

HARMONI: Advancing Science with the Next-Generation Integral Field Spectrograph

University of Oxford
The Queen's College



Laurence Henry Peter Routledge

Supervisor: Prof. Niranjan Thatte

Hilary 2022

Abstract

We start by investigating the noise behaviour of Teledyne HxRG infrared detectors. We look at data from KMOS H2RG darks, along with data from a MOONS engineering grade H4RG detector. From these we determine read noise distributions for the HxRG detectors, and identify a number of other non-uniform noise contributions. We then incorporate this information, along with a pre-existing HxRG noise generation tool, into `HSIM`. We then use this new advanced detector systematics (ADS) mode to produce simulations investigating the point source sensitivities, spatial variation of noise, reproducibility of recovered point source flux, and signal-to-noise ratio (SNR) for a number of different noise components. We determine from this that these additional detector noise effects provide non-trivial contributions to the noise profile of the detectors. In particular we find pink noise to be problematic, as it is a spatially-varying noise component that has a large impact on the overall SNR. We recommend careful use of the detector reference pixels to correct this, accounting for the potential for bad pixels within these.

We also look at the impact of an under-illuminated spectrograph entrance slit on telluric correction. We model the SINFONI 250×125 mas H-band line spread function (LSF) and vary the level of the input illumination. From this we determine an expected reduction in LSF width of 14%. We then use observations of CD-23 13701 and Callisto taken with SINFONI with adaptive optics (AO) correction to confirm a result of $14.0 \pm 2.1\%$ when fitted using `molecfit`. We attempt four methods of telluric correction; directly using CD-23 13701 as a telluric standard star, a synthetic spectrum created from `molecfit` run on CD-23 13701, this synthetic spectrum convolved with the appropriate LSF, and a synthetic spectrum created from `molecfit` run on the Callisto data. We perform this analysis on multiple spaxels within the Callisto data, and determine average residual sum of squares of 0.43, 0.34, 0.30, and 0.29 for the four methods. We determine from this that inadequate knowledge of the LSF leads to systematically worse telluric correction. Provided the LSF is known, the level of correction is similar between using a telluric standard star or the science target itself, with the remaining differences partly explained by variations in the atmospheric profile between the observations.

We end with testing a science case of HARMONI; recovering spatially resolved kinematics from high redshift galaxies. We use the `NUTFB` simulation to create a galaxy at an effective redshift of 1.44. We then produce a number of simulations using `HSIM` to determine how well we can recover the rotation curve and star formation rate (SFR), varying the simulated observation time and spectral resolution. We determine that 1 hour of HARMONI observation is sufficient to recover the kinematics to the point at which we become limited by effects within the galaxy. We also investigate the impact of the AO point spread function (PSF), particularly showing the effects of the PSF wings smearing out the rotation curve, as well as differences in the kinematics as a result of PSF elongation.

Declaration

The research conducted during this work was carried out by myself at the University of Oxford, under the supervision of Prof. Niranjan Thatte. No part of this thesis has been submitted for any degree or qualification here or elsewhere. All work not referenced explicitly in the text, or mentioned in this section, is my own.

Chapter 2 was based off of preliminary investigations by R. Houghton, with additional data collection assisted by E. George and D. Ives at the European Southern Observatory in Garching, Germany.

The Teledyne HxRG noise generation tool which is used in Chapter 3 was developed by Rauscher (2015). Section 3.1.1 summarises this paper. Aspects of the work in this chapter have been published as conference proceedings as Routledge et al. (2020).

Chapter 4 uses data taken as part of an observing proposal using SINFONI, with Co-Is N. Thatte, A. Smette and J. Richard. All data reduction and analysis was conducted by myself. An abstract for this work has been submitted for presentation at SPIE Montréal 2022.

Chapter 5 is based on collaborative work with M. Richardson. Section 5.1 is a report of the input cube created by M. Richardson. The values in Table 5.3 were produced by M. Richardson. All the mock observations and data reduction pipeline are my work. The figures and results were also produced by me. This work has been published as a refereed paper as Richardson et al. (2020).

The `HSIM HARMONI` science prediction tool was developed by Zieleniewski et al. (2015), and is maintained by myself and M. Pereira-Santaella. `Molecfit` was initially developed by Smette et al. (2015).

This work makes use of the open source `Python` programming language¹, versions 2.7 and 3.x, and additional modules `NumPy` (Harris et al., 2020), `SciPy` (Virtanen et al., 2020), `Astropy` (Astropy Collaboration et al., 2013, 2018) and `Matplotlib` (Hunter, 2007).

All copyright rests with the author of this work. Any work derived from this thesis should be acknowledged accordingly. Consent for reproducing parts of this thesis can be sought from the author.

Laurence Routledge
March 2022

¹<https://www.python.org/>

Acknowledgements

A wizard is never late! ... Nor is he early, he arrives precisely when he means to.

— Gandalf,
Lord of the Rings

A statement that holds true as much with theses as it does wizards! Although having a friendly wizard at hand may have made the last couple of years a little smoother. But the worst of it is hopefully behind us, and I've emerged the other side with a thesis complete. I think it's worth taking a few paragraphs to thank everyone who's helped me along the journey.

Firstly I would not have achieved this without the patience and support of Niranjana, particularly during the 'working from home' era that we have found ourselves in. Beyond this, I'd like to thank the wider HARMONI group for welcoming me to the world of instrumentation and integral field spectroscopy. Thanks go to Matthias, Miguel, Kearn, Fraser and Ryan, to name but a few of the vast team I have had the opportunity to be a part of. And of course to Vanessa, for all her tireless work behind the scenes. I also thank Liz, Derek, Benoit and Naidu for their assistance throughout my work, and particularly when I spent time in Germany with them. I'd also like to thank Alain and Mark for their guidance and collaborations.

Beyond this I'd like to say a quick word to the numerous people I've met, shared an office with, or enjoyed a slice of cake with on a Thursday morning! There's far too many people to name here without leaving somebody out, but I'll express an extra thanks to Tom, Andrea and Alvaro, who have all worked alongside me at various times. And of course to Ashling, who keeps everything running no matter what!

Within Oxford, I'd like to thank all my other friends that I've made, with special mention to all those equally as nerdy as me who like to sit down on an evening and roll some dice around pretending to be someone else: Doug, Ige, Jay, Rich, Harry, Harry, Will and Wills. It's been a great time, and I hope it continues to be so for a while to come.

And the wider group of people I've met over the years, either physically or virtually, such as those who will play a 'relaxing' game of *Overwatch* with me at the end of a stressful day. Team mates I've played with on *Oxford Omnics* and *overf5*, I count you amongst some of my good friends.

Lastly, I'd like to extend an extra special thanks to Christina, as without you the last few years would have been considerably tougher. You've always been there to encourage me and support me and that means more than words can say. And of course to my family.

Mum and Dad especially, for putting up with me, even though it's your job! All the pets as well, that are always there whenever I come home. And all my extended family; aunts, uncles, and cousins who I should really keep in better contact with.

"I don't know half of you half as well as I should like; and I like less than half of you half as well as you deserve.", as Bilbo might say.

But, without further ado, I have a thesis to present before I ramble any further! I hope you enjoy reading ***HARMONI: Advancing Science with the Next-Generation Integral Field Spectrograph***. (Rights to the film can be negotiated.)

Contents

1	Introduction	1
1.1	Basics of Integral Field Spectroscopy	1
1.1.1	How does Integral Field Spectroscopy work?	2
1.1.2	Utilisation of Integral Field Spectrographs	5
1.1.3	IFS with Adaptive Optics	6
1.1.4	SINFONI on the VLT	13
1.1.5	HARMONI on the ELT	15
1.2	Infrared Detectors and Noise	16
1.2.1	Principles of Infrared Detectors	17
1.2.2	Readout methods	19
1.2.3	Infrared Detector Noise	23
1.2.4	Signal-to-Noise Ratio (SNR) Equation for Infrared Astronomy . .	28
1.3	Telluric Absorption Features in the Infrared	32
1.3.1	What are Telluric Absorption Features?	32
1.3.2	Correction Methods	34
1.4	High Redshift Galaxy Kinematics	38
1.4.1	Rotation Curves	39
1.4.2	H α as a Tracer for Star Formation Rate and Kinematics	42
1.5	Thesis Outline	44
2	Study of Detector Noise on Teledyne HxRGs	47

2.1	Purpose of Investigation	47
2.2	Analysis of KMOS Darks	48
2.2.1	Choosing Data from KMOS	48
2.2.2	Characterising Noise	49
2.2.3	Comparison with October 2015 Data	55
2.3	Analysis with MOONS Engineering Grade H4RG	58
2.3.1	Data Acquisition	58
2.3.2	Results	60
2.4	Discussion and Conclusions	64
3	Testing Advanced Detector Systematics using Simulated Detectors in HSIM	65
3.1	Creating Simulated HARMONI Detectors	65
3.1.1	Teledyne HxRG Noise Generator	65
3.1.2	Integration into HSIM	69
3.2	Input Cube Creation and Analysis	74
3.2.1	HSIM Input Cube	74
3.2.2	Simulation Methodology	76
3.2.3	Analysis	78
3.3	Results and Discussion	80
3.3.1	Point Source Sensitivities	81
3.3.2	Spatial Variation and Reproducibility	82
3.3.3	Signal to Noise Ratio (SNR) Limits	86
3.4	Individual Noise Components	89
3.4.1	Spatial Variation and Reproducibility	89
3.4.2	Signal to Noise Ratio (SNR) Limits	93
3.5	Conclusions	97
4	Using SINFONI to Understand the Impact of an Under-Illuminated Slit for Telluric Correction	99

4.1	Background of Investigation	100
4.1.1	Motivation	100
4.1.2	Theoretical Models	101
4.1.3	Data Acquisition	106
4.2	Data Reduction and Analysis	107
4.2.1	Determining the Spatial PSF Width	108
4.2.2	Extracting Spectra	109
4.2.3	Telluric Modelling with <code>Molecfit</code>	112
4.3	Results and Discussion	115
4.3.1	<code>Molecfit</code> LSF Profiles	115
4.3.2	Telluric Corrected Callisto Spectrum	116
4.3.3	Best Methods for Telluric Correction	118
4.3.4	Comparison with Other Spaxels	121
4.4	Conclusions	124
5	Recovering High Redshift Gas Kinematics with <code>HSIM</code>	127
5.1	High Redshift Galaxy Simulation	127
5.1.1	Simulation Methodologies	128
5.1.2	<code>NUTFB</code> Galaxy	129
5.1.3	Creating a <code>HSIM</code> Input Cube	131
5.2	Analysis Methods	135
5.2.1	<code>HSIM</code> Observations	135
5.2.2	Velocity Maps	137
5.2.3	Pseudo Long Slit	140
5.2.4	<code>DiskFit</code>	141
5.3	Results and Discussion	142
5.3.1	Comparison with Simulated Galaxy	143
5.3.2	Exposure Time	146
5.3.3	Spectral Resolution	148

CONTENTS

5.3.4	Pseudo Long Slits	150
5.4	PSF Sensitivity	152
5.4.1	AO PSF Behaviour	152
5.4.2	PSF Elongation	155
5.5	Conclusions	158
6	Conclusions	159
6.1	Infrared Detector Noise	159
6.1.1	Conclusions	159
6.1.2	Continuing Work	161
6.2	Impact of Under-Illuminated Slit on Telluric Correction	162
6.2.1	Conclusions	162
6.2.2	Continuing Work	163
6.3	Recovering Gas Kinematics using HSIM	164
6.3.1	Conclusions	164
6.3.2	Continuing Work	165
6.4	Closing Statement	166
A	Running HSIM using ADS	167
A.1	ADS Mode Inputs and Options	167
A.2	ADS Mode Output Cubes	169

List of Figures

1.1	Different Designs of Integral Field Unit	2
1.2	Schematic of an Adaptive Optics System	9
1.3	Illustrative Operation Methods for SCAO and LTAO	12
1.4	Illustrative Diagram of an MCT photodiode	18
1.5	Circuit Diagram of the Readout Electronics for a Photodiode Array	20
1.6	Fowler and Up-The-Ramp Sampling Methods	22
1.7	Intermittent Channel Failure on KMOS H2RG Detector	29
1.8	Absorption Features in the Visible and Near-Infrared	33
1.9	Predicted vs. Observed Galaxy Rotation Curves	41
2.1	Raw KMOS Dark Exposures	49
2.2	Sample Processed KMOS Dark Exposure	51
2.3	Histogram of Values Returned by an Individual KMOS Detector Pixel . .	52
2.4	Base Mean and Standard Deviation Maps for KMOS Detectors	53
2.5	Distribution of KMOS Detector Pixel Means and Standard Deviations . .	54
2.6	Pixel Distributions from October 2015 Dark Exposures	55
2.7	Correlations Between Pixel Means and Standard Deviations on KMOS Detector 1	56
2.8	Example MOONS H4RG Engineering Grade Detector Dark	59
2.9	Base Mean and Standard Deviation Maps for MOONS H4RG Detector . .	61

2.10	Distribution of Detector Pixel Means and Standard Deviations on the MOONS H4RG	62
3.1	Flowchart for Adding Detector Noise using the Noise Generator Tool . . .	67
3.2	Power Spectra in Fourier Space of NIRSpec Detector Noise	68
3.3	Illustration of Conversion from Datacube Space to Detector Space as Im- plemented into <code>HSIM</code>	71
3.4	Distribution of Datacube onto Simulated <code>HSIM</code> H4RGs	72
3.5	Example Simulated HARMONI detector used in <code>HSIM</code>	73
3.6	Distribution of Detector Pixel Means and Standard Deviations on a Sim- ulated HARMONI H4RG	74
3.7	Position of Test Point Sources on HARMONI Field of View	75
3.8	Example Output Cubes from <code>HSIM</code> ADS Mode Testing Runs	78
3.9	Averaged Point Source Flux at 21 Fixed Locations with Enhanced ADS Noise Components	84
3.10	SNR as a Function of Integrations with Varying Enhanced ADS Noise Levels	87
3.11	Averaged Point Source Flux at 21 Fixed Locations with Single ADS Noise Components	91
3.12	SNR as a Function of Integrations with Single ADS Noise Components . .	94
4.1	Components and Convolutions of Model SINFONI LSFs with Full and Under-full Illumination	102
4.2	SINFONI Line Spread Functions in the H-Band at 250 mas Spaxel Scale .	105
4.3	Reduced CD-23 13701 H25 Data and Best Fit Function	109
4.4	Images of CD-23 13701 and Callisto at $1.7\ \mu\text{m}$ at 250 mas Spaxel Scale .	110
4.5	Recovered Spectrum for Telluric Standard Star CD-23 13701	111
4.6	Recovered Spectrum for Extended Science Object Callisto	113

4.7	Molecfit Reduced χ^2 Values for Gaussian and Lorentzian LSFs on CD-23 13701 and Callisto	114
4.8	Reduced Spectra and Molecfit Results Between $1.74\ \mu\text{m}$ and $1.76\ \mu\text{m}$.	115
4.9	Telluric Corrected Callisto Spectrum, Using Four Different Telluric Cor- rection Methods	117
4.10	Residuals of Callisto Spectrum for the Different Telluric Correction Meth- ods	119
5.1	Projected Extinction Along Multiple Lines-of-Sight of the Input Galaxy .	133
5.2	2D Velocity Maps and Sample 1D Spectrum Showing Least-Squares Fit .	139
5.3	H α Intensity Maps	143
5.4	Dependence of Simulation Resolution on Extracted Rotation Curve . . .	146
5.5	Dependence of Exposure Time on Extracted Rotation Curve	147
5.6	Dependence of Spectral Resolution on Extracted Rotation Curve	149
5.7	Comparison Between a Pseudo Long Slit and 2D Velocity Extraction . . .	151
5.8	Dependence of PSF Shape on Extracted Rotation Curve	154
5.9	1D and 2D Profiles of PSFs with Nonaxisymmetric Residual Jitter	156
5.10	Dependence of PSF Elongation on Extracted Rotation Curve	157

List of Tables

1.1	SINFONI parameters	14
1.2	HARMONI parameters	15
2.1	Correlation Coefficients for KMOS Detectors	57
3.1	Nominal ADS Noise Levels in <code>HSIM</code> v.310	66
3.2	Default <code>HSIM</code> Simulation Parameters for ADS Testing	77
3.3	Point Source Sensitivities for Different Noise Components Using ADS in <code>HSIM</code>	81
3.4	Spatially Averaged Flux and IQR of Recovered Point Source Fluxes with Enhanced ADS Noise Components	82
3.5	SNR Scaling with Varying Enhanced ADS Noise Components	86
3.6	Spatially Averaged Flux and IQR of Recovered Point Source Fluxes with Single ADS Noise Components	89
3.7	SNR Scaling with Single ADS Noise Levels	95
3.8	Summary of Detector Noise Effects and Current Inclusion in <code>HSIM</code> . . .	97
4.1	Summary of SINFONI Observations	107
4.2	Residual Sum of Squares for Different Telluric Correction Methods . . .	120
4.3	Relative Molecular Gas Column Densities from <code>Molecfit</code>	121
4.4	FWHM and Gas Column Densities for Varying Spaxels of Callisto H250 Data	122

4.5	Residual Sum of Squares of Different Telluric Correction Methods for Multiple Spaxels of Callisto H250 Data	123
5.1	Simulated Galaxy Properties	131
5.2	HSIM Simulation Parameters for NUTFB Galaxy	136
5.3	Recovered Galaxy Kinematics	145
5.4	Custom HSIM PSFs to Investigate AO PSF Performance	153
5.5	Custom HSIM PSFs with Varying Levels of Jitter	155

Chapter 1

Introduction

In this chapter we introduce the main areas of study of this thesis. We start with an introduction to integral field spectroscopy, as well as covering the basics of adaptive optics. We also highlight the two primary instruments involved in this thesis, SINFONI on the Very Large Telescope (VLT), and the upcoming HARMONI on the Extremely Large Telescope (ELT). We will then cover some of the more technical issues involved with observational astrophysics, which make up a large part of this work. We will review the operation of infrared detectors, with an introduction to a number of noise effects, and then telluric features and calibration methods for observations. Lastly, we will briefly discuss the scientific importance of the work done on high redshift galaxy kinematics.

1.1 Basics of Integral Field Spectroscopy

Integral field spectroscopy (IFS) is a form of spectroscopy in which spectra are simultaneously obtained for all points covering a modest field of view (FoV). This differs from traditional, long-slit spectroscopy, which as the name suggests can only observe along a narrow slit at any given time. The development of IFS has several advantages over long-slit spectroscopy. Being able to utilise all the light incident on the telescope and observe a 2 dimensional area of sky allows you to extract spectra from an extended object, or multiple objects within the FoV. This is a large improvement in observing efficiency, which is

crucial on telescopes with high demand where observing time comes at a premium, such as will be the case on the ELT. Additionally, observing simultaneously means you are removing potential variations between differing slit observations, such as airmass, seeing conditions or precipitable water vapour (e.g. Ren and Allington-Smith, 2002). IFS can also be utilised in conjunction with adaptive optics (AO) systems, which aim to correct for atmospheric turbulence to increase the overall spatial resolution. The AO performance will vary based on the seeing conditions at a particular time, making it an ideal partner to work alongside an IFS, as the correction will be applied simultaneously within the FoV.

1.1.1 How does Integral Field Spectroscopy work?

The first instrument that utilised IFS technology was SILFID (Vanderriest and Lemonnier, 1988), which incorporated multiple optical fibres as part of the integral field unit (IFU) to simultaneously collect light from multiple objects within a large field of view. This is only one of several methods of IFS, alongside lenslet arrays and image slicers, which are shown illustratively in Figure 1.1 (adapted from Allington-Smith et al., 1998).

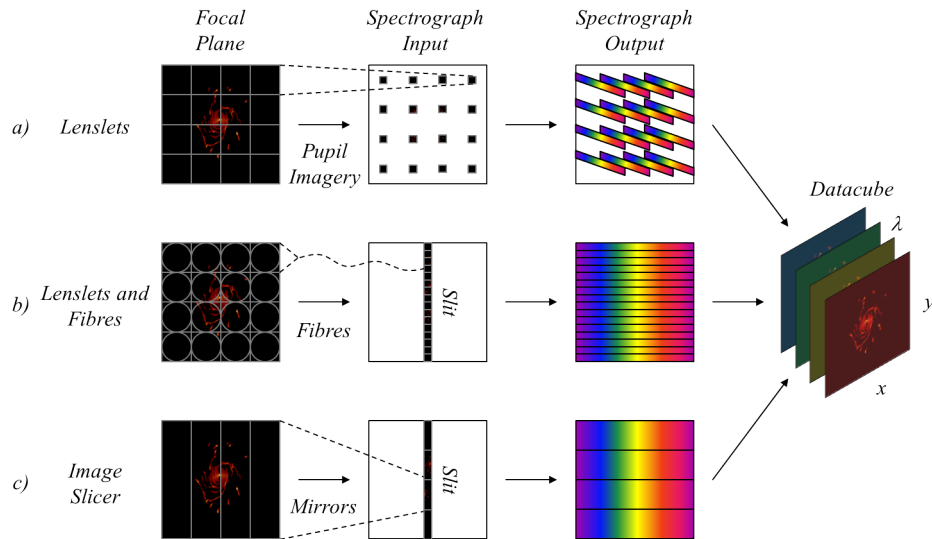


Figure 1.1: **Different Designs of Integral Field Unit.** Varying methods for carrying out integral field spectroscopy, showing *a)* a lenslet array, *b)* a lenslet array with optical fibres and *c)* an image slicer. In each case, the data is reconstructed into a 3-dimensional datacube, with two spatial axes and one spectral axis. *Image adapted from Allington-Smith et al. (1998).*

Optical Fibre Method

This method of IFS uses a series of optical fibres that each receive light from a different part of the sky. These fibres can then be rearranged at their other end to form a line that can be then be spectrally dispersed akin to a long slit, and thus a spectrum is obtained for each fibre. This is a relatively straightforward method conceptually, however does suffer from a couple of issues. Firstly, optical fibres are circular, and as such have a limitation on how closely they can be packed together. In addition, as optical fibres have a layer of cladding around them, this further reduces the amount of light captured for scientific use. Alongside this, they are prone to focal ratio degradation (FRD). FRD is essentially an effect in which light incident on the fibre at a given angle exits the fibre at a range of angles, reducing the signal on the detectors and lowering the resolution (e.g. Carrasco and Parry, 1994.)

Despite these drawbacks, optical fibre IFUs are a popular choice, and are particularly used for large scale surveys such as MaNGA for the Sloan Digital Sky Survey (Bundy et al., 2015; Drory et al., 2015; Gunn et al., 2006), SAMI (Sánchez et al., 2012), or CALIFA (Croom et al., 2012).

Lenslet Arrays

Differing from the optical fibre method, lenslet-based IFUs use an array of small lenses that are placed in the focal plane of the telescope. The light from each lenslet is then dispersed to obtain the spectra, one spectrum per lenslet, and this is incident onto the detector (Courtès, 1981). From a mechanical viewpoint, this is a simple method, relative to the use of optical fibres or image slicers (see next section), however does also have a limitation on the amount of usable light, as a clear gap is needed around the lenslets to avoid crosstalk. The TIGER instrument (Bacon et al., 1995) was the first to utilise this form of IFU, with SAURON (Bacon et al., 2001) advancing the principles for more extensive science results. More recently this method has been adopted for instruments such as SPHERE (Beuzit et al., 2019) on the VLT.

An adaptation to the above two methods uses a combination of lenslets feeding into optical fibres. This increases the complexity of the instrument, however overcomes many of the issues of optical fibres. It increases the filling factor of the optical fibres, as well as speeding up the incident beam, reducing the FRD. The downsides come in terms of the cost of design complexity, as you now need a combination of lenses and optical fibres, and in addition either a secondary output array of lenses or oversized optics is required to accommodate the fast output beam, otherwise there will be a reduction in throughput. Coupling losses, which can arise when coupling light between optical devices, can also be increased as a result of the more intricate design architecture. GMOS (Allington-Smith et al., 2002) on the Gemini Observatory uses a combination of lenslets and optical fibres.

Image Slicers

Also conceptually fairly straightforward but mechanically challenging, image slicer IFUs use a stack of mirrors at slightly different angles to slice the two dimensional field of view into thin strips, that then form a single, long slit. They have the advantage of utilising nearly all of the available field of view, as well as being possible to cool cryogenically, and thus are suitable for use in infrared astronomy. The biggest downside to image slicer IFUs are the technical challenges, with the optics being complex and expensive to manufacture, and can be significantly larger than other IFUs.

Proposed initially in 1938 (Bowen, 1938), it was first utilised for infrared astronomy on the 3D instrument (Weitzel et al., 1996), and both SINFONI (Eisenhauer et al., 2003) on the VLT and the upcoming HARMONI (Thatte et al., 2010) for the ELT utilise image slicer IFUs. MUSE (Bacon et al., 2010), also on the VLT, is another example of an image slicer IFU, which has a large FoV of 1×1 arcmin.

Datacubes

Common to all forms of IFS is that the process involves fitting three dimensions of data onto a two dimensional detector. Traditionally in long-slit spectroscopy, there are two

dimensions of data; one spatial and one spectral, however with IFS there is an additional spatial dimension. This results in a non-trivial distribution of the data on the detector, however this can be disentangled as part of the data reduction pipeline. The resulting three-dimensional datacube is usually returned as $I(x, y, \lambda)$, with I being the intensity of photon flux, x and y axes corresponding to the spatial dimensions of the area of sky, and the λ axis being the spectrum for that point.

One important effect to consider when combining this datacube is the impact of atmospheric differential refraction (ADR). This effect is caused by the wavelength-dependent refraction of light through the atmosphere, which effectively leads to an object appearing in a different position on the telescope focal plane at different wavelengths. Unless corrected for, this would lead to an object moving along the spectral axis of the datacube. This can be corrected by determining the shift caused by ADR as a function of wavelength, for example by theoretical modelling based on the temperature and pressure of the atmosphere, and then resampling and interpolating the data such that each spaxel returns the expected spectrum. One downside to this approach is that any erroneous data from, for example, bad pixels, is now spread across multiple data points (Filippenko, 1982; Walsh and Roy, 1990; Arribas et al., 1999).

1.1.2 Utilisation of Integral Field Spectrographs

Due to the versatility of IFS, it can be used for a range of science goals, with improving technology pushing the boundaries of what is possible with each new generation of instrument. IFUs are designed with particular science goals in mind, and as such there are several different flavours of IFU. Typically these are either:

- (a) Large instruments that have a singular field of view, where spectra are extracted for all the light within the field of view (subject to the limitations described above). These are useful both for large extended objects, as well as science cases such as resolved stellar populations. Both SINFONI and HARMONI fall into this category of IFU.

- (b) Instruments that utilise Multi-Object Spectroscopy (MOS), which have multiple, small IFUs, that can observe different parts of a field of a view simultaneously. These are very useful for large-scale surveys, as spectra can be extracted for many targets simultaneously, and overhead time, such as that to change the telescope pointing, are greatly reduced. The K-band Multi-Object Spectrograph (KMOS; Sharples et al., 2013) consists of 24 image slicer IFUs for example, while MaNGA (Bundy et al., 2015) has 17 optical fibre bundles.
- (c) High-contrast spectrographs are designed to maximise the ability to distinguish between very faint and very bright sources, such as is required for direct imaging of exoplanets, but with a low spectral resolution. SPHERE on the VLT is such an instrument, using the BIGRE IFU design (Antichi et al., 2009) with extreme adaptive optics and a coronagraph.

1.1.3 IFS with Adaptive Optics

As we mentioned earlier, IFS is often conducted in conjunction with adaptive optics (AO) to increase the spatial resolution of the instrument. The goal of AO is to sense and correct for turbulence causing distortions to light as it passes through the Earth's atmosphere. In principle, a point source observed with a circular telescope of diameter D should produce the image of an Airy pattern, with the first minimum located at

$$\theta = 1.22 \frac{\lambda}{D} \text{ rad}, \quad (1.1.1)$$

for a given wavelength λ . With a central obscuration, the FWHM of this PSF (in radians) is given by

$$\text{FWHM} = A \frac{\lambda}{D}, \quad (1.1.2)$$

where A depends on the size of the central obscuration, but is usually close to 1.

However, this assumes that the light is coherent onto the telescope, which for ground-

based astronomy is not the case. This is due to the light passing through the Earth's atmosphere and becoming distorted due to changes in the refractive index caused by different temperatures and molecular gas ratios. This leads to wavefront distortions in the light incident on the telescope, causing the Airy pattern to become broken up into speckles. Additionally, the atmosphere is non-stationary over time, meaning this speckle pattern will change, so that on long exposures this becomes a disc, known as the seeing-limited disc.

The extent of the impact of the atmosphere can be defined in terms of the Fried parameter r_0 (Fried, 1966; Roddier, 1981), and the Greenwood time constant τ_0 (Greenwood, 1977). These define the spatial extent of atmospheric regions that are approximately uniform, and the time-scale over which the atmospheric changes become problematic. Using the Kolmogorov model for turbulence (Kolmogorov, 1941; Tatarski et al., 1961), these wavefront perturbations, and thus phase variations, are caused by changes in the refractive index of the atmosphere. This model can be used to determine the variance in phase, σ^2 , across the aperture of the telescope, given by (Fried, 1965; Noll, 1976)

$$\sigma^2 = 1.0299 \left(\frac{D}{r_0} \right)^{5/3}. \quad (1.1.3)$$

The parameter r_0 can be calculated for a given wavelength if the strength of the turbulence, C_N^2 as a function of height, h , is known, through the equation (Hardy, 1998)

$$r_0 = \left[0.423 \left(\frac{2\pi}{\lambda} \right)^2 \sec(\gamma) \int_0^{\text{inf}} C_N^2(h) dh \right]^{-3/5} \quad (1.1.4)$$

where γ is the angular distance from zenith. This means that r_0 varies depending on the wavelengths being observed, but is typically $\sim$ tens of centimetres. Therefore for large telescopes such that $D \gg r_0$, the resolution is now limited by atmospheric turbulence, meaning

$$\theta = 1.22 \frac{\lambda}{r_0}. \quad (1.1.5)$$

The purpose of AO is to counteract this, so that the performance again begins to approach the diffraction limit of the telescope. This is done through the use of a wavefront sensor (WFS) in combination with a deformable mirror (DM), via a real time control (RTC) system, shown schematically in Figure 1.2. The purpose of the beamsplitter in this system is to divide off some of the light for the WFS, whilst passing the rest to the instrumentation, which for the scope of this work is an IFS.

There are multiple types of WFS, but they all share the common purpose of detecting the phase differences in the incident light. The most common method, known as a Shack-Hartmann WFS (SH-WFS; Collier, 1971), uses a lenslet array to create a series of pupils that are imaged onto a detector. Differences in wavefronts will cause these pupils to be shifted relative to a flat wavefront, and this is measured by the WFS. Increasing in use are pyramid wavefront sensors (PWFS), which use a lens to focus the light onto a pyramidal prism, with modulation to increase the linear range of the sensor, before a second lens forms an image on a detector. The differences in intensity distribution on the detector can then be used to sense the perturbations in the wavefronts. These are increasing in use partly due to their ability to obtain greater sensitivity and SNR for the same amount of incident light, meaning the amount of light required for good correction is smaller. This means that a dimmer star can be used for AO correction and thus a greater proportion of the sky can be covered in comparison to using a SH-WFS (Ragazzoni, 1996; Hutterer et al., 2019; Shatokhina et al., 2020).

The DM is then able to change shape according to the detected wavefront distortions. In a closed-loop system, such as illustrated in Figure 1.2, the corrected light is then again passed to the WFS for an iterative correction. In an open loop system, the DM propagates the light directly to the next stage in the instrument. The level of correction possible depends on the deformability of the mirror, given by the actuator density across it. The actuators allow parts of the mirror to move up and down, therefore the greater the actuator density, the finer the wavefront correction can be.

The RTC acts as the bridge between these two, passing the information from the WFS

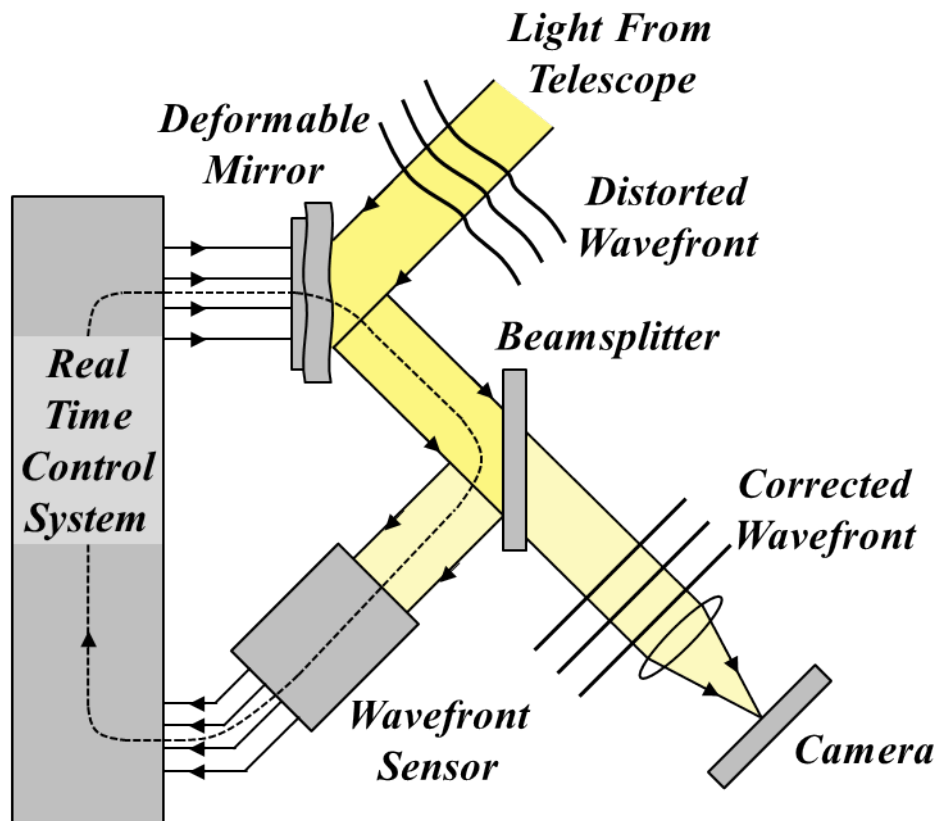


Figure 1.2: **Schematic of an Adaptive Optics System.** Basic layout of an adaptive optics system, operating in a closed loop. Incident, uncorrected, light hits the deformable mirror, before entering a beamsplitter which sends part of the light to the wavefront sensor. This feeds information into the real time control system, which changes the deformable mirror, increasing the correction on the light. The rest of the light is sent off to the scientific instruments. *Image adapted from one by Max, C., Centre for Adaptive Optics, UC Santa Cruz.*

to inform the DM how it needs to adapt. For the best possible correction, this has to run very fast, often into the hundreds of Hz. This is so that the mirror can adapt at the same frequency as the turbulence in the atmosphere is changing, i.e. the reciprocal of t_0 .

There are a number of ways of determining the performance of an AO system, including the Strehl ratio, encircled energy (EE), and FWHM of the resulting PSF. The Strehl ratio, defined by

$$S = \frac{I_{\text{obs, peak}}}{I_{\text{ideal, peak}}}, \quad (1.1.6)$$

where $I_{\text{obs, peak}}$ is the peak observed intensity of the PSF, and $I_{\text{ideal, peak}}$ is the maximum theoretical peak intensity, where typically a Strehl ratio of ≥ 0.8 corresponds to a good correction (Guenther, 1990). This is useful to determine the overall performance of the AO system, however does have some drawbacks, such as it does not infer anything about the shape of the PSF, as well as it not always being trivial to determine the theoretical peak intensity. EE can also be used, which is a measure of the amount of energy contained within increasingly sized apertures from the centre of the PSF. This gives a curve from 0% to 100% of the energy, giving some information about the shape of the PSF. This can also be calculated with square apertures, as can be more appropriate with detector pixels, giving the ensquared energy, rather than encircled energy. The FWHM is the width of the PSF where the flux is equal to half that of the peak flux. This can be determined from a resultant AO-corrected PSF and can be given in terms of detector pixels or corresponding angular resolution. This is a useful quantity for astronomers, as it determines the finest scale on which different features can be resolved.

Whilst modern AO systems are able to provide a very good correction, they are not perfect, and there are a number of issues and errors that are still present. These include:

- (a) **Fitting Error.** As mentioned previously, the ability for the AO to correct the wavefront is limited in part by the capabilities of the DM. If there are insufficient actuators to accurately correct the wavefront on the order of r_0 , residual imperfections will remain. As $r_0 \propto \lambda^{6/5}$ (Equation 1.1.4), these errors will be worse at shorter

wavelengths for a fixed number of actuators. Additionally, the stroke of the actuators needs to be sufficient such that they can correctly form the required shape on the DM to adjust the wavefront as needed.

- (b) **WFS Error.** The ability for the WFS to accurately determine the required correction is dependent on the amount of signal incident upon it. The brighter the target, the higher the SNR, and thus a better correction can be made. Alternatively, the frequency of the AO loop can be lowered to allow for longer integration on the WFS, however this increases servo error as a result.
- (c) **Servo Error.** Even when running at very high frequencies, there is a finite amount of time between determining the wavefront correction, and the DM adopting the new position. However with timescales of the atmospheric turbulence being short, on the order of milliseconds (e.g. Kellerer, 2007), there will be small differences. This can be counteracted by running the loop faster, however this typically comes at the cost of an increased WFS error.
- (d) **Isoplanatic Error.** Using light from a different target other than the science object such that it is off-axis means there is a small variation in the atmosphere through which the light has passed. This will leave some residual errors in the wavefront correction for the light from the science target, which increase the further separated the two objects are. Using a close correcting star, or the science target itself, helps to minimise this error.
- (e) **Non-Common Path Aberrations (NCPA).** These arise as result of differences between the optical path of the light through the instrument, downstream of the WFS. These are therefore not corrected for by the AO, however through careful calibration they can be accounted for (e.g. Mendiña Fernández, 2021).

There are multiple flavours of AO, serving different purposes in terms of level of correction and sky coverage. The main two which we use or simulate in the work in this thesis are shown in Figure 1.3, and explained below:

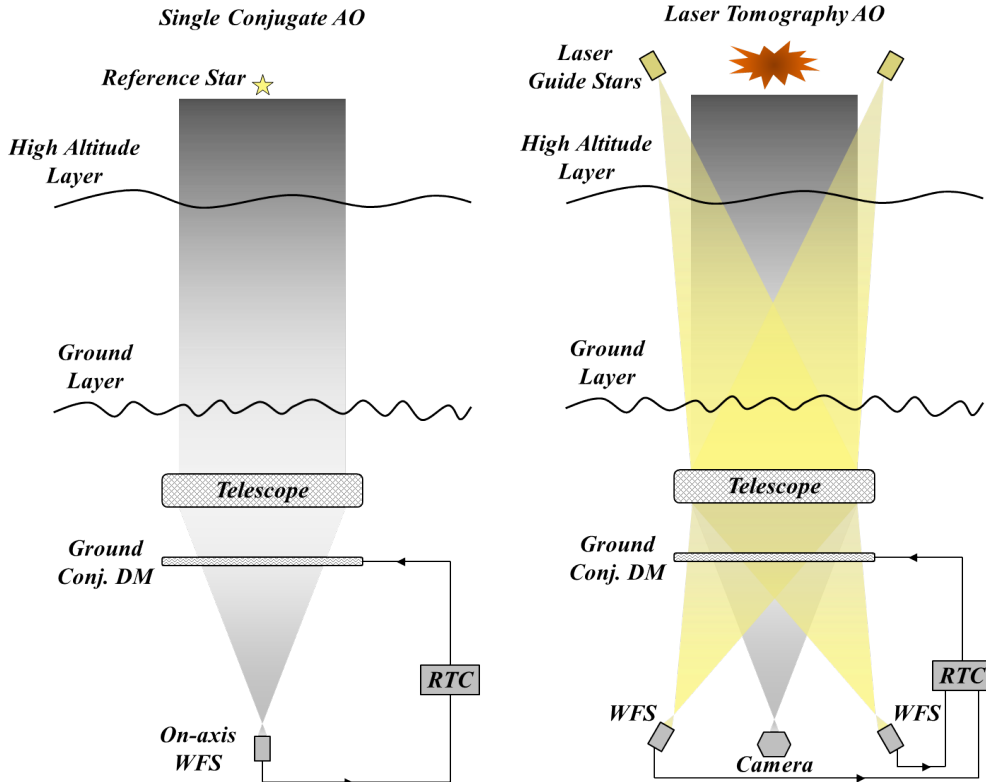


Figure 1.3: **Illustrative Operation Methods for SCAO and LTAO.** *Left:* Illustration of the operation of a single-conjugate AO (SCAO) system using a singular natural guide star and wavefront sensor (WFS) to detect the wavefront distortions, that are then passed to the deformable mirror (DM) via the real time control (RTC) system. *Right:* Operation of a laser tomography AO (LTAO) system, using an asterism of laser guide stars, each with their respective WFS that feed the information to the RTC to reassemble the atmospheric profile and adjust the DM. *Images adapted from ones on the European Southern Observatory's website.*

- (a) **Single-Conjugate AO (SCAO).** This uses a single natural guide star (NGS) to correct the atmospheric turbulence. A DM conjugate to the telescope pupil provides the corrections, however the FoV where this correction is applicable is limited. It is ideal in situations where the science object is suitable for use as the guide star, however when this is not possible a close, bright, star is required. As a result of this, SCAO is only capable of providing a good correction over a few percent of the sky.
- (b) **Laser Tomography AO (LTAO).** This uses an asterism of laser guide stars (LGS) in conjunction with multiple WFSs around the science target. The variations in wavefront error between the WFSs can be used to reconstruct the atmospheric profile in the column beneath the science target. This still requires a single NGS for tip-tilt correction, which is a low-order optical aberration corresponding to horizontal and vertical shifts in the spatial position of the target, which cannot be determined from the laser asterism. The requirements on the brightness of this star are far less stringent than for SCAO however, and therefore LTAO greatly increases the percentage of the sky that can be covered by an AO system. HARMONI is designed to feature both an SCAO and an LTAO system (Neichel et al., 2016).

1.1.4 SINFONI on the VLT

The Spectrograph for INtegral Field Observations in the Near Infrared (SINFONI; Eisenhauer et al., 2003) was installed on the VLT at the European Southern Observatory's (ESO) Paranal Observatory in Chile prior to decommissioning in 2019. While operational, SINFONI offered near-infrared integral field spectroscopy in the wavelength range $1.1\text{--}2.45\text{ }\mu\text{m}$, utilising an image slicer to divide the field of view into 32 slitlets on which spectroscopy could be performed. The height of these slitlets could be varied, offering 250 mas, 100 mas or 25 mas spatial pixel (spaxel) resolution in the direction across the slit, with twice the resolution along the slit. The spectra were then imaged onto a HAWAII-2RG detector, with 2048×2048 pixels.

1.1. BASICS OF INTEGRAL FIELD SPECTROSCOPY

Parameter	Value
Wavelength Range	1.1–2.45 μm
Spectral Resolution (band)	~ 1500 (H+K), ~ 2000 (J), ~ 3000 (H), ~ 4000 (K)
Spaxel Scales (Field of View)	250×125 mas ($8'' \times 8''$), 100×50 mas ($3'' \times 3''$), 25×12.5 mas ($0.8'' \times 0.8''$)
Detector	Teledyne HgCdTe $2k \times 2k$ (HAWAII-2RG)
Adaptive Optics	NOAO, NGS, LGS

Table 1.1: **SINFONI parameters.** Key parameters for the SINFONI instrument on the VLT.

SINFONI could be used either in a seeing-limited NOAO (no adaptive optics) mode, with no additional correction, or by using AO to improve the image quality. An NGS could be used, ideally brighter than $R \sim 11$ mag, however would still provide some correction down to stars as faint as $R \sim 17$ mag with good seeing conditions. An NGS close to the science target, or the science target itself, is always desirable, as otherwise the atmospheric turbulence seen by the AO sensing and science arm will be different. In addition, there was an option to use an LGS in combination with an NGS for tip/tilt correction, however this NGS can be further away and dimmer, as a large proportion of the correction can be determined from the LGS.

Science conducted with SINFONI has been wide-ranging, from large surveys such as SINS (Förster Schreiber et al., 2009) which investigates dynamical and physical properties of distant luminous galaxies, to solar system science such as radiative transfer processes in Uranus and Neptune (Toledo et al., 2019, 2020). Since it has been decommissioned, the integral field spectrograph (SPIFFIER; Iserlohe et al., 2004; George et al., 2016) is being installed as part of the Enhanced Resolution Imager and Spectrograph (ERIS; Amico et al., 2012) that plans for first light in 2022.

This work uses data obtained with SINFONI to investigate the impact of an under-illuminated input slit with regards to telluric correction, which is introduced in Section 1.3, and discussed in detail in Chapter 4.

1.1.5 HARMONI on the ELT

The High Angular Resolution Monolithic Optical and Near-infrared IFS (HARMONI; Thatte et al., 2010) is the upcoming IFS for ESO’s ELT. It is in many ways the ‘spiritual successor’ to SINFONI, offering a similar range of options, albeit at higher spatial and spectral resolution and utilising the larger light-capturing capacity of the ELT. HARMONI will offer spectroscopy over the spectral range $0.47\text{--}2.45\ \mu\text{m}$, also using an image slicer to divide up the field of view. Unlike SINFONI, however, HARMONI will have four individual IFUs each with two HAWAII 4RG detectors and two CCD231-84 detectors to allow for a greater spatial resolution while maintaining a similar field of view to SINFONI. HARMONI will have four spaxel sizes, of 4 mas, 10 mas, 20 mas and 30×60 mas.

Parameter	Value
Wavelength Range	$0.46\text{--}2.45\ \mu\text{m}$
Spectral Resolution (band)	~ 3300 (V+R, I+z+J, H+K), ~ 7000 (I+z, J, H, K), ~ 17000 (z-high, H-high, K-short, K-long)
Spaxel Scales (Field of View)	60×30 mas ($9.1''\times 6.1''$), 20×20 mas ($3''\times 4.1''$), 10×10 mas ($1.5''\times 2''$), 4×4 mas ($0.6''\times 0.8''$)
Detectors	$4 \times$ Te2v CCD231-84 $8 \times$ Teledyne HgCdTe $4\text{k}\times 4\text{k}$ (HAWAII-4RG))
Adaptive Optics	NOAO, SCAO, LTAO, HCAO

Table 1.2: **HARMONI parameters.** Key parameters for the HARMONI instrument designed for the ELT (as of Thatte et al., 2020).

HARMONI will also have the option of observing with adaptive optics, alongside a seeing-limited mode (NOAO), with single conjugate adaptive optics (SCAO), laser-tomography adaptive optics (LTAO) and high contrast adaptive optics (HCAO). SCAO uses a single NGS to provide the correction, providing there is a bright enough star ($R \leq 12$ mag for optimal performance) close to the science target. LTAO uses an asterism of six LGS to allow for correction over a large fraction of the sky, although still requiring one or more NGS for tip/tilt correction. HCAO requires a bright ($R \leq 9$ mag for optimal performance) on-axis star to obtain a contrast greater than 10^6 at 200 mas.

HARMONI is currently in the final design stage of development, with it being a first-light instrument for the ELT, currently planned for 2027. It is designed to be a work-horse instrument, capable of investigating a number of science cases such as observing spatially resolved kinematics of high redshift galaxies (Richardson et al., 2020), resolved stellar populations in nearby galaxies (Gonzalez and Battaglia, 2018), or studying atmospheres of planets and moons within our solar system.

Whilst the instrument is still in development, a HARMONI simulation tool (`HSIM`; Zieleniewski et al., 2015; Routledge et al., 2020) has been created to test the capabilities of the instrument. `HSIM` aims to simulate the expected behaviour of HARMONI to produce a resulting datacube in (x, y, λ) , akin to that obtained from a data reduction pipeline. A large portion of this work involves `HSIM`, with Chapters 2 and 3 improving the existing tool, and Chapter 5 using `HSIM` for quantitative science performance prediction.

In this work we focus primarily on the infrared capabilities of HARMONI. Infrared detectors can be particularly problematic in terms of noise, therefore increasing our understanding of these contributes greatly to improving the science that can be conducted, especially given that the majority of the HARMONI wavelength coverage is in the infrared. Furthermore, telluric absorption features are particularly troublesome in the infrared, which we look at using `SINFONI` in Chapter 4. We then look at galaxy kinematics, the tracer for which we use being $H\alpha$, that falls into the infrared at the redshift we simulate.

1.2 Infrared Detectors and Noise

An important part of an IFS is the detectors used, whether that is a traditional charge-coupled device (CCD) for visible wavelengths, or a Mercury Cadmium Telluride (MCT) detector for the near-infrared, such as will be used in HARMONI. As we rely on these detectors to obtain any data, it is important to understand how they work, as well as potential noise issues that may arise from them. We focus here on infrared detectors,

covering the physical processes behind them. We also introduce different readout methods which can be used to obtain data with infrared detectors. We then catalogue a list of potential issues that arise with them, before ending this section with a discussion on the signal-to-noise ratio (SNR) in the infrared.

1.2.1 Principles of Infrared Detectors

Infrared detectors use the principle of photoconductivity, which is to convert photons into an electric current that can be detected by an amplifier. The most basic form of this would be to use a semiconducting material with electrodes on opposite faces, creating an electric field within it. An incident photon of sufficient energy can then free an electron which creates a current as it passes through the material, which is then detected. From solid state physics, in order to achieve this an electron must be promoted from the valance band to the conduction band, which means the incident photon must have sufficient energy to overcome this band gap, implying a cut-off wavelength λ_c . This is determined by

$$\lambda_c = \frac{hc}{E_g} = \frac{1.24 \mu\text{m}}{E_g(\text{eV})} \quad (1.2.1)$$

where h and c are the fundamental Planck's constant and the speed of light, and E_g is the band gap energy. Typically these band gaps are of the order ~ 1 eV (Kittel, 2005), therefore in a pure semiconductor the cut-off wavelength is only around $1 \mu\text{m}$, above which photons would not have sufficient energy to create a photocurrent (Rieke, 2007).

