Analysing the Performance of SPLITTER

A Noise Removal Algorithm for DESHIMA 2.0

Liam Ligthart

TU Delft

Analysing the Performance of SPLITTER

A Noise Removal Algorithm for DESHIMA 2.0

by

Liam Ligthart

in partial fulfillment of the requirements for the degree of

Bachelor of Science
in Applied Physics

at the Delft University of Technology,
to be defended on Wednesday, July 6, 2022, at 10:00 AM.

Student number:	5107180		
Project duration:	April 19, 2022 – June 22, 2022		
Thesis committee:	Dr. A. (Akira) Endo,	TU Delft	Main supervisor
	Dr. M. (Matus) Rybak,	TU Delft / Leiden University	Daily supervisor
	Dr. A.F. (Sander) Otte,	TU Delft	Examiner

Cover image by NAOJ [1]

TU Delft Delft University of Technology

Abstract

More insight has been gathered on the performance of the SPLITTER (Stationary spectrum Plus Low-rank Iterative TransmiTtance EstimatoR) algorithm as developed by Brackenhoff for denoising data gathered from observations of high-redshift galaxies. By using matrix decomposition to split the gathered data into a low-rank atmosphere matrix and a sparse matrix with the signal and photon noise, the algorithm avoids the subtraction of two noisy signals, and therefore the factor of $\sqrt{2}$ additional noise that comes with it. The algorithm was specifically developed for DESHIMA 2.0 (DEep Spectroscopic HIgh-redshift MApper), a wideband spectrometer that achieves a bandwidth from 220 to 440 GHz, using 347 spectral channels. Previous performance tests of the algorithm included the comparison of the weighted root mean square error between SPLITTER and the usual technique of Direct Subtraction on realistic data simulated using the TiEMPO (Time-dependent End-to-end Model for Post-process Optimization) software package. This resulted in a 1.7 improvement factor for the whole spectrum and 1.3 in the emission line area.

The performance of SPLITTER on two key components of the spectrum have been analyzed separately. First, the measurement of the continuum has been analyzed by using TiEMPO to create realistic simulations of the observation of custom spectra with a linear continuum. SPLITTER showed to be more precise as the noise level was lower, but less accurate, as there was a systematic offset in the estimated continuum. Using a modified black body model for the continuum and assuming the relative offset is independent of the strength of the continuum, the observed offsets and errors were propagated to offsets in estimations of dust temperature T_{dust} and spectral emissivity of the dust β. Because of the offset, SPLITTER also showed a systematic offset in estimated T_{dust}, but as the algorithm is more precise, it performed better at estimating β, since β determines the shape of the spectrum and has less influence on the strength. Second, to test the detection of emission lines, custom spectra have been created containing the same linear continuum and single spectral line at four different frequencies. Each line was set to have a known signal to noise ratio compared to the photon noise in its frequency bin. The retrieved signal to noise ratio as compared to noise of neighboring bins showed an improvement of 1.9 for SPLITTER compared to Direct Subtraction for bright lines. Weak lines did not show any improvement in SNR. There seemed to be no correlation between continuum overestimation and emission line measurements.

Conclusion is that SPLITTER definitely shows improvement in noise reduction, but comes with an overestimation of the continuum. The consequences of this are that Direct Subtraction is still preferred for estimating dust temperature, but for estimations of spectral emissivity and detection of emission lines, SPLITTER is more robust.

Acknowledgements

I would like to thank the creator of the SPLITTER algorithm, Stefanie Brackenhoff, for your helpful attitude during my project and for providing me with the many files and snippets of code you used to run and test the algorithm. Especially during the first weeks of my project, I appreciated the time you spend carefully explaining the algorithm to me and answering my questions even though you were occupied with your own research. I gained a lot of critical knowledge on the algorithm during our meeting in Leiden, and don't think my project would be the same if you hadn't taken the time to meet up.

Secondly, I want to thank my daily supervisor Matus Rybak. Even though you were often working in Leiden, at a conference abroad, or stuck at home because of a broken wrist, you were always very quick to respond to me and widely available for an online meeting. I appreciated the effort you spend during the evenings and weekends answering my questions, commenting on my report, or running simulations on your virtual machine. You always put in more effort than I wanted to ask of you, for which I am grateful.

Furthermore, I want to thank my main supervisor Akira Endo, for providing me with valuable insights and helping me organize the structure of the project, assuring that everything went smoothly. Special thanks to the third member of my assessment committee as well, Sander Otte, for the time and effort you need to put into reading my thesis and for being there for my presentation and defense.

Also, I would like to thank the MKIDs and Terahertz Sensing groups for letting me be part of their comfortable and productive work environment.

Contents

1 Introduction

The practice of astronomy dates back thousands of years. People have always been intrigued by the phenomena happening above our own heads. In ancient times, celestial bodies were associated with gods and spirits, and natural phenomena such as tides, rain, or seasons were linked to the position of the stars. Charts and maps had been made, but for a long time, this was only based on observations with the naked eye.
It wasn't until the early 17th century that Galileo Galilei placed two lenses at a specific distance, creating a telescope that enlarges faraway objects. His first observations were the craters and mountains on the moon, which, before that, was assumed to be smooth [2]. Many more observations were made with this simple telescope of his, and the development of better instrumentation has started to become a field of interest.

Today there are many different ways to look at the sky. Due to advancements in electronics, we are not limited to the bandwidth of visible light anymore but can observe much larger parts of the electromagnetic spectrum. As much of the wavelengths in the visual light spectrum are often absorbed by clouds of gas and dust, many galaxies become obscured from our view.
Thick clouds of gas are a requirement for star formation, and there exist a lot of galaxies with high star formation rates that also contain a lot of dust, therefore making them invisible when we only observe visible light. However, these so-called Dusty Star-Forming Galaxies (DSFGs) re-radiate absorbed energy by dust and gas in the far-infrared wavelengths [3]. Development in observational tools can therefore give us critical knowledge about star formation in these obscured galaxies.

DESHIMA is a wideband spectrometer that will achieve an octave bandwidth from 220 to 440 GHz with its 2.0 version. The DESHIMA 1.0 has been successfully mounted on the Atacama Submillimeter Telescope Experiment (ASTE) and has shown promising results for instantaneous wideband spectroscopy [4]. A key factor that determines the necessary observation time is the signal-to-noise ratio. Galactic signals are often weak and because of the warm atmosphere and the photon noise caused by the instrument, the total noise level is usually much higher than the signal, forcing the necessity for long observation times.
Current noise reduction methods have a critical influence on the noise, as the common technique of subtracting a noisy off-source signal from a noisy on-source signal will increase the noise by a factor of $\sqrt{2}$. Therefore, SPLITTER has been developed by Brackenhoff [5], based on a method from Taniguchi [6], that avoids this subtraction of noisy signals and therefore the additional factor of $\sqrt{2}$. This method, however, has only been tested on a few simulated observations, and more knowledge on the robustness of the algorithm could be obtained by applying more detailed testing.

In this bachelor's project, the reliability and performance of the SPLITTER algorithm are evaluated, using extensive simulated observations. In Brackenhoff's thesis, four versions of SPLITTER have been introduced. The version used in this thesis is an object-oriented version of the algorithm combining Singular Value Decomposition and Alternating Direction Method of Multipliers ([5] for detailed descriptions of the different versions). The analysis is done by looking at the two different components in the spectrum separately.

The first question that will be answered in this project is:

> To what extent can we estimate the shape of the continuum created by the modified black body radiation of the dust clouds, and how well will the physical parameters dust temperature and spectral emissivity be estimated?

The modified black body continuum reveals much information on two important physical parameters of the dust: T_{dust} and β, which, in their turn are related to star formation rate. To gain detailed knowledge of this star formation rate, it is therefore of vital importance that the continuum can be estimated with as little error as possible.

The second question that this project focused on is:

> How much better is SPLITTER in detecting spectral emission lines, and how well does it estimate the flux of such a line?

As one of the main goals of DESHIMA 2.0 is to detect the bright [CII] fine-structure line and the CO rotational lines, it is important to know how sensitive SPLITTER is to detecting a single line.

In chapter 2, the relevant theory for this project will be explained, followed by chapter 3 which describes the path to answer the first research question. Chapter 4 will focus on the second research question, and chapter 5 will discuss the conclusions of the project.

2

Characteristics of the Source, Signal, and Instrument

2.1. Dusty Galaxies at High Redshift

Many galaxies contain large amounts of interstellar dust. In the past, this dust was only seen as an obscuring medium that was absorbing the intensity of the light from galaxies behind it. That mindset has changed, for today the (far-)infrared emission from these dusty galaxies is used to describe physical conditions [7], bringing light to some of the most fundamental aspects of the universe, like star formation. After the violent collapse of high amounts of accreted gas, forming a newborn star, remnants of this gas and formed dust particles get scattered all around these newborn stars [8],[3]. Gaining insight into these areas greatly boosts our ability to describe this phenomenon.
New, hot stars emit thermal radiation on a broad spectrum, but the main source of light, UV radiation (λ = 100 - 400 nm) and visible light (λ = 400 - 750 nm), is greatly absorbed by thick clouds of dust and gas. In return, the dust and gas also re-emit this absorbed light as infrared wavelengths. Due to developments in instruments for infrared detection, thousands of these Dusty Star-Forming Galaxies (DSFGs) have been located [3], which, after redshift of the signal, can be studied in the (sub)mm wavelengths.

The light absorbed by the dust will cause it to heat up, and collisions with gas particles in between the dust cause the temperature of the gas to rise until they are both in thermal equilibrium. The energy is re-emitted in two main ways: the dust emits a smooth continuum over a large bandwidth while the gas cools down by sending radiation in the form of thin emission lines.
As the dust absorbs starlight, it heats up, and all bodies with a temperature emit radiation. Often, a simplified model of a fully absorbing body called a black body is used, that radiates according to

$$B_\nu(T) = \frac{2h\nu^3}{c^2\left(e^{\frac{h\nu}{kT}} - 1\right)}, \tag{2.1}$$

where T is the temperature of the dust (commonly between 10 and 100 K [9, 10]), h, k and c the Planck constant, Boltzmann constant, and speed of light respectively, and ν the emitted frequency.
However, when observing dust clouds it is important to note that they are not perfect black bodies. Not only do they not absorb all incoming radiation, but reflect small parts too, and tend to vary in opacity [11]. The continuum is therefore often approximated by the modified black body model:

$$S(\nu, T) \propto \left(1 - e^{-\tau(\nu)}\right) B_\nu(T) \propto \frac{\left(1 - e^{-\tau(\nu)}\right)\nu^3}{e^{\frac{h\nu}{kT}} - 1}, \tag{2.2}$$

where $S(\nu, T)$ is the emitted flux density by the source (the amount of energy received per unit time, area and frequency), $\tau(\nu) = (\nu/\nu_0)^\beta$, the optical depth, and ν_0 the value where the optical depth is 1, often measured to be 1.5 THz [3]. β is the spectral emissivity, a factor depending on the shape of the dust particles, commonly found within the range of 1 to 2 [12]. Knowing both T_{dust} and β gives vital information on the luminosity of the galaxy in the far infrared region, which in its turn is expected to have a direct relation with the star formation

rate in that galaxy.

The emission lines are only present at some specific frequencies. The gas cools down the most through transitions between the two fine-structure energy levels of single ionized carbon (C^+), commonly denoted as [CII]. The transition between these levels is very rare and requires enormous amounts of gas to detect. The rest frequency of the emitted photon is $\lambda_0 = 158\mu m$ [13]. The high abundance of ionized carbon, combined with the fact that radiation at this frequency is almost unaffected by attenuation from the dust, makes this the brightest emission line from the FIR region to meter wavelengths [14]. Another common process that emits in the FIR region is the transition of a CO molecule to a lower quantized rotational state, denoted by J. As the frequency increases linearly with the rotational state J, the rotational emission lines are often called the "CO ladder", starting at $\nu_0 = 115$ GHz for $J = 1 \rightarrow 0$ [15]. The fact that the spacing between these lines is linear, and the size of the frequency spacing depends on the redshift, makes this another good indicator for measuring the redshift of a galaxy.

The infrared radiation from the dust and gas will be altered by redshift. Redshift elongates the incoming waves, causing the observed frequency to be lower. It is caused in three ways [16]:

1. Doppler effect. As the source and observer move away from each other, the emitted radiation appears to have a longer wavelength.
2. Cosmological redshift (or expansion redshift). Radiation traveling through an expanding universe gets elongated to larger wavelengths.
3. Gravitational redshift. Radiation emitted at a source with stronger gravity gets stretched by the high curvature of nearby space.

All three effects can be combined into one parameter z, where the fraction of emitted and observed wavelength is defined as

$$1 + z = \frac{\lambda_{obs}}{\lambda_{em}}. \tag{2.3}$$

The desired targets of DESHIMA have redshifts around $1 + z \sim$ 1-10 [4]. Due to this wide range of redshifts, the interesting parts of the spectrum (like the [CII] line) are shifted to a broad range of frequencies in the order of 100 GHz. It is therefore important that an instrument with a large bandwidth could be made so that the [CII] line is captured for this wide range of redshifts.

An example of such a spectrum with luminosity $10^{13.6}$ $L_\odot$ (where $L_\odot$ represents the luminosity of the sun, $3.828 \cdot 10^{26}$ W) and redshift $z = 4$, has been plotted in figure 2.1. This figure has been created using the GalSpec python package [17], which simulates a realistic spectrum.

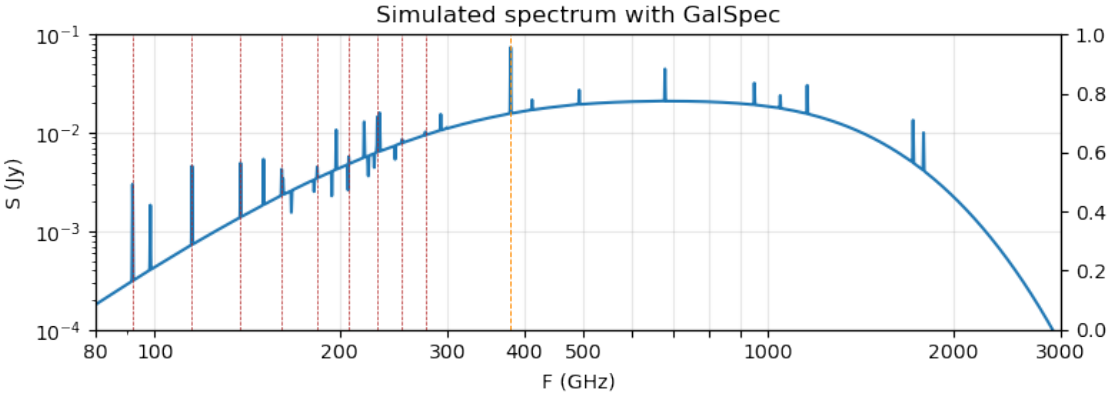

Figure 2.1: Example of a spectrum from a DSFG wit luminosity $10^{13.6}$ $L_\odot$ and redshift $z = 4$, plotted over a very large bandwidth in log-log scale. Red lines indicate the frequencies of the 5th to 12th CO rotational emission lines and the orange line indicates the frequency of the [CII] fine structure line. The rest of the peaks originate from other atoms/molecules.

2.2. The DESHIMA Instrument

Spectrometers capturing high bandwidth spectra in the region of frequencies on the order of 100 GHz are difficult to create. Examples of older instruments report ultra-high bandwidth as 36 GHz (in the range 74 to 110 GHz) [18]. An example of a recent high bandwidth grating spectrometer is Z-Spec [19]. Using a reflective grating, it disperses the incoming signal into an array of 160 channels, achieving an observed frequency band from 190 to 308 GHz, and therefore a bandwidth of 118 GHz. The downside of this instrument, however, is that it is very large (order of meters [20]). This large size makes it harder to cool the instrument and limits the ability to create multiple spatial pixels.

