Aperture-synthesis radar imaging with compressive sensing for ionospheric research

D. L. Hysell[1], P. Sharma[2], M. Urco[3], and M. A. Milla[4]

[1]Earth and Atmospheric Sciences, Cornell

University, Ithaca, NY, USA

[2]Electrical and Computer Engineering,

Cornell University, Ithaca, NY, USA

[3]Leibniz Institute for Atmospheric Physics,

Kuehlungsborn, Germany

[4]Jicamarca Radio Observatory, Lima, Peru

Key Points.

  1. Compressed sensing inverse methods applied to aperture synthesis radar imaging of ionospheric plasma density irregularities. 2. Performance of basis pursuit denoising (BPDN) and orthogonal matching pursuit (OMP) generally inferior to that of maximum-entropy method (MaxENT). 3. Computational speed of OMP is attractive and prompts research into more suitable function library.

Inverse methods involving compressive sensing are tested in the application of two-dimensional aperture-synthesis imaging of radar backscatter from field-aligned plasma density irregularities in the ionosphere. We consider basis pursuit denoising (BPDN), implemented with the FISTA algorithm, and orthogonal matching pursuit (OMP) with a wavelet basis in the evaluation. These methods are compared with two more conventional optimization methods rooted in entropy maximization (MaxENT) and adaptive beamforming (Linear Constrained Minimum Variance or LCMV or often "Capon's Method.") Synthetic data corresponding to an extended ionospheric radar target are considered. We find that MaxENT outperforms the other methods in terms of its ability to recover imagery of an extended target with broad dynamic range. FISTA performs reasonably well but does not reproduce the full dynamic range of the target. It is also the most computationally expensive of the methods tested. OMP is very fast computationally but prone to a high degree of clutter in this application. We also point out that the formulation of MaxENT used here is very similar to OMP in some respects, the difference being that the former reconstructs the logarithm of the

19 image rather than the image itself from basis vectors extracted from the obser-

20 vation matrix. MaxENT could in that regard be considered a form of compres-

21 sive sensing.

1. Introduction

Aperture-synthesis methods have been used in radio astronomy since the 1950s to form images of distant radio sources from spaced-receiver interferometry data. The signals from the receivers are stochastic and exhibit jointly-normal multivariate statistics. All of the information is therefore contained in the second-order statistics or the spatial covariances or "visibilities." The transformation between the measured visibility data and the desired brightness function, which specifies the radiation intensity versus bearing, is linear and closely related to a Fourier transform [*Thompson*, 1986]. Aperture-synthesis methods are inverse methods for computing the transformation on the basis of sparse and noisy data. The problem is generally underdetermined and poorly conditioned.

Only fairly recently have aperture synthesis methods been applied to radar observations of the upper atmosphere and ionosphere (see e.g. *Kudeki and Sürücü* [1991]; *Hysell* [1996]; *Hysell et al.* [2002, 2004]; *Saito et al.* [2006]; *Hysell et al.* [2008]; *Saito et al.* [2008]; *Sommer and Chau* [2016]; *Urco et al.* [2018a, b]). The problem is similar to the one in radio astronomy with a few important differences. For one, the number of receivers used in upper-atmospheric radar applications has so far been relatively small, spaced-receiver imaging being mainly an afterthought in radar design. For another, the required cadence of measurements in upper-atmospheric radar applications can be very high. Images typically need to be formed in tens of Doppler frequency bins and in hundreds or thousands of range gates at an experimental cadence of about once per second. This makes computational efficiency critical.

Most crucially, whereas astronomical images are typically characterized by point sources, radar targets in the upper atmosphere and ionosphere tend to be spatially extended and "blobby." The targets are much weaker at the periphery than at the center, and preserving the boundaries requires methods with high dynamic range. **The faithful reconstruction of edges is also important for inferring velocities from time series.** Metrics for evaluating competing methods should be designed around this requirement.

The most widely-used imaging methods in upper-atmospheric research include adaptive beamforming methods like the one described by *Capon* [1969], iterative deconvolution methods like CLEAN [*Högbom*, 1974], and Bayesian optimization methods rooted in maximum entropy (e.g. [*Skilling and Bryan*, 1984]). In this paper, we consider methods arising from recent advances in the field of compressive sensing. Compressive sensing for one-dimensional imaging was evaluated recently by *Harding and Milla* [2013] who examined coherent backscatter from ionospheric plasma density irregularities at the magnetic equator. We pursue the same problem, this time considering imaging in two dimensions. The results should apply equally well to ionospheric scatter at low, middle, and high latitudes as well as to scatter from index-of-refraction variations in the neutral lower, middle, and upper atmosphere.

2. Compressive sensing and radar imaging

Compressive sensing has occupied a central role in image processing research for more than a decade and has significantly increased the acuity of any number of experimental modalities (see *Mackenzie* [2009] and references therein for review). The idea involves re-

producing undersampled signals accurately by exploiting inherent sparseness of the samples in an appropriately-chosen basis. Two-dimensional photographic images can often be represented very accurately by a small number of nonzero coefficients when expressed in a wavelet basis, for example. This implies not only the efficacy of image compression but also an opportunity for novel sampling strategies that incorporate sparseness into their design. The practical advantages of restoring a signal from sparse samples, as opposed to sampling the signal fully in the Nyquist sense, compressing the results, and then decompressing them later, are obvious. Moreover, the prospect of recovering complex signals from a small number of samples or from a small number of non-adaptive sensors has tremendous appeal in scientific realms where dense sensor arrays may be impractical to deploy to begin with. Aperture-synthesis radar imaging, which seeks the best estimate of the image brightness on the basis of a few measurements of the visibility, belongs in this realm.

The obvious objective function for evaluating sparseness is the $l_\circ$ pseudonorm, $||x||_\circ$, the number of nonzero entries in the vector x. If $x \in \mathbb{R}^m$ is a state vector which is s-sparse (has at most s nonzero entries) and is constrained by a data vector $y \in \mathbb{R}^n$ through an observation model $Ax = y$, where $A \in \mathbb{R}^{n \times m}$ is a non-invertible observation matrix, then a reasonable model for x could be **a vector with the smallest number of non-zero values which satisfies the observation model, viz.:**

$$x = \operatorname*{argmin}_{x} ||x||_\circ : Ax - y = 0 \tag{1}$$

which is a constrained optimization problem. This model yields not just accurate but exact recovery of undersampled state vectors under the condition that A is one-to-one in all

2s-sparse vectors. However, this is prohibitively complex (hard) optimization problem in computational enumeration.

In practice, the problem in (1) may be replaced by a closely related problem known as basis pursuit [*Donoho*, 2006]. Here, the $l_\circ$ pseudonorm is relaxed to the l_1 norm, i.e.

$$x = \underset{x}{\operatorname{argmin}} ||x||_1 : Ax - y = 0 \tag{2}$$

where the l_1 norm enforces sparsity in a manner similar to the $l_\circ$ norm. The simple modification transforms the task from an enumeration problem to a problem in convex optimization which can be solved practically using linear programming methods.

Numerous sufficient conditions for unique recovery of an s-sparse vector x from basis pursuit and related approaches are discussed in the literature including the restricted isometry property (RIP) [*Candés et al.*, 2006; *Candés and Tao*, 2006], the exact recovery condition (ERC) [*Tropp*, 2004, 2006], and the mutual incoherence condition (MIC) [*Donoho and Huo*, 2001; *Tropp*, 2006]. The last of these options is the **most intuitive and** most practical to evaluate. The mutual coherence for a matrix A with columns with unity l_2 norms is defined in terms of

$$\mu_A = \max_{i \neq j} |A_i^T A_j|, \tag{3}$$

i.e., the maximum pairwise column correlation. The sufficient condition for sparse signal recovery is that $\mu_A < (2s - 1)^{-1}$. This is also the worst-case necessary condition [*Tropp*, 2004].

