Remote Sensing for Earth Observation and Surveillance
Homework 1

Adrian Burton Parisi
10710691
940789
adrianburton.parisi@mail.polimi.it

1 Problem 1: Passive Radio-astronomy

1. Assume for the moment that there is just a single source whose power spectrum is constant over the whole available 10Mhz bandwidth. Estimate the equivalent brightness temperature of this source and assess the theoretical accuracy of the estimation.

The assumption that the source is at a constant power across the entire bandwidth effectively creates a rectangular pulse at a given power for a 10Mhz bandwidth centered at 0Hz within the frequency domain. There are several methods in which the power can be calculated, the method used throughout this report is

$$P_x = \frac{1}{N}\sum^{N} x_n(t)x_n^*(t) \tag{1}$$

where $x_n(t)$ is complex envelope of the signal in the time domain. Inputting one of the signals provided into Eq. 1 results in a power of 9.6439×10^{-16}W

The power can also be defined by

$$P_n = k_B B_n T_n \tag{2}$$

where P_n, B_n, T_n, and k_B are the power, bandwidth, and brightness temperature for a given source, n, and the Boltzmann's constant, respectively, results in a brightness temperature of (6.985±0.022)K.

The error of the power was calculated via

$$\sigma_P = \sqrt{\frac{P^2}{BT_{obs}}} \tag{3}$$

where P, B, and T_{obs} are the power, bandwidth, and observation time, respectively. Basic error propagation was conducted, using the error of the power, in order to calculate the corresponding error in the brightness temperature. The observation time was computed from using the provided sampling frequency of 10MHz and the number of samples within the data set leading to a value of 0.01s.

2. Estimate the power spectrum of the signal at one of the antennas and comment on whether the assumption at point 1 is justified. How many distinct sources can you detect?

The power spectra can be calculated by taking the a Fast Fourier Transform (FFT) of one of the signals using the function `dft`. Within the FFT, `dft` also calculates and compensates for delay. Execution time for the FFT is dependent on the length of the transform and the most rapid calculation occurs when the length is a power of 2. Therefore, in order to calculate the next viable power of 2 with the given data set the follow equation was used,

$$N_f = 2^{\text{ceil}(log_2(N_t))} \tag{4}$$

where N_t is the length of the data set. The power spectra, S_x, was calculated with the output of the `dft` using

$$S_x = \frac{1}{T_{obs}}|X(f)|^2 \tag{5}$$

where T_{obs} is the observation time and $X(f)$ is the amplitude as a function of the frequency. The result has been shown in Figure 1.

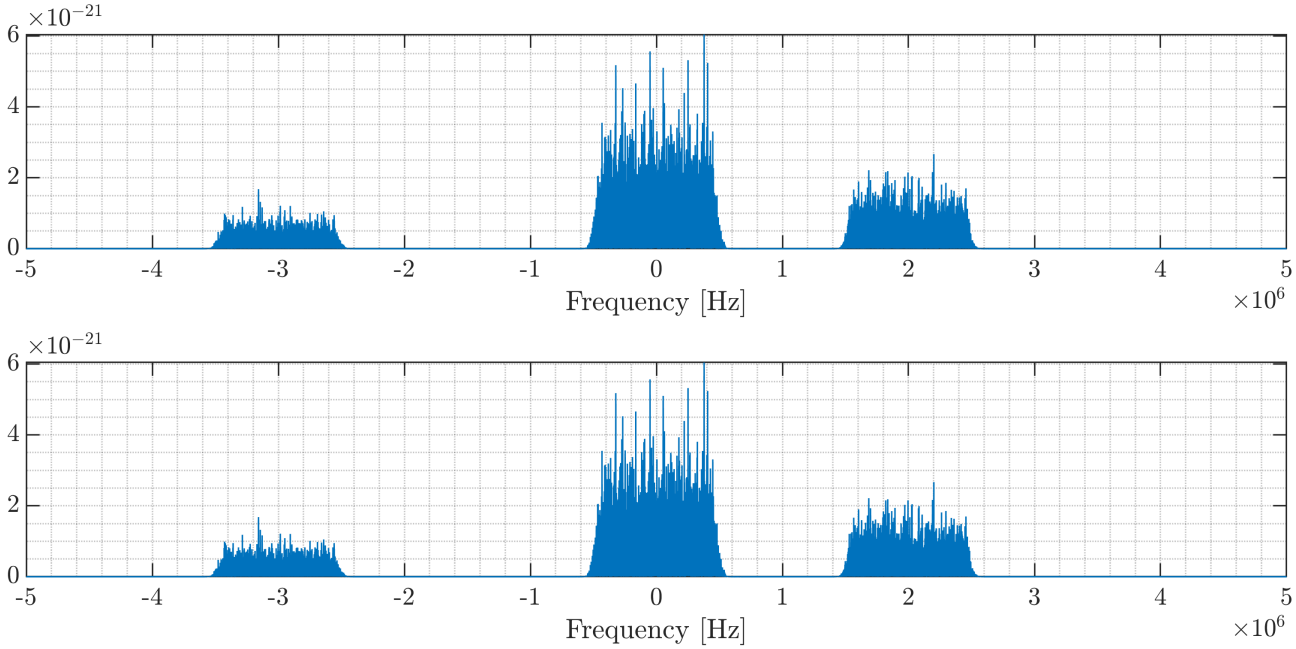

Figure 1: *The calculated power spectra of the two available signals. The upper spectra corresponds to the first signal within the provided data set and the lower is the second.*

Figure 1 suggests that there are three sources centred around 0Mhz, -3MHz, and 2MHz with the 0MHz physically representing 1420MHz. Furthermore, it can be noted that each of the sources have varying amplitudes and, albeit approximately similar bandwidth, collectively do not span across the 10MHz assumed in point 1. Therefore, the initial assumption of a single source with constant power across the entire bandwidth is clearly inadequate.