This can be combated by doping a semiconducting material with impurities to alter their physical properties. Impurities can be either donors or acceptors, indicating whether they have more or less electrons available for conducting than the semiconductor. Donor impurities have additional electrons available for excitation into the conduction band, and create an n-type material, referring to the primary charge carriers being the negatively-charged electrons. Acceptor impurities are so-called as they accept electrons from the valance band to complete covalent bonds, leaving behind holes. These holes act as pos-

itively charged carriers, and thus the material is known as p-type. The result of doping a semiconductor in this way is that it lowers the required energy from E_g to the smaller E_i , the difference in energy between the impurity level and either the bottom of the conduction band (in n-type) or the top of the valance band (in p-type) (Rieke, 2007; Kittel, 2005).

By creating a p-n junction, where one side is p-type and the other n-type, a photodiode can be created, and this is what is used in infrared detectors. The donor n-type material can release electrons that are easily absorbed by the p-type material, and this creates a potential difference across the junction. An external voltage, or bias voltage, can be applied, and when incident photons on the detector cause a movement of charge carries, this can be detected as a change in the bias voltage. In Figure 1.4 we show a diagram of an MCT photodiode. The substrate is used in the creation of the array, but is typically removed prior to astrophysical use. This has numerous benefits, including increasing the quantum efficiency, allowing detection of visible light and removing fringing within the substrate (e.g. Beletic et al., 2008).

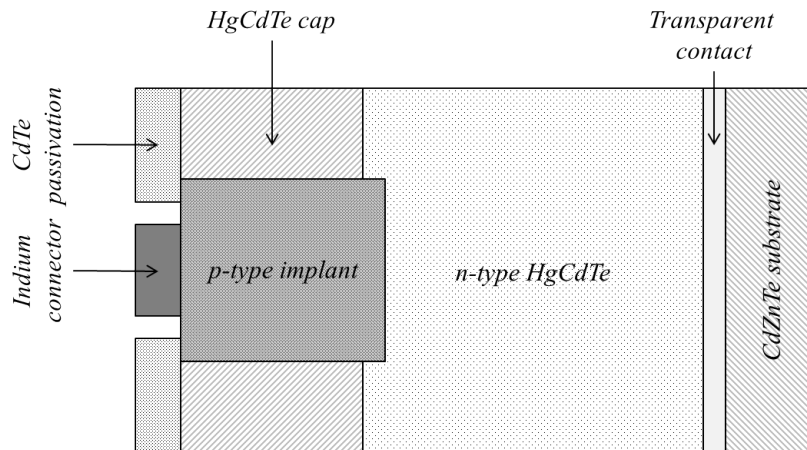


Figure 1.4: **Illustrative Diagram of an MCT photodiode.** A typical p-n junction photodiode manufactured out of Mercury Cadmium Telluride (HgCdTe). Light is incident on the right hand side of the figure, and the signal is detected through contacts on the left. The substrate is used in the creation process, but can be removed prior to use. *Image adapted from Rieke (2007) and based on Garnett et al. (2004).*

In this work we focus on the MCT detectors produced by Teledyne, known as HxRG detectors. The H stands for HAWAII, which is short for HgCdTe Astronomical Wide Area

Infrared Imager. The x is short for the number of 1024×1024 blocks that make up the full array, which in this work we look at arrays with 2048 and 4096 pixels. The R refers to reference pixels being present on the detector, and G denotes guide window capability (Blank et al., 2011).

1.2.2 Readout methods

The infrared detector is only part of the entire setup, as electronics are required to convert the photocurrents into readable signals. This is typically done by combining the photodiode with a readout integrated circuit (ROIC) through indium connectors to create a focal plane array (FPA). A ROIC consists of metal oxide semiconductor field effect transistors (MOSFETs) performing charge-to-voltage conversion for each pixel. MOSFETs were successfully created first in 1959 (Kahng, 1960), building on work conducted by L. E. Lilienfeld in the 1920s and 30s (e.g. Lilienfeld, 1926), and are the foundation of modern electronics. A silicon MOSFET is connected to each pixel of the detector array, and this collects the detected signal to be read out of the detector (Rieke, 2007).

In Figure 1.5 we show a circuit diagram of part of a ROIC. This shows four cells in the array, with the detector pixels represented by diodes. Charge is collected and builds up on the transistor T_1 until the circuit is read out. This is done by applying power to a given row and column simultaneously, for example r_1 and c_1 . This closes the switch transistors T_2 , T_3 and T_4 and applies power to T_1 , connecting the signal to the output amplifier to allow a reading to be taken. The pixel could then be left to accumulate charge further by removing power from c_1 , and c_2 can be powered to read out the next pixel on the row. This can be done for every pixel on the row, before moving onto the next row by powering r_2 . Alternatively, the reset line can be powered, setting the input of T_1 to the reset voltage V_R . This makes it destructive, as the accumulated charge is reset when doing this (Rieke, 2002, 2007).

This versatility in principle could allow for any pixel to be read out in any order, and for any single pixel to be reset. However in practice, detector arrays are controlled by

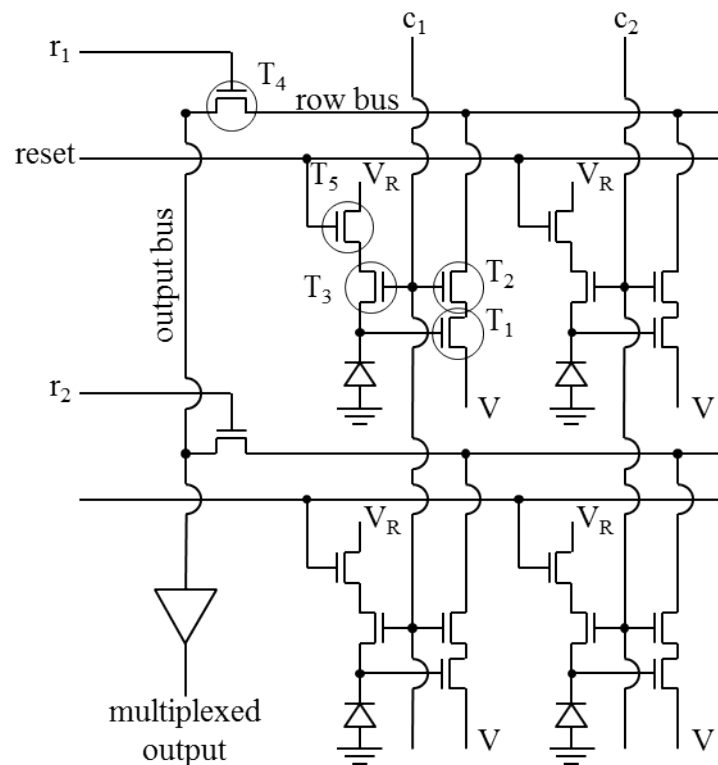


Figure 1.5: Circuit Diagram of the Readout Electronics for a Photodiode Array. Readout circuit for four cells of a detector array, shown schematically as diodes. Transistor T_3 allows for resetting an individual pixel, and could be removed if that behaviour is not required. *Image adapted from Rieke (2002).*

a clocking pulse that reads out the pixels in turn for simplicity. Additionally, with the HxRG devices, the number of individual output amplifiers can be varied, increasing the readout speed at the cost of additional power requirement. For the H2RGs we look at in this work, the detectors have 32 output amplifiers, while the H4RGs have 64.

The HxRGs also have the option to read out a subarray to a separate output, allowing for a guide window to be implemented for an imager on long exposures for example. The outer 4 rows/columns of the HxRG devices are not photosensitive, and are used as reference pixels to track some of the noise that can be introduced during the readout process (MacDougall et al., 2006).

The ability to read the signal on a pixel without affecting the charge allows for a variety of ways in which to read out the array. In principle, each pixel could be read and then reset immediately, with an additional read made directly afterwards, the signal being

the difference between the two reads. However, this read-reset-read method introduces a large amount of additional noise (see Section 1.2.3) and is therefore unfavourable. Given that a pixel will still accumulate charge when power is not provided to the MOSFET, non-destructive sampling methods are preferred, with larger integration times between reads. This relies on the electronics to be stable over large enough timescales such that the fluctuations through bias drifts, for example, are small compared to the signals being measured. This is also where reference pixels can be used to help correct for such drifts (Rieke, 2007).

We outline two common readout methods below, namely Fowler(-N) sampling and sample-up-the-ramp:

Fowler(-N) Sampling

Fowler(-N) sampling is a form of correlated double sampling (CDS) named after an original proponent of this method (Fowler and Gatley, 1990), and is shown diagrammatically in the left panel of Figure 1.6. In this approach a number of reads, N , are taken close together at the start of the integration, known as pedestal reads, denoted by P_x . After the desired exposure time, T_{ex} , another N signal reads are taken, denoted by S_x . For each pair of reads, the difference is taken, and the average of the N reads gives the total signal. This method is advantageous over continuously reading the signal as the array can be turned off for a large amount of the integration time, minimizing the power dissipation. In addition, as the read noise is typically dominated by high-frequency variations, taking multiple reads close together and calculating the average can reduce the total read noise by approximately $\sqrt{N}$ (Rieke, 2007).

Sample-Up-The-Ramp

Sample-up-the-ramp (SUTR) is an alternative non-destructive readout method in which a series of reads are taken at regular intervals, and is shown in the right panel of Figure 1.6. In this method a read is taken every time interval Δt for the duration of the exposure, T_{ex} .

Depending on the readout electronics, a running gradient can be made, or a least-squares fitting method can be applied to the full ‘ramp’ when all reads are taken. This is then used to determine the signal over the length of the integration.

SUTR can be preferred over Fowler(-N) sampling as the signal is sampled periodically, meaning it is less susceptible to transient, unexpected noise sources. For example, an ionizing particle could cause a jump in the signal, however provided the full slope is obtained, this discontinuity can be accounted for, and the remaining data can be used. This is also useful in the event of a pixel reaching saturation before the exposure is completed, as the ramp can be extrapolated and a signal estimated accordingly. This is not ideal, however, as saturated pixels have other problems with them, such as persistence, which is discussed more in Section 1.2.3. In longer exposures the ability to recover data from points along the ‘ramp’ can lead to this method being more desirable, whereas in shorter exposures the theoretical improvement in read noise reduction of Fowler(-N) sampling may be preferred. A composite method can also be adopted, where a number of reads are taken in quick succession at multiple times over the observation, to gain the benefits of both techniques (Rieke, 2007).

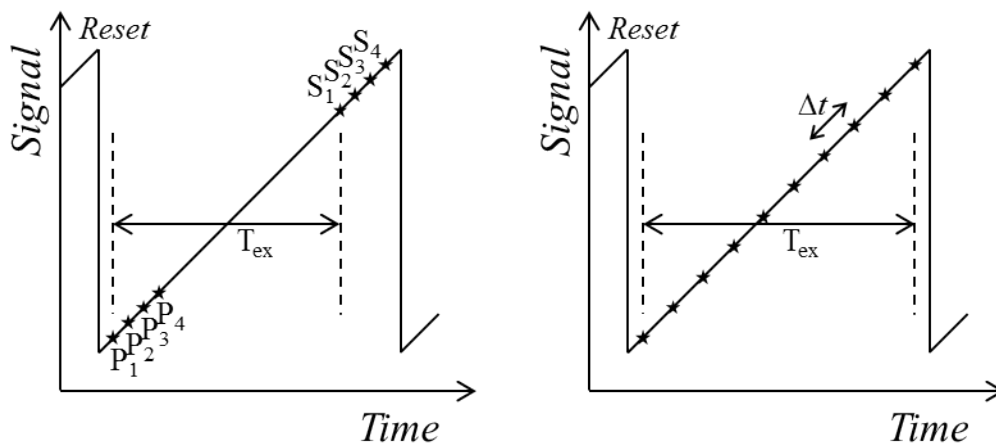


Figure 1.6: **Fowler and Up-The-Ramp Sampling Methods.** Different methods for reading out an infrared detector array showing Fowler(-4) sampling (left) and sample-up-the-ramp (SUTR; right). *Image adapted from Finger et al. (2000).*

1.2.3 Infrared Detector Noise

As with every physical process, it is not possible to have a pure signal with no noise. During the process of obtaining the data, several forms of noise arise, largely as a result of the readout electronics. As many astronomical observations are of faint targets, particularly with an instrument such as HARMONI, it is important to understand where this noise comes from. It is then equally crucial to minimise the impact, either through hardware, observing strategy, or as part of the data reduction process. In this section we introduce a number of sources of noise, along with some other detector issues, explaining their origins, and discussing measures that can be implemented to reduce them.

Johnson-Nyquist Noise

Johnson-Nyquist noise is thermal noise caused by Brownian motion of charge carriers (Johnson, 1928; Nyquist, 1928). This is unavoidable, and the current generated (and hence signal produced) is dependent on the temperature, T , and resistance, R , of each indium connector through the equation

$$I_{rms} = \sqrt{\frac{4k_B T \Delta f}{R}}. \quad (1.2.2)$$

This is the root mean square current, where k_B is the Boltzmann constant and Δf is the frequency bandwidth. As this noise is the result of fundamental physical processes, it is not possible to entirely remove it, however it can be minimised through design and operation. For the detectors that we investigate in this work the resistance is dependent on Teledyne's manufacturing process and cannot be changed. However, by cryogenically cooling the detectors in operation, the temperature of the connectors is reduced, and thus the amount of Johnson-Nyquist noise.

kTC Noise

This noise is another thermal noise, similar to Johnson-Nyquist noise. This occurs as a result of thermal currents causing a randomly fluctuating charge, Q , on the photodiode, equal in magnitude to

$$Q = \sqrt{k_B T C} \quad (1.2.3)$$

where C is the capacitance. This varies over a timescale of $\tau = RC$, therefore in order to keep this noise consistent we want the resistance to be high. If we refer to the circuit diagram in Figure 1.5, if the reset line isn't powered, and thus T_5 is open, then the resistance is high and the variations are low. However, when the reset line is powered and the switch is closed, the resistance is low and thus the fluctuations are rapid. This is why continuously resetting the photodiode adds a large amount of additional noise, and non-destructive readout methods that allow the signal to accumulate are favoured. When using readout methods such as SUTR or Fowler sampling, the kTC noise is almost entirely removed (Rieke, 2007). HARMONI will use SUTR, therefore it is expected that there will be very little kTC noise in the data.

Dark Current

Dark current is a current that flows in the photodiode in the absence of photons. This is related to defects within the semiconducting material. i.e. HgCdTe for the HxRG devices. Through quantum mechanical processes, electrons and holes can change energy bands in the absence of any light, and this is made easier through the presence of defects. This creates a current that is indistinguishable from a signal generated by an incident photon, aside from it also appearing in 'dark' observations, where no light is incident on the detector. In theory the dark current can be measured and removed, however there will be residual Poisson, or shot, noise as a result that will remain after the average dark current is subtracted. As the processes through which dark current occurs are dependent on temperature, cooling also decreases this noise. In typical Teledyne infrared detectors,

operating at cryogenic temperatures with a cutoff wavelength of $\sim 2.5 \mu\text{m}$, dark current introduces additional unwanted signal at the level of $\sim 0.01 \text{ e}^- \text{ s}^{-1} \text{ pixel}^{-1}$ (e.g Jerram and Beletic, 2019).

Pink ($1/f$) Noise

$1/f$ noise is so named as the power spectral density of the noise is inversely proportional to the frequency that noise occurs at. Each electronic component in the architecture of the detector can introduce its own pink noise. Pink noise can manifest in several different ways in the detector, including:

- (a) **Horizontal banding.** Each row is connected to the multiplexed output by a MOS-FET that is subject to its own $1/f$ noise. As this is different for each row, this can introduce a banding effect, where each pixel on that row has a similar level of noise associated with it. One way to counteract this is through the use of the reference pixels around the edge of the detector. These are read out in the same way, although are not photosensitive, therefore they can be used to track the $1/f$ noise present in that row. Ensuring that this is done correctly for each read of the detector is important, to accurately track the amount of $1/f$ noise on each read.
- (b) **Alternating Column Noise (ACN).** ACN is an effect seen on HxRG devices in which odd and even columns on a given output amplifier have a different noise profile. This is not a fundamental result of the array infrastructure, however was found to be the result of a proprietary design choice made by Teledyne with the HxRGs (Rauscher, 2015). The implication of this is that there is additional $1/f$ noise that affects the odd and even columns separately, which causes the striping sometimes seen in images. With traditional readout methods it is hard to negate ACN, however an alternative method known as Improved Reference Sampling and Subtraction (IRS²) interleaves reference pixels with regular pixels, and has been shown to reduce the impact of ACN (Rauscher et al., 2012).

(c) **Picture Frame Effect.** This effect is so-called as it appears as a border or ‘picture frame’ around the edge of the detector. It is driven by small temperature changes during the readout process (Rauscher et al., 2012 and references within) and originates from $1/f$ noise on the bias voltages (Rauscher, 2015). This effect depends on the controller used with the detectors, and can vary between detectors. Controlling the temperature fluctuations is one way of minimising the impact of this effect, as is avoiding the edge of the detectors where possible with observational data.

White ‘Read’ Noise

White noise differs from pink noise in that the power spectral density is the same at all frequencies. This means that at high frequencies it is dominant over pink noise, which drops off as $1/f$. Within the detector architecture, each pixel has a corresponding MOSFET which introduces pink noise. However, as each pixel contributes a different amount of noise independently, when viewed as an overall image, this becomes white noise. Taken overall, the pixels have a read noise distribution that is often approximated by a Gaussian function, however does have a tail to high values (Rieke, 2007) that we explore further in Chapters 2 and 3.

Non-Linearity

This is an effect in which the detected response does not increase linearly with the incoming signal. This is a result of amplifier bias reducing slightly as charge builds up, which becomes more pronounced as a pixel reaches the maximum potential charge, known as full well (Blank et al., 2011). Above this threshold the pixel becomes saturated, at which point additional charge can spread and ‘bleed’ into neighbouring pixels, although this can be corrected for in data reduction. This is usually defined as the point at which linearity deviates by more than 5%, and should be avoided as much as possible. By characterising the linearity of a given detector, the bias voltages can be adjusted to produce the best performance possible. Through careful planning of observations, saturation can be avoided

by ensuring exposure times are not too long so that a typical pixel will not be saturated.

Reciprocity Failure

This is a similar effect to non-linearity, in that the signal does not respond linearly with flux, however it is dependent on the incident photon rate, rather than charge accumulation. It has been seen on H2RG devices that they respond better to higher fluxes than lower ones. This causes issues particularly in understanding the QE of the detector, as it will bias it either higher or lower depending on the flux. This effect is seen to be reduced at cryogenic temperatures, however can still be present to a small degree (e.g. Biesiadzinski et al., 2011, 2014). We hope that in HARMONI this effect will be minimal, or not present.

Electronic Crosstalk

Electronic crosstalk is an effect in which a signal is propagated in an unwanted fashion from one area of the electronics to another. This can occur due to signals travelling in two adjacent, unshielded wires and inducing a copy of the signal in the other. Due to the complexity of the array architecture this phenomenon can occur, creating signals in areas of the detector which should be devoid of signal (e.g. Bushouse et al., 2008). This effect can be minimised through careful design, and additionally through calibration, as once the amount of electronic crosstalk present is known, it should be static with time. This can be done by measuring the responses on each amplifier block when illuminating only a single amplifier at one time, to generate a crosstalk matrix of size $n \times n$, where n is the number of amplifier blocks (George et al., 2018).

Persistence

Persistence is in effect an increased dark current caused by previous exposures. An amount of charge can remain trapped, or ‘persists’, after the pixel is reset, and this slowly leaks out over time. The exact nature of persistence is non-trivial, and can impact some parts of a detector more than others. Much work has been carried out to understand and

mitigate the impact of persistence (e.g. Smith et al., 2008; Smith et al., 2008; McLeod and Smith, 2016) however it remains an issue with infrared astronomy. Persistence can still appear several days after the initial observation, and some methods work better than others for reducing its impact (Tulloch, 2018), however this will be an effect that will impact the HARMONI detectors in some way. Recent work by ESO has developed a persistence correction pipeline using persistence models developed by Tulloch et al. (2019), and is currently undergoing testing, with the goal of correcting $\gtrsim 80\%$ of the persistence (Neeser, 2021).

Intermittent Channel Failure

This is an effect noticed on the KMOS H2RG detectors in which some columns appear to lose a significant portion of their flux. This is seen when observing point sources with low natural seeing, and results in profiles that do not reflect the true nature of the target, as demonstrated in Figure 1.7. This effect does not seem to have an impact on the electronic crosstalk, suggesting that the correct flux is initially recorded, and then becomes lost during the readout process (Houghton, 2016). It is yet to be known whether this is a unique effect to the KMOS H2RGs, or if a similar effect will be seen on the HARMONI H4RGs.

1.2.4 Signal-to-Noise Ratio (SNR) Equation for Infrared Astronomy

In its simplest form, the signal-to-noise ratio is a measure of the total desired signal, S , divided by the total noise, N , such that $\text{SNR} > 1$ means the signal is larger than the noise, whilst an $\text{SNR} < 1$ indicates that the noise is larger than the signal. As we have discussed above, infrared detectors produce a measurable response from incident photons, which comprises the signal. We have already discussed a number of sources of noise, including dark current and different sources of readout noise, however there are additional sources of noise not intrinsic to the detector used. Photon arrival on the detector is a discrete process, and is therefore subject to Poisson statistics and corresponding shot noise. In

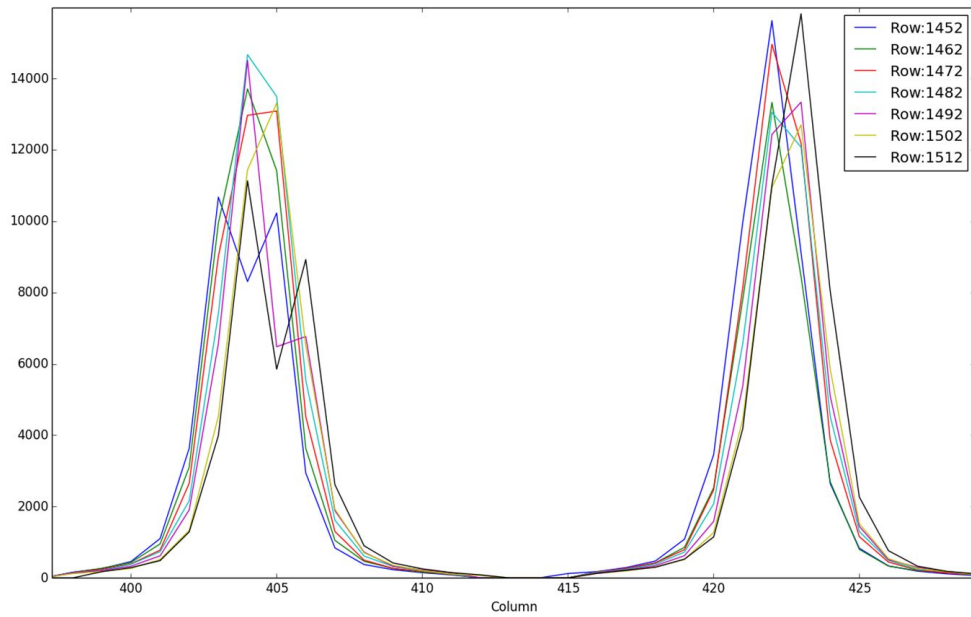


Figure 1.7: **Intermittent Channel Failure on KMOS H2RG Detector.** Exposure KMOS.2013-10-30T23:50:51.778, available on the ESO archive, showing the impact of intermittent channel failure. The right side shows an expected profile across consecutive columns, while the left shows the intermittent channel failure, in which the peak flux is suppressed. *Image reproduced from Houghton (2016).*

addition to that, there are other background sources of photons that can appear as a signal, such as thermal emission from the telescope, or sky emission. In principle then,

$$\text{SNR} = \frac{S}{N} = \frac{\text{Desired Signal}}{\text{Total Noise}} \quad (1.2.4)$$

For processes that are governed by Poisson statistics, including dark current in the detector, the variance is equal to the mean, so the average noise is equal to the square root of the signal. We can also assume that these are uncorrelated, therefore we can add the variances, σ_i^2 , together for each component i . For the non-Poissonian noise sources introduced by the detector, things are a little more complicated, however. Many of these sources are correlated, and varying in physical location on a detector, so there is no simple way of determining the noise. We assume a crude first approximation that we can sum the variance of these too (Rauscher, 2015), however we will show in Chapter 3 that a simulation approach is required for a full understanding of their impact.

We therefore now have

$$\text{SNR} = \frac{S}{\sqrt{\sum_{i=1}^x \sigma_i^2}} = \frac{S}{\sqrt{\overline{S} + \overline{B} + \overline{T} + \overline{D} + \sum_{j=1}^n \sigma_j^2}}, \quad (1.2.5)$$

where σ_j^2 are the variances of the detector noise components, including white and pink readout noise, and $\overline{B}$, $\overline{T}$ and $\overline{D}$ are the average values of the background emission, thermal emission, and dark current respectively. Additionally, in the infrared a second sky exposure is usually taken, which includes all the same noise except for the Poisson noise in the signal, such that

$$\text{SNR} = \frac{S}{\sqrt{\overline{S} + 2(\overline{B} + \overline{T} + \overline{D} + \sum_{j=1}^n \sigma_j^2)}}. \quad (1.2.6)$$

In order to determine how this will vary with longer observations, we have to understand how both the signal and noise sources accumulate. For a source with flux per unit area F_λ , the number of electrons from an extended object incident on area of the detector, N_λ , is given by

$$N_\lambda = \frac{F_\lambda f_{\text{trans}} A \Delta \lambda N_{\text{spax}} d_{\text{spax}}^2}{E_\lambda} \quad (1.2.7)$$

where f_{trans} is the total transmission fraction, A is the light capturing area of the telescope, $\Delta \lambda$ is the wavelength resolution, N_{spax} and d_{spax} are the number and size of the spaxels, and E_λ is the energy of a photon at that wavelength. For a point source, the flux is distributed according the PSF, with the additional fraction f_{PSF} dictating how much flux falls within the detector area.

The number of electrons from background and thermal emissions can also be expressed in the same manner, through

$$N_B \approx \frac{F_B f_{\text{trans}} A \Delta \lambda N_{\text{spax}} d_{\text{spax}}^2}{E_\lambda}, \quad N_T \approx \frac{F_T f_{\text{trans}} A \Delta \lambda N_{\text{spax}} d_{\text{spax}}^2}{E_\lambda}, \quad (1.2.8)$$

where we make an approximation that the transmission fraction will be the same for each emission component. This is not entirely correct, as some emission will only pass through parts of the instrument, however is a good approximation for these calculations. As we established in Section 1.2.3, dark current contributes a constant rate of additional flux, and we assume that the other detector noise components are independent of exposure time.

If we therefore expose for an amount of time T_{exp} a number of times N_{int} , the SNR in the infrared becomes

$$\begin{aligned} \text{SNR} &= \frac{N_{\lambda} T_{\text{exp}} N_{\text{int}}}{\sqrt{N_{\text{int}} [N_{\lambda} T_{\text{exp}} + 2(N_{\text{B}} T_{\text{exp}} + N_{\text{T}} T_{\text{exp}} + D N_{\text{spax}} T_{\text{exp}} + N_{\text{spax}} \sum_{j=1}^n \sigma_j^2)]}} \\ &= \frac{N_{\lambda} T_{\text{exp}} \sqrt{N_{\text{int}}}}{\sqrt{[N_{\lambda} + 2(N_{\text{B}} + N_{\text{T}} + D N_{\text{spax}})] T_{\text{exp}} + 2 N_{\text{spax}} \sum_{j=1}^n \sigma_j^2}}. \end{aligned} \quad (1.2.9)$$

If we are now in a situation where we have a bright object such that the noise is dominated by N_{λ} , we can simplify Equation 1.2.9 to

$$\text{SNR} \propto \sqrt{N_{\lambda} T_{\text{exp}} N_{\text{int}}}, \quad (1.2.10)$$

i.e. the SNR increases either as the square root of the observing time, or by the flux from the source. Therefore for an object of given flux, the SNR can be improved either for observing with longer integrations, or by taking more integration blocks. Usually the latter is easier to increase, so as to reduce the likelihood of detector saturation, and the issues that arise as a result (see Section 1.2.3).

If we instead consider the case where we are limited by readout and detector effects, then Equation 1.2.9 becomes simply

$$\text{SNR} \propto N_{\lambda} T_{\text{exp}} \sqrt{N_{\text{int}}} \quad (1.2.11)$$

since the readout noise does not depend on the incident flux, or the individual exposure

time. This therefore means that observing in longer integration blocks would be more beneficial than observing multiple times.

This does not, however, account for systematic effects in the detectors, and assumes that the noise can be added in quadrature. This is a good approximation in most cases, however we study this in detail in Chapter 3 to show that this is an over-simplistic approach, and the SNR becomes more complicated when including these additional noise parameters.

1.3 Telluric Absorption Features in the Infrared

Another aspect of this work looks at potential issues associated with correcting telluric absorption features in infrared astronomy. In this section we introduce the origin of these features, including typical absorption levels covering the wavelength range of HARMONI. We then cover two ways in which these can be corrected for: using a separate observation of a telluric standard star, or `molecfit`, which uses a radiative transfer code to create a synthetic spectrum to use for correction.

1.3.1 What are Telluric Absorption Features?

Telluric features arise as a result of molecular absorption in the Earth’s atmosphere, and are seen as absorption lines in a spectrum taken from an astrophysical target. The atmosphere is made up of a combination of gases, primarily 78% nitrogen (N_2), 21% oxygen (O_2), 1% argon (Ar), 0.04% carbon dioxide (CO_2), and trace elements of others (Allen, 2000). In addition, water vapour (H_2O) in the atmosphere will also lead to absorption features. The wavelengths and relative amounts of light absorbed depends on the gas molecule, however here we will primarily consider those that affect the near-infrared, as this is the region covered by HARMONI that will be the most impacted.

In Figure 1.8 we show a synthetic absorption spectrum covering the visible and near infrared out to $2.6\ \mu m$ (from Smette et al., 2015, based on Noll et al., 2012). We can see

from this that there are a few absorption features in the visible spectrum, however they increase significantly from $\sim 0.7 \mu\text{m}$. The primary absorption features in the infrared out to $2.45 \mu\text{m}$ (the upper limit of HARMONI) are caused by H_2O , with significant contributions from methane (CH_4), CO_2 , and O_2 .

As we discussed in Section 1.1.3, the atmosphere is not stable, and changes in the relative levels of the different gases will naturally lead to different amounts of absorption at the respective wavelengths. The most variable molecule is H_2O , which is directly related to the weather conditions at the time of the observation (Smette et al., 2015). The relative fraction of other molecules will also vary on different time-scales, from seasonal variations to hourly changes (e.g. Thoning et al., 1989; Schneising et al., 2009). It is therefore not possible to create a ‘one-spectrum-fits-all’ absorption profile, and the correction must be tailored to fit the specific observation.

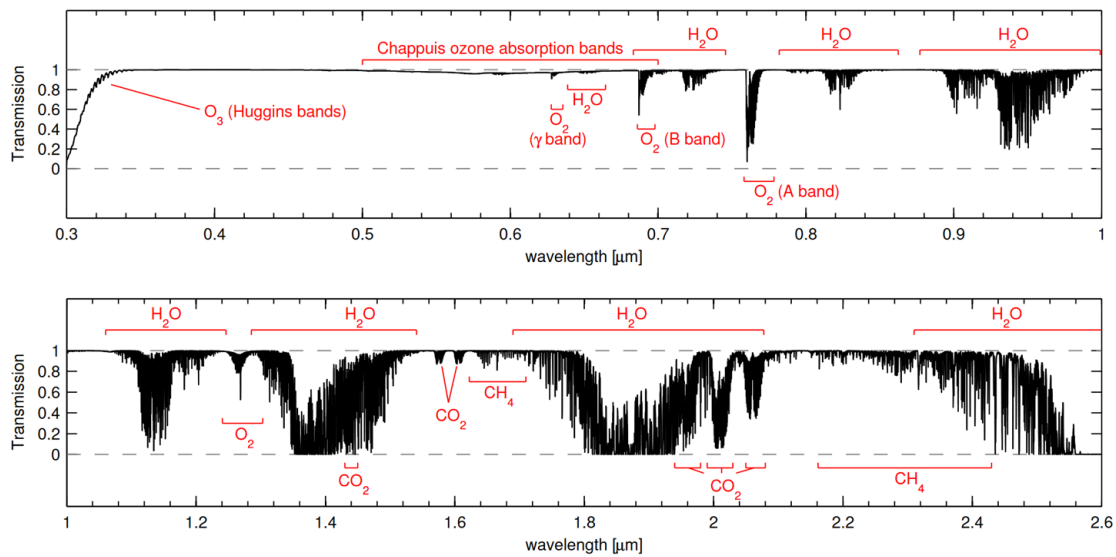


Figure 1.8: Absorption Features in the Visible and Near-Infrared. Synthetic absorption spectrum showing the predominant molecular absorption covering the wavelength range $0.3\text{--}2.6 \mu\text{m}$, at resolution $R \sim 10000$. Image reproduced from Smette et al. (2015), based on data from Noll et al. (2012).

1.3.2 Correction Methods

It is important to correct for this molecular absorption from the atmosphere, otherwise it will impact the recovered spectrum from the astrophysical target being observed. This is typically done by dividing the science spectrum by a telluric spectrum, such that the extraneous absorption features are removed. Here we introduce two methods for obtaining this telluric spectrum: a telluric standard star, and the `molecfit` telluric modelling tool, both of which we use in this work in Chapter 4.

Telluric Standard Star

One approach that can be used is to observe a different target with a known spectrum, ideally one with few intrinsic features, known as a telluric standard. Often this is an early-type star where the spectrum is mostly empty apart from hydrogen lines and some helium features (e.g. Walborn and Fitzpatrick, 1990). In situations where the hydrogen lines are of particular interest to the science, a G-type star is often used for the correction instead (e.g. Allen and Cragg, 1983; Cotton et al., 2014).

In addition to the spectral features, there are a number of additional constraints on the suitability of a star for use as a telluric standard. It must be able to be observed closely with the science target in terms of time, so as to ensure the atmospheric profile is as similar as possible between the two observations. Additionally, it must be close in terms of angular distance, as the absorption lines do not scale with the same function of airmass. The molecular mixing ratios will also vary slightly depending on the direction through the atmosphere at which the observations are taken. The chosen telluric standard must also be bright enough such that the recovered spectrum will be of sufficient SNR to provide a correction to the desired level (Smette et al., 2015).

The benefit of this method is that conceptually it is very straightforward to implement. The data from the observation are taken, reduced, and then divided from the science target, potentially with a small amount of post-processing to mask out the intrinsic features of the star. However, there are a number of drawbacks to this method, including:

- (a) Finding a suitable star close to the science target in terms of airmass;
- (b) Changes in the atmospheric profile between the science and telluric observations;
- (c) Insufficient SNR in the telluric star for adequate correction;
- (d) Unexpected features in the telluric star when reducing the data;
- (e) An additional cost in telescope time to take these observations, reducing the amount of time available to take the science data;
- (f) Differences in the line spread function (LSF) between observations.

The latter of these is the subject of the study in Chapter 4, and can particularly arise as a result of observing with good AO correction at a coarse spaxel scale.

Synthetic Spectra Creation, e.g. `Molecfit`

An alternative telluric correction method uses radiative transfer codes to create a synthetic spectrum that can then be divided out. This method was first shown to be possible by Lallement et al. (1993), who used the molecular data bank GEISA (e.g Chédin et al., 1982; Husson et al., 1992; Jacquinet-Husson et al., 2016) and forward modelling code STRANSAC (Scott, 1974) to correct for H₂O lines around the sodium doublet at 0.5895 μm . It was then demonstrated more extensively by Bailey et al. (2007), who used the Spectral Mapping Atmospheric Radiative Transfer (SMART; Meadows and Crisp, 1996) model to create a full synthetic atmospheric absorption spectrum in the J band (1.1–1.4 μm). This approach was advanced further by Seifahrt et al. (2010) to correct near-infrared spectra obtained with the CRIRES spectrograph (Käufl et al., 2004), who incorporated the radiative transfer code LBLRTM (Clough et al., 2005), based on FASCOD (Clough et al., 1981), along with the line database HITRAN (Rothman et al., 2009), meteorological data, and a model of the instrument LSF.

The tool which we will use in this work is `molecfit` (Smette et al., 2015; Kausch et al., 2015). This can simulate and fit atmospheric absorption features over a large wave-

length range (0.3–30 μm), incorporating a number of inputs to perform an accurate telluric correction. These include:

- (a) an atmospheric profile;
- (b) a molecular database;
- (c) a radiative transfer code;
- (d) an LSF model;
- (e) a fitting algorithm.

We will briefly cover the implementation of each of these components below.

Atmospheric Profile The atmospheric profile describes the temperature, pressure, and molecular mixing ratios at different heights at a given location. As we have determined, the relative mixing of atmospheric gases will affect the relative strengths of the absorption features, as will the temperature and pressure of the gas. A good atmospheric profile is therefore required to accurately determine the telluric features, and this is obtained in `molecfit` through a combination of a reference atmospheric profile, modelled 3-D data, and meteorological measurements.

An equatorial atmospheric profile containing molecular species up to an altitude of 120km (Smette et al., 2015; Dudhia, 2017) is merged with profiles from the Global Data Assimilation System (GDAS¹), which are linearly interpolated to the exact time of the observation. Lastly, `molecfit` can include specific information related to meteorological data at the time of the observation. Measurements taken and included in the FITS header can be incorporated, as can information regarding the precipitable water vapour (PWV) if it is known through independent calibration, for example through radiometer measurements (e.g. Kerber et al., 2012).

¹<http://www.ready.noaa.gov/gdas1.php>

Molecular Database In addition to an atmospheric profile, a database of molecular information is required. `Molecfit` uses an updated version of the HITRAN 2008 database (Rothman et al., 2009), which contains over 2.7 million spectral lines for 42 molecular species. When running `molecfit`, only molecules for which a mixing ratio is available in the atmospheric profile can be fitted, however this includes all the significant features in the SINFONI and HARMONI wavelength ranges. A detailed list of all the available molecules can be found in Smette et al. (2015).

Radiative Transfer Code The atmospheric profile, along with the molecules to include, is then used as an input to the line-by-line radiative transfer model (LBLRTM; Clough et al., 2005). This simulates the effect of atmospheric absorption and emission to produce a spectrum at a resolving power of approximately 4 million (Smette et al., 2015). This code is free and publicly available².

LSF model `Molecfit` can then broaden the output spectrum from the radiative transfer code by an LSF suitable for the instrument with which the observations were taken. This can be a combination of one or more of a boxcar function, Gaussian, or Lorentzian profiles. In addition, a fixed LSF can be provided as an input to `molecfit`, which is then used for the convolution. We utilise the ability of `molecfit` to fit the LSF in Chapter 4 as part of our investigation into the impact of an under-illuminated entrance slit.

Fitting Algorithm The fitting algorithm used in `molecfit` is the C implementation of the least-squares fitting library MPFIT (Markwardt, 2009), based on the FORTRAN fitting routine MINPACK-1 (Moré, 1978). This uses the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963), which, as with many fitting methods, finds local minimums. This does therefore require good initial estimates of the parameters, however enables fast and accurate convergence for complicated functions such as those presented by spectrum fitting.

²http://rtweb.aer.com/lblrtm_frame.html

The result of `molecfit` is therefore a synthetically generated spectrum which should accurately fit the data in terms of molecular abundance and LSF, as well as including the ability for additional corrections to continuum or wavelength calibration should they be needed. This spectrum can then be divided out from the science spectrum to provide the desired telluric correction.

This should provide numerous benefits over using a telluric standard star, although does come with its own drawbacks. If any of the input models are incorrect, then this will impact the overall accuracy of the fit. Additionally, if the science target has intrinsic features that are spectrally close the telluric features, there may be some blending of the two, potentially reducing the accuracy of the telluric correction.

In this work we use both methods to attempt telluric correction of a science target. In particular, we set up an investigation such that we expect a difference in LSF between the science target and the telluric standard star, using data taken with SINFONI. We also adopt an additional method, which uses `molecfit` in conjunction with a telluric standard star and separate LSF convolution to provide the correction. This work is presented in Chapter 4.

1.4 High Redshift Galaxy Kinematics

In this work we look at the capabilities of HARMONI to study high redshift galactic kinematics, in particular between redshifts 1 – 3. This is during an interesting era in the Universe’s history, around the peak of star formation (Madau and Dickinson, 2014). In this section we will introduce the concept of rotation curves, both from a theoretical viewpoint, as well as what is seen observationally. We explain their appearance and the implications that arise as a result of the difference between the observed rotation curves and the theoretical model. We also cover the importance of $H\alpha$ both as a tracer for star formation and for the galactic kinematics, as well as mentioning the limitations of this method.

1.4.1 Rotation Curves

One way in which galactic kinematics can be represented is through their rotation curves. These show the average rotational velocity of stars at increasing radii from the centre of the galaxy. It is possible to make a theoretical model for the expected rotation curve of a galaxy by considering a star of mass m in a circular orbit of radius r from the centre of the galaxy.

From Newtonian mechanics, to be maintaining a circular orbit, then the centripetal force

$$F_c = \frac{mv^2}{r} \quad (1.4.1)$$

must equal the gravitational force on the star, where v is the rotational velocity. We can then consider two primary locations for this star; in the bulge near the centre of the galaxy or far from the centre, in a Keplerian regime where the mass of the galaxy can all be assumed to be at smaller radii than the star.

In the first instance, we can model the bulge as a spherical volume, V of constant density, ρ . In this situation, from the shell theorem, the gravitational force on a star would be

$$F_g = \frac{GM(r)m}{r^2}, \quad (1.4.2)$$

where $M(r)$ is the mass of the galaxy enclosed within radius r . It is then possible to obtain the rotational velocity for the star by equating Equations 1.4.1 and 1.4.2 to get

$$\begin{aligned} \frac{mv^2}{r} &= \frac{GM(r)m}{r^2} \\ v^2 &= \frac{GM(r)}{r} \\ v &= \sqrt{\frac{GM(r)}{r}}. \end{aligned} \quad (1.4.3)$$

Using the equation for the mass of a spherical volume with constant density,

$$M(r) = V(r)\rho = \frac{4}{3}\pi r^3 \rho \quad (1.4.4)$$

we can substitute $M(r)$ in Equation 1.4.3 to give

$$\begin{aligned} v &= \sqrt{\frac{G}{r} \frac{4}{3} \pi r^3 \rho} \\ &= \sqrt{\frac{4\pi G \rho}{3}} r. \end{aligned} \tag{1.4.5}$$

As in this regime we have assumed that ρ is constant, we can see from Equation 1.4.5 that the rotational velocity is linearly proportional to the radius from the centre.

Now considering the second regime, where we assume the star to be in the outermost regions of the disk of the galaxy, with the majority of the stars closer to the centre, and as such we assume the mass to be concentrated closer to the centre as well. In this case, the gravitational force on the star is given by

$$F_g = \frac{GMm}{r^2} \tag{1.4.6}$$

where M is now the total mass of the galaxy and independent of r . We can equate Equation 1.4.6 to Equation 1.4.1 to give

$$\begin{aligned} \frac{mv^2}{r} &= \frac{GMm}{r^2} \\ v^2 &= \frac{GM}{r} \\ v &= \sqrt{GM} \sqrt{\frac{1}{r}}. \end{aligned} \tag{1.4.7}$$

From this we can see that rather than increase linearly with radius as we saw previously, the rotation velocity should decrease as $\sqrt{1/r}$. This therefore also implies a turnover radius at which the rotation velocity peaks, and this is shown as the dashed line in Figure 1.9.

However, observations do not match with this theory. Initially noticed in the 1930s with observations of the Coma cluster (Zwicky, 1933) and followed up in the 1970s with rotation curve measurements of a number of galaxies (Rubin et al., 1978), a discrepancy was found. Whilst the inner part of the rotation curve rises as expected to a turnover

point, it does not then fall off as would be predicted by Equation 1.4.7. Instead it appears to flatten out, as shown by the solid line in Figure 1.9. If we refer back to Equation 1.4.3, for v to be constant then

$$\frac{M(r)}{r} \stackrel{!}{=} C \quad (1.4.8)$$

where C is constant. To satisfy this, then $M(r)$ must be linearly proportional to r .

The conclusions drawn from this imply that there is a large amount of mass that we cannot see, present in the outer parts of the galaxy, known as the dark matter halo. Much work has been done to try and understand the profile of this dark matter halo (e.g. Navarro et al., 1996; Merritt et al., 2006), and it is now understood that every galaxy forms within a dark matter halo (Wechsler and Tinker, 2018).

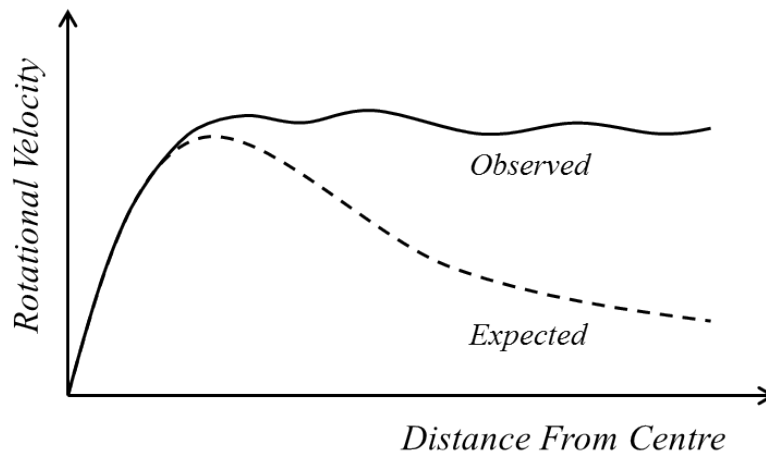


Figure 1.9: **Predicted vs. Observed Galaxy Rotation Curves.** Diagram showing the difference between the expected rotation profile of a galaxy (dashed line) with typical observations (solid line).

An area of active research is the variance in the ratio of baryonic (visible) matter to dark matter in galaxies, as this helps to constrain models for galaxy evolution. With the rotation curves observed, this suggests that the inner regions of spiral galaxies are baryon-dominated, while the outer regions are dark matter-dominated (e.g. Martinsson et al., 2013; Courteau and Dutton, 2015). Recent work has suggested the possibility that this relationship does not hold true at higher redshifts, and that rotation curves may in fact turn over and fall in the outer regions, as would be more expected in the absence of dark matter

(e.g. Lang et al., 2017; Genzel et al., 2017; Genzel et al., 2020). With instruments such as HARMONI, we will be able to obtain spatially and spectrally resolved observations of galaxies around the peak of star formation between redshifts of 1 to 3 (Madau and Dickinson, 2014). This will allow for direct measurement of individual rotation curves, from which our knowledge and models of galaxy evolution can be improved.

1.4.2 $H\alpha$ as a Tracer for Star Formation Rate and Kinematics

Star formation in nearby galaxies mostly occurs in giant molecular clouds (GMCs), large clouds of molecular gas in which small perturbations cause instabilities such that they exceed the Jeans mass (Jeans, 1902). Through processes such as turbulence and self-gravity, these instabilities grow and the GMCs fragment, with clusters of stars beginning to form at the point of collapse. As stars begin to form within these clouds, the largest, most luminous stars begin to ionise the remaining gas in the GMC into HII regions (e.g. Lada and Lada, 2003; McKee and Ostriker, 2007). When this ionised hydrogen recombines, a common de-excitation is from the $n = 3$ energy level to the $n = 2$ energy level and the photon emitted is known as $H\alpha$ emission.

If this $H\alpha$ emission can be measured, then we should be able to infer the star formation rate (SFR) from it. The standard method assumes a constant initial mass function (IMF; a description of the mass distribution of a population of stars) which then provides a relationship between integrated $H\alpha$ luminosity and the SFR (e.g. Kennicutt, 1983; Kennicutt et al., 1994; Osterbrock and Ferland, 2006). At rest, $H\alpha$ has a wavelength of 656.281 nm which falls towards the red end of the visible spectrum, however with observations of distant galaxies this is redshifted to longer wavelengths. We can use the definition of cosmological redshift

$$\lambda_{obs} = \lambda_{em}(1 + z) \quad (1.4.9)$$

with λ_{obs} and λ_{em} the observed and emitted wavelengths respectively, to determine the observed wavelength range for the $H\alpha$ line across galaxies at the peak of star forma-

tion. From this we can see that around the peak of star formation ($z \approx 1-3$; Madau and Dickinson, 2014) the observed wavelength for the $H\alpha$ line increases by a factor of 2–4, i.e. from 1312.562 nm to 2625.123 nm. These wavelengths now fall comfortably in the near-infrared, and all but the longest wavelengths fall within the observational range of HARMONI.

If we can now spectrally resolve the $H\alpha$ line, we can use the knowledge of the expected wavelength of the line to determine any shift from this, corresponding to gas motion within the galaxy. There is a limitation to this as a tracer for the kinematics, however, as it does assume that the gas motion is entirely due to ordered rotation of the galaxy, which is not always the case. If the gas does not follow the circular motion of the galaxy, for example due to shocks (e.g. Mortazavi and Lotz, 2019) or outflows (e.g. Roberts-Borsani et al., 2020), then this can impact the observed velocity, and thus recovered kinematics.

