. 2012). The optimisation problem specified by Eq. (7) can be solved using convex optimisation methods, such as Douglas-Rachford splitting (Combettes & Pesquet 2007).

In addition to the standard synthesis problem specified by Eq. (7), the alternative analysis framework and other variants may also be considered. For example, average sparsity in the analysis setting, with an additional reweighting scheme, was considered by Carrillo et al. (2012), yielding the SARA algorithm, which was shown to provide excellent reconstruction fidelity for interferometric imaging.

For accurate reconstruction it is essential that the measurement basis, in case of radio interferometry essentially the Fourier basis, is incoherent with the sparsity dictionary. The incoherence of the bases is proportional to the inverse of the maximum of the inner products of measurement and sparsity vectors (see, *e.g.*, Candes & Romberg 2007). Consequently, Dirac and Fourier bases are maximally incoherent. For radio interferometry, the Dirac basis (*i.e.* pixel basis) is thus essentially optimal in terms of coherence for signal reconstruction. However, the Dirac basis may not necessary provide a sparse representation of the underlying image. In practice, there is therefore often a trade-off between coherence and sparsity.

This trade-off can be mitigated by the spread spectrum effect (Wiaux et al. 2009b). Since the measurement basis can essentially be identified with the Fourier basis for radio interferometry, coherence is given by the maximum modulus of the Fourier coefficient of the sparsity atoms. The w-modulation corresponds to a norm-preserving convolution in Fourier space, spreading the spectrum of the sparsity atoms, thus reducing the maximum modulus of their Fourier coefficients and increasing incoherence. The increased incoherence due to this spread spectrum effect acts to improve the fidelity of image reconstruction. The universality of the spread spectrum effect has been demonstrated by Wiaux et al. (2009b), such that optimal coherence can be reached (for Gaussian sources) in the limiting case of a large w-component.

3   SPARSE VARIANT OF W-PROJECTION: THEORY AND IMPLEMENTATION

In this section we describe how to support varying w-components in the radio interferometric imaging framework outlined in the preceeding section. Firstly, we describe the idea of the w-projection algorithm (Cornwell et al. 2008) and explain how it is incorporated in our imaging framework. Secondly, an adaptive sparsification procedure is presented to recover a sparse matrix approximation of the Fourier representation of the w-modulation kernel. This approach reduces the memory requirements and computation-time of solving the interferometric imaging problem considerably.

3.1 w-projection

The operator description of Eq. (6) is only valid for a w-modulation with constant w, *i.e.* a constant w-component. For a first examination of the spread spectrum effect this approach is convenient since it reduces the computational burden of a quantitative study substantially; this is the approach that was taken by Wiaux et al. (2009b). In this work, we simulate more realistic survey settings where we include varying w-components in the imaging process. To do this we apply a variant of the w-projection algorithm (Cornwell et al. 2008).

By applying the Fourier convolution theorem to Eq. (6), the w-modulation operator **C** may be moved through the Fourier transform **F**, so that it is applied as a convolution in visibility space, expressed by the operator **Ĉ**. Each row of the matrix operator **Ĉ** is given by the (shifted) Fourier transform of the real space representation of the w-modulation $C(l, m; w)$. Naively, expressing the application of the w-modulation in this manner is computationally less efficient that the original formulation given by Eq. (6). However, applying the w-modulation in this manner via the **Ĉ** operator has two important advantages. Firstly, a different w-component may be used to compute each row of **Ĉ**; hence, a different w-component can be considered when computing the visibility corresponding to a particular (u, v) component, while still exploiting the computational accessibility of a two-dimensional FFT to implement **F**. Secondly, many of the elements of **Ĉ** will be close to zero; hence, a sparse matrix approximation of **Ĉ** can be used to speed up its computational application and reduce memory requirements significantly.

Following this formulation, the radio interferometric measurement equation to be solved may be written

$$y = \hat{\mathbf{C}}\mathbf{F}\mathbf{U}x + n = \Phi'x + n. \qquad (8)$$

where the other operators remain unchanged compared to Eq. (6).

The masking operator **M** is incorporated in $\hat{\mathbf{C}}$, where each row corresponds to one measured visibility. The resulting operator hence has dimension number of measured visibilities M by number of pixels of the image N^2 (where we consider $N \times N$ pixel images). We adopt this new description of the measurement operator $\mathbf{\Phi}'$ when solving the radio interferometric optimisation problem. Note that the convex optimisation algorithms highlighted in Sec. 2.3 are transparent to the measurement operator $\mathbf{\Phi}$ and only require the application of the operator and its adjoint; hence, the w-projection algorithm may be incorporated into the compressed sensing framework easily.

In our description of the w-projection algorithm, every (u, v) visibility component can have a different w-component. Contrast this to the approach taken typically, where, for computational reasons, the w-projection is performed on a relatively small number of w-planes. We do not require such a restriction in our approach since we address computational issues by making a fast and accurate approximation of the application of the $\hat{\mathbf{C}}$ operator, as described subsequently.

3.2 Fast approximation of the w-modulation operator $\hat{\mathbf{C}}$

For every visibility component (u, v, w), the w-modulation kernel is computed in real space via Eq. (5). Then the kernel is Fourier transformed via a two-dimensional FFT. The resulting convolution kernel is collapsed into a vector which defines one (shifted) row of the w-modulation operator $\hat{\mathbf{C}}$. In Fig. 1, some exemplary rows of the absolute value of the complex convolution operator $\hat{\mathbf{C}}$ are shown, where we quantify the amount of modulation by the variable w_d, which is simply a rescaling of w. A value $w_\mathrm{d} = 0$ implies no modulation, while $w_\mathrm{d} = 1.0$ corresponds to maximal modulation (the explicit definition of w_d is given in Sec. 4.1). As expected, it can be seen that the support and amplitude of the w-modulation kernel in Fourier space depends on w_d (since w is related to its instantaneous frequency). Furthermore, the absolute value of a large number of the entries of rows of the convolution operator are very close to zero, particularly for small w-modulation.

In order to speed up computational performance and reduce memory requirements, we sparsify the w-modulation operator $\hat{\mathbf{C}}$. Since the Fourier support of the w-modulation kernel varies for different w_d, it is not efficient to *a priori* limit the support of a row of the matrix operator to a constant size. Instead, the number of non-zero elements of each row of $\hat{\mathbf{C}}$ should be adaptive to w_d. Furthermore, we do not presume to know *a priori* the location of the significant elements of $\hat{\mathbf{C}}$; thus, the location of non-zero elements should also be chosen adaptively. We therefore recover a sparse approximation of $\hat{\mathbf{C}}$ by applying a different threshold for each row of $\hat{\mathbf{C}}$: all entries less than the threshold (in absolute value) are set identically to zero. The threshold is chosen in such way that a certain fraction of the total energy of the w-modulation kernel is preserved. The total energy is calculated by summing the squared elements of one row of the w-modulation kernel. The threshold of each matrix row is determined by considering the preserved energy distribution as a function of the threshold. We invert this relation in a computationally efficient way by using a bisecting method to determine the threshold for a specified level of preserved energy. This is a very general approach to the sparsification of $\hat{\mathbf{C}}$ since we do not make any assumption about the structure of each row; hence, our strategy is applicable for all types of direction-dependent effects and not only w-modulation.