3. Basis Pursuit Denoising (BPDN)

Signal sparsity is approximate rather than exact in practice, and the samples themselves are contaminated by noise. In this case, the basis pursuit problem is further relaxed to the nearby basis pursuit denoising (BPDN) problem which includes the provision for a finite residual which is weighted together with the sparsity penalty in the objective function:

$$x = \operatorname*{argmin}_{x} \left(||Ax - y||_2^2 + \lambda ||x||_1 \right) \tag{4}$$

Eq. 4 has the form of a regularized least-squares regression problem with λ being the regularization parameter. While eq. 4 is sometimes referred to as having Lagrangian form, λ is not an undetermined Lagrange multiplier here since the l_1 norm term does not function as a constraint. The λ parameter merely represents the tradeoff between the fidelity of the experiments and the noise sensitivity.

Note that two other problems are closely related to BPDN:

$$x = \operatorname*{argmin}_{x} ||x||_1 : ||Ax - y||_2^2 \leq \epsilon \tag{5}$$

$$x = \operatorname*{argmin}_{x} ||Ax - y||_2^2 : ||x||_1 \leq \epsilon \tag{6}$$

where ϵ is a tunable parameter like λ. All three problems have the same solutions for the appropriate choices of λ and ϵ which will depend on the data. The variant of the problem in 5 will be discussed further below. The variant in 6 incorporates the least absolute shrinkage and selection operator (Lasso) [*Tibshirani*, 1996]. An "elastic net" variant of the problem adds a regularization term based on the l_2-norm of x to the objective function.

96 While the observation matrix and the data in radar imaging experiments are complex,

97 we consider the real and imaginary parts of the observing model separately, making n

98 twice the number of interferometry baselines. This greatly simplifies the handling of the

99 l_1 norm but necessitates an explicit treatment of the real and imaginary parts of the data in

100 the error analysis [*Hysell and Chau*, 2006]. The positive definite symmetric inverse error

101 covariance matrix C_d^{-1} given there can be factored as $C_d^{-1} = C_d^{-T/2} C_d^{-1/2}$. In order to

102 incorporate error propagation, A and y are pre-whitened through scaling by the square root

103 information matrix $C_d^{-1/2}$, i.e. $(Ax - y) \to C_d^{-1/2}(Ax - y)$.

104 Likewise, a change of basis can be accommodated through the transformation $Ax \to$

105 $(AW^{-1})(Wx)$ where W is a linear transformation from state space to a basis where spar-

106 sity is optimized. **For example, as discussed below, it is commonplace in compressive**

107 **sensing to transform to a wavelet basis in which imagery can sometimes be repre-**

108 **sented very sparsely.** The MIC condition then applies to the new observation matrix

109 AW^{-1}. In this case, the last step of the algorithm is the transformation of the solution

110 vector x back to image space.

111 Note, finally, that the l_1 regularization problem can be cast in the form of a constrained

112 quadratic optimization problem by defining $x = u - v$, where u and v are the parts of

113 x that are positive and negative, respectively, and restricting the solution domain to the

114 region $u \geq 0, v \geq 0$ so that $||x||_1 \to u + v$. This reformulation invites the application of

115 any number of popular optimization methods rooted in linear programming. The penalty

116 is the doubling in length of the solution vector.

A range of approaches can be applied to BPDN and related problems (see *Yang et al.* [2013]; *Rani et al.* [2018] for reviews). Here, we pursue the fast iterative shrinkage thresholding algorithm (FISTA), a fast gradient algorithm which is simple to implement for aperture-synthesis imaging [*Beck and Teboulle*, 2009]. The algorithm depends on the concept of the proximal gradient which is summarized briefly below.

Gradient algorithms must cope with the fact that l_1 norm of a function is nondifferentiable. An option for convex nondifferentiable functions is to substitute the gradient operator with the subgradient operator which bounds the gradient. The subgradient of the l_1 norm of x is simply $\text{sign}(x)$. The subgradient method applied to BPDN can be solved iteratively using a conventional gradient descent method, but convergence will be slow.

In a proximal gradient method, which can be viewed as a generalized gradient descent method, the objective function is divided into two parts, i.e. $x = \text{argmin}_x \ (g(x) + h(x))$, where $g(x)$ is convex and differentiable and $h(x)$ is convex and possibly nondifferentiable. Eq. 4 is obviously in this form. The idea then is to iteratively minimize the sum of h and a quadratic local model of g. Toward this end, define the proximal operator of the function h about z as:

$$\text{prox}_{\alpha h}(z) \ = \ \underset{x}{\text{argmin}} \ \left(h(x) + \frac{1}{2\alpha}||x - z||_2^2 \right) \tag{7}$$

Each iteration proceeds from the results of the last, shifted opposite the direction of the local gradient of g, viz.

$$x^k \ = \ \text{prox}_{\alpha_k h}(x^{k-1} - \alpha_k \nabla g(x^{k-1})) \tag{8}$$

127 where the step size α_k can either be a fixed constant or determined by a line search. For

128 the l_1-regularized quadratic regression problem, stability is guaranteed by setting α_k to a

129 constant equal to the reciprocal of the smallest Lipschitz constant of ∇g which is twice the

130 largest eigenvalue of $A^T A$. This in turn can be calculated using the power method with

131 negligible computational burden.

In the case of the regularized least squares problem, $g(x)$ is quadratic, and its gradient has

an elementary linear form, i.e., $A^T(Ax-y)$. Furthermore, for $h(x) = \lambda||x||_1$, the proximal

operator or shrinkage function is given by the iterative soft thresholding function:

$$\text{prox}_{\lambda h}(y) = \begin{cases} y - \text{sign}(y)\lambda & |y| \geq \lambda \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

132 In the event that only positive values of x are admissible, as is the case for some formula-

133 tions of the radar-imaging problem, that condition can be readily incorporated in eq. 9 as

134 well.

The iterative shrinkage thresholding algorithm (ISTA) implied by eq. 8 can be shown

to converge at a rate that is linear in k like a standard gradient descent method under the

condition that ∇g is continuous in the L-Lipschitz sense, for both fixed and variable step

sizes. The FISTA algorithm, meanwhile, is able to achieve a quadratic rate of convergence

by incorporating Nesterov's acceleration method [*Nesterov*, 1983]. The modified algorithm

for FISTA, including acceleration, is [*Beck and Teboulle*, 2009]

$$z^k = \text{prox}_{\alpha_k \lambda h}(x^{k-1} - \alpha_k \nabla g(x^{k-1})) \tag{10}$$

$$t_k = \frac{1}{2}\left(1 + \sqrt{1 + 4t_{k-1}^2}\right) \tag{11}$$

$$x^k = z^k + \frac{t_{k-1} - 1}{t_k}(z_k - z_{k-1}) \tag{12}$$

135 with eq. 9 giving the proximal gradient. The algorithm is iterated to convergence with

136 convergence criteria typically being based on the norm of $x^k - z^k$.

137   Note that a parallelized version of FISTA (pFISTA) has been developed by *Peng et al.*

138 [3–6 Nov., 2013]. For the aperture-synthesis imaging application, parallelization can be

139 exploited without reformulating the basis algorithm simply by computing images for dif-

140 ferent Doppler spectral bins and range gates in separate processes.