DESHIMA (DEep Spectroscopic HIgh-redshift MApper) is a so-called Integrated Superconduction Spectrometer, ISS in short. An ISS works with a single data readout line where the signal enters the chip. Here, the signal passes an array of bandpass filters, ideally only transmits through one of them and gets detected by this channel [4]. Because of this filter array, the observing bandwidth of each channel is lower compared to a standard Fourier transform spectrometer, which observes the entire bandwidth at once. This causes photon noise to be far lower for each detector [21].
Many spectrometers that directly analyze the incoming frequency spectrum first need to down-convert the high frequency into a lower one by combining it with a 'mixing' frequency close to the frequency of the radio signal. This will create a signal at a frequency that is the difference between the source and mixing frequency, which can be tuned low enough to be read out. DESHIMA doesn't need to do this, because of the type of detectors it uses. These heterodyne receivers amplify the incoming signal while preserving its phase, inevitably resulting in the so-called "quantum noise" [22]. As the ISS only measures the power, without any interest in the phase, the quantum noise gets avoided. The biggest advantage of DESHIMA compared to other devices such as Z-Spec is that it is extremely small. The entire spectrometer fits on a single chip, sized in the order of centimeters. This allows the usage of multiple chips to create more spatial pixels and observe multiple parts of the galaxy at the same time. A key component needed for achieving this small size and the absence of mixing is the Microwave Kinetic Inductance Detector (MKID).

After each bandpass filter, the signal enters an MKID. The MKID is a superconductor capable of detecting breaks of Cooper pairs due to photon absorption. In a superconductor, the charge is transported either through resistive quasiparticles (excited electrons), or resistiveless Cooper pairs. Cooper pairs are paired electrons that behave like a boson as a result of their total spin being an integer (0 or 1). This causes all pairs to be allowed in a single quantum state, unlike single electrons (fermions), causing resistiveless superconduction [23].

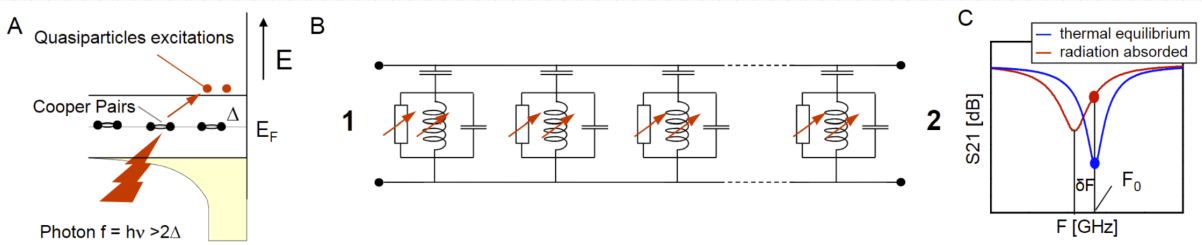

Figure 2.2: A: Incoming radiation can break the superconducting Cooper pairs into quasiparticles if their energy is high enough, changing the impedance of the superconductor as described in 2.4. B: Multiple MKIDs in an array. Changes in resonance frequency can be read out from a single data line using frequency division multiplexing. C: transmission from contact 1 to 2 of a single MKID. The equilibrium position is indicated in blue, and as the KID absorbs a photon, the response changes towards the red curve, therefore changing the response of the readout frequency F_0. Image from Baselmans 2012 [24]

Incoming radiation from photons in the MKID can cause the cooper pairs to be broken into two quasiparticles. The energy of these particles lies Δ above the cooper pairs which are located at the Fermi energy E_F. An incoming photon therefore needs $h\nu > 2\Delta$ in order to excite a cooper pair into two quasiparticles, see figure 2.2a. This change in quasiparticles and Cooper pairs has an effect on the impedance of the superconductor, described by

$$Z_s = R_s + i\omega L_s, \tag{2.4}$$

where R_s is the resistance related to the quasiparticles, and L_s is the kinetic inductance from Cooper pairs

[24]. Therefore, this superconductor has varying resistance and inductance, and combining it with a capacitor creates a circuit with a resonance frequency that changes depending on the input signal. Figure 2.2b shows an array of these superconductors linked with a capacitor, and figure 2.2c has a plot of the response of such an MKID. It is seen that as the resonance frequency changes, the readout signal S21 increases.
A schematic of the DESHIMA 1.0 (produced in 2017) and DESHIMA 2.0 (an improved version currently being worked on) can be seen in figure 2.3. Radiation enters the antenna and propagates through the signal line that has all the bandpass filters connected. After passing through one of these filters with a matching frequency, the signal interacts with the MKID assigned to that frequency band. As each MKID can have a different resonance frequency, the readout signal can be taken from a single line, using frequency division multiplexing [4, 5].

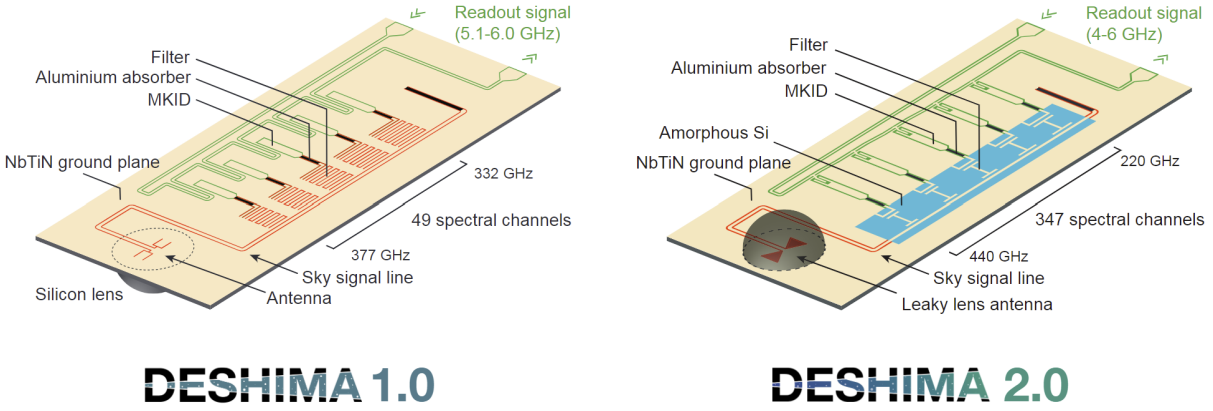

Figure 2.3: Schematics of the DESHIMA 1.0 and 2.0 chip. Radiation enters the antenna and gets filtered by one of the bandpass filters and transmits its power to the connected MKID. All MKIDs are read out at the same time. DESHIMA 1.0 achieves a bandwidth from 332 to 377 GHz using 49 spectral channels, whereas DESHIMA 2.0 will achieve 220 to 440 GHz with 347 spectral channels, improving the frequency resolution $F/\Delta F$ from 380 to 500 for each channel [25]. The 2.0 also uses a leaky-lens antenna, achieving better coupling efficiencies, therefore capturing more light. Image from Endo 2019 [4]

DESHIMA 1.0 was produced and tested in 2017, and with 49 spectral channels, it covered a bandwidth from 377 to 332 GHz. It was mounted on the ASTE telescope (Atacama Submillimeter Telescope Experiment, a telescope using a 10-meter diameter dish antenna to observe sub-mm wavelengths, figure 2.4) during the final months of that year. It showed its fast spectral imaging capabilities and the ability to do on-the-fly mapping, creating composite maps of several galaxies [4].

Figure 2.4: The Atacama Submillimeter Telescope Experiment (ASTE), a 10-meter diameter dish antenna located in the Atacama desert in Chile, at an altitude of 4860 m. Image taken from [1].

One of the observations from DESHIMA 1.0 that was published in the First Light demonstration paper by Endo et al. (2019) [4], of the bright interacting galaxy pair VV114, is shown in figure 2.5. The spectrum shows detection of the CO (3-2) rotational emission line, which is slightly shifted from its rest frequency because of the redshift of $z = 0.02$.

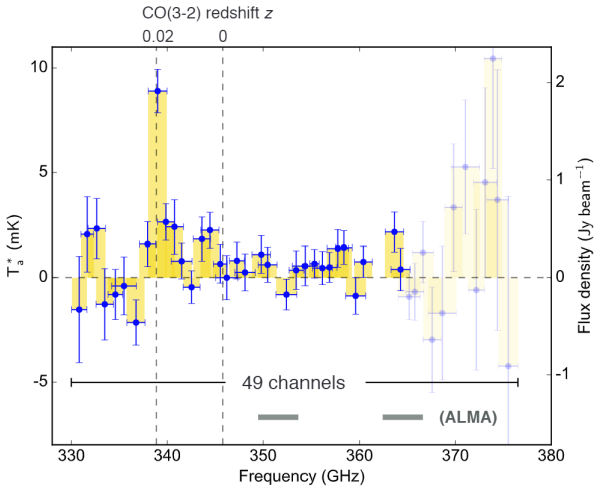

Figure 2.5: DESHIMA 1.0 observations of the VV 114 galaxy. The figure shows a spectrum with a clear peak at the CO (3-2) rotational emission line frequency, after a redshift of 0.02. Horizontal errorbars and yellow area show the full width at half maximum of the filters, vertical errorbars show the standard deviation of the noise in each channel. Channels above 365 GHz are noisy as a cause of the low atmospheric transmittance in that area. Image taken from Endo 2019 [4].

Since then, the development of the improved DESHIMA 2.0 has been continuing, which achieves an octave bandwidth from 220 to 440 GHz with 347 spectral channels. This allows it to detect the [CII] line at redshifts between 3.3 and 7.6, a huge improvement compared to the range of DESHIMA 1.0 (between 4.0 and 4.7) [25]. It also used a leaky-lens antenna, improving coupling efficiency, thus coupling more light to the instrument. Further improvements should allow the performance of DESHIMA 2.0 to be competitive with the current receivers on APEX (Atacama Pathfinder EXperiment) [26].

2.3. Noise and Noise Reduction

Radiation from any object in space travels through the atmosphere. Molecules in the atmosphere (especially water vapor) absorb, scatter, and emit radiation, causing attenuation, disruptions, and the addition of a strong signal compared to the incoming signal from a faraway galaxy. The instrument itself also adds a level of noise to the observed data. In the end, the signal from the galaxy is highly disturbed by both atmospheric attenuation and the large intensity of noise, and on top of that the additional brightness temperature the atmosphere adds by its own radiation, making the weak signal from the galaxy very difficult to observe. To compare the two, the signal from a bright galaxy has a brightness temperature in the order of mK, while the atmosphere adds tens of K for the highly transmissive frequency 'windows', and over 200 K for the highly absorbing frequency bands.

To get a good model for our incoming data, we consider three main sources for our observed brightness temperature:

$$T_{measured}(\nu, t) = T_{sky}(\nu, t) + T_{ph}(\nu, t) + T_{TLS}(\nu, t), \tag{2.5}$$

where $T_{measured}$ is the observed brightness temperature, T_{sky} the brightness temperature of the sky, T_{ph} the photon noise and T_{TLS} the Two Level System noise [5]. Photon noise is in the order of 100 mK, whereas TLS noise is about half a K.
Furthermore, $T_{sky}(\nu, t)$ is composed of the source brightness temperature $T_A^*(\nu)$ and the physical atmospheric temperature $T_{p,atm}(t)$ in the following way:

$$T_{sky}(\nu, t) = T_A^*(\nu)\eta_{atm}(\nu, t) + T_{p,atm}(t)(1 - \eta_{atm}(\nu, t)), \tag{2.6}$$

with $\eta_{atm}(\nu, t)$ the transmittance of the atmosphere, both varying over frequency (because gasses in the atmosphere interact differently with other wavelengths) and time (because of the slowly changing properties of the atmosphere). Note that the brightness temperature of the source is a function of frequency only, as the time for events to occur in the galaxy is so large that we are looking at a static image. $T_{p,atm}$ is, however, only a function of time, as it is a physical temperature. Since it usually takes a few hours for this temperature to change, we can use that it varies very slow with time [5].

The limiting noise of the DESHIMA instrument [4], photon noise, T_{ph} is caused by both shot noise, a consequence of photons being discrete and entering the detector in discrete timesteps, and photon bunching, caused by the random time interval between the arrival of a photon in the instrument, causing them to arrive in bunches. Both are Poisson-distributed effects, and as we have many photons detected and the sampling frequency of DESHIMA (160 Hz) is far below the bandwidth of a channel, we can well approximate this by a Gaussian distribution. This noise is not correlated with either time or frequency of the signal, and its power spectral density (a measure of the strength of the noise) does not depend on sampling frequency f [5],[27].

TLS noise originates in the MKID devices. As a result of excitations from electrons from the ground state to their first excited state, the dielectric constant of the substrate is altered, changing the resonance frequency of the MKID. As this noise arises in each channel separately, the channels are uncorrelated. However, it is correlated in time and shows a slow varying offset in the mean unlike the Gaussian scatter that photon noise adds. It behaves as $1/f$ noise when plotted against sampling frequency f, thus the noise is reduced at higher sampling frequencies [5].
Of course, there are more sources for fluctuations in the signal received, like the noise from the atmosphere, the imperfect shape of the filters, the slightly different observing angle for on- and off-source measurements, and noise originating in the readout line of the DESHIMA, but they are very small compared to the ones given above and are therefore disregarded [5].

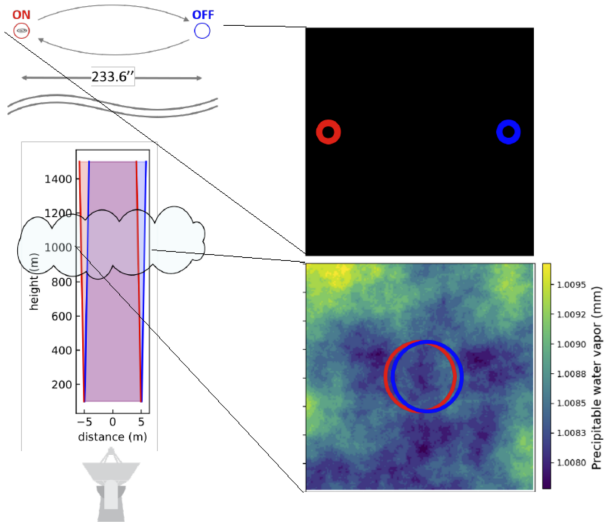

Figure 2.6: Position SWitching (psw) chopping. A quickly rotating chopping wheel is used to switch between on- and off-source observing angles with a high frequency (10 Hz). Left: observing angles of on-source (red) and off-source (blue) samples. Top right: spatial difference between the on- and off-source beam. The size of the circle corresponds to the outline of the beam. Bottom right: spatial difference between the on- and off-source beam, again the outline corresponds to the beam size. As the beams almost completely overlap, the estimated atmospheric effects from the off-source samples are a good approximation for the atmospheric effects at the on-source samples. Image taken from Brackenhoff 2021 [5].

A standard way of denoising the atmospheric noise of an observed spectrum is by both taking on-source measurements (resulting in $T_{sky,on}$ as the sky temperature) and taking a measurement with a pointing angle right next to the source, to take an off-source measurement (with $T_{sky,off}$ as sky temperature). This is done using a position switching (psw) chopper wheel that slightly changes the pointing angle of the telescope with a frequency of 10 Hz. Details can be seen in figure 2.6. The sky temperature as described in equation 2.6 will then be [6]:

$$T_{sky,on}(\nu,t) = T_A^*(\nu)\eta_{atm}(\nu,t) + T_{p,atm}(t)(1-\eta_{atm}(\nu,t)), \tag{2.7}$$

$$T_{sky,off}(\nu,t) = T_{p,atm}(t)(1-\eta_{atm}(\nu,t)). \tag{2.8}$$

Assuming constant atmospheric temperature, combining the two equations and averaging over time yields an estimate of the source brightness temperature:

$$T_A^*(\nu) = T_{p,atm}\langle\frac{T_{sky,on}(\nu,t) - T_{sky,off}(\nu,t)}{T_{p,atm} - T_{sky,off}(\nu,t)}\rangle, \tag{2.9}$$

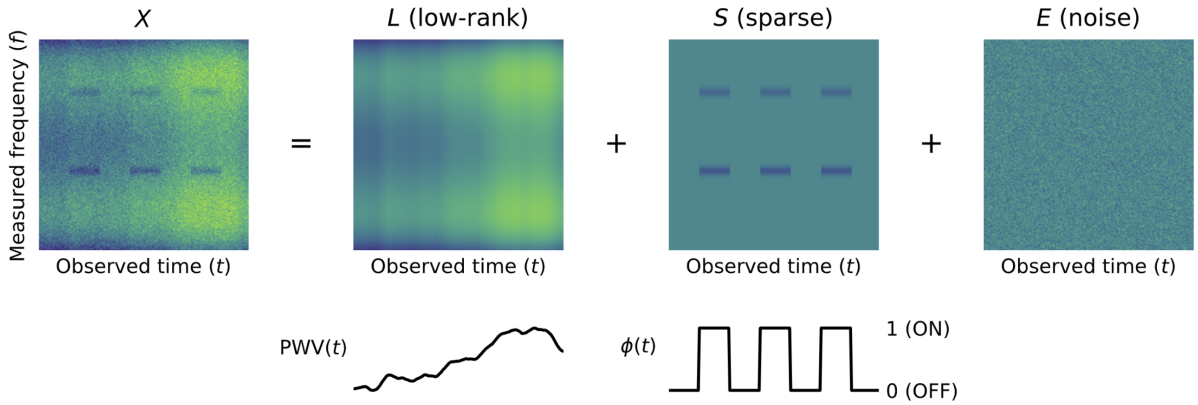

Figure 2.7: The data scientific noise removal approach as proposed by Taniguchi et al. [6]. The data matrix **X** is decomposed into a low rank atmosphere matrix **L**, a sparse signal matrix **S** and a noise matrix **E**. The matrix **L** depends on the PWV (precipitable water vapor) in the atmosphere, and the matrix **S** is chopped because of the chopping wheel, altering the signal to be on and off.

where $\langle\cdot\rangle$ denotes the time average. As the instrument cannot measure on-source and off-source at the same time, the times at which the on- and off-source samples have been taken are slightly different from one another (with on-off switching of 10 Hz, that would mean there's only $1/(10\cdot 2) = 0.05$ seconds delay between them). For simplicity, the same t has been used in the equation above. This technique is often used in astronomy as it is a very straightforward method and does not require long computation times [5].

