Astron. Astrophys. 324, 735–742 (1997)

ASTRONOMY AND ASTROPHYSICS

Fast maximum entropy restoration of low-noise solar images

K. Arzner and A. Magun

Institute of Applied Physics, University of Berne Sidlerstrasse 5, CH-3012 Bern, Switzerland

Received 31 July 1996 / Accepted 17 January 1997

Abstract. Maximum Entropy (ME) deconvolution has become a widely used approach for image restoration. Although it has many attractive features, its computational burden is considerably larger than for competing linear inversion techniques. The present work accelerates ME restoration by using linear techniques within the ME framework and proposes a simple and fast algorithm which is applicable to low-noise ($< 3\%$) images. The algorithm is tested in numerical simulations, compared to well-known algorithms and applied to solar radio maps obtained by a scanning antenna.

Key words: techniques: image processing – methods: numerical – Sun: general

1. Introduction

Many observational situations in optical or radioastronomical imaging can be described in the following way: the measurement **m** is the convolution of the true brightness distribution **t** with some point spread function (PSF) plus additive noise **n**:

$$\mathbf{m} = P\,\mathbf{t} + \mathbf{n}\,. \tag{1}$$

Here, all boldface quantities are arrays (1D distributions or 2D-images) and P denotes a linear operator convolving the true image with the PSF. For simplicity, we assume that all arrays have the same number of pixels N, although this demand is not essential and could be relaxed.

The PSF describes the instrumental response and its shape may vary from a simple smoothing kernel to a comb-like form in interferometric applications. In this paper, we intend to deconvolve solar maps obtained by a single scanning antenna. Its PSF consists of an extended main lobe with moderate sidelobes. The scan convolves the true image with the PSF and filters out high spatial frequencies. This may be illustrated by a $sinc^2$-shaped PSF with a triangular filter passband. As the power at frequencies above the triangle's cutoff is zeroed, P becomes singular. These frequencies build P's nullspace (i.e. those contributions which are mapped by P onto zero). A second example is a PSF with a gaussian filter shape. The power at high frequencies becomes not exactly zero but is so strongly supressed that it is dominated by the noise **n** or by numerical errors. The operator P is then called ill-conditioned.

A direct inversion fails for a singular as well as for an ill-conditioned P because the measurement allows no conclusion on the high frequency part of the true brightness distribution **t**. Therefore other techniques are needed to restore the true image from the measurement – i.e. to select the best representative out of all true images allowed by the measurement.

In the second part of our century, various selection criteria have been developed such as smoothness of restoration (curvature, correlations) or best agreement with the true distribution in a least-square sense (Wiener filter). For a short overview, see Numerical Recipes (Press et al. 1988) and for more detailed information and applications Tikhonov & Arsenin (1977), Menke (1984), Molina (1992), Van Trees (1986), Helstrom (1966).

All these methods have an important advantage. They are fast because the restoration is obtained from the measurement by a linear, inverse operation. But for imaging applications, they have also a severe disadvantage as they cannot preserve positivity of the restored brightness in a natural way. An approach which intrinsically overcomes this problem identifies the brightness distributions with probability distributions and applies the methods of information theory to them. In Maximum Entropy (ME) methods (Ables 1974, Narayan 1986) which we will use here, the restoration is selected by the demand of minimum – possibly wrong – information content, measured by the entropy. There have been extended discussions on the correct choice of the entropy function (Narayan, Kikuchi 1977); we prefer for axiomatic demands (Tikochinski et al. 1984; Eriksson 1987), the following, classical expression for the entropy S:

$$S = \sum_{i=1}^{N} r_i \, ln \, \frac{r_i}{a_i}\,, \tag{2}$$

where $\mathbf{r} = (r_1, ..., r_N)$ represents the restoration and the a priori distribution **a** has been introduced to account our prior knowledge on the true image. For many applications, for which we don't have any prior information on the measured object, **a** becomes flat. For the deconvolution of solar images however, **a** can be used to bring in the knowledge that interesting finestructures are superimposed on a quiet and constant solar disk. Note, that due to the probabilistic interpretation, all distribuitions and P

Send offprint requests to: K. Arzner

must be normalized ($\sum t_i = \sum r_i = \sum a_i = 1$ and $\sum_i P_{ij} = 1$; assuming $\langle \mathbf{n} \rangle = 0$ and hence $\sum n_i = 0$).

According to the ME principle we must find the restoration $\mathbf{r}$ which has the highest entropy and which is compatible with the measurements. If the image is noise-free, each pixel of $\mathbf{m}$ must agree with $P\mathbf{r}$, leading to N constraints to be imposed on $\mathbf{r}$. On the other hand, if there is some noise in the measured image, we cannot demand a strict agreement but require that the residuals $\mathbf{m} - P\mathbf{r}$ follow error statistics of the measurement. Several criteria for this have been proposed. Bryan & Skilling (1980) enforce the residual distribution to follow the expected error distribution; but often the residuals are more simply assumed to be gaussian (Ables 1974; Gull & Daniell 1978; Cornwell & Evans 1985), and the compatibility with the measurements is imposed by a single χ^2-constraint.

It is intuitively clear that a single χ^2-constraint becomes somewhat inefficient if the noise level tends to zero and we are very close to a noise-free measurement (which would require N constraints). This may be illustrated by noting that for spatially uniform noise, the χ^2-constraint leads in the ME solution to $\partial/\partial\mathbf{r} \sum_i (\mathbf{m} - P\mathbf{r})_i^2 \propto P(\mathbf{m} - P\mathbf{r})$. Compared to the noise-free constraint $(\mathbf{m} - P\mathbf{r}) = 0$, an additional operator P appears which blurs the measurement constraint. This becomes especially important when a high agreement of the restoration with the measurement is demanded. The convergence of iterative χ^2-algorithms is therefore expected to be slow if the noise level is small. In fact, numerical simulations showed that all tested χ^2-algorithms became rather inefficient if the noise level $\sigma_m / \langle \mathbf{m} \rangle$ is below 1 %.

