Max–Planck–Institut für biologische Kybernetik
Max Planck Institute for Biological Cybernetics

Technical Report No. 186

Spectral Stacking: Unbiased Shear Estimation for Weak Gravitational Lensing

Reshad Hosseini,[1] Matthias Bethge,[1]

Oktober 2009

This Technical Report has been approved by:

Director at MPIK

Postdoc at MPIK

Max–Planck–Institut für biologische Kybernetik
Max Planck Institute for Biological Cybernetics

Technical Report No. 186

Spectral Stacking: Unbiased Shear Estimation for Weak Gravitational Lensing

Reshad Hosseini,[1] Matthias Bethge,[1]

Oktober 2009

[1] Computational Vision and Neuroscience Group, email: hosseini;mbethge@tuebingen.mpg.de

Spectral Stacking: Unbiased Shear Estimation for Weak Gravitational Lensing

Reshad Hosseini, Matthias Bethge

Abstract. We present a new method for the estimation of shear in gravitational lensing from a set of galaxy images with unknown distribution of shapes. Common procedures first compute an estimate of some characteristic feature for each individual galaxy and then average over these. The average can be used to estimate the shear as it becomes independent of the individual galaxy shapes with increasing number of images. A common problem of the previous methods is that the estimators of the features are biased. Here we introduce "*spectral stacking*" which uses the power spectrum as a characteristic feature of the individual galaxies. If the galaxy images are contaminated by Poisson noise, an unbiased estimator of the power spectrum exists which is used in the analysis. Furthermore, the power spectrum is independent of the location of the individual galaxy centers provided the smoothed galaxy intensities decay sufficiently fast. No further assumptions are necessary. The algorithm won the main contest of the Great08 challenge.

1 Introduction

Due to dark matter in the universe, the light of galaxies is bended and as a result we can only image the distorted versions of the original galaxies. In the case of weak gravitational lensing the distortion of the galaxy image can be described by a shear transformation:

$$\begin{pmatrix} \mathrm{x_u} \\ \mathrm{y_u} \end{pmatrix} = \underbrace{\begin{pmatrix} 1-a & -b \\ -b & 1+a \end{pmatrix}}_{A} \begin{pmatrix} \mathrm{x}_s \\ \mathrm{y}_s \end{pmatrix}, \tag{1}$$

where $\mathrm{x_u}$ and $\mathrm{y_u}$ are the coordinates of the undistorted galaxy image and x_s and y_s are those of the sheared image. The variables a and b determine the shear transformation matrix.

The generation of the image data set provided in the Great08 challenge can be described as follows: Images of 10^4 galaxies of unknown shapes are sheared using a unique shear parameter pair (a, b). Subsequently, the images are blurred with a smoothing kernel and then pixelized. The resulting pixel intensities are distorted by Poisson noise. This is accomplished by drawing from a Poisson distribution whose mean is given by the pixel intensity plus some constant background.

The shear parameter estimation problem in the challenge is a simplified version of the real problem because here the shear is assumed to be the same for all galaxies while in reality the shear is not constant even if we only consider galaxies in some part of the sky that are located in close vicinity to each other. Also the smoothing kernel is assumed to be constant across the image and the same for all galaxies while in reality it is not. The noise model considered in the challenge is Poisson distributed but the true noise is better described as a combination of Poisson and Gaussian noise [1].

One assumption that is fundamental to solving the shear estimation problem is *isotropy* which says that all rotations of the same galaxy around the center of mass should be equally likely. Most procedures exploit this assumption in the following way: they first compute an estimate of some characteristic feature for each individual galaxy and then average over these. By virtue of the isotropy, the average becomes independent of the individual orientations of the galaxies with increasing number of images. This allows one to estimate the shear from the orientation of the average image.

An important factor for the success of this approach is to have an unbiased estimator of the characteristic feature under the given noise model. All previous methods, however, require additional assumptions on the shape of galaxies in order to guarantee that the estimates of the feature are unbiased. The method that worked second best in the challenge [2] requires only two parameters to be estimated which rely on prior knowledge about the intrinsic galaxy model: The x- and y-coordinate of the center of mass of the galaxy before smoothing.