4. Orthogonal Matching Pursuit (OMP)

141   A computationally expedient and highly intuitive alternative for sparse least-squares es-

142 timation is orthogonal matching pursuit (see e.g. *Tropp and Gilbert* [2007]; *Cai et al.*

143 [2010]; *Cai and Wang* [2011]). This is a greedy algorithm that formulates estimates of the

144 state vector by applying the least-squares pseudoinverse to a submatrix of the original mea-

145 surement matrix A. The submatrix starts from a null matrix and grows with each iteration

146 with the addition of a column selected and copied from A. The column selected is the one

147 with the highest correlation with the current residual. Iteration continues until a stopping

148 criterion is met.

149   **At this point, the state estimate is consistent with the data while being maximally**

150 **sparse. In terms of the discussion above, OMP solves the compressive sensing prob-**

151 **lem stated in 5 which, like BPDN, makes allowance for observation noise. OMP be-**

152 **gins from an essentially minimalist representation of the state vector and expands it in**

153 **the l_1-norm sense until the residual falls below a specified threshold. As with BPDN,**

154 **it is often expedient to transorm the problem using a basis in which the solution is**

155 **especially sparse.**

Define $A_k \in \mathbb{R}^{n \times k}$ to be a submatrix of A with its k columns having been copied from A. At each iteration, k is increased as the column of A most highly correlated with the current residual r is appended to A_k. The residual at the kth iteration is defined as:

$$r^k = y - A\hat{x}^k \tag{13}$$

$$= (I - A\tilde{A}_k)y \tag{14}$$

$$\tilde{A} \equiv (A^T A)^{-1} A^T \tag{15}$$

156 Here, $\tilde{A}$ is the standard least-squares pseudoinverse, and the estimator $\hat{x}^k = \tilde{A}_k y \in \mathbb{R}^m$ is

157 defined so as to have $k < m$ nonzero elements with indices that correspond to the indices

158 of the columns copied from A in the order in which they were copied. The index of the

159 selected column is the index of the term in $A^T r$ with the largest modulus.

160 Allowances for statistical errors can be introduced through incorporation of the square-

161 root data covariance matrix $C_d^{-1/2}$ described earlier. Note also that additional regularization

162 can be introduced here through the appropriate augmentation of the least-squares pseudoin-

163 verse. Stopping criteria may be based either on the size of the residual or the anticipated

164 sparsity of the solution. Absent the explicit introduction of regularization, OMP has no

165 tuning parameters comparable to the λ parameter in BPDN. The choice of stopping con-

166 ditions in the standard algorithm is subjective, however, particularly if neither the sparsity

167 nor the data error covariances are known a priori [*Kallummil and Kalyani*, 2017]. Note that

168 the number of iterations can be no larger than the number of data n. When the number of

169 basis functions is larger than the number of data, the problem becomes underdetermined,

170 and the least-squares pseudoinverse ceases to exist. After n iterations, the problem is even

171 determined, and the residual is identically zero.

5. Comparison using synthetic data

172 It is illustrative to benchmark the most common algorithms applied to the aperture syn-

173 thesis imaging problem. Among them is the adaptive beamforming method often attributed

174 to *Capon* [1969] and referred to as the linearly-constrained minimum variance (LCMV) or

175 the minimum variance distortionless response (MVDR) method **(or simply as "Capon's**

176 **method.")** We also consider the maximum entropy method (MaxENT) as formulated by

177 *Wilczek and Drapatz* [1985]. The relative performance of the two methods applied to the

178 aperture-synthesis imaging problem has been evaluated by *Yu et al.* [2000]. Another popu-

179 lar algorithm for aperture synthesis imaging is CLEAN [*Högbom*, 1974]. In hindsight, this

180 can be viewed as a special case of OMP and so will not be considered explicitly here.

181 We consider synthetic data comparable to what would be acquired by the Jicamarca

182 Radio Observatory measuring coherent scatter from field-aligned plasma density irregular-

183 ities in the equatorial ionosphere. The typical receiving antenna configuration is shown in

184 Fig. 1. Reception is performed using eight antenna sub-modules. Eight antennas imply 28

185 non-redundant interferometry baselines or 29 including the zero baseline. (The number of

186 data, n, is therefore 58). The radar operates at 50 MHz, and so the longest interferometry

187 baseline, 569 m, is approximately 95 wavelengths long.

Since the scattering irregularities are aligned with the geomagnetic field, the spatial correlation length of the backscatter in the north-south direction is very long and just barely measurable at Jicamarca [*Farley et al.*, 1981]. The most important information is to be found in the correlation length of the backscatter in the east-west direction, but there is utility in measuring the centroid of the backscatter in the north-south direction which varies somewhat with range.

We consider an imaging domain spanning ± 0.1 rad. in the plane of the magnetic equator and ± 0.02 rad. in the plane of the magnetic meridian. The former figure is roughly the effective field of view illuminated by the Jicamarca radar in imaging mode. The synthetic target is an elongated Gaussian ellipsoid with a half-width in the plane of the magnetic equator 10 times wider than in the plane of the magnetic meridian. The image will be constructed in a domain 128 pixels in the plane of the magnetic equator by 32 pixels in the plane of the magnetic meridian. The observing matrix for the problem is given by *Hysell and Chau* [2006].

Normally distributed independent noise is added to the synthetic visibility data for nonzero lags at the 2% level. This is a simplified treatment for observation noise and represents an upper bound for the experimental uncertainty associated with a signal-to-noise ratio larger than unity and averages of 2500 statistically-independent samples. In actual experiments, noise bias in the zero-baseline data is estimated and removed, and so no noise bias will be added here. For an exhaustive treatment of error analysis and propagation in aperture-synthesis imaging, see *Hysell and Chau* [2006].

209   Fig. 2 shows the results of image recovery using the LCMV method which serves here as

210   a baseline. Grayscales indicate relative backscatter intensity in dB relative to the intensity

211   at the center of the target. We plot 30 dB of dynamic range in the figure. In practical radar

212   experiments, radar clutter from pulse coding is generally present at the level of about -22

213   to -25 dB. We desire about 30 dB of usable dynamic range from imaging methods so that

214   pulse coding rather than radar imaging will be the limiting factor for radar clutter. The

215   elliptical contours represent the truth model and indicate backscatter at the -10 dB, -20 dB,

216   and -30 dB levels, respectively.

217   LCMV has no tuning parameters and also no means of incorporating estimates of mea-

218   surement confidence levels in the analysis. The method has accurately recovered the

219   strongest intensity region within the 10-dB contour. However, the method has done a

220   poor job of rejecting clutter outside the 30-dB contour. (We define clutter as extraneous in-

221   tensity outside the 30-dB boundary of the truth model.) Strong artifacts are present across

222   the image. The clutter level rises sharply as the level of random fluctuations added to the

223   synthetic data is increased.

224   Fig. 3 shows the results of image recovery using the FISTA algorithm. For this test,

225   we do not incorporate wavelet transforms and simply minimize the l_1 norm of the im-

226   age brightness itself. Since the intensity is non-negative, we can and do incorporate that

227   information in the computation of the proximal gradient to speed computation.

228   There is one tunable parameter in FISTA – the regularization parameter λ. The larger

229   the value of λ, the greater the tendency toward sparse solutions. Here and in the methods

230   to follow, λ has been adjusted so as to make the chi-squared parameter equal to the number

231 of data n. This has the effect of limiting clutter while adequately filling the truth-model

232 contours.

233 FISTA recovers an elliptical target with approximately the same width in the zonal di-

234 rection as the truth model but with more than twice the width in the meridional direction.