Also important to consider is the scale on which these processes occur. In the local Universe, these star-forming regions are found to have sizes of roughly 100 pc (Blitz, 1993; González Delgado and Pérez, 1997; Hughes et al., 2010). There is no issue spatially resolving galaxies to these scales in the local Universe, however things are more difficult at higher redshifts. If we consider a galaxy at a redshift of 2, we can determine the scale on sky of that galaxy. Assuming a flat Universe with Hubble constant $H_0 = 67.4 \text{ km s}^{-1}$ and matter density fraction $\Omega_m = 0.315$ (Planck Collaboration et al., 2020) we obtain a scale of $8.583 \text{ kpc ''}^{-1}$ (from Wright, 2006).

This means that the typical scale of a star-forming region would subtend of the order of 10 mas on sky. However, we are helped by the SFR being higher, meaning large regions of the galaxies will have active star formation, and as such current-generation IFUs are able to recover kinematics for galaxies at these redshifts (e.g. Genzel et al., 2006; Law et al., 2007; Glazebrook, 2013; Wisnioski et al., 2015). These are limited though, and even assuming the best seeing conditions and AO performance, are limited by the diffraction limit of the telescope, in the case of SINFONI being $\sim 50 \text{ mas}$ for the VLT, and to be able

to obtain a sufficient signal would require a large integration time. This imposes an upper limit on the spatial resolution of any recovered rotation curve, and doesn't provide the opportunity to investigate behaviour at the scale of the star-forming regions. Whilst these do not need to be resolved to obtain a rotation curve, they are of great interest for our understanding of star formation throughout the history of the Universe.

This is one of the major benefits of HARMONI and the ELT. With a larger light-gathering area and better spatial resolution it will be able to spatially resolve features within galaxies that were otherwise impossible. Being able to spectrally resolve the $H\alpha$ line at this scale will allow for investigation of star-forming processes at higher redshifts than has previously been possible, as well as obtaining improved rotation curves. As this is one of the science goals of HARMONI, we have tested the capabilities of this by simulating observations of a mock galaxy at redshift 1.44 and extracting kinematics and SFRs. This work is presented in Chapter 5.

1.5 Thesis Outline

There are three parts to this thesis, that all work towards the goal of improving the science that will be conducted with HARMONI. The first half of this work focusses on the infrared detectors that will be used for data collection, primarily looking at noise effects that will impact observations. In Chapter 2 we present an investigation into noise on the H2RG devices in place on KMOS, along with an engineering grade H4RG detector for MOONS (Cirasuolo et al., 2012). For both of these we show our analysis techniques and present our findings, highlighting the similarities and differences between the two.

In Chapter 3 we then use these results, and combine them with a pre-existing tool to improve the capabilities of `HSIM` to simulate detector noise. We then detail a number of simulations that we conducted to test this additional noise, along with our data analysis process. At the end of this chapter we give our results, alongside a discussion of the dominant noise components that are of most importance to consider when reducing real

HARMONI data.

We shift our focus slightly in Chapter 4 to investigate another potential issue that could arise when observing with HARMONI. When observing with the ELT, the input slit of HARMONI could be significantly under-illuminated when using adaptive optics to observe a point source. This could result in a telluric standard star under-illuminating the slit, but these data then being used to correct an extended object that does fill the slit. In this chapter we discuss work we have done to test this, by using SINFONI to take data that replicates this effect. We then show our data reduction and analysis, and present our findings at the end, discussing the potential impact for HARMONI observations.

In Chapter 5 we look at the potential performance of HARMONI as a whole, using HSIM to create simulated observations of a galaxy at redshift 1.44. We outline the input cube that we use, and then the various analysis methods and data reduction we perform. We show the recovered kinematic profiles for a range of simulations, along with a discussion of the impact of the AO PSF on our results.

Finally, we give our overall conclusions in Chapter 6, along with potential extensions to the work that can be continued in the future. We end with a brief guide detailing how to use the new features of HSIM in Appendix A.

Chapter 2

Study of Detector Noise on Teledyne HxRGs

In this chapter we present the work carried out to more accurately characterise the noise profiles of Teledyne HxRG detectors. Initially we explain the reasoning behind using data from KMOS, and then describe the analysis methods used to obtain noise profiles for some of the detector noise. We then present the results of this, primarily in the form of distribution maps of means and standard deviations on a pixel-by-pixel basis. We then repeat this analysis with data taken using an engineering grade H4RG and compare with our findings from the H2RGs. At the end of this chapter, we give our conclusions from this study, highlighting the importance of this in regards to science prediction with HARMONI.

2.1 Purpose of Investigation

There are two primary purposes for conducting this investigation:

- (a) to enhance our understanding of detector noise. As we build larger and more powerful instruments, it is important to understand issues that can arise throughout the entire instrument so that we can conduct the best science. A key part of this chain

is the detectors, with which all the data are collected. HxRG detectors, which are going to be used in HARMONI, are known to have different forms of noise, some of which can not be reduced by simply observing for longer. In this study we will focus on looking at these, and consider ways in which we can account for this.

- (b) to improve the `HSIM` science simulation tool. Prior to this work, `HSIM` only included detector noise in a very simple sense, merely adding stochastic read noise and dark current evenly across the detector. We know this is overly simplified, therefore one focus of this work is to improve this through a combination of the results of this study, and incorporating elements of the Teledyne noise generator into `HSIM`. This allows for more accurate science prediction, where the impact of non-stationary detector noise terms can be accounted for.

2.2 Analysis of KMOS Darks

2.2.1 Choosing Data from KMOS

To begin this study we first needed to obtain a sufficiently large dataset to analyse. Although HARMONI will use H4RG detectors, at the time of this work there were no instruments using these, therefore we looked at H2RGs instead. We made an assumption here that they will behave in a similar way to the H4RGs, and in Section 2.3 we tested this assumption with an engineering grade H4RG. As we were interested in detector noise we chose dark exposures, where any ‘signal’ detected should be noise. The K-band Multi Object Spectrograph (KMOS) is an instrument installed on the European Southern Observatory’s (ESO’s) Very Large Telescope (VLT) and with three cryogenically cooled H2RGs has similar operation to that planned for HARMONI. As this instrument has been operational for many years, there were sufficient data available in the ESO archive that we could download and analyse for this study.

To get a sufficient understanding of the noise, we downloaded every 60 s dark expo-

sure taken with KMOS in October 2016 totalling 165 darks. This gave us a sufficient sample size such that we could account for stochastic processes, which we know will also be present, and focus on systematic effects. We also downloaded every exposure from October 2015, totalling 195 exposures, to conduct the same analysis, so as to ensure that our findings are consistent over long time scales.

2.2.2 Characterising Noise

We show in Figure 2.1 an example of one of the KMOS darks that made up our dataset, with each of the three detectors shown. Despite this being a dark observation, a number

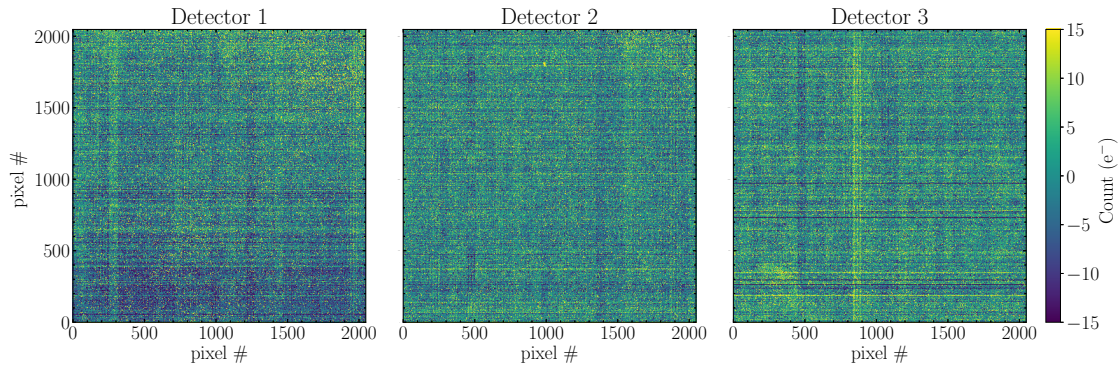


Figure 2.1: **Raw KMOS Dark Exposures.** An example of one of the sets of darks used in this analysis, namely `KMOS.2016-10-04T09:55:31.789`. This is a 60 s exposure with no post-processing from the ESO archive, aside from converting from ADUs to electrons. All detectors are shown on the same scale.

of features can clearly be seen:

- (a) **Horizontal banding.** Small $\sim 1/f$ drifts in bias voltages cause some horizontal banding, that the reference pixels then attempt to correct for. However, bad pixels amongst the reference pixels can also cause more pronounced horizontal features. These are prominent across all detectors, but particularly the lower part of detector 1.
- (b) **Vertical banding.** Pedestal drifts, which are residual bias offsets between integrations, cause slight variations between amplifier blocks. This leads to the 64-column

wide vertical features seen, which will vary between integrations. There are amplifier blocks on all detectors which exhibit this behaviour.

- (c) **Alternating Column Noise (ACN).** The effect where alternating columns within an amplifier block have higher/lower base noise values. This is particularly noticeable on one of the amplifier blocks on detector 3.
- (d) **Hot pixels.** Some pixels are ‘hot’, meaning they always return a significantly higher signal than would be expected. Pixels that are always ‘hot’ can be mapped and excluded in data reduction, but therefore reduce the usable data obtained. The upper right corner of detector 1 has a high concentration of hot pixels.
- (e) **Hotspot.** There is a cluster of hot pixels in the upper centre of detector 2, known as a hotspot. All these pixels can be removed in data reduction, but any science data that falls in that region will be lost.

KMOS uses reference pixel subtraction to attempt to correct for the horizontal $1/f$ banding, however this is done during the readout process on each read. This means that it is not possible to perfectly undo the reference pixel correction, as the data we have are the end result of the sample-up-the-ramp (SUTR) process. However, as we are looking at darks there should be no signal on the detectors, thus we can make an approximation of the total horizontal reference pixel correction by subtracting the median value of each row in turn. We repeated this for the columns, instead taking each block of 64 columns corresponding to a single amplifier block, and subtracting off the median of all the odd columns and the even columns separately to account for ACN. We then multiplied through by the gain on the detectors, equal to $2.10 \text{ e}^-/\text{ADU}$ to obtain the number of electrons per detector pixel.

An example of a resulting dark image is shown in Figure 2.2. This has removed almost all the horizontal structure, and the majority of the vertical features. We do see that one amplifier block still has a lot of vertical structure, therefore to ensure that the rest of the analysis was as accurate as possible, we excluded this. However, we keep the data from

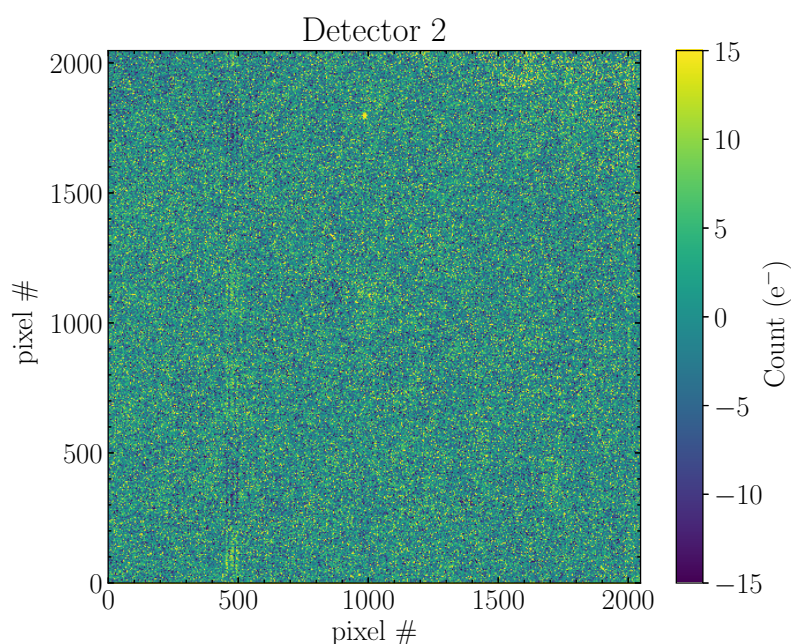


Figure 2.2: **Sample Processed KMOS Dark Exposure.** An example of a KMOS dark exposure with our post-processing applied, showing detector 2 of the `KMOS.2016-10-04T09:55:31.789` observation. This has removed the majority of the features seen in Figure 2.1, although we do see that one amplifier block around the 500 pixel column has not subtracted very well. To ensure that our results are as accurate as possible, we therefore excluded this amplifier block from further analysis.

the hotspot, as it is not uncommon for HxRG detectors to have such grouped areas of hot pixels. With this complete, the next step was to ascertain the behaviour of the detector pixels.

To do this, we took each pixel on each detector and compared it with the same pixel on the other dark exposures in our sample. In Figure 2.3 we show a typical distribution of values returned by a pixel on the detectors. To make this we took the counts from the same pixel on a detector across each of the 160 frames and binned them into 2 e^- wide bins, where the uncertainties on each bin are given by counting statistics. We can see visually that this follows a Gaussian distribution, and we then used a non-linear least-squares fitting method from `SciPy` to obtain a best fitting Gaussian function, with the amplitude, standard deviation, and central value as free parameters. As we see, it provides a good fit to the data, and from this we could extract a mean and standard deviation for the pixel.

However, due to the volume of data, it was not possible to store all the exposures

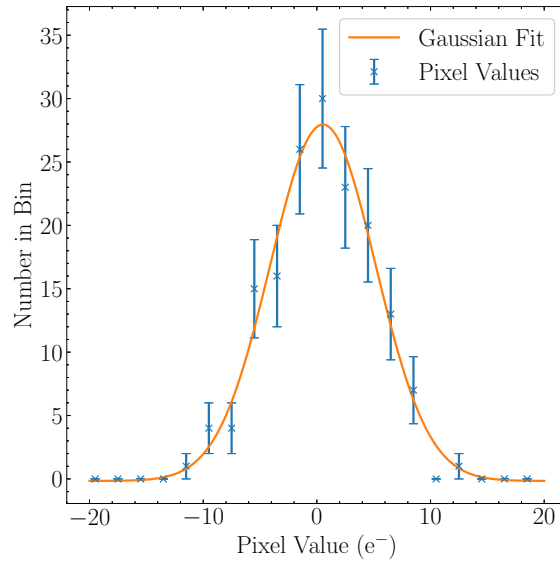


Figure 2.3: **Histogram of Values Returned by an Individual KMOS Detector Pixel.** Typical distribution of electron counts binned into bins of width 2 e^- showing the counts (blue crosses) with uncertainties given by the square root of the number of pixels in each bin, and a best-fitting Gaussian (orange line)

in memory simultaneously, and doing this process for each pixel would be very time consuming. Therefore an implementation of Welford’s algorithm (Welford, 1962) was used to calculate a running mean, μ , and standard deviation, σ , for each pixel.

These define ‘base maps’ for the KMOS detectors, and are shown in Figure 2.4. The mean value here is the average value a pixel will return, which given that we have subtracted off the median value we would expect to return around 0 e^- . The standard deviation measures the spread around the mean value, and thus gives us an indication of how close to the mean value a pixel will be each time it is measured. This is similar to the determination of read noise of the detector, which for the KMOS detectors is 9 electrons for short integrations, dropping to 2.6 electrons for long exposures with Fowler sampling. We would expect the standard deviation of the pixels to be around this value, likely a little higher than the minimum due to the exposure time, sampling method, and residual noise sources such as dark current.

In an ideal world, and as is often considered when modelling infrared detectors, each pixel is assumed to have the same read noise and on average subtract to 0 e^- . However, we know in practice that this is not the case, as we can clearly see from Figure 2.4 there are a

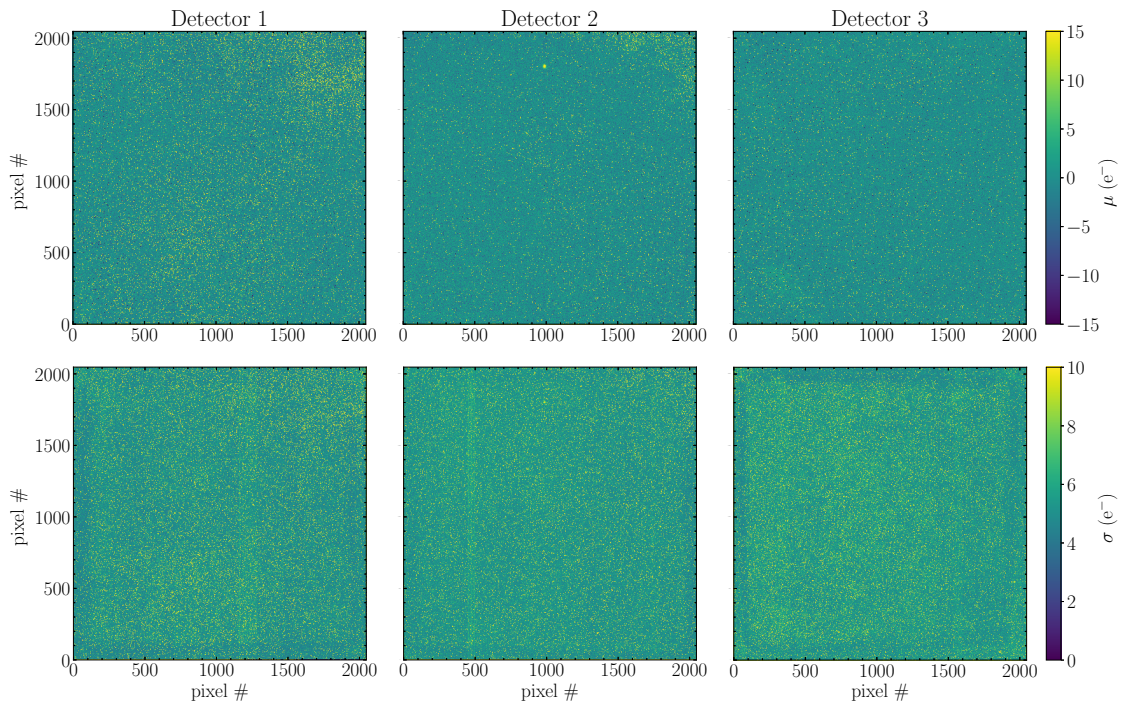


Figure 2.4: Base Mean and Standard Deviation Maps for KMOS Detectors.
Top: Base maps for the mean pixel values of the KMOS detectors. The hotspot on detector 2 can be seen clearly, along with regions of hot pixels on each detector. *Bottom:* Base maps for the standard deviations of each pixel. This brings out the picture-frame effect, particularly on detectors 1 and 3. Along with the amplifier block around column 500 on detector 2, we also see that two amplifier blocks around column 1250 on detector 1 have additional noise that has not removed, therefore those are also excluded from further analysis.

number of pixels that do not behave like this. From these maps, we made a distribution of means and standard deviations for the photosensitive pixels on each detector, which are shown in Figure 2.5. This has excluded the columns mentioned in the figure captions, but kept hot pixels and hot spots.

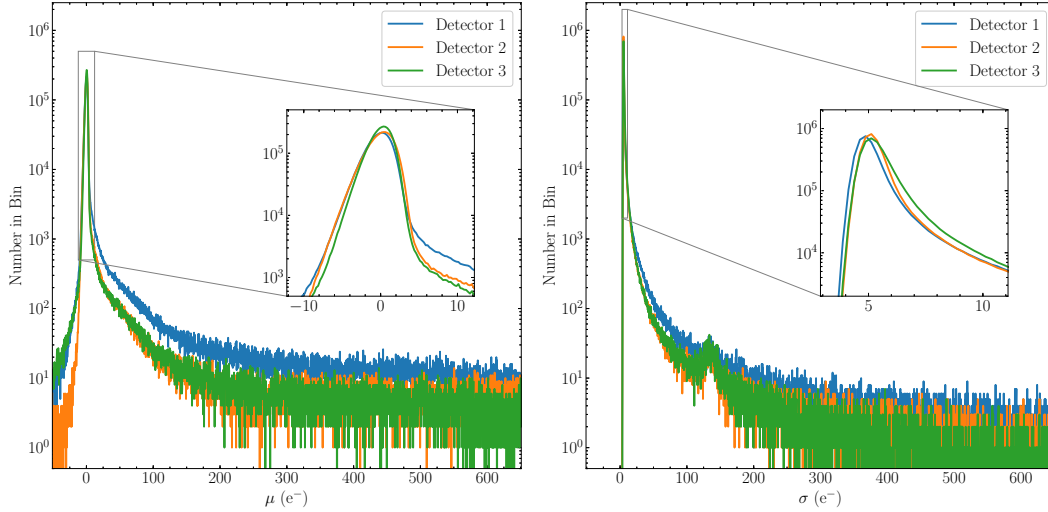


Figure 2.5: Distribution of KMOS Detector Pixel Means and Standard Deviations. *Left:* Distribution of pixel means on KMOS detectors, binned into 0.25 e^- bins. Inset shows the shape of the distribution around the peak. *Right:* Same, for the pixel standard deviations. Bin sizes are fixed at 0.25 e^- in both panels.

We see here that as expected not every pixel behaves in the same way. What is interesting however, is that each detector follows a similar distribution. For the mean pixel values, they are distributed around 0 e^- as we would expect. However they have a long tail to high values which is reflective of a number of hot pixels on the detectors. We also see that detector 1 has a larger number of pixels with high average counts, and this can be also be seen from Figure 2.4 with a greater density of hot pixels.

We also see that the standard deviation of each pixel follow a distribution of values that is very similar for each detector. On each of them they have a peak value of approximately 5 e^- which is a little higher than the best-case read noise of the KMOS detectors, but in agreement with our expectations. Interestingly, we also see a small bump in the distribution around 130 e^- which is present in all detectors. These pixels are randomly distributed across the detector, therefore this is not a result of a hotspot or other such

feature.

These distributions are important as they give us an understanding of the detectors on a pixel-by-pixel basis, which we could then use to construct our own simulated detectors for incorporation into HSIM. We take from these that each detector, whilst independent, behaves in a similar way when looked at for an individual pixel. Pixels that do not return to 0 e^- will always have some residual noise that remains after averaging multiple exposures, while pixels with high standard deviations will vary greatly between exposures and thus return unreliable data.

2.2.3 Comparison with October 2015 Data

In order to ensure that the distributions obtained in Section 2.2.2 were truly reflective of the detectors and not a result of chance, or were strongly varying with time, we repeated the analysis with another set of dark exposures from October 2015. We once again excluded the same amplifier blocks to ensure consistency, and show the resulting mean and standard deviation distributions in Figure 2.6.

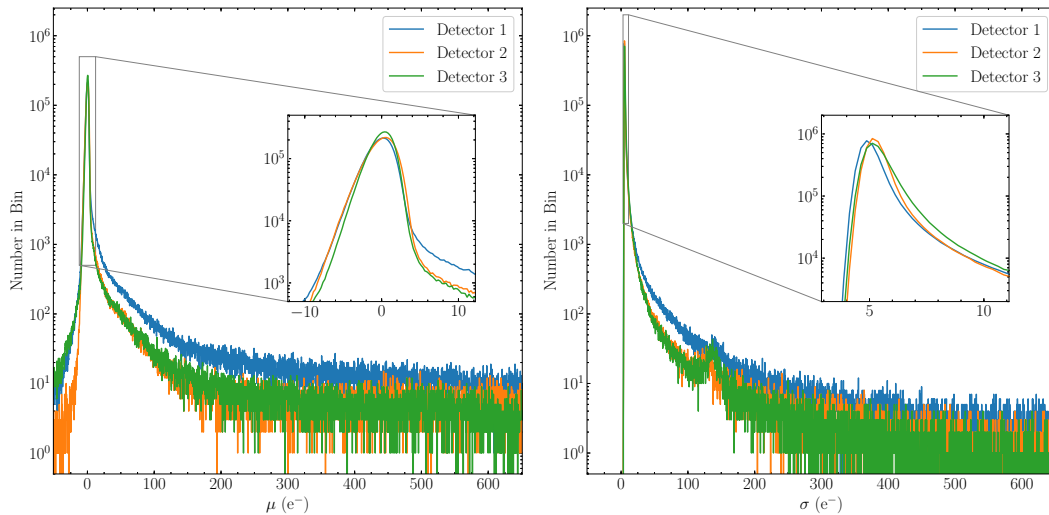


Figure 2.6: Pixel Distributions from October 2015 Dark Exposures. *Left:* Distribution of pixel means on KMOS detectors from the set of 60s dark exposures from October 2015, binned into 0.25 e^- bins. Inset shows the shape of the distribution around the peak. *Right:* Same, for the pixel standard deviations.

These distributions are almost identical to those in Figure 2.5, with the average mean

value being around 0 e^- and the average standard deviation being around 5 e^- . We once again see long tails to high values of both means and standard deviations as well. This tells us that the distributions we have obtained are valid for these detectors over long time scales, and are therefore a good starting point to use for our simulated detectors. We also investigated whether the pixels themselves remain statistically the same over that time, or if each pixel behaves differently, but when taken together return those distributions. To do this we plotted the values obtained from the October 2015 dataset against the values obtained from the October 2016 dataset, with an example of the resulting plots shown in Figure 2.7. For the means and standard deviations of each detector, we also calculated the correlation coefficient, R , given by

$$R = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2}} \quad (2.2.1)$$

where x_i and y_i denote the October 2015 and 2016 datasets, and $\bar{}$ denotes the mean of that sample. These are summarised in Table 2.1.

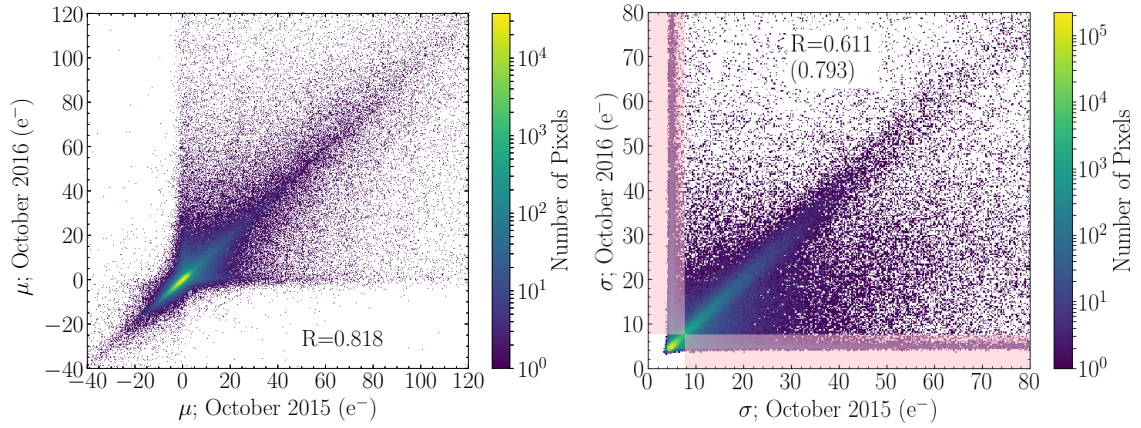


Figure 2.7: Correlations Between Pixel Means and Standard Deviations on KMOS Detector 1. *Left:* 2D histogram showing the correlation between the mean pixel values from detector 1 in the October 2015 dataset with those from the October 2016 dataset. Lighter colours denote higher density. The correlation coefficient is calculated as $R_\mu=0.818$. *Right:* Same, for the pixel standard deviations, with correlation coefficient $R_\sigma=0.611$. When the region shaded in pink is excluded from the calculation, we recover $R_\sigma=0.793$.

We can see in the comparison of the means between the two datasets that the majority

of data lies along the line $x=y$, meaning they return the same value in both months. This is true for both the pixels around 0 e^- , as well as those with higher and lower values. There are a smaller number that return close to 0 e^- in one dataset, but a different value in the other, which is likely due to spontaneous noise events such as cosmic hits grossly skewing the average value of a pixel in a given month. There is a similar number in both wings, therefore from this we can say that a given distribution such as that in Figure 2.5 represents a snapshot of the pixel means on a detector.

Detector #	R_μ	R_σ
1	0.818	0.611 (0.793)
2	0.799	0.595 (0.859)
3	0.844	0.610 (0.927)

Table 2.1: **Correlation Coefficients for KMOS Detectors.** Correlation coefficients for the means (R_μ) and standard deviations (R_σ) of each KMOS detector between October 2015 and October 2016. Values in brackets denote correlation coefficients as a result of removing outlying values from the standard deviation data.

Regarding the comparison between the standard deviations, we see that the correlation is tri-modal, with a large proportion falling along the $x=y$ line, but also a significant proportion that fall into wings where they return close to 5 e^- in one month and not the other. This tells us that as with the means, largely a pixel will have the same standard deviation over time. This allows for identifying pixels which always have a high standard deviation to be flagged as bad pixels. Pixels that are harder to account for, however, are those that have a high standard deviation in one dataset and not the other. These are possibly pixels which have cosmic hits which skew the distribution, or could be ‘transient’ bad pixels, which behave like regular pixels sometimes, and bad pixels at others. What is important for this work however, is that there are a similar number of these in each dataset, so as with the means we can say that these distributions represent a snapshot of the detectors.

As for the correlation coefficients quoted in Table 2.1, the values are similar for each detector at around 0.8 for R_μ and 0.6 for R_σ . These are the same for all detectors, suggesting that they all behave in a similar manner. This is in line with what we see in Figure 2.7

as the means have a tighter correlation than the standard deviations. However, as we have discussed, a large proportion of the pixels not falling along the $x=y$ correlation fall along a line corresponding to the average value of the dataset within a given month.

To determine the impact of these transient bad pixels on the correlation, we removed certain pixels on the basis of two selection criteria and re-evaluated the correlation coefficient, quoted in brackets in Table 2.1. These criteria were either a pixel with extremely high standard deviation in a given month (greater than 20 times the median for that month), or a pixel which returned a value of less than 1.5 times the median standard deviation in one month, but greater than 1.5 times the median in the other. This corresponded to 4.0%, 2.6% and 3.0% of all pixels on detectors 1 through 3 respectively. We can see that as a result the correlation coefficients increase significantly, with all but detector 1 now being greater than 0.85. This confirms our hypothesis that a small number of outlying pixels were skewing the correlation, and therefore we can safely say that the pixels are strongly correlated over the timescale of at least a year.

2.3 Analysis with MOONS Engineering Grade H4RG

2.3.1 Data Acquisition

In addition to the data downloaded from the ESO archive for H2RGs, we also obtained data using an engineering grade H4RG device for the Multi Object Optical and Near-infrared Spectrograph (MOONS) instrument, also for the VLT (Cirasuolo et al., 2012). MOONS is a fibre fed multi-object spectrograph covering a wavelength range from 0.6-1.8 μm with two spectral resolutions, one of $R \sim 5000$ and one of $R \sim 20000$. Due to the large number of fibres (~ 1000) MOONS has two spectrographs, and for the infrared YJ and H bands the collected light is fed to two H4RG devices on each spectrograph. Prior to receiving and characterising the science grade devices, Teledyne supplied an engineering grade H4RG detector on which we were able to obtain data for this work.

With this, we were able to recreate darks taken using the same acquisition process as

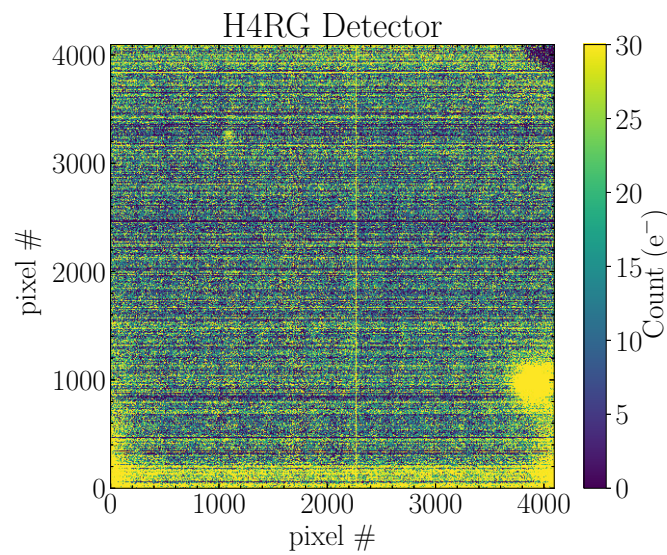


Figure 2.8: **Example MOONS H4RG Engineering Grade Detector Dark.** An example of one of the dark images taken using the MOONS engineering grade H4RG detector. It is evident from this that the set-up was imperfect, and the slightly raised average counts suggests a low level of global illumination across the detector. In addition to this, features such as a large glowing hotspot and a defective corner can also be seen.

KMOS, and for the same duration. These were taken with the detector operating cryogenically at 40K, the same temperature at which the KMOS detectors are run at. This should ensure that temperature dependent effects such as dark current, kTC noise or reciprocity failure (see Section 1.2.3), which can alter the appearance of the detector, are similar to what would be seen in the KMOS detectors. We then took 70 60 s dark exposures, repeating the readout method and reference pixel subtraction carried out by KMOS. We show in Figure 2.8 an example of one of these exposures, and we see in this a number of notable features corresponding to issues with the detector:

- (a) **Large Hotspot.** A very bright hotspot in the lower right corner of the detector spanning a region a few hundred pixels across.
- (b) **Small Hotspot/Column Defect.** A smaller hotspot in the upper left. The column corresponding to the centre of this hotspot is also affected, with it returning higher values than the neighbouring columns.
- (c) **Damaged Corner.** The top right corner of the detector is damaged, leading to a

region with no photosensitive pixels, and a larger concentration of hotspots in the whole corner.

- (d) **Read Glow.** The bottom of the detector has increased glow near the readout circuitry, increasing the level of noise in that region.
- (e) **Other Features.** In addition to these features that are unique to this detector, other effects common to HxRG devices are clearly present, such as significant horizontal banding, along with a number of individual hot pixels.

2.3.2 Results

We repeated the analysis described in Section 2.2.2 to create a base map for this detector, and this is shown in Figure 2.9. What is immediately clear from this is that the detector is not as clean as the KMOS detectors. This is in line with expectations however, as the experimental set-up was not perfect, and as an engineering grade detector has a number of issues as described above. We see from this base map that the regions around the hotspots and towards the bottom of the detector return pixels with higher means, and to a slightly lesser extent higher standard deviations. In addition we have no usable data from the upper right corner which is damaged. We also see that systematically across the detector we see lower mean values and higher standard deviations than on the KMOS detectors. Higher standard deviations means a greater variance between observations, which is to be expected with both the set-up and detector. Lower means, many below zero, is likely a result of our analysis pipeline, as we subtract off the median of each row of photosensitive pixels. Whilst the median is more stable than the mean, in a detector with a large number of hot pixels, stray light, or other residual noise, the median will be skewed slightly higher and thus we get an overall lowering of the mean in our maps.

From the base mean map of this detector, there are number features that we see that were not immediately expected. We see an unexplained discontinuity around the upper left hotspot, with the base map having lower mean values above it then below it. We

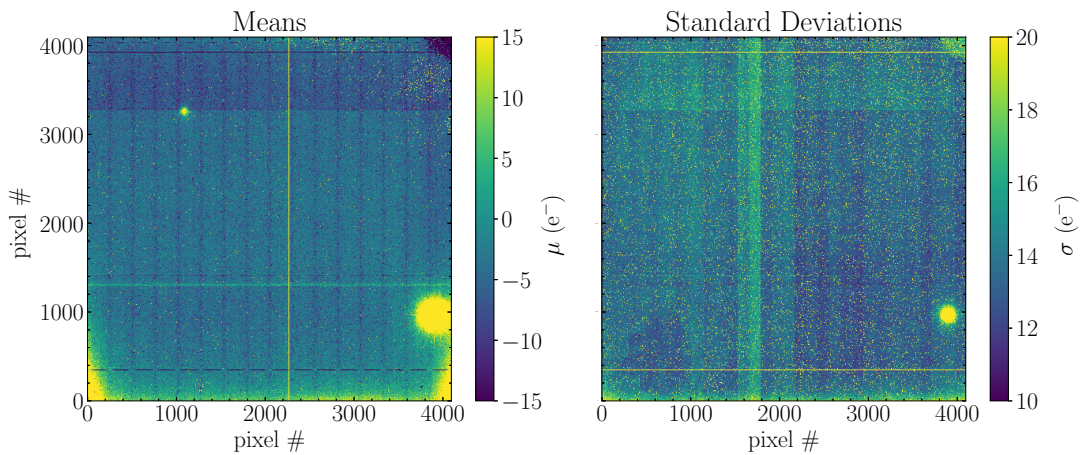


Figure 2.9: **Base Mean and Standard Deviation Maps for MOONS H4RG Detector.** *Left:* Base map for the pixel means on the engineering grade MOONS H4RG detector. Many features are prominent here, such as the hotspots, damaged corner and readout glow, along with various horizontal and vertical residual features. *Right:* Base maps for the standard deviations of each pixel. The larger hotspot can clearly be seen, as well as residual differences in amplifier bias and some rows of pixels that are bad.

also see another horizontal feature around row 1300 with a smaller band just above it, and a particularly hot column around 2250. We also see vertical ‘striping’ that is not correspondent to the amplifier blocks. This is likely due to the detector not being entirely stable, as is seen in Ives et al. (2020), as these were not taken with current boost operation, the detector is very unstable in the first half an hour after powering on, and many of these features are seen as a result.

We do see some additional residual features in the standard deviation map as well, beyond that caused by the large hotspot and readout glow. We see that even with the pedestal drift correction we apply, different amplifier blocks have systematically higher or lower standard deviations. This again is likely due to stability, where some amplifier blocks have greater drifts than others, and thus the data from each column block varies by larger amounts between reads. Interestingly, we do not see the hotspot in the upper left of the detector present in the standard deviations, although we still see a slight discontinuity above and below the line of the central pixel of it. This means that this hotspot is consistently returning the same value on each pixel, and so whilst it is hot, data taken on this region of the detector could still be used somewhat reliably, albeit with higher shot noise

as a result of the increased noise. This is dependent on a pixel not being saturated, which is the case for this hotspot.

Despite these issues, the base maps are significantly cleaner than the raw observations, and we therefore continued the analysis in the same manner as on the KMOS H2RGs, creating distributions for the means and standard deviations, which can be seen in Figure 2.10. However, we do exclude a large amount of the detector when doing this to ensure our results are as reliable and consistent as possible. Initially, we only use data from the central 2044×2044 region of the detector.

This serves two purposes: firstly this excludes the amplifier glow, damaged corner, large hotspot and the horizontal discontinuity, and secondly this means we are using approximately the same number of pixels as for each KMOS detector, so we can directly compare our distributions. We then mask out further regions that are clearly not representative of the detector as a whole, which are two columns each side of column 2265, thirty rows each side of row 1310, and five rows each side of row 1410. We also remove column 1093, as this is the column containing the upper left hotspot, and returns anomalously high readings.

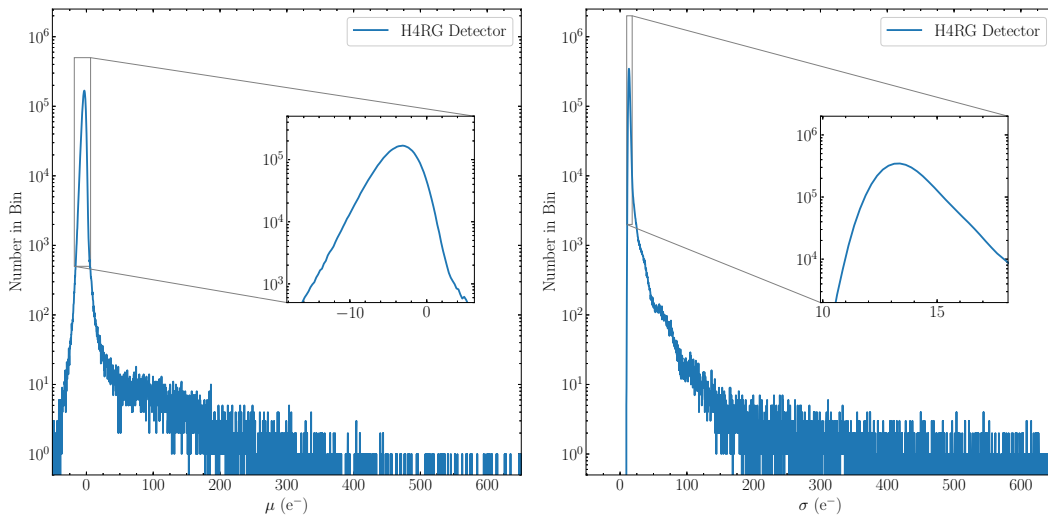


Figure 2.10: Distribution of Detector Pixel Means and Standard Deviations on the MOONS H4RG. *Left:* Distribution of pixel means on the engineering grade MOONS H4RG detector, binned into $0.25 e^-$ bins. Inset shows the shape of the distribution around the peak. *Right:* Same, for the pixel standard deviations.

What we see from these distributions is that despite the issues and instabilities with the detector, the overall shape of these distributions is very similar to those seen in Figure 2.5. For the means, as we saw from Figure 2.9 the average value is less than 0 e^- , at around -3 e^- , however this is less than 1 ADU and as explained, is likely an artefact of a large amount of noise that is skewed higher than in the KMOS detectors. The shape of the distribution, however, is almost identical to that of the KMOS detectors, with a spread either side of the peak, and a tail to high values. This tail is actually noticeably smaller than that for the KMOS detectors, which is potentially due to only using the central portion of the detector for this analysis, as we avoid issues such as the picture frame effect, that affect the edges of HxRG devices.

For the standard deviations, we see a much higher peak than for the KMOS H2RGs, at approximately 13 e^- . If we are using this as a proxy for read noise, this is higher but not dissimilar to 60 s unbuffered sample-up-the-ramp exposures using the same detector, which is approximately 8 e^- (Ives et al., 2020). Given that these data were taken with an inferior set-up to that in Ives et al. (2020), prior to full characterisation of the detector, we can be confident that we are indeed seeing a distribution that is reflective of the read noise on individual pixels. We again see a peaked distribution, with a large tail to high standard deviations. There is a small ‘hump’ around 100 e^- , although it is barely distinguishable from the noise, suggesting whatever feature seen in the KMOS detectors is likely an artefact of either that set-up or those particular detectors, rather than a feature of the HxRG devices.

There are a couple of slight bumps in this distribution that we do not see in the KMOS devices, one around 30 e^- and one around 60 e^- , corresponding to a few thousand pixels at each. We can see from Figure 2.9 that there are a few amplifier blocks which consistently return slightly higher standard deviations, therefore we would expect to see corresponding features in our distribution.

2.4 Discussion and Conclusions

In this investigation, we have developed a method of producing cleaner base detector maps for the purposes of investigating their noise properties. By taking a number of darks such that we can perform statistical analysis on them, we extracted the mean and standard deviation of each pixel on each of the three KMOS H2RG devices, as well as an engineering grade H4RG device for the MOONS instrument.

For the KMOS detectors, we have seen that on a pixel level the detectors behave very similarly and in particular the standard deviations of the pixels follow almost identical distributions for each detector. We have also seen that these distributions are stable over time frames of at least a year, with each pixel returning a similar value to that on data from a year previously. This is also reflected in high correlation coefficients, with the majority of pixels that don't fall along the line of correlation likely being due to transient effects.

When we looked at the H4RG engineering grade device, despite a number of issues with the detector itself, as well as the set-up, we were able to obtain a sufficient dataset to perform the same analysis. Repeating the same data reduction pipeline, and removing the most affected areas of the detector from our analysis, we again obtained distributions for the means and standard deviations of the detector pixels. Whilst the values of the peaks of these distributions differed, as would be expected, the shapes were very similar to that of the H2RG detectors.

The implication of this is that a random pixel on a HxRG device can be assumed to follow the distributions seen above, for both its mean and standard deviation, which we assume to be a proxy for read noise. This means that when modelling detectors and creating mock observations with HARMONI, to obtain a more reliable simulation we should consider the impact of this on our science results. Typically it is assumed that all the pixels across a detector behave the same, with read noise values following a normal distribution around the average. We have shown here that this is not the case, and therefore in the next chapter we will discuss our methods for improving the existing HxRG detector simulation tool, and its integration into `HSIM`.

Chapter 3

Testing Advanced Detector Systematics using Simulated Detectors in **HSIM**

In this chapter we summarise an existing tool for HxRG detector simulations, the Noise Generator (NG) tool. We will then explain the improvements that we have made to this tool, and describe the process involved in modifying **HSIM** in order to incorporate this into it. We will then show tests that we conducted to ascertain the impact the advanced detector systematics (ADS) may have on the science with **HARMONI**. We simulate point sources at different places within the field of view, altering the levels of ADS noise. We test the limiting sensitivity, repeatability of recovering the same signal over multiple observations, and variations in the SNR. At the end of this chapter, we summarise our findings, focussing on the implications for **HARMONI** and emphasising the importance of characterising the detectors in the instrument in order to achieve the best possible results.

3.1 Creating Simulated **HARMONI** Detectors

3.1.1 Teledyne HxRG Noise Generator

There is a standalone tool for simulating infrared detectors, known as the noise generator (NG; Rauscher, 2015). This is written in `Python-3` and aims to simulate a number of

3.1. CREATING SIMULATED HARMONI DETECTORS

the effects that we have previously described, namely ACN, correlated and uncorrelated $1/f$ noise and picture frame noise. It was initially developed for the Near Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope (JWST), however was built such that parameters could be tuned accordingly for other HxRG detectors. NG has the option of creating either a 2-dimensional noise image for a detector, or a 3-dimensional cube that simulates every read. For the purposes of this work, we chose the former, as this provides a good estimation of the noise and is considerably faster to produce, which is important when implementing this within `HSIM`.

Noise Component	Nominal Value (e^-)
Read Noise (long exp.)	2.845
Read Noise (short exp.)	12.0
Uncorrelated Pink Noise	1
Correlated Pink Noise	3
ACN	0.5
Picture Frame	0.2

Table 3.1: **Nominal ADS Noise Levels in `HSIM` v.310.** Default amounts of each noise component included in the ADS mode of `HSIM`. Noise component levels other than read noise are based off of investigation of the JWST detectors, as presented in Rauscher (2015). The average read noise is taken as the technical specification requirement as of September 2015. All components are implemented such that they can be adjusted to suit a particular simulation, however, and can be modified if more up-to-date data becomes available. Full details for how to do this can be found in Appendix A.

In Figure 3.1 we show the internal steps within NG that are followed to create a 2-dimensional detector image (adapted from Rauscher, 2015). We show the nominal values for each of the noise components in Table 3.1. The value for the read noise on long and short exposures is taken as the technical specification requirement as of September 2015. This is likely more optimistic than work by Ives et al. (2020) has shown is possible with the H4RG devices, however the analysis in this chapter was completed prior to this. The nominal values for the other components are based off of detailed principle components analysis (PCA) of the devices for NIRSpec as described in Moseley et al. (2010) and Rauscher (2015). These are assumed to be realistic for all HxRG devices, particularly for the purpose of this work, and can be updated accordingly when such analysis can be

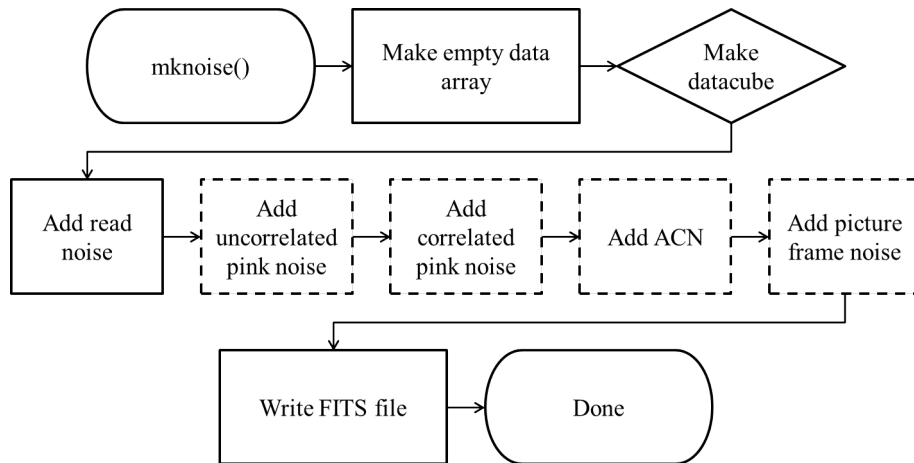


Figure 3.1: **Flowchart for Adding Detector Noise using the Noise Generator Tool.** Flowchart for the `mknoise()` method used from NG. Boxes with solid borders use the time domain, while dashed borders indicate processes that are added in Fourier space. We only consider the aspects that are involved in making a 2-dimensional detector image, rather than a detector datacube. *Image adapted from Rauscher (2015).*

conducted on the HARMONI H4RG detectors.

The implementation of noise components within NG primarily uses Fourier space to add noise in the frequency domain. In Figure 3.2 we reproduce Fig. 2 from Rauscher (2015) showing the power spectrum in Fourier space for the NIRSpec H2RG. The values in the time domain, taken as successive individual pixels along a row, before moving to the next row, have been Fourier transformed to give a power spectrum in the frequency domain. From this the $1/f$ noise is very clear, as is the appearance of white noise at high frequencies where it becomes dominant. The spike at the Nyquist frequency (equal to 50 kHz for detectors with 1×10^{-5} s delay time between pixels) corresponds to the ACN. The picture frame effect is harder to see directly here, however has been shown to be approximated by modulating the 0^{th} component of the PCA by $1/f$ (Rauscher, 2015).

