Parameter Uncertainties in IMFIT
CASA Requirements Document
Brian S. Mason (NRAO)
March 5, 2014

The CASA task IMFIT fits elliptical Gaussians to image data. As reported by several users in CASA tickets (CAS-3476, CAS-5879; see also the parent ticket CAS-6124) parameter uncertainties reported in CASA v4.2 are currently unsatisfactory in several ways:

1. Uncertainties relating to the position and size of the Gaussian are generally too small;
2. the uncertainty for peak and integrated flux density are in some cases too large;
3. the error estimates appear to be based on the RMS of fit residuals in the designated region of interest (possibly the main cause of #1).

The errors purport to be

```
longitude error = (FWHM beam width in longitude direction)/(2*S)
latitude error = (FWHM beam width in latitude direction)/(2*S)
FWHM of major axis error = (FWHM beam along major axis)/sqrt(S)
FWHM of minor axis error = (FWHM beam along minor axis)/sqrt(S)
```

where $S =$the fitted component peak amplitude divided by the noise level. Mention is also made of expressions given in § 2.1.1 of AIPS++ note 224, which only apply for the case of normally distributed, uncorrelated pixel noise. The basis of the error in the normalization is not clear, and the errors are clearly too large in some cases. There are several problems with the given expressions: 1) it is claimed that the FWHM errors are too large by $\sqrt{2}$, and have the wrong scaling with S; 2) they do not appear to explicitly account for noise correlations; 3) the flux errors are wrong; 4) since the pixel size does not appear, the positional and size errors given are manifestly incorrect for the *uncorrelated* pixel noise. In this note I attempt to specify a satisfactory solution for these issues.

Condon (1997, hereafter C97) has derived expressions which address most of these shortcomings. I use them as the basis to specify expressions which will give more useful parameter uncertainties in IMFIT . I mostly follow the notation of C97. These issues have also been discussed in AIPS memos 92 and 97, discussing what has been implemented[1] in the classic AIPS task JMFIT.

1 Assumptions

Key assumptions made at the outset are:

- The given model (elliptical Gaussian, or elliptical Gaussian plus constant) is an adequate representation of the data.
- An accurate estimate of the pixel noise μ is provided to, or can be derived by, IMFIT . For Gaussian, uncorrelated noise, the current IMFIT approach should be reasonable. For the case of correlated noise (e.g., a CLEAN map), the fit region will need to contain many "beams" or an independent value of RMS will need to be provided to IMFIT . In the future, better methods of direct and automatic estimation should be implemented.
- The SNR of the Gaussian component is high. This comes about because C97 uses a Taylor series to linearize the problem; he states that the fractional bias in the fitted amplitude A due to this assumption is of order $1/\rho^2$, where ρ is the total SNR of the Gaussian with respect to the given data set (defined more precisely below). For a 5σ "detection" of the Gaussian, this is a 4% effect.
- All (or practically all) of the flux in the component being fit falls within the region of interest. If a constant term is free in the fit, the region of interest should be larger still. The derivations of the expressions summarized in this note assume an effectively infinite region.

[1] The AIPS *MFIT routines implement an earlier and slightly less comprehensive version of the C97 equations; they do not account for the effect of fixing parameters on free parameter uncertainties or the effect of a free DC term on the uncertainties of other parameters.

Symbols used in this note are as follows: ρ = overall signal to noise ratio of the Gaussian (defined below); A = normalization of the Gaussian component in question; I is the integrated flux density of the Gaussian component; h = the angular size of an image pixel along one side; θ_M = the major axis (FWHM) of the fitted Gaussian component; θ_m = the minor axis (FWHM) of the fitted Gaussian component; μ = the RMS image noise (regardless of whether this noise is correlated or uncorrelated between pixels); x_o = the center of the Gaussian in the hypothetical coordinate axis *along the fitted major axis of the Gaussian* (*not* the R.A. axis!); y_o = the center of the Gaussian in the hypothetical coordinate axis along the fitted *minor axis* of the Gaussian; ϕ = the position angle of the Gaussian component; θ_N the FWHM of the noise correlation (*e.g.*, the synthesized beam FWHM). $S = A/\mu$ is the estimated peak SNR in terms of surface brightness.

2 Straightforward Approach

2.1 Uncorrelated Pixel Noise

For the case of uncorrelated pixel noise (which should be the default assumption), the overall SNR ρ is found to be:

$$\rho = \frac{A}{h\mu}\sqrt{\frac{\pi\theta_M\theta_m}{8\,ln2}}. \tag{1}$$

This is equation 20 of C97.