In this work we introduce *spectral stacking* as a method that under mild assumptions, is completely independent of the shape of galaxies by using the power spectrum as a feature that can be estimated in an unbiased way. In the next section, we explain the details of the method. In section 3, we show simulation results and in section 4 we discuss possibilities how to further improve the method proposed in this report.

2 Method

We begin with a summary of the image generation process. The original unsheared galaxy image is $i_u(\mathrm{x_u}, \mathrm{y_u})$ and the sheared image is $i_s(\mathrm{x}_s, \mathrm{y}_s)$ where the shear is modeled by the linear coordinate transformation of (1). Afterwards, the image is blurred. This yields

$$i_f(\mathrm{x}_s, \mathrm{y}_s) = i_s(\mathrm{x}_s, \mathrm{y}_s) * h_s(\mathrm{x}_s, \mathrm{y}_s), \tag{2}$$

where i_f is the filtered image, $h_s(\mathrm{x}_s, \mathrm{y}_s)$ is the filter kernel, and $*$ denotes the convolution operator. Then the image is pixelized which is equivalent to a convolution with a box kernel $h_f(\mathrm{x}_s, \mathrm{y}_s)$ and subsequent sampling. Furthermore, the range of the image is limited which is equivalent to a multiplication with a box window function $w(\mathrm{x}_s, \mathrm{y}_s)$. Thus, we obtain

$$i_r(\mathrm{x}_s, \mathrm{y}_s) = [i_f(\mathrm{x}_s, \mathrm{y}_s) * h_f(\mathrm{x}_s, \mathrm{y}_s)]w(\mathrm{x}_s, \mathrm{y}_s) \tag{3}$$

$$i_r[\mathrm{m}_s, \mathrm{n}_s] = i_r(\mathrm{m}_s T, \mathrm{n}_s T) \;\; \forall\{\mathrm{m}_s, \mathrm{n}_s\} \in \mathbb{Z}, \tag{4}$$

where T is the sampling period. Without loss of generality we can set $T = 1$. Furthermore, we can concatenate the two linear filtering steps into one $h(\mathrm{x}_s, \mathrm{y}_s) = h_s(\mathrm{x}_s, \mathrm{y}_s) * h_f(\mathrm{x}_s, \mathrm{y}_s)$. This yields

$$i_r(\mathrm{x}_s, \mathrm{y}_s) = [i_s(\mathrm{x}_s, \mathrm{y}_s) * h(\mathrm{x}_s, \mathrm{y}_s)]w(\mathrm{x}_s, \mathrm{y}_s) \tag{5}$$

$$i_r[\mathrm{m}_s, \mathrm{n}_s] = i_r(\mathrm{m}_s, \mathrm{n}_s) \;\; \forall\{\mathrm{m}_s, \mathrm{n}_s\} \in \mathbb{Z}. \tag{6}$$

Taking the Fourier transform of the two equations yields:

$$I_r(\Omega_x, \Omega_y) = [I_s(\Omega_x, \Omega_y)H(\Omega_x, \Omega_y)] * W(\Omega_x, \Omega_y) \tag{7}$$

$$\tilde{I}_r(\omega_x, \omega_y) = \sum_{k_x, k_y = -\infty}^{\infty} I_r(\omega_x + 2k_x\pi, \omega_y + 2k_y\pi), \;\; -\pi \leq \{\omega_x, \omega_y\} \leq \pi, \tag{8}$$

where capital letters denote Fourier transform, Ω is the frequency of the continuous signal and ω is the frequency of the discrete signal which ranges from $-\pi$ to π. For our analysis we assume that the filtered image is confined to the domain of the rectangular function and zero outside. Then we can drop the convolution with the sinc function $W(\Omega_x, \Omega_y)$:

$$I_r(\Omega_x, \Omega_y) \approx I_s(\Omega_x, \Omega_y)H(\Omega_x, \Omega_y). \tag{9}$$

The range of $I_r(\Omega_x, \Omega_y)$ is less than 2π (because of the filtering with the smoothing kernel). Therefore, we have

$$\tilde{I}_r(\omega_x, \omega_y) = I_r(\omega_x, \omega_y), \;\; -\pi \leq \{\omega_x, \omega_y\} \leq \pi. \tag{10}$$

