Draft version November 27, 2017
Preprint typeset using LaTeX style emulateapj v. 12/16/11

AN ADAPTIVE EXTENSION OF BRIGGS WEIGHTING FOR OPTIMAL IMAGING OF RESOLVED SOURCES

Ryan A. Loomis[1,2]

Draft version November 27, 2017

ABSTRACT

In his PhD thesis, Dan Briggs showed that a weighting scheme which tunes between the two extremes of natural and uniform weighting can recover most of the benefits of each (aka 'robust' weighting). This method is now essentially the default weighting for most interferometric imaging. The derivation of robust weighting assumes, however, that the observations are of a point source, an assumption which is rapidly becoming inaccurate for modern observations (e.g. those of protoplanetary disks presented in this thesis). We present a more general form of traditional Briggs 'robust' weighting, which is adaptive to the source being observed and properly incorporates the source visibility contributions into the weight of any given grid cell, maximizing SNR. As with robust weighting, the weights are tunable between this matched filter 'natural' weighting and uniform weighting. In the limit of a point source, robust weighting is recovered, while the weighting scheme outperforms robust weighting for all other source structures and offers significant benefits over the conventional Gaussian tapering of weak sources. We demonstrate the weighting scheme on both synthetic and real observations and discuss the limitations of source structure estimation.

1. INTRODUCTION

Interferometric observations incompletely (and non-uniformly) sample the UV plane, due to inherent limitations of both aperture synthesis and array design. Additionally, each interferometric visibility has unique noise properties, with longer baselines often being noisier. These two properties make the problem of weighting during interferometric imaging non-trivial. When combining data with varied noise properties, data weights are often set to be inversely proportional to the variance ('natural' weighting). This maximizes signal to noise (SNR) when imaging a point source, but also results in undesirable beam properties. Conversely, when the weights of the gridded data are set such that the uv plane is effectively uniformly sampled ('uniform' weighting), the sidelobes of dirty beam are minimized and resolution is improved, while image noise properties are significantly degraded. In his PhD thesis (Briggs 1995), Dan Briggs showed that a weighting scheme which tunes between these two extremes can recover most of the benefits of each ('robust' weighting). This method is now essentially the default weighting for most radio interferometric imaging. The derivation of robust weighting assumes, however, that the observations are of a point source, an assumption which is often not true in practice.

In this paper we present a more general form of robust weighting. This new weighting scheme is adaptive to the source being observed, and properly incorporates the source visibility contributions into the weight of any given grid cell, maximizing SNR. As with robust weighting, the weights are tunable between this matched filter 'natural' weighting and uniform weighting. In the limit of a point source, robust weighting is recovered, while the weighting scheme outperforms robust weighting for all other source structures and offers significant benefits over the conventional Gaussian tapering of weak sources. We demonstrate the weighting scheme on both synthetic and real observations and discuss the limitations of source structure estimation.

[1] Corresponding author: rloomis@cfa.harvard.edu
[2] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138

2. ADAPTIVE ROBUST WEIGHTING

2.1. *Natural weighting for resolved sources*

Natural weighting is derived by maximizing the signal to noise ratio (SNR) for a point source. The SNR can be written as:

$$\mathrm{SNR} = \frac{I^D}{\Delta I^D} = \frac{\sum\limits_k W_k V_k}{(\sum\limits_k W_k^2 \sigma_k^2)^{1/2}} \tag{1}$$

where there are k visibilities in the dataset, each with a source visibility contribution V_k and total weight W_k. If the data is gridded into cells $\{p, q\}$ and the noise properties of all visibilities within each cell are similar, an approximate form can be written:

$$\mathrm{SNR} = \frac{\sum\limits_{p,q} W_{pq} V_{pq}}{(\sum\limits_{p,q} W_{pq}^2 \sigma_{pq}^2)^{1/2}} \tag{2}$$

As shown in Briggs (1995), this equation for SNR can then be maximized for a point source ($V_{pq} = 1\ \forall p, q$) to yield the condition:

$$W_{pq} \propto \frac{1}{\sigma_{pq}^2} \tag{3}$$

which is the definition of natural weighting. In a similar manner, the SNR of the central pixel of a resolved source that has been phase shifted to the center of the map can be written as:

$$\mathrm{SNR}(0,0) = \frac{\sum_{p,q} W_{pq}\mathrm{Re}(V_{pq})}{(\sum_{p,q} W_{pq}^2\sigma_{pq}^2)^{1/2}} \tag{4}$$