The fractional parameter errors for the 6 parameters of the fit (plus the one derived parameter I) are then given by the following relations:

$$\begin{aligned}\frac{\sigma(A)}{A} &= \frac{\sigma(I)}{I} = \frac{\sigma(\theta_M)}{\theta_M} = \frac{\sigma(\theta_m)}{\theta_m} && (2)\\ &= \sqrt{8\,ln2}\frac{\sigma(x_o)}{\theta_M} = \sqrt{8\,ln2}\frac{\sigma(y_o)}{\theta_m} = \frac{\sigma(\phi)}{\sqrt{2}}\left(\frac{\theta_M^2-\theta_m^2}{\theta_M\,\theta_m}\right) && (3)\\ &= \frac{\sqrt{2}}{\rho} && (4)\end{aligned}$$

This is equation 21 of C97.

Note: The reported positional uncertainties must be in the coordinate system of the image, not the hypothetical coordinate system defined by the fitted Gaussian. The proper positional errors are easily derived from $\sigma(x_o)$ and $\sigma(y_o)$ by propagation of errors using the 2D rotation matrix which enacts the rotation ϕ.

2.1.1 Heuristic

This section is not part of the specification

It is instructive to evaluate equations 1 and 2-4 in terms of the volume of the Gaussian component $\Omega_g = 2\pi\sigma_x\sigma_y$ and $S = A/\mu$. The total SNR ρ is

$$\rho = \sqrt{\frac{\pi\sigma_x\sigma_y}{h^2}}\,S = \sqrt{\frac{\Omega_g/2}{h^2}}\,S = \sqrt{\frac{N_p}{2}}\,S \tag{5}$$

where N_p is the number of pixels encompassed by the volume of the Gaussian component. We find for the positional error

$$\sigma(x_o) = \frac{2}{\sqrt{N_p}}\frac{\sigma_x}{S} = \frac{2}{\sqrt{8\,ln2}}\frac{1}{\sqrt{N_p}}\frac{\theta_M}{S} \sim \frac{1}{\sqrt{N_p}}\frac{\theta_M}{S} \tag{6}$$

although the last approximation in this line is only good to 15%. For the major axis width error you find

$$\sigma(\theta_M) = \frac{2}{\sqrt{N_p}}\frac{\theta_M}{S} \tag{7}$$

or approximately 2× the position error (exactly $\sqrt{8\,ln2} = 2.35\times$ the position error).

2.2 Correlated Pixel Noise

For the case that the noise is correlated due to smoothing uncorrelated pixel noise by an approximately Gaussian smoothing kernel of FWHM θ_N, similar expressions (as Eqs. 2 through 4 above) are used but the equation for the SNR ρ is modified to:

$$\rho(\alpha_M, \alpha_m) = \frac{A}{\mu} \frac{\sqrt{\theta_M \theta_m}}{2\theta_N} \left[1 + \left(\frac{\theta_N}{\theta_M}\right)^2\right]^{\alpha_M/2} \left[1 + \left(\frac{\theta_N}{\theta_m}\right)^2\right]^{\alpha_m/2} \tag{8}$$

This is Eq.41 of C97. The power-law exponents α_M and α_m have empirically calibrated values, and different values should be used when calculating uncertainties for different fit parameters:

- For the amplitude A and integrated flux density I, use $\alpha_M = \alpha_m = 3/2$.
- For θ_M and x_o, use $\alpha_M = 5/2$, $\alpha_m = 1/2$.
- For θ_m, ϕ and y_o, use $\alpha_M = 1/2$, $\alpha_m = 5/2$.

Then for *all parameters except the integrated flux density I*, use the relationships in Eq. 2 - 4 along with the appropriate value of $\rho(\alpha_M, \alpha_m)$. For integrated flux density, use this:

$$\frac{\sigma(I)}{I} = \sqrt{\left(\frac{\sigma(A)}{A}\right)^2 + \left(\frac{\theta_N^2}{\theta_N \theta_m}\right)\left[\left(\frac{\sigma(\theta_M)}{\theta_M}\right)^2 + \left(\frac{\sigma(\theta_m)}{\theta_m}\right)^2\right]} \tag{9}$$

Again, positional uncertainties in x_o and y_o will need to be propagated back to the image native coordinate system by rotation.