The amplitude of the Fourier transform is invariant with respect to shifts. That is, both a signal and a shifted version of it have the same amplitude spectrum. More generally, we can write

$$|\tilde{I}_r(\omega_x, \omega_y)|^p \approx |I_s(\omega_x, \omega_y)|^p |H(\omega_x, \omega_y)|^p, \tag{11}$$

where p is an arbitray power. Since the convolution kernel $H(\omega_x, \omega_y)$ is assumed to be the same for all galaxies, taking the sum over all galaxy images yields

$$\sum_{k=1}^{N} |\tilde{I}_{r_k}(\omega_x, \omega_y)|^p \approx |H(\omega_x, \omega_y)|^p \sum_{k=1}^{N} |I_{s_k}(\omega_x, \omega_y)|^p, \tag{12}$$

where N is the total number of galaxy images. If N becomes sufficiently large, the second term in (12) converges $\sum_{k=1}^{N} |I_{s_k}(\omega_x, \omega_y)|^p \rightarrow G(\omega_x, \omega_y)$ and the limiting function has an elliptical shape being of the form

$$G(\omega_x, \omega_y) = f([\omega_x, \omega_y]\Sigma^{-1}[\omega_x, \omega_y]^\top), \tag{13}$$

where the covariance matrix Σ is dependent on the shear matrix and $f : [0, \infty) \rightarrow \mathbb{R}$ is an arbitrary function. The elliptical shape of the limiting function $G(\omega_x, \omega_y)$ originates from the assumed isotropy. This can be explained as follows: Due to the isotropy assumption the sum over a large number of images converges to a spherically symmetric function. Since shearing is a linear transformation first shearing each individual image and then averaging yields the same result as first averaging and then shearing. Thus the effect of the shear can be described to transform the spherically symmetric function into an elliptically contoured function. This property also holds true for the magnitude of the Fourier transform because for the magnitude taking the Fourier transform of a rotated image is the same as rotating the Fourier transformed image. The relationship between the coordinate systems in the frequency domain between sheared and original galaxies is given by:

$$\begin{pmatrix} \omega_x \\ \omega_y \end{pmatrix} = \underbrace{\begin{pmatrix} 1-a & -b \\ -b & 1+a \end{pmatrix}^{-1}}_{M=A^{-1}} \begin{pmatrix} \omega'_x \\ \omega'_y \end{pmatrix}. \tag{14}$$

The covariance matrix of a spherically symmetric function is proportional to the identity matrix:

$$\int \begin{pmatrix} \omega'_x \\ \omega'_y \end{pmatrix} \begin{pmatrix} \omega'_x & \omega'_y \end{pmatrix} G(\omega'_x, \omega'_y)\, d\omega'_x\, d\omega'_y = k \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{15}$$

Hence, applying the shear transformation yields:

$$\begin{aligned} \Sigma &= \int \begin{pmatrix} \omega_x \\ \omega_y \end{pmatrix} \begin{pmatrix} \omega_x & \omega_y \end{pmatrix} G(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \\ &= M(1-a^2-b^2) \left[\int \begin{pmatrix} \omega'_x \\ \omega'_y \end{pmatrix} \begin{pmatrix} \omega'_x & \omega'_y \end{pmatrix} G(\omega'_x, \omega'_y)\, d\omega'_x\, d\omega'_y \right] M^\top \\ &= k(1-a^2-b^2)\, MM^\top. \end{aligned} \tag{16}$$

Using the fact that the shear matrix is symmetric it follows from (16) that $A = \lambda\, \Sigma^{-1/2}$ where λ is a constant. In order to fulfill the constraint that the trace of the shear matrix A must be equal to $(1+a)+(1-a) = 2$ we need to set $\lambda = \frac{2}{\text{trace}(\Sigma^{-1/2})}$. Now, we can write $G(\omega_x, \omega_y)$ as a function of the shear matrix A:

$$G_A(\omega_x, \omega_y) = f(\lambda^{-2}[\omega_x, \omega_y]AA^\top[\omega_x, \omega_y]^\top). \tag{17}$$