Following this strategy, we plot in Fig. 2 examples of the two-dimensional w-modulation kernel in real space for different values of w_d and different preserved energy levels E. It can be clearly seen that a lower preserved energy proportion results in a greater distortion to the w-modulation kernel in real space. However, for a preserved energy level of $E = 0.75$ (shown in the third column of Fig. 2), the distortion is moderate.

We study the relative sparsity of the sparse approximation of $\hat{\mathbf{C}}$ by plotting in Fig. 3 its proportion of non-zero entries. The size of the initial matrix operator $\hat{\mathbf{C}}$ with dimensions $M \times N^2$ corresponds to 100 percent of non-zero entries. Sparsifying the matrix with a preserved energy of $E = 0.999$ already decreases the proportion of non-zero entries to ~ 20 percent. This case effectively corresponds to no approximation and shows the sparse nature of the exact w-modulation kernel in Fourier space. Applying lower energy levels reduces the number of non-zero entries of $\hat{\mathbf{C}}$ further, to 7 percent for $E = 0.9$ and to 0.6 percent for $E = 0.1$.

In the following, we address the question to what degree the approximation degrades reconstruction quality. In order to do so, we simulate full image reconstructions. The experimental set-up of these simulations is described in detail in Sec. 4, where a thorough study of the effect of w-modulation is performed. Here we study the impact of the sparse approximation of the matrix operator $\hat{\mathbf{C}}$ only, in order to select a suitable preserved energy proportion E for the subsequent analysis performed in Sec. 4.

We vary the preserved energy level E used to construct the sparse approximation of $\hat{\mathbf{C}}$ and examine the image reconstruction quality of the two test images shown in Fig. 6 (these images and the experiment set-up will be introduced in detail in Sec. 4.1). We measure the effectiveness of the sparse matrix approximation of $\hat{\mathbf{C}}$ through the signal-to-noise ratio (SNR) of the reconstructed image (defined explicitly in Sec. 4.1) and the runtime of image reconstruction relative to the runtime of a full energy reconstruction, both as a function of the preserved energy level E. The SNR of the reconstructed images as a function of preserved energy proportion is shown in Fig. 4 for visibility coverages of 10 percent and 50 percent, where w-components vary. It can be seen clearly that the quality of reconstruction remains close to the case where no sparsification is made ($E = 1$) for relatively low energy proportions. Below $E \sim 0.5$ for 10 percent coverage and below $E \sim 0.7$ for 50 percent coverage, reconstruction quality slowly starts to decrease.

The relative runtime of image reconstruction implemented in MATLAB and run on a state-of-the-art 6 core machine is plotted in Fig. 5. It can be seen clearly that the runtime is significantly decreased by reducing E. For example, for M31 with 10 percent visibility coverage, the computation-time is already reduced from 4h12min for $E = 1$ to 60min for $E = 0.999$ (where effectively no approximation of $\hat{\mathbf{C}}$ is made). Reducing the preserved energy proportion further, computation-time is reduced to 29min for $E = 0.9$, and to 21min for $E = 0.6$ where reconstruction quality is still not markedly degraded.

The computational cost of image reconstruction is related closely to the asymptotic complexity of application of the measurement operator, which we consider next for various algorithms. In the case each visibility has a different w-component, if the w-projection algorithm is not applied, then there is no advantage in applying an FFT. In this setting one would apply a discrete Fourier transform; thus, application of the measurement operator would be of order $O(MN^2)$. If w-projection is applied, using an FFT but without any sparsification of $\hat{\mathbf{C}}$, then the complexity becomes $O(MN^2 + N^2 \log N)$, where the first term corresponds to multiplication by a dense matrix $\hat{\mathbf{C}}$ of size $M \times N^2$ and the second to the FFT. By sparsifying $\hat{\mathbf{C}}$, as described above, then the complexity is reduced to $O(S + N^2 \log N)$, where S specifies the number of non-
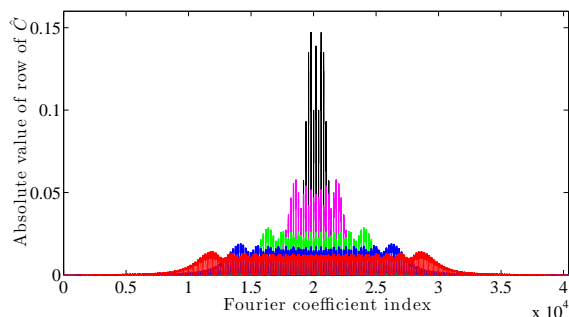
Figure 1. Rows of the *w*-modulation kernel $\hat{\mathbf{C}}$ for different values of $w_d = \{0.10, 0.25, 0.50, 0.75, 1.00\}$, plotted in black, magenta, green, blue and red, respectively. The number of significant kernel entries and their support depends strongly on the level of *w*-modulation, *i.e.* on the value of w_d. Consequently, the approximate kernel support should be chosen adaptively for each w_d.
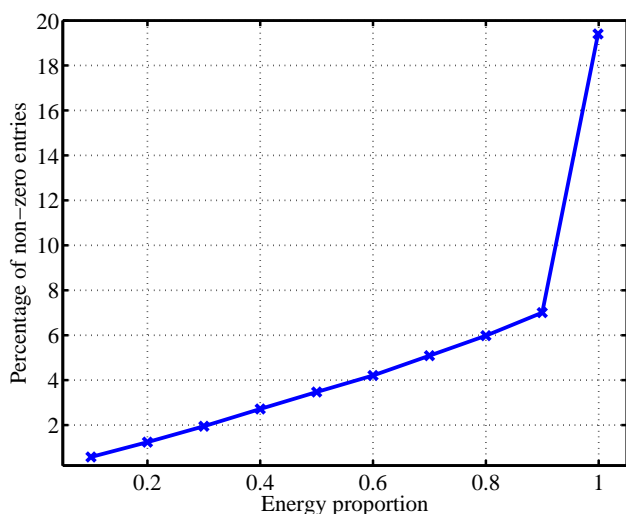

Figure 3. Percentage of non-zero entries of the approximate *w*-modulation kernel $\hat{\mathbf{C}}$ as a function of preserved energy proportion. Note that the point near $E = 1$ is in fact $E = 0.999$, highlighting the fact that $\hat{\mathbf{C}}$ is sparse even when effectively no approximation is made. Applying lower energy levels reduces the number of non-zero entries of $\hat{\mathbf{C}}$ further.

zero entries of $\hat{\mathbf{C}}$. Sparsifying $\hat{\mathbf{C}}$ can thus reduce the computational cost of imaging considerably, as also demonstrated by the numerical experiments performed above.