235 This is a consequence of the fact that the interferometry baselines in the meridional di-

236 rection are relatively short while the target itself is relatively narrow. **(Note that highly**

237 **field-aligned targets have long spatial correlation lengths and elongated visibilities in**

238 **the direction of the magnetic meridian and, consequently, narrow angular meridional**

239 **widths in brightness imagery.)** This is not an important shortcoming in practice where

240 1D images are normally extracted from the 2D images either by taking a horizontal cut or

241 by averaging over a narrow range of meridional coordinates. The imaging resolution in

242 the meridional direction need only be sufficient to allow us to estimate the centroid of the

243 backscatter.

244 Clutter is mainly limited to bleedthrough in the direction of the magnetic meridian. Clut-

245 ter is largely excluded from the horizontal bisector of the image corresponding to the mag-

246 netic equator. Clutter associated with interferometry sidelobes is completely suppressed.

247 However, the dynamic range of the FISTA image is also limited; pixels appear to be either

248 fully set or fully unset throughout most of the image. Gradation in image intensity appears

249 to be largely lost using this **and any method rooted in compressive sensing**.

250 Fig. 4 shows the results of image recovery using the OMP algorithm. For this test,

251 we have used a transformation W consistent with a two-dimensional wavelet transform.

252 **Wavelet transforms are widely used for image analysis generally and for analysis of**

astronomical data in particular (e.g. *Starck and Bobin* **[2009] and references therein). The reason is that wavelets are efficient in reproducing the kind of hierarchical, fractal-like features that typify astronomical datasets. The experience base in iono-spheric imaging is much narrower, but the similarity with astronomy in terms of the targets at issue and the methods in use prompts trials with wavelets for the same reason.**

Here, we implement the pyramidal scheme of *Press et al.* [1988]. Daubechies d20 wavelets were used, but qualitatively similar results were obtained using d12 and d4 wavelets. This can be attributed to the fact that most of the information in the synthetic image is contained in the low-frequency components and so using larger filters makes little difference.

The tunable parameter for OMP is the number of iterations. The example shown here used 13 iterations, yielding a chi-squared value close to the number of data n. Using fewer iterations leads to images composed of simple two-dimensional shapes with a blocky appearance. More iterations leads to a somewhat more elliptically-shaped center image and to clutter forms with different and more complicated shapes. In the noiseless case, the additional iterations beyond the expected number of nonzero values increases the probability of signal recovery [*Sahoo and Makur*, 2015]. However, in cases with noise, additional stopping criteria are required to avoid the selection of the zero components. *Cai and Wang* [2011] (eq. 5) propose a stopping condition to ensure signal recovery in the case of Gaussian noise. In this work, a chi-squared value equal to the number of data is used as the

274 stopping condition. **This is readily enforced for all the methods considered here and is**

275 **consistent with Morozov's discrepancy principle for regularization [*Morozov*, 1966].**

276 The algorithm has done a reasonable job recovering the elliptical target although the

277 stretching in the meridional direction is more severe than with the other methods. How-

278 ever, considerable clutter is evident across the image. Like LCMV and unlike FISTA

279 (and MaxENT as we will see), the results of OMP **with Daubechies wavelets** are not im-

280 proved significantly by taking 1D cuts or averages since the clutter is widespread. Unlike

281 all the other methods, OMP with Daubechies wavelets produces clutter which is not odd

282 symmetric. When the number of iterations is increased, the clutter becomes completely

283 asymmetric. This is particularly problematic since clutter lacking odd symmetry would be

284 more likely to be mistaken for an interesting feature in actual imagery.

285 Note that qualitatively similar results were obtained using Daubechies wavelets with the

286 FISTA algorithm albeit with much greater computational cost. Using the OMP algorithm

287 without wavelet transforms, meanwhile, produces very unsatisfactory results. The algo-

288 rithm simply selects and activated a number of pixels equal to the number of iterations in

289 that case. Most but not all of the pixels fall within the 30-dB contours in the figure. Given

290 a number of iterations equal to the number of data, the residual is identically zero, but the

291 recovery of the truth image is poor. OMP **with the Dirac (or identity) basis** evidently

292 does not perform well with distributed targets. It is known that BPDN requires fewer mea-

293 surements than OMP to recover a signal with the same probability. Whereas OMP needs

294 of the order of $k\ln(m)$ measurements, BPDN needs $k\ln(m/k)$ measurements [*Tropp and*

295 *Gilbert*, 2007]. This can explain why FISTA works without transforming the domain while

OMP does not. The synthetic image is not sparse enough for the original basis to apply to OMP, **i.e., the MIC is not met**.

It is noteworthy also that the non-negativity of the image solution is not enforced when wavelet bases are incorporated in the manner described here. For references on preserving the non-negativity condition, see *Yaghoobi et al.* [2015] and *Nguyen et al.* [Sep., 2017].

Finally, Fig. 5 shows the results of image reconstruction using the MaxENT algorithm, following the prescription of *Hysell and Chau* [2006]. The MaxENT algorithm has a tuning parameter which is the design value for the chi-squared parameter. Here and in actual imaging experiments, we again force chi-squared to equal the number of data n.

Like FISTA, the MaxENT algorithm recovers an elliptical target with approximately the same width in the zonal direction as the truth model but with more than twice the width in the meridional direction. Overall, the algorithm recovers the dynamic range of the truth model with a minimum of clutter. A small amount of clutter associated with the interferometry sidelobes is visible in the upper-left and lower-right corners of the image.

A quantitative comparison of the four imaging methods is made in Tab. 1. For each of the methods, the relative computational cost is shown. This number is the computation time in ms on a single i7 960 CPU core with a 1733 MHz CPU speed. The four methods were implemented in C and compiled with gcc with full optimization. The implementations make use of LAPACK for linear algebra operations. Computation times vary with different algorithm parameters and truth models, and the figures shown are meant to be only representative.

317 Tab. 1 also shows the RMS discrepancy between the recovered image and the truth

318 model in dB. Assessing penalties in terms of dB quantities prevents the overemphasis of

319 just the strongest portions of the targets, a common practice that undervalues the impor-

320 tance of dynamic range in image construction.

321 The RMS figure of merit was computed by first thresholding the recovered image and

322 the truth model and then summing the squares of the differences between the model and

323 the image in dB across pixels. Thresholding means setting a -30 dB floor for both the

324 model and the image prior to computing the metric. The rationale for thresholding is to

325 de-emphasize discrepancies in regions of the image that are too weak to be meaningful.

326 The two RMS error terms for each method refer to the result for the entire image and to

327 a 1D cut through the horizontal bisector through the image, respectively. The latter is the

328 better metric for applications involving radar backscatter from field-aligned irregularities.

329 Most of the RMS error in the tests is due to spurious clutter, although FISTA and OMP

330 also suffer from underpredicting brightness at the periphery of the radar target. Tab. 1

331 suggests that the imaging methods fall into two categories: slow methods that resist clutter

332 and fast methods that are prone to clutter. In the case of the slow methods, most of the

333 clutter in the tests is removed by considering horizontal cuts or averages through the 2D

334 images. This is not true for the fast methods where the clutter is more widespread in the 2D

335 images. Overall, neither FISTA nor OMP appear to offer improved performance over the

336 more conventional methods, LCMV and MaxENT, at least as they have been formulated

337 here.