This therefore means the pink noise, ACN, and picture frame effect are easier to include by generating the noise in Fourier space, and converting it back into a time domain for inclusion in the simulated detectors. The $1/f$ noise is generated using the class `pink_noise()` within NG. This class creates pink noise to the desired level whilst ensuring that it does not appear periodic as a result of applying the Fourier transform. Twice

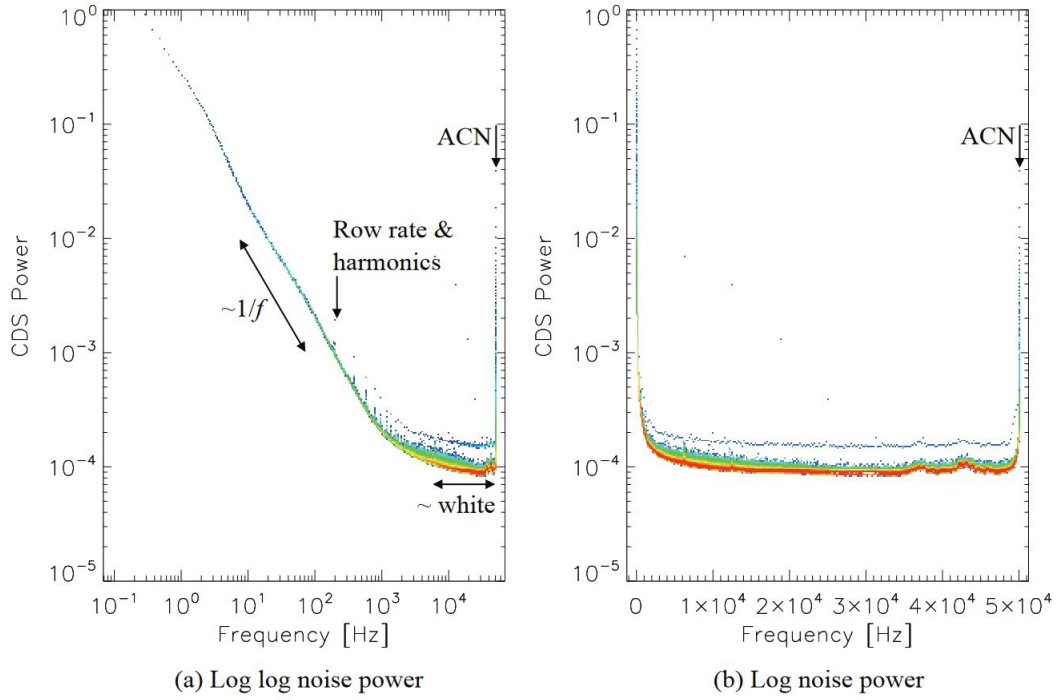


Figure 3.2: Power Spectra in Fourier Space of NIRSpec Detector Noise. *Left:* Averaged power spectrum on a logarithmic scale for the NIRSpec H2RG detector in Fourier space, where the time domain has been taken as successive pixel reads along a row, before moving to the next row. The $1/f$ noise can clearly be seen, as can the white noise at higher frequencies after which the $1/f$ noise is no longer dominant. At the Nyquist sampling frequency, equal to 50 kHz, an additional feature can be seen corresponding to the ACN. Colours denote low density (blue) to high density (red) of data points. *Right:* Same, shown on a linear scale to emphasise the ACN. There are some other small features, such as harmonics, which are not modelled by NG (and hence HSIM) as they contain little power with respect to the other features. *Image reproduced from Rauscher (2015).*

as many random values as required are generated, prior to the Fourier transform. These values are then Fourier transformed, and the $1/f$ nature applied based on 1×10^{-5} s between pixel reads. This is then inverted back into time domain, and the first half of the data kept, to ensure that the noise is more reflective of a physical system, and does not trend back towards the starting point. The data are re-scaled such that the initial mean and standard deviation of the pink noise is retained. Two types of pink noise are simulated: correlated and uncorrelated. The difference is that ‘correlated’ means that the same pink noise appears on all outputs, whereas ‘uncorrelated’ noise can be unique to a particular output (Rauscher, 2015).

As Figure 3.2 shows, the ACN appears as a feature at the Nyquist frequency, with $1/f$ -like behaviour on the low-frequency side. In order to simulate this, additional non-periodic pink noise is generated for the odd and even columns within an output separately. As mentioned previously, the picture frame effect can be simulated by incorporating the 0th component from PCA. We only use NG to simulate a 2-dimensional detector image, rather than a full detector cube with every read, therefore we use the NIRSpec 0th PCA component to add a stationary picture frame effect to the desired level (Rauscher, 2015).

3.1.2 Integration into HSIM

Prior to this work HSIM was designed as an entirely ‘cube-in-cube-out’ simulator, where the user supplies an (x, y, λ) datacube to the simulator. All the internal processing was applied to it in the datacube plane, and a similar (x, y, λ) datacube was returned at the end as the mock observation. However, as we have seen in Chapter 2 detectors are non-uniform across their physical extent and thus in order to simulate all the potential detector effects we needed to make an adjustment to the simulator. To do this, we modified HSIM to be able to reformat the datacube to create a detector plane that we could then co-add to simulated detector noise. This can then be reconverted back into a datacube, with the detector noise included, to maintain the ‘cube-in-cube-out’ nature of HSIM. This option is now provided as a choice for the user, by selecting the Advanced Detector Systematics

(ADS) mode.

In Figure 3.3 we show this new transformation diagrammatically. The precise manner in which the observed light from a source is passed to the detector is non-trivial, and is dependent on the optical arrangement of the instrument. In order to simplify this, whilst still providing a reasonable approximation for the ADS mode, we split the field of view into eight, with two octants corresponding to each of the four spectrographs. This limits the input datacube size to the maximum field of view for HARMONI, being 204×152 spatial pixels and 3700 wavelength channels. The light incident on each of these spectrographs is divided into narrow ‘slitlets’ by the image slicer, and are arranged such that each detector has 38 horizontally adjacent slitlets, in order to convert the 2 spatial dimensions into 1 on the detector. The spectra are then dispersed along the second axis of the detector for each slitlet.

Each quarter is distributed over a pair of detectors as shown in Figure 3.4. Doing this allows us to simulate the slit stagger that will be present in HARMONI, as well as including a ‘V’ shaped slit pattern, and a few pixels padding between each slitlet. This is greatly simplified to the way in which the light will be distributed across the detectors, however this does ensure that features that occur horizontally on the detectors are not directly translated to horizontal features in the output cubes of the simulator. This also means that spatial variations on the detector, such as that caused by the picture frame effect, can show up as increased noise in the centre of the field of view.

As we mentioned previously, there are also limitations to what the NG can do. Primarily, as originally developed it uses a Gaussian read noise distribution, where each pixel behaves in the same way. We have shown in Chapter 2 that this is not the case either, and pixels have read noise that follows a distinct distribution that is common to both H2RGs and H4RGs. In order to include this, we modified the NG such that we create eight 4096×4096 pixel detectors, one for each that will be present in HARMONI, where each pixel is assigned a read noise value drawn from the distributions we have previously identified. We normalise the distribution such that it peaks at 1 e^- read noise, which is

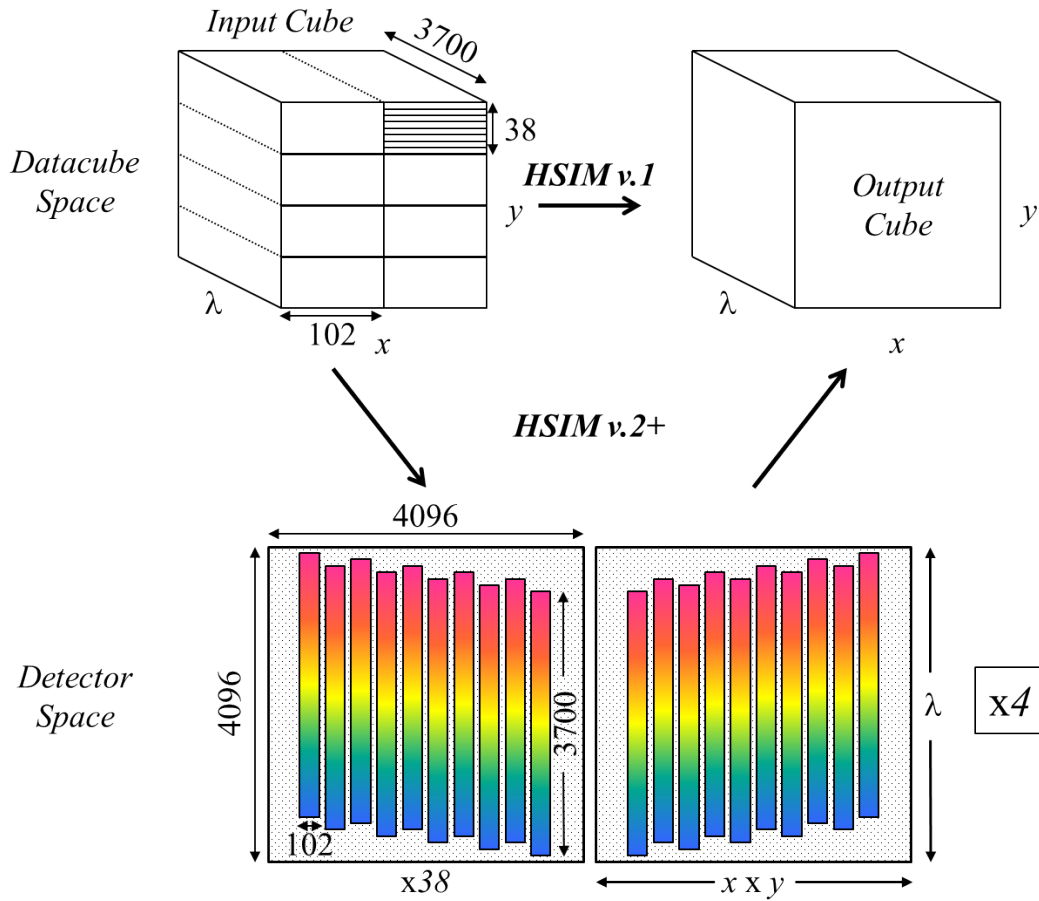


Figure 3.3: Illustration of Conversion from Datacube Space to Detector Space as Implemented into HSIM. HSIM has been modified from a ‘cube-in-cube-out’ simulator (*top*) to one that includes a simulated conversion onto the detector focal plane (*bottom*). The input datacube is cropped to a maximum size that of 204×152 spatial pixels, with 3700 wavelength channels. This is divided between 8 octants, with 38 slitlets of 102 spatial pixels each, which are then passed to pairs of simulated detectors representing each spectrograph. The 38 slitlets are placed sequentially on a simulated detector, with a simulated slit stagger and offset between slitlets, mirrored between the paired detectors. This conversion is then reversed such that HSIM returns a datacube as was originally designed.

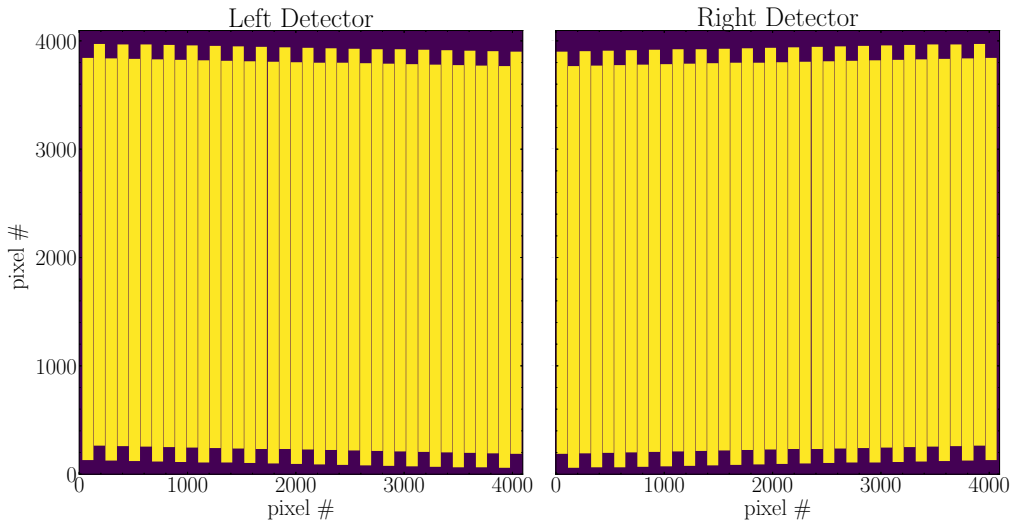


Figure 3.4: **Distribution of Datacube onto Simulated *HARMONI* H4RGs.** Illustration of the distribution of the datacube into the detector plane within *HARMONI*. The pixels over which the datacube is distributed are plotted in yellow, showing the slit offsets and stagger that we simulate.

then be scaled by the appropriate value in Table 3.1 to add the relevant read noise in a simulation. Within *HARMONI* the NG then runs using these read noise maps, rather than drawing from a Gaussian distribution, to create more realistic simulated detector noise that is then added to the simulation. We include a full breakdown of how to run *HARMONI* in ADS mode in Appendix A.

We show an example of a resulting detector created using this process, with no simulated spectra, in Figure 3.5. This uses the nominal ADS noise components, meaning we are primarily dominated by read noise and pink noise, both correlated and uncorrelated. The latter is responsible for the majority of the banding seen across the detector. In addition we can see a number of hot pixels, as well as the presence of some ACN.

In order to further confirm that our implementation was working as desired, we repeated the analysis conducted for Section 2.2.2. To do this we took 40 simulated dark images from *HARMONI* where the pixel base map and average noise component values remained the same on each, but with a different noise seed between them to allow for stochastic processes to vary. *HARMONI* now generates many of the horizontal and vertical noise features seen previously, therefore we conducted the same reference pixel subtraction as done for

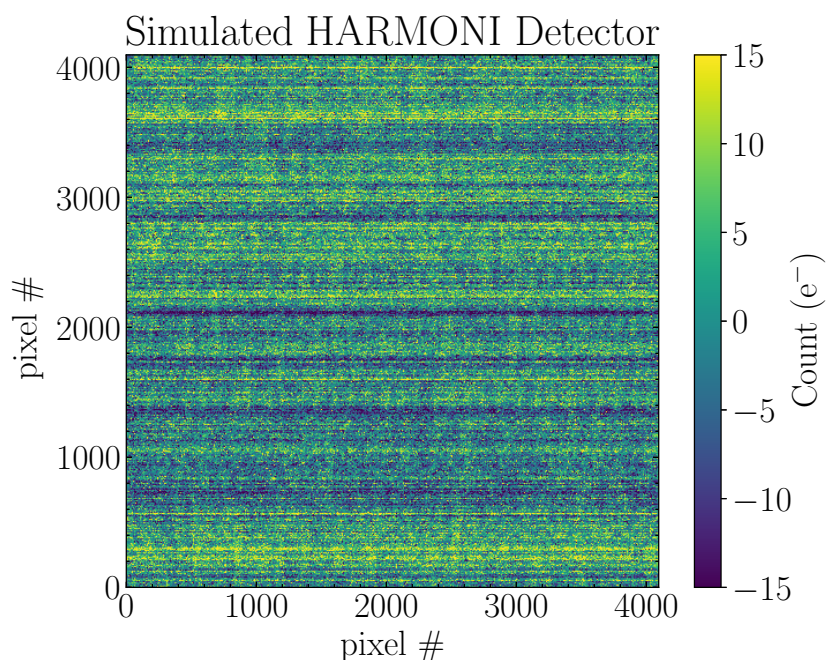


Figure 3.5: **Example Simulated HARMONI detector used in HSIM.** An example of one of the simulated detectors used in HSIM, with nominal ADS used and no additional correction. The noise we see here is dominated by a combination of pink noise and read noise, although there are other effects such as hot pixels and some ACN that is also visible.

KMOS for a direct comparison. As before, we then determined the mean and standard deviation on a pixel-by-pixel basis, and we show the resulting plots in Figure 3.6.

We see from this that the distributions of the means and standard deviations follow expected profiles, and are similar to those for the KMOS H2RG devices and the MOONS H4RG device. The distribution of the means is much cleaner than that seen in the observational data, as we have simulated ‘perfect’ darks, with no stray light, or potential random effects such as cosmic hits, so we have a distribution peaked around 0 e^- . For the standard deviations, we see the same shape as we have seen previously, with a sharp rise to a peak, then a steady fall off with a tail towards high values. The peak of this distribution falls approximately at the injected read noise level of 2.845 e^- , as we would expect to see. This confirms that our implementation into HSIM is producing realistic simulated detectors.

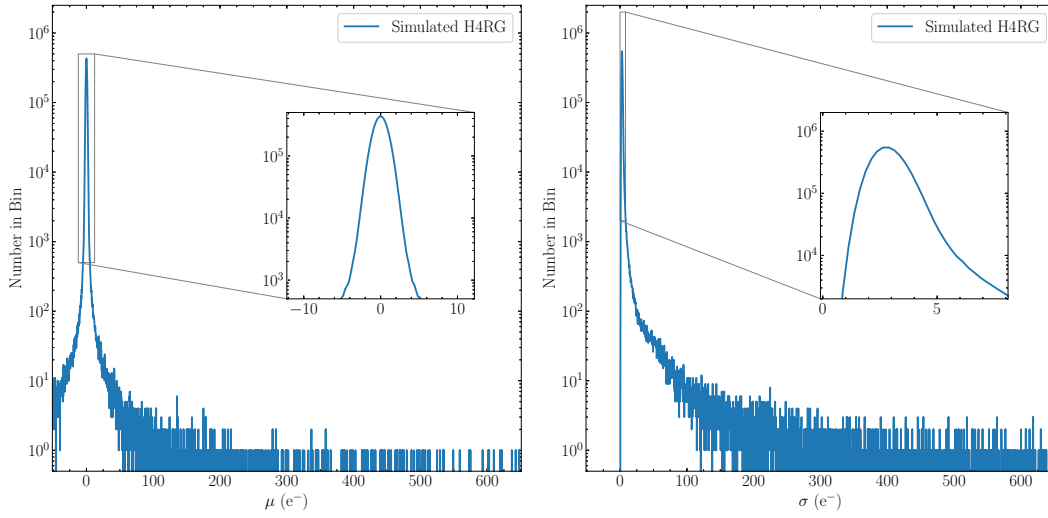


Figure 3.6: **Distribution of Detector Pixel Means and Standard Deviations on a Simulated HARMONI H4RG.** *Left:* Distribution of pixel means on a simulated HARMONI H4RG detector, binned into 0.25 e^- bins. Inset shows the shape of the distribution around the peak. *Right:* Same, for the pixel standard deviations.

3.2 Input Cube Creation and Analysis

3.2.1 HSIM Input Cube

In order to test the impact of the detector systematic effects we first needed to create a suitable input cube for HSIM that we could then run in different configurations. To do this, we generated a point source using the same methodology as in the HSIM-compatible CUBEGEN tool¹. For the purposes of this, the precise nature of the point source was not of priority, therefore we used a flat spectrum with no spectral features. However, of more importance was the brightness of the source, as we wished to investigate the impact of the detector systematic effects. This therefore meant we desired a point source close to the limiting magnitude of detection in the regime in which we conducted our simulations. We therefore created a cube with the central wavelength at $1.66 \mu\text{m}$ to fall firmly within the infrared, and an AB magnitude of 24.

In addition to this we had to choose the wavelength range to cover. Simulating an entire datacube that filled the range spatially as well as spectrally was impossible, due to

¹Publically available at <https://github.com/HARMONI-ELT/CUBEGEN>

the computational time that would be required, particularly spectrally. After testing, we determined that including the first 500 wavelength channels was the best option for both ensuring we were covering a large proportion of the detector, whilst also ensuring we could complete our simulations. Spatially, we created the input cube with 1 mas pixels, a factor of ten greater than the sampling in our simulations, using 10×10 mas spaxel simulations. Once we had created this input point source, we then placed it at 21 locations within a full spatial field of view of HARMONI. These were positioned as shown in Figure 3.7, such that we had 21 point sources that we number accordingly. These were placed such that when the field of view was divided across the eight detectors, some of the point sources fell entirely on one detector, whilst others were split across two or four detectors.

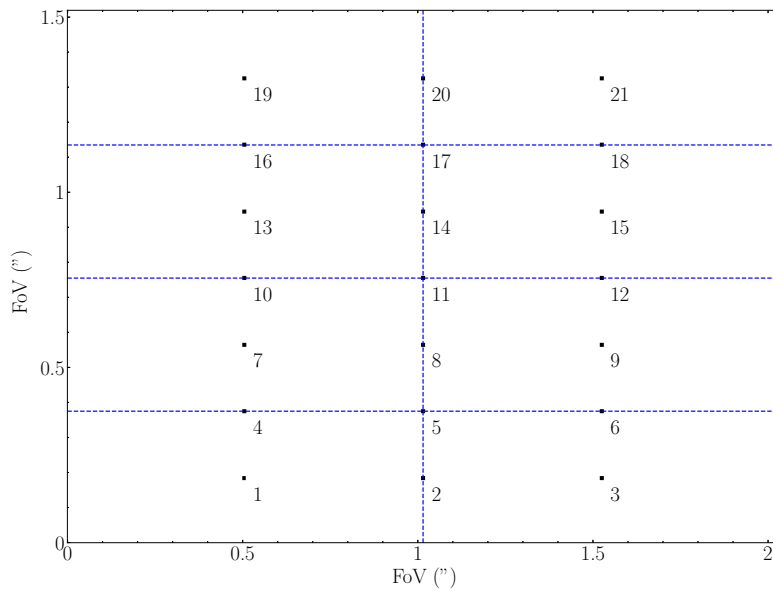


Figure 3.7: Position of Test Point Sources on HARMONI Field of View. Locations of the 21 point sources simulated during our ADS testing situated within a field of view (FoV) equal in size to that of HARMONI when observing at 10 mas scale. Point source sizes are enlarged for illustrative purposes, and the numbers to which we refer to them are shown adjacent to each point source. We also show the portions of each field of view that are assigned to a single detector, with the boundaries of each shown (blue dashes).

When we created our first simulations, we found that the sources nearer the centre of the field of view had systematically higher flux than those at the edges. We identified the source of this being the large wings of the H_{SIM} point spread function (PSF) blending

a small amount of light from all the point sources together. Whilst this is an important effect to consider in observations, this was not the focus of this study. In order to prevent this, we took the returned PSF from these preliminary simulations and cropped it such that we kept the central 100 mas but set the wings of the PSF to zero. This ensured that no light from one point source blended with another, whilst having minimal impact on our simulations, providing we only extract flux from within a 100 mas radius around each point source. The other main drawback of this method is that by providing `HSIM` with a user-defined PSF it does not generate a specific PSF corresponding to each wavelength in our simulation. However, again as the purpose of these simulations was to test the impact of the ADS in `HSIM`, this was not a concern for us.

3.2.2 Simulation Methodology

We were interested in testing the point source sensitivity with ADS, as well as potential spatial variation due to non-uniform noise sources, and also the reproducibility of recovering the same point source flux when using ADS. We therefore produced a number of simulations using the input cube shown in Figure 3.7 to explore a range of parameters, with Table 3.2 showing the default `HSIM` inputs. Within these input parameters, we varied the exposure time in order to test the signal-to-noise behaviour of the ADS, and we used a random noise seed on each simulation. This ensured that the instance of detector noise differed between observation, as this is determined by the noise seed provided. The noise seed generated for a given run is returned in the `HSIM` log files, therefore an exact simulation could be repeated if necessary.

To test the behaviour of each of the noise components included in the ADS, we produced multiple sets of simulations, where one component was enhanced by a factor of five over the nominal noise level. This is to simulate a situation in which the detectors are particularly susceptible to a given contributing noise source. We also produced a set of simulations where we only included an individual ADS noise component, which is less physically realistic, but allowed for more direct comparisons to be drawn. We repeated

HSIM Parameter	Value	Description
ver	310	Version of HSI M
DIT	900	Integration time per exposure (seconds)
NINT ^a	20	Number of exposures
spax	10x10	Spaxel size (mas)
Grating	H	Chosen grating, equal to R $\sim$ 7000
Seeing	0.72	Simulated seeing conditions
Air Mass	1.3	Simulated air mass
Temp	280.5	Temperature of telescope (K)
Moon	0	Moon illumination
jit	0	Residual telescope jitter
ADR	False	Atmospheric Differential Refraction correction
Dets	True	Advanced detector systematics on
PSF ^b	User	User Defined Input PSF
Seed ^c	\$RANDOM	Random noise seed

Table 3.2: **Default HSI M Simulation Parameters for ADS Testing.** Default parameters used as the HSI M inputs for ADS testing. ^aWe also simulated exposures with longer and shorter times as part of our investigation. ^bThe input PSF is explained in Section 3.2.1 and is based on the LTAO PSF. ^cWe used a random noise seed on each run to ensure variation between the output cubes.

this for a range of total integrations as well, namely 4, 20, 60, 100, 200 and 400 for exposure times of between 1 hour and 100 hours. For each ADS configuration and exposure time, we produced 10 simulations, so that we could investigate the behaviour between multiple observations of the same input cube.

In Figure 3.8 we show a comparison between one frame of a simulated output cube with no ADS, with a simulation with nominal ADS, both with 900×20 s exposure time. In both we can clearly see the broadening of the point source due to the PSF, however there are a number of features we see in the simulation that uses ADS that are not present in the one without.

Firstly, we can see the regions of the FoV that correspond to the eight different detectors, as they have slight differences in the level of background noise in each octant. In addition to that, we see some roughly horizontal features appearing, due to non-uniform noise effects on the detectors. It is important to consider that this is shown in datacube space, rather than detector space, therefore the effects we have seen previously are folded in and not easily associated together. We do also see a number of hot pixels, which comes

naturally from having a distribution of read noise with a high tail.

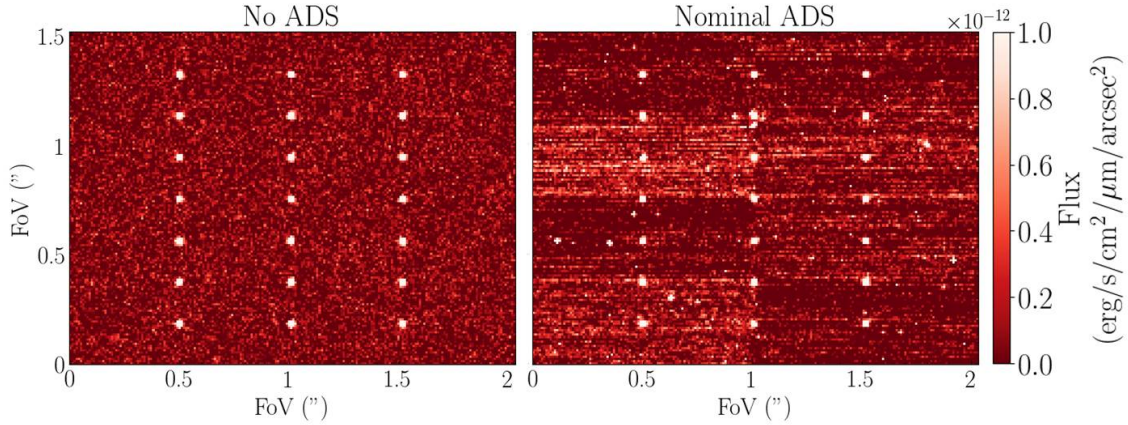


Figure 3.8: **Example Output Cubes from `HSIM` ADS Mode Testing Runs.** *Left:* A single wavelength slice from a flux-calibrated `HSIM` simulation with 900×20 s integration time not using the ADS mode. *Right:* Same, using the ADS mode with nominal detector noise. Differences can clearly be seen, such as the variation between detector noise levels, and a number of hot pixels.

3.2.3 Analysis

There are a number of output cubes returned from `HSIM` that we utilised for our analysis, as well as the information provided in the `.log` file. `HSIM` version 310 incorporates the 5σ point source sensitivity limit predictions used in Zieleniewski et al. (2015), which we adapted to include the noise contributions from the different components of the ADS. In this, we have assumed a ‘worst-case scenario’, in which we know the magnitude of the varying components but do not assume any knowledge of their spatial dependence, and we assume no additional pipeline corrections on the datacubes.

To investigate the impact of the ADS on the recovered flux we used the flux-calibrated reduced datacube returned by `HSIM`. We chose to use the 900×4 data for this, as we have sufficient SNR that our results are reliable, whilst also ensuring we still capture the nature of the noise processes. The reduced datacube subtracts off a separate instance of noise, although it does assume a perfect knowledge of the sky, and only varies the Poisson noise from it. This does use the same base detectors, with the same read noise associated with each pixel, but again this is a different instance of noise that is subtracted off from the

‘observed’ datacube.

With this, we then extracted the flux from a three spaxel box at each wavelength channel, centred on the location of each point source. This was chosen as this is the limit to which we would only include pixels that have an SNR greater than 1, thus are contributing more signal than noise for this exposure time. We summed these for each point source in the field of view, and then repeated this for every simulation to obtain a set of ten total flux values for each point source. We took the median of these values to use as the expected flux for a given point source to minimise the impact of outliers, with the uncertainty on this value determined by the standard deviation.

To determine the spatially averaged flux, we then calculated the mean of these 21 values, adding the uncertainties in quadrature. We also determined the interquartile range (IQR) for the flux at each of the 21 locations, and calculated the average IQR for these locations, with an uncertainty given by the standard deviation of these values. This gave us a measure of how reliably the same flux is recovered between simulations.

This analysis was then repeated for each of the seven simulation runs, one for each enhanced noise component, along with the nominal ADS and no ADS simulations. To ensure that we had correctly accounted for PSF effects in the output cubes, we also repeated this analysis with the noiseless object cube. As expected each point source returned the same flux with no variation between simulations, meaning any differences are due to variations in the noise.

We also utilised the reduced SNR cube to investigate the behaviour of the SNR as a function of time with the different noise components. For this, we extracted the SNR from the same 3 spaxel box as we did for the flux. We then averaged the SNR spectrally, and took the peak from within the extraction region to obtain a single SNR value corresponding to each point source in every exposure. We took the average of these 21 values to obtain a single SNR for a given exposure, with the standard deviation giving us an uncertainty predominately due to spatial variations within the FoV. We then further averaged the 10 exposures for each simulation set together, giving us 210 datapoints correspond-

ing to each exposure time and noise component. We added the standard deviations in quadrature to give an overall estimation of the uncertainty.

We could then use these values to fit a functional form to the SNR. We have set up these simulations such that we are in a regime where read noise is the dominant component, therefore we use the form

$$\text{SNR}(N) = A\sqrt{N} \quad (3.2.1)$$

where N is the number of individual integrations, and A is the amplitude factor, corresponding to the SNR in a single integration. To perform this fit we used the `curve_fit` non-linear least squares fitting method from `SciPy`, using uncertainties obtained in the manner described above, and A as a free parameter. We could then determine the SNR after one hour, SNR_{1h} by multiplying this value by 2, since we have 4 separate integrations per hour of simulated observation.

Comparing these expected SNR values after 1 hour gives us a direct comparison for the implications of the noise components on the recoverable SNR, with an uncertainty returned from the least-squares fit. Dividing this by the expected SNR without using ADS, $\text{SNR}_{1h,N}$, allowed us to obtain a relative noise contribution when enhancing different noise components.

3.3 Results and Discussion

The primary goal from these simulations was to determine the impact of ADS when observing with HARMONI, assuming a ‘worst-case scenario’ in which no additional correction is applied in the instrument pipeline. Initially we present here the results of the point source sensitivities as determined from `HSIM`. We then show the potential spatial variations and repeatability issues that come with this additional noise. At the end of this section, we show the SNR limitations that these additional noise components could introduce into observations.

3.3.1 Point Source Sensitivities

For each simulation, we could obtain the 5σ point source sensitivity per resolution element, using the values obtained from the 900×4 simulations, corresponding to 4 exposures of 15 minutes each. These are calculated within `HSIM` by taking the median of the noise cube and multiplying this by 5 to give the 5σ limit. This is then divided by the appropriate flux calibration factor for the central wavelength of the desired grating, which for the H-band is $1.625 \mu\text{m}$ with `HARMONI`. This uses the internal PSF to determine the fraction of the light which falls into the core of the PSF to account for the spatial distribution of the light effectively dimming the peak flux. These values are shown in Table 3.3 for each of the five noise components included in the ADS mode alongside the corresponding values for the nominal- and no-ADS simulations. Each noise component is enhanced by a factor of five over the nominal values, as described in Section 3.2.2.

Simulation Type	5σ Sensitivity (AB mag)
No ADS	23.89
Nominal ADS	23.34
$5 \times 2.845 \text{ e}^-$ Read Noise	22.40
$5 \times 0.5 \text{ e}^-$ ACN	23.27
$5 \times 0.2 \text{ e}^-$ Picture Frame	23.33
$5 \times 3 \text{ e}^-$ Correlated Pink Noise	22.35
$5 \times 1 \text{ e}^-$ Uncorrelated Pink Noise	23.10

Table 3.3: **Point Source Sensitivities for Different Noise Components Using ADS in `HSIM`.** 5σ sensitivities per resolution element, in AB magnitude, for 900×4 s simulated exposures with `HSIM` for the systematic noise components included with the ADS mode. Each noise component is enhanced by a factor of 5 over the nominal values.

We see initially from these results that as expected we do lose a little bit in 5σ sensitivity when we include ADS, going from $M_{\text{AB}} = 23.89$ to $M_{\text{AB}} = 23.34$ corresponding to a drop in limiting magnitude of 0.55. When we look at the simulations that had one component enhanced relative to the others, we see that the largest drops in 5σ sensitivity come when we have very high correlated pink noise and read noise, with AB magnitudes of 22.35 and 22.40 respectively. As these are individually the largest contributing noise components, it makes sense that enhancing these would correspond to a greater loss in

limiting magnitude. Similarly for enhanced ACN, picture frame effect and uncorrelated pink noise the drop is less significant, with magnitudes of 23.27, 23.33 and 23.10.

However, it is important to consider that these point source sensitivities do not take into account spatial variation of the noise, and elements such as the picture frame effect are very spatially dependent. These noise components are each added in quadrature to give an overall estimate of the levels of noise within an exposure, assuming there is no additional pipeline correction. To better understand the impact of the ADS noise, the next sections present the results of more detailed investigations into the spatial dependence and repeatability of observations.

3.3.2 Spatial Variation and Reproducibility

Using the methodology described in Section 3.2.3 on the 900×4 s simulations, we could plot the average extracted flux from each point source. This is shown in Figure 3.9 with each panel representing a different enhanced noise component, with panels also showing the nominal- and no-ADS simulations. We also tabulate the spatially averaged flux and IQR for the 21 point sources, which is shown in Table 3.4.

Simulation Type	^a Spatially Averaged Flux	^a Average Point Source IQR
No ADS	5.28 ± 0.01	0.07 ± 0.03
Nominal ADS	5.23 ± 0.13	0.58 ± 0.48
	(5.24 ± 0.09)	(0.48 ± 0.18)
$5 \times 2.845 \text{ e}^-$ Read Noise	5.44 ± 0.28	1.31 ± 1.59
	(5.27 ± 0.12)	(0.71 ± 0.32)
$5 \times 0.5 \text{ e}^-$ ACN	5.31 ± 0.10	0.53 ± 0.19
$5 \times 0.2 \text{ e}^-$ Picture Frame	5.32 ± 0.12	0.54 ± 0.25
$5 \times 3 \text{ e}^-$ Correlated Pink Noise	5.14 ± 0.44	2.51 ± 1.09
$5 \times 1 \text{ e}^-$ Uncorrelated Pink Noise	5.32 ± 0.15	0.91 ± 0.39

^a $10^{-9} \text{ erg s}^{-1}$

Table 3.4: Spatially Averaged Flux and IQR of Recovered Point Source Fluxes with Enhanced ADS Noise Components. Spatially averaged flux across the 21 locations, along with the averaged IQR across each location for the seven simulation runs completed with `HSIM` using ADS noise components enhanced by a factor of 5. Values in brackets have outlier points excluded from the calculation.

From Figure 3.9 we can immediately see that there is a difference between using (nominal) detector systematics, and not using any. We can see that there is increased variability in the flux recovery, indicated by the larger uncertainties in the nominal ADS panel. We can also see some differences spatially, particularly at location 6, where the uncertainty is much larger than the others. When we calculate the spatially averaged flux, we recover a value of $(5.28 \pm 0.01) \times 10^{-9} \text{ erg s}^{-1}$ when we do not include ADS, compared with $(5.23 \pm 0.13) \times 10^{-9} \text{ erg s}^{-1}$ when we do. In terms of the IQR we find a value of $(0.07 \pm 0.03) \times 10^{-9} \text{ erg s}^{-1}$ without ADS, in comparison with $(0.58 \pm 0.48) \times 10^{-9} \text{ erg s}^{-1}$ when we do.

The recovered flux from using ADS is consistent within 1σ of that without, as we would hope to find, however there is a much larger uncertainty on this. This comes from adding the additional noise, as well as larger spatial variation in the determined flux value. The IQR when we do not use ADS is understandably very small, and very consistent, as this only includes ‘flat’ read noise in each exposure. This increases when we include ADS, as the noise profile is more complex and spatially varying. The flux from the point source at location 6 corresponds to the location of a bad pixel at one spectral channel, greatly increasing the overall noise at that point. If we mask out this location, and recalculate the results, we obtain an average flux of $(5.24 \pm 0.09) \times 10^{-9} \text{ erg s}^{-1}$, with an IQR of $(0.48 \pm 0.18) \times 10^{-9} \text{ erg s}^{-1}$, which is in tighter agreement with the no-ADS simulations. This shows that if we do not properly account for detector effects, whilst we may be able to recover accurate fluxes over a number of observations, an individual observation can have a high chance of being skewed due to the detector noise.

With the other simulations, the recovered flux varies depending on which noise component is enhanced, but all of them recover a value that agrees within one sigma of the flux without ADS, as with the nominal levels. We see that the uncertainties on enhanced ACN, picture frame effect, and uncorrelated pink noise are all similar, with larger uncertainties for read noise and correlated pink noise. Given that these are individually the largest contributing noise components in the nominal ADS simulations, when we enhance

3.3. RESULTS AND DISCUSSION

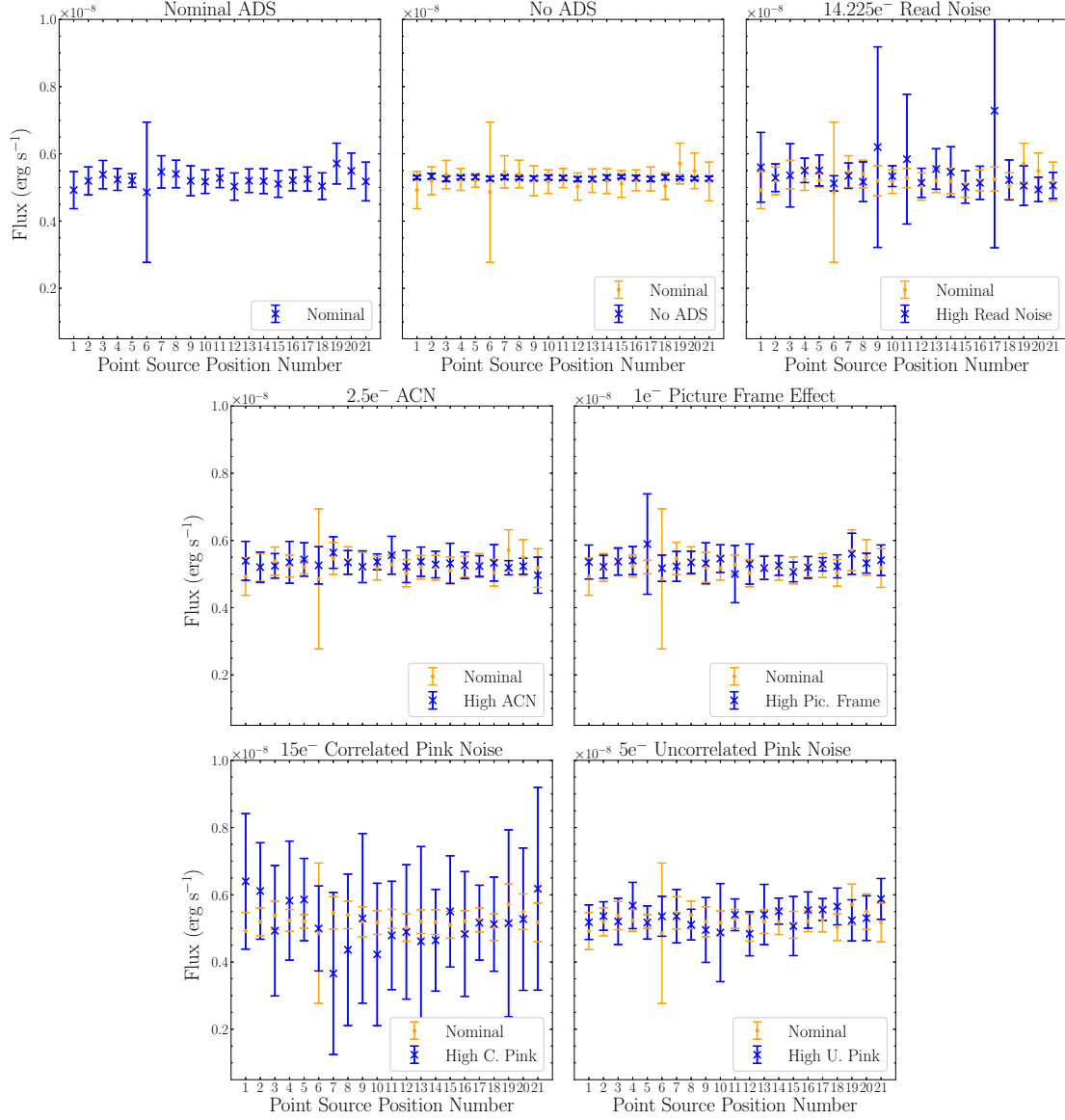


Figure 3.9: Averaged Point Source Flux at 21 Fixed Locations with Enhanced ADS Noise Components. Seven panels showing median recovered point source fluxes from the 900×4 s simulation sets using *HSIM*. *From top left to bottom right:* nominal ADS values, no ADS used, enhanced read noise, enhanced ACN, enhanced picture frame effect, enhanced correlated pink noise, and enhanced uncorrelated pink noise. We also plot the nominal values in orange on every panel for quick comparison. The numbering convention used for the point sources is shown in Figure 3.7. The uncertainties on these values are obtained by taking the standard deviation of the ten simulations run for each noise component.

them further it would be expected that these lead to the largest variations in flux recovery.

The average IQR, measuring how closely the same flux is recovered across the ten simulations, follows a similar pattern. Enhanced read noise and correlated pink noise have the highest IQR, as well as the highest uncertainty, implying large spatial variations in the recoverability of the flux. However when we look at Figure 3.9 we can see this is largely driven by three particular locations for the read noise, 9, 11, and 17. This is again the result of bad pixels, and we would expect there to be more present when we increase the average read noise by a factor of five.

For uncorrelated pink noise we also see an increase in the error on the IQR, which comes as a result of the noise being different on each amplifier block, thus it affects some locations more than others. We see for ACN and picture frame effect that when we enhance the noise, the IQR goes down but the uncertainty goes up. As these in themselves are a small effect relative to the total noise, this is indicative of an effect that is particularly localised, as the average value, i.e. the majority of the detector, is unaffected, however certain locations increase significantly, so the variance on the IQR increases.

From these results, we can say that incorporating these additional noise effects will introduce some spatial variation across the field of view, as well as impacting the reproducibility of recovering the same flux from an exposure. In practice, this would mean longer observing times (or more integrations) to be able to recover data to the same level of reliability as without them. As for enhancing the separate noise components, those that have the largest impact in this test are read noise and pink noise, particularly correlated pink noise. Minimising and characterising read noise and hot pixels is part of the usual process when testing a detector for astronomy, but from this we can show that it is equally, or even more important to control or account for the pink noise as well. In terms of spatial variation and reproducibility, the other components have similar impacts when enhanced by a factor of five over their nominal levels, however we will look at them on an individual level again in Section 3.4, where we show simulations with just one noise component present.

3.3.3 Signal to Noise Ratio (SNR) Limits

In addition to the spatial variations and reproducibility, we were interested in the impact on the SNR when introducing additional detector noise effects. In Figure 3.10 we show the SNR as a function of total integration time for each of the enhanced noise components as well as the nominal- and no-ADS simulations. We also calculate the factor A as defined in Equation 3.2.1 for each curve and use this to determine the SNR_{1h} values. These are presented in Table 3.5, along with the relative factor $\text{SNR}_{1h}/\text{SNR}_{1h,N}$ with $\text{SNR}_{1h,N}$ the SNR after 1 hour without using ADS.

Simulation Type	SNR_{1h}	$\text{SNR}_{1h}/\text{SNR}_{1h,N}$
No ADS	6.895 ± 0.000	1
Nominal ADS	6.643 ± 0.006	0.963
$5 \times 2.845 \text{ e}^-$ Read Noise	4.637 ± 0.005	0.673
$5 \times 0.5 \text{ e}^-$ ACN	6.356 ± 0.005	0.922
$5 \times 0.2 \text{ e}^-$ Picture Frame	6.587 ± 0.006	0.955
$5 \times 3 \text{ e}^-$ Correlated Pink Noise	4.000 ± 0.019	0.580
$5 \times 1 \text{ e}^-$ Uncorrelated Pink Noise	5.825 ± 0.008	0.845

Table 3.5: **SNR Scaling with Varying Enhanced ADS Noise Components.** The SNR at 1 hour, as determined by multiplying the $\text{SNR}(N)$ as defined by Equation 3.2.1 by 2, for enhanced ADS noise component simulations. We also give the normalised $\text{SNR}_{1h}/\text{SNR}_{1h,N}$ with $\text{SNR}_{1h,N}$ being the SNR after 1 hour without using ADS. To the level of precision of these simulations, there is no uncertainty associated with the no-ADS runs, as the noise cube used in the determination of the SNR is the same in each run.

Looking at these results, we see that the impact of nominal detector systematics on the overall SNR is small but noticeable, with an expected SNR at 1 hour of 6.643 ± 0.006 , compared with 6.895. This corresponds to a ratio of 0.963, meaning there is roughly a 4% drop in SNR when using ADS. This is in line with expectations, as the ADS noise levels are chosen such that they approximate the overall noise level, but with increased complexity in their implementation. There is also a small variation in recovered SNR between simulations, leading to the uncertainty on the value, coming from spatial variations and different detector instances. These are only present when ADS is used, therefore without them the SNR will always return the same value.

When we enhance the different noise components we see a reduction in the recov-

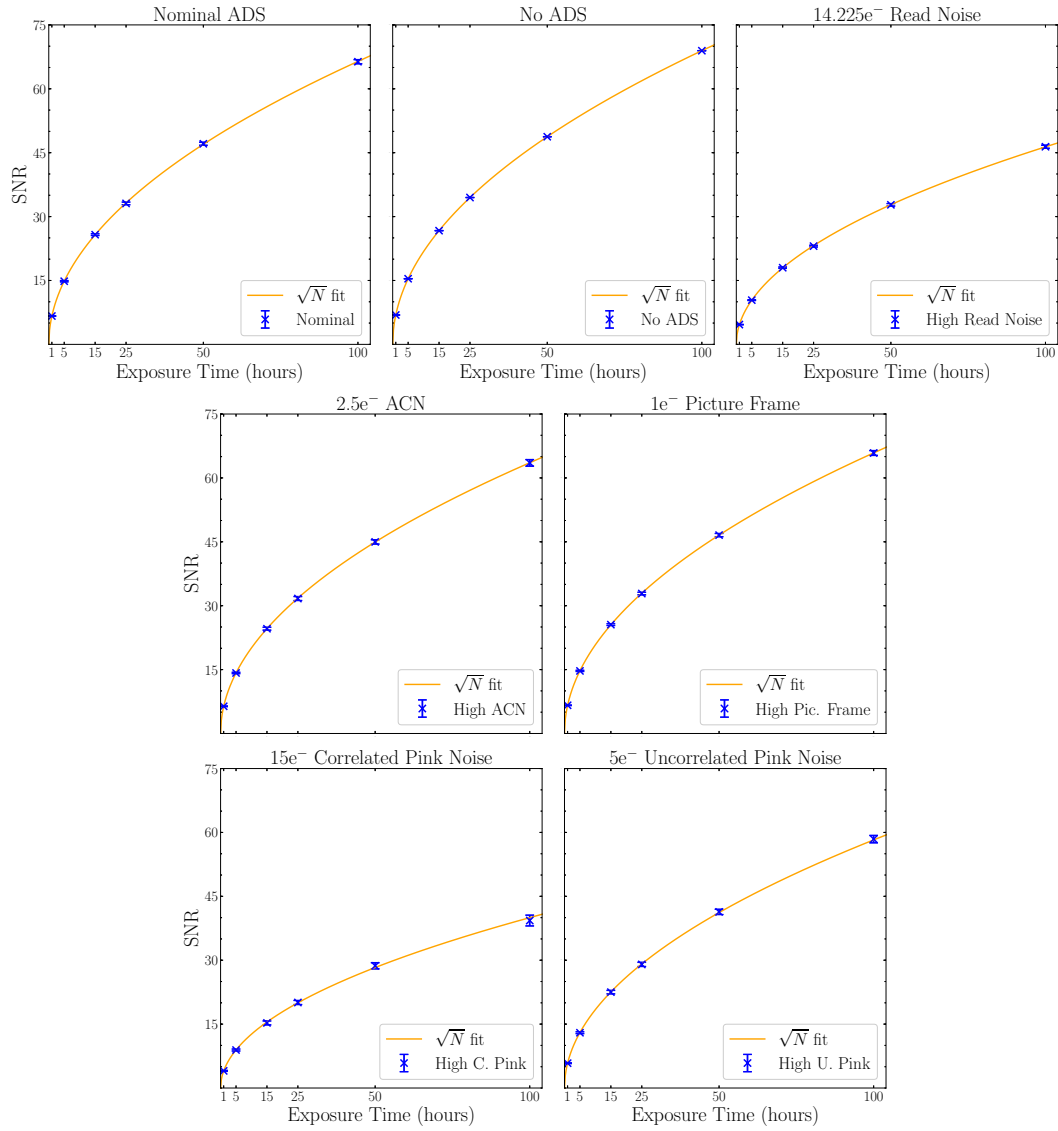


Figure 3.10: SNR as a Function of Integrations with Varying Enhanced ADS Noise Levels. Seven panels showing the relative SNR values as a function of integrations for each enhanced ADS noise component. *From top left to bottom right:* nominal ADS values, no ADS used, enhanced read noise, enhanced ACN, enhanced picture frame effect, enhanced correlated pink noise, and enhanced uncorrelated pink noise. We also show the best fitting $\sqrt{N}$ function, from which we derive the parameter A and corresponding SNR after 1 hour of observation.

ered SNR, as we would expect by including additional noise. The greatest impact comes from enhancing correlated pink noise, where the SNR is reduced by a factor of 0.580 to 4.000 ± 0.019 . Enhancing the read noise also has a large reduction in the SNR, with a factor of 0.673 giving an expected SNR of 4.637 ± 0.005 after one hour. As these are the largest individual contributing factors, it makes sense that enhancing these causes the largest drop in SNR. What we do see is that the variance between the two is very different, with increased read noise causing a more uniform drop in SNR at every location compared to correlated pink noise.