In our subsequent analysis, we take an extremely conservative approach and select a preserved energy level of $E = 0.75$. This is sufficient to render a study of the spread spectrum effect for varying *w*-component computationally tractable in the experimental setting described in Sec. 4.1, while we can also be confident that reconstruction quality is not degraded. For future applications, the preserved energy proportion could be further reduced, provided that the implications for image reconstruction are considered carefully.

4 ADVANTAGE OF *W*-COMPONENTS: SIMULATIONS AND RESULTS

In this section we study the impact of the spread spectrum effect for various *w*-modulation scenarios, including constant and varying *w*-components, using extensive numerical simulations. Firstly, the
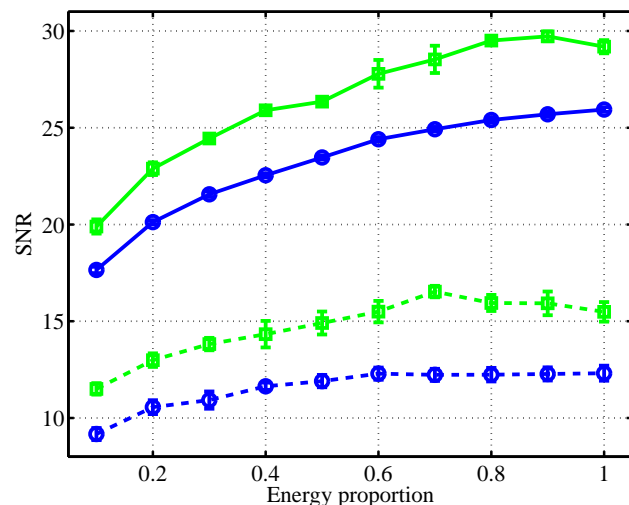

Figure 4. Reconstruction quality of M31 (green lines marked with squares) and 30Dor (blue lines marked with circles) as a function of preserved energy proportion of the approximate *w*-modulation operator $\hat{\mathbf{C}}$. Daubechies 8 wavelets are used as the sparsity basis, while a visibility coverage of 10 percent is considered in the dashed lines and a coverage of 50 percent is considered in the solid lines (see text of Sec. 4 for further details). The data points are the mean of 10 simulations and the error bars show one-standard-deviation. Notice that the level of preserved energy can be reduced considerably without markedly reducing reconstruction quality.

experimental parameters of our simulations are outlined. Secondly, image reconstruction results are presented and discussed for the different *w*-modulation scenarios, where we also consider different images and choices of sparsifying dictionaries.

4.1 Experimental set-up

We investigate the effect of *w*-modulation on reconstruction quality using two test images: an HII region in M31 and 30 Doradus (30Dor) in the Large Magellanic Cloud, as shown in Fig. 6. We choose two images to demonstrate reconstruction quality on images with different spatial distribution characteristics. For computational reasons, we downsample the original test images[3] (yielding an $N \times N = 125 \times 125$ pixel image for M31 and an $N \times N = 127 \times 127$ pixel image for 30Dor).

The opening angle of the FoV is chosen constant throughout our analysis at $\theta = 2°$, with corresponding planar FoV $L = 2\sin(\frac{\theta}{2})$. Note, however, that our analysis is insensitive to the choice of θ since different choices just rescale the visibility coordinates (u, v, w).

The band-limit B of a pixelated image is related to its FoV L and the number of pixels along each dimension N via $B = \frac{N}{2L}$. The band-limit of the linear chirp approximation of the *w*-modulation, as specified by Eq. (5), is given approximately by its instantaneous frequency: $B^{\text{Chirp}} \simeq w_d \frac{N}{2L} = w_d B$, where we introduce the variable $w_d = w\frac{L^2}{N}$ (Wiaux et al. 2009b; McEwen & Wiaux 2011). As commented already, the *w*-modulation can be seen as a convolution in Fourier space; thus, the band-limit of the convolved image is given by $B + B^{\text{Chirp}} \simeq (1 + w_d)B$. We consider the maximal case of $w_d = 1$, where the band-limit is doubled. In order to support the spreading due to the maximal *w*-component, it is necessary to double the

---

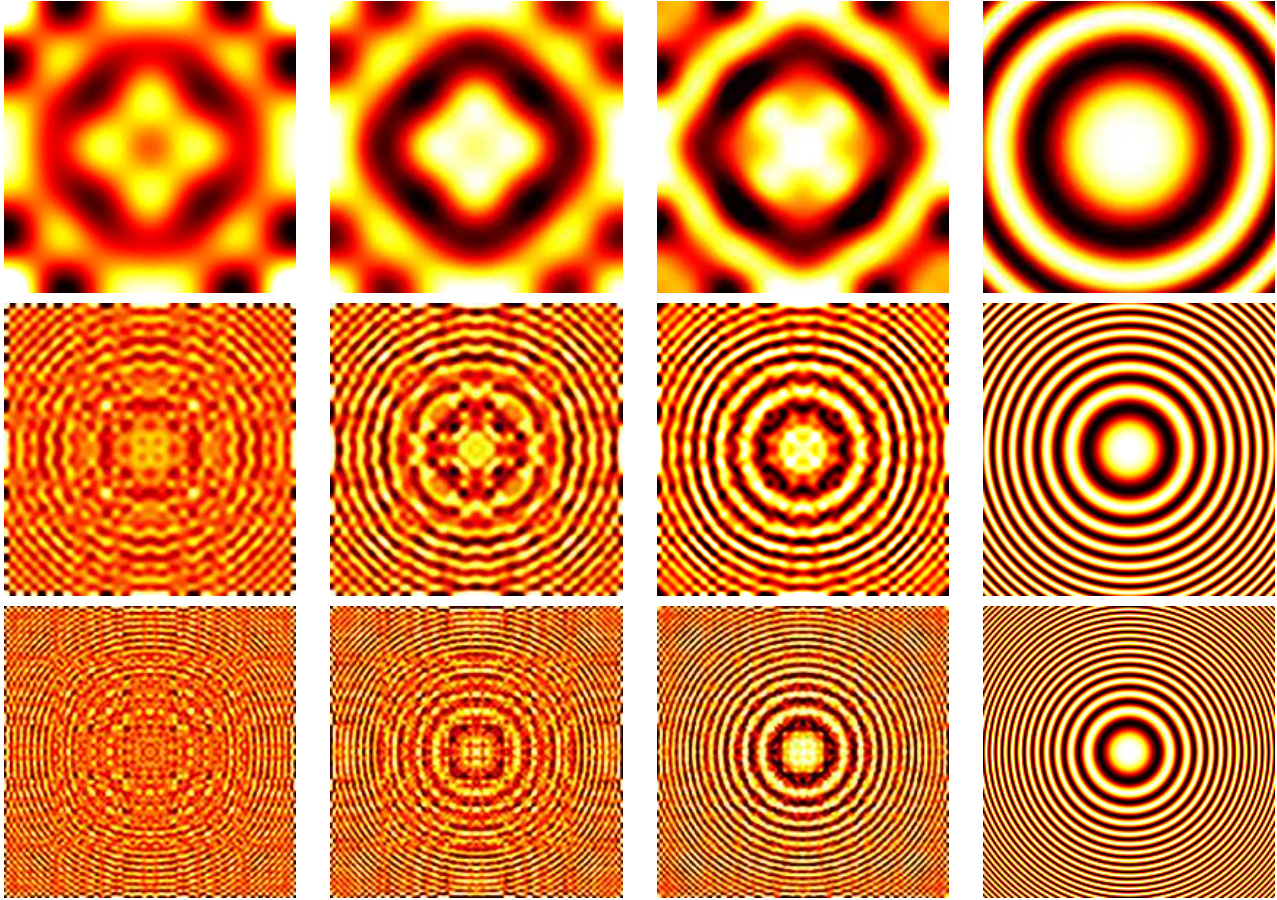[3] Available at http://casaguides.nrao.edu/index.php.