We point out, however, that our formulation of MaxENT is actually similar to OMP in some respects and so may also be considered a form of compressive sensing in that regard. OMP attempts to form an image from a superposition of basis vectors (atoms) drawn from the observation matrix such as to restrain the norm of the residual in a least-squares sense. No more than n vectors may be involved. MaxENT, meanwhile, attempts to reconstructs the logarithm of the image from exactly n basis vectors also drawn from the observation matrix. The logarithm relationship derives from the form of Shannon's entropy. To reproduce a Gaussian ellipsoid, MaxENT need therefore only construct a paraboloid from the available basis vectors. This helps to explain its comparative success in the tests posed in this study.

We conclude this section with an analysis of the performance of the various radar imaging methods under different levels of statistical uncertainty. Radar backscatter from soft targets is stochastic, and uncertainty is associated with visibility estimates based on finite time averages. The uncertainty is a function of the number of statistically independent samples, the signal-to-noise ratio, and the data themselves [*Hysell and Chau*, 2006]. The associated fluctuations in the experimental visibility estimates contribute to imaging errors.

Regularization in imaging methods is used to limit the growth of fluctuations as they propagate from visibility estimates to brightness estimates. The penalty for regularization is bias, and the balance between fluctuations and bias is exemplified in Fig. 6 which shows the 1D RMS error parameter as a function of the relative standard deviation of the visibility estimates (standard deviation relative to mean value).

360 **The results for the four imaging methods considered in this paper are shown. In each**

361 **case, bias dominates fluctuations at small values of σ/μ where the error parameter**

362 **tends to reach a floor asymptotically. The situation reverses at large values where the**

363 **error parameter is an increasing function of σ/μ.**

364 **The transition from bias to fluctuation is different for the four methods considered.**

365 **The LCMV method shows performance comparable to FISTA and MaxENT for small**

366 **values of σ/μ, but performance degrades immediately as σ/μ increases. LCMV in-**

367 **cludes no explicit regularization and so is nowhere bias-limited. It functions best**

368 **when fluctuations in the data are small, in the large signal-to-noise ratio limit for ex-**

369 **ample, but is prone to instability when fluctuations are large and the data covariance**

370 **matrix becomes poorly conditioned. The results become erratic for large σ/μ.**

371 **OMP with a wavelet basis, meanwhile, appears to be bias-limited up through large**

372 **values of σ/μ. This prevents the method from capitalizing on high-quality data that**

373 **might be available, for example, in the high signal-to-noise ratio limit. The preference**

374 **for sparse solutions appears to represent a strong bias that conceals image features**

375 **even when they have support in the data. MaxENT and FISTA are superior in this**

376 **regard in that their performance benefits from values of σ/μ as small as about 1%.**

377 **Even smaller values would be difficult to achieve in practical experimental applica-**

378 **tions.**

Summary

This paper set out to compare and contrast four methods for inverting spaced-receiver radar observations of field-aligned plasma density irregularities in the ionosphere. **Measured visibility data are related to desired radar images through a linear transformation similar to a Fourier transform, but the sparse and incomplete sampling of the former necessitates the use of inverse methods in forming the latter.** Two new methods, basis pursuit denoising (using FISTA) and orthogonal matching pursuit, were considered along with two conventional methods based in entropy maximization (MaxENT) and adaptive beamforming (LCMV). Realistic synthetic data as might be acquired by the imaging array used at the Jicamarca Radio Observatory were used for the study.

Choosing the right metric evaluation is important. The goal of aperture-synthesis imaging, an underdetermined problem, is not the minimization of the norm of the residual. For the MaxENT, FISTA, and OMP runs conducted here, the chi-squared parameter was made to be equal to the number of data elements n in each case. The goal instead is the accurate recovery of the $m \gg n$ elements of the truth model. However, the RMS discrepancy between the truth model and the recovered image would not be a very appropriate metric given the importance of dynamic range in radar imaging. All of the methods considered here would have scored well by this metric since they were all able to reproduce the strongest parts of the truth model. We instead considered the RMS discrepancy in the logarithms of the truth model and the recovered image. A threshold was set at -30 dB in the calculation to avoid overemphasizing discrepancies in parts of the images that were negligibly small, however. By this metric, MaxENT outperformed the other methods. Basis

pursuit denoising performed adequately but did not reproduce the desired gradation at the boundary of the truth-model target. Neither LCMV nor OMP provided adequate clutter suppression outside the boundary of the target.

We note again that none of the methods tested here were able to resolve the synthetic radar target in the direction of the magnetic meridian. Radar backscatter from field-aligned irregularities occurs within a very narrow range of meridional angles, and we do not expect or even attempt to measure magnetic aspect width using imaging in practice. The important metric is the ability to recover the shape of the backscatter in the appropriate 1D cut through the 2D imagery. There are advantages in formulating the imaging problem in 2D, however, including the ability to measure the direction of the centroid of the backscatter.

Moreover, overestimating the width of targets in the meridional direction could pose a problem to the extent that it makes the method prone to underestimating the width in the zonal direction, a feature exhibited to some extent by all of the test cases considered here. **Target widths in the two directions are coupled since the moments are approximately conserved.** This observation may warrant a redesign of the imaging array used routinely at Jicamarca which presently does not include long meridional baselines.

Methods rooted in BPDN and OMP may yet be able to perform well in the aperture-synthesis imaging problem. The impressive speed of OMP warrants an examination of different basis functions which may be more suitable for extended 2D radar targets. One intriguing possibility is the use of curvelet transforms which have proven expedient in applications similar to this one [*Ma and Plonka*, 2010].

References

Beck, A., and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imag. Sci.*, *2*, 183–202, 2009.

Cai, T., and L. Wang, Orthogonal matching pursuit for sparse signal recovery with noise, *IEEE Trans. Inf. Theory*, *57*, 4680–4688, 2011.

Cai, T. T., L. Wang, and G. Xu, Stable recovery of sparse signals and an oracle inequality, *IEEE Trans. Inf. Theory*, *56*, 3516–3522, 10.1109/TIT.2010.2048,506, 2010.

Candés, E. J., and T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies?, *IEEE Trans. Inf. Theory*, *52*, 5406–5425, 2006.

Candés, E. J., K. Romberg, and T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory*, *52*, 489–509, 2006.

Capon, J., High-resolution frequency-wavenumber spectrum analysis, *Proc. IEEE*, *57*, 1408, 1969.

Donoho, D., Compressed sensing, *IEEE Trans. Inform. Theory*, *52*, 1289–1306, doi:10.1109/TIT.2006.871,582, 2006.

Donoho, D. L., and X. Huo, Uncertainty principles and ideal atmnoic decomposition, *IEEE Trans. Inf. Theory*, *47*, 2845–2862, 2001.

Farley, D. T., H. M. Ierkic, and B. G. Fejer, Radar interferometry: A new technique for studying plasma turbulence in the ionosphere, *J. Geophys. Res.*, *86*, 1467–1472, 1981.

Harding, B. J., and M. A. Milla, Radar imaging with compressed sensing, *Radio Sci.*, pp. 582–588, DOI: 10.1002/rds.20,063, 2013.

Högbom, J. A., Aperture synthesis with a non-regular distribution of interferometer baselines, *Astron. Astrophys. Supp.*, *15*, 417–426, 1974.

Hysell, D. L., Radar imaging of equatorial *F* region irregularities with maximum entropy interferometry, *Radio Sci.*, *31*, 1567, 1996.

Hysell, D. L., and J. L. Chau, Optimal aperture synthesis radar imaging, *Radio Sci.*, *41*, 10.1029/2005RS003,383, RS2003, 2006.