As the actual observed brightness temperature also contains photon noise and TLS noise, using this technique directly with these values is often done with long observation times, as an effort to let these noise components converge to their mean.
However, this method suffers from the fact that subtracting two noisy signals causes an increase in noise by $\sqrt{2}$. Also, as noise variance decreases proportionally to the square root of the observation time ($\sqrt{t}$), very long integration times are often needed [5].
A method proposed by Taniguchi et al. is to use statistical matrix decomposition to separate the astronomical signal from the noise [6]. The idea is to split a matrix of the observed data (one axis being frequency, the other being time) into three matrices, as seen in figure 2.7:

1. A low-rank matrix representing the atmospheric effects
2. A sparse matrix representing the astronomical signal
3. A noise matrix containing stochastic noise effects

This method does not use the subtraction of two noisy signals and therefore prevents the extra factor $\sqrt{2}$ arising in the noise.

A similar technique was created to use for data generated by DESHIMA. Taniguchi et al. assumed the data was collected from heterodyne receivers, which not only have a higher spectral resolution than DESHIMA, but also a different noise characteristic (it doesn't have the time-correlated TLS noise that DESHIMA has). Also, Taniguchi only estimated sparse sources with only a few emission lines, creating many channels that do not have line emissions and thus help determine the atmospheric effects. However, this doesn't allow detection for the continuum, that spreads across the entire bandwidth [5].
SPLITTER (Stationary spectrum Plus Low-rank Iterative TransmiTtance EstimatoR) is an algorithm created by Brackenhoff [5], that adapted Taniguchi's method of matrix decomposition to be used on data gained from DESHIMA 2.0.

The essence is the same, the data matrix **X** is split up into a low-rank atmospheric matrix **L** containing elements from atmospheric effects and TLS noise, a source matrix **S** from which $T_A^*(\nu)$ can be estimated, and a residual noise matrix **N** that contains elements from TLS and photon noise. As the TLS noise varies relatively slowly in time, the matrices can be split into several chunks in the time domain, allowing the TLS noise to be

assumed constant for each chunk. Because the TLS noise gets eliminated this way, it hasn't been simulated in this thesis.
How SPLITTER estimates the source is by taking a n amount of estimates, each using different chunks of the full data. These estimates first try to estimate a third-order polynomial continuum, and from the residual noise, they pick the location of the Θ highest emission/absorption peaks. As these peaks could be accidental peaks in noise, instead of using the peak height of their own estimate, they pass on the locations of their peaks to a random other estimate of the signal and use the peak intensities of that estimate to add to the total signal. Because of the randomness of noise, the intensity taken from the other estimate will most likely be lower, and less noise is propagated through the algorithm. Signal peaks will always stay high, so they are fully passed on through the algorithm. Thorough information on the algorithm of SPLITTER is described in Brackenhoff's thesis [5].

After developing SPLITTER, the algorithm has been tested by Brackenhoff on a realistic simulation of DESHIMA 2.0 data from a galaxy with luminosity $10^{13.8}$ $L_\odot$ and redshift $z = 3$. The results were a better sensitivity than the usual method of standard subtraction of data. When calculating the weighted root mean square error of both methods, the whole spectrum was estimated 1.7 times better, and the region with spectral lines showed an improvement of 1.3. The weighted root mean square error as used by Brackenhoff is determined by the following equation [5]:

$$\sqrt{\frac{1}{M}\Sigma_\nu[\hat{\eta}_{atm}(\nu)(T_A^*(\nu)-\hat{T}_A^*(\nu))]^2} \tag{2.10}$$

where M is the number of channels, $T_A^*(\nu)$ is the true value for the brightness temperature of the signal, $\hat{T}_A^*(\nu)$ is the estimated value for the brightness temperature, and $\hat{\eta}_{atm}$ is the estimated atmospheric transmission coefficient, computed from the low rank atmosphere matrix: $\hat{\eta}_{atm} = \exp -\mathbf{L}^k\mathbf{1}$ after k iterations.

To test the SPLITTER algorithm on realistic data, a software package called TiEMPO (Time-dependent End-to-end Model for Post-process Optimization) was used, which simulates observations of the DESHIMA 1.0 and 2.0 [28]. It incorporates the entire signal flow from the source, through the atmosphere, through the telescope and the instrument, and outputs the noisy signal in an on-source and off-source timestream of brightness temperature observations. It uses the package GalSpec [17] to simulate a spectrum of a galaxy with a continuum and atomic and molecular emission lines, taking luminosity and redshift of the galaxy as input. Other important input parameters include precipitable water vapor (pwv), which is the height of a resulting water column if all water vapor in a column of air has condensed into rain. This pwv is modeled using existing data from ARIS (Astronomical Radio Interferometer Simulator), the data being a two-dimensional map of the pwv. The parameter called windspeed determines how fast the map of the atmosphere moves over our observation beam from telescope to source and the gridsize is used to determine the resolution of this atmosphere. Finally, an important parameter is the total observation time of a simulated observation.

3

Measuring the Continuum

The continuum emission of a galaxy is modeled by the modified black body equation 2.2. Two vital parameters defining the strength and shape of this emission are the temperature of the dust, and the spectral emissivity index, defined by the shape of the dust particles. When researching these physical properties of the galaxy, the continuum must be measured precisely and accurately. Although these terms seem like synonyms, there is an important difference between them. Precise means little spread in the measurements, and accurate means little offset in the average of the measurements. Figure 3.1 visualises these differences.

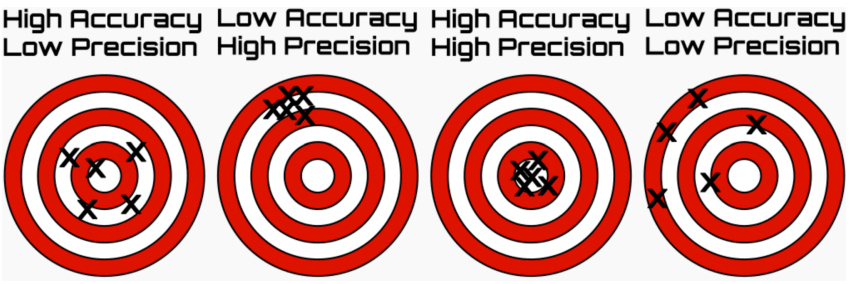

Figure 3.1: Precision versus accuracy. Precise means little spread in the measurements, while accurate means little offset in the average of the measurements. Image taken from [29].

3.1. Offset and Noise in Continuum

Trials have been done to test SPLITTER's performance in estimating the height of the continuum. TiEMPO [28] was used to simulate observations of a custom galaxy, with a continuum that has a linear increase in brightness temperature between the points (ν, T) = [(220 GHz, 0.25 mK), (440 GHz, 0.75 mK)], as seen in figure 3.2. This roughly corresponds to the continuum of a galaxy at redshift 4 with luminosity $10^{13.6} L_{\odot}$ (as plotted over a huge bandwidth in figure 2.1), with the difference that the real one has slight curvature instead of being a straight line.

As TiEMPO takes a spectrum with its units in flux density as input (with the units Jansky (10^{-26} W m^{-2} Hz^{-1})), the created spectrum had to be converted first. For this, the following equation was used: [28]

$$S(T_{sky}) = \frac{2\Omega k T_{sky}}{\eta_{MB} \lambda^2}, \tag{3.1}$$

with T_{sky} the brightness temperature in K, Ω the main beam solid angle, η_{MB} the main beam efficiency (<1 because of side lobes, which are angles at which the antenna also receives a signal, pointed away from the main beam). λ the wavelength and k Boltzmann's constant.

The resulting flux density spectrum could be used as input for TiEMPO. To simulate the response of DESHIMA in the case of no atmospheric and instrumental noise, the response of each of the 347 filters, evenly spaced

in logspace, was calculated separately. To do this, the filter response was modeled by a Lorentzian, with the Full-Width at Half Maximum equal to the width of a bin ($\Delta\nu = \nu/R$, where the spectral resolution R equals 500). It was explicitly chosen not to use a convolution as this cannot make use of a variable Lorentzian.

TiEMPO used a precipitable water vapor of 1 mm (a realistic value for the observation site of the telescope), a windspeed of 10 m/s, a gridsize of the atmosphere model of 1 m, and an observation time of 30 minutes, meaning an on-source observation time of 15 minutes, due to on- and off-source chopping. Usually, observations of high redshift galaxies take many hours, but to reduce simulation runtime and data size, it was chosen to run shorter observations and increase the brightness of the galaxy. To get the same signal-to-noise ratio during an 8-hour (4-hour on-source) observation, a galaxy would be $\sqrt{8/0.5} = 4$ times less luminous. For a galaxy at redshift 4, the corresponding luminosity would be $10^{13.0} L_\odot$, which according to Rybak 2022 [26] is on the lower end of the target galaxy luminosities. Together with the sampling rate of 160 Hz, this generates 288 thousand samples, half of them being on-source due to the on/off chopping. Instead of passing luminosity and redshift to let GalSpec generate a spectrum, the custom spectra were passed on into TiEMPO, which required slight modifications to its code.

On each simulation from TiEMPO, both SPLITTER and Direct Subtraction were used to denoise the data. The parameters used as input for SPLITTER were kept the same for all runs and can be found in table 3.1.

Threshold on η_{atm}	ρ	n	Θ	Iterations	t_l
20%	$5 \cdot 10^3$	10	3	60	1600

Table 3.1: Parameters used to run SPLITTER. Data at frequencies below the threshold η_{atm} is ignored in the algorithm. ρ describes how large the estimate is adjusted each iteration, n describes the amount of estimates for signal vector s are computed, Θ describes the amount of peaks the algorithm will be looking for (set to 3 to keep consistency with the spectra that included emission lines), t_l is the amount of time samples in each chunk the big matrix is split in. At 60 iterations, the algorithm has already converged.

Four longer runs of TiEMPO have been done too, with a total observation time of 1 hour, to see if there is any difference in continuum estimation for longer observation time. To reduce computation time, the support time t_l of SPLITTER has been doubled for these observations. The resulting spectrum using SPLITTER and Direct Subtraction on one of these runs has been plotted in figure 3.2. There seemed to be no noticeable difference in offset for these particular cases.

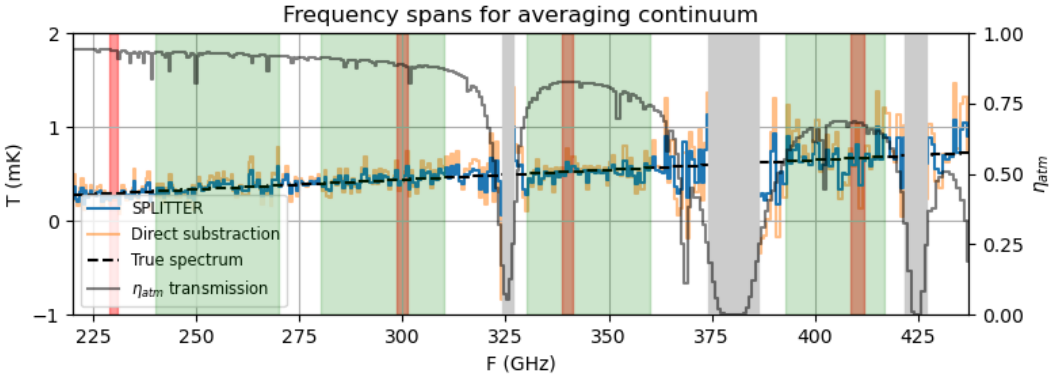

Figure 3.2: An example of SPLITTER (blue) and Direct Subtraction (orange) used to denoise the observations. Here the on-source observation time was 30 minutes. The green areas are the four bands over which the average continuum has been calculated for comparison, centered around 252, 295, 345, and 405 GHz with bandwidths of 30, 30, 30, and 24 GHz. The red areas are frequencies around the simulated emission lines in the spectra with 15 minutes on-source time, which are therefore ignored when averaging the continuum. The grey areas are frequencies for which the atmospheric transmittance was so low that they are ignored in the algorithm.

In order to comment on the performance of both SPLITTER and Direct Subtraction, the brightness temperature of the continuum has been estimated in four frequency areas by taking the average of this area. These areas had a bandwidth of around 30 GHz, and are visualized in figure 3.2. The same simulated observations of spectra with custom emission lines were used for continuum estimation, so the regions around those lines were ignored when calculating the average continuum height.

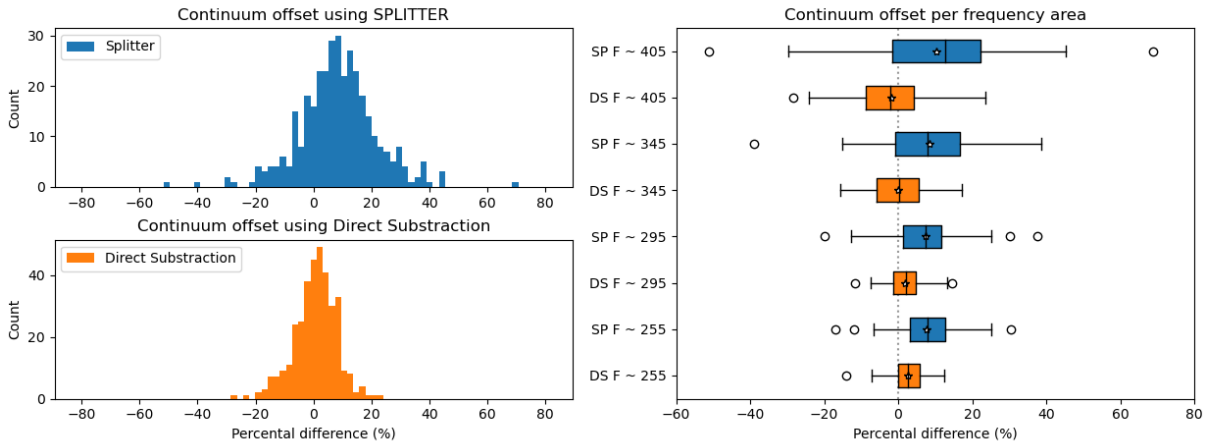

Figure 3.3: Relative offset in average continuum height. Left: histograms of all 4 spans from figure 3.2 combined for SPLITTER (blue) and Direct Subtraction (orange). Right: box plots of the relative offset distribution for each central frequency of the four areas for SPLITTER (blue) and Direct Subtraction (orange).

In total, the relative offset in the 4 frequency spans was calculated for 88 simulations with an on-source observation time of 15 minutes, and the results are visualized in figure 3.3. With all the data from the four frequency spans combined in the two plots on the left side of figure 3.3, it appears that both relative offsets converge to a normal distribution. Direct Subtraction has an average offset of (0.7 ± 7.2) %, which is in line with the expected mean of zero. SPLITTER, however, has a mean at (8.5 ± 13.5) %, showing both a systematic overestimation in the continuum and more spread in this than Direct Subtraction.
Looking at the box plots on the right of figure 3.3, it can be noticed that the spread in continuum offset is higher for higher frequencies. This can be explained by the fact that the atmospheric transmittance is lower in this area (figure 3.2 for reference), therefore adding more noise, resulting in a greater possibility of misestimating the signal.