2.2.1 Definition of the Noise Correlation Scale

This section is not part of the specification

For the case that the noise correlations are Gaussian, θ_N is the FWHM of this noise correlation function. This is probably a good enough approximation for most cases of interest. According to equations 37 and 38 of C97, however, θ_N may be defined for the more general case of a noise-smoothing function $C(x, y)$ as:

$$\theta_N = \sqrt{\frac{2\,ln2}{\pi}} \frac{\int\int\, dx\, dy\, C(x,y)}{\sqrt{\int\int\, dx\, dy\, C^2(x,y)}} \sim 1.50 \times \frac{\int\int\, dx\, dy\, C(x,y)}{\sqrt{\int\int\, dx\, dy\, C^2(x,y)}} \tag{10}$$

2.2.2 Special Case: Point Source

This section is not part of the specification

It is instructive to consider the special case that $\theta_M = \theta_m = \theta_N$. Then Eq. 8 becomes

$$\rho = \sqrt{2}\frac{A}{\mu} = \sqrt{2} S \tag{11}$$

for all parameters. The major axis position error is

$$\sigma(x_o) = \frac{2}{\rho}\sigma_x = \frac{\sqrt{2}}{S}\sigma_x = \frac{\sqrt{2}}{S\sqrt{8\,ln2}}\theta_M \sim \frac{\theta_M}{2\,S} \tag{12}$$

The last approximate equality is the equation currently implemented in `IMFIT` . Note that the approximation is only valid to $\sim 20\%$, and only for a source that has the same apparent size as the noise correlation in the map (an unresolved source). For $\theta_N << \theta_M$ this expression will considerably overestimate the errors.

The major axis error is seen to be

$$\sigma(\theta_M) = \frac{\sqrt{2}}{S}\theta_M \tag{13}$$

Note that the overall factor aside, the current `IMFIT` has a $1/\sqrt{S}$ scaling instead of the correct $1/S$ scaling.

2.3 Implementation

- Perform fit by current methods (e.g., Levenberg-Marquardt minimization of χ^2).
- The noise estimate μ should be obtained from one of the following methods, in order of preference:
  - (*most preferred*) If a positive, non-zero value for the `RMS` parameter has been supplied to `IMFIT` , set μ equal to to this value. Note it is important that `RMS` should default to an undefined or otherwise illegal value.
  - If no value has been specified, look in the image header for a keyword `CHANRMS`; if present with a legal, non-zero value, then set μ equal to this value[2].
  - If no `RMS` parameter has been supplied to `IMFIT` , and no CHANRMS keyword/value is present in the image header, then measure it by taking it equal to the RMS of the residual image in the given region of interest (i.e., default to the current method). Issue a warning.
- There will be a `NOISEFWHM` parameter in `IMFIT` which specifies the noise-correlation beam FWHM θ_N as an angle with specifiable units. Two example legal values of `NOISEFWHM` are `'4.5arcsec'` and `'2arcmin'`. Its default value is `''` (i.e., the null string).
- Support use of both the "correlated-noise" expressions (§ 2.2) and the "uncorrelated-noise" expressions (§ 2.1), according to the following logic:
  - If `NOISEFWHM` has a non-null, positive value, set θ_N equal to this value and use the *correlated noise* equations (§ 2.2).
  - If `NOISEFWHM` is set to a value smaller than the image pixel size (e.g., `0arcsec`, `-1arcmin`, or any positive value smaller than the geometric mean of the pixel dimensions), then force use of the *uncorrelated noise* expressions (§ 2.1). Issue a warning.
  - If `NOISEFWHM` has a *null* value, look for beam information in the image header by standard means (BMAJ, BMIN). If beam information is present in the image header, compute θ_N as the geometric mean of the given beam major and minor axis FWHM values. Compute parameter uncertainties as in § 2.2 for *correlated noise*. If no beam information is in the header, issue a warning and compute parameter uncertainties using the *uncorrelated noise* expressions(§ 2.1).
  - Note: the units of the image (Jy, K, Jy/bm, Jy/pixel) are irrelevant to the choice of how to compute the parameter uncertainties.