Hysell, D. L., M. Yamamoto, and S. Fukao, Imaging radar observations and theory of type I and type II quasi-periodic echoes, *J. Geophys. Res.*, *107 (A11)*, 1360, 2002.

Hysell, D. L., M. F. Larsen, and Q. H. Zhou, Common volume coherent and incoherent scatter radar observations of mid-latitude sporadic *E*-layers and QP echoes, *Ann. Geophys.*, *22*, 3277–3290, 2004.

Hysell, D. L., G. Michhue, M. F. Larsen, R. Pfaff, M. Nicolls, C. Heinselman, and H. Bahcivan, Imaging radar observations of Farley Buneman waves during the JOULE II experiment, *Ann. Geophys.*, *26*, 1837–1850, 2008.

Kallummil, S., and S. Kalyani, Tuning free orthogonal matching pursuit, *arXiv.org*, p. arXiv:1703.05080v1, 2017.

Kudeki, E., and F. Sürücü, Radar interferometric imaging of field-aligned plasma irregularities in the equatorial electrojet, *Geophys. Res. Lett.*, *18*, 41, 1991.

Ma, J., and G. Plonka, The curvelet transform, *IEEE Sig. Proc. Mag.*, *2*, 118–133, 2010.

Mackenzie, D., *What's Happening in the Mathematical Sciences*, vol. 7, American Mathematical Society, Providence, RI, 2009.

Morozov, V. A., On the solution of functional equations by the method of regularization, *Soviet Math. Dokl.*, *7*, 414–417, 1966.

Nesterov, Y. E., A method for solving a convex programming problem with convergence rate $O(1/k^2)$, *Soviet Math. Doklady*, *27*, 372–376, 1983.

Nguyen, T. T., C. Soussen, J. Idier, and E.-H. Djermoune, An optimized version of nonnegative OMP, XXVIe Colloque GRETSI Traitement du Signal & des Images, GRETSI 2017, Juan-les-Pins, France, Sep., 2017.

Peng, Z., M. Yan, and W. Yin, Parallel and distributed sparse optimization, Paper presented at the 2013 Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 3–6 Nov., 2013.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge Univ. Press, New York, 1988.

Rani, M., S. B. Dhok, and R. B. Deshmukh, A systematic review of compressive sensing: Concepts, implications, and applications, *IEEE Access*, *6*, 4875–4894, DOI:10.1109/ACCESS.2018/2793,851, 2018.

Sahoo, S. K., and A. Makur, Signal recovery from random measurements via extended orthogonal matching pursuit, *IEE Trans. Sig. Proc.*, p. doi: 10.1109/TSP.2015.2413384, 2015.

Saito, S., M. Yamamoto, H. Hashiguchi, and A. Maegawa, Observation of three-dimensional signatures of quasi-periodic echoes associated with mid-latitude sporadic-E layers by MU radar ultra-multi-channel system, *Geophys. Res. Lett.*, *33, L14109*, doi:10.1029/2005GL025,526, 2006.

Saito, S., M. Yamamoto, and H. Hashiguchi, Imaging observations of nighttime mid-latitude F-region field-aligned irregularities by an MU radar ultra-multi-channel system, *Ann. Geophys.*, *26*, 2345–2352, 2008.

Skilling, J., and R. K. Bryan, Maximum entropy image reconstruction: General algorithm, *Mon. Not. R. Astron. Soc.*, *211*, 111, 1984.

Sommer, S., and J. L. Chau, Patches of polar mesospheric summer echoes characterized from radar imaging observations with MAARSY, *Ann. Geophys.*, *34*, 1231–1241, 2016.

Starck, J.-L., and J. Bobin, Astronomical data anaysis and sparsity: from wavelets to compressed sensing, *Astronomical Data Analysis*, p. https://arxiv.org/pdf/0903.3383.pdf, 2009.

Thompson, A. R., *Interferometry and Synthesis in Radio Astronomy*, John Wiley, New York, 1986.

Tibshirani, R., Regression shrinkage and selection via the Lasso, *J. Roy. Stat. Soc.*, *58*, 267–288, 1996.

Tropp, J., Greed is good: Algorithmic results for sparse approximation, *IEEE Trans. Inf. Theory.*, *50*, 2231–2242, 2004.

Tropp, J., Just relax: Convex programming methods for identifying sparse signals in noise, *IEEE Trans. Inf. Theory.*, *52*, 1030–1051, 2006.

Tropp, J., and C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inf. Theory.*, *53*, 4655–4666, 2007.

Urco, J. M., J. L. Chau, M. A. Milla, J. P. Vierinen, and T. Weber, Coherent MIMO to improve aperture synthesis radar imaging of field-aligned irregularities: First results at Jicamarca, *IEEE Trans. Geosci. and Remote Sens.*, *56*, 2980–2990, doi:10.1109/TGRS.2017.2788,425, 2018a.

Urco, J. M., J. L. Chau, T. Weber, and R. Lateck, Enhancing the spatio-temporal features of polar mesosphere summer echoes using coherent MIMO and radar imaging at MAARSY, *Atmos. Meas. Tech.*, pp. https://doi.org/10.5194/amt–2018–258, 2018b.

Wilczek, R., and S. Drapatz, A high accuracy algorithm for maximum entropy image restoration in the case of small data sets, *Astron. Astrophys.*, *142*, 9, 1985.

Yaghoobi, M., D. Wu, and M. E. Davies, Fast non-negative orthogonal matching pursuit, *IEEE Signal Proc. Lett.*, *22*, 1229–1233, 10.1109/LSP.2015.2393,637, 2015.

Yang, A. Y., Z. Zhou, A. G. Balasubramanian, S. S. Sastry, and Y. Ma, Fast l_1-minimization algorithms for robust face recognition, *IEE Trans. Img. Proc.*, *22*, 3234–3246, doi:10.1109/TIP.2013.2262,292, 2013.