Moreover, the frequency shift corresponds to a Doppler shift on account of the emitter moving relatively towards or away from the observer. The left source, as the frequency decreases, indicates the source is moving away from the observer, and vice versa for the right source. The relative speeds required for these sources to produce their corresponding Doppler shifts suggest that they are within the Milky Way as their velocities are within the range of 400-600km/s.

3. Quantify bandwidth and frequency-shift of each detected source.

The bandwidth and frequency shifts were found utilising built-in matlab functions `obw` and `medfreq` applied to the previously computed power spectra. The plot output of these functions can be seen in Figure 2 and the tabulated results in Table 1. All of these shifted are relative to the 0MHz central point. The bandwidth, as calculated with `obw`, is the 99% of the occupied bandwidth of the spectrum between two limiting frequencies which were ±1MHz of the median frequency of the proposed bandwidth.

Source	Freq. Shift [MHz]	Bandwidth [MHz]
Left	3.0039360	1.0047150
Central	0.0088962	1.0018062
Right	1.9981876	1.0048913

Table 1: *Results of the analysis showing in Figure 2.*

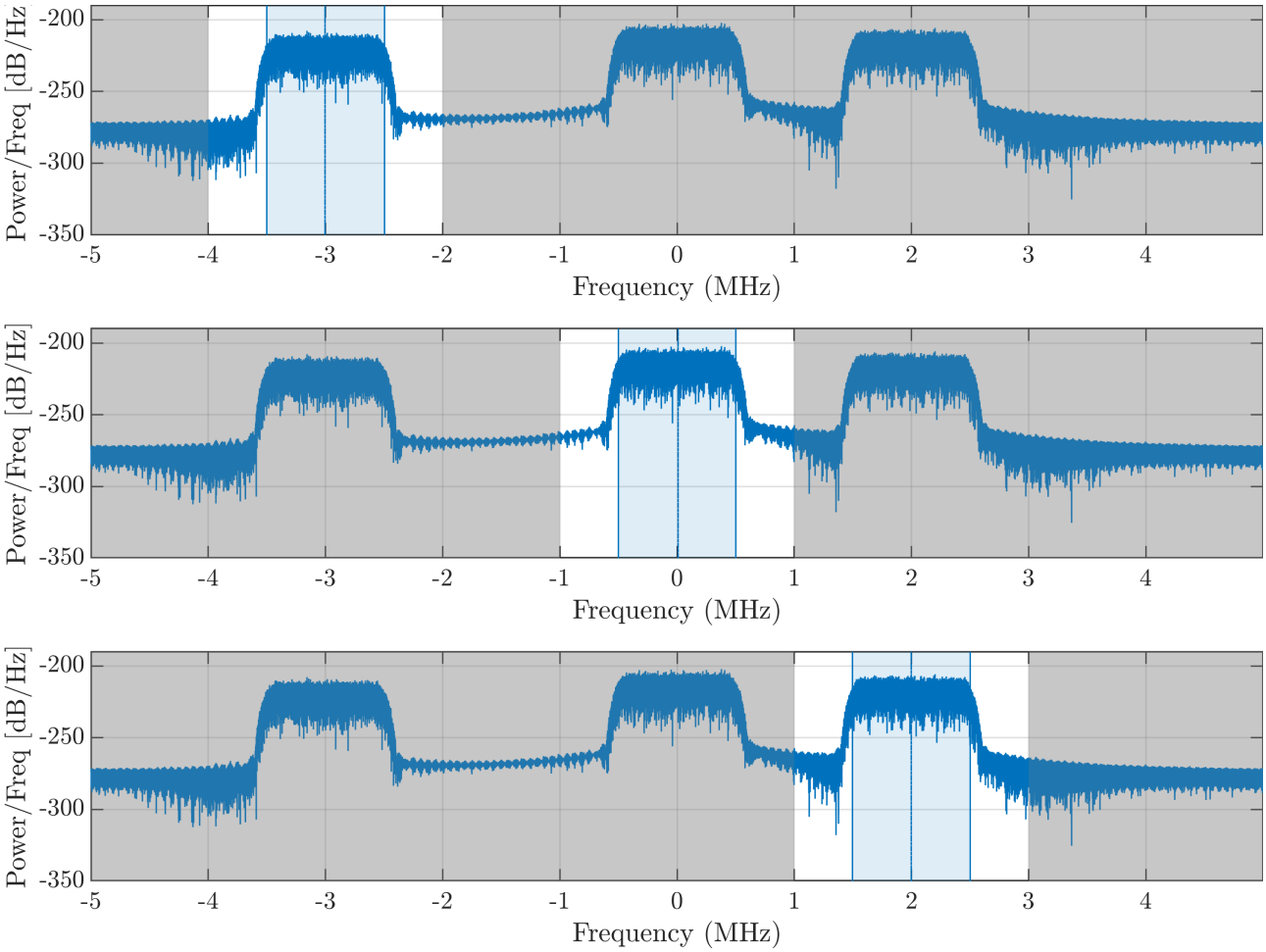

Figure 2: *Individual analysis of each of the sources found in the power spectra using the in-built matlab functions* ***obw*** *and* ***medfreq****.*

4. Propose and implement an algorithm for the extraction of the signal associated with each detected source, and validate it by estimating its power spectrum.