Figure 2. The (real part of the) w-modulation kernel plotted in real space. For each row, a different w_d is considered: $w_d = 0.1$ for the upper row; $w_d = 0.5$ for the second row; and $w_d = 1.0$ for the lower row. Notice the higher frequency content of the w-modulation kernel with increasing w_d, *i.e.* from highest to lowest row. For each column, a different approximation of the w-modulation kernel is plotted, for preserved energy levels $E = \{0.25, 0.50, 0.75, 1.00\}$, respectively, from the left to the right. Notice that the kernels for $E < 1$ more accurately approximate the exact kernel ($E = 1$) as the energy preserved increases, *i.e.* from left-to-right columns.

band-limit, which requires doubling the pixel support of the image. Upsampling of the image by a factor of two is applied, in order to support this increased band-limit, by the application of the operator **U**, which implements zero-padding in Fourier space (as described in Sec. 2.2). In this scenario, the maximal w-component considered is $\frac{2}{L}$ larger than the maximal u or v component considered.

We sample the visibilities on a $2N \times 2N$ grid and reconstruct images from visibility coverages ranging from 10 percent to 90 percent of the full-set of discrete visibilities. We assume uniformly random sampling distributions in the visibility coordinates (u, v, w), in order to disregard considerations due to particular telescope configurations and to perform large numbers of simulations.

We simulate instrumental noise n under the assumption of independent identically distributed Gaussian noise. The width of the Gaussian distribution is calculated via $\sigma(n) = 10^{-\text{SNR}_{\text{in}}/20}\sigma(x)$, where $\sigma(x)$ is the standard deviation of the input image. The input signal-to-noise ratio SNR_{in} is chosen as 30 dB throughout the article. Reconstruction quality is measured by the SNR_{out} of the recovered image, calculated via $\text{SNR}_{\text{out}} = 20 \log_{10}(\sigma(x)/\sigma(x - x_{\text{rec}}))$, where $\sigma(x - x_{\text{rec}})$ is the standard deviation of the difference of original and recovered image. Our results are shown in units of dB and can be easily compared to the input noise level of 30 dB.

4.2   Results

We reconstruct images from simulated visibility measurements for various w-modulation scenarios: no w-component (*i.e.* constant $w_d = 0$); a uniformly random varying w-component (with $w_d \sim \mathcal{U}(0, 1)$); and a constant, maximal w-component (*i.e.* constant $w_d = 1$). Our main focus is to study the more realistic varying w-component scenario, since the impact of the spread spectrum effect for constant w-component has been studied by Wiaux et al. (2009b), where the improvement in reconstruction quality due to a non-zero w-modulation was demonstrated already. Furthermore, we also vary the visibility coverage, image and sparsity dictionary, in order to demonstrate the robustness of our results. The presented results are simulated using an energy threshold of $E = 0.75$ for the sparse approximation of the w-modulation operator.

In our analysis we consider the Dirac and Daubechies 8 wavelet sparsity bases, and also the SARA algorithm (Carrillo et al. 2012). Reconstruction with the Dirac basis can be considered as a proxy for the CLEAN algorithm (Högbom 1974); the close correspondence was shown previously by Wiaux et al. (2009a). Daubechies 8 wavelets can be considered as a proxy for multi-scale CLEAN (Cornwell 2008); the Daubechies 8 wavelet reconstruction here can in fact be considered as an upper bound on the reconstruction quality of multi-scale CLEAN (since Li et al. (2011a)
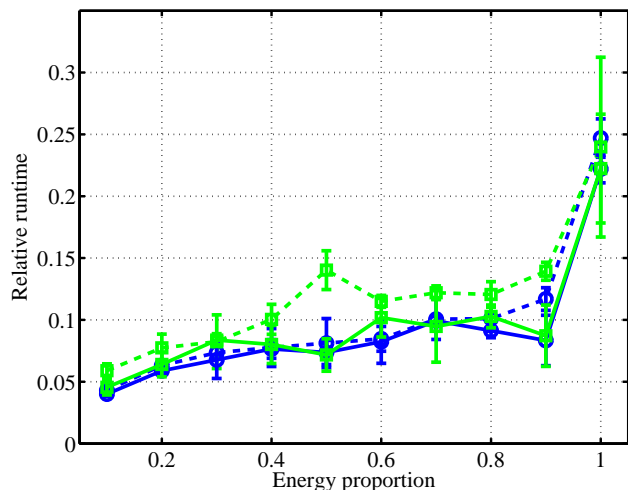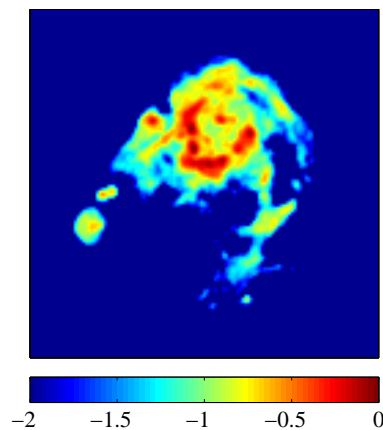
Figure 5. Relative computation-runtime as a function of the preserved energy with visibility coverage of 10 percent in the dashed lines and a coverage of 50 percent in the solid lines. The green curves refer to the simulations using M31 (marked with squares) and the blue curves to 30Dor (marked with circles). The data points are the mean of 10 simulations and the error bars show one-standard-deviation. The point close to $E = 1$ is in fact $E = 0.999$, highlighting the fact that the natural sparsity of $\hat{\mathbf{C}}$ can be exploited even when effectively no approximation is made. Notice that by reducing the level of preserved energy, runtime can be reduced considerably.
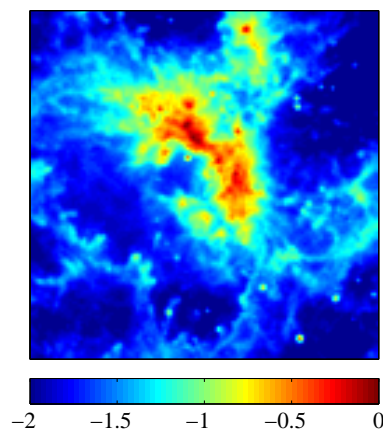
showed ℓ_1 reconstruction using the isotropic undecimated wavelet transform to be superior to multi-scale CLEAN using M31 as a test image, and Carrillo et al. (2012) in turn showed Daubechies 8 wavelets to be superior to the isotropic undecimated wavelet transform for both the test images used in this work). We also consider the SARA algorithm developed recently by Carrillo et al. (2012), which exhibits reconstruction performance substantially superior to both CLEAN- and multi-scale CLEAN-like algorithms.