The same plots can be made for the relative Root Mean Square Error in the continuum at the same spans of figure 3.2. This parameter is defined by the root mean square of the offset each bin has from the actual signal, and dividing this by the average value it should have across the span. It gives a measure of the noise in the spectrum, rather than the offset.

$$RMSE_{relative} = \sqrt{\frac{1}{N}\Sigma_i \left(\frac{T_i - T_{true}}{T_{true}}\right)^2}, \tag{3.2}$$

where i iterates over the N amount of filters in the selected span.

With all the data from the four frequency spans as defined in figure 3.4, SPLITTER scores a median value of 23.8%, whereas Direct Subtraction has a median error of 41.0%. Both curves peak early and fall off more slowly, as visible in the left histograms of figure 3.4. The right plot of that same figure shows that both denoising methods have a noticeably higher relative error in the high-frequency areas. This could be explained again by the lower atmospheric transmittance for these frequencies.
Considering the results from both the offset and relative RMSE, it can be said that SPLITTER is indeed filtering random noise better than Direct Subtraction, as shown by the smaller RMSE. This is also seen in figure 3.2, where SPLITTER shows less jitter than Direct Subtraction. However, it is of utmost importance to realize that the algorithm seems to add a systematic offset to the height of the continuum, while also showing a larger variation in this offset. This leaves the algorithm to (greatly) overestimate the continuum in most cases, with a chance to underestimate it as well.

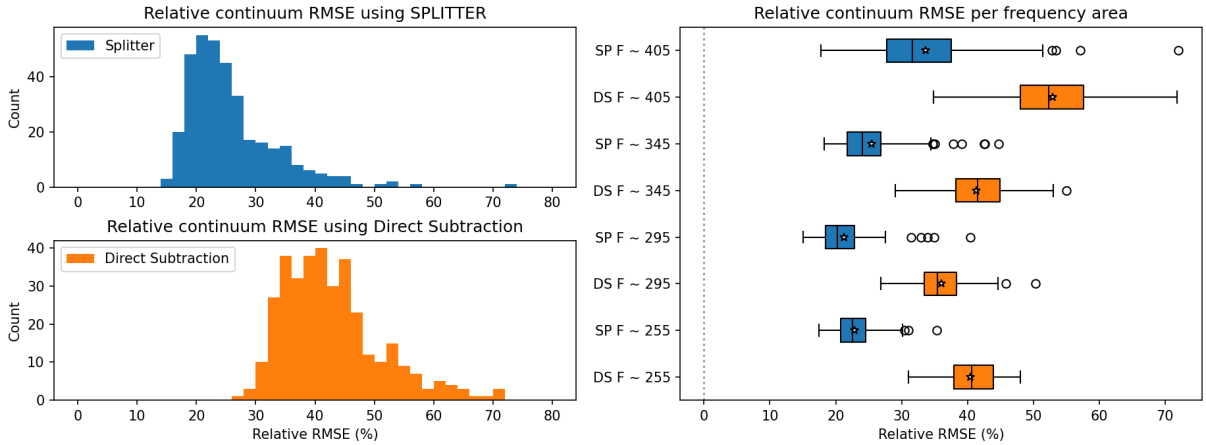

Figure 3.4: Relative Root Mean Square Error (defined in equation 3.2) of the continuum. Left: histograms of the 4 spans from figure 3.2 combined for SPLITTER (blue) and Direct Subtraction (orange). Right: box plots of the relative Root Mean Square Error for each of the four frequency spans for SPLITTER (blue) and Direct Subtraction (orange).

It is interesting to see what effect this overestimation has on estimating the physical parameters of the dust, as is done in the next part.

3.2. Consequences for Estimated T_{dust} and β

The continuum emission of DSFGs that we are interested in is modeled by a modified black body (equation 2.2), with two important parameters: T_{dust}, describing the physical temperature of the dust, and β, describing the shape of the particles.

It is important to note that for this section it has been assumed that the relative offset and noise as previously shown will stay constant for galaxies with different dust temperatures, redshift, and spectral emissivity index. This assumption has been made based on the idea that farther and less luminous galaxies can be observed for a longer time to get the same signal-to-noise ratio as close and bright galaxies when observed for shorter periods.

As the modified black body equation is not a linear function, instead of propagating the error through the equation, it has been chosen to simulate many observations to get the error in the estimate. To do this, a modified black body curve with known T_{dust}, z and β has been taken, and random offset plus noise have been added to four points that are sampled from the curve, with the frequencies of these samples at the center frequencies of the spans from figure 3.2. This simulates how the black body curve is being estimated if we just have our observed data. The observed flux density, therefore, is modeled using two normal distributions for the total relative offset:

$$S_{obs} = S_{true}[N(\mu_{off}, \sigma_{off}) + N(0, \sigma_{noise})] \tag{3.3}$$

It has been assumed that offset is fully correlated across the four frequencies, meaning that if the continuum in one frequency span is overestimated by the mean offset plus one standard deviation, the other bins will overestimate it by their mean plus standard deviation too. Offset as a result of noise is assumed to be completely uncorrelated across all points, and the standard error of the averaged continuum ($\sigma/\sqrt{N}$) has been taken as σ_{noise}. This value was between 4 and 5 % for SPLITTER and 7 to 10 % for Direct Subtraction.

For each iteration, the process looked like figure 3.5. The error that was used as input for the curve fitting function was the standard error in the noise average. This was done because realistically, only this value is obtained from the measurements.

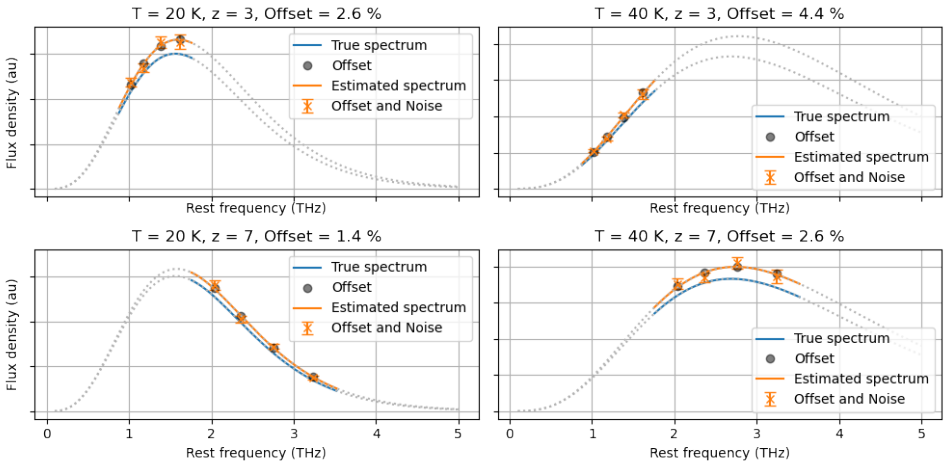

Figure 3.5: Four examples of how the estimation of T_{dust} has been tested (this particular one for SPLITTER). The true spectrum (blue) is sampled by taking the average continuum height over four frequency ranges. Both methods add a correlated offset (grey) and random noise (orange). The errorbars on the orange samples are the error due to random noise and are also used in the curve fit function. The orange line is the estimated spectrum through the curve fit function. The observed bandwidth (colored) shifts to higher rest frequencies as the redshift z increases. The y-axis of plots in different columns is not shared. The effect of a higher temperature is a much stronger spectrum that peaks at a higher frequency, and higher redshift results in an observed rest frequency that is at a higher frequency and has a larger bandwidth. Relative offset in the estimated dust temperature decreases as more of the shape of the spectrum is covered, which is achieved at high redshifts and low temperatures.

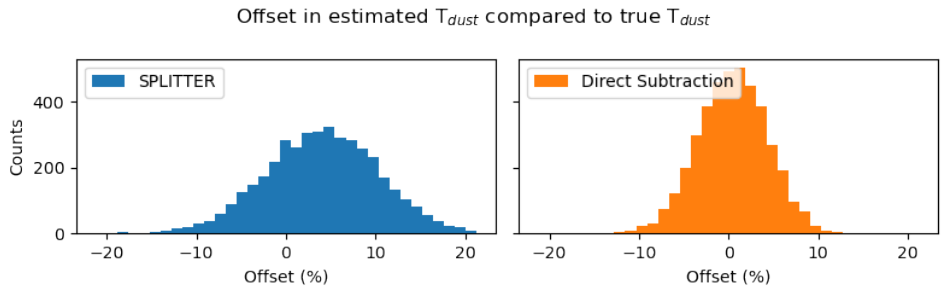

Figure 3.6: Histograms showing the relative offsets in estimated dust temperature when using SPLITTER (blue) or Direct Subtraction (orange) to estimate the continuum, when using a curve fitting tool on 4096 samples with a random offset and noise (equation 3.3). Values are calculated for a galaxy that emits a modified black body spectrum with a dust temperature of 40 K and has a redshift of 3. SPLITTER shows an offset of (4.1 ± 6.2) %, whereas Direct Subtraction centers around (0.4 ± 3.9) %.

3.2.1. β as a Fixed Parameter

First, the parameter β was assumed to have a fixed value of 1.5, as this is often assumed for dusty galaxies [3]. Doing N = 4096 iterations of estimating a spectrum with true dust temperature T_{dust} = 40 and redshift z = 3, results in offsets as plotted in figure 3.6. Here, SPLITTER has an average offset of (4.1 ± 6.2) %, and Direct Subtraction peaks at (0.4 ± 3.9) %. So, for this particular case, the offset in the estimated continuum has indeed propagated towards the estimated dust temperature. It is, however, a smaller relative offset than what was found for the continuum. This can be explained by the exponential dependence of T_{dust} in the modified black body function (equation 2.2), which means that a small temperature change already has a lot of influence on the strength of the emitted spectrum.

The same analysis has been done for a range of temperatures (from 10 to 100 K) and redshifts (from 0 to 6), and has been plotted in figure 3.7. In the left plot, the temperature was fixed at 40 K, and in the right plot, the redshift was set at a fixed value of 3. The median has been plotted as a line, and the shaded region around it represents the offset between the 16th and 84th percentile of the 4096 samples. If it were a perfect Gaussian distribution, these values would translate to the mean and the region between one standard deviation offset from the mean.
It is seen that Direct Subtraction is mostly centered around 0% offset. SPLITTER seems to decrease its offset for measurements at high redshifts and low temperatures, and the standard deviations in this offset have the

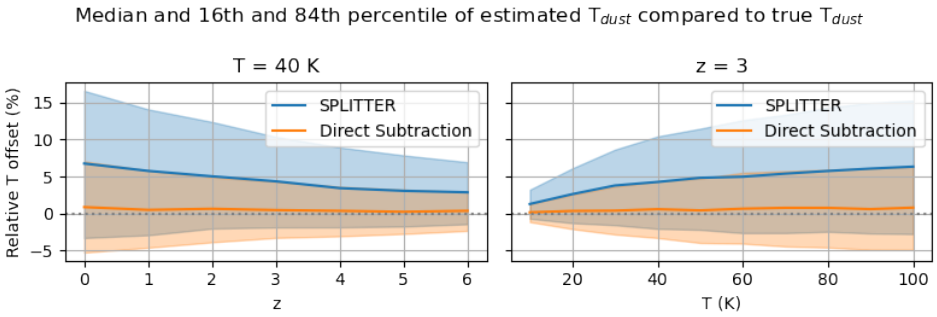

Figure 3.7: Relative offsets in estimated dust temperatures when using SPLITTER (blue) and Direct Subtraction (orange). Left: redshift is varied from 0 to 7 with increments of 1, true dust temperature is set at 40 K. Right: true dust temperature is varied from 10 to 100 with increments of 10, and redshift is set at 3. Plots show the systematic overestimation in SPLITTER, and an increase in this overestimation and standard deviation of both methods at low redshifts and high temperatures.

same behavior for both techniques.
The reason why both the spread and the mean offset for SPLITTER (and if observed closely, also the mean offset of Direct Subtraction) are lower for these values of dust temperature and redshift, is because in these regions, more of the shape of the spectrum is captured by the instrument. To explain this, figure 3.5 will be helpful again. For galaxies at higher redshifts, the observed rest frequency will be shifted towards a higher frequency, and a larger part of the shape will be captured, making it easier for the curve fitting tool to estimate the true spectrum. The same thing happens with a decrease in dust temperature, as the peak of the curve will move towards a lower frequency according to Wien's displacement law ($\lambda_{peak} T = b$) [8], and a larger portion of the shape is captured by the device's bandwidth.
To get a feeling of the full dependence of the offset in the estimated T_{dust} on the true dust temperature and spectral emissivity index, heatmaps have been made, showing the mean offset and its standard deviation, along with the error that is given by the curve fitting function. This error is what is calculated as the error in your estimation when accounting for the noise when fitting a continuum through the retrieved data. It will not account for a systematic offset in this fitted continuum, as this offset cannot be known without knowing the true spectrum.

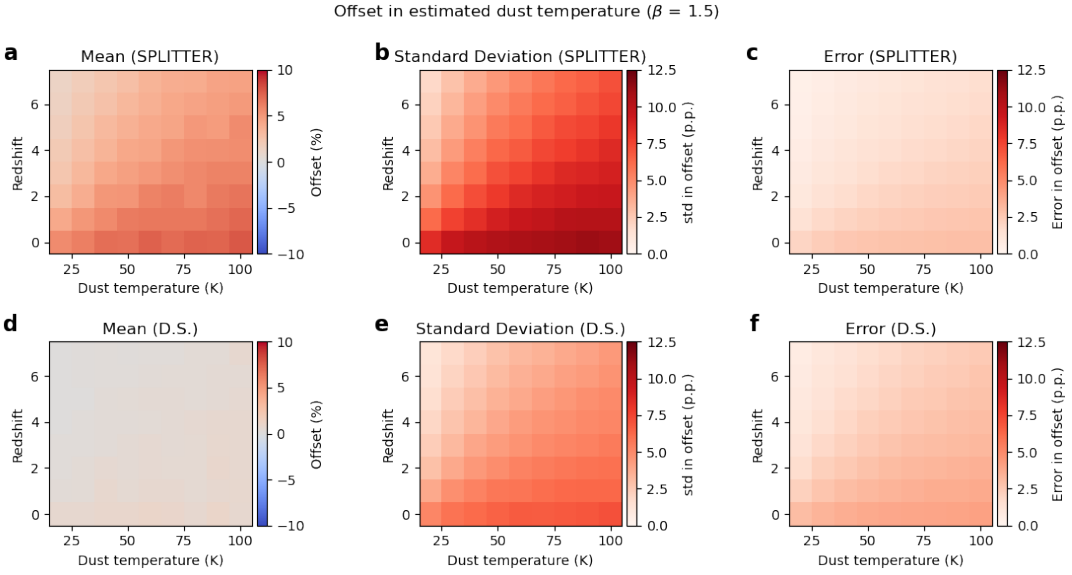

Figure 3.8: Heatmaps with true dust temperature on the x-axis and redshift on the y-axis. a,d: mean relative offset. Direct Subtraction shows no systematic offset, as expected from the math, whereas SPLITTER has a mean offset increasing with temperature and decreasing with redshift. b,e: Standard deviation in the relative offset. SPLITTER shows a larger spread than Direct Subtraction, and both show an increase with higher temperature and lower redshift. c,f: Error in estimated dust temperature as given by the curve fitting function. This only takes the standard deviation of the noise as input as this is the only parameter known when doing observations. Therefore, it shows too low of an error as the overestimation is not accounted for.

Figure 3.8 shows these heatmaps. Immediately visible in the first column is the mean close to 0 in offset for Direct Subtraction. SPLITTER shows more relative offset at high temperatures and lower redshifts, and the standard deviations of both are also larger in those regions, as explained before. A notable difference in the second column is the higher standard deviation in offset that SPLITTER has. Therefore, besides having an average overestimation, there is also more spread in this, being so high to allow slight underestimations as well. In the third column, it is seen that the error as a result of the curve fit tool is much lower than the mean and standard deviation of the offset, being even lower for SPLITTER than for Direct Subtraction because SPLITTER shows less noisy jitter in the spectrum. As a result, it would seem as if SPLITTER is more accurately giving the estimate of the continuum, whereas this estimate might have a large offset of which the amount is unknown when doing a single observation.

Until now, β has been ignored as a free variable and was forced to a fixed value of 1.5. However, when observing a galaxy, this value is unknown too. Figure 3.9 shows what the parameter does to the spectrum. This effect is different from the dust temperature, so they should be able to be measured at the same time. this is discussed in the next section.