In our application to scanned solar maps, the resolution is limited since the PSF filters out high spatial frequencies. They can partly be recovered by the restoration described below that leads to an improvement of spatial resolution. In the imaging of the dynamic sun, a large number of images with typically $> 10^4$ pixels is to be restored with only little noise present. As this is expected for planned submillimeter observations, we developed a fast ME algorithm which works optimally at small noise levels. Altough the algorithm is based on zero-noise constraints it also gives good results in the presence of small noise. It takes advantage of linear inversion techniques and accelerates the restoration of nearly noise free data by a factor 10 compared to earlier published algorithms.

2. Fast algorithm

As the ME solution is non-linear and the image size is typically $> 10^4$ pixel, the numerical treatment must be iterative. Although there are powerful χ^2-algorithms available such as those of Cornwell & Evans (1985) or Skilling & Bryan (1984), there were only few attempts to improve existing zero-noise algorithms.

The zero-noise restoration implies the introduction of N Lagrange multipliers $\boldsymbol{\lambda} = \lambda_1 .. \lambda_N$ (one for each constraint), and the quantity to be maximized is: $S(\mathbf{r}) + \sum \lambda_i (\mathbf{m} - P\mathbf{r})_i$. For the case of a flat prior $\mathbf{a}$ and ignoring normalization for the moment, the noise-free ME solution reads:

$$\left.\begin{array}{rcl} \mathbf{r} & = & e^{P\boldsymbol{\lambda}} \\ P\mathbf{r} & = & \mathbf{m} \end{array}\right\} \quad \text{unpriored ME solution} \tag{3}$$

where the Lagrange multipliers $\boldsymbol{\lambda}$ must be adapted such that $P\mathbf{r} = \mathbf{m}$ is satisfied. In the following derivation, all notations of the form $\mathbf{x} = \mathbf{yz}$, $\mathbf{y}/\mathbf{z}$, $e^{\mathbf{y}}$ etc. mean elementwise operations.

Existence and uniqueness of the ME solution. Before starting the design of an iterative algorithm it is necessary to make sure that there is an unique solution to be found by the algorithm. This can be done by generalizing a convexity argument from Skilling & Bryan to the noise-free constraint $\mathbf{m} = P\mathbf{r}$.

Although the restoration itself is unique, this is not the case for the Lagrange multipliers. From Eq. (3) we notice that the ME solution has less degrees of freedom than Lagrange multipliers as it is restricted to a smooth class of images given by $\mathbf{r} \propto e^{P\boldsymbol{\lambda}}$. There is always one (and only one) member of this class which satisfies the measurement constraints, but the non-uniqueness of the Lagrange multipliers will have consequences for the convergence of our algorithm. We shall come back to this point later.

Fast algorithm. Let us assume now that we have some initial guess $\mathbf{r}_0(\boldsymbol{\lambda}_0) \propto e^{P\boldsymbol{\lambda}_0}$ and want to modify $\boldsymbol{\lambda}_0$ in order to improve the agreement with the measurement. A test variation

$$\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_0 + \delta\boldsymbol{\lambda}$$

produces a new restoration $\mathbf{r}_1$ according to Eq. (3):

$$\mathbf{r}_1 = \mathbf{r}_0 \, e^{P\,\delta\boldsymbol{\lambda}} \tag{4}$$

The optimal choice of $\delta\boldsymbol{\lambda}$ must also hold near the solution, where the corrections are small:

$$\delta\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_0 \simeq \mathbf{r}_0 \, P \, \delta\boldsymbol{\lambda} \, . \tag{5}$$

For our application to single-aperture antennas, the operator P is usually positive semidefinite or numerically close to 'semi' implying that $\mathbf{x}(P\mathbf{x}) \geq 0$ for any test vector $\mathbf{x} \neq 0$. Geometrically, this indicates that the image $P\mathbf{x}$ of $\mathbf{x}$ points into the half-space spanned by $\mathbf{x}$, including its hyperplane boundary $P\mathbf{x} = 0$. Hence, a trial variation of the Lagrange multipliers

$$\delta\boldsymbol{\lambda} \propto (\mathbf{m} - P\mathbf{r}_0) \tag{6}$$

ensures that the relative correction $\delta\mathbf{r}/\mathbf{r}_0$ (Eq. 5) is not opposite to the residual which is an approximate condition for uniform convergence. Its direction however is not optimum and iterating the above trial variation converges slowly. This has been noted by Hollis et al. (1992), who suggest another empirical trial variation, namely $\delta\boldsymbol{\lambda} = ln(\mathbf{m}/P\mathbf{r}_0)$ that still needs some 100 iterations to converge.

Our algorithm searches for the optimum direction of $\delta\boldsymbol{\lambda}$ by demanding that the new residuals must vanish:

$$\mathbf{m} - P\,(\mathbf{r}_0 + \delta\mathbf{r}) = 0$$

and therefore $P\delta\mathbf{r} = \mathbf{m} - P\mathbf{r}_0$. Since $\delta\mathbf{r} \simeq \mathbf{r}_0 P\delta\boldsymbol{\lambda}$, the Lagrange multipliers should change according to

$$\delta\boldsymbol{\lambda} = P^{-1}\frac{1}{\mathbf{r}_0}P^{-1}(\mathbf{m} - P\mathbf{r}_0)\,. \tag{7}$$

Although P is singular or ill-conditioned, we can replace P^{-1} by a – not yet specified – pseudoinverse (regularized inverse) $\hat{P}^{-1}$. This represents the optimum choice of the trial variation $\delta\boldsymbol{\lambda}$ in the framework of linear solutions to the linearized inversion problem (5) and we can expect to approach the ME solution in a direct and efficient way.