Yu, T. Y., R. D. Palmer, and D. L. Hysell, A simulation study of coherent radar imaging, *Radio Sci.*, *35*, 1129, 2000.
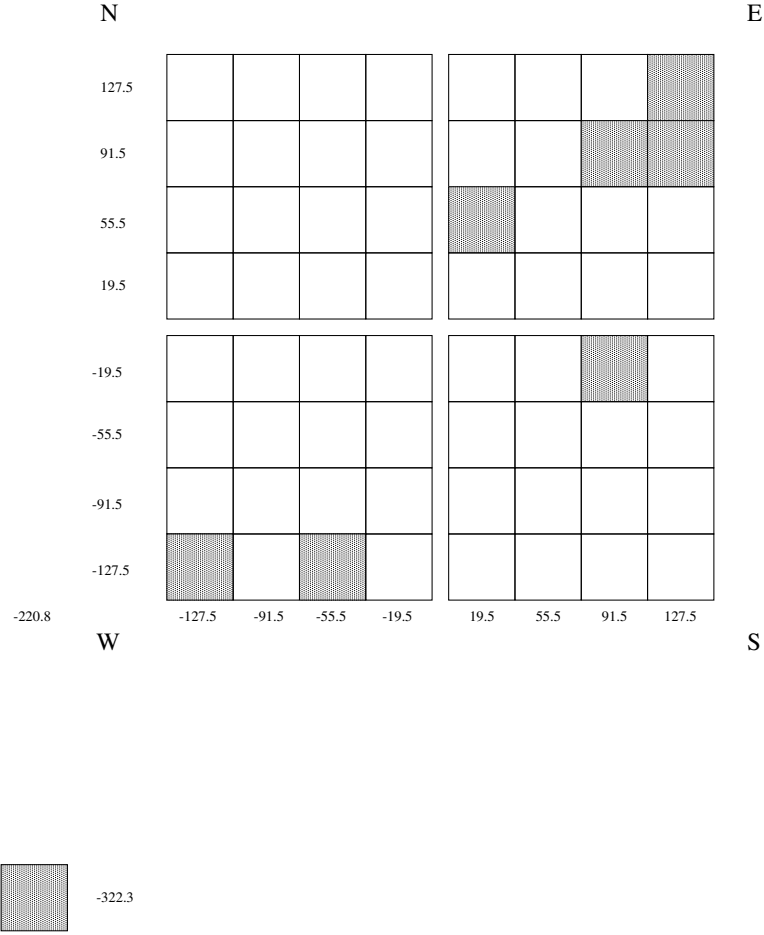
Figure 1. Configuration of the antenna array of the Jicamarca Radio Observatory in imaging mode. The corners of the main array are nearly aligned with the cardinal directions, and the north-south line is closely aligned with the direction of the geomagnetic field at present. Module coordinates with respect to the array center are indicated. The distance between adjoining modules is 36 m, and the longest baseline is 569 m. Seven modules of the main array plus an eighth outrigger array are used for imaging.
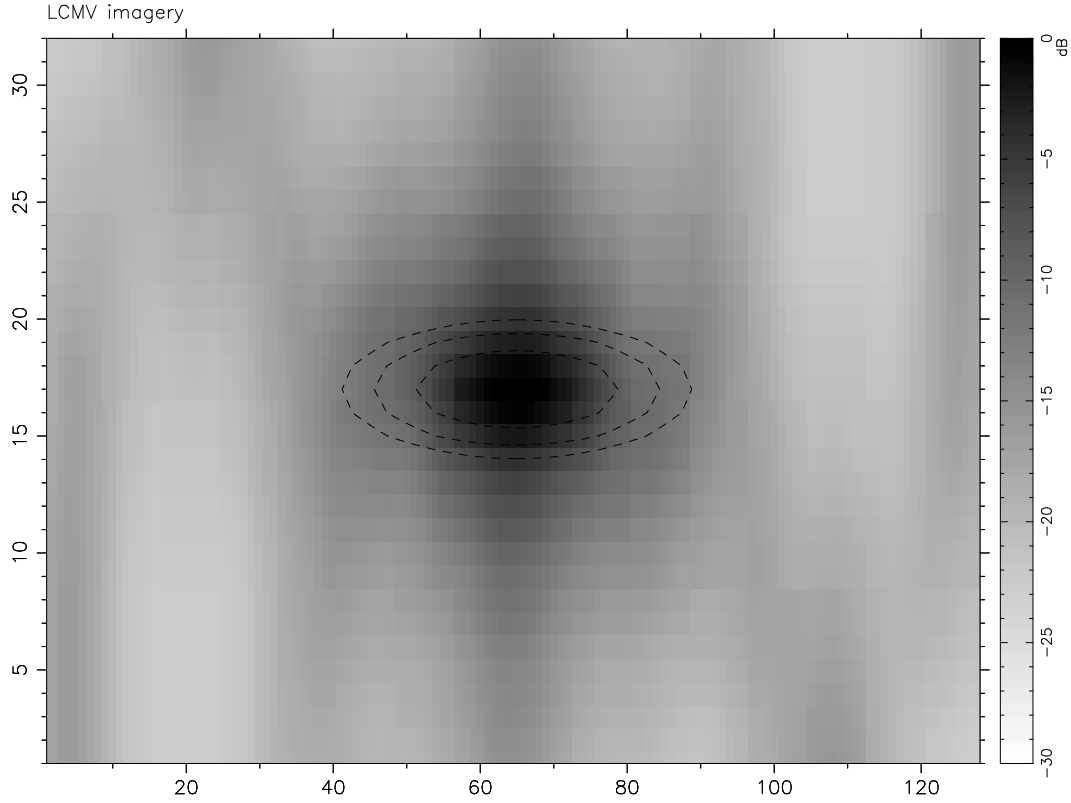
Figure 2. Image recovered from synthetic data using the LCMV algorithm. The imaging domain spans ±0.1 rad in the plane of the magnetic equator (horizontal direction) and ±0.02 rad in the plane of them magnetic meridian (vertical). The resolution of the image is 128×32. Grayscales denote brightness relative to the maximum in dB. Red dashed contours represent the -10 dB, -20 dB, and -30 dB contours of the truth model, a Gaussian ellipsoid.
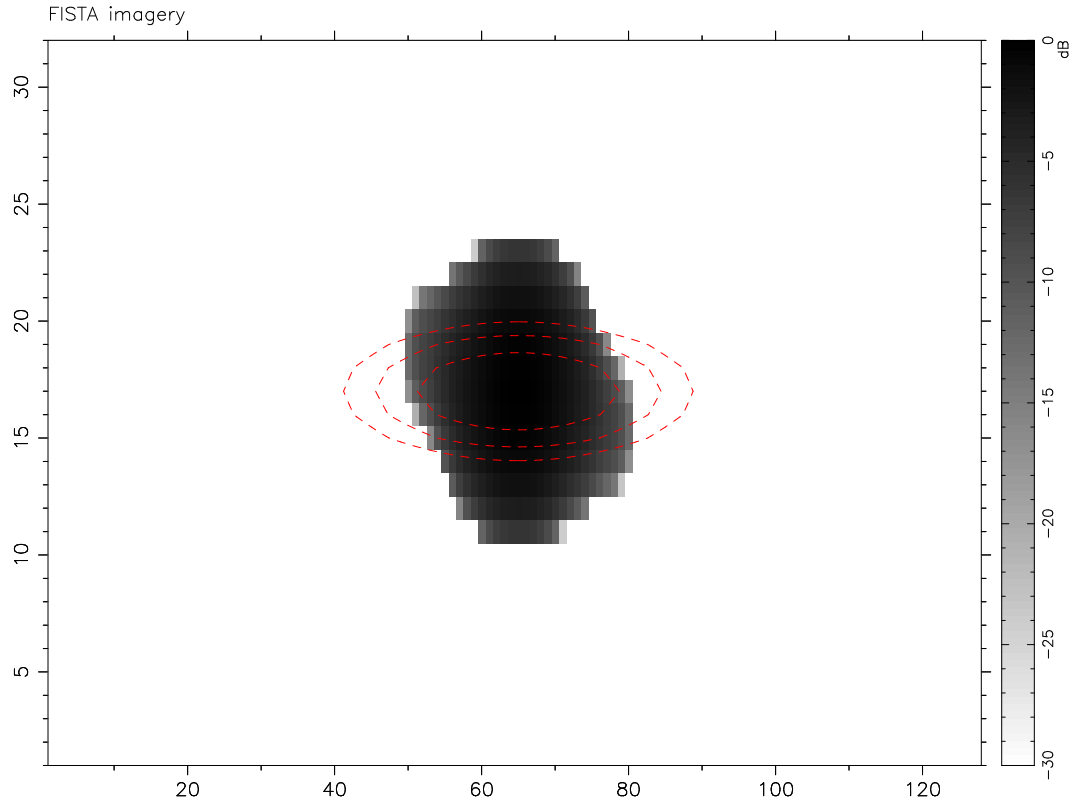
Figure 3. Image recovered from synthetic data using the FISTA algorithm.
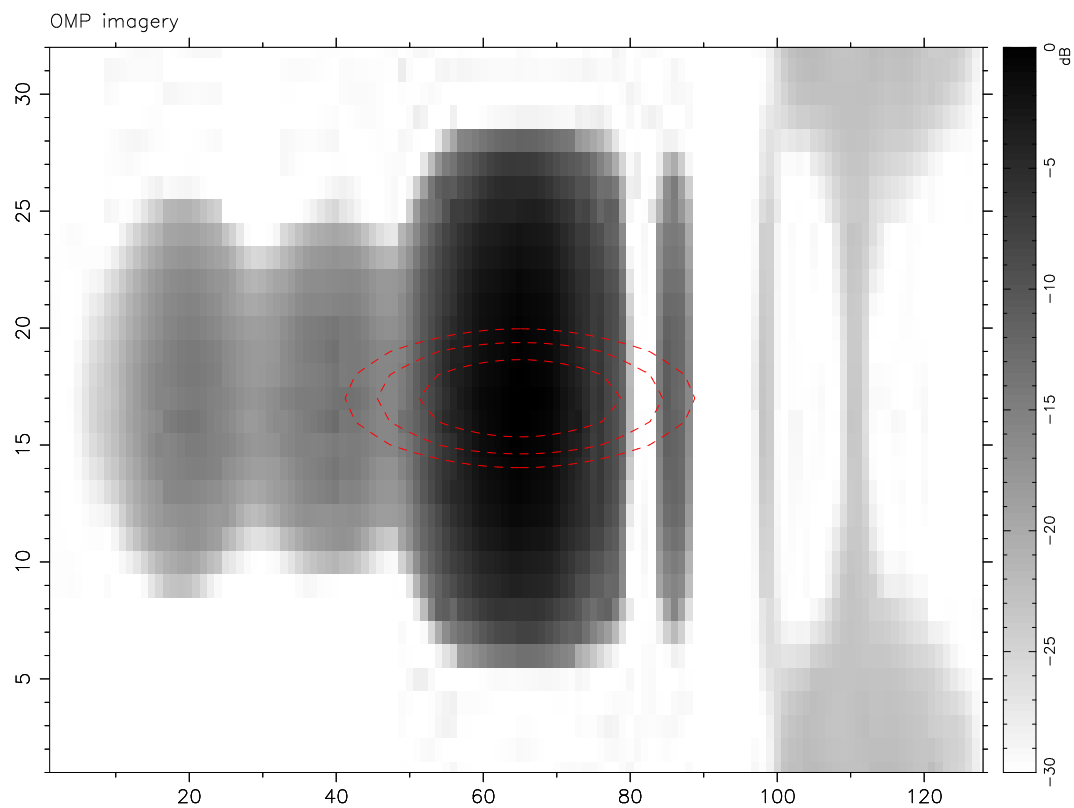
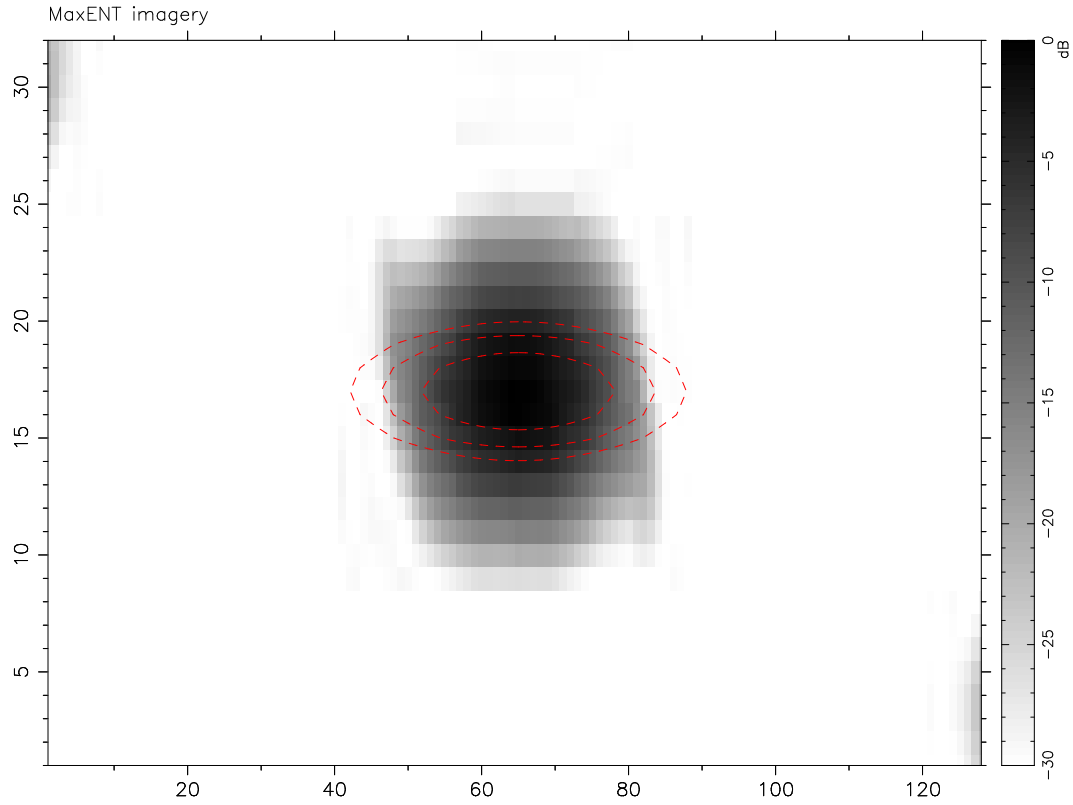**Figure 4.** Image recovered from synthetic data using the OMP model.