Finally, we have to take into account that the images $i_r[\mathrm{m}_s, \mathrm{n}_s]$ are contaminated by Poisson noise. That is, the intensities of the observed pixels $i_o[\mathrm{m}_s, \mathrm{n}_s]$ are drawn from Poisson distributions whose means are given by $i_r[\mathrm{m}_s, \mathrm{n}_s]$. If we choose $p = 2$ we can use

$$S(\omega_x, \omega_y) := |\tilde{I}_o(\omega_x, \omega_y)|^2 - C \tag{18}$$

as an unbiased estimator of the power spectrum under Poisson noise [3]:

$$E\{S(\omega_x, \omega_y)\} = |\tilde{I}_r(\omega_x, \omega_y)|^2 \tag{19}$$

Here, $\tilde{I}_o(\omega_x, \omega_y)$ is the Fourier transform of the noisy image $i_o[\mathrm{m}_s, \mathrm{n}_s]$ and $C = \sum_{\mathrm{m}_s, \mathrm{n}_s} i_o[\mathrm{m}_s, \mathrm{n}_s]$. If we sum over this unbiased estimate of the power spectra of the galaxies (see (12)), we obtain an unbiased estimate of the filtered limiting function $|H(\omega_x, \omega_y)|^2 G(\omega_x, \omega_y)$. Due to the central limit theorem, the resulting noise is approximately Gaussian. Fortunately, an unbiased estimator for the variance of the resulting Gaussian noise can also be obtained (see appendix). Taken together, we obtain the following model:

$$\sum_{k=1}^{N} S_k(\omega_x, \omega_y) = |H(\omega_x, \omega_y)|^2 G_A(\omega_x, \omega_y) + \eta(\omega_x, \omega_y), \tag{20}$$

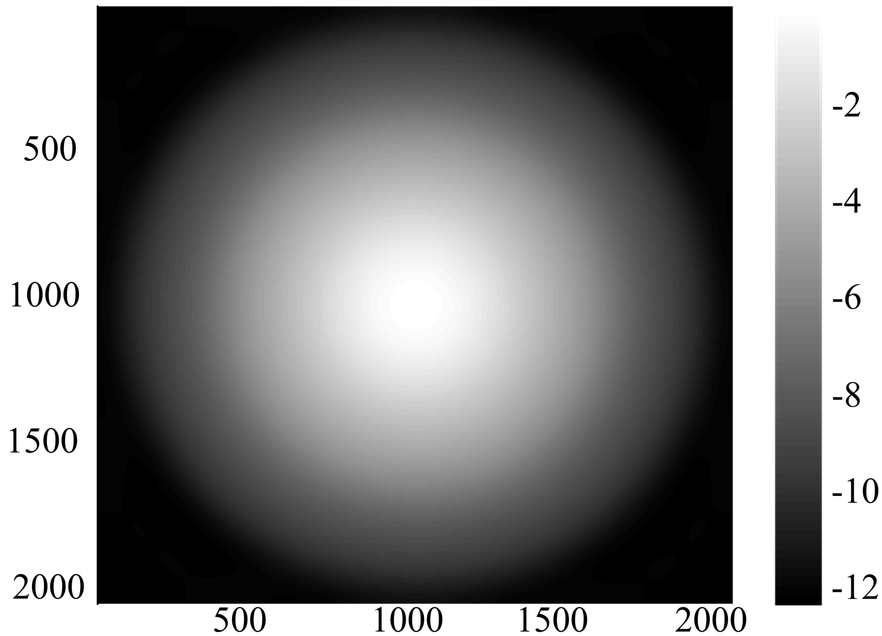


Figure 1: Logarithm of the power spectrum of the smoothing filter. It is estimated by summing over the power spectrum of the stars which is given as a separate set of 10^4 stars.

where $G_A(\omega_x, \omega_y)$ is an elliptically contoured function given in (17) which depends on the shear matrix A and where $\eta(\omega_x, \omega_y)$ denotes Gaussian noise whose an unbiased estimator of its variance is given by (A-4) in the appendix. Hence, for estimating the shear matrix A we can use a nonlinear least square procedure to optimize over the limiting function using the inverse square root of the the noise variance as a weighting function. In other words, we want to minimize the following loss function over the range of possible shear matrices A:

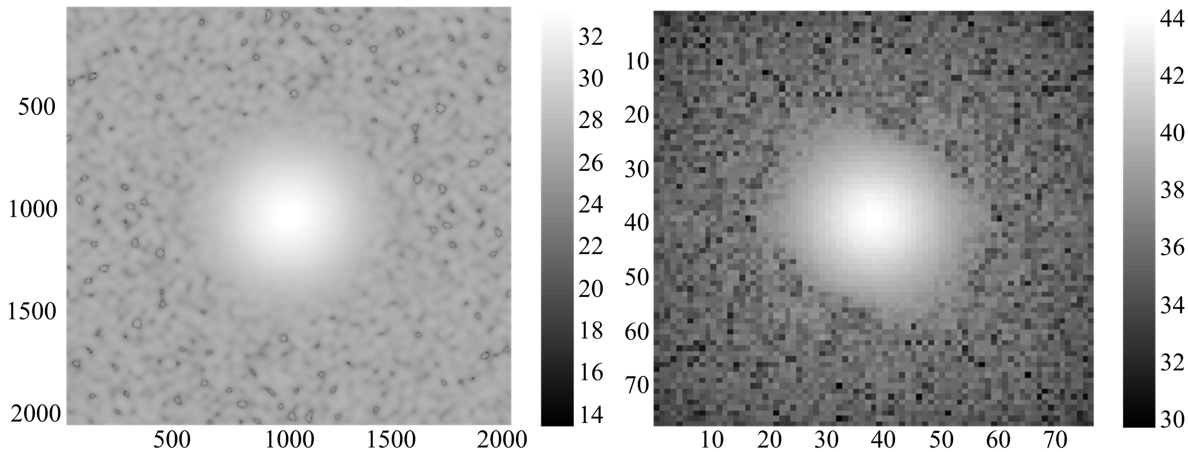
$$E_A = \sum_{\omega_x, \omega_y} \left[\frac{1}{\sigma(\omega_x, \omega_y)} \left(\sum_{k=1}^{N} S_k(\omega_x, \omega_y) - |H(\omega_x, \omega_y)|^2 G_A(\omega_x, \omega_y) \right) \right]^2 . \tag{21}$$

In order to carry out the minimization, we further need to specify a parametric model for the radial function f that we used in the definition of $G_A(\omega_x, \omega_y)$. In our case, we used a polynomial to model the radial function in the log domain.

3 Results

We applied the algorithm to the main contest of the Great08 challenge. In the challenge a data set of smoothed star images was provided in addition to the galaxy images. We used this data set to estimate the term $|H(\omega_x, \omega_y)|^2$ by first computing the power spectrum for each individual star image and then averaging over the entire ensemble. All galaxy images have been smoothed with one out of three different smoothing filters. Our estimate of one of the smoothing filters is shown in Figure 1. The magnitude decays very fast such that the pixelization does not lead to aliasing in the Fourier domain.

In order to solve the nonlinear least square problem of (21) we used the Levenberg-Marquardt method [4] implemented in the MATLAB optimization toolbox. We sample the Fourier transform over a rectangular grid of 2048×2048 points (This is equivalnet to zero-pading in the pixel domain and then applying Fourier transform). For the nonlinear least square fit, we only used a fraction of the sample points in the fourier domain: We only kept those points for which the power spectrum of the filter is greater than 1.5 percent of its maximum value. As a radial function we used



(a)

(b)

Figure 2: (a) The sum of the unbiased estimates of the galaxy spectra in the frequency domain. (b) The inverse Fourier transform of (a) in the desired range. Both figures show the logarithm of the absolute value.

a polynomial of degree 9 in the log domain:

$$f(r) = \sum_{k=0}^{9} a_k \log(r)^k \tag{22}$$

In Fig. 2(a), we have shown the result of the estimate of the summed power spectra for a set of galaxies chosen from the main challenge ("real noise"). In Fig. 2(b), we applied the inverse Fourier transform to get the pixelized representation of the signal and display only the meaningful part of 78x78 pixels. (The original support of the image is 39x39 and squaring in the Fourier domain is equivalent to convolution in the pixel domain). As one can see, the image intensity decays rapidly in the pixel domain. This finding justifies our assumption that a signal is present only in the range of the box.