There is one technical point to be considered: as the pseudounity $\hat{P}^{-1}P$ is not equal to the identity operator and the linearization (5) does not hold for too large exponents in Eq. (4), it is necessary to stabilize the iteration. This can be done by adding a small, positive vector $\boldsymbol{\epsilon} = (\epsilon, ..., \epsilon)$ to the denominator in Eq. (7). The iteration step reads then

$$\begin{aligned}\mathbf{r}_{n+1} &= \mathbf{r}_n\, e^{P\,\delta\boldsymbol{\lambda}_n}\\ \delta\boldsymbol{\lambda}_n &= \hat{P}^{-1}\frac{1}{\mathbf{r}_n + \boldsymbol{\epsilon}}\,\hat{P}^{-1}\,(\mathbf{m} - P\mathbf{r}_n)\,.\end{aligned} \tag{8}$$

A small ϵ leads to faster convergence with increasing tendency to instability. It has been found empirically that the fastest, yet stable, trade-off is achieved when

$$\epsilon \lesssim \langle\mathbf{r}\rangle = 1/N\,. \tag{9}$$

Due to Eq. (8), $\mathbf{r}_{n+1}$ has the requested form of a ME expression assuming that $\mathbf{r}_n$ itself could have been written as $e^{P\boldsymbol{\lambda}}$. This is easily ensured by a sufficiently smooth initial guess such as a flat distribution or a low-pass filtered measurement.

Positivity and Normalization. Positivity of the restoration is obviously preserved by (4). Normalization should be introduced by a further Lagrange multiplier η into the ME expression (3), yielding $\mathbf{r} = e^{P\boldsymbol{\lambda}+\eta}$. The normalization multiplier must be adapted such that $\sum r_i = 1$, leading to $\mathbf{r} = e^{P\boldsymbol{\lambda}}/\sum e^{P\boldsymbol{\lambda}}$. This suggests a re-normalization of the restoration after every iteration cycle. Numerical tests show that this simple method can be applied.

Convergence to the ME solution. The algorithm (8) converges if $P\delta\boldsymbol{\lambda} = 0$. To ensure that the unique ME solution was found, this must occur only if the residuals vanish. By multiplying Eq. (5) by P we obtain

$$P\delta\boldsymbol{\lambda} = P\hat{P}^{-1}\underbrace{\frac{1}{\mathbf{r}+\epsilon}\underbrace{\hat{P}^{-1}P(\mathbf{t}-\mathbf{r})}_{(1)}}_{(2)}$$

and analyze the terms (1) and (2). With the zero-noise assumption $\mathbf{m} = P\mathbf{t}$ and vanishing residuals $\mathbf{m} - P\mathbf{r} = P(\mathbf{t}-\mathbf{r}) = 0$ it follows that $(\mathbf{t}-\mathbf{r})$ lies in the nullspace $\mathcal{N}(P)$ of P. As only the latter must be mapped by $\hat{P}^{-1}P$ onto zero (term 1), we demand that the nullspaces of $\hat{P}^{-1}P$ and P coincide. Next it is plausible (although not proven in mathematical rigor) that the term (2) is either zero or it contains components (called the range of $\hat{P}^{-1}P$) which have passed $\hat{P}^{-1}P$. Finally, we have to ensure that these components – and thus non-vanishing residuals – are not zeroed by the subsequent application of $P\hat{P}^{-1}$. This requires that the range of $\hat{P}^{-1}P$ is not within the nullspace of $P\hat{P}^{-1}$. In summary, the conditions for finding the unique ME solution are:

$$\mathcal{N}(\hat{P}^{-1}P) = \mathcal{N}(P) \text{ and } range(\hat{P}^{-1}P) \not\subset \mathcal{N}(P\hat{P}^{-1}) \tag{10}$$

where $\mathcal{N}(\hat{P}^{-1}P)$ and $\mathcal{N}(P\hat{P}^{-1})$ are the nullspaces of $\hat{P}^{-1}P$ and $P\hat{P}^{-1}$, respectively. A construction of $\hat{P}^{-1}$ which allows to control its resolving power and which satisfies Eq. 10 is discussed later.

Stability to small Noise. Real measurements are not noise-free. Small-scaled features in the noise are not in the range of the smoothing operator Pe^P so that the inverse problem $\mathbf{m} + \mathbf{n} = Pe^{P\boldsymbol{\lambda}}$ (Eq. 1 and 3) is ill-posed.

However, numerical simulations show that the zero-noise algorithm gives a good approximation to the χ^2-constrained solution when it is stopped after a small number of iterations. Testing with gaussian noise, we found that the termination criterion $\chi^2 \le 4\chi^2_{expected}$ yields the best agreement of the fast ME approximation with the correctly χ^2-constrained ME restoration. It turned also out that this termination criterion makes the fast ME restoration almost insensitive to the choice of $\hat{P}^{-1}$, because the resolving power of $\hat{P}^{-1}$ and the number of needed iterations compensate.

In order to economize computation time one tends to use a highly-resolving pseudoinverse $\hat{P}^{-1}$ with a high convergence speed. However, this amplifies small-scaled noise and can lead to instability. A compromize has to be made. To use the highest speed of convergence, the noise growth rate has to be estimated. Since we are dealing with small noise levels, the effect of the measurement noise $\mathbf{n}$ in one iteration step can be linearized to yield the relative error $\frac{\Delta\mathbf{r}_{n+1}}{\mathbf{r}_{n+1}} \simeq P\hat{P}^{-1}\frac{1}{\mathbf{r}_n+\boldsymbol{\epsilon}}\hat{P}^{-1}\mathbf{n}$. A worst-case estimate for the maximum noise growth rate q is obtained if we neglect the $\mathbf{r}_n$-term in the denominator and estimate the maximum value of the other terms:

$$\left|\frac{\Delta\mathbf{r}}{\mathbf{r}}\right| \le \epsilon^{-1}\,M\,|\mathbf{n}| \doteq q\,, \tag{11}$$

where $M = \sqrt{\max|\lambda_i|}$ with the eigenvalues λ_i of $(\hat{P}^{-1}\hat{P}^{-1}P)(\hat{P}^{-1}\hat{P}^{-1}P)^T$ and where the absolute signs denote the L_2-norm (for $\langle\mathbf{n}\rangle = 0$, $|\mathbf{n}|$ becomes $\sqrt{N}\sigma_n$) and ϵ the previously defined stability parameter (Eq. 9). From numerical simulations we obtained the empirical criterion

$$q \lesssim 0.7\sqrt{N} \tag{12}$$

which ensures that the residuals become not dominated by noise during the restoration. Combining our definition of the noise level ($\sigma_n/\langle\mathbf{m}\rangle$) and the empirical estimate of Eq. (9), the stability criterion may be written as $M\,\sigma_n/\langle\mathbf{m}\rangle \lesssim 0.7$.