Figure 5. Image recovered from synthetic data using the MaxENT algorithm.

Method	Time (ms)	RMS error (dB)	1D cut (dB)
LCMV	20	12.73	9.99
FISTA	1920	5.16	3.82
OMP	10	9.83	8.11
MaxENT	1090	4.66	1.86

Table 1.   Comparison of aperture synthesis imaging methods. The execution time is measured in ms, and the RMS error is measured in dB. The two RMS error figures refer to the entire 2D image and to a 1D cut through the horizontal bisector through the image, respectively (see text).

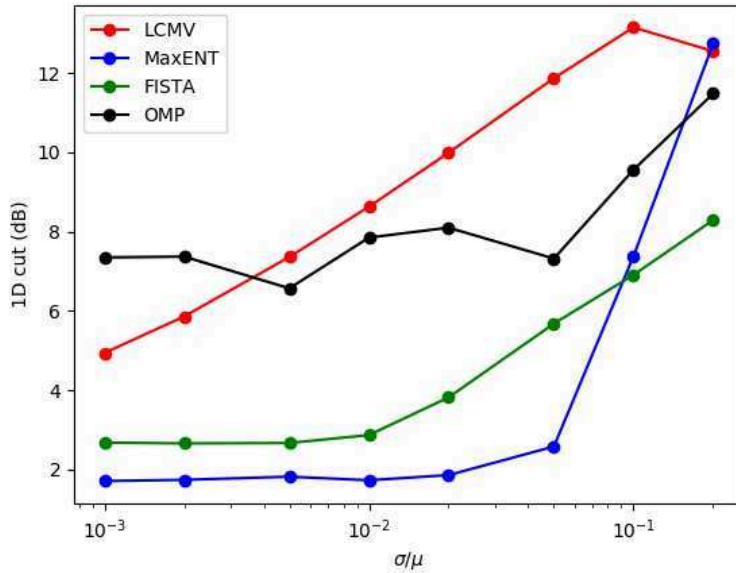**Figure 6.**   RMS errors for 1D cuts versus the relative standard deviation of visibility measurements for four aperture synthesis imaging methods.