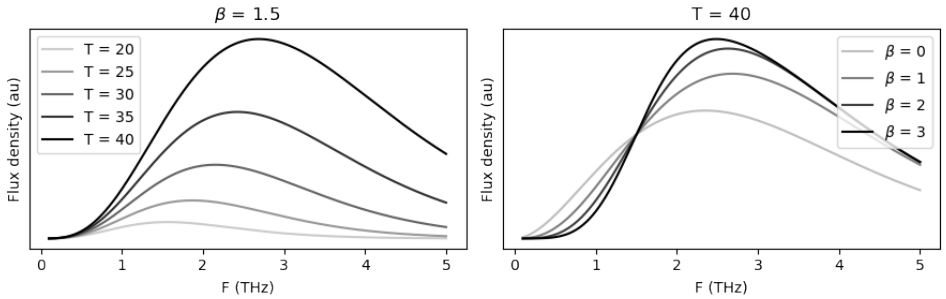

Figure 3.9: Two plots showing the influence of physical dust temperature and spectral emissivity index to the emitted spectrum as described by the modified black body function (equation 2.2). The frequency on the x-axis is rest frequency and the frequency at redshift $z = 0$. An increase in dust temperature greatly increases the flux density of the emission and moves the frequency at which the emission peaks (loosely described by Wien's displacement law ($\lambda_{peak} T = b$) [8]. β has the most effect on the shape of the spectrum, applying a frequency-dependent scaling factor around the frequency for which the optical depth equals unity, in this thesis assumed to be 1.5 THz based on existing observations [3].

3.2.2. β as a Free Parameter

The same analysis has been repeated, but this time with the spectral emissivity index β as a free parameter. Unlike dust temperature, which almost simply scales the spectrum, the parameter β mostly changes its shape. As seen in figure 3.9, a higher spectral emissivity index means a steeper curvature, especially in the region at lower frequencies than the peak. β doesn't seem to have an effect at the frequency of 1.5 THz. This is the assumed frequency for which the optical depth of the dusty galaxy has a value of one. Note that this is yet another assumption based on existing measurements [3]. As the process is similar to when β was fixed, the random seed has been set the same as before in order to compare the differences.

Just as previously done, we first look at the histograms showing the relative offset when using fixed values for relevant parameters. Again we set dust temperature to 40 K, measure at a redshift of 3, and beta is set to 1.5 just like all previous runs. The resulting 4096 offsets for estimated dust temperature and spectral emissivity index have been plotted in figure 3.10. With a mean and standard deviation of (4.0 ± 7.6) %, SPLITTER seems to have a little more spread in temperature offset as before, when β was a fixed parameter. Direct Subtraction has a slight change in its mean and also shows more spread with an offset of (-0.4 ± 4.9) %. Now that β is a free parameter, the fitting tool has more trouble estimating dust temperature correctly. This makes sense, as the algorithm now has to search for two optimal parameters, while previously it already knew the exact value for β.

A noteworthy observation seen from the bottom histograms of figure 3.10, is that SPLITTER is better in estimating the spectral emissivity index than Direct Subtraction, as the latter seems to underestimate the value for β systematically. With an offset of (-1.7 ± 17.5) % when using SPLITTER, compared to (-9.5 ± 24.0) % with Direct Subtraction, it is not only more accurate but also more precise.

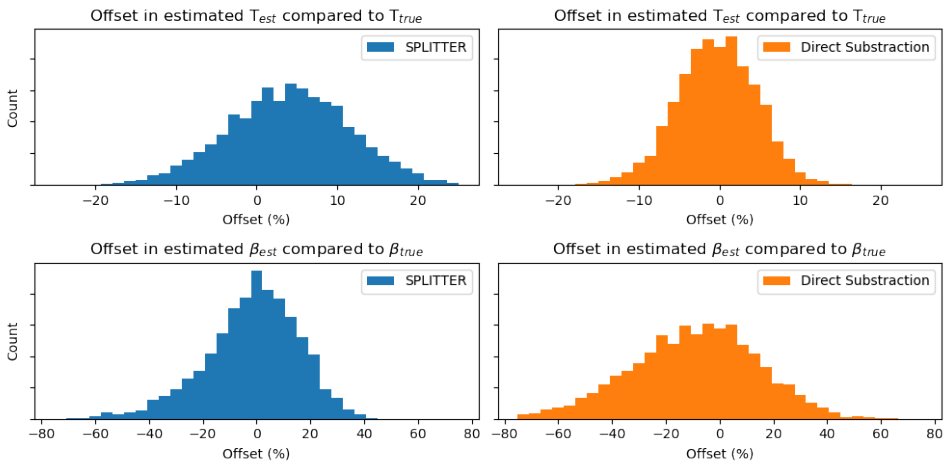

Figure 3.10: Histograms showing the relative offsets in estimated dust temperature and spectral emissivity index β when using SPLITTER (blue) or Direct Subtraction (orange). Values are calculated for a galaxy that emits a modified black body spectrum with a dust temperature of 40 K, a spectral emissivity index of 1.5, and a redshift of 3. SPLITTER shows an offset in the estimated temperature of (4.0 ± 7.6) % and Direct Subtraction of (-0.4 ± 4.9) %, showing an increase in standard deviation compared to the case when β was a fixed variable. Offsets in estimated β are (-1.7 ± 17.5) % when using SPLITTER and (-9.5 ± 24.0) % when using Direct Subtraction. Here, SPLITTER shows the more accurate estimate.

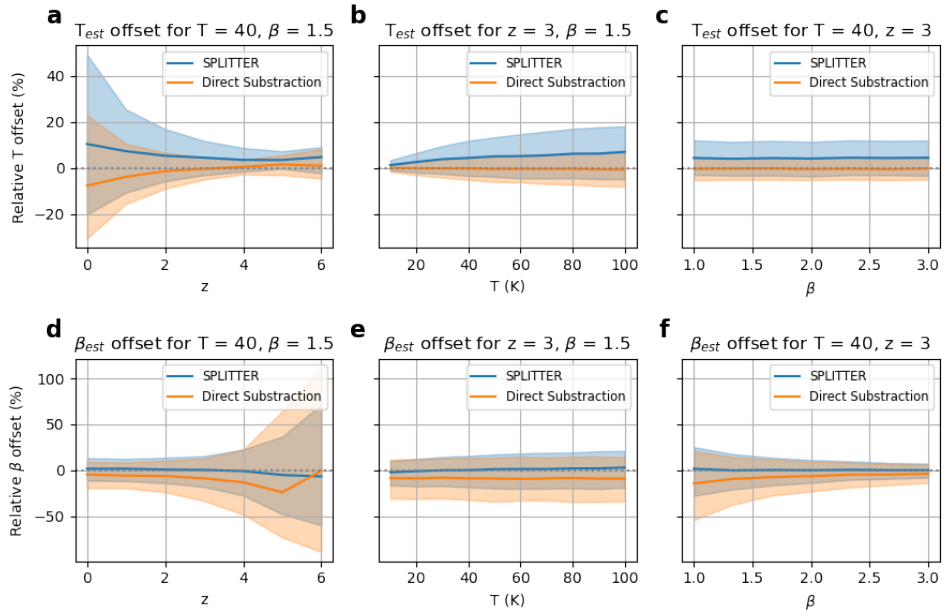

Figure 3.11: Relative offsets in estimated dust temperatures and β when using SPLITTER (blue) and Direct Subtraction (orange), when varying one parameter. Left: redshift is varied from 0 to 6 with step size 1. Middle: dust temperature is varied from 10 to 100 with step size 10. Right: β is varied from 1 to 3 with step size $\frac{1}{3}$. Plots show better estimation for T_{dust} when using Direct Subtraction, but better estimation for β when using SPLITTER.

Plotting these estimates as a function of redshift, dust temperature, and β results in figure 3.11, where again the median has been plotted with a shaded area between the 16th and 84th percentile (corresponding to one standard deviation offset if it were perfect Gaussian distributions).

Figure 3.11a shows that with β as a free parameter, at low redshifts the dust temperature estimate is much worse than before. This could be explained because of the extra freedom now that β is a free parameter too. A small error in the estimate for this parameter causes a relatively high offset in flux density for observations at low frequency as seen in figure 3.9 (0 redshift means an observational bandwidth from 220 to 440 GHz).
In figure 3.11b, we see almost the same temperature dependence for offset in dust temperature as before, with a slightly larger standard deviation, and figure 3.11c shows that the offset in the estimated dust temperature does not depend on the value of β. This builds on the hypothesis that the performance of dust temperature estimate is based on how much of the shape of the spectrum is observed, as a change in beta doesn't allow more or less coverage of the shape, unlike dust temperature and redshift of the galaxy do.
The plot in figure 3.11d shows severe error in estimated β for measurements at high redshifts. This can be explained with the nature of the two parameters. Figure 3.9 shows that a change of β mostly influences the slope at low frequencies, but at high frequencies, the dominant effect that a change of β causes is an offset in strength. Therefore, for higher redshift measurements, when only the high frequencies are observed, the offset has much more effect on the estimate of β than at low frequencies.
Slight temperature dependence of the estimate for β is visible in figure 3.11e, but as a different dust temperature won't drastically change the shape of the spectrum, this effect is almost zero. The plot in figure 3.11f shows that the estimated β is more precise for higher values of its true value. This might seem counterintuitive at first, as figure 3.9 shows that a small change in β has far more effect on the shape of the emitted spectrum at low values compared to higher ones. This would mean that this parameter should be more accurately estimated for lower values of it. However, an error in the estimate is mostly caused by the random noise of the points, as this influences the shape of the spectrum far more than a systematic offset factor that is the same among each point. For high values of β, the slope is large, and the impact of this noise becomes smaller as the observed points spread out more on the y axis, making it easier for the fitting algorithm to find the right fit.

To fully investigate the effect of all parameters, several heatmaps have again been plotted. As they can only map two free parameters, the redshift has been set at a fixed value of z = 3.

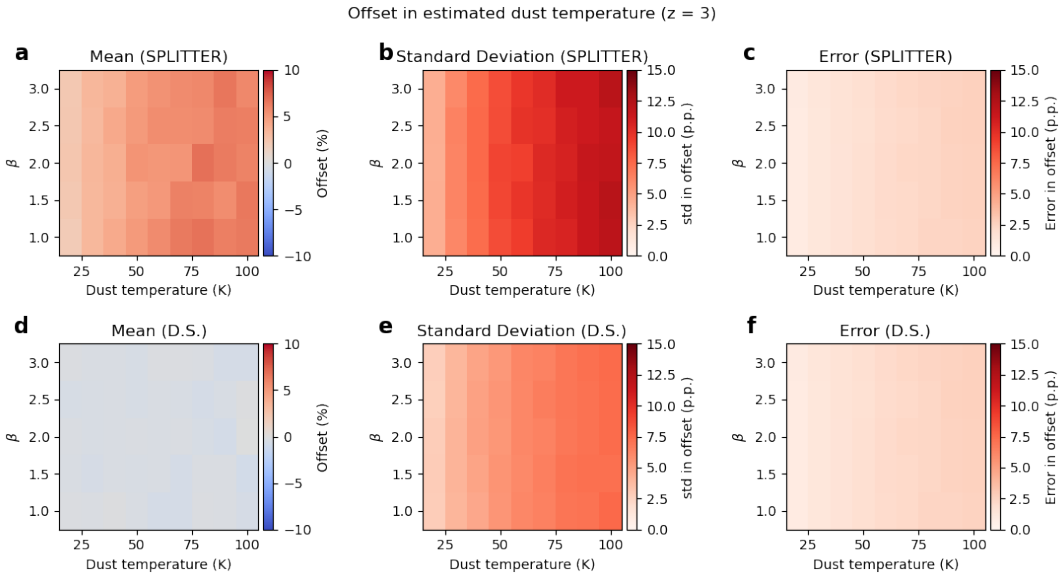

Figure 3.12: Heatmaps with true dust temperature on the x*axis and β on the y-axis, for z = 3. Left: mean relative offset of T_{dust}. Middle: standard deviation in this relative offset. Right: relative error in estimated dust temperature as returned by the fitting function. Unit p.p. means percentage point, the arithmetic difference between two percentages. Direct Subtraction shows a more accurate and precise estimate for T_{dust}, and both methods seem to be independent of the value for β, therefore the shape of the spectrum does not influence the estimate of the temperature.

Immediately visible in figure 3.12, showing the offset in dust temperature estimate, is the mean offset of 0 when using Direct Subtraction, and a systematic overestimate plus a higher standard deviation for SPLITTER. When compared to the case of a fixed value for β, in figure 3.8, there isn't much difference to be seen. Again, the error given from the fitting function is drastically lower than the offset and its standard deviation, implying yet again that a result will behave as if it were accurate, while it isn't. The independence of β for temperature estimate is

seen from the equally colored vertical lines.

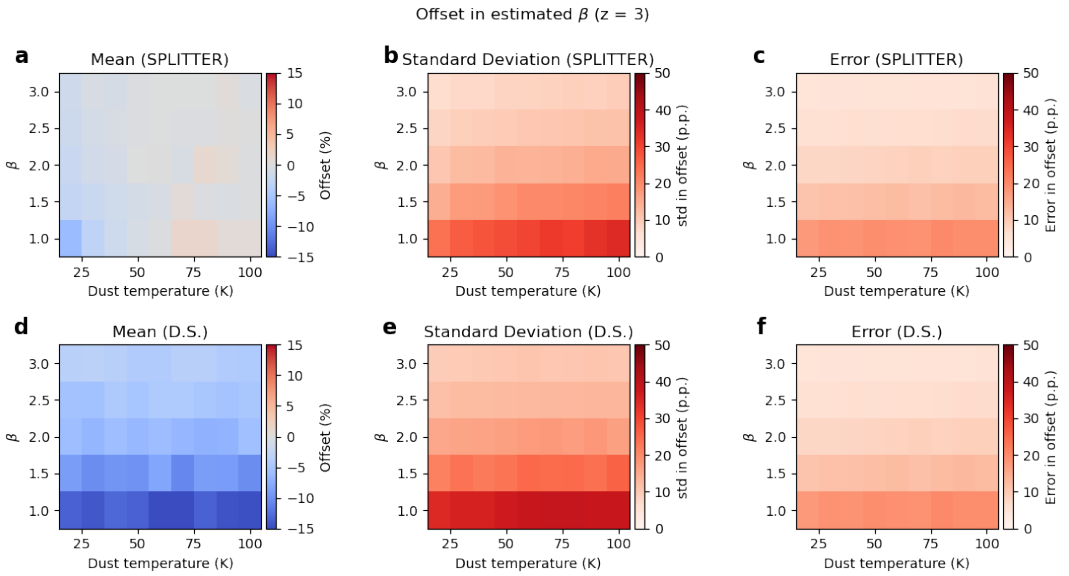

Figure 3.13: Heatmaps with true dust temperature on x axis and β on y axis, for z = 3. Left: mean relative offset of β. Middle: standard deviation in this relative offset. Right: relative error in estimated value for β as returned by the fitting function. Unit p.p. is percentage point, the arithmetic difference between two percentages. SPLITTER shows a more accurate and precise estimate for β, and both methods seem to be independent of the value for T_{dust}, implying that the strength of the spectrum does not influence the estimate of β.

Figure 3.13 displays the mean and standard deviation in the estimate of β, along with the error returned by the fitting algorithm. Clearly standing out in parts a and d of this figure is the fact that SPLITTER returns a better estimate for the spectral emissivity index than Direct Subtraction, which systematically seems to underestimate this value. Parts b and e also show that SPLITTER is a bit more precise, as the standard deviation is slightly lower. The error returned by the fitting function in c and f seems to be roughly the same for both methods, and overall the influence of the dust temperature seems to be close to zero, as indicated by the approximately equally colored horizontal lines.

The fact that SPLITTER seems to perform better at estimating the β index and worse at the dust temperature can be explained by the behavior of the two, shown in figure 3.9. As seen in the previous section, SPLITTER shows a reduction in noise, therefore being more precise, but this is at the cost of a systematic overestimation, making it less accurate. This accuracy is what determines the mean offset in dust temperature, as this dust temperature changes mostly the strength of the emitted spectrum. Therefore, systematically overestimating the continuum translates to overestimating the dust temperature. However, the β factor, determined by the shape of the dust, changes the shape of the spectrum rather than the strength. As SPLITTER returns a less noisy continuum and the overestimation is assumed to be fully correlated between each frequency point, this shape is more precisely determined and β will be estimated better than when using Direct Subtraction.