With enhanced uncorrelated pink noise, ACN, and picture frame we see a smaller reduction in SNR, with expected values of 5.825 ± 0.008 , 6.356 ± 0.005 , and 6.587 ± 0.006 respectively, corresponding to reductions of factors 0.845, 0.922, and 0.955. As these are smaller contributing factors, we would expect them to have a lesser impact on the overall SNR. Additionally, noise such as the picture frame effect is very spatially dependent, therefore the impact may be significant at one location in the FoV, but minimal on the rest. These have small uncertainties, particularly ACN and picture frame noise, meaning the SNR is fairly consistent each time. Uncorrelated pink noise has a slightly higher uncertainty, which is to be expected for a noise component which varies between each amplifier block.

From these results, we can confirm the expected behaviour that increasing the overall noise levels reduces the SNR which is further proof that the simulated detectors are functioning correctly. We can see that for the most part, the SNR does not vary greatly over the 210 point sources for a given observation, and is well described by a $\sqrt{N}$ fit. As with the spatial variation and reproducibility tests, we also look at all these components on an individual basis in Section 3.4 and repeat this analysis to isolate particular noise components that have the greatest impact on overall SNR.

3.4 Individual Noise Components

Previously, we have looked at simulations where one noise component is significantly enhanced over the others in the ADS mode, which is a more realistic potential scenario for a HxRG detector, however does make it harder to disentangle exactly which noise components are dominant. In this section, we present a separate investigation in which we repeat much of the same analysis, however with simulated detectors that only contain one noise component in the ADS mode. In addition, we set the relevant component to be 5 e^- rms each time, to allow for direct comparisons to be drawn.

3.4.1 Spatial Variation and Reproducibility

We repeated the simulation set to obtain results in which only one component from the ADS is present. We show these results in Figure 3.11, with the spatially averaged fluxes and IQRs tabulated in Table 3.6. We include the same nominal- and no-ADS simulations as previously described for comparison.

Simulation Type	^a Spatially Averaged Flux	^a Average Point Source IQR
No ADS	5.28 ± 0.01	0.07 ± 0.03
Nominal ADS	5.23 ± 0.13	0.58 ± 0.48
	(5.24 ± 0.09)	(0.48 ± 0.18)
Only 5 e^- Read Noise	5.18 ± 0.07	0.33 ± 0.36
	(5.25 ± 0.06)	(0.26 ± 0.17)
Only 5 e^- ACN	5.29 ± 0.07	0.38 ± 0.16
Only 5 e^- Picture Frame	5.35 ± 0.09	0.51 ± 0.13
Only 5 e^- Correlated Pink Noise	5.28 ± 0.13	0.80 ± 0.32
Only 5 e^- Uncorrelated Pink Noise	5.26 ± 0.08	0.50 ± 0.15

^a $10^{-9}\text{ erg s}^{-1}$

Table 3.6: **Spatially Averaged Flux and IQR of Recovered Point Source Fluxes with Single ADS Noise Components.** Spatially averaged flux across the 21 locations, along with the averaged IQR across each location for the seven simulation runs completed with `HSIM` using single ADS noise components set at 5 e^- rms. Values in brackets have outlier points excluded from the calculation.

Overall the average noise level is lower for a given simulation set, as would be expected given that we are including less noise sources. What we do see, however, is that

the different components do not add noise in the same way. If this was the case, we would expect to recover almost identical errors on the flux no matter which component we simulate, which we do not see.

With 5 e^- average read noise, using the distributions identified in Chapter 2, we see that for most point sources we recover the same flux, with a large outlier at position 8. The average flux is $(5.18 \pm 0.07) \times 10^{-9} \text{ erg s}^{-1}$, which is lower than the recovered flux from not using ADS, however this is being heavily skewed by the outlier. This is again the result of an uncorrected bad pixel contributing levels of flux that vary considerably between observations. If we remove this from our analysis, we recover an average flux of $(5.25 \pm 0.06) \times 10^{-9} \text{ erg s}^{-1}$, which does now agree with the no-ADS flux. Prior to masking this point the IQR suggested something was unusual, as we recovered a value of $(0.33 \pm 0.36) \times 10^{-9} \text{ erg s}^{-1}$, implying a very large spatial variation. When we repeat the calculation excluding the bad pixel data, we get $(0.26 \pm 0.17) \times 10^{-9} \text{ erg s}^{-1}$, still suggesting there is reasonable variation in the flux recovery, but more in line with expectations. This highlights the importance of good instrument pipeline correction or masking, to negate the impact of bad pixels.

With only 5 e^- ACN in our simulations, we recover a spatially averaged flux of $(5.29 \pm 0.07) \times 10^{-9} \text{ erg s}^{-1}$ consistent with the recovered flux without using ADS. There is very little uncertainty associated with this flux, although it is noted that this implementation assumes that each amplifier block experiences the same average level of ACN, which is not always the case on real detectors. We do also see that with an IQR of $(0.38 \pm 0.16) \times 10^{-9} \text{ erg s}^{-1}$ the reproducibility of recovering the same flux is impacted by ACN, compared to not including ADS, although fairly consistently across the field of view. Therefore for observations in which ACN is present, it appears that background subtraction should provide a reasonable correction, however there will be additional noise associated with it. As such it is important to consider in which parts of the detector it is present, and ideally correct for it either in the instrument pipeline or in further data reduction.

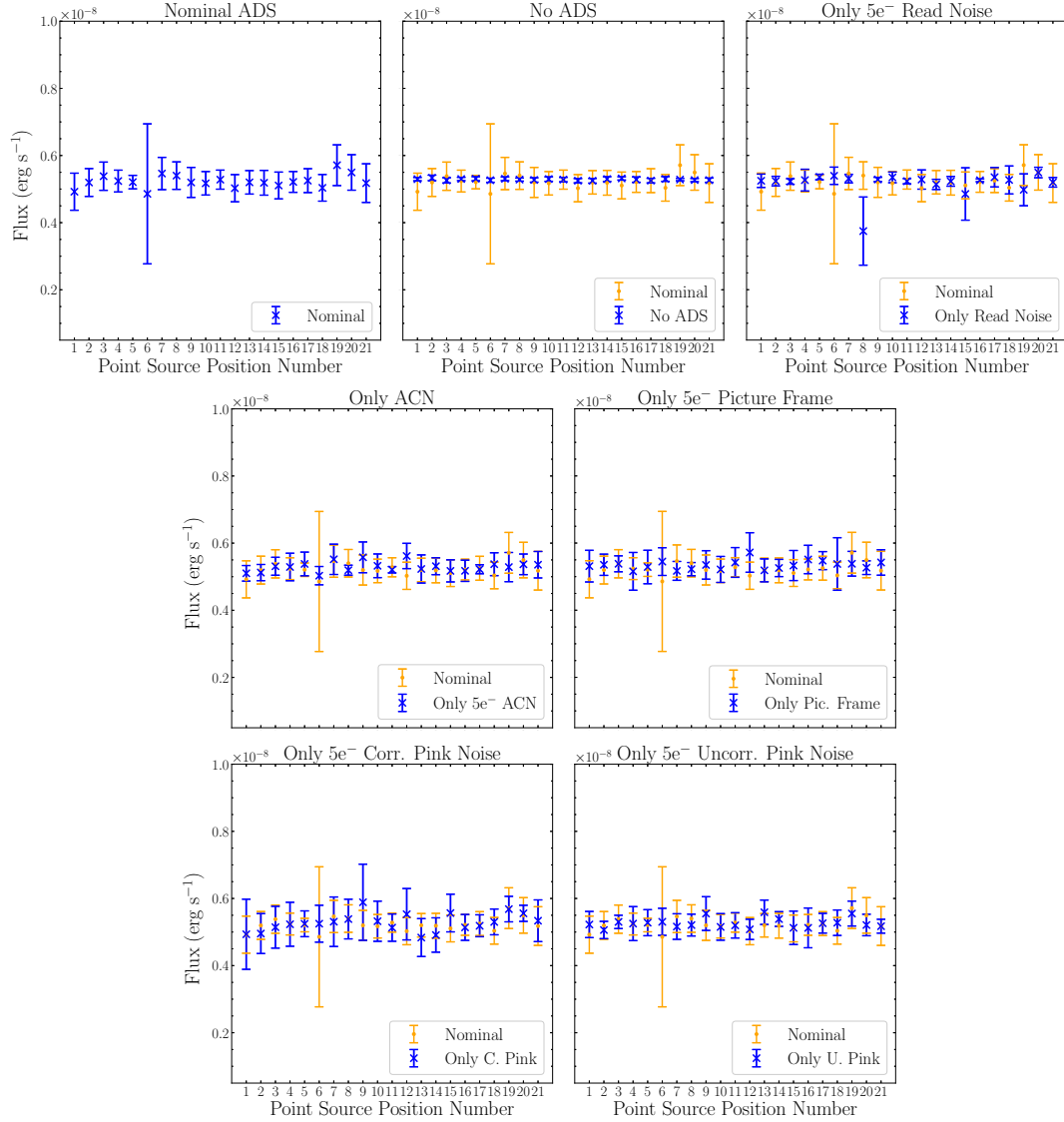


Figure 3.11: Averaged Point Source Flux at 21 Fixed Locations with Single ADS Noise Components. Seven panels showing median recovered point source fluxes from the 900×4 s simulation sets using HS IM. *From top left to bottom right:* nominal ADS values, no ADS used, only read noise, only ACN, only picture frame effect, only correlated pink noise, and only uncorrelated pink noise. We also plot the nominal values in orange on every panel for quick comparison. The numbering convention used for the point sources is shown in Figure 3.7. The uncertainties on these values are obtained by taking the standard deviation of the ten simulations run for each noise component.

When we only include picture frame noise in our simulations, the spatially averaged flux is $(5.35 \pm 0.09) \times 10^{-9} \text{ erg s}^{-1}$ which is consistent with the results from when we do not include ADS. This suggests that the average picture frame effect is being subtracted off with reasonable accuracy, however with an IQR of $(0.51 \pm 0.13) \times 10^{-9} \text{ erg s}^{-1}$ there is a difference in how well this is done at different locations. This would be expected, as the picture frame effect predominantly impacts the edges of detectors. If we look at Figure 3.11, the worst affected point sources are numbers 5, 12, and 18. These correspond to point sources that lie on horizontal boundaries between detectors. As the distribution of flux from the FoV to the detectors is non-trivial, and spectrally we only use a portion of the total available detectors as described in Section 3.2.1, these point sources correspond to detectors in which a larger proportion of the flux falls near the edge. This effect is therefore important to consider, and can primarily be mitigated by ensuring targets are positioned such that the spectra of interest do not fall close to the edge of a detector. With HARMONI, this may lead to offsetting a target slightly from the centre of the field of view.

With only correlated pink noise present in our simulations, we recover a spatially averaged flux of $(5.28 \pm 0.13) \times 10^{-9} \text{ erg s}^{-1}$ which is the same average as that from when we do not use ADS. However, the larger uncertainty in this value implies significant variations and this can be seen clearly in Figure 3.11. An IQR of $(0.80 \pm 0.32) \times 10^{-9} \text{ erg s}^{-1}$ also confirms that it is both harder to recover the same flux, as well as a large spatial difference in this reproducibility. For uncorrelated pink noise, the spatially averaged flux is $(5.26 \pm 0.08) \times 10^{-9} \text{ erg s}^{-1}$, in agreement with the other fluxes. The lower uncertainty on this value implies that it is more even spatially, and this is seen visually from Figure 3.11 as well. The IQR of $(0.50 \pm 0.15) \times 10^{-9} \text{ erg s}^{-1}$ also suggests that this impacts the reproducibility of recovering the flux, albeit with less variation across the entire FoV.

We see correlated pink noise having a larger impact, as this appears as the same noise on all amplifier blocks, therefore every location will be impacted by it. Uncorrelated pink noise means each amplifier block has some that is unique to itself, therefore it would

be expected that some locations are affected worse than others. In terms of removing pink noise, the reference pixels at the edge of the detector are used, and this can work particularly well to remove what we simulate as correlated pink noise, providing there are no issues within the reference pixels.

If there are bad pixels within the reference pixels, or transient effects such as cosmic ray impacts, these should be accounted for prior to reference pixel subtraction, otherwise horizontal effects such as those seen on the KMOS detectors in Figure 2.1 will be present. One way to counteract such events is to save every read of the detector, and apply the reference pixel correction in the instrument pipeline, rather than as part of the electronics. This would allow for additional exclusion of bad pixels, negating some of the effects seen on the KMOS detectors.

3.4.2 Signal to Noise Ratio (SNR) Limits

Looking now at the recovered SNR with individual ADS noise components each set to 5 e^- , we show the recovered SNRs as a function of time in Figure 3.12. As before, we also plot the best fitting $\sqrt{N}$ fit, as would be expected for the SNR scaling as a function of integrations N . From this, we again calculate $\text{SNR}_{1\text{h}}$, the expected SNR after one hour of observation, and show these in Table 3.7 along with the normalised $\text{SNR}_{1\text{h}}/\text{SNR}_{1\text{h},N}$. We include the same nominal- and no-ADS simulations as we generated previously for comparison.

With each component we are now simulating the same amount of average noise, therefore an initial assumption would be to assume they all affect the SNR in the same way, as would be the case for a consistent, uniform noise source. However, we can see from our results that this is not the case.

When we only include 5 e^- read noise, we find an expected SNR of 6.689 ± 0.001 after one hour, with a ratio of 0.970 compared to not using ADS. We see a slight increase in SNR compared to nominal levels, as with a single component at 5 e^- we include a very similar level of noise to the combined effect of all the ADS noise if it was all spatially

3.4. INDIVIDUAL NOISE COMPONENTS

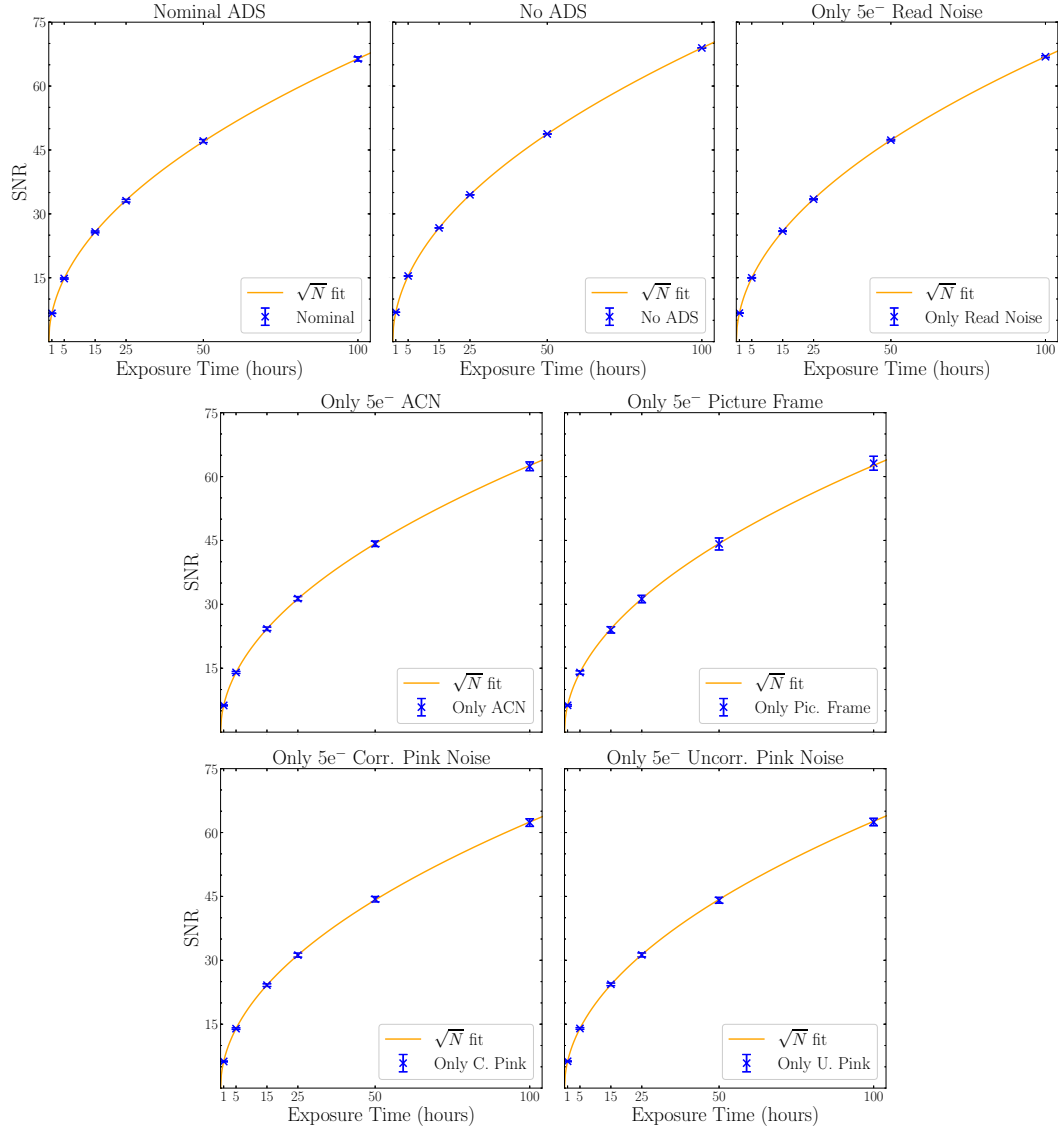


Figure 3.12: SNR as a Function of Integrations with Single ADS Noise Components. Seven panels showing the relative SNR values as a function of integrations for each ADS noise component in turn. *From top left to bottom right:* nominal ADS values, no ADS used, only read noise, only ACN, only picture frame effect, only correlated pink noise, and only uncorrelated pink noise. We also show the best fitting $\sqrt{N}$ function, from which we derive the parameter A and corresponding SNR after 1 hour of observations.

Simulation Type	SNR_{1h}	$\text{SNR}_{1h}/\text{SNR}_{1h,N}$
No ADS	6.895 ± 0.00	1
Nominal ADS	6.643 ± 0.006	0.963
Only $5 e^-$ Read Noise	6.689 ± 0.001	0.970
Only $5 e^-$ ACN	6.260 ± 0.005	0.908
Only $5 e^-$ Picture Frame	6.262 ± 0.016	0.908
Only $5 e^-$ Correlated Pink Noise	6.248 ± 0.005	0.906
Only $5 e^-$ Uncorrelated Pink Noise	6.264 ± 0.009	0.908

Table 3.7: **SNR Scaling with Single ADS Noise Levels.** The SNR at 1 hour, as determined by multiplying the $\text{SNR}(N)$ as defined by Equation 3.2.1 by 2, for single-component ADS simulations. We also give the normalised $\text{SNR}_{1h}/\text{SNR}_{1h,N}$ with $\text{SNR}_{1h,N}$ being the scaling factor without using ADS. To the level of precision of these simulations, there is no uncertainty associated with the no-ADS runs, as the noise cube used in the determination of the SNR is the same in each run.

uniform. As read noise is the most uniform across a detector, reflected in the small uncertainty on the expected SNR, we would expect the overall noise to be similar to that for the nominal ADS parameters. We do see, as expected, that only including read noise in this manner is a reduction in SNR compared to not using ADS, which incorporates read noise in a more simplistic manner. This confirms that increasing read noise is a flat effect across the detectors, corresponding to an overall reduction in SNR, and thus increased exposure time to reach the same level of confidence in an observation.

With $5 e^-$ ACN we do see a reduction in SNR, with 6.260 ± 0.005 expected after one hour, a factor of 0.908 of the no-ADS levels. ACN does not subtract out as uniformly, which is why we see it have a greater impact on the SNR than the read noise. Averaging over the 210 point sources included in our analysis, we see a small uncertainty, implying that the impact of the ACN is relatively consistently affecting all the point sources, which is expected as a result of the way in which it is simulated. Combined with our results from Section 3.4.1 we can say that the impact of ACN is to cause a significant, consistent lowering of SNR. However, in real detectors ACN can affect different amplifier blocks in different ways, therefore characterising the detectors will be very important to ensure that this effect is correctly accounted for in data reduction.

Only including picture frame noise we find an expected SNR of 6.262 ± 0.016 after one

hour, a factor of 0.908 of the no-ADS levels. We would expect the picture frame effect to manifest as a spatially varying noise component, and this is reflected in the results we see here, as well as in Section 3.4.1. We see that the overall SNR is very similar to when we enhance the other components aside from read noise, however the variance is the largest of any single component. This means that for the most part the impact of the picture frame effect is low, but for point sources that are impacted by it, the increased noise is significant. This reinforces the importance of understanding to what extent this effect manifests on the HARMONI detectors, and planning observations such that targets of interest do not fall into badly affected regions.

With pink noise, for both correlated and uncorrelated we see a similar impact on the expected SNR after one hour. We recover an SNR of 6.248 ± 0.005 for correlated pink noise, and 6.264 ± 0.009 for uncorrelated, corresponding to reductions of 0.906 and 0.908 respectively over the no-ADS noise levels. Pink noise is a spatially varying noise, but not one that necessarily manifests in the same way on the same part of a detector in subsequent exposures. This is why we see a greater reduction in SNR compared to the same amount of read noise, and a higher variance.

Correlated pink noise occurs on every amplifier block, whilst uncorrelated pink noise can vary between them, however both behave in the same way with respect to an individual pixel, therefore simulating them both to the same level returns almost the same reduction in SNR. Uncorrelated pink noise varies between amplifier blocks, which leads to the higher uncertainty on the overall SNR compared to correlated pink noise. In practice, correcting for the correlated pink noise can be done through use of the reference pixels at the edge of the detector, whilst uncorrelated pink noise is harder to account for. Ensuring the best possible pink noise correction will be crucial however, to maximise the potential of HARMONI.

3.5 Conclusions

In conclusion, we have created the architecture within `HSIM` to create simulated detectors of the type that will be used in `HARMONI`. These are included in such a way that they can be tweaked and improved to increase the accuracy of the simulations as more information is known about the actual `HARMONI` detectors. As part of this, we have created a new option for running `HSIM`, known as the Advanced Detector Systematics mode (ADS). In Table 3.8 we summarise the noise components that are now included in `HSIM` as a result of this work.

Noise Component	Prior to <code>HSIM</code> v.310	Included in <code>HSIM</code> v.310
Read Noise	✓ ^a	✓
Varying Read Noise	✓ ^b	✓ ^b
Dark Current	✓	✓
Horizontal Bias	✗	✓
ACN	✗	✓
Picture Frame Effect	✗	✓
Hot Pixels	✗	✓
Pedestal Drift	✗	✗ ^c
Persistence	✗	✗
Intermittent Channel Failure	✗	✗
Non-Linearity	✗	✗
Electronic Crosstalk	✗	✗

Table 3.8: Summary of Detector Noise Effects and Current Inclusion in `HSIM`.

A list of identified issues with infrared detectors, and the ability of `HSIM` to simulate these prior to and with version 310. ^aOnly includes flat read noise that is the same for each pixel. ^bIncludes a short and long exposure time read noise. Could easily be modified to include further variations. ^cCould easily be included with the current architecture.

We then created an input cube containing simulated point sources positioned at different places within the `HARMONI` field of view, and ran a number of tests to assess the impact of using the ADS mode. We initially looked at the average impact of the noise components on the point source sensitivity, before looking at spatial variations and the ability to recover the same flux over multiple observations. The last suite of simulations aimed to understand the impact of these detector effects on the recovered SNR of the point

sources.

We produced simulations which included enhanced noise components, as well as single noise components to fully understand their impact. From this, we could determine that all the noise components simulated have an impact on HARMONI observations in different ways. Read noise provides a fairly uniform increase in noise across the entire field of view, with hot pixels potentially skewing data intermittently if not sufficiently accounted for. ACN and picture frame noise provide small but spatially varying effects that impact different parts of the field of view depending on what areas of the detector the light falls on. Pink noise, both correlated and uncorrelated, provides both a significant reduction in SNR as well as being spatially varying, making it the most complicated form of noise simulated.

In order to counteract these, a number of options can be utilised, but must be done carefully so as to not introduce additional problems. Utilising the reference pixels to minimise the effect of pink noise can be done, however as we have seen from observations with KMOS, doing this without knowledge of the reference pixels being subtracted can add more features onto the detectors. Characterising the HARMONI detectors will be very important to understand which parts are susceptible to ACN, and to what extent we see the picture frame effect. This should be considered when observing with HARMONI, and where possible avoid particularly problematic areas of the detectors to be able to best utilise the instrument for the desired science.

Chapter 4

Using SINFONI to Understand the Impact of an Under-Illuminated Slit for Telluric Correction

In this chapter we present an investigation into telluric correction using SINFONI, in particular with the instrument set up such that the slit is under-illuminated when observing the telluric standard star. Initially we cover the motivation behind doing this, and present theoretical models for the expected behaviour of the line spread function (LSF) with both a fully-filled and under-filled slit. We then describe the data that were acquired for this investigation, and detail the analysis we undertook. This covers reducing the SINFONI data to obtain spectra for the standard star and science target, in our case Callisto. We also describe our use of `molecfits` to determine the LSF shape and telluric absorption features.

In our results, we show our Callisto spectrum, corrected in four different ways: with the telluric standard star directly, with `molecfits` correction based on the Callisto spectrum, with `molecfits` correction based on the telluric standard spectrum, and finally with `molecfits` correction based on the telluric standard and broadened to the full LSF. We also divide this by a solar spectrum, assuming Callisto to be primarily reflected sunlight,

and determine the residual sum of squares (RSS) of each of these correction methods. We also investigate differences in the LSF and telluric correction between different spaxels of Callisto data. At the end of the chapter we give our conclusions from this study, highlighting issues with understanding the LSF, and the relevance for HARMONI.

4.1 Background of Investigation

4.1.1 Motivation

Good telluric correction is very important for ground-based instruments, so that we can be certain that features we see in the science spectra are genuine features of that target, and not a result of molecular absorption in the atmosphere. As we described in Section 1.3, one way in which this can be done is through observation of a star with few intrinsic features, known as a telluric standard star. When this spectrum is divided out from the science spectrum, the result should only contain the intrinsic features of that science target.

An assumption of this approach is that the shape of the spectral lines, the line spread function (LSF), is the same between the telluric standard and the science target. Provided the observing spaxel scale is finer than the spatial resolving power of the instrument for both targets, such that the input slit is fully illuminated, then this assumption should hold true. However with good adaptive optics correction and coarse spaxel scales, it is possible to have a point source that does not entirely illuminate the slit. With the VLT the diffraction limit at $1.6\ \mu\text{m}$ has a PSF FWHM of $\sim 50\ \text{mas}$, however the coarsest spaxel scale of SINFONI is $125 \times 250\ \text{mas}$. For the ELT, with a 39 m primary mirror, the FWHM of the diffraction limited PSF at $1.6\ \mu\text{m}$ is almost 10 mas, with HARMONI offering $30 \times 60\ \text{mas}$ spaxel scales. In both these cases, with good adaptive optics correction such that it is approaching the diffraction limit of the telescope, a telluric standard star observed at a coarse spaxel scale will not illuminate the entire slit of the spectrograph uniformly.

In this situation, we would expect the LSF to then vary between the telluric standard

star and an extended science target that fills the entrance slit. This is because the LSF is a convolution of a number of factors that can be represented by

$$\text{LSF} = F_{\text{in}} * F_{\text{spec}} * F_{\text{grat}} * F_{\text{out}} \quad (4.1.1)$$

where F_{in} , F_{spec} , F_{grat} and F_{out} are functions representing the input slit, spectrograph, grating and output slit respectively. The output slit represents the pixellation at the detector. In the next section we investigate this in more detail, and produce theoretical models for SINFONI with both a fully illuminated, and under-illuminated entrance slit.

The work in this chapter is to determine the impact of having an under-illuminated slit, however there could be methods adopted such that it is guaranteed the telluric standard star always fully illuminates the slit of the IFU. This could be done by defocussing the star such that the PSF is sufficiently broad to illuminate the slit evenly, or it could be done by moving the star uniformly across the entrance slit during the exposure. This does have to be done with care, to ensure the resulting telluric spectrum is still of sufficient SNR to use for correction (Uttenthaler et al., 2010) .

4.1.2 Theoretical Models

As we have mentioned, the LSF by which a spectral line is broadened is a convolution of a number of functions, each representing a different component in the entire spectrograph. In order to understand the potential impact of having an under-filled input slit on the LSF, we have created a model LSF, where each of the components given in Equation 4.1.1 are described by an appropriate functional form. In Figure 4.1 we show the functional forms used, and the resulting convolutions for both a full and under-full input slit. The reasoning for these choices for each component is as follows:

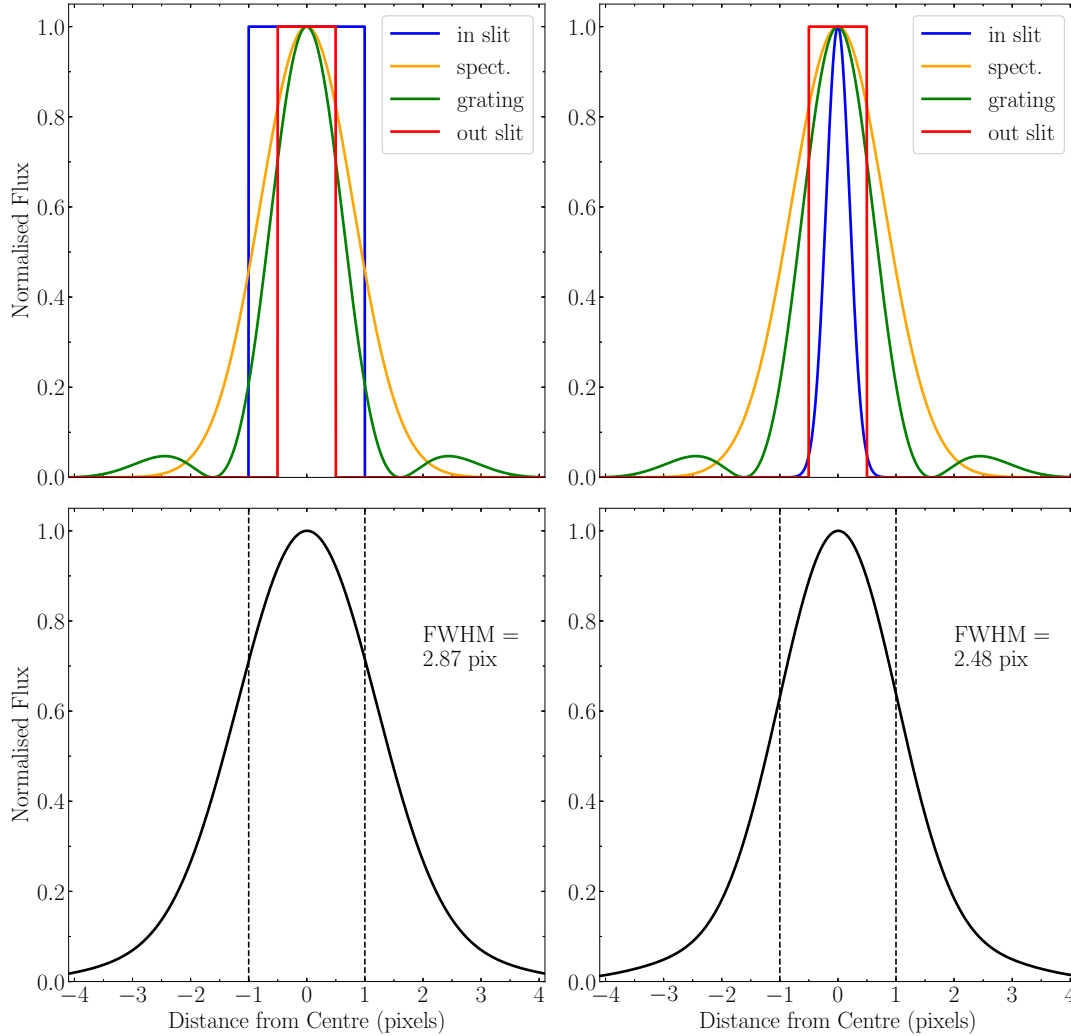


Figure 4.1: Components and Convolutions of Model SINFONI LSFs with Full and Under-full Illumination. *Top:* Components that make up the model SINFONI LSFs for the fully illuminated input slit (*left*) and under-filled slit (*right*). These consist of the input slit (blue), spectrograph (orange), grating (green), and output slit (red). The input slit is changed to a 0.48 pixel FWHM Gaussian function, as opposed to a 2 pixel boxcar function in the latter panel. *Bottom:* Total convolutions of the four components in the upper panels for the fully illuminated (*left*) and under-filled (*right*) input slit. The result of under-illuminating the input slit is a reduction in the FWHM of 14%. Dashed lines denote ± 1 pixel from the centre.

Input Slit.

The input slit is the only function that varies between the two modelled cases, as this is the function that determines the level of illumination. In the case of a fully illuminated input slit, this is equal in width to two pixels, and thus we model this using a boxcar function, where

$$F_{\text{in}}(x) = \begin{cases} 1, & -1 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (4.1.2)$$

and x is given in terms of pixels.

For the under-illuminated input slit we instead model the spectrum using a normalised Gaussian function centred on 0, given by

$$F_{\text{in}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (4.1.3)$$

where σ is the standard deviation of the function, related to the full width at half maxima (FWHM) through

$$\text{FWHM} = \sigma\sqrt{8\ln 2}. \quad (4.1.4)$$

The value chosen here for the FWHM was 0.48 pixels, equal spatially to 60 mas. This is a little broader than the diffraction limit of the VLT at the wavelengths we are investigating, which would be 43 mas at $1.7\ \mu\text{m}$. However, in Section 4.2 we show that this is the performance of the instrument at the time of observation.

Spectrograph.

The next component which we model is the impact of the spectrograph. To a first approximation, we assume that the effect of the spectrograph is Gaussian and circular, meaning it broadens the same both spatially and spectrally. We therefore modelled this using the same functional form as Equation 4.1.3, adjusting the width accordingly. By looking at the raw observed data frame, we could identify the slitlet in which the majority of the

light falls. We then used this data to fit a Gaussian function across the spatial axis using a non-linear least squares method from `SciPy`, with the amplitude, width, central position and constant background as free parameters. We took the average of 100 wavelength channels to recover a corresponding FWHM of 1.9 pixels, equal to a σ of 0.8 pixels. We then used this value for the width of the spectrograph functional form in the model.

Grating.

For the grating, initially we modelled this as a perfect grating, with a form given by the diffraction equation

$$F_{\text{grat}} = \frac{\sin^2(N\delta/2)}{(\delta/2)^2} \quad (4.1.5)$$

with

$$\delta = \frac{2\pi}{\lambda} b \sin(\theta)$$

for N slits with spacing b at wavelength λ . However, it has been seen for SINFONI that the grating functions are imperfect as a result of cryogenic deformation (George et al., 2017). The exact nature of the deformation causes a non-trivial effect on the shape of the grating function, therefore we have implemented an approximation in this model. We used the reference material contained within the SINFONI User Manual¹, which contains line profiles for each grating and spaxel scale. In Figure 4.2 we reproduce the section showing the H-band, 250 mas line profiles. From these we digitised the data and determined the overall FWHM for the LSF at the upper end of the H band.

In addition, we used the SINFONI wavemap to determine the exact wavelength spacing for the slitlet in which our data falls, so that we could convert this measured FWHM into pixels. This gave us FWHM values of 2.87 at $\lambda \sim 1.7 \mu\text{m}$ and 2.58 at $\lambda \sim 1.8 \mu\text{m}$. This variation is of a similar level to the effect we wish to investigate, therefore it is important that we are consistent in the treatment of the gratings, and the wavelength ranges used. We then manually adjusted the width of the grating functions such that the overall

¹Available at: <https://www.eso.org/sci/facilities/paranal/decommissioned/sinfoni/doc/VLT-MAN-ESO-14700-3517-P103.pdf>

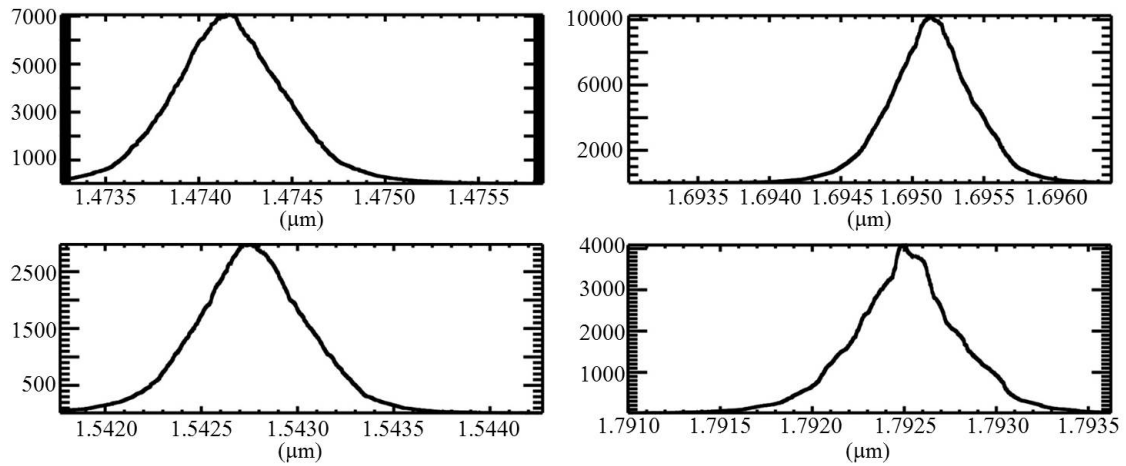


Figure 4.2: **SINFONI Line Spread Functions in the H-Band at 250 mas Spaxel Scale.** Four panels showing the LSF at different wavelengths within the H-band with SINFONI at the 250×125 mas spaxel scale. *Image reproduced from the SINFONI User Manual.*

convolution equalled the desired width for the fully illuminated entrance slit.

Output Slit.

The output slit corresponds to the detector pixellation, and therefore with the scaling used can also be represented by a boxcar of form

$$F_{\text{out}}(x) = \begin{cases} 1, & -0.5 \leq x \leq 0.5 \\ 0, & \text{otherwise,} \end{cases} \quad (4.1.6)$$

where the width is half that of the fully illuminated input slit.

Overall Convolution

The lower panels in Figure 4.1 show the results of convolving these functional forms together. We have fixed the width of the fully illuminated LSF to be equal to 2.87 pixels at $\lambda=1.7 \mu\text{m}$ by adjusting the grating function as described above, which corresponds to a spectral resolving power, R , of 3040 at $1.7 \mu\text{m}$. When we then use these same functional forms, but instead convolve with a Gaussian input function representing an under-filled

slit, we obtain a narrower LSF as expected. We calculate a FWHM for this convolution of 2.48 pixels, corresponding to an R of 3520. The reduction in FWHM is 0.39 pixels, equal to 14% of the width of the fully illuminated LSF.

Importantly, the resulting LSF is still Nyquist sampled, as there are more than 2 sampled points per FWHM of any observed feature. However it does mean that dividing sets of data that differ in the slit illumination should result in an improper telluric correction.

4.1.3 Data Acquisition

To investigate the impact of this effect on telluric correction, we used SINFONI on the VLT to observe the star CD-23 13701, an O9III star suitable for use as a telluric standard, and Callisto as a bright, extended science object. These observations were taken the night of May 30th 2019 as a calibration proposal under Program ID 4102.L-0887(A), and were taken close to consecutively so as to ensure the atmospheric profile was as similar as possible between them. The seeing varied between 0.8'' and 1.3'' for the observations, however all were taken using NGS AO correction. The airmass varied between 1.082 and 1.172 across the observations. The list of observations are summarised in Table 4.1. All the data taken are available for download from the ESO archive.

We took both J-band and H-band data for CD-23 13701 and Callisto, however for this work we only consider the H-band. This is due to the LSF being narrower, thus the spectral resolution is higher, so we would expect to see a greater difference as a result of varying slit illumination. For CD-23 13701 we took data at each of the spaxel scales of SINFONI, 250×125 mas, 100×50 mas and 25×12.5 mas, whilst for Callisto we took data at the coarsest and finest scales.

Our main interest was seeing how significant the impact the entrance slit being under-filled is for telluric correction. We therefore primarily used the H-band, 250×125 mas data, hereafter denoted by H250. We additionally used the H-band 25×12.5 mas (H25) data for the determination of the spatial extent of CD-23 13701, as described in Section 4.2.1.

Target	Grating	^a Spaxel Scale	Start Time	^b T _{exp}	Airmass
Sky	H	250×125	06:37:37	50	1.169-1.167
Callisto	J	25×12.5	07:05:50	5	1.082-1.085
Callisto	J	250×125	07:07:36	5	1.085-1.087
Callisto	H	250×125	07:08:51	5	1.095-1.097
Callisto	H	25×12.5	07:13:58	5	1.097-1.100
CD-23 13701	J	25×12.5	07:55:11	50	1.101-1.120
CD-23 13701	J	100×50	08:04:27	5	1.121-1.124
CD-23 13701	J	250×125	08:06:16	5	1.124-1.129
CD-23 13701	H	250×125	08:08:10	5	1.139-1.142
CD-23 13701	H	100×50	08:13:49	5	1.143-1.147
CD-23 13701	H	25×12.5	08:15:43	50	1.148-1.172

^a mas ^b s

Table 4.1: **Summary of SINFONI Observations.** List of targets observed using SINFONI, with the respective gratings and spaxel scales chosen, along with the time of observation, exposure time T_{exp}, and airmass range over the observation. All data were taken the night of May 30th 2019, and are available for download from the ESO archive as Program IDs 60.A-9235(A) (sky) and 4102.L-0887(A) (Callisto and CD-23 13701).

4.2 Data Reduction and Analysis

All the data were reduced using the SINFONI data reduction pipeline, using the `Reflux` environment. This associates appropriate calibration files together and provides many of the required corrections to the data. This includes flat fielding, dark subtraction, wavelength calibration, bad pixel masking, and datacube reconstruction. In addition to these data, we also reduced a separate exposure for use in sky subtraction which was taken earlier that night, `SINFO.2019-05-31T06:41:49.758`. This was a 50 s exposure taken using the H-band grating at 250×125 mas spaxel scale, under Program ID 60.A-9235(A).

For a comparison spectrum with CD-23 13701 we used a star of the same spectral type, HR 1899, using data from Meyer et al. (1998), which used a Fourier Transform Spectrometer (FTS) on the Mayall 4m telescope at the Kitt Peak National Observatory to obtain spectra at a resolving power of 3000. We used this to identify areas in the spectrum which contain intrinsic features from the star, which must be masked out prior to using it for telluric correction.

We modelled Callisto as a reflected solar spectrum, using a computed solar irradiance spectrum (Kurucz, 2005; Chance and Kurucz, 2010). We took a spectrum at $R=30000$ and used a Gaussian function to convolve it to a resolving power of $R=2920$. This was to match the average resolving power of the SINFONI H-band, as quoted in the SINFONI User Manual, such that we could divide it out of the recovered Callisto spectrum.

4.2.1 Determining the Spatial PSF Width

After reducing the data, and prior to extracting the spectra, we used the H25 observation of CD-23 13701 to confirm the performance of the AO and to determine the extent to which the H250 data was under-illuminated. In Figure 4.3 we show the raw data at $\lambda=1.7\ \mu\text{m}$, along with the best fitting results. The H25 data is ten times finer sampled than the H250 data, and with a diffraction limit of $\sim 40\ \text{mas}$, should be sufficiently sampled such that we could fit the data to determine the FWHM of the PSF, provided we had good AO correction.

We used the non-linear least squares fitting method `curve_fit` in the `SciPy` package to fit a 2-dimensional Gaussian function to the data. We allowed this fit to have a different standard deviation along the two axes of the fit, and for these axes to be rotated with respect to the slit orientation, to account for effects such as AO PSF elongation. We see from this that a Gaussian function fits the core of the PSF very well, although it is not representative of the wings, as we would expect. We also see the effect of the rectangular pixels, where the data across the slit is half as well sampled as that along it.

As the core is well-represented, we could use the recovered best-fitting parameters to determine the FWHM of CD-23 13701 in the H25 data. We then assumed this to be the width of the star in the H250 data for determining the fraction by which the slit is under-illuminated. We recovered a FWHM in the direction along the slit of $58\ \text{mas}$, and a FWHM across the slit of $59\ \text{mas}$. With a spaxel being $25 \times 12.5\ \text{mas}$, and taking the FWHM to one significant figure at $60\ \text{mas}$, this corresponds to a FWHM of 4.8 pixels in the H25 data. Thus we assume the same performance on the H250 data, such that

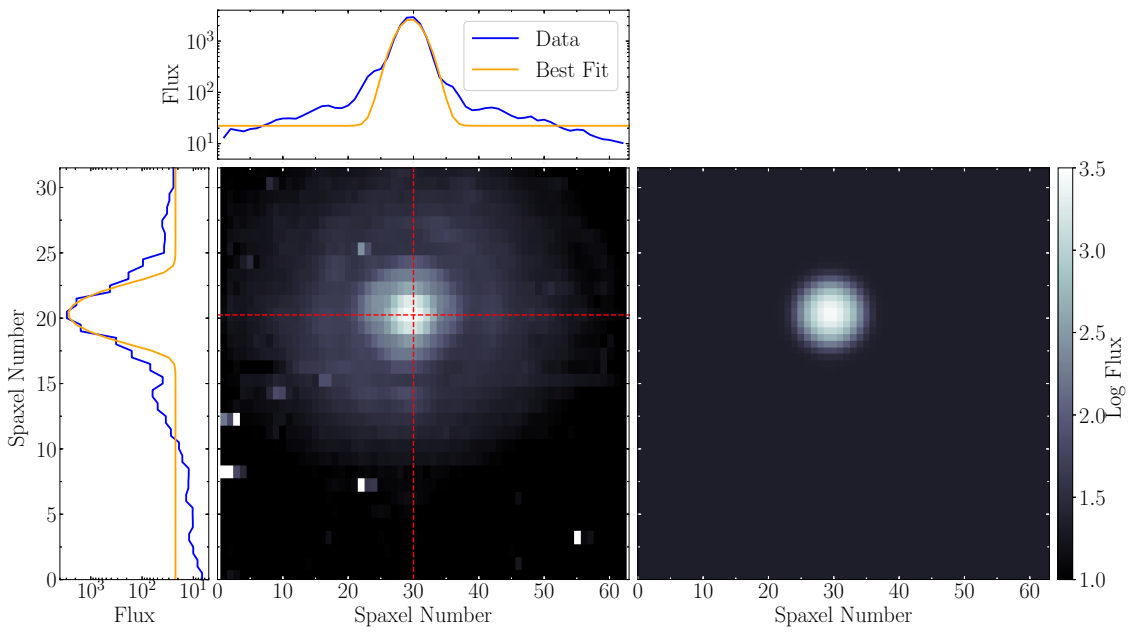


Figure 4.3: Reduced CD-23 13701 H25 Data and Best Fit Function. *Bottom, centre:* Reduced image showing CD-23 13701 at $1.7\ \mu\text{m}$ and observed at the 25×12.5 mas spaxel scale. Red dashes indicate the lines along which the data in the popout plots are taken along. *Bottom right:* Best fitting 2D Gaussian function to the reduced data. *Top:* 1D cut along the horizontal dashed line, showing the real data (blue) and best-fitting 2D Gaussian (orange) along the same cut. *Bottom, left:* Same as previous, for the vertical dashed line. All data are shown on a logarithmic scale.

the FWHM of CD-23 13701 is 0.48 pixels. This is the justification for our value in Section 4.1.2, and we can be confident that the AO is performing well, and the input slit of the spectrograph is significantly under-illuminated in the H250 CD-23 13701 data.

4.2.2 Extracting Spectra

In Figure 4.4 we show the reduced H250 datacubes at a wavelength of $1.7\ \mu\text{m}$ for CD-23 13701 and Callisto. As we expected, the majority of light from the star falls into one spaxel, whilst Callisto is extended across several. In both cases we took the light from a single spaxel, indicated by the red boxes in Figure 4.4, and extracted the corresponding spectra for each. We extracted the spectra from the same spaxel in each dataset to ensure we minimise additional variations. We would expect small variations in the LSF between spaxels as the light will take slightly different paths through the instrument, which we consider in Section 4.3.4. There may also be some differences in the Callisto spectrum due

to surface properties and intrinsic features of Callisto. Using only one spaxel minimises the impact that this may have on our findings.