Choosing the pseudoinverse. In the previous paragraphs we discussed two criteria for the choice of $\hat{P}^{-1}$ involving the nullspace and the maximum noise growth rate. Controlling of the nullspace becomes particularly simple when $\hat{P}^{-1}$ is constructed in P's eigensystem or by singular value decomposition (SVD). The operators P and $\hat{P}^{-1}$ are characterized by their eigen- or singular values and their nullspaces are spanned by base vectors corresponding to zero or small values below the noise level. The first condition in Eq. (10) can be verified by testing if the eigen- or singular values of P and $\hat{P}^{-1}P$ are pairwise below or above the noise level. Another advantage of constructing $\hat{P}^{-1}$ in this manner is that the error growth rate (Eq. 11) is directly obtainable.

If the number of pixels is small enough (< 200 for numerical calulations on our Sparc Workstation), P can be represented by a matrix and its pseudoinverse is found by standard methods known from the solution of ill-posed or ill-conditioned linear equations (see, e.g. Numerical Recipes, Menke 1984, Mammone & Rothacker 1987, Lannes 1987). The SVD-method is computational efficient for one-dimensional simulations and the nullspace and noise growth rate are easily controlled (see above). Therefore we chose it for the examples shown in Figs. 1 and 2.

The SVD pseudoinverse is obtained by ignoring all contributions from singular values w_i smaller than $max(w_i)/C$, where C is the maximum allowed singular value ratio. C determines the resolving power of $\hat{P}^{-1}$ and the maximum noise growth rate by substituting $M = C$ in Eq. (11). The SVD-pseudoinverse acts like a filter with a sharp cutoff for small singular values which form P's nullspace and it is consistent with the first condition in Eq. (10). The demand of a limited maximum noise growth rate (Eq. 12) means that the cutoff occurs before the signal-to-noise ratio has fallen to 1.4 (assuming flat signal- and noise spectra).

In dealing with large 2D-images it is numerically more efficient to define the pseudoinverse as a convolution with a pseudoinverse kernel. This means that the deconvolution becomes a filtering in the Fourier-space and numerical calculations become fast by the use of FFT-based routines. This principle was applied to the solar maps shown in Sect. 4, where the pseudoinverse kernel was calculated by few (< 10) Schultz-Hotteling type iterations from the PSF (see Numerical Recipes 1988):

$$\hat{P}^{-1}_{2n+1} = 2\hat{P}^{-1}_n - \hat{P}^{-1}_n P \hat{P}^{-1}_n \,. \tag{13}$$

The pseudoinverse is determined by the initial guess $\hat{P}^{-1}_0$ and the number of iterations. It was found that for our examples of positive and normalized P, the inclusion relation (10) is satisfied for $\hat{P}^{-1}_0 = 1$ and only few (< 10) iterations are necessary. The maximum noise growth is continually computed from the filter coefficients and the Schultz iteration is stopped as soon as Eq. (12) is violated.

Although the SVD- and Schultz-Hotteling constructions imply the a successful choice of the pseudoinverse, it is advisable to verify the inclusion relation (10) by applying all operators to a white noise test data set and by comparing the spectral cutoffs of its images. In addition, the residuals should be displayed during the iterations to control convergence.

Introducing a prior into the zero-noise restoration. If a prior $\mathbf{a}$ is introduced into the linearized scheme Eq. (5), a fundamental problem will occur. The priored ME-solution is

$$\left.\begin{array}{rcl} \mathbf{r} &=& \mathbf{a}\, e^{P\boldsymbol{\lambda}} \\ P\mathbf{r} &=& \mathbf{m} \end{array}\right\} \quad \text{priored ME solution}$$

A test variation in the Lagrange multipliers leads again to $\delta\mathbf{r} = \mathbf{r}\, e^{P\,\delta\boldsymbol{\lambda}}$, so that the prior has no influence on $\delta\mathbf{r}$. We solved this problem by the substitution

$$\mathbf{r}' = \frac{\mathbf{r}}{\mathbf{a}}$$

and iterating $\mathbf{r}'$ instead of $\mathbf{r}$. Then the non-normalized ME solution reads $\mathbf{r}' = e^{P\boldsymbol{\lambda}}$ and $P(\mathbf{a}\mathbf{r}') = \mathbf{m}$. A test variation $\delta\boldsymbol{\lambda}$ produces again a variation $\delta\mathbf{r}' = \mathbf{r}' P \delta\boldsymbol{\lambda}$, the prior now taken into account by the measurement constraint $\mathbf{m} - P(\mathbf{a}\mathbf{r}') = 0$. The remaining derivation, including the estimate of the noise growth rate, is done as before. Of course, normalization must now be enforced with respect to $\mathbf{r} = \mathbf{a}\mathbf{r}'$.

3. Numerical simulations

In order to test our fast ME algorithm different types of model brightness distributions and point spread functions were convolved and different amounts of random noise added to generate simulated 'measurements'. These were then restored and compared to the original distributions and to the results of other deconvolution methods. As we want to apply our algorithm to solar scans made by a radio telescope, the input assumptions for the simulations were as follows.