The accuracy of the results for the challenge has been evaluated by the following score:

$$Q = \frac{10^{-4}}{\frac{1}{2}\left(\left\langle\left(<\hat{a}_i - a_i>_{i\in k}\right)^2\right\rangle_k + \left\langle\left(<\hat{b}_i - b_i>_{i\in k}\right)^2\right\rangle_k\right)}, \tag{23}$$

where a_i and b_i are the true shear matrix parameters for a galaxy set i and $\hat{a}_i$ and $\hat{b}_i$ are the estimates, respectively. The inner brackets denote averaging over sets with similar shear value and observing condition $j \in k$ and the outer brackets denote averaging over different observing condition. The value is calculated from 2700 different sets of galaxy images. The method introduced in this paper achieved the best score of $Q = 210.9$. The second best method based on stacking in the pixel domain [2] achieved a score of $Q = 131.4$.

4 Discussion and Future Work

In this paper, we introduced "spectral stacking" as a new method for the estimation of shear from a set of distorted galaxies with unknown shape distribution. Unlike existing methods, this procedure is completely independent of the intrinsic galaxy shapes. The characteristic feature that we use is an unbiased estimator of the power spectrum for each galaxy. The resulting method achieved the best score of $Q = 210.9$ in the challenge.

Various modifications and extensions are possible that may lead to further improvements of spectral stacking. First, We would like to mention that it can easily be modified to incorporate both Poisson and Gaussian noise. If the smoothing kernel is changing from one galaxy to another, one could use an averaged smoothing kernel in the analysis.

Second, one could get rid of the *ad hoc* assumption of using only those pixels for the nonlinear least square minimization which are greater than a certain threshold. To this end, one can apply the inverse Fourier transform to the power spectrum and reformulate the problem in that domain because there we know the range to which the nonlinear least squares method has to be applied (see Fig. 2). After applying the inverse Fourier transform, the estimation problem of (20) reads

$$f(x,y) = \tilde{h}(x,y) * g(x,y) + \eta(x,y), \tag{24}$$

where $f(x,y)$ and $\tilde{h}(x,y)$ are the inverse Fourier transform of $\sum_{k=1}^{N} S_k(\omega_x,\omega_y)$ and $|H(\omega_x,\omega_y)|^2$, respectively. Note that $g(x,y)$, the inverse Fourier transform of $G(\omega_x,\omega_y)$, is again an elliptically contoured function and one can use the same parametric model as in (22). Because the inverse Fourier transform is a linear operator the noise is again Gaussian. Therefore, the same nonlinear least square method can be used to search for the optimal shear parameters.

Another possible modification would be to use the square of the power spectrum (i.e. to choose p=4) as characteristic feature of the galaxy images. For this feature an unbiased estimate is given by:

$$\tilde{S}(\omega_x,\omega_y) = \frac{1}{2}S(\omega_x,\omega_y)^2 - \frac{1}{2}C^2 - CS(\omega_x,\omega_y) - \frac{1}{2}S(2\omega_x,2\omega_y). \tag{25}$$

Using 19 and A-1, it is straightforward to see that $\tilde{S}(\omega_x,\omega_y)$ is an unbiased estimator of the square of power spectrum $|\tilde{I}_r(\omega_x,\omega_y)|^4$:

$$E\{\tilde{S}(\omega_x,\omega_y)\} = \frac{1}{2}E\{S(\omega_x,\omega_y)^2\} - \frac{1}{2}C^2 - CE\{S(\omega_x,\omega_y)\} - \frac{1}{2}E\{S(2\omega_x,2\omega_y)\} = |\tilde{I}_r(\omega_x,\omega_y)|^4. \tag{26}$$

It is an empirical question whether this feature would lead to better results.