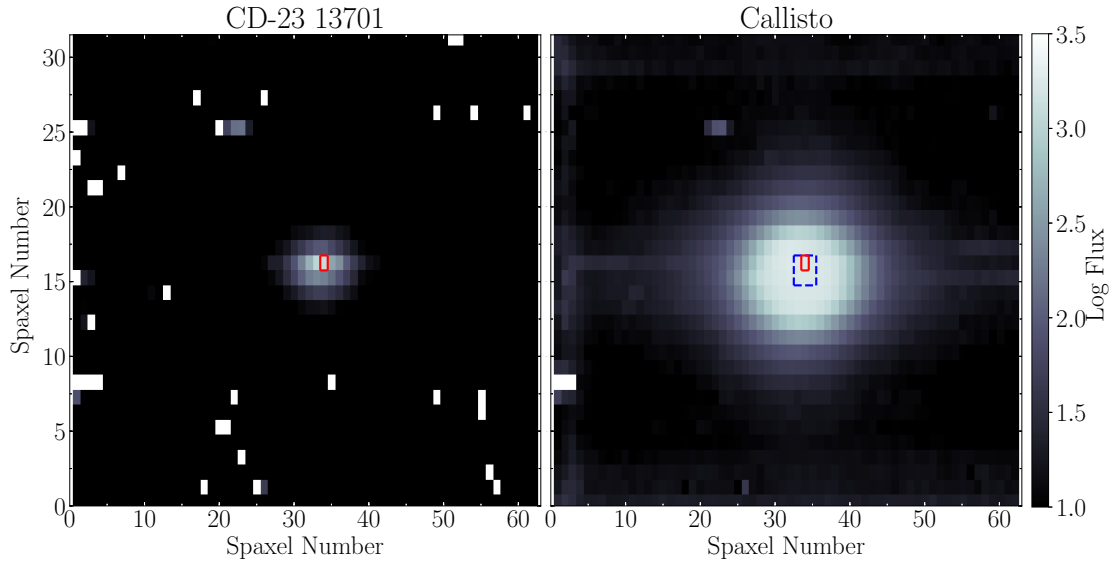


Figure 4.4: Images of CD-23 13701 and Callisto at $1.7 \mu\text{m}$ at 250 mas Spaxel Scale. Single-wavelength images of CD-23 13702 (*left*) and Callisto (*right*) at $1.7 \mu\text{m}$ when observed with the 250×125 mas scale of SINFONI. Data are shown on the same logarithmic scale. Red boxes indicate the primary spaxels selected for extracting the spectra. The blue dashed box in the right panel indicates the six spaxels used in Section 4.3.4.

The correct wavelengths for the spectral data were obtained using the appropriate values from the FITS header. For sky subtraction we used the sky image mentioned previously, extracting the information from the same spaxel as for the spectra. As the sky observation had a longer exposure time, we corrected the sky flux by a factor of ten accordingly, and visually confirmed that we have removed the sky emission effectively.

The resulting spectra, with continuum normalised, are shown in Figures 4.5 and 4.6 for CD-23 13701 and Callisto respectively. We also show the sky spectrum that has been subtracted out, along with appropriate reference spectra of HR 1899 and solar irradiance. The fourth spectrum in each plot is the result of using `molecfit` on the data, the process for which we will discuss in Section 4.2.3.

We can see in Figure 4.5 that the recovered spectrum from CD-23 13701 primarily consists of telluric features, however there are some features intrinsic to the star that are present. We would expect some hydrogen lines, and have marked the wavelengths

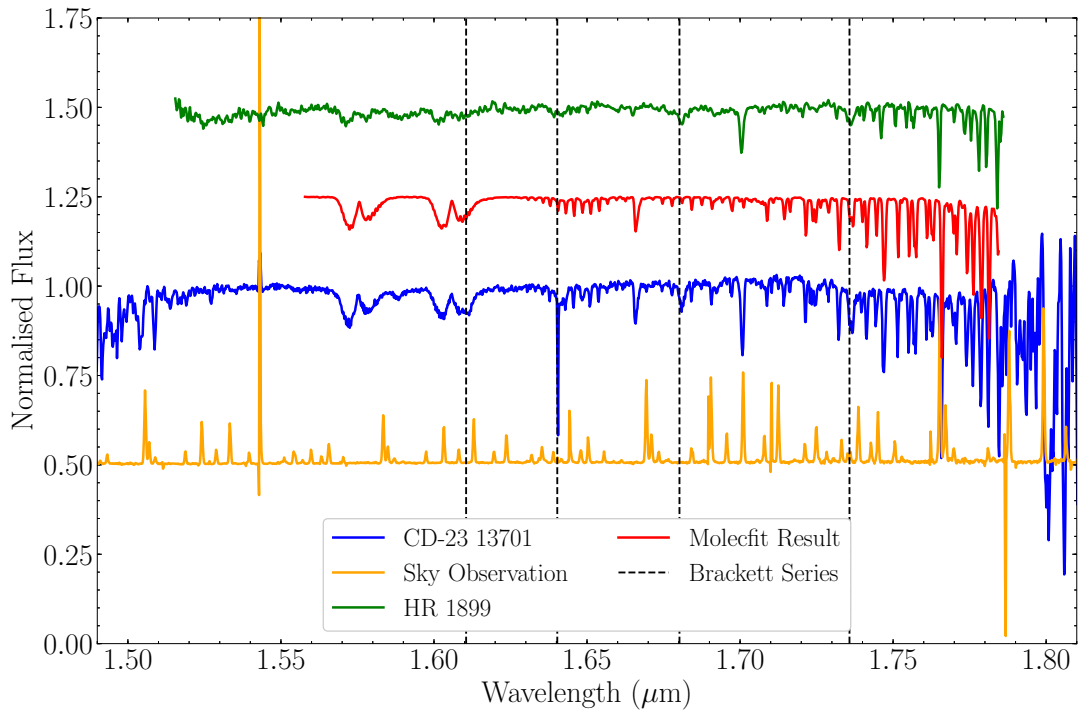


Figure 4.5: Recovered Spectrum for Telluric Standard Star CD-23 13701. The recovered spectrum from CD-23 13701 (blue) with continuum removed and sky subtracted, but with molecular absorption present. We also show the subtracted sky (orange, normalised to 0.5) and best fitting `molecfit` result (red, normalised to 1.25). For comparison we show the spectrum of another O9III star, HR 1899 (green, normalised to 1.5), which has had some telluric correction applied. Lastly we indicate the location of the hydrogen Brackett series lines, where we would expect intrinsic features to be present.

at which the strongest Brackett series lines within this wavelength range would appear, and there is some corresponding absorption features in the spectrum. There is also an absorption feature at $\sim 1.7 \mu\text{m}$, corresponding to a helium line (Litzén, 1970), that must be masked when using the spectrum for telluric correction. There is an anomalous, narrow feature around $1.65 \mu\text{m}$, however this falls within a wavelength region corresponding to a Brackett line, thus will be masked out prior to telluric correction.

With the Callisto spectrum in Figure 4.6 we see a lot more features, as we would expect. We can associate these with both telluric absorption, along with features seen in the solar irradiance spectrum. With these extracted spectra, we could then look at correcting for the telluric absorption both with direct spectrum division, and with using `molecfit`.

4.2.3 Telluric Modelling with `Molecfit`

As we introduced in Section 1.3, `molecfit` (Smette et al., 2015; Kausch et al., 2015) is a tool for determining the molecular absorption in the atmosphere, and correcting spectra accordingly. `Molecfit` is versatile, and can be used either on a telluric standard star for correcting a science spectrum, or directly to the science target. It can also fit the LSF of the given spectrum.

We converted the spectra to the input format required by `molecfit`, and set up the parameter file with the relevant information regarding observing location and time. For the wavelength extent we chose the range $1.5575 \mu\text{m}$ to $1.7843 \mu\text{m}$. This covers the majority of the telluric features in the H-band where we are confident in our spectrum recovery. We then chose the relevant molecules to include in the fit, which for the H-band was H_2O , CO_2 , and CH_4 , and allowed the molecular abundances to be determined by `molecfit`. We used the precipitable water vapour (PWV) information from the weather report from the night of observation as the initial input guess.

For the LSF, we experimented with a number of different line profiles, primarily either a Gaussian or a Lorentzian function. For our final results, we chose to use a Gaussian LSF

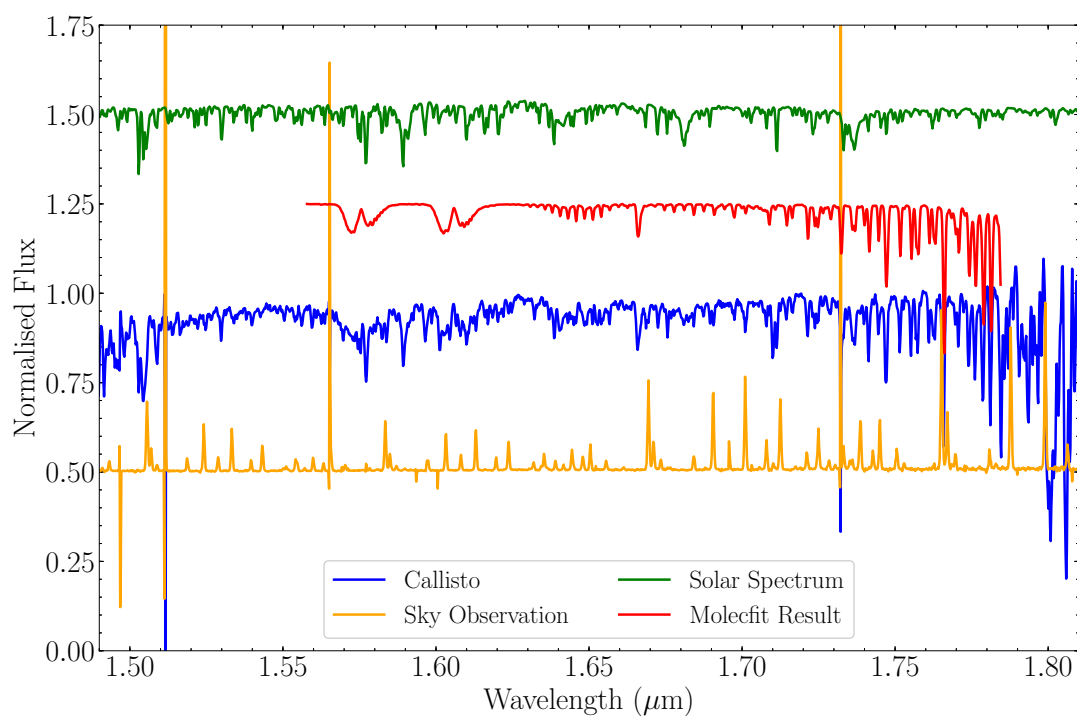


Figure 4.6: Recovered Spectrum for Extended Science Object Callisto. The recovered spectrum from Callisto (blue) with continuum removed and sky subtracted, but with molecular absorption present. We also show the subtracted sky (orange, normalised to 0.5) and best fitting `molecfit` result (red, normalised to 1.25). For comparison we show a solar irradiance spectrum convolved with a Gaussian function to $R=2920$ (green, normalised to 1.5).

profile, where we allow `molecfits` to determine the best-fitting width of this profile. In Figure 4.7 we present the justification for this, by showing an investigation of the reduced χ^2 of the fit as a function of LSF FWHM for both a Gaussian and a Lorentzian function.

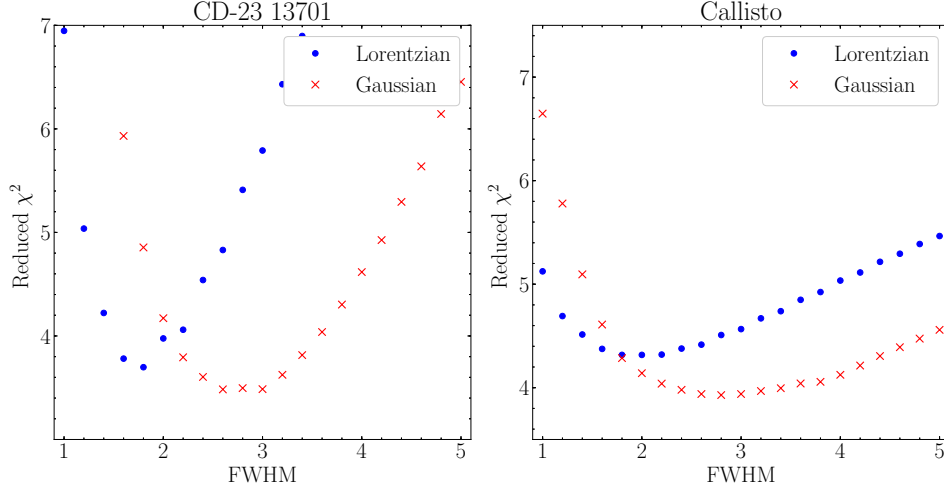


Figure 4.7: **Molecfits χ^2 Values for Gaussian and Lorentzian LSFs on CD-23 13701 and Callisto.** *Left:* Reduced χ^2 values returned from `molecfits` when using a Gaussian (blue circles) or Lorentzian (red crosses) LSF profile to fit the CD-23 13701 H250 data. *Right:* Same, for the Callisto H250 data.

To determine this, we took the best-fitting molecular profiles from the initial runs, and then used an LSF of fixed width for a given functional form to run `molecfits`. We then adjusted the width of the provided LSF, and observed how the reduced χ^2 of the fit changed. We repeated this with both sets of data, and as can be seen, in each case a Gaussian form returns the better fit. We see that in general `molecfits` is providing a better fit to the CD-23 13701 data, compared to Callisto, both in terms of overall reduced χ^2 , as well as a more defined global minimum. We would expect this, as the majority of features in CD-23 13701 are the telluric features themselves, whereas Callisto contains a large number of intrinsic features that can reduce the confidence of the direct fit.

We also confirm from this that the recovered LSF is indeed the best fit in terms of reduced χ^2 , corresponding to the global minimum. We see that the χ^2 increases more rapidly at smaller FWHMs, with a shallower gradient towards larger values. This is to be expected, as a narrow line profile will always trend away from the true data, whereas broader LSFs will sometimes align with adjacent absorption features, particularly when

multiple lines are close together. We see this effect with both datasets, and with both choices of LSF.

4.3 Results and Discussion

4.3.1 Molecfit LSF Profiles

We have shown the `molecfit` best fitting absorption spectra in Figures 4.5 and 4.6, and in Figure 4.8 we show a section of this result over-plotted against the same section for both CD-23 13701 and Callisto, along with the residuals. In this wavelength range of $1.740\ \mu\text{m}$ to $1.760\ \mu\text{m}$ we would expect the majority of features present to be H_2O absorption lines, with minimal intrinsic features from either target, although there are a few shallow features present in the reflected solar spectrum. Here we can directly see how well `molecfit` has recovered the depth and width of the telluric features.

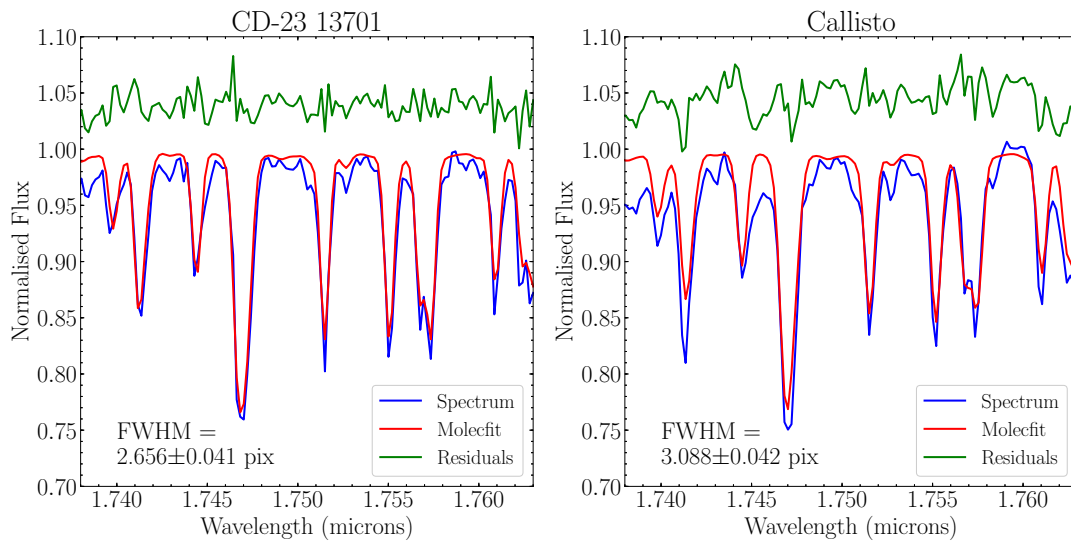


Figure 4.8: Reduced Spectra and Molecfit Results Between $1.74\ \mu\text{m}$ and $1.76\ \mu\text{m}$. *Left:* Reduced spectrum (blue) and best fitting `molecfit` result (red) for the CD-23 13701 H250 data. Shown above are the residual differences between the two spectra (green). *Right:* Same, for the Callisto H250 data. Quoted FWHM values are based on the `molecfit` best-fitting LSF width.

We see in Figure 4.8 that the absorption features are well reproduced by `molecfit` for both CD-23 13701 and Callisto. In each of these cases, the `molecfit` output is a

result of the molecular absorption lines convolved by the best-fitting LSF for each dataset. As we explained previously, we have used a Gaussian LSF for both, with the width as a free parameter to the fit.

For CD-23 13701 the width of this LSF is returned as 2.656 ± 0.041 pixels, with an overall reduced χ^2 of 3.45. For Callisto, the LSF width is 3.088 ± 0.042 pixels, with a reduced χ^2 of 4.32 for the full fit. This corresponds to spectral resolutions of $R_{\text{CD}} = 3282^{+52}_{-50}$ and $R_{\text{Cal}} = 2823^{+39}_{-38}$. In terms of the difference in FWHM, we find a result of 0.432 ± 0.059 pixels, corresponding to a reduction in FWHM of $14.0 \pm 2.1\%$ between the two LSFs.

These LSFs are slightly broader, and thus we recover a slightly worse spectral resolution than predicted by our models. This is likely to due to a combination of the models being slightly simplified in comparison to the true LSF, imperfections in the data skewing the fit, and variations in the LSF between spaxels. If we refer to the SINFONI User Manual, the averaged spectral resolution across the H250 grating is 2920, with which we are in much closer agreement with. We also show in Section 4.3.4 that there is variation in the LSF with between spaxels, therefore it is not unexpected that our recovered LSF is not exactly equal to the average spectral resolution.

What we confirm from these results, however, is that we are seeing a difference in LSF between the under-illuminated slit and the fully-illuminated slit. With a value of $14.0 \pm 2.1\%$ we are in agreement with our modelling in terms of the difference we would expect. This gives us confidence that our modelling is close to accurate, with a small systematic difference in the data that is leading to a broader recovered LSF for both.

4.3.2 Telluric Corrected Callisto Spectrum

In Figure 4.9 we show the resulting Callisto spectrum, where we have applied four different telluric correction methods. These are: direct correction with CD-23 13701, correction with the `molecfit` result from fitting the Callisto spectrum, correction with the `molecfit` result from fitting the CD-23 13701 spectrum, and this latter correction with

a modification to the LSF. We compare this with the same solar irradiance spectrum which we are using as a model for Callisto's intrinsic spectrum.

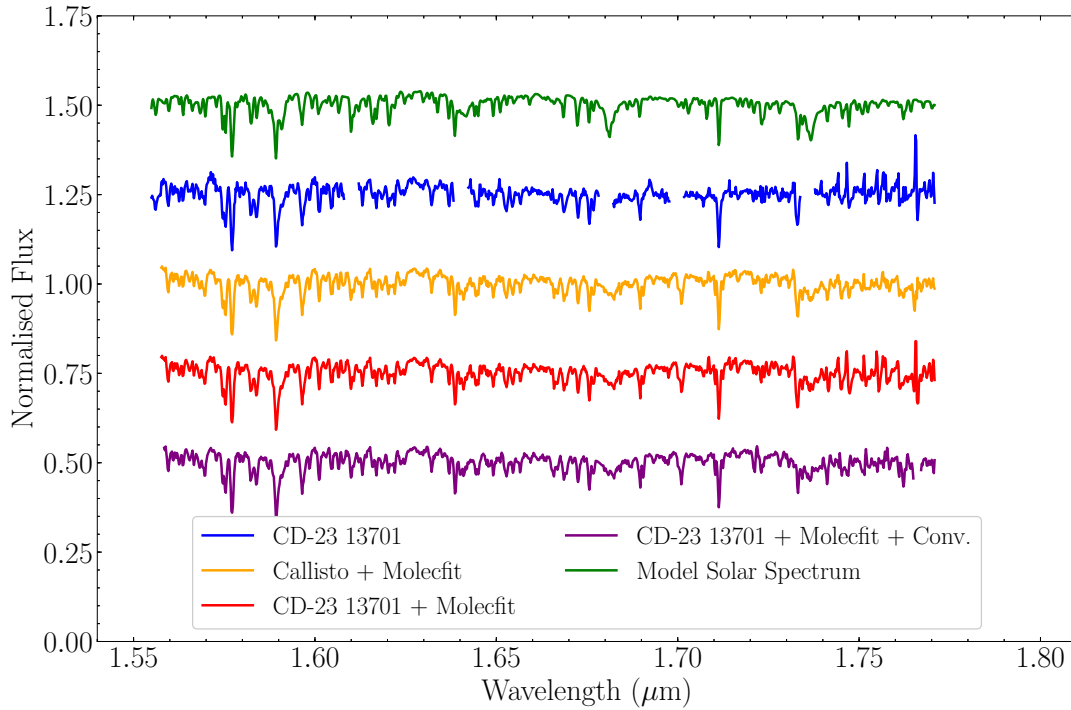


Figure 4.9: Telluric Corrected Callisto Spectrum, Using Four Different Telluric Correction Methods. Four spectra with different methods for telluric correction, along with the model solar irradiance spectrum (top, green). In order below this we show the spectrum corrected using: just the telluric standard star, CD-23 13702 (blue), `molecfit` direct to the Callisto spectrum (orange), `molecfit` to the CD-23 13701 spectrum (red), and `molecfit` to the CD-23 13701 spectrum convolved to the Callisto LSF (purple). Spectra are all normalised to a flux of 1, and manually shifted higher or lower to accommodate displaying them together.

With CD-23 13701, we initially mask out the areas we have identified as containing intrinsic features, which are the Brackett lines at $1.6105\ \mu\text{m}$, $1.6403\ \mu\text{m}$, $1.6802\ \mu\text{m}$, and $1.7357\ \mu\text{m}$, as well as the He line at $1.7002\ \mu\text{m}$. We then divide the recovered Callisto spectrum by this, obtaining the resultant spectrum shown by the blue line in Figure 4.9.

For the `molecfit` division we use the returned absorption spectrum, which is convolved to the fitted LSF for each dataset. With this, we do not need to mask out any parts of the spectrum due to intrinsic features, although we restrict our correction to the spectral range over which we have run `molecfit`. These resultant spectra are shown as the orange (`molecfit` on Callisto) and red (`molecfit` on CD-23 13701) lines in

Figure 4.9. We include the fit to CD-23 13701 despite knowing the LSF is incorrect, as both a comparison, and to determine the extent by which the correction is inaccurate.

Finally, we consider the case where we use `molecfits` on the telluric standard star, CD-23 13701, and then combine that with knowledge of the LSF prior to correction. We convolve the `molecfits` result from CD-23 13701 with a Gaussian function such that it has a FWHM of 3.088 pixels, and then divide this out from the Callisto spectrum. This is shown by the purple line in Figure 4.9.

Visually, these methods all appear to provide a reasonable telluric correction, although there are some differences, particularly where telluric features are abundant, such as around $1.75 \mu\text{m}$. With the CD-23 13701 division, and corresponding `molecfits` run, we can see features that are not present in the solar irradiance, or the other correction methods. In the next section we look in more detail at the adequacy of each method.

4.3.3 Best Methods for Telluric Correction

In order to establish the quality of these telluric correction methods, we then further divided these Callisto spectra by the model solar irradiance, shown in Figure 4.10. The residuals we see here are from a combination of insufficient telluric correction, any intrinsic Callisto features that differ from the solar irradiance, and noise in the observations. We assume that the residuals are dominated by the telluric correction, with the Callisto features providing a small contribution that will be constant between each spectrum.

If we had a perfect correction, and there were no residual intrinsic features, we would expect the spectrum to be entirely flat, corresponding to a normalised flux of 1. We can use this as a model to determine the residual sum of squares (RSS) of each of the telluric correction methods, which are summarised in Table 4.2. This is a measure of the discrepancy between the data and the model, given by (Archdeacon, 1994)

$$\text{RSS} = \sum_{i=1}^n (y_i - f(x_i))^2 \quad (4.3.1)$$

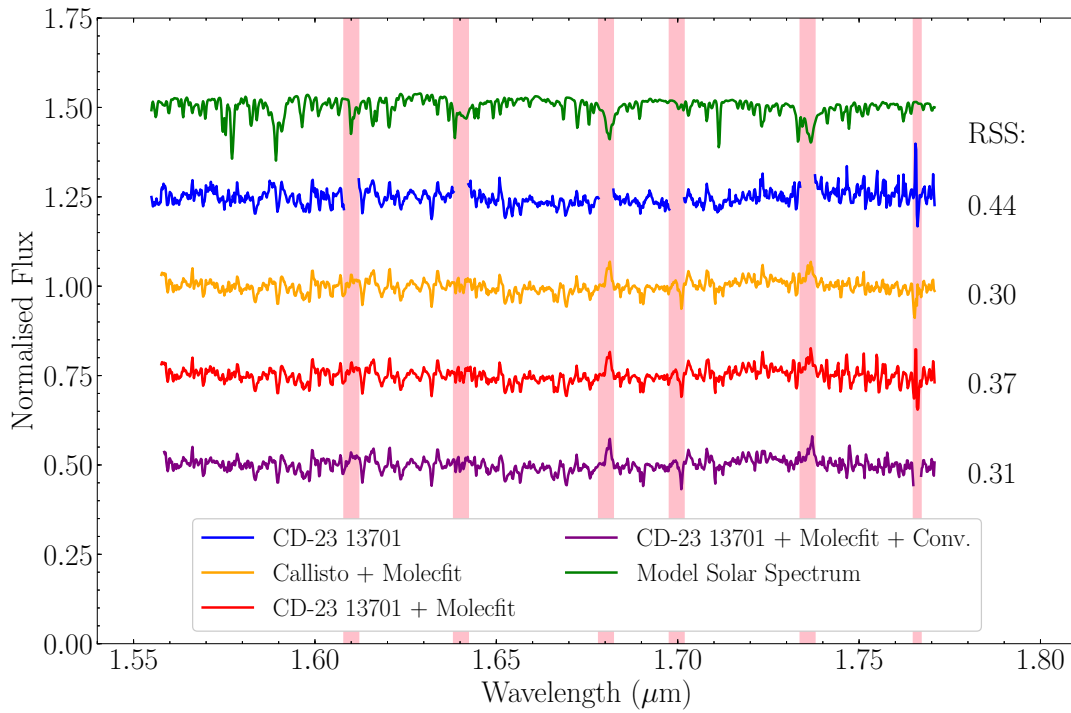


Figure 4.10: Residuals of Callisto Spectrum for the Different Telluric Correction Methods. The residuals for the four spectra shown in Figure 4.9 when divided by the model solar irradiance spectrum (green, top). In order below this spectrum we show the residuals from using: just the telluric standard star, CD-23 13702 (blue), `molecfit` direct to the Callisto spectrum (orange), `molecfit` to the CD-23 13701 spectrum (red), and `molecfit` to the CD-23 13701 spectrum convolved to the Callisto LSF (purple). Spectra are all normalised to a flux of 1, and manually shifted higher or lower to accommodate displaying them together. The RSS values are determined from assuming a perfect correction corresponding to a normalised flux of 1 along the entire spectral range. Areas marked in pink are excluded from the RSS calculation.

where y_i is the recovered spectrum, and $f(x)$ is equal to 1 at all points. We ensure in the determination of this value that we include data from the same regions, excluding areas in which one correction method does not provide any spectral information. These are indicated by the pink regions in Figure 4.10.

For direct correction with CD-23 13701, we recover an RSS of 0.44, which is the worst of the four methods used. As we know the LSF varies between the two datasets, this is what would expect to see. As well as this, there will be some residual noise within the CD-23 13701 spectrum that is propagated through, additionally increasing the RSS.

When we instead use the `molecfit` result from CD-23 13701 but do not adjust

Correction Method	RSS
CD-23 13701	0.44
CD-23 13701 + <code>Molecfit</code>	0.37
CD-23 13701 + <code>Molecfit</code> + LSF Conv.	0.31
Callisto + <code>Molecfit</code>	0.30

Table 4.2: **Residual Sum of Squares for Different Telluric Correction Methods.** Residual sum of squares from dividing a solar irradiance spectrum from the Callisto spectra that were telluric corrected by: direct telluric standard division, using `molecfit` on the telluric standard star, convolving that result to the expected LSF, and using `molecfit` directly to the science data.

the LSF, we find an RSS of 0.37. This is an improvement over using the CD-23 13701 spectrum directly, as a result of the `molecfit` spectrum containing just the modelled absorption, free of any residual intrinsic features. However, this is still an inferior correction compared to the remaining two methods, as we are still assuming the wrong LSF in the correction.

Using the same `molecfit` result, but convolving it to the fully illuminated LSF we recover an RSS of 0.31. This is a further improvement over using `molecfit` on CD-23 13701 as expected, since we are now more accurately representing the LSF of the absorption features. In the absence of the ability to use the science data itself, this is the best telluric correction we could achieve.

When we use `molecfit` directly to the Callisto data we recover a very similar, although slightly smaller, RSS of 0.30. We would expect this to provide a good correction, as we are modelling the correct LSF and using the `molecfit` spectrum for the correction. In addition to this, as we are fitting the science data directly, we are fitting the exact molecular gas profile at the time of observation. Whilst the data were taken almost consecutively, there was approximately one hour between the H250 Callisto data and the H250 CD-23 13701 data (Table 4.1), and a slight difference in airmass. The relative molecular gas column densities can vary over these time scales (e.g. Thoning et al., 1989), and we show in Table 4.3 the recovered best fitting values for each of the observations.

We can see from this that best-fitting relative molecular gas column densities vary slightly between the two exposures. Of the three molecules that are prevalent within the

Molecule	Relative Gas Column Densities	
	CD-23 13701	Callisto
H ₂ O	0.819±0.014	0.828±0.013
CO ₂	1.306±0.026	1.234±0.026
CH ₄	0.808±0.036	0.868±0.036

Table 4.3: **Relative Molecular Gas Column Densities from `molecfit`.** Relative molecular gas column densities for the atmospheric profiles at the time of the CD-23 13701 and Callisto H250 observations, as determined by `molecfit`.

spectral range we have looked at, the H₂O and CH₄ gas columns are the same to within 1σ , whilst the CO₂ agrees within 2σ . At the level of correction that we are achieving, however, these small differences will correspond to slight variations in the divided spectra. It is therefore expected that we would not recover exactly the same correction using CD-23 13701 as using Callisto, even if the LSF was the same. This furthers the importance of understanding the LSF when using a different target for telluric correction, so as to minimise the residual telluric features that can remain.

4.3.4 Comparison with Other Spaxels

With Callisto being an extended object at the 250×125 mas spaxel scale, we can apply this same analysis to other spaxels which are also fully illuminated, to gain a more thorough understanding of the optimal telluric correction methods. We consider the six spaxels within the blue dashed box in Figure 4.4, which we number 1–6 from top left to bottom right. These are the adjacent spaxels to the one we have primarily investigated, #2, which contain similar levels of flux such that we are confident they are also fully-illuminated. We then repeat our analysis individually on each of these spaxels, extracting the spectra, applying `molecfit`, and then determining the RSS.

In Table 4.4 we present the `molecfit` results for the best-fitting Gaussian FWHM for each spaxel, along with the relative gas column densities for H₂O, CO₂, and CH₄. We also give the weighted average, μ , and uncertainty, σ for each parameter, given by

$$\mu = \frac{\sum (x_i w_i)}{\sum w_i} \quad (4.3.2)$$

and

$$\sigma = \frac{1}{\sum w_i} \quad (4.3.3)$$

where the weights w are given by the inverse square of the uncertainties (Bevington and Robinson, 2003). We see from this that the FWHM of the LSF varies between spaxels, which is to be expected due to light following different paths through the instrument. Spaxel #2, in which the majority of the light from CD-23 13701 falls, corresponds to the broadest LSF that we find, which explains some of the difference in LSF between that and our models.

Spaxel (pixels)	FWHM	Relative Gas Column Densities		
		H ₂ O	CO ₂	CH ₄
1	2.999±0.041	0.824±0.013	1.221±0.025	0.885±0.035
2	3.088±0.042	0.828±0.013	1.234±0.026	0.868±0.036
3	3.072±0.043	0.815±0.014	1.276±0.027	0.859±0.036
4	3.029±0.046	0.820±0.014	1.246±0.028	0.823±0.039
5	3.019±0.047	0.803±0.015	1.267±0.028	0.806±0.038
6	2.967±0.045	0.814±0.014	1.240±0.028	0.824±0.038
Average	3.030±0.018	0.818±0.006	1.246±0.011	0.846±0.015

Table 4.4: **FWHM and Gas Column Densities for Varying Spaxels of Callisto H250 Data.** FWHM values and relative gas column densities for 6 spaxels, as returned by `molecfits`. Spaxel #2 is the one for which we have focussed on, and corresponds to the central spaxel of the CD-23 13701 data.

When we consider the weighted average of these six spaxels we have an LSF FWHM of 3.030 ± 0.018 pixels, corresponding to a spectral resolution of $R_{\text{Cal}} = 2877 \pm 17$. We do also see that there are small variations within the recovered gas column densities depending on which spaxel the spectrum is recovered from. The recovered gas column densities can have a dependence on the LSF, as well as the SNR of the provided spectrum. The SNR can be improved by increasing the aperture over which the flux extraction is conducted, however this will lead to adding spectra with different LSFs.

Despite these small variations, we find that the weighted averaged gas column densities of 0.818 ± 0.006 , 1.246 ± 0.011 , and 0.846 ± 0.015 for H₂O, CO₂, and CH₄ respectively support our previous conclusion that there are small differences in the atmospheric profile between the CD-23 13701 and Callisto observations. With H₂O and CH₄, the levels are

Spaxel	RSS Using:			
	CD-23 13701	CD-23 13701 + Molecfit	CD-23 13701 + Molecfit + LSF Conv.	Callisto + Molecfit
1	0.45	0.36	0.30	0.30
2	0.43	0.35	0.30	0.29
3	0.43	0.36	0.31	0.31
4	0.43	0.33	0.30	0.29
5	0.43	0.33	0.30	0.29
6	0.42	0.32	0.30	0.28
Average	0.43	0.34	0.30	0.29

Table 4.5: Residual Sum of Squares of Different Telluric Correction Methods for Multiple Spaxels of Callisto H250 Data. Residual sum of squares from dividing a solar irradiance spectrum out of the recovered Callisto spectra from six fully-illuminated spaxels when using four different telluric correction methods. These are: direct telluric standard division, using `molecfit` on the telluric standard star, convolving that result to the expected LSF, and using `molecfit` directly to the science data.

within the statistical uncertainty from `molecfit` between the two observations, however there is a difference in CO_2 column density, corresponding primarily to the absorption features around $1.6\ \mu\text{m}$.

We then repeated the determination of the RSS for the telluric correction methods described previously, with the results shown in Table 4.5². For the LSF convolution, we use the appropriate FWHM value from Table 4.4 for the relevant spaxel. We also give the average RSS across the 6 spaxels for each method.

From this we can see that there is some variation between the RSS depending on which spaxel we are considering. As there will be slight variations in the noise of the spectra, and potentially intrinsic features to Callisto, this is what we would expect. What we do confirm from this is that for each spaxel we see the same pattern as previously, where the worst correction comes from using just CD-23 13701, with an average RSS of 0.43. Including `molecfit` reduces this to 0.34, and when we further convolve the spectra to the appropriate LSF this becomes 0.30. There are some spaxels for which this

²Whilst spaxel #2 is the same for which we calculate the values in Table 4.2, we recover slightly different values here. This is a result of some additional wavelength channels being masked out due to high noise levels in other spaxels, thus for consistency in the RSS calculations we mask these from all spectra.

recovers the same level of telluric correction as applying `molecfits` directly to Callisto, although on average this is slightly better with an RSS of 0.29.

This has confirmed that knowledge of the LSF is important for optimal telluric correction, and has additionally shown that the LSF can vary between spaxels of a given instrument configuration. Provided that the LSF is known, this shows that we can still use a telluric standard star in conjunction with `molecfits` for correction to the same level as direct application of `molecfits` to the science data. Other issues related to using a telluric standard star remain however, such as differences in gas column densities between observations, or variations in the airmass.

4.4 Conclusions

We have investigated the impact of telluric correction when the standard star used under-illuminates the entrance slit of the spectrograph. We created model LSFs, convolving the impact of the entrance slit, spectrograph, grating and output slit together to create an overall line profile. With this we adjusted the input slit from a two-pixel boxcar to a 0.48 FWHM Gaussian function, and determined that we would expect to see a difference in line width of 14% as a result. We also determined from this that in both instances the data were Nyquist sampled, however with a narrower PSF or grating function there could be issues where the under-illuminated slit also causes under-sampling of the spectral lines.

We then reduced data taken with SINFONI, with the observations set up such that we should see the impact of this effect. We confirmed that the telluric standard star, CD-23 13701, was under-filling the input slit, and extracted spectra from the H-band 250×125 mas observations for both this and our extended object, Callisto.

With these spectra, we used `molecfits` to determine the molecular abundances in the atmosphere, as well as to provide a best fit to the LSF. From this we found that the LSF width was slightly broader than in our models, however we still saw a difference between the fully illuminated and under illuminated LSFs. This was equal to $14.0 \pm 2.1\%$,

in agreement with the amount expected from our models.

We then adopted four different methods for telluric correction on the Callisto spectrum. We divided out the telluric features using CD-23 13701 directly, using `molecfit` directly to the Callisto spectrum, using `molecfit` on CD-23 13701, and finally using `molecfit` on CD-23 13701 and adapting the LSF to align with that of the fully illuminated slit. In each case we additionally removed the same model solar spectrum and determined the RSS for each of the telluric correction methods. These were 0.44, 0.37, 0.31 and 0.30 respectively for the four different methods.

We found that we recovered approximately the same level of correction using the telluric standard star, `molecfit`, and knowledge of the LSF, as we did with applying `molecfit` directly to Callisto. The slight difference that we saw is likely a result of small observational variations between the two, such as the relative gas column densities or airmass. We found that when we did not account for the difference in LSF, the telluric correction was inferior as a result.

We therefore conclude that using a telluric standard which does not fully illuminate the input slit will result in imperfect telluric correction. In this situation, using `molecfit` applied to the science data directly can account for this, assuming that a reliable fit can be obtained. In situations where this is not possible, then we found we could almost achieve the same result by using `molecfit` on the telluric standard star, and combining it with knowledge of the LSF to correct the telluric features.

This shows the importance of having a good understanding of the LSFs of HARMONI, since there will also be situations in which this effect can occur. We showed in Section 4.3.4 that the SINFONI LSF can vary between spaxels for a given instrument configuration, and HARMONI will offer more choices for observations. The more accurately the LSF is known, the greater versatility there can be in the telluric correction, therefore fully characterising the LSFs of HARMONI will assist in obtaining the best science possible with the instrument.

Chapter 5

Recovering High Redshift Gas

Kinematics with **HSIM**

One of the science goals of HARMONI is to investigate spatially resolved kinematics of high redshift galaxies by using the $H\alpha$ line up to $z \sim 2.5$, after which it falls out of the observable range of the instrument. To test the capabilities of HARMONI, we used **HSIM** to undertake a mock observation of such a case using a simulated galaxy. In this chapter we will start by introducing the simulated galaxy that we are using, and then describe the analysis that we carry out.

We will show our mock observation process and the analysis pipeline created to recover kinematics and rotation curves from our data. We then present the results and discuss the consequences of varying parameters such as exposure time and spectral resolution on our mock observations. Finally, we end with an investigation into the impact of potential variation in the point spread function (PSF) on recovered kinematics before summarising our findings with the conclusions.

5.1 High Redshift Galaxy Simulation

5.1.1 Simulation Methodologies

As we discussed in Section 1.4, there is currently a large amount of research into galaxy formation and evolution. One way in which this is done is through simulation, comparing recovered parameters with those seen in observations, often accounting for any observational biases that arise (e.g. Kaviraj et al., 2017; Guidi et al., 2018). This has both the advantage of improving our understanding of the Universe through cosmological models, as well as providing observational astronomers with realistic expectations for the quality of data that can be obtained.

In order to make a mock observation however, an input cube is required of sufficient resolution such that it can then be ‘observed’, with the only biases being added through that process. There are three primary ways for going about this:

- (a) **Redshifted local objects.** This involves taking an object observed at high spatial and spectral resolution in the local Universe and artificially placing it at the desired redshift (e.g. García-Lorenzo et al., 2019). This has the advantage of being a real galaxy, and with more than sufficient signal-to-noise ratio (SNR) to provide an input for a high redshift mock observation. However, the primary drawback of this method is that it assumes galaxies at high redshift have the same properties as local Universe galaxies, which is both known to be incorrect, and also doesn’t help in understanding the properties of high redshift galaxies. In addition, unless carefully accounted for, observational biases from the data can be propagated through, and, in effect, be seen twice when creating a mock observation.
- (b) **Ad-hoc galaxies.** This method involves using available high redshift data, primarily from gravitationally lensed objects (such as those in Jones et al., 2010; Livermore et al., 2012; Livermore et al., 2015), to create mock galaxies that can then be ‘observed’ (e.g. Zieleniewski et al., 2015). The benefit of this method is that the properties of the galaxies will match what is seen at high redshift, however since the data from which they are created is incomplete, there are likely to be inaccuracies

in the overall galaxy. They may not be dynamically stable, or representative of all galaxies at that redshift.

- (c) **Cosmological simulations.** With this method, a cosmological simulation is used to create the galaxy for mock observation (e.g. Kendrew et al., 2016). These use primordial density fluctuations from the cosmic microwave background (CMB), propagated forward in time to create individual galaxies with high spatial resolution that represent the physics input into the cosmological simulation. The advantage of this method is that the physics is well understood, leading to reliable galaxies that can be used for observation. The downside is that the direct outputs from cosmological simulations require some post-processing to create a galaxy that is suitable for use, and in addition, if there are any inaccuracies in the input physics, these will be present in the galaxy. However, this can be of use when comparing cosmologically simulated galaxies with real observations, as differences enhance our understanding of the constraints on the parameters used in creating the simulation. This is the method used for creating the galaxy that we then ‘observe’ with `HSIM`.

5.1.2 NUTFB Galaxy

For this work, we selected the largest galaxy from the NUTFB simulation in the NUT simulation suite (Powell et al., 2011). NUT uses the RAMSES (Teyssier, 2002) adaptive mesh refinement (AMR) code to create cosmological zoom simulations. The lowest redshift at which this galaxy is simulated at high redshift was $z = 3$, due to computational constraints of simulating the evolution to a lower redshift.

The NUTFB simulation uses a star formation rate (SFR; dM_*/dt) that is dependent only on the gas density, and has an average value of (from Kay et al., 2002; Rasera and Teyssier, 2006 and references within)

$$\frac{dM_*}{dt} = \epsilon \frac{V \rho_{\text{gas}}}{t_{\text{ff}}} = 57 \left(\frac{\epsilon}{0.01} \right) \left(\frac{V}{\text{kpc}^3} \right) \left(\frac{n_{\text{H}}}{400 \text{ cm}^{-3}} \right)^{3/2} \left(\frac{X}{0.76} \right)^{-1} \text{M}_{\odot} \text{yr}^{-1}. \quad (5.1.1)$$

This is based on the empirical SFR from Schmidt (1959), where the SFR is proportional to the density of gas, ρ_{gas} , and the characteristic time over which a body collapses under its own gravitational attraction, known as the free-fall time, t_{ff} , using a specific efficiency ϵ . As the freefall time is proportional to the gas density through (e.g. Carraro et al., 1998)

$$t_{\text{ff}} \propto \rho_{\text{gas}}^{-1/2} \quad (5.1.2)$$

the overall SFR is proportional to $\rho_{\text{gas}}^{3/2}$. This gas density is given in terms of the number density of hydrogen, n_{H} , and the mass fraction in hydrogen, X . The total gas mass, M_{gas} , within an individual AMR cell of volume V is equal to $V \times \rho_{\text{gas}}$, where X is set to 0.76 and only gas with $n_{\text{H}} > 400 \text{ cm}^{-3}$ forms stars over the freefall time, with the efficiency $\epsilon = 0.008$. The total stellar mass created, M_* , is equal to $dM_*/dt \times \Delta t$, where Δt is the simulation time step.

The stellar feedback prescription used in NUTFB uses a model that only includes energy from type-II supernovae (following Dubois and Teyssier, 2008). This was primarily due to computational simplicity, as type-II supernovae are a result of an isolated high-mass star running out of fuel. After a star particle reaches an age of 10 Myr, $\eta_{\text{SN}} \times 10^{50} \text{ erg M}_{\odot}^{-1}$ of specific internal energy are injected into the simulation. As NUTFB uses a Salpeter initial mass function (IMF; Salpeter, 1955), $\eta_{\text{SN}} = 0.106$. More recent cosmological simulations have included a more rigorous feedback method which also includes energy from Type-Ia supernovae, such as Kessler et al. (2019), however these were created after the analysis in this chapter was under way.

The parameters of the resulting galaxy are summarised in Table 5.1. The purpose of this galaxy was to provide a high resolution input for observation with HSIM, rather than to create a perfectly representative galaxy. In addition to this, we then artificially took this galaxy and placed it at redshift 1.44, at which we would expect the processes involved to be different. Therefore our focus was not on the recovered parameters themselves, but rather their comparison with the input parameters, to both examine the ability of HAR-

HSIM Parameter	Value	Description
Name	NUTFB	Simulation name
Δx	11.9 pc	Simulation resolution
r_{200}	34 kpc	Halo virial radius
$M_{\text{DM},200}$	$1.3 \times 10^{11} M_{\odot}$	Halo dark matter mass
$r_{\text{c-h}}$	1.7 kpc	Galaxy cold-hot gas transition radius
r_{e}	0.37 kpc	Galaxy stellar half-mass radius
M_{*}	$1.28 \times 10^{10} M_{\odot}$	Galaxy stellar mass within $r_{\text{c-h}}$
M_{gas}	$1.8 \times 10^9 M_{\odot}$	Galaxy gas mass within $r_{\text{c-h}}$
M_{total}	$2.06 \times 10^{10} M_{\odot}$	Galaxy total mass including DM within $r_{\text{c-h}}$
dM_{*}/dt (inst)	$3.27 M_{\odot} \text{ yr}^{-1}$	Galaxy instantaneous star formation rate
dM_{*}/dt (avg)	$2.53 M_{\odot} \text{ yr}^{-1}$	Galaxy star formation rate over 50 Myr
θ_i	51°	Galaxy angular momentum vector inclination to the line of sight

Table 5.1: **Simulated Galaxy Properties.** Properties of the NUTFB galaxy chosen for mock observation with HSIM (reproduced from Richardson et al., 2020).

MONI to observe such an object, as well as to investigate observational biases that are introduced.

5.1.3 Creating a HSIM Input Cube

From this cosmological simulation, we took a 14 kpc width cube, where the three dimensions are the spatial (x, y, z) dimensions. In order to convert this into a HSIM input cube, we needed to add spectral data in, and collapse the simulated cube along a chosen line of sight. We initially created an emission datacube, with the fiducial cube having a resolution of 24 pc, twice that of the native RAMSES simulation. We also created emission cubes of half and double this resolution for comparison. Each native simulation gas cell was matched to a respective cell in the emission datacube using the physical (x, y, z) position.

The total $\text{H}\alpha$ emission in a cell was determined based on the SFR by following Kennicutt et al. (1994), where ionizing photons from young, massive stars are reprocessed by the neutral medium and a fraction are then re-emitted as $\text{H}\alpha$. This relationship is given by

$$L_{\text{H}\alpha} = 1.26 \times 10^{41} \left(\frac{dM_{*}}{dt} M_{\odot} \text{ yr}^{-1} \right) \text{ erg s}^{-1}, \quad (5.1.3)$$

where the SFR, dM_*/dt , is given by Equation 5.1.1. Each cell in the emission datacube was given the appropriate $H\alpha$ luminosity, along with the average metallicity of the relevant native simulation cells that contribute to it. The averaged velocity and local gas turbulence were also recovered and the information saved for use in creating the `HSIM` input cube.

The gas turbulence, σ_{gas} , was estimated from the sum of the squares of the velocity gradient matrix, $\mathbf{v}$, through

$$\sigma_{\text{gas}}^2 = \frac{1}{3} \sum_{a=1}^n \sum_{b=1}^n v_{ab}^2 \equiv \frac{1}{3} \text{tr}(\mathbf{v}^T \mathbf{v}), \quad (5.1.4)$$

where $\mathbf{v}^T$ is the transpose of the velocity gradient matrix, a and b denote the components of the matrix and the factor of 3 accounts for only turbulence along the line of sight. Each of the components in the matrix were estimated from the differences in the velocity at the midpoint of adjacent cells in each of the three dimensions such that for example

$$v_{xx}^2 = \left(\frac{\delta v_x}{\delta x} \right)^2 \approx \left(\frac{v_{x,i+1/2} - v_{x,i-1/2}}{\Delta x} \right)^2 \quad (5.1.5)$$

with

$$v_{x,i+1/2} = \frac{1}{2}(v_{x,i} + v_{x,i+1}), \quad v_{x,i-1/2} = \frac{1}{2}(v_{x,i} + v_{x,i-1}). \quad (5.1.6)$$

Here $v_{x,i}$ corresponds to the x velocity within a given cell i , with adjacent cells denoted by $i+1$ and $i-1$, and Δx is the simulation resolution along the x axis. This is repeated for each v_x , v_y , and v_z , with $\Delta x = \Delta y = \Delta z$, giving a 3×3 matrix for each cell, which are then summed and divided by 3 to give the square of the turbulence along the line of sight for that cell. This is then repeated for all cells to give the local gas turbulence at every location in the datacube.