Therefore, based on the results gathered from these tests, SPLITTER seems to more precisely denoise the data, at the cost of adding a systematic overestimation of the continuum, making it less accurate. As a result, in order to get a better estimate of the strength of the continuum, and therefore a better estimate of the physical temperature of the dust, it would seem more beneficial to use the method of Direct Subtraction. However, when the shape of the continuum, and therefore the value of the spectral emissivity index needs to be calculated, it is preferred to use the SPLITTER algorithm, as it is better at eliminating the noise from the true spectrum. For higher redshift, Temperature is estimated better, while for lower redshift, the spectral emissivity can be estimated more precisely.

4

Emission Line Detection

Emission lines are tools to obtain vital information on a galaxy. For dusty galaxies in the far-infrared region, the emission line of [CII] is usually the brightest and is often used for determining the redshift of these galaxies. Another target for DESHIMA 2.0 are rotational lines of CO. It is therefore important that the incoming flux of this line is retrieved with good accuracy.

4.1. Decisions on the Creation of Custom Galaxy Spectra

To test the performance of SPLITTER against the usual technique of Direct Subtraction on this retrieval of line flux, custom spectra have been created that simulate a galaxy with a continuum and a single narrow emission line. For convenience, these lines have all been placed at exactly the center frequency of four selected channels of the DESHIMA 2.0. The center frequencies of the four selected channels were 229.886, 299.861, 340.084, and 410.348 GHz, hereafter mentioned as the frequency rounded to the nearest GHz.

Rotations of a galaxy cause a slight redshift of the line due to the Doppler effect, resulting in the emission line being spread out over a small bandwidth of frequencies. This is often modeled with a Gaussian, where the integrated area represents the total flux of the emission line. Usual Full Width at Half Maximum (FWHM) of such a Gaussian is around 600 km/s (conversion to bandwidth via $\Delta\nu = \nu 600/(3 \cdot 10^5)$). To reduce leakage to neighboring frequency channels, the FWHM of the Gaussian was set to 200 km/s instead. It was chosen not to be lower than this as the resolution of the custom spectrum would need to become too high, causing TiEMPO simulations to take even longer.

The total flux of such a simulated emission line was determined by the received level of noise in its frequency channel. This noise level was obtained by looking at the off-source samples of an existing TiEMPO simulation. This timestream of samples only contained noise from the atmosphere, as the galaxy was artificially turned off to obtain this data. As a result of time-correlated TLS noise and slowly changing atmospheric effects, a moving average in the offset was present, as seen in figure 4.1a. Therefore, to get the level of the Gaussian photon noise, the timestream was split into many small chunks (just like how SPLITTER eliminates this effect) from which the standard deviation was calculated. As seen in figure 4.1b, the mean can be approximated to be constant when taking chunks in the order of 6 seconds. The total standard deviation of the noise is calculated by taking the root mean square of the standard deviation of each chunk. A mathematical explanation of why this was done can be found in appendix A.2. Figure 4.1c shows that this value indeed goes down as smaller and more chunks are selected for this calculation. When having a known on-source observation time, the standard error caused by the noise can be calculated simply by dividing the standard deviation by the number of samples taken ($\sigma_{st} = \sigma/\sqrt{N}$). For 15 minutes on-source at a sampling rate of 160 Hz, this would mean N = 144000 samples.

The total noise power received by a specific frequency bin was calculated as the standard error in flux density of the photon noise σ_{st}, multiplied by the bandwidth of the channel. This value would be multiplied by an x amount to calculate the total flux of the emission line to achieve a signal-to-noise ratio of x inside that channel. As the channels of the DESHIMA are not perfect square bandpass filters but modeled by Lorentzians, part

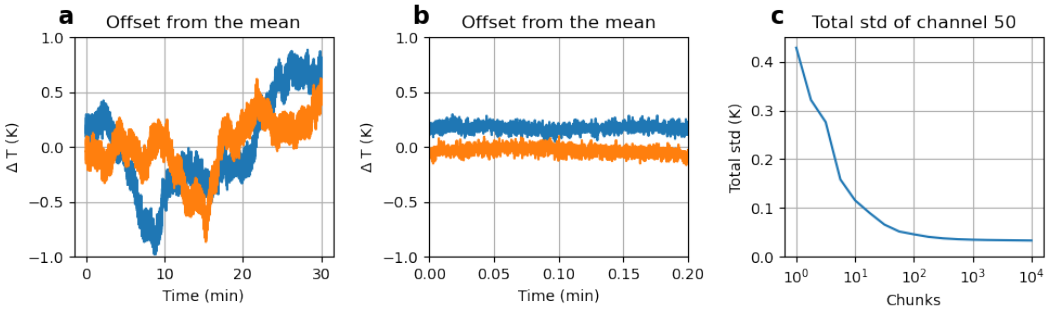

Figure 4.1: a: Difference from the mean of the brightness temperature of two neighboring channels (channel 50 and 51). The whole time span of 30 minutes has been shown. b: chunk of (a) showing only 12 seconds of observation time. c: total standard deviation plotted against the number of timestream chunks. This value is calculated by the root mean square of the standard deviation of a single chunk, as explained in appendix A.2. The value goes down as the time-correlated TLS noise is dampened the more chunks are taken.

of the signal from the narrow peak still leaks into neighboring bins. This is unavoidable, even if the emission line was a perfect delta function. The fraction of signal ending up in the correct bin was applied as a correction factor to the total flux of the line. This makes sure that the total flux received in the center channel is indeed x times stronger than the noise. The correction factor used had a value of 0.60. Instead of using this correction factor, the analysis could also be done by determining the received flux on the three or five channels around the emission line, but it has been chosen to use the correction factor instead and make sure that the SNR as received by the central channel would be x.

To simulate a pseudo-realistic galaxy, a continuum linear between 220 GHz, 0.25 mK and 440 GHz, 0.75 mK has been added to each spectrum (the same as used for continuum estimations, as these simulations were used for that as well). Note that this does not correspond to a linear continuum in flux density as the relation between flux density and temperature is not linear. The resulting spectrum with a line of x SNR at a frequency of 230, 300, 340, or 410 GHz was then run through TiEMPO with an observation time of 30 minutes. Different SNRs in the range from 3 to 15 have been used for each line. Each combination of line frequency and SNR has been run through TiEMPO 3 times to simulate three observations with a different atmosphere model. To calculate the DESHIMA response when there is no noise and attenuation present, each of the 347 channel's response was modeled by a Lorentzian with a FWHM equal to the spacing between the bins. As this spacing is a constant value in logspace, the width of each bin will be larger for higher frequencies. Because of this non-constant width, it has been chosen to calculate the response for each bin separately, instead of convolving a single Lorentzian over the entire spectrum.

In total, 69 simulations have been analyzed for this test. The TiEMPO simulated data has been denoised by SPLITTER using input parameters as noted in table 4.1. Note that the parameter Θ, which determines the amount of emission lines SPLITTER is searching for, has been set to 3. There is a small chance that there is a bright noise peak present, and setting Θ to 3 prevents SPLITTER from ignoring the real emission line when this happens.

Threshold on η_{atm}	ρ	n	Θ	Iterations	t_l
20%	$5 \cdot 10^3$	10	3	60	1600

Table 4.1: Parameters used to run SPLITTER. Data at frequencies below the threshold η_{atm} is ignored in the algorithm. ρ describes how large the estimate is adjusted each iteration, n describes the amount of estimates for signal vector s are computed, Θ describes the amount of peaks the algorithm will be looking for (set to 3 to prevent dampening of the emission line in the presence of a higher noise peak), t_l is the amount of time samples in each chunk the big matrix is split in. At 60 iterations, the algorithm has already converged.

An example of the denoising done with both SPLITTER and Direct Subtraction can be seen in figure 4.2, where the 1st run of the line at 230 GHz with SNR 14 has been analyzed.

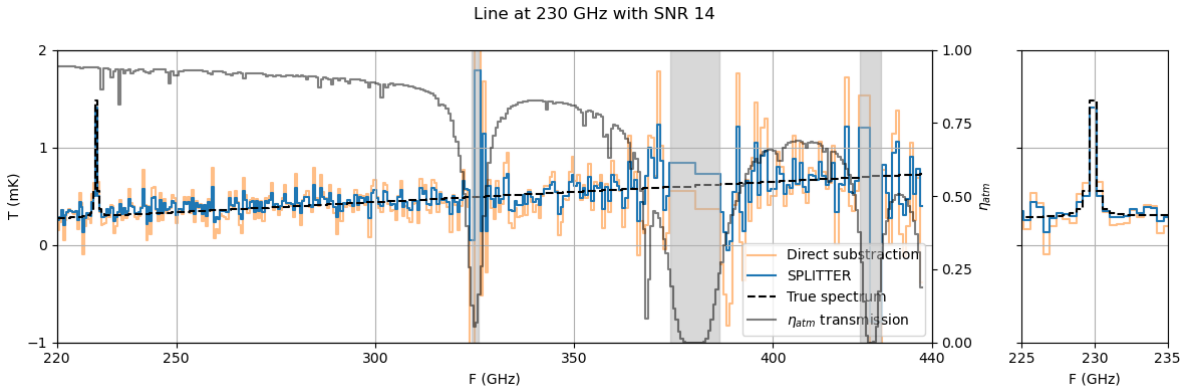

Figure 4.2: Example of using SPLITTER and Direct Subtraction on a TiEMPO simulation with a line at 230 GHz with a set SNR of 14. The zoom-in on the right shows that both methods slightly underestimate the brightness temperature of the line, which can be directly translated to a slight underestimation in flux. After subtraction of the average continuum for both methods, SPLITTER underestimates the flux by a 6.0 % offset and Direct Subtraction by only 2.8 %. Direct Subtraction is more accurate in this case, even though the channel has the same brightness temperature for both methods. Since Direct Subtraction underestimates the continuum a bit in this area, the underestimated line flux will be compensated slightly.

4.2. Performance on Recovered Emission Line Flux

First, the flux of the line was calculated by subtracting the average continuum of the neighboring 30 bins from the total flux detected by the channel. Note that the five central bins around the line were not used to calculate the continuum, as they contained leaked flux from the emission line. This value is then compared to the known true line flux. The fractions of recovered flux against true flux have been plotted in figure 4.3.

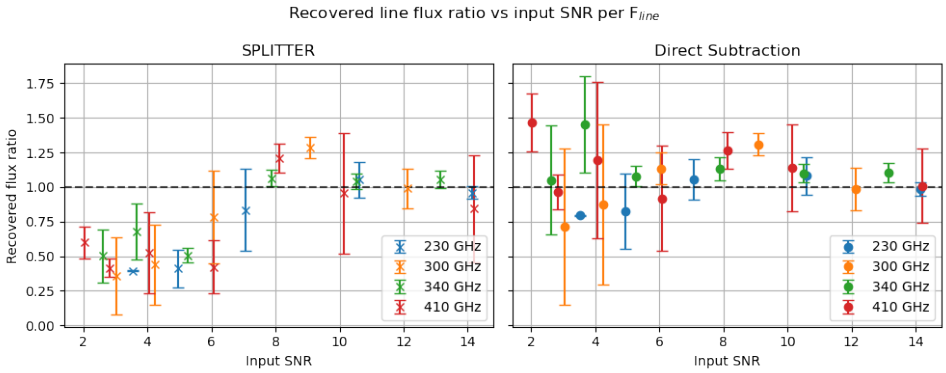

Figure 4.3: Recovered line flux for SPLITTER (left) and Direct Subtraction (right). The x-axis shows the SNR that the line flux was set at for the input spectrum in TiEMPO, based on the level of the photon noise in off-source measurements, and the different colors indicate the four different line positions in the spectrum. There is no visible difference in flux recovery for different line locations, but an obvious result is that SPLITTER tends to underestimate weak lines, whereas Direct Subtraction just gets a larger spread in this ratio. The underestimation for SPLITTER is most likely caused by the fact that the algorithm suppresses noise, and has seen the weak line as a noise peak.

It can be seen that for bright lines, in the region with photon SNR > 10, the total flux of the line is approximated very well, showing less than 20 % offset on each iteration. As the strength of the line gets lowered, and therefore the input SNR too, different behaviors are seen for both denoising techniques.
Direct Subtraction shows a larger standard deviation in its offset the closer the line SNR goes to 0. This is explained by the fact that the strength of the noise does not change (as this does not depend on the input spectrum), while the strength of the emission line is lowered. The noise will therefore have a greater effect on the fraction of recovered line flux, as it becomes relatively larger compared to the line strength.
SPLITTER shows interesting behavior for lower input SNR. Instead of showing an increased standard deviation around ratio 1, the recovered fraction systematically becomes lower. SPLITTER is designed to search for a set number of peaks each iteration, in this analysis set to three lines. As the emission line flux decreases, this line might fall more and more between the noise peaks, and therefore is not detected as a line, but rather suppressed as the algorithm thinks it is noise.

To investigate how much the difference between the two methods is, the recovered flux using SPLITTER has been divided by the flux recovered using Direct Subtraction, to make a relative difference factor. They have been plotted in figure 4.4.

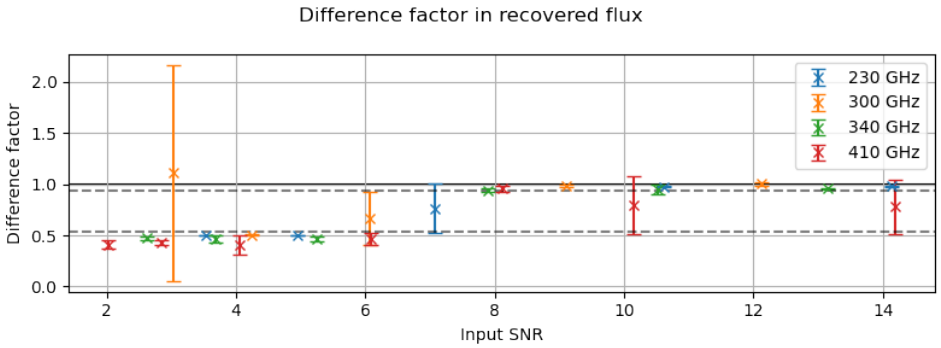

Figure 4.4: Recovered line flux using SPLITTER divided by recovered line flux using Direct Subtraction. This factor was 0.93 ± 0.07 for bright emission lines (SNR compared to photon noise > 7.5), and 0.46 ± 0.03 for weak lines (SNR < 6.5). Slight underestimation compared to Direct Subtraction is therefore seen for bright lines.

The difference seems to be systematically around a factor of 0.93 ± 0.07 for bright lines (SNR > 7.5 compared to photon noise) and 0.46 ± 0.03 for weak lines (SNR < 6.5). Note that the outlier of F_{line} = 300 and SNR ≈ 3 has been left out for this calculation. Therefore, there does seem to be a slight underestimation of the emission line flux for SPLITTER, even for bright lines.

Both methods seem to start a change in behavior (SPLITTER underestimating the flux and Direct Subtraction gaining a higher standard deviation) around an SNR lower than 7. It is interesting to see if the actual emission line is still being detected at these small SNRs. In the next section, the recovered signal-to-noise ratio as compared to the noise in neighboring channels is discussed.

4.3. Recovered Signal to Noise Ratio of Lines

The recovered SNR was calculated for each simulation, both for SPLITTER and Direct Subtraction. This value should not be confused with the input SNR used to determine the total flux of the line, as that one is based on the raw photon noise and the other is the SNR that the line has when compared to the noisy output from both denoising methods. The recovered line flux as calculated previously has been used again. This line flux was then compared to the standard deviation of the recovered signal in the same channels around the line.

This recovered SNR has been plotted against the input SNR in figure 4.5, where the average and standard deviation for lines at different frequencies have different colors. A standard often used is that lines with an SNR > 5 are taken as well detected, whereas lines with 3 < SNR < 5 are so-called 'tentative' detections, meaning that it is unsure if it is a high noise peak or a real emission line. Anything with SNR < 3 is said to be noise. The black line shows the minimum for good detection and the red line the minimum for tentative detection. It can be observed from both plots that SPLITTER is definitely recovering a higher SNR than Direct Subtraction, which can be explained due to the higher precision of SPLITTER, as the noise level is lower. This lower noise level also allows for fainter lines to be detected, as the threshold of line intensity for achieving > 5 SNR is lower. It can be seen in the figure that lines between an SNR of 6 to 8 are only tentative detections for Direct Subtraction, but become good detections for SPLITTER. Noticeable is that the line at 410 GHz often gets a much lower SNR than the other lines. A big difference is that at this frequency, the atmospheric transmittance is much lower than for the other three lines, and this is probably why we see this difference here.