The model PSF was chosen as gaussian or a standard beamshape of one- or twodimensional apertures such as $sinc^2$ resp. $J_1(\rho)^2/\rho$ where ρ is the wavelength normalized distance from the beamcenter. The latter results from an uniform aperture illumination and shows the strongest sidelobes expected from a single-beam antenna. We found empirically that our fast ME restoration was unaffected by the presence of these high sidelobes.

True brightness distributions similar to solar ones were modelled by the superposition of rectangular and gaussian peaks whose number, separation and height was varied in order to test the sensitivity of our restoration to small-amplitude and small-scaled features. All model (true) distributions were tapered by P^2 at the image's edge in order to avoid problems with non-restorable edge points.

The measurement noise is determined by the observing system which, in our application, is an antenna connected to a radio receiver. The output fluctuations of the receiver contain

contributions from the antenna as well as from the receiver system and are gaussian. As the solar maps contain finestructures whose amplitude is relatively small compared to the background (quiet sun plus receiver) intensity, the noise level was assumed to be constant across the brightness distributions, so that χ^2 becomes proportional to the variance. By varying the gaussian noise level, we found that a useful application of the zero-noise approximation is possible for noise levels below 3%. Above this limit, traditional least-square methods such as the Cornwell algorithm become superior with regard to computational efficiency.

For simplicity the same regular grid was used for $\mathbf{m}$, $\mathbf{r}$, and the PSF. In the one-dimensional simulations, the number of data points N was typically set to 100. This allows a rapid (a few seconds) SVD-inversion by our SUN Workstations. The gridsize should be such that the PSF has a width of ≈ 8 pixels, not less than 4 pixel but also not more than 12 pixel because the expected resolution enhancement (see Cornwell & Evans) in the restored images is in the order of a factor 2 and an oversampling over the obtainable resolution limit should be avoided in order to economize computation time.

As had been discussed in Sect. 2, our fast ME algorithm relies on a zero-noise assumption and gives only an approximate ME restoration if a small amount of noise is present in the measurement. Fig. 1 compares results of the fast ME restoration with those for the correctly χ^2-constrained ME restoration obtained by the Cornwell algorithm (Cornwell & Evans 1985). The true distribution (top left) is an irregular superposition of several overlapping gaussian peaks separated by roughly the half power beamwidth (HPBW) of the gaussian PSF (dotted). The simulated measurement (top right) is the convolution of the true distribution with the PSF plus a noise level of 0.8 %. The quality of the fast ME approximation may be judged from its difference (dotted line) to the correct χ^2-restoration whose rms is 4%. A SVD pseudoinverse with an allowed singular value ratio of $C = 50$, $\epsilon = 0.6 \times N$ and a maximum noise growth rate of $q = 0.7\sqrt{N}$ was chosen. The algorithm was stopped after 3 iterations when the best (rms) agreement with the correct χ^2-restoration was reached. The exact χ^2-solution itself was found by continuating the Cornwell algorithm up to several thousands of iterations, the convergence controlled by the test parameter proposed in (Skilling & Bryan, 1984). The Cornwell algorithm needed approximatively 30 iterations to reach the same residual level as the fast ME restoration.

In order to find the optimum termination, extended numerical simulations as in Fig. 1 were carried out with different model (true) distributions and PSF's mentioned above. Using the measurement as initial guess and working with SVD-pseudoinverses, it was found that the best (rms) fit to the correctly χ^2-constrained Cornwell restoration is obtained with the termination criterion $|\mathbf{m} - P\mathbf{r}| \leq 2|\mathbf{n}|$. In average, this criterion was reached after 4 ± 2 iterations and the fast ME approximation fits the correctly χ^2-constrained solution within 5 % rms. Compared to the Cornwell algorithm, when applied to the same low-noise measurements, the computational burden for an equivalent approximation was by a factor 10 smaller. It

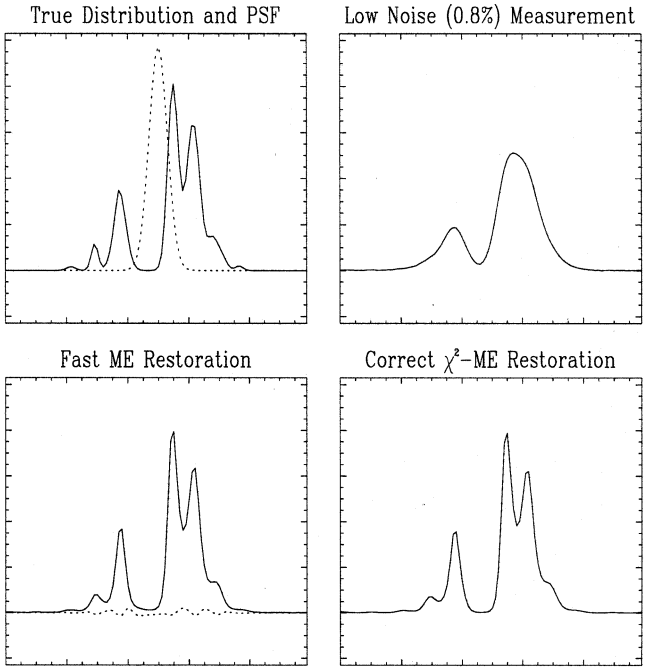

Fig. 1. Computer simulation of the fast ME approximation (bottom left) after 3 iterations compared to the correctly χ^2-constrained ME restoration (bottom right). Both restorations were obtained from a small-noise (0.8 %) measuerement (top right). The reliability of our algorithm may be judged by the difference between the fast and correctly χ^2-constrained restorations (dotted curve, bottom left).

turned out that the number of iterations needed by the fast ME algorithm is inverse to the resolving power of $\hat{P}^{-1}$ and that the choice of a low-resolving pseudoinverse must be compensated by a larger number of iterations and that the result of the fast ME restoration is almost unaffected by the choice of $\hat{P}^{-1}$.