Then in order to remove a spatial dimension from the datacube, we integrated through it along the y -axis of the simulation, giving a two dimensional array in which the galaxy was inclined at 51° to the line of sight. Doing this allowed us to account for extinction

such that each cell along the line of sight is only extinguished by dust in front of it, rather than behind it. Assuming the extinction scales with the density of gas-phase metals along the line-of-sight column prior to the cell being extinguished, Σ_Z we can define the extinction as

$$A_V = 1.086 \left(\frac{3}{4\rho_d} \right) f_d \Sigma_Z \frac{Q_0}{\lambda_V} \quad (5.1.7)$$

assuming an extinction coefficient $Q_V = Q_0 a / \lambda_V \simeq Q_0 a / \lambda_{H\alpha}$ (Carroll and Ostlie, 2017, Section 12.1) and the dust grains of size a follow $dn/da \sim a^{-3.5}$ (Mathis et al., 1977). We also take a single dust grain to have density $\rho_d \simeq 3 \text{ g cm}^{-3}$ and a dust fraction f_d of 8% (a little lower than Peeples et al., 2014), and $A_V \lesssim 1$ (e.g. Calzetti et al., 2000; Sobral et al., 2012). A given cell also has self-absorption, making the assumption that the $H\alpha$ emission and the dust mass are uniformly distributed. We show the total extinction in the galaxy along different lines-of-sight in Figure 5.1 (reproduced from Richardson et al., 2020).

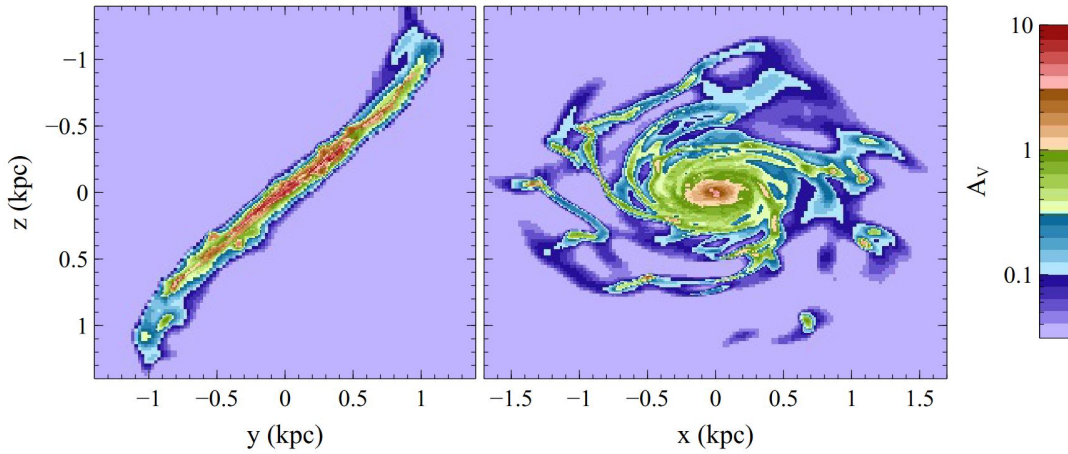


Figure 5.1: Projected Extinction Along Multiple Lines-of-Sight of the Input Galaxy. *Left:* Projected extinction along the x axis of the NUTFB galaxy. *Right:* Same, for the y axis. Both are shown on the same logarithmic scale in A_V , with each hue representing 0.5 dex, and each brightness step being 0.1 dex. *Image reproduced from Richardson et al. (2020).*

The emission cube was initially created at a frame of rest, with velocity resolution of 10 km s^{-1} . Each emission cell then contributed a Gaussian emission line to the input cube

for `HSIM`, where the width of the line was determined by

$$\sigma = \sqrt{\sigma_{\text{gas}}^2/3 + \sigma_{\text{th}}^2} \quad (5.1.8)$$

where σ_{gas} is the local gas turbulence and σ_{th} is the thermal width given by

$$\sigma_{\text{th}} = \sqrt{2P/\rho}, \quad (5.1.9)$$

with P as the mean gas pressure and ρ is the gas density. Typically the turbulence is much greater than the thermal width for this galaxy. In order to convert from velocity, v , into wavelength, the Doppler effect equation

$$\frac{\lambda_o}{\lambda_s} = \sqrt{\frac{1 + \beta}{1 - \beta}} \quad (5.1.10)$$

was used, with λ_o the observed wavelength, λ_s the source wavelength and $\beta = v/c$ with c being the speed of light. This was inverted to solve for β as a function of uniformly spaced wavelength, and then the value of β interpolated from the velocity cube.

Now this has given the spectral axis to the data, we then artificially adjusted the $z = 3$ galaxy such that observationally it would be at $z = 1.44$. At the same time, the distance units were converted from physical parsecs into angular milli arc seconds, assuming a flat cosmology and using $H_0 = 70.4 \text{ km s}^{-1}$ and $\Omega_M = 0.272$ (Komatsu et al., 2011). This was done primarily to bring the $\text{H}\alpha$ line into the wavelength range covered by `HARMONI` and at $z = 1.44$ the rest frame wavelength falls in a gap between sky emission lines. The luminosity was converted into flux through

$$F_{\text{H}\alpha} = \frac{L_{\text{H}\alpha, z=1.44}}{4\pi D_l^2} \quad (5.1.11)$$

where D_l is the luminosity distance to the galaxy at $z = 1.44$, and the luminosity has been scaled by a factor of $1/(1+z)^2$ to account for the reduction in photon energy due to

redshift, and the effect of them arriving more slowly due to cosmological time dilation.

As mentioned previously, this does result in a galaxy in which the processes involved are not entirely physically accurate for the apparent redshift. In addition, the `NUTFB` simulation has inefficient feedback, creating a large central bulge and small effective radius, r_e , compared to typical galaxies at this redshift (Price et al., 2016). Furthermore, this leads to a fast rotation speed, peaking at 480 km s^{-1} at 60 pc, a little faster than would be expected of a galaxy at this redshift (Übler et al., 2017). However, our primary focus of this investigation was to assess the ability of `HARMONI` to observe a galaxy at this redshift, and look at observational biases introduced in this process, and the galaxy that we have created here is suitable for this purpose.

The primary input cube which we used for `HSIM` has a resolution of 24 pc, corresponding to a resolution of 2.78 mas, equal to half the resolution of the simulation, as this allowed for faster mock observations. We also generated an input cube at the finer, intrinsic simulation resolution of 1.39 mas (denoted by $\Delta 1$) and coarser resolution of 5.56 mas ($\Delta 6$) to ensure we were not introducing bias from this, and this is presented in Section 5.3.1. The pipeline used to convert from the `RAMSES` simulation to a `HSIM` input cube is available through the `RAMSES2HSIM` pipeline¹ (Richardson et al., 2020).

5.2 Analysis Methods

5.2.1 `HSIM` Observations

The simulated input cubes were then mock observed with `HSIM` to produce realistic exposures such as would be expected to be seen with `HARMONI`. The common parameters used are detailed in Table 5.2, where we primarily adjusted the value of `NINT`, the number of repeated exposures, to understand the impact of exposure time on our results. Our observed exposures had integration times of $2\times$, $4\times$, $20\times$ and $60\times$ 900 seconds, for total integration times of 0.5, 1, 5 and 15 hours respectively. We also produced one observa-

¹Available at <https://github.com/mlarichardson/ramses2hsim>.

5.2. ANALYSIS METHODS

HSIM Parameter	Value	Description
ver	300	Version of HSIM
DIT	900	Integration time per exposure (seconds)
NINT ^a	20	Number of exposures
spax	10x10	Spaxel size (mas)
Grating ^b	H+K	Chosen grating, equal to $R \sim 3500$ ($\Delta\lambda \sim 4.6\text{\AA}$)
Seeing	0.43	Simulated seeing conditions
Air Mass	1.1	Simulated air mass
Temp	280.5	Temperature of telescope (K)
Moon	0	Moon illumination
jitter	0	Residual telescope jitter
ADR	False	Atmospheric Differential Refraction correction
Dets	False	Advanced detector systematics off
AO	LTAO	Use Laser Tomographic Adaptive Optics
Seed	1	Random noise seed

Table 5.2: **HSIM Simulation Parameters for NUTFB Galaxy.** Default parameters used in HSIM for our simulated observations. ^aWe also simulated exposures with longer and shorter times as part of our investigation. ^bWe also simulated an exposure with the higher resolution H grating.

tion using the higher resolution H grating, corresponding to $R \sim 7000$ ($\Delta\lambda \sim 2.3\text{\AA}$) to investigate if spectral resolution affected the recovered kinematics.

The LTAO PSF used in the simulations is created from a power spectral density (PSD) created for use with HSIM based off of the 6 LGS asterism design for HARMONI (see Neichel et al., 2008) and has seeing conditions of 0.43 at a zenith angle of 25° . This can generate a unique PSF appropriate for each wavelength within the simulation. Here we assume an optimal AO PSF, with no residual jitter, to reduce the number of variables in our simulations. This corresponds to an instrumental AO PSF of ~ 11 mas FWHM over the wavelength range simulated.

As this work was completed simultaneously with the work on developing HSIM, the advanced detector systematics were not fully implemented at the time of creating the simulated exposures, therefore our simulations do not use them.

5.2.2 Velocity Maps

In order to recover the kinematics of the galaxy from the resultant data, we needed to create velocity maps from the three dimensional data cubes. Prior to making these, the data were masked on a per-spaxel basis with an integrated SNR over the spectral range of the emission line, with a threshold of 7. This had two functions; firstly it increased the reliability of the data by removing noise, such as that from skyline residuals, to increase the confidence in our fit. Secondly, it also removed spaxels which contain no, or very little, light from the galaxy. Removing them at this point increased the efficiency of the fitting pipeline, as it did not have to attempt to fit data for which the $H\alpha$ line is non-existent. In spaxels where there was an SNR below 7, a second attempt was made by binning the surrounding pixels into a 2×2 box, to increase SNR at the expense of spatial resolution (akin to the process in Zieleniewski et al., 2015).

We also considered two alternative methods for maximising the amount of useful data we could extract from low-SNR regimes, namely evaluating the first moment and Voronoi binning. In the first case, the first moment along the spectral axis was calculated, excluding the regions liable to contain skyline residuals. In this instance, the recovery of the $H\alpha$ line (and subsequent velocity maps) were often skewed by residual noise, particularly when the SNR was low. We therefore discounted this method for recovering our velocity maps. In the case of Voronoi binning, we used the `VorBin` method described in Cappellari and Copin (2003) and created bins with a cut-off of 7, as with the initial threshold on the data. Here we didn't see a significant improvement in our velocity maps for the short exposures due to the clumpy nature of the gas, and no difference in long exposures with high SNR per spaxel as would be expected. As the benefits were very marginal, particularly with the increased computational time of creating the Voronoi bins, we decided not to pursue this method further.

Once we had masked the data to isolate only spaxels with adequate SNR for analysis, we converted the wavelengths into velocities, prior to attempting the fit to the $H\alpha$ line. We

brought the observed wavelengths, λ_o , back to their rest-frame wavelengths, λ_r , through

$$\lambda_r = \frac{\lambda_o}{1+z}, \quad (5.2.1)$$

where z is the cosmological redshift, which we have simulated to be 1.44. We then used the relation

$$\frac{\Delta v}{c} = \Delta \ln \lambda \quad (5.2.2)$$

and interpolated our data to create a linearly spaced velocity axis, knowing that the rest frame wavelength of H α is 656.46 nm. We could then fit each spaxel, or grouped 2×2 spaxels, to obtain the velocity corresponding the centre of the H α line for. In order to do this a Gaussian fit was applied, incorporating the entire range contained within the datacube, 1.595 μm to 1.606 μm prior to converting into velocity. There were 4 free parameters in this fit: central velocity, amplitude, width, and constant background. To obtain the initial guesses for these parameters a first order Savitzky-Golay filter was applied to the data with a five velocity channel window, and then the maximum amplitude of the smoothed data was found, along with the corresponding velocity. These were then used as initial guesses for the amplitude and central velocity for the Gaussian fit. A non-linear least squares fitting method from `SciPy` was used, with uncertainties given by the corresponding standard deviation cube returned from `HSIM`.

Much of this uncertainty comes from sky line residuals, which are a result of a difference in Poisson noise on the sky lines between the simulated observation cube, and the simulated sky cube. The reduced cube, which we used in this analysis, contains a ‘perfect’ sky subtraction in that the sky flux is the same between the two, however differences in the Poisson noise lead to skyline residuals such as those seen in Figure 5.2. From this fit, we obtained the line of sight velocity per spaxel or 2×2 spaxel box, as well as the dispersion from the width of the Gaussian function fitted. This allowed us to produce velocity maps such as those shown in Figure 5.2 for each of our simulated observations.

After this fitting algorithm was run, the following checks on the returned parameters

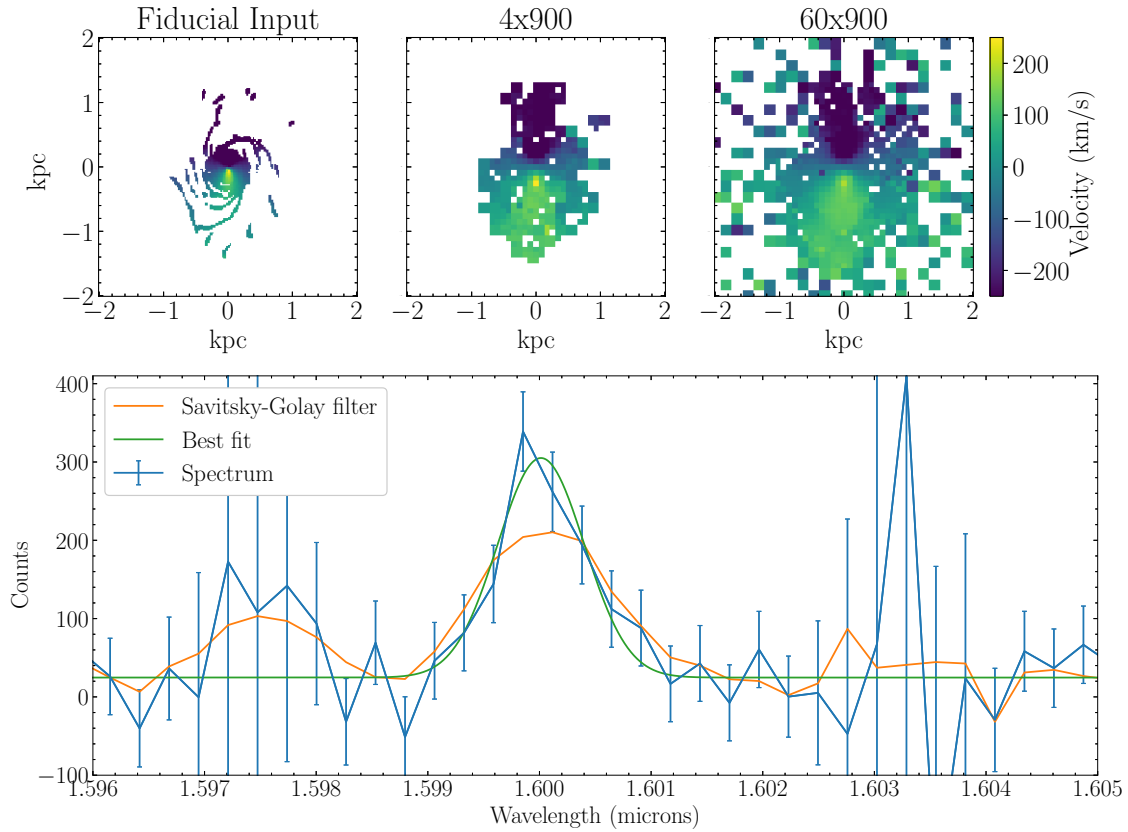


Figure 5.2: 2D Velocity Maps and Sample 1D Spectrum Showing Least-Squares Fit. *Top:* 2D velocity maps resulting from our $H\alpha$ fitting algorithm for the fiducial input cube (*left*), 4×900 cube (*middle*) and 60×900 cube (*right*). White areas in the input cube have no emission. *Bottom:* 1D spectrum showing a sample of our fitting algorithm with the reduced output spectrum (blue), Savitsky-Golay filter for peak-finding (orange) and best-fit Gaussian (green). Skyline residual noise around $1.603\ \mu\text{m}$ is present, but has been rejected by our algorithm.

were implemented to ensure the data used in further analysis were as reliable as possible, and to remove false positive fits, such as those to noise peaks. Data with extremely high dispersion (greater than 600 km s^{-1}) and those with high residual background (greater than an absolute value of 50 counts) were rejected. We also removed data with an uncertainty on the magnitude of the peak greater than the peak value (implying the peak value was consistent with zero), and data where the uncertainty on the dispersion was greater than 0.8 times the value of the dispersion. The final check removed data where the maximum amplitude was less than twice that of the background, as such a line was considered too weak to be reliable. These values were chosen to retain the maximum amount of reliable data, whilst ensuring very few false positive fits were included.

5.2.3 Pseudo Long Slit

As part of our preliminary analysis, we constructed a pseudo long slit placed along the major axis of the galaxy, simulating the effect of a traditional long-slit spectrograph. In order to do this, we initially de-rotated the galaxy such that we could place a slit across it, using an inbuilt function within the `SciPy` package to conduct spline interpolation. This was primarily to align the galaxy with a Cartesian grid to make the analysis easier to perform. Alternatively, we could have placed the pseudo-slit at an angle across the data, and instead perform the interpolation for the regions of the galaxy which are covered by the slit. We then co-added the flux from spaxels across the slit up to the number required for the chosen width to simulate all the flux being incident in one resolution element, adding the uncertainties in quadrature. The spectra from these summed spaxels were now fitted in the same way as described in Section 5.2.2 to extract the rotation velocity at each point. Here the uncertainties obtained are solely on the least-squares fit using the provided uncertainty on the flux, and do not take into account the dispersion. Once all the velocities were obtained, the negative half of the rotation curve was flipped so that it can be compared to the other extracted rotation curves.

We created two pseudo slits, the first was of the minimum size of 10 mas, equal to

the spatial resolution of the simulation. Our second slit was of width 110 mas, roughly equivalent to 1 kpc whilst ensuring a symmetric slit around the centre, which encompasses the majority of the flux from the galaxy. We present the results in Section 5.3.4 as an interesting comparison between long-slit spectroscopy and IFU spectroscopy, however do not draw any conclusions from this in regards to the performance of HARMONI.

5.2.4 DiskFit

Since HARMONI will be an integral-field spectrograph, it makes more sense to incorporate the full 2D data from the velocity maps, and to do this we used the software package `DiskFit` (Spekkens and Sellwood, 2007; Sellwood and Spekkens, 2015). This software requires an input data file containing the 2D velocity profile of the galaxy, along with a number of parameters relating to the galaxy, namely; initial guesses for the centre, inclination and position angles, and the systemic velocity of the galaxy. It then fits circular rotation velocities to specified annuli of a given width and spacing, minimising χ^2 for all annuli. As `DiskFit` does not have information about the emission line flux morphology, it would not be able to produce a reliable deconvolution from the smearing of the instrumental PSF, so we turned off this feature for our analysis.

We converted the 2D velocity maps to the correct format for `DiskFit`, and ran it for our data. Due to the size of some of the datacubes, in particular the input cubes, we constrained the fitting region to approximately the central 2 kpc. This truncation has no impact on our findings, as only in the input cube and very long exposure (and hence high SNR) cubes could we get data that far out, and our primary comparisons are within the first 1 to 1.5 kpc. The uncertainties on the provided velocity values, σ_v , are not immediately trivial to determine. Giving only the statistical uncertainty on the velocity from the Gaussian fit can fail to accurately account for physical variation within the galaxy (e.g. Sellwood et al., 2021). In this work, we take a conservative approach to the uncertainties, where we combine the errors from the Gaussian fit to the emission line, σ_G , and the

dispersion of the line, Δv , such that

$$\sigma_v = \sqrt{\sigma_G^2 + (\Delta v)^2}. \quad (5.2.3)$$

This follows methodology used in de Blok et al. (2008), and incorporates physical variations in the velocity of the gas along the line of sight.

The outputs from `DiskFit` depend on the parameters chosen to fit, which for us were the centre of the galaxy, systemic velocity and rotational velocity at each annulus. We fixed the inclination to the line of sight and the position angle, as we could recover those from the simulation data. We obtained uncertainties on the fit parameters through a bootstrapping procedure inbuilt into `DiskFit`. The uncertainties on the peak velocity and corresponding radius were determined by taking the bin containing the peak rotational velocity, along with the neighbouring bins, and fitting a quadratic function to these points. The uncertainties from the `DiskFit` fits were propagated through into this, to give the overall estimate of the uncertainties on the parameters in Table 5.3.

We found that `DiskFit` is extremely sensitive to the centre co-ordinates, and when we had low-SNR data in the outer parts of the input velocities, this could skew where it identified the centre, and thus skewed the resulting kinematic profile. To counter this we adopted a two-stage procedure, whereby we initially constrained the outer radius to 1 kpc and determined the centre accordingly. We then held this centre fixed in a subsequent run of `DiskFit`, and carried out a second fit where the maximum radius was limited by the SNR of the data, up to a maximum of 2 kpc as mentioned above. We chose this point as it matches the well-ordered rotation of the galaxy in the input cube.

5.3 Results and Discussion

Our primary goal of these simulations and subsequent analysis was to understand how well HARMONI will allow us to recover the characteristics of our simulated galaxy using HSIM. Initially we will present the summary of the comparisons between our input galaxy

and the various `HSIM` simulations. Then we will discuss in more detail the impact of the exposure time, and the spectral resolution, before ending with a brief comparison to the results obtained from the pseudo long slit. In Section 5.4 we investigate the impact of the AO PSF on our simulations.

5.3.1 Comparison with Simulated Galaxy

Figure 5.3 shows the $H\alpha$ intensity in the input cube compared to output cube observed after 20×900 seconds of integration time. This clearly shows the spiral arms with clumps of cold, star-forming gas. We can see that in the input cube we recover a rotational velocity from each point that has $H\alpha$ present, as can be seen by comparing with Figure 5.2. For the output cube, we are naturally at a coarser resolution, and in addition to that we can see the impact of the instrument PSF blurring the $H\alpha$ further. We do, however, still maintain some ability to discern the centre of the galaxy, the spiral arms, and clumps of star-forming regions.

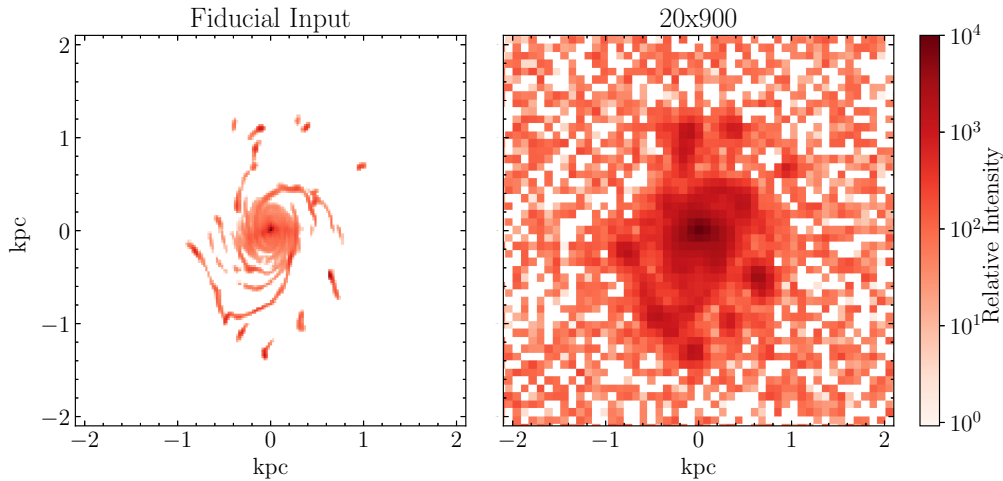


Figure 5.3: **$H\alpha$ Intensity Maps.** *Left:* $H\alpha$ intensity map for the input cube, with a resolution of 2.78 mas. *Right:* Same, for the 20×900 mock observation, at a resolution of 10 mas. The brightness scale is constant for both panels, and is shown logarithmically to enhance the features in the galaxy.

There were three primary metrics we were interested in recovering and comparing with our input simulation. Namely, the peak rotational velocity, the radius at which this velocity occurred, and the expected SFR. These results are tabulated in Table 5.3, where

5.3. RESULTS AND DISCUSSION

we show these values for each of our output cubes, along with the input cube, and the input cube smoothed by a convolution with a Gaussian of both 10 mas and 20 mas FWHM. We used the results from `DiskFit` to recover the peak rotational velocity, accounting for the inclination to the line of sight, θ_i of the galaxy through the equation

$$v_p = \frac{v_D}{\sin(\theta_i)}, \quad (5.3.1)$$

where v_D is the peak rotational velocity returned from `DiskFit` and θ_i is recovered from the simulation and set to 51° . The radius at which this occurs was then obtained and denoted by R_p . The quoted SNR was the average integrated SNR along the spectral axis between $1.5954 \mu\text{m}$ and $1.6060 \mu\text{m}$ over a circular aperture of diameter 60 mas. The SFR was determined by extracting the flux from an aperture given by $r_{c-h}=1.7 \text{ kpc}$ and using the equation

$$F_{H\alpha} = \frac{L_{H\alpha}}{4\pi D_L^2} \quad (5.3.2)$$

to calculate the $H\alpha$ luminosity, $L_{H\alpha}$ where $F_{H\alpha}$ is the $H\alpha$ flux and D_L is the luminosity distance to the galaxy. We could then use Equation 5.1.3 to determine the SFR. We additionally calculated an extinction-corrected SFR, where we assume $A_V=1$, and the ratio of total to selective absorption at V, $R_V=3.1$ (e.g. Fitzpatrick, 1999; Calzetti et al., 2000; Domínguez et al., 2013).

We initially compared the intrinsic rotation profile of the galaxy, determined from the simulation by calculating the velocity components of the gas along the line of sight, with our analysis of the input cubes to `HSIM`. This gives a result of $476 \pm 19 \text{ km s}^{-1}$ at a radius of $87 \pm 19 \text{ pc}$ for the fiducial input cube, with the double resolution $\Delta 1$ cube yielding $479 \pm 20 \text{ km s}^{-1}$ at $71 \pm 9 \text{ pc}$, and the half resolution $\Delta 6$ cube giving $493 \pm 10 \text{ km s}^{-1}$ at $100 \pm 24 \text{ pc}$. These are all very close to the intrinsically determined peak of rotation of 480 km s^{-1} at 60 pc . As may be expected, the coarser the input cube is sampled, the slightly worse we are able to recover the intrinsic characteristics, but as can be seen in Figure 5.4 we recover the rotation curve of the galaxy very well. Between 1-1.5 kpc the gas begins

Intrinsic	$^a\text{R}_\text{p}$		$^b\text{v}_\text{p}$	$^c\text{SFR}_\text{inst}$	$^c\text{SFR}_50$		
Native res	60		480	3.27	2.53		
– 10 mas	100		410	-	-		
– 20 mas	150		350	-	-		
Analysis	Input Cubes			Output Cubes			SNR
	$^a\text{R}_\text{p}$	$^b\text{v}_\text{p}$	^cSFR	$^a\text{R}_\text{p}$	$^b\text{v}_\text{p}$	^cSFR	
Fiducial	87±19	476±19	1.49	269±48	334±24	1.41 (3.00)	30
Δ1	71±9	479 ± 20	0.96	232±129	314±11	0.93 (1.98)	19
Δ6	100±24	493±10	1.81	269±56	319±10	1.78 (3.80)	37
2×900	-	-	-	253±286	327±106	1.36 (2.91)	10
4×900	-	-	-	260±48	324±13	1.55 (3.31)	14
60×900	-	-	-	254±54	318±13	1.46 (3.11)	53
R7000	-	-	-	252±59	325±11	1.47 (3.15)	26
$^a \text{ pc} \quad ^b \text{ km s}^{-1} \quad ^c \text{ M}_\odot \text{ yr}^{-1}$							

Table 5.3: **Recovered Galaxy Kinematics.** *Intrinsic:* Galaxy kinematic properties and SFR as determined from the cosmological simulation. –10 mas and –20 mas denote the intrinsic profile smoothed by those amounts respectively. *Analysis:* Values obtained for the galaxy kinematics as a result of our analysis pipeline. SFR in brackets assumes an extinction of $A_V=1$.

to no longer follow the ordered rotation of the galaxy, and is seen by the deviation from the black and magenta lines. The similarity between the recovered parameters and the intrinsic kinematic profile gives us confidence that our analysis pipeline is reliable.

The right panel in Figure 5.4 shows the rotation curves obtained from the simulated HSIM observations. Our primary result, the recovered rotation profile of the fiducial simulation is shown as green circles. For this, we recover a peak rotation speed of $334 \pm 24 \text{ km s}^{-1}$ at $269 \pm 48 \text{ pc}$, which is both lower in magnitude and further from the centre than the rotational profile of the intrinsic galaxy. This is in line with expectations however, as we expect a broadening primarily due the limitations on resolution from the instrument and AO PSF. This is explored in more detail in Section 5.4.

Also in Figure 5.4 we show the results from our simulated exposures with input cubes of coarser and finer spatial resolution. All of these were simulated with a total exposure time of 5 hours to enable a direct comparison, thus we could ascertain if the resolution of the input cube had an impact on our results. For the double-resolution $\Delta 1$ cube we recover a peak rotation speed of $314 \pm 11 \text{ km s}^{-1}$ at $232 \pm 129 \text{ pc}$, and for the half-resolution $\Delta 6$

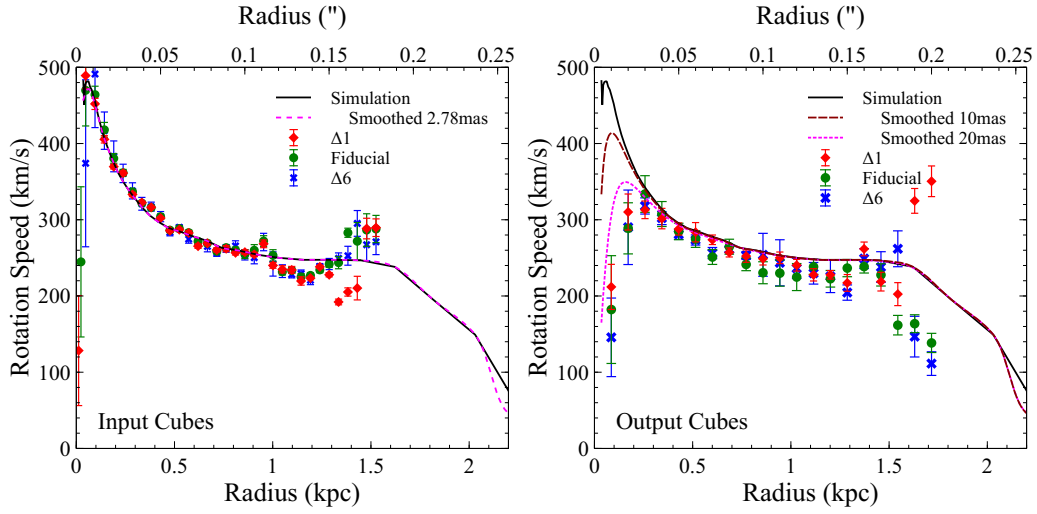


Figure 5.4: Dependence of Simulation Resolution on Extracted Rotation Curve. *Left:* Extracted rotation curves from the input cubes at the fiducial 2.78 mas resolution (green circles), double resolution 1.39 mas ($\Delta 1$, red diamonds) and half resolution 5.56 mas ($\Delta 6$, blue crosses). Also shown is the intrinsic rotation curve extracted from the angular momentum of the galaxy in simulation (black solid line), and this curve smoothed by the resolution of the fiducial cube (magenta dashed line). *Right:* Same, for the output cubes from the HSIM runs. Shown here for comparison is the intrinsic rotation smoothed by the spatial resolution of our exposures at 10 mas (dark red dashed line) and by twice the spatial resolution at 20 mas (magenta dotted line).

cube $319 \pm 10 \text{ km s}^{-1}$ at $269 \pm 56 \text{ pc}$. These values are consistent within 1σ of each other, which suggests that the spatial resolution chosen is not affecting the recovered kinematics.

From these results, we can be happy that neither our analysis method, nor the chosen spatial resolution of the input cube have had a systematic effect on our results. This therefore allowed us to investigate the impact of longer or shorter exposure times (Section 5.3.2), and higher spectral resolution (Section 5.3.3). We could then draw some conclusions about the constraints of HARMONI for recovering kinematics of a galaxy at this redshift.

5.3.2 Exposure Time

As part of our investigation we produced simulated observations at a range of exposure times. This allowed us to check the scaling of the SNR with exposure time, as well as to see how the recovered kinematic parameters varied with these different exposure

times. Alongside the 20×900 s fiducial cube, we produced 3 other cubes of exposure times 2×900 s, 4×900 s and 60×900 s for total integration times of 30 minutes, 1 hour and 15 hours respectively. We would expect the SNR for these cubes to scale by $\sqrt{t}$, where t is the exposure time, and as we can see from Table 5.3 this is the case. For the four cubes in which we only varied the exposure time we recover SNRs of 10, 14, 30 and 53 in increasing order of exposure time, which is very consistent with the expected scaling.

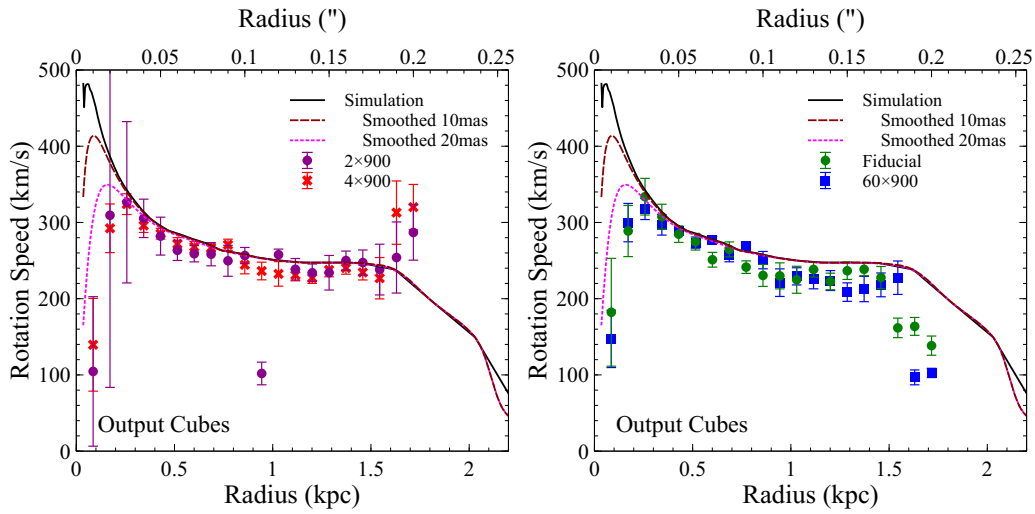


Figure 5.5: **Dependence of Exposure Time on Extracted Rotation Curve.** *Left:* Extracted rotation curves from the shorter simulated exposures of 2×900 (purple circles) and 4×900 (red crosses), simulating 30 minutes and 1 hour respectively of exposure time. *Right:* Same, for the longer simulated exposures of the fiducial cube (20×900 , green circles) and 60×900 (blue squares), simulating 5 hours and 15 hours respectively of exposure time.

In Figure 5.5 we show the full rotation curves for these output cubes. For the shortest simulated observation, the 2×900 cube, we have very large uncertainties on the rotation profile returned from `DiskFit`. This is because at this scale there are very few pixels with high enough SNR that are reliable enough to use `DiskFit`. From these, we can recover peak rotation speeds at the different simulated exposure times. For the shortest exposure, 2×900 , we recover a value of $327 \pm 106 \text{ km s}^{-1}$ at a radius of $253 \pm 286 \text{ pc}$, which would be consistent with the rotation peaking at the centre of the galaxy.

Clearly with only 30 minutes simulated exposure time, we are unable to obtain a re-

liable estimate of the kinematic properties of this galaxy. However, with an increased exposure time of 1 hour in the 4×900 cube, we recover a peak rotation velocity of $324\pm 13 \text{ km s}^{-1}$ at $260\pm 48 \text{ pc}$, which is now consistent with the fiducial cube. For the longest simulated exposure, the 60×900 cube, we recover a peak rotation velocity of $318\pm 13 \text{ km s}^{-1}$ at $254\pm 54 \text{ pc}$, also consistent with the fiducial cube.

As expected, the shortest exposure suffers from insufficient SNR and as a result we are unable to recover reliable estimates for the peak values of rotation, however we are still able to recover the shape of rotation curve reasonably accurately. All of the recovered rotation curves agree within 1σ for all but a couple of outlier points in the region of ordered rotation, up to around 1.5 kpc, after which we see the impact of the more turbulent gas having a more pronounced effect.

Interestingly, we do not see a significant improvement in either the recovered rotation curve, or the recovered kinematic parameters, between 1 hour and 15 hours of simulated exposure time. We can see from Figure 5.5 that the error bars do not increase on the data points, although the SNR does (from Table 5.3). This is likely due to the uncertainties on the velocity after this time being dominated by the dispersion of the gas, rather than inaccuracies in the fit to the $\text{H}\alpha$ line. This suggests that 1 hour of observation time with HARMONI is sufficient to accurately determine the line of sight velocity from the $\text{H}\alpha$ line of a galaxy at redshift 1.44. We are also limited in our accuracy by the resolution and PSF effects, which are discussed further in Section 5.4.

5.3.3 Spectral Resolution

We also wished to investigate the impact of spectral resolution on our recovered kinematic parameters. For the fiducial cube we observed the galaxy using the H+K grating, which has a spectral resolving power of $R \sim 3500$, so we also produced an output cube using the H grating, which has a spectral resolving power of $R \sim 7000$. We kept all the other parameters constant for this R7000 cube and the recovered rotation curve can be seen in Figure 5.6. In regards to the SNR for this cube, we have twice as many data points within

the spectral range included, but half as much flux per spaxel for the same integration time. In the regime of our simulations we are primarily dominated by shot noise processes, therefore we would expect the noise per spaxel to decrease by the square root of the signal, giving us

$$\text{SNR}_{\text{per spaxel}} = \frac{S/2}{N/\sqrt{2}} = \frac{1}{\sqrt{2}} \frac{S}{N} \quad (5.3.3)$$

compared to the signal (S) and noise (N) in the $R \sim 3500$ data. However, since we maintain the same spectral range, the total number of spaxels included is doubled, therefore the spectrally integrated noise would be expected to increase by a factor of $\sqrt{2}$, which should lead us to recovering the same overall SNR. This is confirmed by a recovered SNR of 26, which is similar to that from the fiducial cube, although slightly lower due to non-Poissonian noise processes such as read noise that are increased by a factor of 2 rather than $\sqrt{2}$.

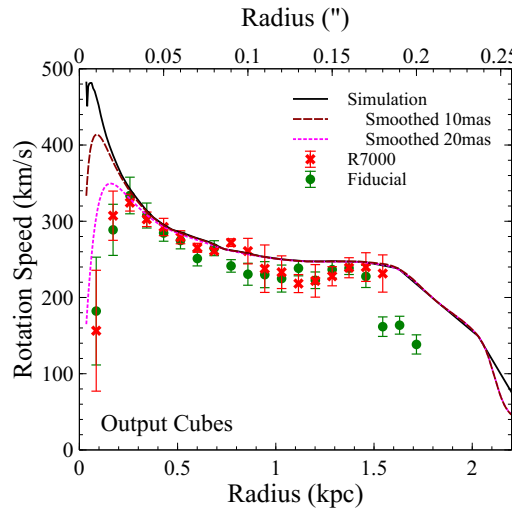


Figure 5.6: **Dependence of Spectral Resolution on Extracted Rotation Curve.** Extracted rotation curves from the fiducial cube with spectral resolution $R \sim 3500$ (green circles) and the R7000 cube with spectral resolution $R \sim 7000$ (red crosses)

We can see from Figure 5.6 that the recovered rotation curves are very similar in shape, with similar uncertainties as well. A recovered peak rotation speed of $325 \pm 11 \text{ km s}^{-1}$ at $252 \pm 59 \text{ pc}$ is also consistent with those recovered from the fiducial cube. This is also expected, as providing the spectral line is resolved there should be no difference in the estimate of the centroid of the line, and thus no difference in the velocity map produced

and provided to `DiskFit`. The only slight variation is that by having more data points sampling the line, we have a smaller uncertainty on the central wavelength, although this is a small contribution to the overall error in comparison to the width of the line due to dispersion in the galaxy.

The main difference we do see between the two curves is that we recover a slightly higher rotational velocity around 0.85 kpc in the R7000 cube. This could potentially be due to the wings of the spatial PSF having a smaller impact on the recovered line of sight velocity in a given spaxel. With wide wings, flux from other radii is convolved with that from a spaxel of interest, potentially biasing the recovered velocity, however a well resolved line may be less susceptible to this effect. This could be investigated further through additional simulations that are better set up to see this effect.

From this, we can ascertain that there is no significant advantage of using a higher spectral resolution grating for a galaxy such as this where the spectral line is already sufficiently resolved at the lower resolution. Despite the line being twice as well sampled, the recovered rotation curve and kinematic parameters were largely unaffected.

5.3.4 Pseudo Long Slits

We also looked at rotation curves extracted from pseudo long slits placed across the galaxy, to illustrate the benefits of integral field spectroscopy. We used the 20×900 s fiducial cube and constructed two pseudo slits, one of the minimum width of 10 mas (~ 100 pc) and one of width 110 mas (~ 1 kpc). We would expect that the broader the slit, the less effectively we recover the rotation profile, and this is seen in Figure 5.7.

The 10 mas slit does a reasonable job at recovering the rotation profile, in comparison with the fiducial cube, which uses the full two dimensions of spatial data. We recover a peak velocity of $329 \pm 16 \text{ km s}^{-1}$ at $257 \pm 43 \text{ pc}$ for the positive curve, and $314 \pm 4 \text{ km s}^{-1}$ at $257 \pm 43 \text{ pc}$ for the negative curve. This is consistent with the values recovered for the fiducial cube with `DiskFit`, and this is to be expected, as in the central region we have sufficient SNR within one spaxel to recover a rotational velocity, and with the narrowest

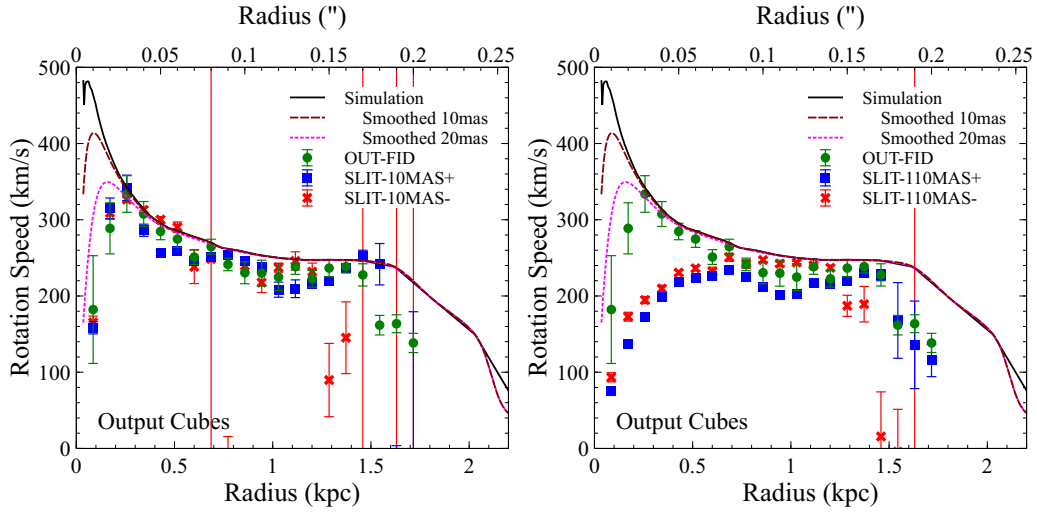


Figure 5.7: Comparison Between a Pseudo Long Slit and 2D Velocity Extraction. *Left:* Extracted rotation curve from a 10 mas pseudo slit placed along the major axis of the galaxy with both the positive half (blue squares) and negative half (red crosses) plotted on the same axes. Shown in comparison is the `DiskFit` result from the same cube (green circles). *Right:* Same, for a 110 mas pseudo slit.

possible slit we are not blending in any data from larger radii. However, where we do see differences is as we go further from the centre, and the SNR gets lower, the rotation curve becomes very sensitive to minor variations, and particularly with the negative half of the curve, which no longer traces the rotation profile after about 1.2 kpc. This is likely due to less of the gas following the ordered rotation in the outer regions of the galaxy, and with only sampling a single spaxel, there is not sufficient information to extract a reliable rotation curve in these regions.

For the 110 mas slit, we see a very different rotation curve, which is not reflective of that recovered with `DiskFit`. Here the entire central peak has been smoothed and flattened, meaning that the peak velocity is $226 \pm 4 \text{ km s}^{-1}$ at $687 \pm 43 \text{ pc}$ for the positive curve and $241 \pm 5 \text{ km s}^{-1}$ at $687 \pm 43 \text{ pc}$ for the negative curve. This is entirely to be expected, and is a consequence of co-adding spaxels across the width of the slit, essentially smearing out the peak. In contrast however, we do recover a better rotation profile in the regime between 0.8 and 1.3 kpc, where the rotation profile is primarily flat. This is because the flux being added here is from a part of the galaxy with a similar rotational velocity, thus does not have a major impact on the overall rotation curve. We still see, however, that

in the outer parts beyond 1.5 kpc the long slit still struggles, particularly for the negative curve, as would be expected as we have a lower SNR there in comparison to that which can be obtained by incorporating all the available data.

We see from this that while long slit spectroscopy is a very valuable tool, particularly with high spatial resolution, it is inferior to a full two dimensional extraction of the kinematics. Some of the issues that arise with the narrower slit could be negated with a longer integration time, by increasing the SNR per spaxel, but the coarser the sampling the worse the recovered profile will be. There are other issues associated with long slit spectroscopy for this work as well, for example misaligning the slit with the major axis of the galaxy, or missing the centre of the galaxy entirely. This would lead to systematic errors in the derivation of the kinematics. These findings are entirely in line with expectations, and are presented here to highlight the importance of IFUs, and the advantages that instruments such as HARMONI bring.

5.4 PSF Sensitivity

One point of particular interest that arose during our investigation was the impact of the shape of the PSF on the kinematics and rotation curve recovered. There were two main factors that we looked at in more detail, the impact of large wings on the AO assisted PSF, and PSF elongation, such as that which would occur with an off-axis natural guide star. In this section we will outline our investigative process and present the results, with their relevance for HARMONI.

5.4.1 AO PSF Behaviour

As we have seen previously in Figure 5.4 for example, the recovered rotation curve suffers from two major differences when compared to the intrinsic rotation of the galaxy. The first of these, the flattening and shifting away from the centre of the peak rotation is a result of beam smearing, which has been studied and discussed extensively in previous

PSF Name	^a Core Width	^a Wing Width	Notes
1MAS	1	-	Gaussian function.
15MAS	15	-	Gaussian function.
AIRY	10	-	Airy function.
DOUBLE*	1	100	Two Gaussian functions.
HSIM-LIKE**	15	88	Two Gaussian functions chosen to match the HSIM PSF.

^a mas

Table 5.4: **Custom HSIM PSFs to Investigate AO PSF Performance.** Additional PSFs created for the purpose of understanding the impact of the AO PSF on the recovered kinematics. Values quoted are the FWHMs of the respective PSF components. *DOUBLE PSF contains 10% of the flux in the core. **HSIM-LIKE PSF contains 70% of the flux in the core.

works (e.g. Davies et al., 2011; Burkert et al., 2016; Wuyts et al., 2016; Genzel et al., 2017 and references within). However, we also see a small but systematic lowering of the ‘flat’ part of the rotation curve, between around 0.6 and 1 kpc. We have eliminated a systematic bias in our analysis, such as assuming an incorrect inclination angle, as being the cause of this in Section 5.3.1. This therefore leads to the conclusion that this systematic lowering is a result of the observation process.²

One possible explanation for this systematic lowering of the rotation curve was that large wings of the AO PSF were convolving in light from larger radii, thus lowering our estimate of the rotation. For median seeing conditions in the H-band, these PSFs can have wings of widths up to 500 mas. To test this, we took advantage of the functionality within HSIM to provide user-defined custom PSFs. This allowed us to create simulated observations with five custom PSFs of varying core and wing sizes, which are tabulated in Table 5.4. Two of these PSFs (1MAS and 15MAS) were single-component Gaussian functions, while another two (DOUBLE and HSIM-LIKE) were made from a summation of two Gaussian functions to simulate a central core and wide wings. The final PSF was an Airy function with FWHM of 10 mas to simulate a pure PSF from a circular aperture with a central obscuration, such as will be present on the ELT.

²The lower rotation curve from around 1 to 1.3 kpc is also seen in the input cube compared to the intrinsically recovered velocity profile. This is due to some of the gas no longer following the ordered rotation of the galaxy at this point, which becomes even more evident after 1.4 kpc, at which point the gas no longer becomes a good tracer for the rotation profile of the galaxy.