Firstly, we provide a qualitative comparison of image reconstruction for the various w-modulation scenarios considered. In Fig. 7 and Fig. 8 we show images reconstructed using Daubechies 8 wavelets as sparsity dictionary for different w-modulation scenarios. The first row of images corresponds to no w-modulation, the second to the varying w-component case and the third row shows the reconstructed images with a maximal, constant w, *i.e.* $w_d = 1$. Images in the first column are reconstructed images, while those in the second row are the difference of the reconstructed and original image. The residual images in the third column are given by the difference of the dirty images of the original and those of the reconstruction, calculated via $\boldsymbol{x}_{\text{res}} = \boldsymbol{\Phi}'^{\text{T}}\boldsymbol{\Phi}'\boldsymbol{x}_{\text{rec}} - \boldsymbol{\Phi}'^{\text{T}}\boldsymbol{\Phi}'\boldsymbol{x}$. A relatively low visibility coverage of 10 percent is chosen to illustrate differences between reconstruction scenarios clearly.

In Fig. 7 it can be seen that for M31 the reconstructed image without w-modulation suffers from considerable contamination, while the reconstructed image for maximal w-modulation is relatively free of contamination. Interestingly, the reconstruction for the more realistic scenario of varying w-component (second row) achieves almost the same reconstruction quality as maximal, constant w-component (third row). In Fig. 8, the same recovery behaviour as in Fig. 7 can be seen for 30Dor. Again, the reconstruction for the more realistic scenario of varying w-component achieves almost the same reconstruction quality as maximal w-modulation with constant w.

(a) M31

(b) 30Dor

Figure 6. Original images in logarithmic scale.

Secondly, we provide a more complete quantitative comparison of reconstruction quality for the different w-modulation scenarios, where in addition we consider different sparsity dictionaries and visibility coverages, and average over 10 simulations. In Fig. 9, the SNR_{out} of the reconstructed images are shown as a function of visibility coverage for the two test images.

When using the Dirac basis as the sparsity dictionary, we do not expect any significant difference in reconstruction quality between the various w-modulation scenarios since the measurement basis (essentially Fourier) and the sparsity basis (Dirac) are maximally incoherent. Hence, w-modulation will not further improve the recovery of the image. In Fig. 9(a) and Fig. 9(b), this intuitive understanding is largely confirmed. In the case of M31, shown in Fig. 9(a), the curves for three w-modulation cases are close together and there is no clear tendency that one case performs better than the other over the entire visibility coverage range. In the case of 30Dor, shown in Fig. 9(b), the case of a maximal, constant w-component seems to perform marginally better than the other two cases; nevertheless, reconstruction quality remains broadly similar.