The influence of an a priori distribution in the presence of small and large noise is illustrated in Fig. 2. The true distribution consists of a superposition of 4 gaussian peaks placed on top of a bright, rectangular box modelling a known background sun. The PSF is a $sinc^2$-function (dotted line) with a HPBW of 1.5 times the distance between the two narrowest peaks. The left column shows the low-noise (0.3%) case where our fast ME algorithm is applied for unpriored and priored restoration. During our investigation of faster restoration algorithms we also found a modified Cornwell algorithm for noise levels above 4 %. The modification concerns the controlling of the normalization- and χ^2-constraints and speeds up convergence by a factor two if the noise level is above 4 %. Its results are shown in the right column for a noise level of 3.9 %. As expected, the higher noise level in the right column reduces the number of distinguishable details in the restoration. The third row contains the unpriored restorations for low and heigh level noise. Note that overshoot artefacts at the edges of the rectangular box become smaller for the priored restorations (bottom line Fig. 2) as expected.

As can be seen from Fig. 2, the resolution enhancement from the measurement to the restoration depends on the noise level. It was found from simulations with point sources and gaussian

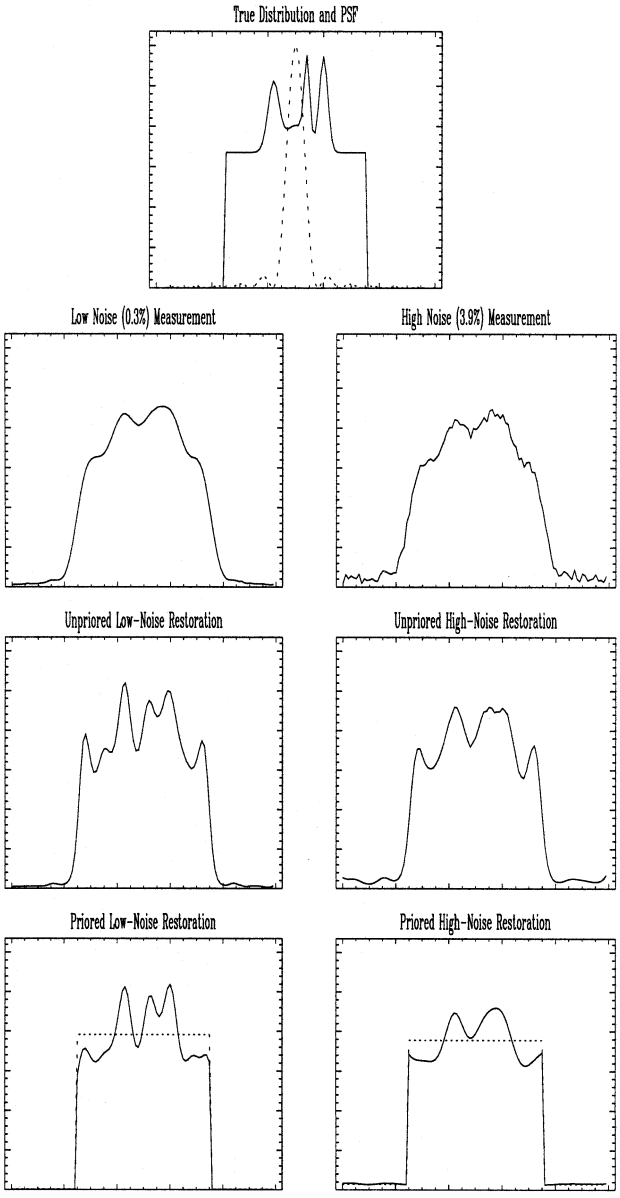

Fig. 2. Computer simulations for different noise levels and an a priori box distribution. Left column: low noise (0.3%) measurement and fast ME restorations. Right column: High-Noise (3.9 %) measurement and χ^2-restorations.

PSF's that for low noise levels ($< 1\%$) resolution enhancement by a factor ≈ 2 is possible, the resolution defined as point-source separability. It was also confirmed what had been noted by many authors before (e.g. Narayan 1986) that peaks with higher amplitudes are generally better restored.

Finally, the restorations were compared to those obtained by the classical Lucy algorithm (Lucy, 1974). The results generally agree if the pseudoinverse is low-resolving (e.g. $q = 0.3$).The Lucy algorithm produces slightly less artefacts if the true distribution contains sharp-edged structures, but converges (about one order of magnitude) slower than the fast ME algorithm.

4. Application to solar images

Our fast ME algorithm was also applied to real observational data, consisting of solar scans recorded by the Nobeyama Radio Observatory (NRO), Japan (Irimajiri et al. 1995). The 45 m antenna of the Nobeyama Observatory recorded maps simultaneously at 36, 89 and 110 GHz. As the beamsize scales linearly with the observed frequency, the maps have different resolutions. This allows the comparison of the ME restored low-resolution map with corresponding higher resolved observations and a qualitative judgement of the restoration quality.

The HPBW of the 45-m telescope is 46" at 36 GHz and 19" at 89 GHz. The scan velocity is small enough that additional beam-smearing due to the scan movement can be neglected. The scans were recorded in a radial scanning mode and calibrated for atmospherical absorption fluctuations (Kosugi et al. 1986). After rebinning the data to cartesian coordinates, the maps were smoothed (Irimajiri et al. 1995) by convolving them with a gaussian beam whose width equals to the HPBW at each observing frequency. By this procedure, possible sidelobes of the antenna beam were smoothed out and the total PSF (antenna beam plus additional smoothing) can be approximated by a gaussian of width $1.5'$ for 36 GHz and 27" at 89 GHz, respectively. The solar maps (Fig. 3) contain 256×256 pixels and cover $20' \times 20'$. Due to the long integration time and the additional smoothing, the images carry only 0.2% noise and allow the application of the fast ME algorithm. The mentioned noise level was estimated by subtracting a running mean over 4 pixels from the 36 GHz-measurement.