In order to isolate an individual source within the power spectra a filter is required. The filter was created using a cardinal sine within the time domain with a bandwidth of 1.5MHz and time axis of 500 points between the limits $[-2.5\ \ 2.5] \times 10^{-5}$s. The filter bandwidth was chosen to reflect the previously found bandwidths for each source is approximately $\approx$ 1MHz.

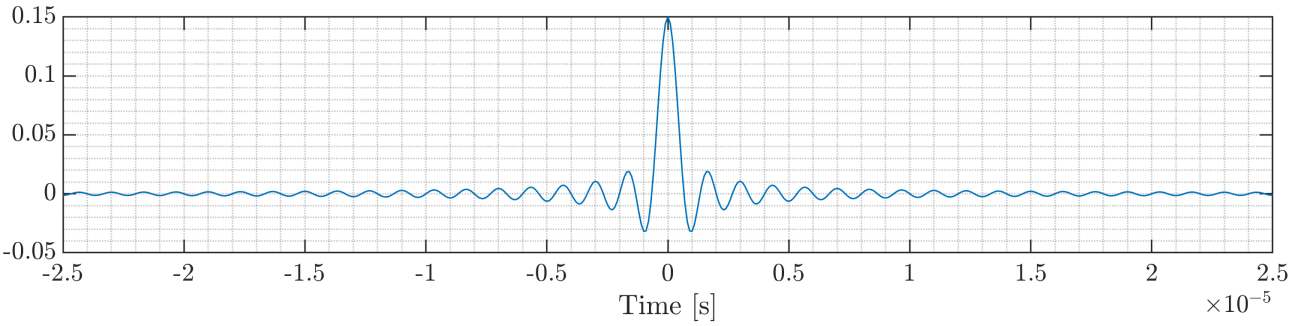

Figure 3: *The cardinal sine of bandwidth 1.5MHz used for the isolation of the individual sources within the signal.*

However, this filter will solely isolate the central source therefore, the filter needs to be shifted in either direction to further isolate the other two sources within the power spectra. To shift the filter, a phase shift given by the exponential,

$$\psi = \exp\left(2\pi f_0 t i\right) \tag{6}$$

where t is the time axis of the filter, and f_0 is the frequency shift the filter will experience. Phase shifts were created for the other sources depending on the frequency shift found in the previous point and applied to the central filter. The superposition the three filters on the power spectra in the frequency domain can be seen in Figure 4. Smaller bandwidth filter could have been used however, as the individual sources are not close to each other there is no need for fine tuning of the bandwidth of the filter.

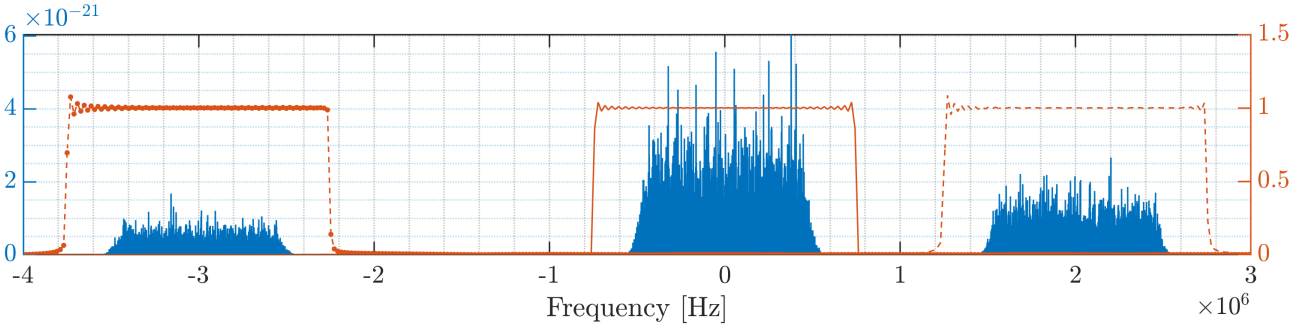

Figure 4: *The power spectra of the signal with the three filters that will be used to isolate each individual source from the signal. Note the three different line styles of the filters to indicate that each style represents a different filter.*

Figure 4 provided a diagnostic tool into the positions and widths of each of the filters in the frequency domain. The signal was convoluted with the filters, individually, within the time domain in order to finally isolate each source. The isolated sources can be seen in Figure 5.

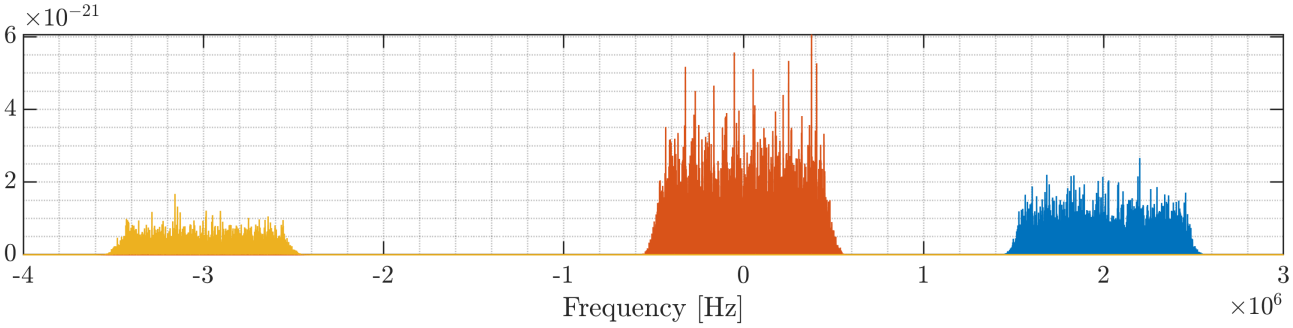

Figure 5: *The isolated signals within the frequency domain after the filtering.*

The power for each of the, now isolated, signals was computed with the aforementioned Eq. 1. The powers for each of the sources are shown in Table 2. It should be noted that the sum of the powers of each of the isolated signals is in very close correlation with the original power calculated in point 1. The error in the power were calculated in a similar fashion as done in point 1 with Eq. 3.