For Daubechies 8 wavelets and the SARA algorithm on the other hand, we expect an improvement in reconstruction quality due to the w-modulation since the sparsity dictionaries are no longer maximally incoherent with the Fourier basis. When using Daubechies 8 wavelets as the sparsity basis, reconstruction quality is improved in the presence of w-modulation, as apparent in Fig. 9(c) and Fig. 9(d). Moreover, the curve for the vary-
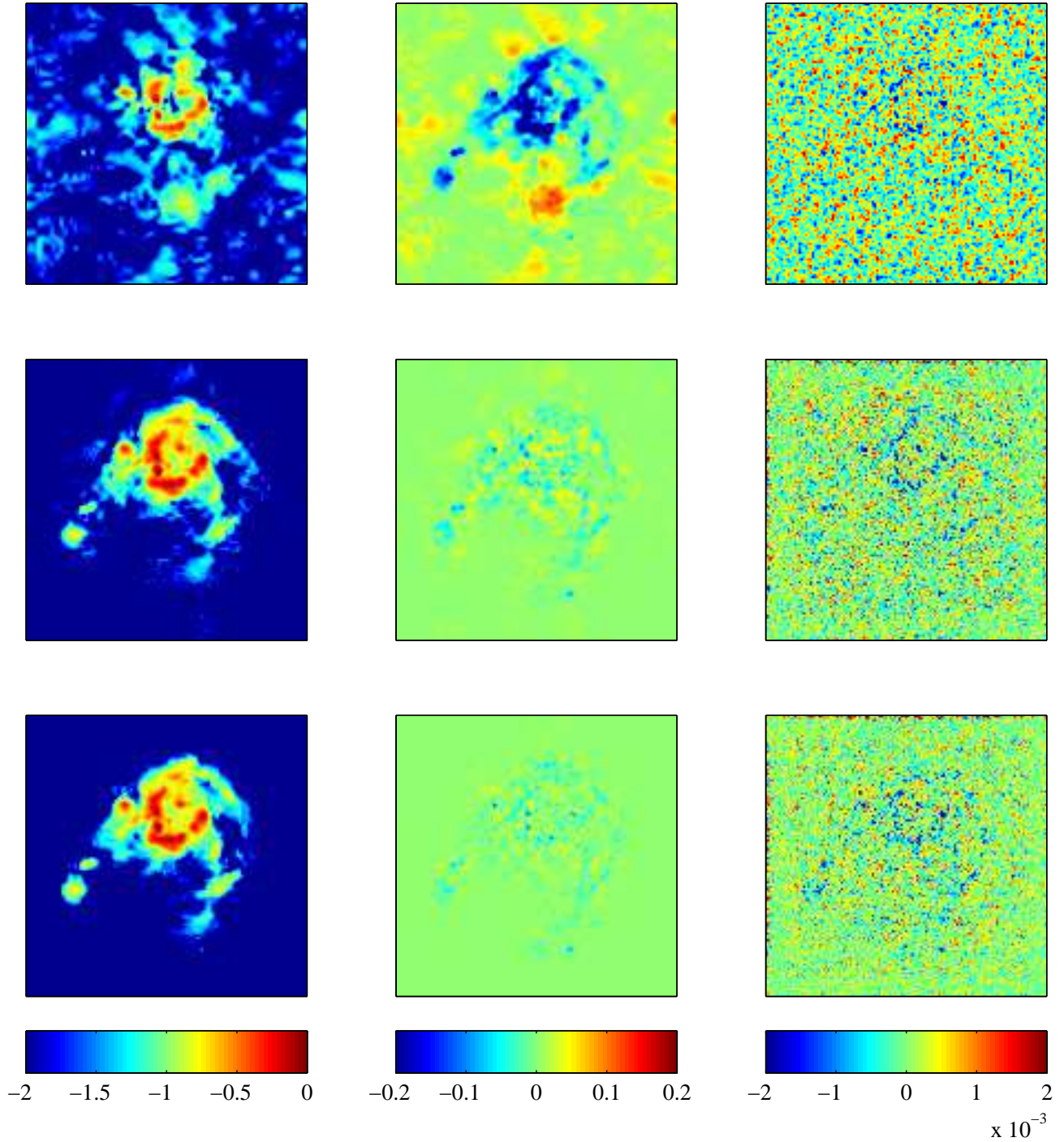
Figure 7. Image reconstruction of M31 using Daubechies 8 wavelets with visibility coverage of 10 percent. The input noise level is 30 dB and the energy threshold is chosen as $E = 0.75$. Each row shows a reconstructed image in $\log_{10}$ scaling (left panel), and in linear scaling the difference of the original and the reconstructed image (middle panel) and the residual image (right panel). The first row corresponds to no w-modulation (*i.e.* $w_d = 0$), the second one to the uniformly random w-component case (*i.e.* $w_d \sim \mathcal{U}(0, 1)$), and the third row to maximal, constant w-component (*i.e.* $w_d = 1$). The $\mathrm{SNR_{out}}$ of the reconstructed images is, respectively, $\mathrm{SNR_{out}} = 5$ dB for no w-modulation, $\mathrm{SNR_{out}} = 16$ dB for varying w-component, and $\mathrm{SNR_{out}} = 19$ dB for maximal, constant w-component. Clearly, the quality of reconstruction for the varying w-component scenario (second row) is very close to the reconstruction quality for maximal, constant w-component (third row).
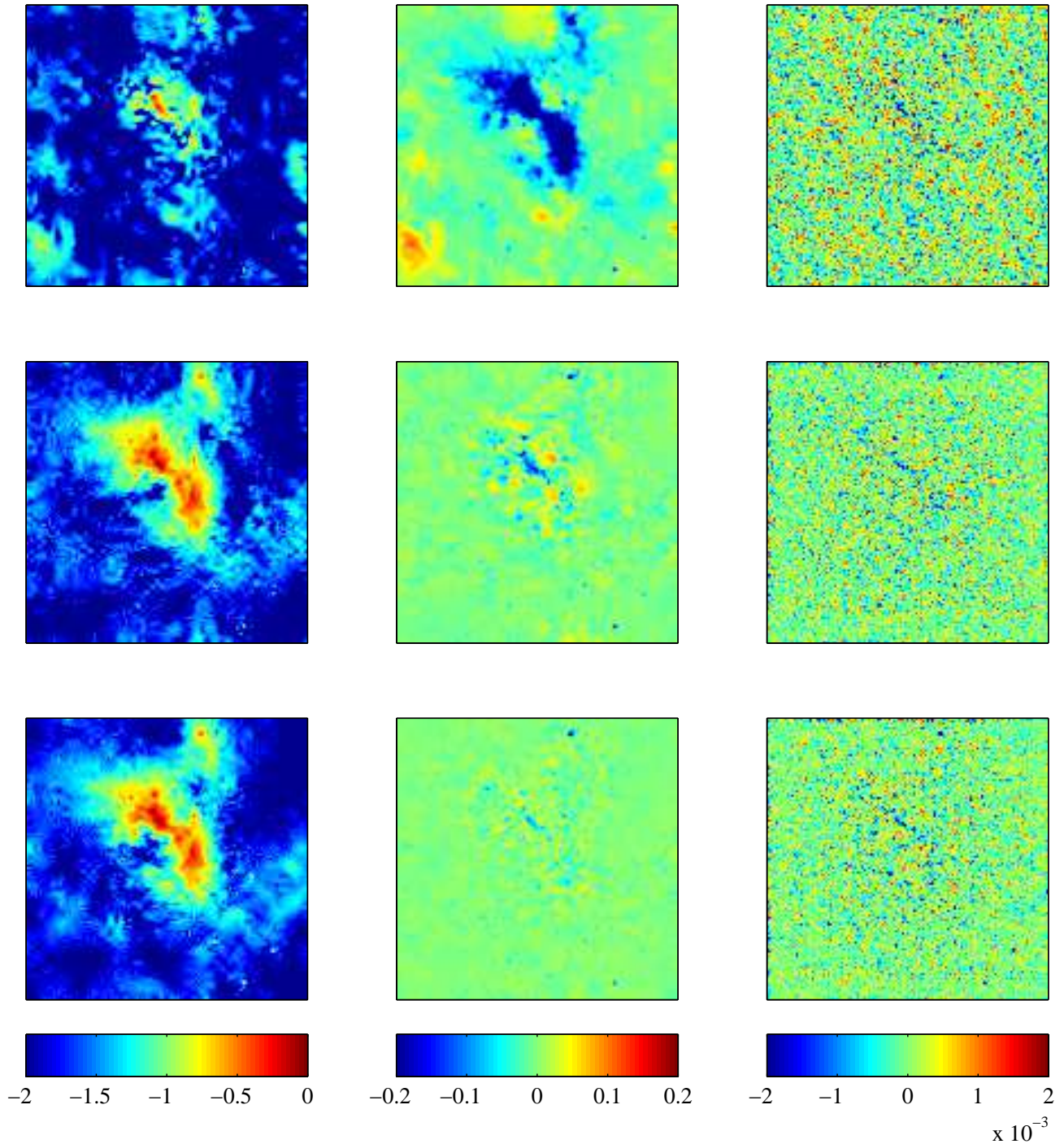
Figure 8. Image reconstruction of 30Dor using Daubechies 8 wavelets with visibility coverage of 10 percent. The input noise level is 30 dB and the energy threshold is chosen as $E = 0.75$. Each row shows a reconstructed image in $\log_{10}$ scaling (left panel), and in linear scaling the difference of the original and the reconstructed image (middle panel) and the residual image (right panel). The first row corresponds to no w-modulation (*i.e.* $w_{\rm d} = 0$), the second one to the uniformly random w-component case (*i.e.* $w_{\rm d} \sim \mathcal{U}(0,1)$), and the third row to maximal, constant w-component (*i.e.* $w_{\rm d} = 1$). The $\rm SNR_{out}$ of the reconstructed images is, respectively, $\rm SNR_{out} = 2$ dB for no w-modulation, $\rm SNR_{out} = 12$ dB for varying w-component, and $\rm SNR_{out} = 15$ dB for maximal, constant w-component. Clearly, the quality of reconstruction for the varying w-component scenario (second row) is very close to the reconstruction quality for maximal, constant w-component (third row).