Following the same maximization process, Briggs (1995) showed that a condition can be derived which maximizes SNR for this general resolved source:

$$W_{pq} \propto \frac{\mathrm{Re}(V_{pq})}{\sigma_{pq}^2} \tag{5}$$

This condition is equivalent to a matched filter (North 1943), an intuitive and expected result, as the matched filter is the optimal linear filter for maximizing SNR. As Briggs (1995) points out, however, the meaning of SNR for a resolved source is not a well defined concept. Crane & Napier (1989) suggest using SNR at a point within the image (such as (0,0) selected above) as the object of merit, but this is not necessarily an obvious choice for non-centrally peaked sources (i.e. a ring). An alternative approach would be to treat the source signal term of the SNR equation as proportional to sum of the source amplitudes (i.e. the signal present in each visibility is maximized, with no attention paid to its sign or where it contributes within the dirty map). In this case, the general 'resolved natural weighting' is written as:

$$W_{pq} \propto \frac{|V_{pq}|}{\sigma_{pq}^2} \tag{6}$$

Although either of these forms of weighting yield optimal SNR in the dirty map, they both significantly alter the dirty beam, which is the Fourier transform of the weights. In the former case, the dirty beam takes on the shape of the original source (convolved with the uv sampling function). When convolved with the dirty image, the center of the source will show the maximal response (Figure 1). This is obviously not beneficial for imaging non-centrally peaked sources (i.e. a ring now shows a peak at the ring center). In contrast, the latter approach convolves the dirty image with the square root of the power spectrum of the source. Signal is enhanced, but flux is distributed across the image in a similar shape to the actual source, as the dirty beam is centrally peaked (Figure 1). The beam does gain some undesirable properties, however, such as large shelves.

2.2. *Tunable weighting*

We are therefore left wondering if these undesireable beam properties can be reduced, while still retaining the signal enhancing properties of natural weighting. Robust weighting is able to achieve these goals for a point source by cleverly connecting natural and uniform weighting, and it follows that a more general form may be derivable for an arbitrary source structure. In the derivation presented in Briggs (1995), a point source of tunable strength S is introduced: when the source is infinitely strong the dirty image approximates the dirty beam, while when the source is infinitely weak, the dirty image is simply the thermal noise field. The dirty image is then minimized through the use of Lagrangian multipliers, resulting in a tunable weighting scheme that retains benefits of both natural and uniform weighting.

Unfortunately, it is difficult to formulate a similarly clever physical interpretation of the problem when a source visibility contribution must be included. In the natural weighting limit we want the weighting scheme to know about the source contribution, while in the uniform limit there should be no knowledge of the source structure. Thus we approach the problem from a different angle: the task is one of multi-objective optimization, with the goal of minimizing beam power (and therefore sidelobes/shelves) while maximizing SNR. More specifically, the problem is Pareto optimal, in that SNR cannot be increased without simultaneously increasing beam power. We therefore adopt the common technique of parametrizing the objective function, i.e. minimizing a function resembling:

$$f(W) \propto \frac{1}{\mathrm{SNR}} + \alpha\,\|B\|^2 \tag{7}$$

When α is 0, SNR is optimized, and when $\alpha \gg \frac{1}{\mathrm{SNR}}$, the beam shape is optimized. Taking this approach, we can easily reproduce classic robust weighting with a few assumptions. First we assume a point source, such that the SNR can be written:

$$\mathrm{SNR} = \frac{\sum_{p,q} W_{pq}}{(\sum_{p,q} W_{pq}^2\sigma_{pq}^2)^{1/2}} \tag{8}$$

Next, we hold the sum of the weights constant:

$$\sum_{p,q} W_{pq} = const = W \tag{9}$$

This normalization allows us to use the same Lagrangian multiplier technique as Briggs (1995). When the sum of the weights and therefore the signal is held constant, SNR can be maximized by minimizing the square of the denominator:

$$\sum_{p,q} W_{pq}^2\sigma_{pq}^2 \tag{10}$$

Combining this with the equation for beam power:

$$\|B^2\| \propto \sum_{p,q} |W_{pq}|^2 \propto \sum_{p,q} W_{pq}^2 \tag{11}$$

the total metric to be minimized is:

$$\alpha\sum_{p,q} W_{pq}^2 + \sum_{p,q} W_{pq}^2\sigma_{pq}^2 \tag{12}$$

which simplifies to:

$$\sum_{p,q} W_{pq}^2(\alpha + \sigma_{pq}^2) \tag{13}$$

We then minimize this while holding the total weights constant:

$$\frac{\partial}{\partial W_{pq}}\Big[\sum_{p,q} W_{pq}^2(\alpha+\sigma_{pq}^2)+\lambda\Big(\sum_{p,q} W_{pq}-W\Big)\Big]=0 \quad (14)$$

$$2W_{pq}(\alpha+\sigma_{pq}^2)+\lambda=0 \quad (15)$$

$$W_{pq}=\frac{\lambda}{2(\alpha+\sigma_{pq}^2)}\propto\frac{1}{\alpha+\sigma_{pq}^2} \quad (16)$$

This is clearly identical to robust weighting, with our tunable parameter α being analogous to the point source power S^2 present in Briggs (1995).