5. Estimate the equivalent brightness temperature of each detected source and assess the theoretical accuracy of the estimation.

Source	**Power** [W]	**Bandwidth** [MHz]	**Brightness Temperature** [K]
Left	$(2.7216 \pm 0.0138) \times 10^{-16}$	1.00471	9.978 ± 0.099
Central	$(5.5381 \pm 0.0553) \times 10^{-16}$	1.00180	40.041 ± 0.400
Right	$(1.3842 \pm 0.0271) \times 10^{-16}$	1.00489	19.616 ± 0.195

Table 2: *The calculated powers of each of the sources present within the signal and their corresponding brightness temperatures.*

The brightness temperature of each of the detected sources can also be calculated in a similar manner to the procedure shown in point 1 with Eq. 2. The bandwidths and powers for the individual sources calculated in point 3 and point 4, respectively, were employed. The results can be seen in Table

2. The brightness temperatures calculated are reasonable compared to absorption studies conducted on the 21cm Hydrogen line.[Roy13] However, without further knowledge of what these sources are and their approximate distance from the observer limited conclusions can be formulated.

6. Propose and implement an algorithm to estimate the direction of arrival of the radiated signals.

Until now only one of the signals were being utilised for the analysis. The other signal was carried through the same procedure as the first to isolate each of the sources embedded in the original signal. Focusing initially on the isolated signal which contains the central source, the cross-spectrum of the two signals can be taken in order to calculate the phase difference between the two signals. This can be done as the two signals are assumed to be generally identical but shifted slightly by a delay due to the different positions of the antennae. Therefore, in order to calculate the interferogram the cross spectrum can be modelled as

$$z_1(t) * z_2^*(t) = |z_1(t)|^2 \exp(-2\pi f_0 \tau i) \tag{7}$$

where τ is the delay between the two signals. `angle`, a matlab function, was applied to the results of the interferogram to obtain the angular difference for one wavelength at the central frequency of a given source. The physical frequency was used for each of the sources rather than solely the frequency shift therefore, all of the frequencies were shifted by 1420MHz. This angular difference spans between -π and π and the results of this procedure with the three sources and their signal corresponding counterparts are shown in Figure 6.

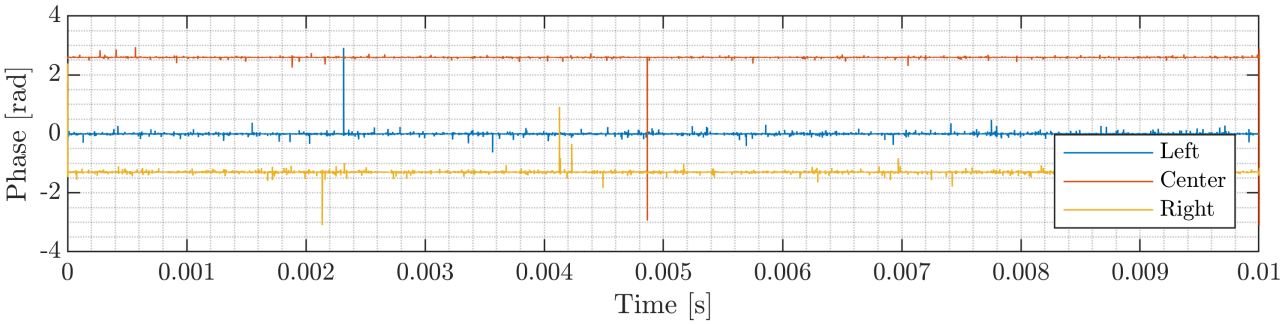

Figure 6: *Angles of the interferogram between the signals of the first and second antennae for each of the individual sources.*

There is some *noise* evident in the each of the results, most likely due to sensor noise of each of the antennae propagating through the calculations. Accordingly, the mean and standard deviation were calculated employing the `mean` and `std` matlab functions.

Source	**Angle** [°]
Left	$(0.3983 \pm 1.1543) \times 10^{-5}$
Central	$(9.9990 \pm 0.0023) \times 10^{-2}$
Right	$(-4.9919 \pm 0.0021) \times 10^{-2}$

Table 3: *Angles of the sources relative to the central axis between the two antennae.*

The distance travelled of a signal over time τ, assuming a refractive index $n = 1$, is simply $d = \tau c$ with c being the speed of light. Therefore, the difference in the distance travelled by the signal between the two antenna is given by

$$d = \tau c = \frac{c}{2\pi f}\phi \tag{8}$$

where f is the frequency of one of the sources in question and ϕ corresponds to the phase angle calculated via the inteferogram. Finally, as presented in Figure 7 the relative angle to the source depending on delay between the two antennae can be represented by simple trigonometry. As the separation of the

two antennae is 50m, the calculated values for the angle of the sources relative to the central axis has been tabulated within Table 3.