ing w-component case is very similar to the maximal, constant w-component case. For M31 in Fig. 9(c), the varying and maximal, constant w-component cases exhibit improved reconstruction quality for low visibility coverages. At coverages above ~ 50 percent, performance approaches a saturation level approximately equal to the input SNR of 30 dB. For 30Dor in Fig. 9(d), the varying and maximal, constant w-component cases show superior reconstruction quality to the absence of any w-modulation over a larger range of visibility coverages, until saturation at approximately the input SNR is again reached. Furthermore, the varying w-component and the maximal, constant w-component scenarios exhibit very similar reconstruction quality. When using the SARA algorithm, the same features are observed as with Daubechies 8 wavelets, as shown in Fig. 9(e) and Fig. 9(f). Again, it is interesting to note that the varying w-component and the maximal, constant w-component scenarios exhibit very similar reconstruction quality. The reconstruction quality with SARA is in general improved compared to Daubechies 8 and Dirac sparsity basis, particularly for very low visibility coverages, as demonstrated by Carrillo et al. (2012) previously.

Fig. 9 shows that reconstruction quality is improved considerably by the spread spectrum effect, caused by the varying w-component, for cases where the sparsity dictionary is not the Dirac basis. For M31 when using Daubechies 8 wavelets (Fig. 9(c)), for example, the same SNR achieved by a scenario without w-modulation for a visibility coverage of 40 percent can be achieved by 20 percent coverage when the spread spectrum effect is considered. For 30Dor (Fig. 9(d)), the effect can be seen more significantly where 60 percent visibility coverage is necessary to accomplish the same SNR as 20 percent coverage when including w-modulation. Similar results are seen for the SARA algorithm although saturation about the input SNR is reached quickly for M31. If the non-coplanar baseline and wide FoV setting is thus modeled accurately, the same image reconstruction quality can be achieved with considerably fewer baselines due to the spread spectrum effect.

5 CONCLUSIONS

We have revisited the spread spectrum effect for non-coplanar baseline and wide FoV settings in radio interferometric imaging. We confirmed previous studies (but with more realistic images and alternative sparsifying dictionaries), showing that for w-modulation with constant w-component, image reconstruction quality in the context of compressed sensing is enhanced compared to the setting where there is no w-modulation (*i.e.* $w = 0$). Extending this analysis to the more realistic scenario of varying w-component is computationally demanding and is the main focus of the current article. We developed a variant of the w-projection algorithm to support different w-components for each visibility measurement, which is based on an sparsification procedure that adapts to the support of each kernel. Our algorithm is generic and thus suitable for all types of direction-dependent effects. This approach renders a study of the spread spectrum effect in the varying w-component setting computationally tractable. Through extensive numerical experiments, we find that the image reconstruction quality achieved for the varying w-component case outperforms the quality achieved in the absence of w-modulation, and is nearly as good as in the case of constant, maximal w-component. The enhancement is particularly clear for very low visibility coverages of order a few percent.

If the non-coplanar baseline and wide FoV setting is thus modeled accurately, the same image reconstruction quality can be achieved with considerably fewer baselines due to the spread spectrum effect. Alternatively, for a given number of baselines, reconstruction quality is improved by the spread spectrum effect. This confirms that one may seek to optimise future telescope configurations to promote large w-components, thus enhancing the spread spectrum effect and consequently the fidelity of image reconstruction. Similar suggestions to optimise future telescope configurations to promote large w-components have also been made by Carozzi (2012) recently, but from an information theoretic point-of-view.

In a recent separate work, the scenario is considered where visibilities are measured at continuous (u, v, w) components (Carrillo et al. 2013), as is the case for real interferometric telescopes. After integrating these approaches it will be possible to apply our methods and to study the impact of the spread spectrum effect on observations made by real radio interferometric telescopes.

REFERENCES

Bhatnagar S., Cornwell T.J., Golap K., Uson J.M., 2008, Astron. & Astrophys., 487, 419, arXiv:0805.0834

Candes E., Romberg J., 2007, Inverse Problems, 23, 3, 969, ISSN 0266-5611

Candes E., Romberg J., Tao T., 2006, Information Theory, IEEE Transactions on, 52, 2, 489, ISSN 0018-9448

Candes E., Tao T., 2006, Information Theory, IEEE Transactions on, 52, 12, 5406, ISSN 0018-9448

Candes E., Wakin M., 2008, Signal Processing Magazine, IEEE, 25, 2, 21, ISSN 1053-5888

Carilli C.L., Rawlings S., 2004, New Astron.Rev., astro-ph/0409274

Carozzi T.D., 2012, in *Proceedings of the International BASP Frontiers Workshop 2013, Information capacity of radio interferometers*

Carozzi T.D., Woan G., 2009, Mon. Not. Roy. Astron. Soc., 395, 1558, arXiv:0812.0141

Carrillo R., McEwen J.D., Wiaux Y., 2012, Mon.Not.Roy.Astron.Soc., 426, 1223, arXiv:1205.3123

Carrillo R.E., McEwen J.D., Wiaux Y., 2013, submitted to Mon.Not.Roy.Astron.Soc., arXiv:1307.4370

Combettes P.L., Pesquet J.C., 2007, IEEE Journal of Selected Topics in Signal Processing, 1, 564

Condon J.J., Cotton W., Greisen E., Yin Q., Perley R., et al., 1998, Astron.J., 115, 1693

Cornwell T. J.; Perley R.A., 1992, Astronomy and Astrophysics, 261, 353

Cornwell T.J., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793, arXiv:0806.2228