Briggs (1995) claims that W_{pq} approaches $\frac{1}{\sigma_{pq}^2}$ as α approaches 0, and approaches a constant when $\alpha \gg \sigma_{pq}$. Thus is true, but inconveniently the constant W_{pq} approaches is 0. As normalization of the weights is arbitrary and α never approaches infinity in practical implementations of the algorithm, this is not a problem, but it does suggest an incomplete formulation of the solution. We will return to this point after deriving the general solution for an arbitrary source structure.

In the general case, holding the sum of the weights constant is not sufficient and the full signal term must be held constant:

$$\sum_{p,q} W_{pq}|V_{pq}|=const=S \quad (17)$$

With this term held constant, we can once again maximize SNR by minimizing equation 10. We still hold the total weights constant, and equation 13 becomes:

$$\frac{\partial}{\partial W_{pq}}\Big[\sum_{p,q} W_{pq}^2(\alpha+\sigma_{pq}^2)+\lambda_1\Big(\sum_{p,q} W_{pq}-W\Big) + \lambda_2\Big(\sum_{p,q} W_{pq}|V_{pq}|-S\Big)\Big]=0 \quad (18)$$

$$2W_{pq}(\alpha+\sigma_{pq}^2)+\lambda_1+\lambda_2|V_{pq}|=0 \quad (19)$$

$$W_{pq}=\frac{\lambda_1+\lambda_2|V_{pq}|}{-2(\alpha+\sigma_{pq}^2)} \quad (20)$$

λ_1 and λ_2 are arbitrary constants, so we choose λ_2 to be -2, and substitute β for $\frac{-\lambda_1}{2}$

$$W_{pq}=\frac{\beta+|V_{pq}|}{\alpha+\sigma_{pq}^2} \quad (21)$$

This form reduces to classic robust weighting with an assumption of a point source ($|V_{pq}|=1$) and choice of β as 0. In this more general form, the weights become $\frac{\beta+|V_{pq}|}{\alpha}$ in the limit $\alpha \gg \sigma_{pq}$, rather than approaching 0 like classic robust weighting. To optimize our choice of β, we repeat our maximization of SNR, this time taking the partial derivative with respect to β.

$$\frac{\partial}{\partial \beta}\Big[\sum_{p,q} W_{pq}^2\sigma_{pq}^2+\lambda\Big(\sum_{p,q} W_{pq}|V_{pq}|-S\Big)\Big]=0 \quad (22)$$

Substituting in equation 20:

$$\frac{\partial}{\partial \beta}\Big[\sum_{p,q}\frac{(\beta+|V_{pq}|)^2\sigma_{pq}^2}{(\alpha+\sigma_{pq}^2)^2} - \lambda\Big(\sum_{p,q}\frac{(\beta+|V_{pq}|)|V_{pq}|}{\alpha+\sigma_{pq}^2}-S\Big)\Big]=0 \quad (23)$$

$$\sum_{p,q}\frac{2(\beta+|V_{pq}|)\sigma_{pq}^2}{(\alpha+\sigma_{pq}^2)^2}-\lambda\sum_{p,q}\frac{|V_{pq}|}{\alpha+\sigma_{pq}^2}=0 \quad (24)$$

We choose $\lambda=2$ and solve for β:

$$\sum_{p,q}\frac{(\beta+|V_{pq}|)\sigma_{pq}^2}{(\alpha+\sigma_{pq}^2)^2}=\sum_{p,q}\frac{|V_{pq}|}{\alpha+\sigma_{pq}^2} \quad (25)$$

$$\sum_{p,q}(\beta+|V_{pq}|)\sigma_{pq}^2=\sum_{p,q}|V_{pq}|(\alpha+\sigma_{pq}^2) \quad (26)$$

$$\sum_{p,q}\beta\sigma_{pq}^2+|V_{pq}|\sigma_{pq}^2=\sum_{p,q}\alpha|V_{pq}|+|V_{pq}|\sigma_{pq}^2 \quad (27)$$

$$\sum_{p,q}\beta\sigma_{pq}^2=\sum_{p,q}\alpha|V_{pq}| \quad (28)$$

$$\beta=\sum_{p,q}\frac{\alpha|V_{pq}|}{\sigma_{pq}^2} \quad (29)$$

$$\beta=\frac{\alpha\overline{|V_{pq}|}}{\overline{\sigma_{pq}^2}} \quad (30)$$

As in Briggs (1995), we choose a convenient function for α, with useful values of r ranging from -2 to 2.

$$\alpha=(5\times10^{-r})^2\,\overline{\sigma_{pq}^2} \quad (31)$$

Thus the total weights become:

$$W_{pq}=\frac{(5\times10^{-r})^2\,\overline{|V_{pq}|}+|V_{pq}|}{(5\times10^{-r})^2\,\overline{\sigma_{pq}^2}+\sigma_{pq}^2} \quad (32)$$

This functional form intuitively aligns the incorporation of information about the source structure and information about the uv sampling function, such that their inflection points are matched.