It should be noted that the *first* signal was used as the reference signal to base the cross-spectrum from and is assumed to be to the left antenna. Therefore, as a positive phase corresponds to the signal being ahead of the reference signal then a positive value coincides with the source being to the right of the central axis, in a similar fashion to Figure 7.

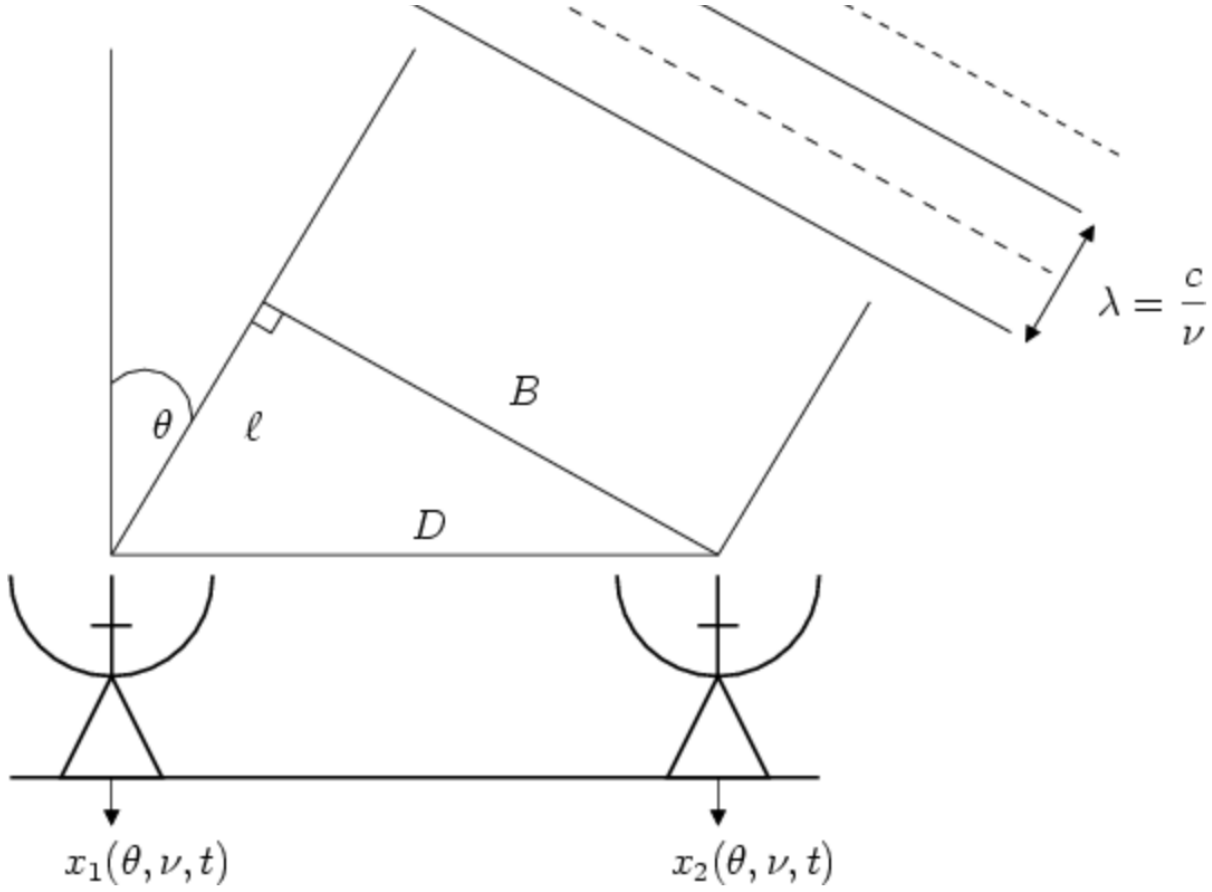

Figure 7: *Simple illustration of the angle of a given source between two antenna using the signal delay.[Int20]*

7. Evaluate empirically the accuracy of the estimation of the angle of arrival of the previous point.

The error of the angle of arrival was computed using error propagation from the result of `std` during the determination of the phase difference of the two different signals for the three corresponding sources. The only error was assumed to be that of the phase and all other potential errors, for example the separation of the antenna, to be 0. Standard error propagation was conducted throughout the calculations required to reach the final result for the angle. The results of the error have been presented alongside final values in Table 3.

2 Problem 2: Estimation in SAR Interferometry

1. Propose an algorithm to estimate terrain topography and displacement rate at any pixel location from the set of interferometric phases.

The interferometric phases can be modelled as

$$\varphi(n) = k_z(n) \cdot z + k_v(n) \cdot v + w(n) - w(0) \tag{9}$$

The height-to-phase, $k_z(n)$, and velocity-to-phase, $k_v(n)$, conversion factors have been provided and consist of 6 different values creating two column vectors. The noise has been stated to be un-correlated therefore, implementations of the Best Linear Unbiased Estimator (BLUE) or the Least Squares (LS) when constructed result with the exact same values. However, even though this is the case the estimator will be constructed following the BLUE algorithm as for future uses if the noise becomes correlated the estimator would require minimal changes potentially future proofing the estimator.