Overall, there seems to be much improvement in the retrieved signal-to-noise ratio, and before looking at the factor of this improvement, first, the improvement factor in noise is shown in figure 4.6. This factor was calculated by determining the ratio between the standard deviations of the 30 channels around the central

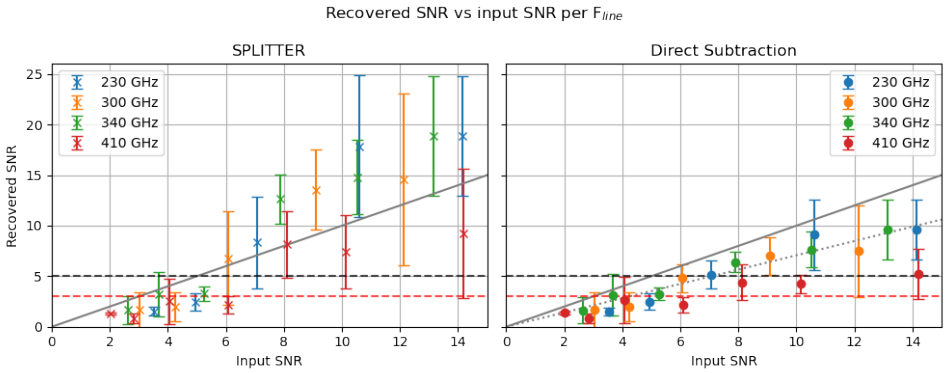

Figure 4.5: Recovered Signal to Noise Ratio for SPLITTER (left) and Direct Subtraction (right). This recovered SNR has been calculated by comparing the line flux to the standard deviation in the flux density of the neighboring 30 bins (excluding the 5 bins around the line to reduce the influence of leakage from the emission line). The red dotted line indicates the barrier from no detection to tentative detection, and the black dotted line indicates the barrier from tentative detection to good detection. Clearly visible is that SPLITTER recovers a greater SNR than Direct Subtraction, and can change a tentative detection from Direct Subtraction to well detected. Below an input SNR of 6, both methods seem to be unable to detect the line. The grey line is the y=x line, showing the expected recovered SNR based on the photon noise. SPLITTER seems to retrieve a higher SNR, implying that it reduced photon noise by a bit. Direct Subtraction shows a linear increase around the line y=x/$\sqrt{2}$, showing that this increased factor of $\sqrt{2}$ on the photon noise as a result of subtracting two noisy signals is indeed present.

emission line channel, excluding the five closest channels to the line.

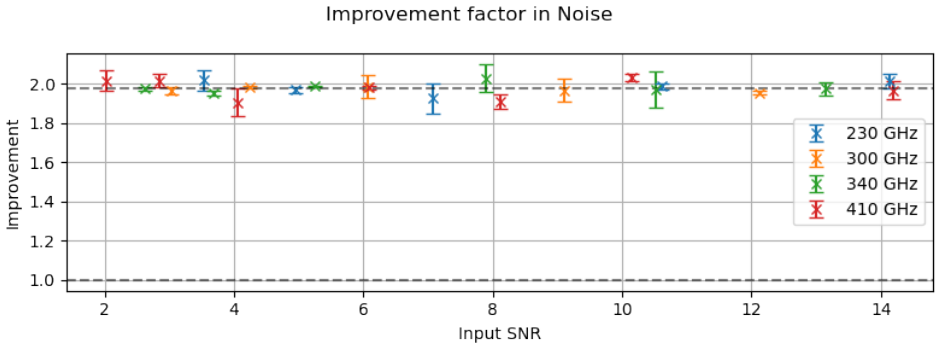

Figure 4.6: Standard deviation of noise from Direct Subtraction divided by the standard deviation of noise from SPLITTER of the 30 neighboring channels of the emission line. A systematic improvement of 1.97 ± 0.03 is visible, independent of the SNR of the line compared to the photon noise. This is what we should expect, as the five channels around the line are excluded from this noise calculation and therefore line flux shouldn't make a difference in the noise standard deviation.

It is obvious that this value should not depend on the signal-to-noise ratio of the line, as the noise level is only determined by the neighboring channels around the line, excluding the five channels closest to the emission line. This improvement factor also seems to be independent of the frequency around which we calculate the noise, and has a value of 1.97 ± 0.03. Therefore, SPLITTER shows an improvement of almost twice as little noise. This factor can be combined with the difference factor in the retrieved signal from figure 4.4, to get the improvement factor in signal-to-noise ratio. This has been plotted in figure 4.7.
For bright lines (SNR > 7.5), this improvement factor is 1.9 ± 0.1 on average, while for weak lines (SNR < 6.5), the factor is just 1.0 ± 0.1. Note that these weak lines are not detected using both methods and are therefore just part of the noise.
If SPLITTER purely splits the atmosphere from the signal, it is expected to see an improvement factor of $\sqrt{2}$, as the only difference would be the absence of this factor due to the subtraction of noisy signals. Getting a value higher than $\sqrt{2}$, therefore, means that SPLITTER is also reducing part of the photon noise. This was also seen in Brackenhoff's improvement factor of 1.7 [5], based on the weighted root mean square error (equation 2.10). These results are very promising for SPLITTER as a denoising algorithm.

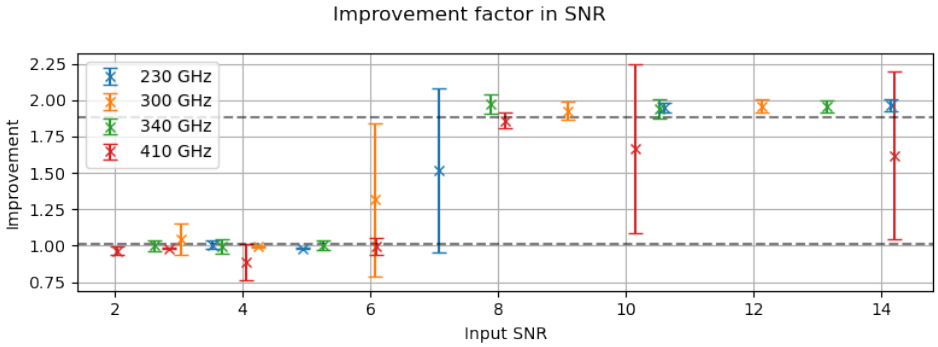

Figure 4.7: Total improvement factor in retrieved signal-to-noise ratio. The signal-to-noise ratio was calculated by comparing the total line flux in the center frequency channel to the standard deviation of the 30 channels around it, excluding the five channels closest to the central channel. Improvement factor is 1.9 ± 0.1 for bright lines (SNR > 7.5) and 1.0 ± 0.1 for weak lines (SNR < 6.5). The improvement factor higher than $\sqrt{2}$ implies that SPLITTER reduces part of the photon noise as well.

Reflecting on the results shown above, SPLITTER definitely shows improvement for detecting spectral emission lines. As the noise from SPLITTER is lower, the lines are detected with higher significance and tentative line detection using Direct Subtraction can become a robust detection after using SPLITTER. There doesn't seem to be a noticeable difference in recovered flux from bright emission lines, which means we do not over- or underestimate the flux from a line. Faint lines will, however, suffer from suppression in flux as SPLITTER sees it as a peak in noise. This only seems to happen for the very faint lines that both methods cannot even detect. Lines in between these regions (6 < input SNR < 8) do seem to show a little underestimation in flux when using SPLITTER, so this possible underestimation should be kept in mind when observing such a barely detected line. Overall, SPLITTER shows an improvement factor in retrieved SNR of 1.9 ± 0.1, and since this is above $\sqrt{2}$, this implies that the algorithm also reduces photon noise a little.

4.4. Correlation between Continuum and Emission Line Errors

It would be interesting to see if the overestimation of the continuum influences the estimation in line flux. To investigate this, a scatterplot has been plotted in figure 4.8 comparing continuum offset versus line flux offset. The Pearson correlation coefficient has been calculated for both methods. This is a value from -1 to 1 that determines the correlation between the points. This coefficient was 0.12 for SPLITTER and -0.10 for Direct Subtraction. As they are both very close to 0, we consider continuum and line estimates to be uncorrelated. In this plot, it is again visible that SPLITTER overestimates the continuum systematically, as most points are to the right of the 1.0 factor in the continuum ratio.

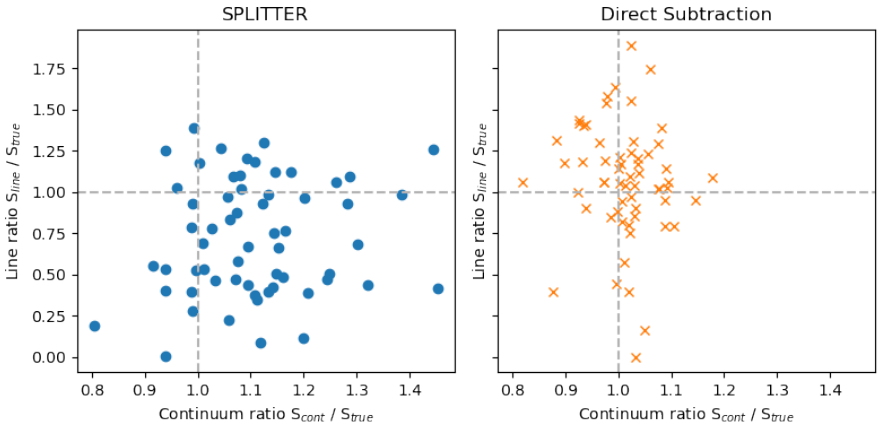

Figure 4.8: Scatterplot showing the relation between continuum estimation and line flux estimation for data at all line frequencies and SNRs. Pearson correlation coefficient is 0.12 for SPLITTER and -0.10 for Direct Subtraction, and as they are close to 0 it can be said that there is no correlation between the two. This implies that an overestimation in the continuum does not affect the recovered flux of the line.

5

Conclusions and Future Work

5.1. Conclusions

This project focused on comparing the performance of the SPLITTER algorithm to the technique of Direct Subtraction when used for noise reduction on simulated DESHIMA 2.0 data. The simulation package TiEMPO was used to create realistic output data of the DESHIMA 2.0, after which SPLITTER and Direct Subtraction were used separately on the noisy data. Performance on two key features of the spectrum has been examined, being the continuum emission and the estimation of the physical parameters dust temperature and spectral emissivity behind it, along with the ability to detect spectral emission lines and the amount of flux recovered from them. The experiments yield the following answers to the research questions:

> To what extent can we estimate the shape of the continuum created by the black body radiation of the dust clouds, and how well will the physical parameters dust temperature and spectral emissivity be estimated?

The continuum has been estimated for a total of 88 simulations of spectra by averaging the spectrum over four frequency spans. Comparing this average to the true spectrum showed a systematic offset of 8.5 % with a standard deviation of 13.5 p.p. (percentage points). Direct Subtraction seemed to estimate the continuum with an offset of 0.7 % on average, with a standard deviation of 7.2 p.p.. Estimates at higher frequencies show a larger standard deviation, most likely due to the lower atmospheric transmittance, causing higher noise in these regions. Calculations of the relative Root Mean Square Error in the same frequency spans show lower errors for SPLITTER, with a median error of 23.8 % compared to Direct Subtraction with 41.0 %. This shows that SPLITTER is more precise than Direct Subtraction, as it reduces noise, but it is less accurate, as the average value shows a systematic overestimation.

It has been assumed that this relative offset behaves with the same Gaussian distribution for each galaxy, and many estimations have quickly been modeled by using this assumed behavior on custom modified black body spectra with varied dust temperatures for galaxies at different redshifts. Consequences are that physical dust temperature is estimated with a mean offset and standard deviation of (4.1 ± 6.2) % for SPLITTER and (0.4 ± 3.9) % for Direct Subtraction on a galaxy with $T_{dust} = 40$ at $z = 3$. For lower redshifts and higher temperatures, both the mean and standard deviation of this offset increase. When taking the spectral emissivity, β, as a free parameter that can also be estimated, this introduced a slight increase in the standard deviation of the offset in temperature estimate, but more importantly, the estimate of this β seemed to be both more accurate and precise when using SPLITTER. With a mean and standard deviation in offset of (-1.7 ± 17.5) % compared to (-9.5 ± 24.0) % for Direct Subtraction when using $T_{dust} = 40$, $z = 3$, and the true value of $\beta = 1.5$, it proved better to use SPLITTER to estimate this parameter. Therefore, as SPLITTER is more precise, it can better estimate the shape of the continuum and therefore the parameter β, but as it is less accurate, it causes an overestimation in the temperature of the dust. So both noise removal techniques show better performance on only one of the two parameters, and for optimal estimation of both of them, it is best to use SPLITTER for spectral emissivity and Direct Subtraction for dust temperature.

> How much better is SPLITTER in detecting spectral emission lines, and how well does it estimate the flux of such a line?

Many simulations of custom spectra have been done with a weak continuum and a single emission line at a certain center frequency of a channel and with a certain input signal-to-noise ratio compared to the photon noise of that channel averaged over time. For bright emission lines, where the signal is about 8 times larger than photon noise, there seems to be little difference between SPLITTER and Direct Subtraction on the retrieval of the total line flux. Both seem to estimate the flux of the line pretty well, with limited offsets due to noise. As for weaker emission lines, Direct Subtraction has the behavior of having a larger standard deviation in the estimated total flux, which makes sense, as the noise is relatively larger compared to the line flux. SPLITTER seems to systematically underestimate the line flux in this case, which is most likely because it has the habit of suppressing anything it thinks is noise, and as the line comes closer and closer to the noise level, SPLITTER is more likely to assume it is noise.

When the retrieved flux of the line channel is compared to the standard deviation of the neighboring bins, the retrieved signal-to-noise ratio is calculated. This was done for both methods, and after comparison, SPLITTER showed an average improvement factor of 1.9 ± 0.1 for bright lines (SNR compared to photon noise > 7.5), but for weaker lines (SNR < 6) this factor got down to 1.0 ± 0.1, meaning that there is no improvement in retrieved SNR of the line. For the weak lines, there was also no detection for both methods, as the recovered SNR was too low. The intermediate region around SNR of 7 shows a highly variable improvement in retrieved SNR, along with a tendency to underestimate the total line flux. This possible underestimation should therefore be thought of when tentative detection of a line occurs.

To summarise, the SPLITTER algorithm shows much improvement in the reduction of noise, which makes it better at detecting spectral emission lines than the standard technique of Direct Subtraction. As this reduction in noise means it can more precisely estimate the shape of the continuum, it can get better estimates for the spectral emissivity β of the dust in a galaxy. However, the algorithm does add a systematic overestimation to the continuum, which propagates towards the estimation of the dust temperature. Therefore, Direct Subtraction remains a better option for estimating this dust temperature, as it is more accurate in estimating the continuum. This overestimation seems to have no correlation with the estimated flux of an emission line.

5.2. Future Work

For further analysis of the performance of the algorithm, it might be interesting to further examine the behavior of the algorithm on lines that the algorithm is just able to detect, with an SNR compared to the photon noise of 6 < SNR < 8, as this shows a transition region between being able to fully estimate line flux and reduce SNR, and being unable to detect the line.
Change in the estimated spectrum when an error in the value for T_{atm} is given is also interesting to research, as this value is not a constant during a long observation. SPLITTER might even be improved when instead of a constant value, the atmospheric temperature is made time dependent.

In order to fully optimize SPLITTER for reliable usage, it will be worth it to find out what causes the algorithm to systematically overestimate the continuum, and possibly fix the issue.
Most importantly, before the use of SPLITTER on DESHIMA 2.0 data, it has to be made compatible with the current filter response of the DESHIMA 2.0, as the filters are not perfect Lorentzians and not equally spaced in logspace as currently assumed by SPLITTER.

Bibliography

[1] Aste telescope. URL https://www.nao.ac.jp/en/research/telescope/aste.html.

[2] NASA, Mar 2019. URL https://solarsystem.nasa.gov/news/307/galileos-observations-of-the-moon-jupiter-venus-and-the-sun/.

[3] Caitlin M. Casey, Desika Narayanan, and Asantha Cooray. Dusty star-forming galaxies at high redshift. *Physics Reports*, 541(2):45–161, 2014. doi: 10.1016/j.physrep.2014.02.009.