Finally, one could replace the objective function of the nonlinear least squares method with a more principled approach: Instead of minimizing (21) we are actually interested in minimizing the mean square error of the estimate of the shear parameters which is reflected in Q (see (23)). To this end, one could use Bayesian sampling methods instead of the nonlinear least square method in order to estimate the shear parameters under the true loss function.

Acknowledgements

We would like to thank Michael Hirsch for helpful comments on an earlier version of the manuscript.

References

[1] Sarah Bridle et al. Handbook of the great08 challenge: An image analysis competition for cosmological lensing. http://zuuserver2.star.ucl.ac.uk/ sarah/GREAT08_Challenge_Document.pdf.

[2] Antony Lewis. Galaxy shear estimation from stacked images. http://cosmologist.info/notes/ShapeEst.pdf, 2009.

[3] J. C. Dainty and A. H. Greenaway. Estimation of spatial power spectra in speckle interferometry. *Journal of the Optical Society of America*, 69(5):786–790, May 1979.

[4] Donald W. Marquardt. An algorithm for Least-Squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, June 1963.

Appendix

We derive an unbiased estimator of the noise variance of $\eta(\omega_x, \omega_y)$ for the model given in (20). In [3] it has been shown that the second moment of the estimator in (18) yields

$$E\{S(\omega_x, \omega_y)^2\} = 2C|\tilde{I}_r(\omega_x, \omega_y)|^2 + C^2 + 2|\tilde{I}_r(\omega_x, \omega_y)|^4 + |\tilde{I}_r(2\omega_x, 2\omega_y)|^2. \tag{A-1}$$

Therefore, the variance of the estimator for the kth galaxy image reads:

$$\begin{aligned} \sigma_k(\omega_x, \omega_y)^2 &= E\{S_k(\omega_x, \omega_y)^2\} - E\{S_k(\omega_x, \omega_y)\}^2 \\ &= 2C_k|\tilde{I}_{r_k}(\omega_x, \omega_y)|^2 + C_k^2 + |\tilde{I}_{r_k}(\omega_x, \omega_y)|^4 + |\tilde{I}_{r_k}(2\omega_x, 2\omega_y)|^2. \end{aligned} \tag{A-2}$$

The total noise variance is the sum of the individual variances:

$$\sigma(\omega_x, \omega_y)^2 = \sum_{k=1}^{N} \sigma_k(\omega_x, \omega_y)^2 = \sum_{k=1}^{N} 2C_k|\tilde{I}_{r_k}(\omega_x, \omega_y)|^2 + C_k^2 + |\tilde{I}_{r_k}(\omega_x, \omega_y)|^4 + |\tilde{I}_{r_k}(2\omega_x, 2\omega_y)|^2. \tag{A-3}$$

From this we can derive that

$$V(\omega_x, \omega_y) = 1/2 \sum_{k=1}^{N} S_k(\omega_x, \omega_y)^2 + 1/2 \sum_{k=1}^{N} C_k^2 + \sum_{k=1}^{N} C_k S_k(\omega_x, \omega_y) + 1/2 \sum_{k=1}^{N} S_k(2\omega_x, 2\omega_y) \tag{A-4}$$

is an unbiased estimator of the noise variance since by taking the expectation of $V(\omega_x, \omega_y)$ we find

$$\begin{aligned} E\{V(\omega_x, \omega_y)\} &= 1/2 \sum_{k=1}^{N} E\{S_k(\omega_x, \omega_y)^2\} + 1/2 \sum_{k=1}^{N} C_k^2 + \sum_{k=1}^{N} C_k E\{S_k(\omega_x, \omega_y)\} + 1/2 \sum_{k=1}^{N} E\{S_k(2\omega_x, 2\omega_y)\} \\ &= \sum_{k=1}^{N} 2C_k|\tilde{I}_{r_k}(\omega_x, \omega_y)|^2 + C_k^2 + |\tilde{I}_{r_k}(\omega_x, \omega_y)|^4 + |\tilde{I}_{r_k}(2\omega_x, 2\omega_y)|^2 = \sigma(\omega_x, \omega_y)^2. \end{aligned} \tag{A-5}$$

Note that here we do not investigate the convergence rate of the estimator because this rate depends on the distribution shape of the galaxies which is not known.