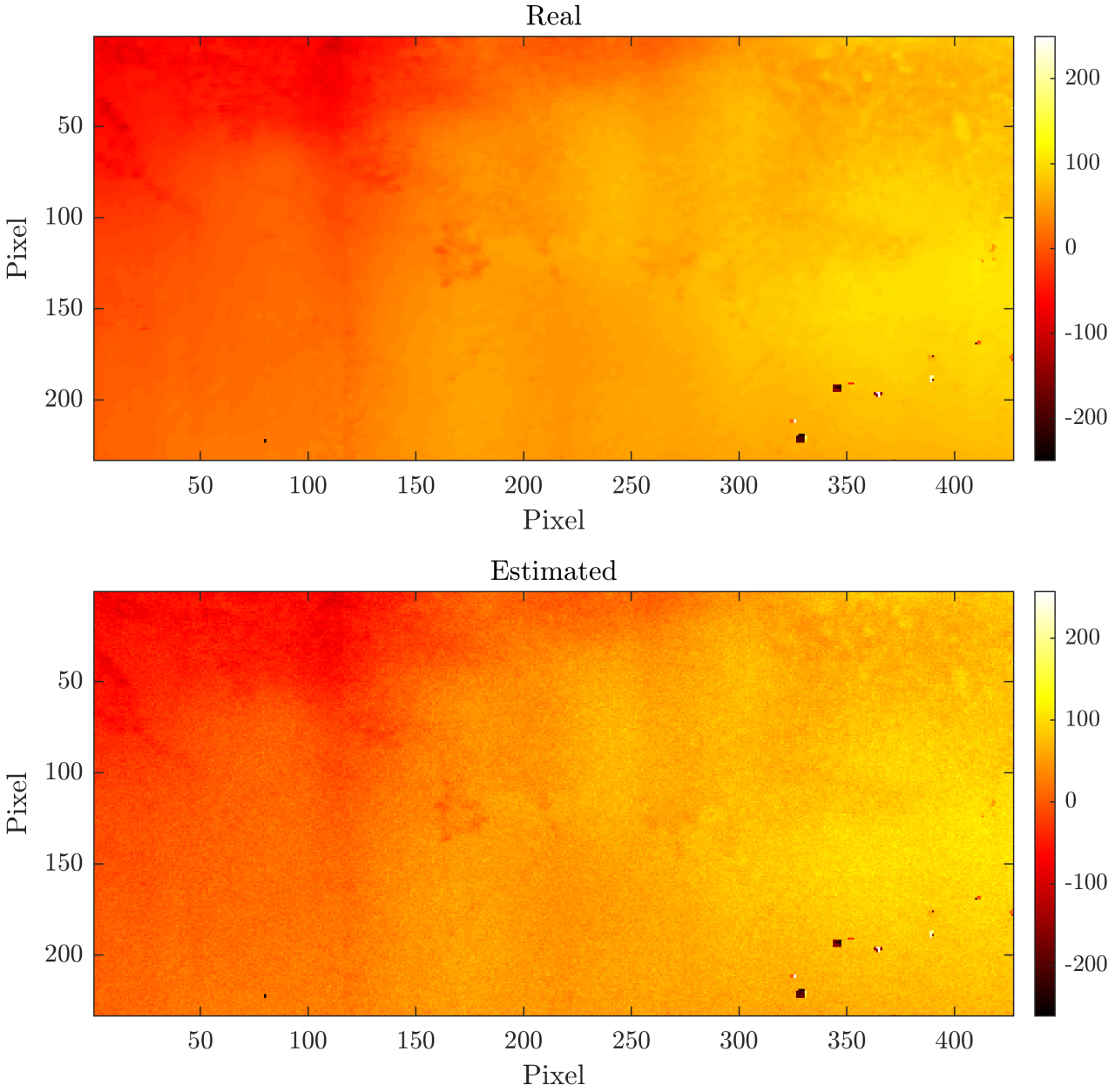

Figure 8: *A comparison between the real provided data and the estimated topology calculated using the estimator proposed in point 1.*

As there is data for two parameters, z and v, embedded in the interferometric phase the matrix A was created. This is a 6×2 matrix with k_z as one column and k_v as the other. The noise matrix, C_w,

was then constructed as a 6×6 diagonal square matrix with $\sigma_w^2 = 9$ at the diagonal elements. The BLUE algorithm then requires the following matrix to be constructed

$$B^T = \left(A^T C_w^{-1} A\right)^{-1} A^T C_w^{-1} \tag{10}$$

where B^T results in a 6×2 matrix.

The interferometric phases consist of a 3D matrix of size $233 \times 427 \times 6$ which corresponds to 6 individual values for a given pixel of a picture with resolution of 233×427. Extracting these 6 values into a single 6×1 matrix coincides to the y vector to be used in aid of estimation of the z and v values for a given pixel.

Finally, the BLUE algorithm can now be used to estimate the values of z and v with

$$\hat{x}_{BLUE} = B^T y \tag{11}$$

which provides, in this case, a 2×1 matrix with one value corresponding to z and the other to v.