2.3.1 Summary of Caveats and Limitations

The following caveats will be highlighted in the `IMFIT` "Help" documentation:

- Fixing Gaussian component parameters will tend to cause the parameter uncertainties reported for free parameters to be overestimated.
- Allowing a free mean level in the fit will tend to cause the reported parameter uncertainties to be slightly underestimated.
- The parameter uncertainties will be inaccurate at low SNR (a $\sim 10\%$ effect at SNR$= 3$).
- If the fitted region is not considerably larger than the largest component that is fit, parameter uncertainties may be mis-estimated.
- An accurate noise measurement μ for the region in question must be supplied. Alternatively, a sufficiently large signal-free region must be present in the region of interest (at least ~ 25 noise beams in area) to derive such an estimate.
- If the image noise is not statistically independent from pixel to pixel, a reasonably accurate noise correlation scale θ_N must be provided. If the noise correlation function is not approximately Gaussian, the correlation length can be estimated using Eq. 10.

[2] The ALMA science archive FITS data products (v1.5 M. Lacy & E. Muller) nominally aim to provide a single number CHANRMS, in Jy/bm, to characterize the noise in an image or cube. This is similar to the AIPS keyword `REALRMS` (AIPS Memo 117) and `ACTNOISE` (which is used by the `*MFIT` routines).

- If fitted model components have significant spatial overlap, the parameter uncertainties are likely to be mis-estimated (i.e., correlations between the parameters of separate *components* are not accounted for).
- If the image being analyzed is an interferometric image with poor *uv* sampling, the parameter uncertainties may be significantly underestimated.[3]

3 More General Approach

This section is not currently part of the specification

The expressions given in § 2 assume a) that all six parameters are fitted for; b) no constant term in the fit. In fact, `IMFIT` allows any number of the fit parameters to be fixed at some given values, and it also allows an optional constant term in the fit. The approach of § 2.2 can be generalized to deal with these cases as follows: express the design matrix of the problem (equation 8 of C97) in terms of ρ, leaving out rows and columns that correspond to fixed parameters and adding in the mean (*ansatz*: Condon's semi-numerical expressions for the correlated noise case work well when some of the parameters are fixed– Jim seemed to think this is a reasonable thing to do); numerically invert this matrix to obtain $\mathbf{D}^{-1}$; and obtain the parameter uncertainties from the square roots of the diagonal elements of this matrix times the given noise estimate μ. This is analagous to the approach of Brouw (AIPS++ note 224), in whose terminology $\mu = \sigma_o$ and $D = A$. The main difference is in accounting for correlated noise via the tweaked definition of the SNR ρ.

To proceed with this agenda refactor $\mathbf{D}$ (eq. 8 of C97) as $\mathbf{D} = \mathbf{\Gamma}\mu^2\rho^2$. Then $\mathbf{\Gamma}$ is found to be:

$$\mathbf{\Gamma} = \begin{pmatrix} 1/A^2 & 0 & 0 & 1/(2A\sigma_x) & 1/(2A\sigma_y) & 0 \\ 0 & 1/(2\sigma_x^2) & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/(2\sigma_y^2) & 0 & 0 & 0 \\ 1/(2A\sigma_x) & 0 & 0 & 3/(4\sigma_x^2) & 1/(4\sigma_x\sigma_y) & 0 \\ 1/(2A\sigma_y) & 0 & 0 & 1/(4\sigma_x\sigma_y) & 3/(4\sigma_y^2) & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/4 \end{pmatrix} \tag{14}$$

The rows and columns of this matrix are in the following order: p_1 = amplitude; p_2 = x position; p_3 = y position; p_4 = x size; p_5 = y size; p_6 = orientation. Rows/columns of fixed parameters should be excluded, ie, a matrix of smaller dimensionality formed. Then this smaller matrix can be inverted numerically and individual parameter uncertainties evaluated as:

$$\sigma(p_i) = \sqrt{(\mathbf{D}^{-1})_{i,i}}\,\mu = \sqrt{(\mathbf{\Gamma}^{-1})_{i,i}}\,\frac{1}{\rho_i} \tag{15}$$

where ρ_i is computed according to the prescription in Eq. 8 for the parameter i in question.

Some areas of future work that may be useful are:

- work out the matrix elements involving the constant term in the fit — it would be worth checking to see if this has already been worked out for JMFIT. I'm not sure if this will break the ability to factor out $\rho\mu$, and therefore make the noise correlation correction of C97 tricky or impossible to implement. Maybe ignoring the influence of the DC term is not too bad if the region of interest is big enough
- Test the applicability of the generalized noise correlation correction proposed here with Monte-Carlo simulations of several cases (fix component position; fix component size; fix component size and position).
- Test parameter estimate accuracy in the poor *uv* coverage case, using some approximate dirty beam to compute a value for θ_N.

[3]It would be interesting to see if using the value of θ_N computed via Eq 10, and using some approximation of the dirty beam, does better.