The left column of Fig. 3 on page 741 shows a section of the solar 36 and 89 GHz-maps from May 28, 1991 with the PSF's at the left bottom of the images. The fast ME restoration of the 36-GHz map (top right) was obtained after 4 iterations with the fast ME algorithm (Eq 8). The pseudoinverse was performed by convolving the data with a pseudoinverse kernel obtained by 5 Schultz type iterations from the 36 GHz-PSF. The pseudo-unity $\hat{P}^{-1}P$ is shown at the left bottom of the fast ME restoration and compared to the 36 GHz PSF, it gives a measure for the resolving power of $\hat{P}^{-1}$.

For comparison, also the result of 500 iterations with the Hollis & Dorbrand - algorithm (see Sect. 2) is shown in the bottom right image. The differences to the fast ME restoration are small (2% of the image's dynamic range). To give an impression on the needed calculation time, we mention that the computation of the fast ME restoration took about 2 minutes on a Sparc 10 Workstation.

The image shown in Fig. 3 was chosen not to contain the solar limb for two reasons: (a) to emphasize the details of the finestructures which reach only 10% of the solar disk's background; (b) to avoid artefacts occuring at the limb of the solar disk that could not be damped by a flat disk prior analog to the one-dimensional example shown in Fig. 2. This different behaviour is not well understood yet but is assumed to be due to an insufficient prioring by a simple circular disk (ignoring limb-brightening and uncertainty of the exact position and ra-

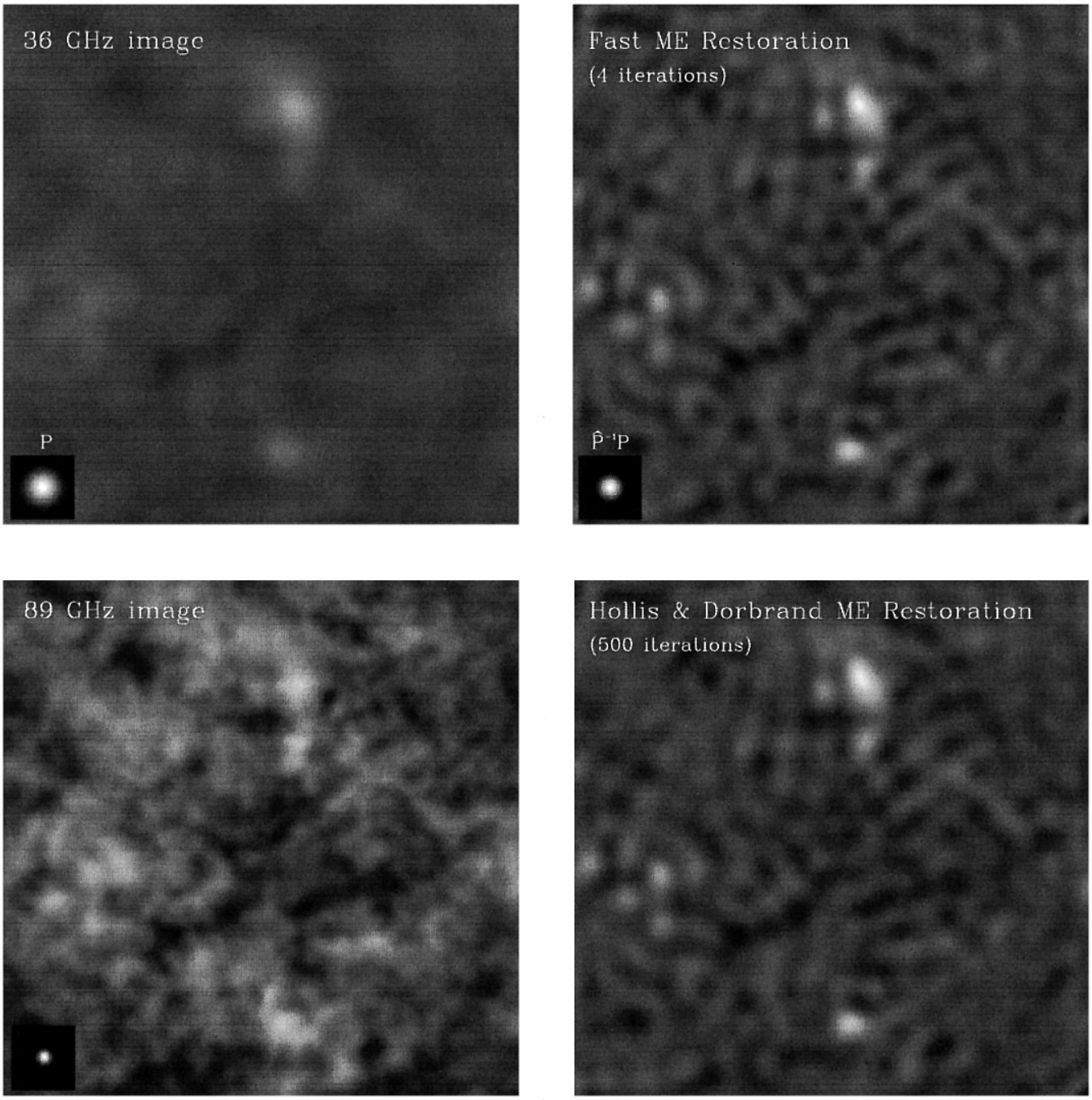

Fig. 3. A $20' \times 20'$-section of a solar radio map recorded by the Nobeyama 45 m telescope on May 28. 1991. The top left Fig. shows a map at 36 GHz and the corresponding point spread function P, the bottom left image the (higher resolved) 89 GHz map. The top right image results from the fast ME restoration of the 36 GHz image after 4 iterations with a pseudounity $\hat{P}^{-1}P$ shown at the bottom of the image. For comparison the bottom right image shows the result of 500 iterations with the Hollis & Dorbrand algorithm also applied to the 36 GHz image. The color scales are equal for the 36 GHz map and the restorations derived from it and zoomed to the fill available range as is the 89 GHz map.

dius of the radio sun). We hope to solve this problem in future by a better choice of the pseudoinverse and the prior.