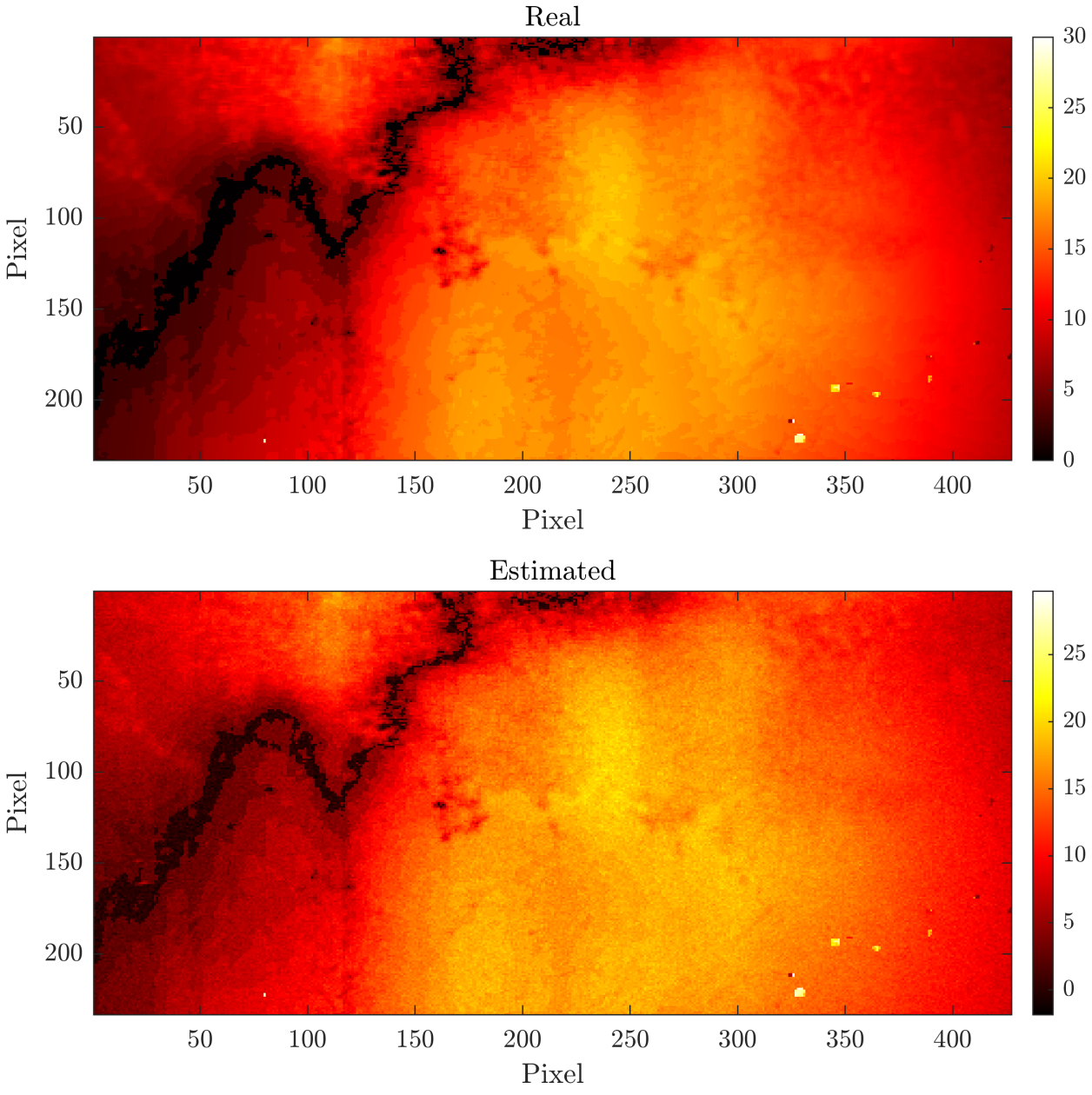

Figure 9: *A comparison between the real provided data and the estimated displacement rate calculated using the estimator proposed in point 1.*

2. Evaluate estimation accuracy on a theoretical basis.

The estimation accuracy of the estimator can be extracted from the co-variance matrix. For the BLUE algorithm this is defined as,

$$Cov_{BLUE} = B^T C_w B = \begin{bmatrix} 51.8638 & -0.8237 \\ -0.8237 & 0.0667 \end{bmatrix} \tag{12}$$

Extracting and taking the square root of the diagonal components results in the theoretical standard deviation for z and v values being 7.2017 and 0.2582, respectively.