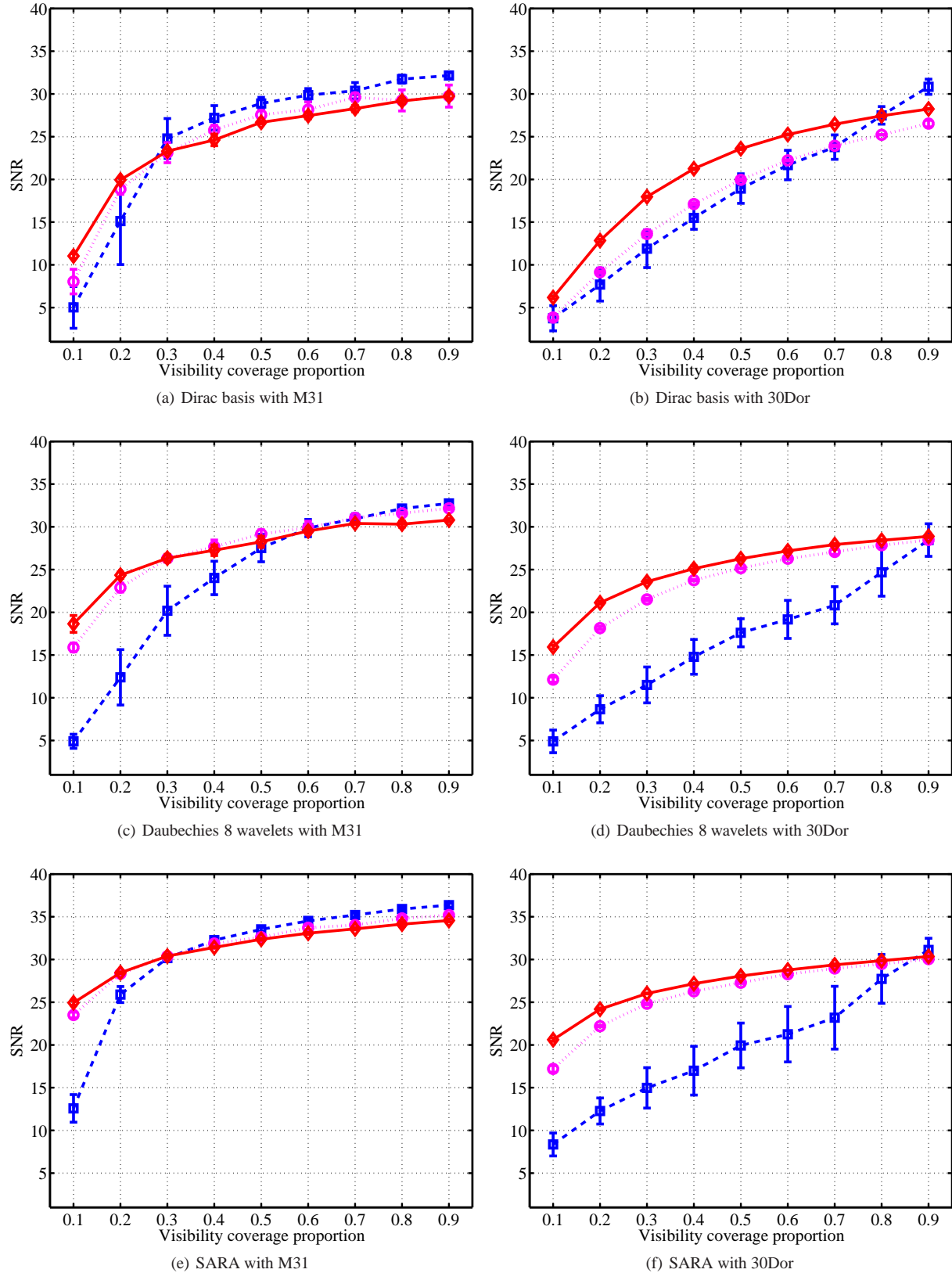**Figure 9.** Signal-to-noise ratios of the recovered image for no w-modulation (blue, dashed lines marked with squares), a varying w-component (magenta, dotted lines marked with circles) and a maximal, constant w-component (red, solid lines marked with diamonds) as a function of visibility coverage. The input noise level is 30 dB and the energy threshold is chosen as $E = 0.75$. The data points are the mean of 10 simulations and the error bars show one standard error. For the Dirac basis in the first row (which is maximally incoherent with the Fourier basis) reconstruction quality remains broadly similar for the various w-modulation scenarios, as expected. The quality of reconstructed images using Daubechies 8 wavelets (second row) and SARA (third row) with constant and varying w-modulation is higher than that without w-modulation. The varying w-component performance in these cases is very close to the maximal, constant w-component.

Cornwell T.J., Golap K., Bhatnagar S., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 647

Cornwell T.J., Voronkov M.A., Humphreys B., 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8500 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, arXiv:1207.5861

Donoho D., 2006, IEEE Trans. Inform. Theory, 52, 1289

Högbom J.A., 1974, A&AS, 15, 417 A&AS, 15, 417

Li F., Brown S., Cornwell T.J., de Hoog F., 2011a, Astron. & Astrophys., 531, A126, arXiv:1106.1709

Li F., Cornwell T.J., de Hoog F., 2011b, Astron. & Astrophys., 528, A31, arXiv:1106.1711

McEwen J.D., Scaife A., 2008, arXiv:0803.2165

McEwen J.D., Wiaux Y., 2011, Mon. Not. Roy. Astron. Soc., 413, 1318, arXiv:1010.3658

Mueller H., 1948, J. Opt. Soc. Am., 36, 661

Smirnov O.M., 2011, Astron. & Astrophys., 527, A106, arXiv:1101.1764

Suksmono A., 2009, 110, arXiv:0904.0508

Tasse C., van der Tol S., van Zwieten J., van Diepen G., Bhatnagar S., 2013, Astron. & Astrophys., 553, A105, arXiv:1212.6178

Thompson A.R., Moran J.M., Swenson G.W.J., 2004, *Interferometry and Synthesis in Radio Astronomy*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Wenger S., Magnor M., Pihlström Y., Bhatnagar S., Rau U., 2010, Publications of the Astronomical Society of the Pacific (PASP), 122, 897, 1367

Wiaux Y., Jacques L., Puy G., Scaife A.M.M., Vandergheynst P., 2009a, Mon. Not. Roy. Astron. Soc., 395, 1733, arXiv:0812.4933

Wiaux Y., Puy G., Boursier Y., Vandergheynst P., 2009b, Mon. Not. Roy. Astron. Soc., 400, 1029, arXiv:0907.0944

Wiaux Y., Puy G., Vandergheynst P., 2010, Mon. Not. Roy. Astron. Soc., 402, 2626, arXiv:0908.4179