[4] Akira Endo, Kenichi Karatsu, Yoichi Tamura, Tai Oshima, Akio Taniguchi, Tatsuya Takekoshi, Shin'ichiro Asayama, Tom J. Bakx, Sjoerd Bosma, Juan Bueno, and et al. First light demonstration of the integrated superconducting spectrometer. *Nature Astronomy*, 3(11):989–996, 2019. doi: 10.1038/s41550-019-0850-8.

[5] S. A. Brackenhoff. Splitter a data model and algorithm for detecting spectral lines and continuum emission of high redshift galaxies using deshima 2.0, 2021. URL http://repository.tudelft.nl/.

[6] Akio Taniguchi, Yoichi Tamura, Shiro Ikeda, Tatsuya Takekoshi, and Ryohei Kawabe. A data-scientific noise-removal method for efficient submillimeter spectroscopy with single-dish telescopes. *The Astronomical Journal*, 162(3):111, 2021. doi: 10.3847/1538-3881/ac11f7.

[7] Bruce T. Draine. *Interstellar Dust: Observed Properties*, page 235–247. Princeton University Press, 2011.

[8] Arnab Rai Choudhuri. *Astrophysics for physicists*. Cambridge University Press, 2010.

[9] C.L. Carilli and F. Walter. Cool gas in high-redshift galaxies. *Annual Review of Astronomy and Astrophysics*, 51(1):105–161, 2013. doi: 10.1146/annurev-astro-082812-140953.

[10] D. J. Hollenbach and A. G. Tielens. Photodissociation regions in the interstellar medium of galaxies. *Reviews of Modern Physics*, 71(1):173–230, 1999. doi: 10.1103/revmodphys.71.173.

[11] Caitlin M. Casey. Far-infrared spectral energy distribution fitting for galaxies near and far. *Monthly Notices of the Royal Astronomical Society*, 425(4):3094–3103, Sep 2012. doi: 10.1111/j.1365-2966.2012.21455.x.

[12] A. Kovács, A. Omont, A. Beelen, C. Lonsdale, M. Polletta, N. Fiolet, T. R. Greve, C. Borys, P. Cox, C. De Breuck, and et al. Far-infrared properties ofspitzer-selected luminous starbursts. *The Astrophysical Journal*, 717(1):29–39, 2010. doi: 10.1088/0004-637x/717/1/29.

[13] Gordon J. Stacey. Thz low resolution spectroscopy for astronomy. *IEEE Transactions on Terahertz Science and Technology*, 1(1):241–255, 2011. doi: 10.1109/tthz.2011.2159649.

[14] G. Lagache, M. Cousin, and M. Chatzikos. The [cii] 158 m line emission in high-redshift galaxies. *Astronomy amp; Astrophysics*, 609, 2018. doi: 10.1051/0004-6361/201732019.

[15] James J. Condon and Scott M. Ransom. *7: Spectral Lines*. Princeton University Press, 2016.

[16] Edward R. Harrison. *15. Redshifts*, page 302–322. Cambridge University Press, 2nd edition, 2000.

[17] Tom Bakx and Stefanie Anna Brackenhoff. deshima-dev/galspec: Update release, November 2020. URL https://doi.org/10.5281/zenodo.4279062.

[18] Erickson N., Narayanan G., Goeller R., and Grosslein R. An ultra-wideband receiver and spectrometer for 74–110 ghz. *Z-Machines to ALMA: (Sub)millimeter Spectroscopy of Galaxies*, 375:71–81, 2007.

[19] C. M. Bradford, J. E. Aguirre, R. Aikin, J. J. Bock, L. Earle, J. Glenn, H. Inami, P. R. Maloney, H. Matsuhara, B. J. Naylor, H. T. Nguyen, and J. Zmuidzinas. The warm molecular gas around the cloverleaf quasar. *The Astrophysical Journal*, 705(1):112–122, Oct 2009. doi: 10.1088/0004-637x/705/1/112. URL https://iopscience.iop.org/article/10.1088/0004-637X/705/1/112.

[20] C. Matt Bradford, Peter A. R. Ade, James E. Aguirre, James J. Bock, Mark Dragovan, Lionel Duband, Lieko Earle, Jason Glenn, Hideo Matsuhara, Bret J. Naylor, Hien T. Nguyen, Minhee Yun, and Jonas Zmuidzinas. Z-spec: a broadband millimeter-wave grating spectrometer: design, construction, and first cryogenic measurements. *Millimeter and Submillimeter Detectors for Astronomy II*, Oct 2004. doi: 10.1117/12.552182. URL https://www.spiedigitallibrary.org/conference-proceedings-of-spie/5498/1/Z-Spec-a-broadband-millimeter-wave-grating-spectrometer--design/10.1117/12.552182.full?SSO=1.

[21] Bruce Sibthorpe and Willem Jellema. Relative performance of dispersive and non-dispersive far-infrared spectrometer instrument architectures. *SPIE Proceedings*, 2014. doi: 10.1117/12.2056524.

[22] Ernest A. Michael and Felipe E. Besser. On the possibility of breaking the heterodyne detection quantum noise limit with cross-correlation. *IEEE Access*, 6:45299–45316, 2018. doi: 10.1109/access.2018.2855405.

[23] R.P. Feynman, R.B. Leighton, and M. Sands. *The Feynman Lectures on Physics, Vol. III: The New Millennium Edition: Quantum Mechanics*, chapter 21-5: Superconductivity.

[24] Jochem Baselmans. Kinetic inductance detectors. *Journal of Low Temperature Physics*, 167(3-4):292–304, 2012. doi: 10.1007/s10909-011-0448-8.

[25] Baselmans J Huiting R Karatsu K Llombart N Rybak M Takekoshi T Tamura Y Akamatsu H Brackenhoff S Bueno J Buijtendorp B Dabironezare S Doing A Fujii Y Fujita K Gouwerok M Hähnle S Ishida T Ishii S Kawabe R Kitayama T Kohno K Kouchi A Maekawa J Matsuda K Murugesan V Nakatsubo S Oshima T Laguna A Thoen D van der Werf P Yates S Endo A Taniguchi A, Bakx T. Deshima 2.0: development of an integrated superconducting spectrometer for science-grade astronomical observations. *Journal of Low Temperature Physics*, 2021.

[26] M. Rybak, T. Bakx, J. Baselmans, K. Karatsu, K. Kohno, T. Takekoshi, Y. Tamura, A. Taniguchi, P. van der Werf, A. Endo, and et al. Deshima 2.0: Rapid redshift surveys and multi-line spectroscopy of dusty galaxies. *Journal of Low Temperature Physics*, 2022. doi: 10.1007/s10909-022-02730-y.

[27] J. Tiebosch. Photon statistics, Jan 2022. URL https://joristiebosch.github.io/thesis/photon-statistics.

[28] Esmee Huijten, Yannick Roelvink, Stefanie A. Brackenhoff, Akio Taniguchi, Tom Bakx, Kaushal B. Marthi, Stan Zaalberg, Jochem J. A. Baselmans, Kah Wuy Chin, Robert Huiting, Kenichi Karatsu, Alejandro Pascual Laguna, Yoichi Tamura, Tatsuya Takekoshi, Stephen J. C. Yates, Maarten van Hoven, and Akira Endo. Tiempo: Open-source time-dependent end-to-end model for simulating ground-based submillimeter astronomical observations. *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy X*, Dec 2020. doi: 10.1117/12.2561014. URL https://research.tudelft.nl/en/publications/tiempo-open-source-time-dependent-end-to-end-model-for-simulating.

[29] Oncology medical physics. URL https://oncologymedicalphysics.com/quantifying-accuracy-precision-and-error/.

A

Appendix

A.1. Influence of Input Parameters of SPLITTER

In the early weeks of the project, the influence of the input parameters of SPLITTER has been tested. This was mostly done to get a feeling of what they are and how the output responds to a change in them, and no rigorous analysis has been done for them like the analysis for emission line and continuum estimation.

One parameter that showed quite a difference was the atmospheric temperature. This is assumed to be a constant by SPLITTER and has a default value of 273 K. Since this value is also used by TiEMPO, it should theoretically give the right spectrum output. As the atmospheric temperature can change a few Kelvins during the timespan of a few hours of observation time, it was tested to see what this parameter does to the continuum estimation. figure A.1 shows the relation between an offset in T_{atm} compared to its true value. As a value below 268 K is tried, the algorithm gives an error since this value is lower than the observed brightness temperature at some time and frequency, and a negative logarithm is calculated. The plot shows a clear increase in continuum offset for a higher T_{atm} input, and a decrease for a few Kelvins lower. It seems to give a right estimate for a passed T_{atm} about one Kelvin lower than its true value, but as this was only tested on one TiEMPO simulation, it cannot be said for sure that this fixes the overestimation.

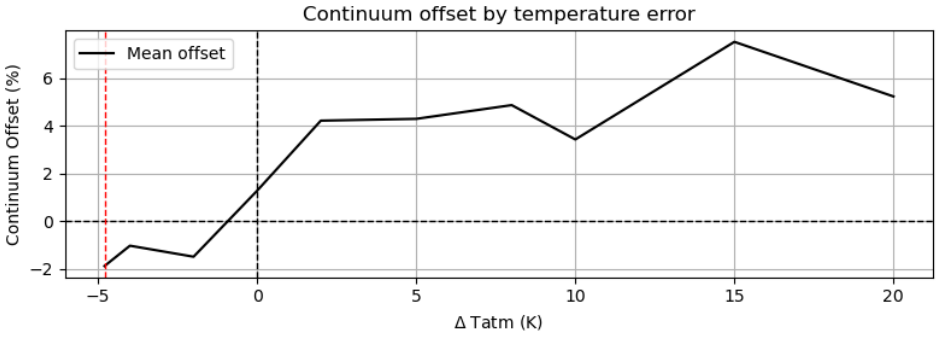

Figure A.1: Continuum offset as a result of a misestimated atmospheric temperature T_{atm}. Red dotted line gives the limit for underestimation as SPLITTER tries to calculate a negative logarithm below this value, since an observed brightness temperature somewhere in the data is larger than the given T_{atm}. Increase with T_{atm} is shown for continuum offset, and it seems to estimate the continuum correctly for a T_{atm} about one Kelvin lower than its true value

The parameter numpeaks, determining the number of peaks that SPLITTER will try to extract, showed to respond to what it says. Running the same TiEMPO simulation on SPLITTER with numpeaks = 1 and numpeaks = 5 results in the plot of figure A.2. It is seen that the red line with numpeaks = 1 skips a few important peaks at 285 GHz, but also reduces noise peaks at 320 GHz and 395 GHz.

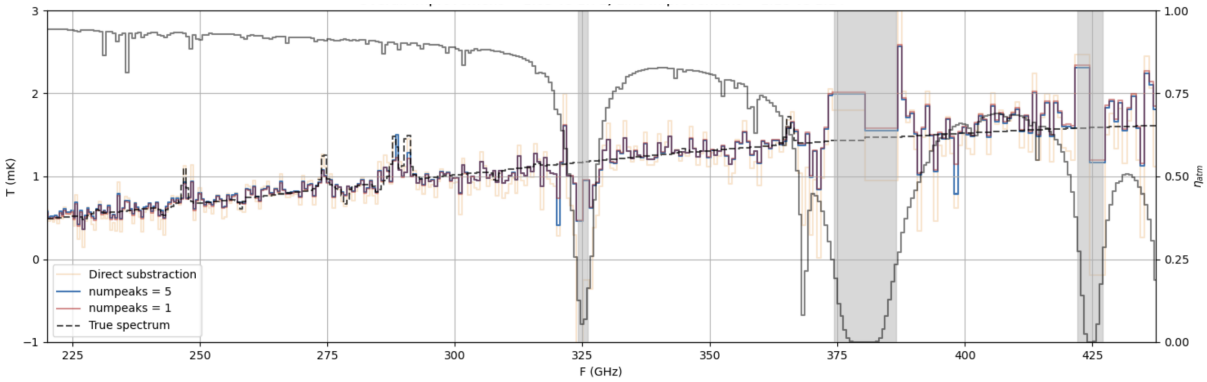

Figure A.2: Resulting spectrum using SPLITTER with numpeaks = 5 (blue) and numpeaks = 1 (red). Direct Subtraction estimate is also visible in the background (orange). Changing the numpeaks parameter to a lower value (red line) showed that some peaks were underestimated, as we would expect from its behaviour.

The parameter contcomps determines the components of the polynomial that SPLITTER tries to fit to the data. Changing this to 1 showed even more overestimation for the lower and higher frequency areas as seen in figure A.3, although it didn't have that much of an effect.

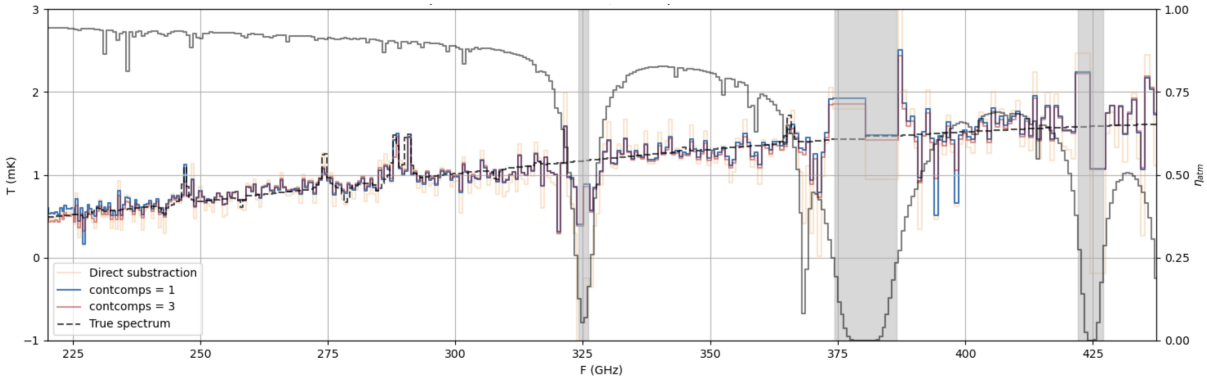

Figure A.3: Resulting spectrum using SPLITTER with contcomps = 1 (blue) and contcomps = 3 (red). Direct Subtraction estimate is also visible in the background (orange). Changing the contcomps parameter to a lower value gave a linear continuum, but made the overestimation worse for low and high frequencies.

A.2. Calculating the Standard Deviation over Multiple Chunks

To find the standard deviation of the noise, the signal needed to be split up into several chunks in time domain in order to cancel out the slow varying atmospheric noise, which made the average move slowly over time. The standard deviation was calculated over these chunks and to get the total standard deviation, the root mean square of all the standard deviations had to be taken. This is shown here.
Start with the equation for standard deviation of the whole data set:

$$\sigma = \sqrt{\frac{\Sigma_{i=0}^{N}(x_i - \mu(t))^2}{N}} \tag{A.1}$$

This sum can be split up into N_{ch} chunks of equal length. Each chunk will now have n samples ($N = N_{ch}n$). Enough chunks should be taken so that the mean can be approximated as a constant in each chunk, μ_j.

$$\sigma = \sqrt{\frac{\Sigma_{j=0}^{N_{ch}}\Sigma_{i=0}^{n}(x_{ji} - \mu_j)^2}{N_{ch}n}} = \sqrt{\frac{1}{N_{ch}}\Sigma_{j=0}^{N_{ch}}\frac{\Sigma_{i=0}^{n}(x_{ji} - \mu_j)^2}{n}} \tag{A.2}$$

where x_{ji} represents a sample in chunk j at index i. We recognise $\sigma_j^2 = \frac{\Sigma_{i=0}^{n}(x_{ji}-\mu_j)^2}{n}$, and can substitute to get

$$\sigma = \sqrt{\frac{1}{N_{ch}}\Sigma_{j=0}^{N_{ch}}\sigma_j^2} \tag{A.3}$$

Which is the root mean square of all the chunk standard deviations σ_j.
Note that if all the σ_j are the same (which should be the case if you take a high enough number of chunks and assuming the Gaussian noise effect of the atmosphere doesn't change), this equation will converge to

$$\sigma \approx \sqrt{\frac{1}{N_{ch}}N_{ch}\sigma_j^2} = \sigma_j \tag{A.4}$$

Therefore, for high amounts of bins, taking the root mean square is approximately similar to taking the regular mean of all standard deviations, or even just taking one value as they are all approximately similar.