In the absence of known point sources in the true image, the resolution cannot be defined by the separability of point sources as had been done in the simulations. We therefore suggest the following definition, which uses the baricenter of the distribution's power-spectra to define a typical resolution Res:

$$Res := \frac{2\pi}{< k >} \quad \text{with} \quad < k >= \frac{\sum |\mathbf{k}| P(\mathbf{k})}{\sum P(\mathbf{k})} .$$

P is the spatial power spectrum of the images after removing their DC part and the sums have to be taken from zero to the Nyquist frequency. With this definition, the resolution enhancement of the fast ME is a factor 2.1, whereas the Hollis & Dorbrand restoration yields a factor 2.0.

Comparing the images in Fig. 3, it becomes obvious that many of the features present in the highly resolved 89-GHz map (bottom left) reappear in the restored 36 GHz maps. However, we cannot expect to see a one-to-one correspondance between the brighness distributions at 36 GHz and 89 GHz, as their emission originates from different heights in the solar atmosphere. Morover, the PSF is only approximatively known and the influence of possible sidelobes – although smoothed before the restoration – might slightly change the result. In spite of this, the coincidence of many restored 36GHz-finestructures with the 89 GHz map indicates at least that they are likely to be real.

5. Summary

A fast Maximum Entropy algorithm was developed which is applicable to low-noise images such as the Nobeyama solar maps. The idea is to use restrictive zero-noise constraining and to take advantage of linear inversion methods for the choice of an optimum trial variation in the Lagrange multiplier space.

The linear inversion involves the choice of a pseudoinverse according to well-defined criteria. Its properties determine the trade-off between stability against noise and convergence speed. Numerical simulations show that a good trade-off between convergence speed and noise-amplification is possible for noise levels below 3%. Above this limit, traditional χ^2-constraining becomes superior regarding computational efficiency. When applied to low-noise measurements and terminated according to a residual criterion described in Sect. 2, our fast ME algorithm (using zero-noise constraining) yields a <5%-approximation to the correctly χ^2-constrained ME solution. This approximation can be used for itself or as initial guess for a classical χ^2-algorithm, thereby accelerating its convergence.

The algorithms and support programs were written in IDL. The algorithms used in the calculation of $\hat{P}^{-1}$ are from Numerical Recipes 1988.

Acknowledgements. We thank Drs. T. Takano and H. Nakajima from the NRO for supplying solar radio maps; J. Costa (University Sao Paulo, Brazil), Eric Rolli and Th. Inaebnit (University Bern, Switzerland) for further (unpublished) test-data and discussions. This work was supplied by the Swiss National Foundation under the project number 20-42265.94.

References

Ables J.G. 1974, Maximum entropy Spectyral analysis, A&AS **15**, 383-393.

Bryan R.K., Skilling J. 1980, Deconvolution by maximum entropy, as illustrated by applications to the jet of M87, MNRAS **191**, 69-79.

Cornwell T.J., Evans K.F. 1985, A simple ME deconvolution algorithm, A&A **143**, 77-83.

Eriksson K.-E., Lindgren K., Månsson B.A. 1987, Structure, Context, Complexity, Organization, World Scientific Publishing Co. Pte. Ltd., ISBN 9971-50-023-4.

Gull S.F., Daniell G.J. 1978, Image reconstruction from incomplete and noisy data, Nat Vol. **272**, 686-690, April 1978.

Helstrom, C.W. 1966, Image Restoration by the Method of Least Squares (1966), J. Opt. Soc. Am. **57**, 3, 297-303.

Hollis, J.M., Dorband, J.,E., Yusef-Zadeh, F. 1992, Comparing restored HST and VLA imaginary of R Aquarii, ApJ **386**, 293-298.

Irimajiri Y., Takano T., Nakajima, H., et al. 1995, Simultaneous Multifrequency Observations of an Eruptive Prominence at mm Wavelengths, Solar Phys. **156**, 363-375.

Kikuchi R., Soffer H.B. 1977, Maximum Entropy Restoration. I. The entropy expression, J. Opt. Soc. Am. Vol **67**, Nr. 12.

Kosugi, T, Ishiguro, M., Shibasaki, K. 1986, Polar-Cap and Coronal-Hole-Associated Brightening of the Sun at Millimeter Wavelengths, PASJ **38**, 1-11.

Lannes A., Roques S., Casanove M.J. 1987, Resolution and robustness in image processing: a new regularization principle, J. Opt. Soc. Am., Vol. **4** No. 1, 189-199.

Lucy, L.B. 1974, An iterative technique for the rectification of observed distributions, AJ Vol. **79**, 746-754.

Mammone R.J., Rothacker, R.J. 1987, General iterative method of restoring linearly degraded images, J. Opt. Soc. Am A / Vol **4**, No.1, Jan. 1987

Menke, W. 1984, Geophys. Data Analysis: Discrete Inverse Theory, Internat. Geophys. Series Vol. 45, Academic Press. ISBN 0-12-490921-3.

Molina R., Del Olmo A., Perea J., Ripley B.D. 1992, Bayseian deconvolution in optical astronomy, AJ Vol. **103** 2, 666-675.

Narayan R. 1986, Maximum Entropy Image Restoration in Astronomy, ARA&AR Vol. **24**, 127-170.

Press W.H., Teulkolsky, S.A, Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C $2^n d$ ed, Cambridge Univ. Press (1988), ISBN 0-521-43108-5.

Skilling J., Bryan R.B. 1984, Maximum entropy image reconstruction: general algorithm, MNRAS vol. **211**, 111-124.

Tikhonov, Arsenin, 1977, Solution of ill-posed problems (Wilney, New York).

Tikochinsky Y, Tishby N.Z., Levine, R.D. 1984, Consistent Interference of Probabilities for Reproducible Experiments, Phys. Rev. Letters **52** No 16, 1357-1360.

Van Trees H. 1968, Detection, Estimation, and Modulation Theory Part 1, p. 481 ff., John Wilney & Sons, New York, 1992.

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.