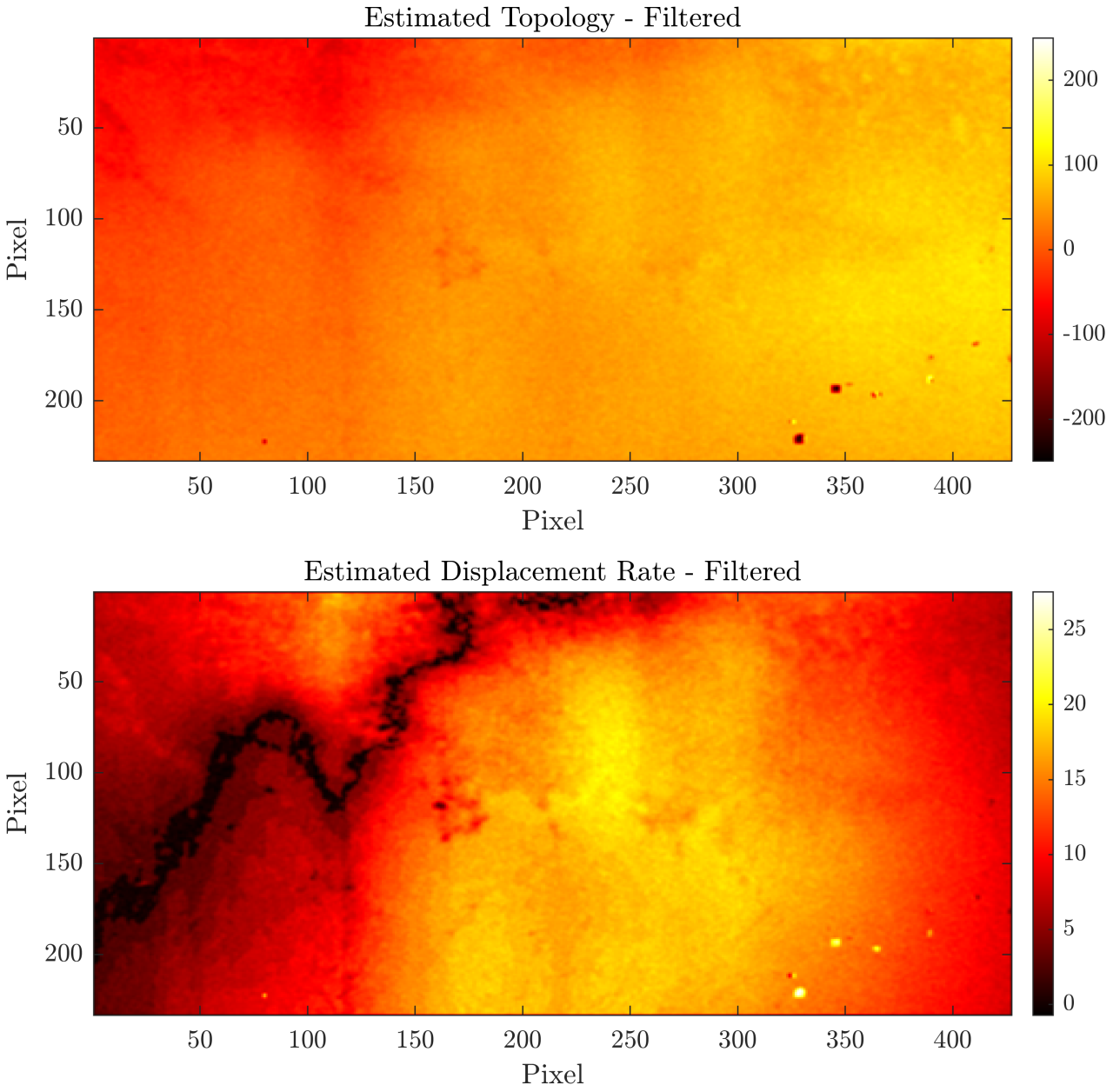

Figure 10: *The estimated plots for z and v after a low pass filter has been applied.*

3. Apply the estimation algorithm on the data-set, and assess estimation accuracy empirically based on the knowledge of true topography and displacement rate. Comment on the result.

A double nested for loop was constructed so that the 6 interferometric phases for a single pixel were extracted and utilised within Eq. 11 so that the estimates of z and v for that pixel can be determined. The following Figures 8 and 9 show the results of the real data-set compared to the estimated one for the z and v, respectively.

From an initial visual inspection of these graphs it is evident that the estimator provides a generally good result following most of the features of the real data. However, in order to define the accuracy numerically further analysis is required. The difference between the real and estimated value was calculated to provide an equivalent size matrix corresponding to the error at a given pixel. The standard deviation of this *error* matrix was determined using the in-built matlab function `std2` and resulted in 7.3697 and 0.5480 for z and v, respectively.

These values are in close agreement with the theoretical estimation accuracy of the estimator calculated in the previous point. However, these discrepancy between the two values arises from actually taking into account real data resulting in an experimental or empirical error rather than just a purely theoretical one and should be prioritised over the theoretical error when discussing about the estimation accuracy. This can be analogous to comparing experimental data to a predefined model where although a theoretical estimate on the accuracy can be made there should be a preference on the experimental error as this provides a direct performance value of the model/estimator.

4. Discuss whether estimation accuracy can be further reduced by spatial filtering, and to what extent.

From 8 and 9 it can be noticed that for the estimated plots there is so additional *speckling* within the data that is not present with the true values. Therefore, as this takes up the majority of the plot there should be a focus on smoothing out these features. Several low pass filters were trialled with a 6-th order `fir1` function providing the best results. The spatial filter was performed after the estimation procedure was completed as these artefacts are assumed to be generated by the estimator itself.

Performing the same procedure as in the previous point to calculate the standard deviation of the filtered results compared to the real data results in values of 4.7356 and 0.4826 for z and v, respectively. This is a reasonable improvement in the error of the estimated values, especially for z, the topological map. The filtered plots are shown in Figure 10 and from a visual inspection it can noticed that the noise previously found within the estimated values are now reduced.

However, in the south-east corner of the image the high resolution filters are now also smoothed out. Therefore, if further refinement of the estimation accuracy is required an algorithm could be developed to filter out the lower spatial frequencies whilst maintaining the higher spatial frequencies in certain locations, thus reducing overall noise in the image but maintaining detail. Exclusion masks could be created around areas of high spatial resolutions and could be re-introduced after the low-pass filtering. This kind of filter would require more time for further experimentation and development to be created and to be able handle general data.

5. Assume that the $k_z(n)$ and $k_v(n)$ conversion factors are proportional to each other. What would be the impact on estimation accuracy on topography and displacement rate?

As the two vectors within the aforementioned A matrix, defined in point 1, are now parallel issues arise when attempting to create the B^T matrix as before. The A matrix is now singular therefore, it can not be inverted completely removing the ability to even create the estimator matrix. The impact this has on the estimation accuracy is then a difficult question to answer as there is no estimator thus, there can not be a corresponding accuracy for that estimator. Boh.

References

[Int20] Interferomertry. 2020. URL: http://www.astro.puc.cl/~rparra/tools/ASTRONOMY/node2.html.

[Roy13] Nirupam Roy et al. "The temperature of the diffuse Hi in the Milky Way – I. High resolution Hi-21cm absorption studies". In: *Monthly Notices of the Royal Astronomical Society* 436.3 (Oct. 2013), pp. 2352–2365. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1743. URL: http://dx.doi.org/10.1093/mnras/stt1743.