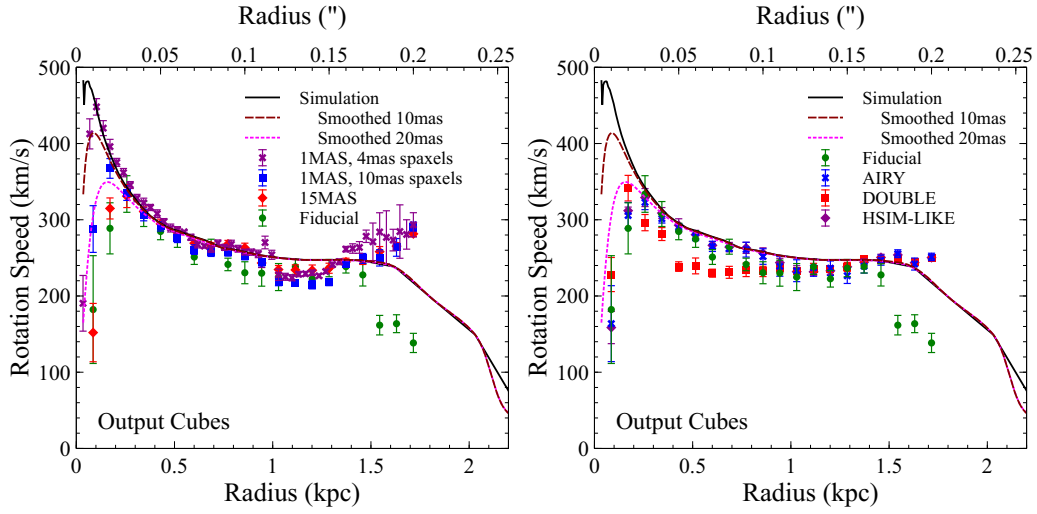


Figure 5.8: **Dependence of PSF Shape on Extracted Rotation Curve.** *Left:* Extracted rotation curves from simulated observations with the 1MAS PSF (purple crosses at 4 mas spaxel scale, blue squares at 10 mas spaxel scale), 15MAS PSF (red diamonds) and comparison with our fiducial cube (green circles). *Right:* Same, for the AIRY PSF (blue crosses), DOUBLE PSF (red squares) and HSIM-LIKE PSF (purple diamonds) which have extended wings.

In the left panel of Figure 5.8 we show the rotation curves recovered from `DiskFit` after simulated observations using the single-component Gaussian PSFs. Here we have also observed our galaxy at a spatial resolution of 4 mas with the 1MAS PSF (shown as purple crosses). This has almost perfectly recovered our input cube, with only a slight impact on the initial few points following the steepness of the curve. We do not see any systematic lowering of the ‘flat’ part of the rotation curve using the 1MAS PSF at this scale, or indeed at the 10 mas spaxel scale we have used throughout. The broader 15MAS PSF also does not show this behaviour, although as would be expected from observing at a coarser resolution we lose the true shape of the rotation curve in the innermost regions.

To determine if the wings of the PSF were responsible for this lowering, we created simulated observations using the DOUBLE and HSIM-LIKE PSFs. The DOUBLE PSF was created with a very narrow 1 mas FWHM core and 100 mas FWHM wings, created such that only 10% of the total flux was within the core of the PSF. The HSIM-LIKE PSF was a model created from the best fitting parameters of a double Gaussian function fitted to the PSF used with `HSIM`. This led to a core of FWHM of 15 mas and wings of

FWHM 88 mas, with 70% of the flux in the core. We can see from the right panel in Figure 5.8 that the AIRY PSF does show a slight lowering of the velocity curve, although is predominately still within 1σ up until the point at which the gas no longer traces the rotation of the galaxy. However for the other two PSFs, and in particular the DOUBLE PSF, with the widest wings, we see a strong systematic lowering of the recovered velocity profile.

This shows that it is indeed the broad wings of the PSF that caused the systematic lowering of the velocity curve. This is due to the extended component of the PSF convolving kinematic information from large radii into the central region, thus systematically lowering the rotation curve. Combined with the effect of sampling the rotation curve at the 10 mas spaxel scale, we are able to reproduce both effects that we have seen in our previous analysis.

5.4.2 PSF Elongation

Whilst the PSFs used in our simulated exposures were rotationally symmetric, this is often not the case with AO-assisted observations, and there is commonly an elongation of the PSF in the direction of the tip/tilt natural guide star (NGS). This is caused by the tip/tilt correction being weakest in the direction of the NGS, as the atmosphere common to both the target and the NGS is smallest in that direction. To simulate this, we created two custom PSFs with different levels of residual jitter; $4\times 2\text{JIT}$ and $6\times 2\text{JIT}$, summarised in Table 5.5, and shown in Figure 5.9.

PSF Name	x-jitter (mas)	y-jitter (mas)	Notes
$4\times 2\text{JIT}$	2	4	2:1 residual jitter ratio.
$6\times 2\text{JIT}$	2	6	3:1 residual jitter ratio.

Table 5.5: **Custom *HSIM* PSFs with Varying Levels of Jitter.** PSFs created with increasing levels of jitter along one axis, to simulate the impact of the NGS being off-axis from the target.

Our fiducial cube was simulated with no residual jitter along either axis, which whilst

it would never be physically possible to obtain a PSF like this, it provides a good basis for comparison. The $4\times 2\text{JIT}$ and $6\times 2\text{JIT}$, with residual jitter in a 2:1 and 3:1 ratio respectively, were used in simulated exposures at angles of 0° , 45° and 90° .

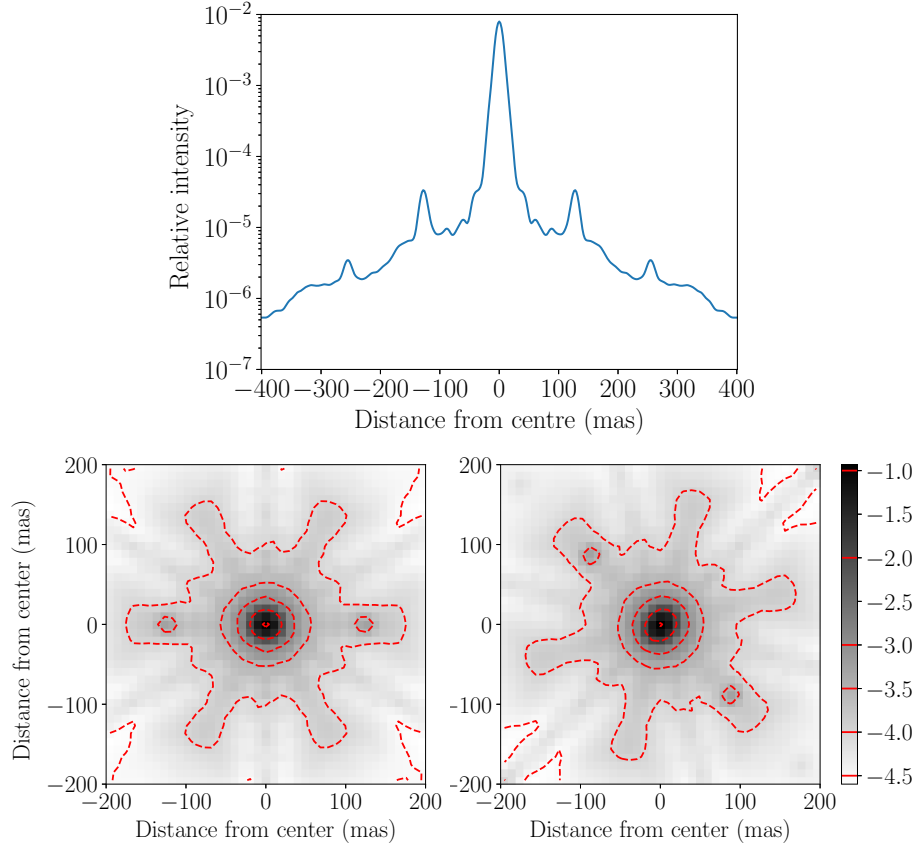


Figure 5.9: 1D and 2D Profiles of PSFs with Non-axisymmetric Residual Jitter.

Top: 1D slice across the centre of the fiducial PSF showing the relative intensity of the peak to the wings of the PSF. *Bottom left:* 2D plot in logarithmic scale of the central part of the 4×2 PSF at $1.6\ \mu\text{m}$ with 2 mas residual jitter vertically and 4 mas residual jitter horizontally. *Bottom right:* Same for the 6×2 PSF, with the larger residual jitter (6 mas) at 45° . The contours shown are both indicated by the logarithmic gradient scale on the right.

We show the recovered rotation curves in Figure 5.10 with comparisons to the fiducial cube, with the $4\times 2\text{JIT}$ PSF used for the left panel and the $6\times 2\text{JIT}$ PSF used for the right panel. All other simulation parameters were kept fixed. Here we see that in both cases with the elongated PSFs we do not track the rise of the rotation curve as well, consistent with elongation of the PSF convolving light from different radii together. We also see large deviation from the fiducial cube in the outer regions of the galaxy, where the gas

density is low, and convolution of different radii will have a larger impact on the recovered line of sight velocity. We see that in the 6×2 JIT PSF cubes the profile of the velocity curve varies more significantly than the fiducial case, and is particularly bad around 0.8-1 kpc when the PSF elongation is at 45° . Outside of these regions, we find that all the curves do still agree to within 1σ of each other, however the divergence increases as the elongation of the PSF increases.

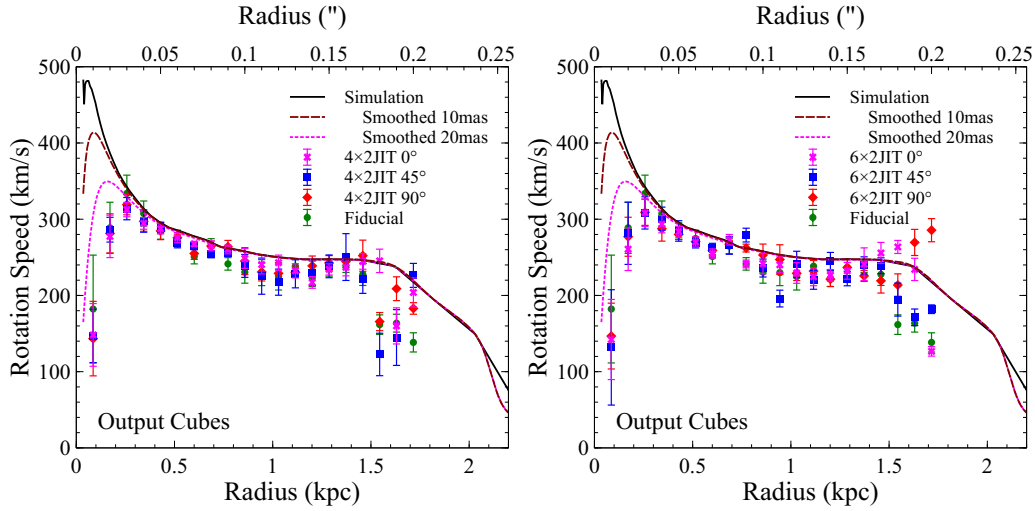


Figure 5.10: Dependence of PSF Elongation on Extracted Rotation Curve. *Left:* Extracted rotation curves from simulated observations with the 4×2 JIT PSF at angles of 0° (magenta crosses), 45° (blue squares) and 90° (red diamonds), and comparison with our fiducial cube (green circles). *Right:* Same, for the 6×2 PSF at 0° (magenta crosses), 45° (blue squares) and 90° (red diamonds).

This is important to understand, as it shows that good knowledge of the instrument and AO PSF is critical for correctly interpreting results of observations made with HARMONI. This is even more important when the residual jitter is not aligned with an axis of the IFU. We have shown here that asymmetries introduced into the PSF reduce our ability to extract an accurate rotation curve from the data, and this is heightened as the level of jitter is increased. As HARMONI can utilise a tip/tilt star of up to $60''$ separation when observing with Laser Tomographic Adaptive Optics (LTAO), it is vitally important that the shape of the PSF is well understood.

5.5 Conclusions

From this investigation we find that we can derive kinematic parameters of this galaxy and obtain a rotation curve after simulating realistic observations with HARMONI. Using a 2D kinematic fit with `DiskFit` the rotation curve recovered in the fiducial case closely matches the intrinsic rotation of the galaxy when we account for the impact of our finite spatial resolution.

We investigated the impact of exposure time, and thus integrated SNR of the galaxy, on the ability to extract a reliable rotation curve for our galaxy. We conclude that for a galaxy of this type at redshift 1.44, we can obtain this in 1 hour of integrated exposure time on HARMONI. This gave us sufficient data above our spectrally integrated SNR threshold of 7 per spaxel, below which we found our fits to be unreliable.

We do see the impact of beam smearing in our results, smoothing out our rotation curves and flattening the peak of rotation, which has been seen previously. However additionally we saw a systematic reduction along the flat part of the rotation curve, which we conclude is caused by large wings of the AO-assisted PSF convolving data from larger radii and reducing the magnitude of our observed line of sight velocity.

Related to this, we finally find that our observed rotation curve is strongly dependent on the instrument and AO PSF. In addition to lowering the flat part of the rotation curve, having asymmetries in the shape of our PSF affects the rotation curve we recover. This is found to be worse when the elongation is not aligned with an axis of the IFU. We therefore conclude that detailed knowledge of the instrument and AO PSF is crucial to correctly account for biases in data of this type taken with HARMONI.

Chapter 6

Conclusions

We present conclusions from the three investigations that we have completed during this work, all of which have revolved around improving the science that will be conducted with HARMONI. Initially we conclude the work conducted on infrared detector noise, then we summarise the findings from the telluric correction work, and end with a summary of the gas kinematic simulations. Our closing paragraphs summarise this entire thesis, and look forwards to the future of integral field spectroscopy.

6.1 Infrared Detector Noise

Chapters 2 and 3 focus on investigating the impact of infrared detector noise. Through a study of KMOS H2RG detectors and a MOONS engineering grade H4RG detector, we showed results of the types of noise present on these devices. We then implemented these into `HSIM` and ran numerous simulations to understand the dominant noise impacts.

6.1.1 Conclusions

Through analysing multiple KMOS dark exposures, our primary findings were to obtain a distribution of pixel read noise, that was similar for each detector. We were also able to confirm that this distribution held true over long timescales. Furthermore, we were able

to test our initial assumption that H4RG devices, such as will be used by HARMONI, behave the same as the H2RGs on KMOS. We repeated our analysis on an engineering grade H4RG detector procured for the MOONS instrument. From this we were able to recover a read noise distribution of the same shape as those found on the KMOS detectors.

We then combined this information, along with an existing HxRG noise generation tool, to make a substantial improvement to `HSIM`'s capabilities. As a result of this, it now simulates read noise in a more realistic manner, as well as including sporadic bad pixels, correlated and uncorrelated pink noise, picture frame effect, and alternating column noise. It also now creates simulated detectors and a method for converting from the (x, y, λ) datacube to the spectrograph focal plane, mimicking the dispensed spectral format on the infrared detectors, and vice versa, allowing for future upgrades to be easily implemented.

With this in place, we conducted simulations to understand the contributions from each of the new noise processes included. Our tests focussed on point source sensitivity, spatial variation, ability to recover the true flux, and impact on the SNR. From this, our findings showed that all these additional noise processes contribute in different ways and should be accounted for accordingly. Aside from read noise, which causes an overall reduction in sensitivity, all these effects are spatially varying. Pink noise was found to be particularly problematic, as it can vary significantly between different observations. ACN and picture frame noise can also cause issues in areas of the detector in which they occur.

We have suggested a number of options for minimising these effects should they manifest in the HARMONI detectors. Read noise is hard to mitigate, however as this is a flat effect, it primarily costs additional exposure time to counteract. Bad pixels need to be carefully considered, as these can skew data disproportionately on areas in which they are present. However, we have seen that there are some pixels that return 'hot' values that are consistent, and therefore could potentially be used, providing they are considered correctly in the noise.

Pink noise should be removed as much as possible, and we suggest reference pixel subtraction is the best method for this. This should be done in combination with identi-

fying bad pixels, both static and transient, as bad pixels within the reference columns can cause further issues. Where possible, observations should avoid placing targets of interest such that they fall near the edge of a detector, to minimise the impact of the picture frame effect. With HARMONI, this may mean that targets are placed off-centre from the field of view.

6.1.2 Continuing Work

Advancements are continually being made in infrared detector technology, and during the course of this work the engineering grade H4RG for MOONS was further characterised to reduce the amount of noise present (Ives et al., 2020). As further science grade H4RGs become available, the understanding of the levels of noise present will be improved, along with potential new issues that haven't yet been considered. `HSIM` has been set up such that these can be incorporated into the simulator as required.

In addition, there are known sources of noise that are seen on infrared detectors that are not yet simulated in `HSIM`. One such effect is persistence, which is where trapped charge from previous exposures slowly leaks out over time, potentially appearing in future observations. There are ways to mitigate the impact of persistence, and potentially correct for it afterwards (e.g. McLeod and Smith, 2016; Neeser, 2021), however a procedure for including this in `HSIM` could be considered in future.

A major upcoming milestone will be characterising the HARMONI H4RG devices to understand the extent to which we see the noise sources that we have investigated. This can then be used to refine the predictions that we have made, and make informed decisions about minimising the impact of the noise on the science observations. Part of this is working in alignment with the pipeline development team, so as to ensure that the relevant processes are implemented for optimal data extraction.

6.2 Impact of Under-Illuminated Slit on Telluric Correction

Chapter 4 looked at issues associated with telluric correction, with a focus on the impact of the LSF. We modelled this effect, and took data with SINFONI to investigate it empirically. We then reduced our data and attempted four different telluric correction methods to determine which was the most effective.

6.2.1 Conclusions

We produced detailed models of the SINFONI LSF, convolving functional forms for the input slit, spectrograph, grating and output slit to obtain an idealised but representative LSF. We then changed the input slit from a 2 pixel wide boxcar function to a 0.48 pixel FWHM Gaussian function, to simulate the impact of an under-illuminated entrance slit, such as would be the case when observing a point source with good AO correction at a coarse spaxel scale. We determined that this would result in a 14% reduction in the overall LSF.

We then obtained time with SINFONI on the VLT to take a series of observations set up to reproduce this effect and analyse its impact. We observed a star suitable for use as a telluric standard, CD-23 13701, and an extended science object, Callisto, both with AO correction at the coarsest 250×125 mas spaxel scale. We additionally used 25×12.5 mas data to confirm the performance of the AO, and to determine the amount by which the entrance slit was under-illuminated.

We reduced this data and extracted spectra for both targets. We then used `molecfit` to fit the data and recover a corresponding LSF. We found that there was a difference between the CD-23 13701 data and the Callisto data of $14.0 \pm 2.1\%$, in agreement with that predicted by our models.

We then used four methods to attempt telluric correction on our Callisto data. We used the telluric standard, CD-23 13701, directly, `molecfit` to both CD-23 13701 and Cal-

listo, and finally the `molecfit` results for CD-23 13701, but broadened to the Callisto LSF. To determine the effectiveness of each of these methods, we divided out a solar irradiance spectrum from Callisto, and determined the corresponding residual sum of squares (RSS) for each of our approaches. We also repeated this for multiple spaxels of Callisto data. From this we determined that the best correction could be obtained by directly fitting to Callisto, however with good LSF knowledge we could almost replicate that by convolving the fit using CD-23 13701.

From this we conclude that where possible, correcting the telluric features directly on the science data has the potential for the best results. This is because in this case there is no difference in the atmospheric profile between the model telluric spectrum and the science spectrum. However in cases where this is not possible, we conclude that when the entrance slit is under-illuminated, correcting ‘blind’, without knowledge of the LSF, will give a worse result. However, when we add in the knowledge of the LSF, we can almost recover the same correction as the direct fit to the data. It is therefore of great importance to have a good understanding of the HARMONI LSFs, to enable accurate telluric correction.

6.2.2 Continuing Work

The importance of good telluric correction will always be relevant to ground-based astronomy in the infrared. Tools such as `molecfit` (Smette et al., 2015) assist in this by fitting the molecular absorption features to the spectra so they can be divided out. This software is continuously being developed and improved to increase the reliability and accuracy of the results. One of the current aims of this software is to be able to perform an entirely ‘blind’ correction. This requires accurate wavelength calibration, relatively stable or well-known molecular abundances, and good knowledge of the LSF (Smette, 2020).

As we have seen in this work, using an incorrect LSF when correcting for telluric absorption will cause additional residuals in the science data. In order for a ‘blind’ correction to be possible, one of the future tasks would be to accurately characterise the LSF

of HARMONI in every offered configuration. This process also requires good external wavelength calibrations, particularly if correcting a large range simultaneously. The abundance of each molecule to be corrected for must also then be known with high accuracy, either through radiometry data, or through separate calibrations. All of these must be taken into consideration during the development process of HARMONI.

The goal of this is for fast, accurate telluric correction without the need for long calibration exposures, which reduce the amount of potential science time with the ELT. `Molecfit` will then be integrated as a telluric correction method into the instrument data reduction pipelines, as it already is for KMOS (Coccato et al., 2019).

6.3 Recovering Gas Kinematics using `HSIM`

In Chapter 5 we used `HSIM` to simulate a galaxy at redshift 1.44, and determine how well we could recover the kinematic properties. We also conducted an investigation into the impacts of the AO PSF when determining these characteristics.

6.3.1 Conclusions

We conducted a full simulated observation of a mock galaxy, including data reduction and analysis. We took a galaxy from the `NUTFB` suite and added $H\alpha$ emission to create an input cube that could be simulated with `HSIM`. We artificially placed this galaxy at a redshift of 1.44 so that the $H\alpha$ line fell within the observable range of HARMONI. The pipeline through which this was done has been made publicly available for use.

We then ran multiple simulations with `HSIM`, primarily varying the exposure time, however we also investigated the impact of the spectral resolution. From these, a data analysis pipeline was developed to convert our output cubes into velocity maps from which we could extract the galaxy kinematics. This was primarily done using `DiskFit`, a software package designed to extract rotation curves from a 2D velocity profile, although we also applied a traditional long-slit approach for comparison.

Our results showed that we were able to closely recover the intrinsic rotation profile of the galaxy when we account for beam smearing effects in our observations. These served to flatten out the peak of rotation, lowering it and causing it to appear at larger radii. One hour of simulated observation time was sufficient to extract a reliable rotation curve, with longer exposures increasing the SNR of our data. We did not see a significant difference when observing at a higher spectral resolution, as would be expected when we were already sufficiently sampling the $H\alpha$ line. As would also be expected, not including the full 2D data by using a pseudo-slit returned less reliable kinematics.

We also found that our results were strongly dependent on the instrument and AO PSF. Large wings in the PSF convolve data from larger radii, reducing the magnitude of the observed line of sight velocity. In addition, we found that asymmetries in the AO PSF affected the recovered rotation curve. This was worse when our PSF was elongated in a direction not aligned with an axis of the IFU. We therefore conclude at the end of this work that it is vital to understand the instrument and AO PSF. Any biases that are present should be compensated for in order to be able to conduct the best possible science with HARMONI.

6.3.2 Continuing Work

In the time prior to the completion of HARMONI and first light on the ELT, planned at the time of writing for 2027, there is the potential for more science simulation and prediction. Within the scope of the work we have carried out here, we could expand our study to a more physically realistic galaxy, or consider a galaxy at other redshifts that still fall within the wavelength range of HARMONI. Additionally, we could use different emission lines to determine the capabilities of the instrument to achieve other science goals, such as the work conducted by Grisdale et al. (2021) on Population III stars.

HSIM is ideally situated to work alongside the James Webb Space Telescope (JWST; Gardner et al., 2006), which was successfully launched in 2021. JWST will be operational for many years prior to the completion of the ELT, and features an IFU covering much

of the spectral range of HARMONI. JWST will enhance our overall understanding of infrared astronomy, and it is predicted that we will discover things that were entirely unexpected. Through tools such as `HSIM`, it is possible to determine the capabilities of the ELT to act as a follow-up instrument for the new and exciting discoveries that will be made with JWST.

6.4 Closing Statement

This work has covered a number of elements of infrared spectroscopy, both using existing integral field spectrographs and looking forwards to the upcoming HARMONI IFS for the ELT. As part of this we have shown the importance of understanding the noise behaviour of the infrared detectors, minimising it where possible, and characterising it where not. We have also shown that we must be careful when performing telluric correction using a star which under-illuminates the slit of the spectrograph. These are both issues that could potentially impact the science that will be carried out with HARMONI, therefore studying and understanding these whilst it is still in development is crucial for the final performance of the instrument.

For an understanding of the potential of HARMONI, we used `HSIM` to perform a set of simulations investigating the ability to recover spatially resolved gas kinematics at redshifts of $\sim 1-2$. We determined that HARMONI will be able to utilise the greater spatial resolving power of the ELT to recover the kinematics of a galaxy at $z=1.44$ with 1 hour of observation time. However, in the process of this, we identified that our recovered kinematics were strongly dependent on the AO PSF, and thus good knowledge of this is also very important to be able to conduct the best science that HARMONI will be capable of.

We are entering a very exciting era for visible and near-infrared spectroscopy, and it is hoped that this work will help towards achieving the science goals of HARMONI on the ELT.

Appendix A

Running HSIM using ADS

In this appendix we explain the new inputs and options required to run HSIM using the advanced detector systematics (ADS) option. This is entirely contained within HSIM, and requires no additional dependencies or modules to be installed. When selected, HSIM creates eight detectors for use in the simulation, and divides the HARMONI field of view across them. All the same outputs will be created as before, along with a number that are unique to the ADS mode.

A.1 ADS Mode Inputs and Options

There are two new options that must be passed to HSIM, whether it is run from the command line or through GUI. These are:

- **Using ADS Mode [True/False]:** The flag to tell HSIM to use advanced detector systematics.
- **Detector Temporary Path:** A file path to which the user has write permissions, to save temporary detector files to.

In addition, there are a number of parameters that define the nature of the created detectors, that are contained within the `config.py` file. These are as follows, with default values in brackets:

- **rd (2.845 e⁻)**: The peak amount of read noise to include in a standard observation.
- **rd_lowexp (12.0 e⁻)**: The peak amount of read noise to include in a short exposure.
- **rn_file (kmos_rn.fits)**: A fits file containing the distribution of read noise to use. The default file is included with the `HSIM` distribution and is based off of analysis of the KMOS detectors (see Chapter 2).
- **pedestal (4 e⁻)**: Magnitude of pedestal drift. Not currently used.
- **c_pink (3 e⁻)**: Standard deviation of correlated pink noise.
- **u_pink (1 e⁻)**: Standard deviation of uncorrelated pink noise.
- **acn (0.5 e⁻)**: Standard deviation of ACN.
- **pca0_amp (0.2 e⁻)**: Standard deviation of `pca0`, used for incorporating the picture frame effect.
- **kTC_noise (29 e⁻)**: Amount of kTC noise to include if creating a SUTR detector cube. Not currently used.
- **bias_amp (500)**: Factor by which to multiply the `pca0` file by to include a bias pattern in SUTR detector cubes. Not currently used.
- **bias_offset (5000 e⁻)**: Level at which to start integrating pixels if creating a SUTR detector cube. Not currently used.
- **force_new (False)**: Option to force `HSIM` to create new detectors even if it finds existing ones in the detector temporary path. By default, `HSIM` will use these if it can, for both computational efficiency and to ensure the same detectors are used between multiple simulations. If this is not desired, set this flag to ‘True’ to create a new set of detectors on subsequent simulations.

A `pca0` file is required to simulate the picture frame effect, with `nirspec_pca0.fits` included as a default file with the distribution. If the user wishes to change this, or the read noise distribution, they can be found in the folder ‘`sim_data/detectors`’ within `HSIM`.

A.2 ADS Mode Output Cubes

Alongside the standard outputs of `HSIM`, there are a number that are unique to the ADS mode, and can be utilised to give further insight into the simulation results. These are:

- **`_all_dets.fits`**: A cube of size (8, 4096, 4096) containing all of the detector noise used, in detector space. This is multiplied by $\sqrt{\text{NDIT}}$ to approximate the accumulative noise over multiple integrations.
- **`_detector_SNR.fits`**: The signal to noise ratio, including the exact amount of noise taken from the used detectors in the simulation. This is better than would be realistically possible to know in a real simulation, but gives an indication of spatial differences in the SNR as a result of detector noise.
- **`_det_std.fits`**: The noise standard deviation, using the detector noise in the same manner as the `_detector_SNR` output.
- **`_used_dets.fits`**: A cube containing just the detector noise included in the simulation, of the same dimensions as the input datacube.

If the input datacube was larger than the `HARMONI` field of view, or covered a spectral range greater than is possible to fit on the detector, the ADS mode will crop it such that only the central region is simulated, and a warning will be raised in the `.log` file. In this case, all the returned outputs will be equal in size to the spatial and/or spectral limits of `HARMONI`.

Bibliography

- Allen, C. W. (2000). *Allen's Astrophysical Quantities*. Springer, New York.
- Allen, D. A. and Cragg, T. A. (1983). The AAOJHKL photometric standards. *MNRAS*, 203:777–783.
- Allington-Smith, J., Murray, G., Content, R., Dodsworth, G., Davies, R., Miller, B. W., et al. (2002). Integral Field Spectroscopy with the Gemini Multiobject Spectrograph. I. Design, Construction, and Testing. *PASP*, 114(798):892–912.
- Allington-Smith, J. R., Content, R., and Haynes, R. (1998). New developments in integral field spectroscopy. In D'Odorico, S., editor, *Optical Astronomical Instrumentation*, volume 3355, pages 196 – 205. International Society for Optics and Photonics, SPIE.
- Amico, P., Marchetti, E., Pedichini, F., Baruffolo, A., Delabre, B., Duchateau, M., et al. (2012). The design of ERIS for the VLT. In McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446, pages 703 – 719. International Society for Optics and Photonics, SPIE.
- Antichi, J., Dohlen, K., Gratton, R. G., Mesa, D., Claudi, R. U., Giro, E., et al. (2009). BIGRE: A Low Cross-Talk Integral Field Unit Tailored for Extrasolar Planets Imaging Spectroscopy. *ApJ*, 695(2):1042–1057.
- Archdeacon, T. J. (1994). *Correlation and regression analysis: A historian's guide*, chapter 9, pages 161–162. University of Wisconsin Press.

- Arribas, S., Mediavilla, E., García-Lorenzo, B., del Burgo, C., and Fuensalida, J. J. (1999). Differential atmospheric refraction in integral-field spectroscopy: Effects and correction. *Atmospheric refraction in IFS. AApS*, 136:189–192.
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., et al. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *AJ*, 156(3):123.
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., et al. (2013). Astropy: A community Python package for astronomy. *A&A*, 558:A33.
- Bacon, R., Accardo, M., Adjali, L., Anwand, H., Bauer, S., Biswas, I., et al. (2010). The MUSE second-generation VLT instrument. In McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 773508.
- Bacon, R., Adam, G., Baranne, A., Courtès, G., Dubet, D., Dubois, J. P., et al. (1995). 3D spectrography at high spatial resolution. I. Concept and realization of the integral field spectrograph TIGER. *AApS*, 113:347.
- Bacon, R., Copin, Y., Monnet, G., Miller, B. W., Allington-Smith, J. R., Bureau, M., et al. (2001). The SAURON project - I. The panoramic integral-field spectrograph. *MNRAS*, 326(1):23–35.
- Bailey, J., Simpson, A., and Crisp, D. (2007). Correcting Infrared Spectra for Atmospheric Transmission. *PASP*, 119(852):228–236.
- Beletic, J. W., Blank, R., Gulbransen, D., Lee, D., Loose, M., Piquette, E. C., et al. (2008). Teledyne Imaging Sensors: infrared imaging technologies for astronomy and civil space. In Dorn, D. A. and Holland, A. D., editors, *High Energy, Optical, and*

- Infrared Detectors for Astronomy III*, volume 7021 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 70210H.
- Beuzit, J. L., Vigan, A., Mouillet, D., Dohlen, K., Gratton, R., Boccaletti, A., et al. (2019). SPHERE: the exoplanet imager for the Very Large Telescope. *A&A*, 631:A155.
- Bevington, P. R. and Robinson, D. K. (2003). *Data reduction and error analysis for the physical sciences*, chapter 4, pages 56–57. McGraw-Hill, New York, NY, 3 edition.
- Biesiadzinski, T., Lorenzon, W., Newman, R., Schubnell, M., Tarlé, G., and Weaverdyck, C. (2011). Reciprocity Failure in HgCdTe Detectors: Measurements and Mitigation. *Publications of the Astronomical Society of the Pacific*, 123(906):958–963.
- Biesiadzinski, T., Lorenzon, W., Schubnell, M., Tarlé, G., and Weaverdyck, C. (2014). NIR Detector Nonlinearity and Quantum Efficiency. *Publications of the Astronomical Society of the Pacific*, 126(937):243–249.
- Blank, R., Anglin, S., Beletic, J. W., Baia, Y., Buck, S., Bhargava, S., et al. (2011). The HxRG Family of High Performance Image Sensors for Astronomy. In Kuhn, J. R., Harrington, D. M., Lin, H., Berdyugina, S. V., Trujillo-Bueno, J., Keil, S. L., and Rimmele, T., editors, *Solar Polarization 6*, volume 437 of *Astronomical Society of the Pacific Conference Series*, page 383.
- Blitz, L. (1993). *Protostars and Planets III*. University of Arizona Press.
- Bowen, I. S. (1938). The Image-Slicer, a Device for Reducing Loss of Light at Slit of Stellar Spectrograph. *ApJ*, 88:113.
- Bundy, K., Bershady, M. A., Law, D. R., Yan, R., Drory, N., MacDonald, N., et al. (2015). Overview of the SDSS-IV MaNGA Survey: Mapping nearby Galaxies at Apache Point Observatory. *ApJ*, 798:7.
- Burkert, A., Förster Schreiber, N. M., Genzel, R., Lang, P., Tacconi, L. J., Wisnioski,

- E., et al. (2016). The Angular Momentum Distribution and Baryon Content of Star-forming Galaxies at $z \sim 1-3$. *ApJ*, 826(2):214.
- Bushouse, H., Baggett, S., Brown, T., Hartig, G., Hilbert, B., Hill, R., et al. (2008). Wide-Field Camera 3 ground testing and calibration - art. no. 62650J. *Proceedings of SPIE - The International Society for Optical Engineering*.
- Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., and Storchi-Bergmann, T. (2000). The Dust Content and Opacity of Actively Star-forming Galaxies. *The Astrophysical Journal*, 533(2):682–695.
- Cappellari, M. and Copin, Y. (2003). Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations. *MNRAS*, 342(2):345–354.
- Carraro, G., Lia, C., and Chiosi, C. (1998). Galaxy formation and evolution - I. The Padua tree-sph code (pd-sph). *MNRAS*, 297(4):1021–1040.
- Carrasco, E. and Parry, I. R. (1994). A method for determining the focal ratio degradation of optical fibres for astronomy. *Monthly Notices of the Royal Astronomical Society*, 271(1):1–12.
- Carroll, B. W. and Ostlie, D. A. (2017). *An Introduction to Modern Astrophysics*, chapter 12. Cambridge University Press, 2 edition.
- Chance, K. and Kurucz, R. L. (2010). An improved high-resolution solar reference spectrum for earth’s atmosphere measurements in the ultraviolet, visible, and near infrared. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 111(9):1289–1295.
- Chédin, A., Husson, N., and Scott, N. A. (1982). Une banque de données pour l’étude des phénomènes de transfert radiatif dans les atmosphères planétaires: la banque GEISA. *Bulletin d’Information du Centre de Données Stellaires*.
- Cirasuolo, M., Afonso, J., Bender, R., Bonifacio, P., Evans, C., Kaper, L., et al. (2012). MOONS: a multi-object optical and near-infrared spectrograph for the VLT. In

- McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 84460S.
- Clough, S. A., Kneizys, F. X., Rothman, L. S., and Gallery, W. O. (1981). Atmospheric Spectral Transmittance And Radiance: FASCOD1 B. In Fan, R. W., editor, *Atmospheric Transmission*, volume 0277, pages 152 – 167. International Society for Optics and Photonics, SPIE.
- Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al. (2005). Atmospheric radiative transfer modeling: a summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(2):233–244.
- Coccatto, L., Freudling, W., Smette, A., Sani, E., Escartin, J. A., Jung, Y., and Bazin, G. (2019). On the Telluric Correction of KMOS Spectra. *The Messenger*, 177:14–18.
- Collier, R. J. (1971). Program of the 1971 Spring Meeting of the Optical Society of America. *J. Opt. Soc. Am.*, 61(5):648–697.
- Cotton, D. V., Bailey, J., and Kedziora-Chudczer, L. (2014). Atmospheric modelling for the removal of telluric features from infrared planetary spectra. *MNRAS*, 439(1):387–399.
- Courteau, S. and Dutton, A. A. (2015). On the Global Mass Distribution in Disk Galaxies. *ApJL*, 801(2):L20.
- Courtès, G. (1981). An Integral Field Spectrograph (IFS) for Large Telescopes. In Humphries, C. M., editor, *Instrumentation for Astronomy with Large Optical Telescopes*, pages 123 – 129. Springer Netherlands.
- Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., Bryant, J. J., Fogarty, L., Richards,

- S., et al. (2012). The Sydney-AAO Multi-object Integral field spectrograph. *MNRAS*, 421(1):872–893.
- Davies, R., Förster Schreiber, N. M., Cresci, G., Genzel, R., Bouché, N., Burkert, A., et al. (2011). How Well Can We Measure the Intrinsic Velocity Dispersion of Distant Disk Galaxies? *ApJ*, 741(2):69.
- de Blok, W. J. G., Walter, F., Brinks, E., Trachternach, C., Oh, S. H., and Kennicutt, R. C., J. (2008). High-Resolution Rotation Curves and Galaxy Mass Models from THINGS. *AJ*, 136(6):2648–2719.
- Domínguez, A., Siana, B., Henry, A. L., Scarlata, C., Bedregal, A. G., Malkan, M., et al. (2013). Dust Extinction from Balmer Decrements of Star-forming Galaxies at $0.75 \leq z \leq 1.5$ with Hubble Space Telescope/Wide-Field-Camera 3 Spectroscopy from the WFC3 Infrared Spectroscopic Parallel Survey. *ApJ*, 763(2):145.
- Drory, N., MacDonald, N., Bershady, M. A., Bundy, K., Gunn, J., Law, D. R., et al. (2015). The MaNGA Integral Field Unit Fiber Feed System for the Sloan 2.5 m Telescope. *AJ*, 149(2):77.
- Dubois, Y. and Teyssier, R. (2008). On the onset of galactic winds in quiescent star forming galaxies. *A&A*, 477(1):79–94.
- Dudhia, A. (2017). The Reference Forward Model (RFM). *Journal of Quantitative Spectroscopy and Radiative Transfer*, 186:243–253. Satellite Remote Sensing and Spectroscopy: Joint ACE-Odin Meeting, October 2015.
- Eisenhauer, F., Abuter, R., Bickert, K., Biancat-Marchet, F., Bonnet, H., Brynnel, J., et al. (2003). SINFONI: integral field spectroscopy at 50-milli-arcsecond resolution with the ESO VLT. In Iye, M. and Moorwood, A. F. M., editors, *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*, volume 4841, pages 1548 – 1561. International Society for Optics and Photonics, SPIE.

- Filippenko, A. V. (1982). The importance of atmospheric differential refraction in spectrophotometry. *PASP*, 94:715–721.
- Finger, G., Mehrgan, H., Meyer, M., Moorwood, A. F., Nicolini, G., and Stegmeier, J. (2000). Performance of large-format HgCdTe and InSb arrays for low-background applications. In Iye, M. and Moorwood, A. F., editors, *Optical and IR Telescope Instrumentation and Detectors*, volume 4008 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1280–1297.
- Fitzpatrick, E. L. (1999). Correcting for the Effects of Interstellar Extinction. *PASP*, 111(755):63–75.
- Förster Schreiber, N. M., Genzel, R., Bouché, N., Cresci, G., Davies, R., Buschkamp, P., et al. (2009). The SINS Survey: SINFONI Integral Field Spectroscopy of $z \sim 2$ Star-forming Galaxies. *ApJ*, 706(2):1364–1428.
- Fowler, A. M. and Gatley, I. (1990). Demonstration of an Algorithm for Read-Noise Reduction in Infrared Arrays. *ApJL*, 353:L33.
- Fried, D. L. (1965). Statistics of a Geometric Representation of Wavefront Distortion. *Journal of the Optical Society of America (1917-1983)*, 55(11):1427–1431.
- Fried, D. L. (1966). Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures. *Journal of the Optical Society of America (1917-1983)*, 56(10):1372.
- García-Lorenzo, B., Monreal-Ibero, A., Mediavilla, E., Pereira-Santaella, M., and Thatte, N. (2019). Black Hole-Galaxy Scaling Relation Evolution From $z \sim 2.5$: Simulated Observations With HARMONI on the ELT. *Frontiers in Astronomy and Space Sciences*, 6:73.
- Gardner, J. P., Mather, J. C., Clampin, M., Doyon, R., Greenhouse, M. A., Hammel, H. B.,

- et al. (2006). The James Webb Space Telescope. *Space Science Reviews*, 123(4):485–606.
- Garnett, J. D., Farris, M. C., Wong, S. S., Zandian, M., Hall, D. N., Jacobson, S., et al. (2004). 2K X 2K molecular beam epitaxy HgCdTe detectors for the James Webb Space Telescope NIRCам instrument. In Garnett, J. D. and Beletic, J. W., editors, *Optical and Infrared Detectors for Astronomy*, volume 5499 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 35–46.
- Genzel, R., Förster Schreiber, N. M., Übler, H., Lang, P., Naab, T., Bender, R., et al. (2017). Strongly baryon-dominated disk galaxies at the peak of galaxy formation ten billion years ago. *Nature*, 543(7645):397–401.
- Genzel, R., Price, S. H., Übler, H., Schreiber, N. M. F., Shimizu, T. T., Tacconi, L. J., et al. (2020). Rotation curves in $z \sim 1\text{--}2$ star-forming disks: Evidence for cored dark matter distributions. *The Astrophysical Journal*, 902(2):98.
- Genzel, R., Tacconi, L. J., Eisenhauer, F., Förster Schreiber, N. M., Cimatti, A., Daddi, E., et al. (2006). The rapid formation of a large rotating disk galaxy three billion years after the Big Bang. *Nature*, 442(7104):786–789.
- George, E. M., Gräff, D., Hartl, M., Huber, H., Eisenhauer, F., and Feuchtgruber, H. (2017). Complex spectral line profiles resulting from cryogenic deformation of the SIN-FONI/SPIFFI diffraction gratings. *Journal of Astronomical Telescopes, Instruments, and Systems*, 3:035002.
- George, E. M., Gräff, D., Feuchtgruber, H., Hartl, M., Eisenhauer, F., Buron, A., et al. (2016). Making SPIFFI SPIFFIER: upgrade of the SPIFFI instrument for use in ERIS and performance analysis from re-commissioning. In Evans, C. J., Simard, L., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 9908, pages 123 – 142. International Society for Optics and Photonics, SPIE.

- George, E. M., Tulloch, S. M., and Ives, D. J. (2018). Fast method of crosstalk characterization for HxRG detectors. *arXiv e-prints*, page arXiv:1808.00790.
- Glazebrook, K. (2013). The Dawes Review 1: Kinematic Studies of Star-Forming Galaxies Across Cosmic Time. *Publications of the Astronomical Society of Australia*, 30.
- Gonzalez, O. A. and Battaglia, G. (2018). From the inner Milky Way to Local Volume galaxies: resolved stellar populations with ELT-HARMONI. *arXiv e-prints*, page arXiv:1810.04422.
- González Delgado, R. M. and Pérez, E. (1997). H II Region Population in a Sample of Nearby Galaxies with Nuclear Activity. II. Luminosity Function, Size, and Radial Distributions. *ApJS*, 108(1):199–228.
- Greenwood, D. P. (1977). Bandwidth specification for adaptive optics systems. *Journal of the Optical Society of America (1917-1983)*, 67:390–393.
- Grisdale, K., Thatte, N., Devriendt, J., Pereira-Santaella, M., Slyz, A., Kimm, T., et al. (2021). Predicting the observability of population III stars with ELT-HARMONI via the helium 1640 Å emission line. *MNRAS*, 501(4):5517–5537.
- Guenther, R. D. (1990). *Modern Optics*. Oxford University Press (OUP).
- Guidi, G., Casado, J., Ascasibar, Y., García-Benito, R., Galbany, L., Sánchez-Blázquez, P., et al. (2018). The SELGIFS data challenge: generating synthetic observations of CALIFA galaxies from hydrodynamical simulations. *MNRAS*, 479(1):917–931.
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., et al. (2006). The 2.5 m Telescope of the Sloan Digital Sky Survey. *AJ*, 131(4):2332–2359.
- Hardy, J. W. (1998). *Adaptive Optics for Astronomical Telescopes*. Oxford University Press (OUP).
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., et al. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

- Houghton, R. (2016). NIR Detector Systematics and Data Reduction Wishlist. Technical Note, University of Oxford.
- Hughes, A., Wong, T., Ott, J., Muller, E., Pineda, J. L., Mizuno, Y., et al. (2010). Physical properties of giant molecular clouds in the Large Magellanic Cloud. *Monthly Notices of the Royal Astronomical Society*, 406(3):2065–2086.
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3):90–95.
- Husson, N., Bonnet, B., Scott, N., and Chedin, A. (1992). Management and study of spectroscopic information: The GEISA program. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 48(5):509–518. Special Issue Conference on Molecular Spectroscopic Databases.
- Hutterer, V., Ramlau, R., and Shatokhina, I. (2019). Real-time adaptive optics with pyramid wavefront sensors: part I. A theoretical analysis of the pyramid sensor model. *Inverse Problems*, 35(4):045007.
- Iserlohe, C., Tecza, M., Eisenhauer, F., Genzel, R., Thatte, N. A., Abuter, R., et al. (2004). On-sky performance of SPIFFI: the integral field spectrometer for SINFONI at the VLT. In Moorwood, A. F. M. and Iye, M., editors, *Ground-based Instrumentation for Astronomy*, volume 5492, pages 1123 – 1134. International Society for Optics and Photonics, SPIE.
- Ives, D., Alvarez, D., Bezawada, N., George, E., and Serra, B. (2020). Characterisation, performance and operational aspects of the H4RG-15 near infrared detectors for the MOONS instrument. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 11454 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 114541N.
- Jacquinet-Husson, N., Armante, R., Scott, N., Chédin, A., Crépeau, L., Boutamine, C., et al. (2016). The 2015 edition of the GEISA spectroscopic database. *Journal of*

- Molecular Spectroscopy*, 327:31–72. New Visions of Spectroscopic Databases, Volume II.
- Jeans, J. H. (1902). I. The Stability of a Spherical Nebula. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 199:1–53.
- Jerram, P. and Beletic, J. (2019). Teledyne’s high performance infrared detectors for Space missions. In *International Conference on Space Optics — ICSO 2018*, volume 11180 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 111803D.
- Johnson, J. B. (1928). Thermal Agitation of Electricity in Conductors. *Physical Review*, 32(1):97–109.
- Jones, T. A., Swinbank, A. M., Ellis, R. S., Richard, J., and Stark, D. P. (2010). Resolved spectroscopy of gravitationally lensed galaxies: recovering coherent velocity fields in subluminal $z \sim 2$ –3 galaxies. *MNRAS*, 404(3):1247–1262.
- Kahng, D. (1960). Electric Field Controlled Semiconductor Device. U.S. Patent No. 3,102,230.
- Käufel, H.-U., Ballester, P., Biereichel, P., Delabre, B., Donaldson, R., Dorn, R., et al. (2004). CRIRES: a high-resolution infrared spectrograph for ESO’s VLT. In Moorwood, A. F. M. and Iye, M., editors, *Ground-based Instrumentation for Astronomy*, volume 5492 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1218–1227.
- Kausch, W., Noll, S., Smette, A., Kimeswenger, S., Barden, M., Szyszka, C., et al. (2015). Molecfit: A general tool for telluric absorption correction. II. Quantitative evaluation on ESO-VLT/X-Shooterspectra. *A&A*, 576:A78.

- Kaviraj, S., Laigle, C., Kimm, T., Devriendt, J. E. G., Dubois, Y., Pichon, C., et al. (2017). The Horizon-AGN simulation: evolution of galaxy properties over cosmic time. *MNRAS*, 467(4):4739–4752.
- Kay, S. T., Pearce, F. R., Frenk, C. S., and Jenkins, A. (2002). Including star formation and supernova feedback within cosmological simulations of galaxy formation. *MNRAS*, 330(1):113–128.
- Kellerer, A. (2007). *Assessing time scales of atmospheric turbulence at observatory sites*. PhD thesis, Université Paris-Diderot.
- Kendrew, S., Zieleniewski, S., Houghton, R. C. W., Thatte, N., Devriendt, J., Tecza, M., et al. (2016). Simulated stellar kinematics studies of high-redshift galaxies with the HARMONI Integral Field Spectrograph. *Monthly Notices of the Royal Astronomical Society*, 458(3):2405–2422.
- Kennicutt, R. C., J. (1983). The rate of star formation in normal disk galaxies. *ApJ*, 272:54–67.
- Kennicutt, Robert C., J., Tamblyn, P., and Congdon, C. E. (1994). Past and Future Star Formation in Disk Galaxies. *ApJ*, 435:22.
- Kerber, F., Rose, T., Chacón, A., Cuevas, O., Czekala, H., Hanuschik, R., et al. (2012). A water vapour monitor at Paranal Observatory. In McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 84463N.
- Kessler, R., Brout, D., D’Andrea, C. B., Davis, T. M., Hinton, S. R., Kim, A. G., et al. (2019). First cosmology results using Type Ia supernova from the Dark Energy Survey: simulations to correct supernova distance biases. *MNRAS*, 485(1):1171–1187.
- Kittel, C. (2005). *Introduction to Solid State Physics*. Wiley, New York, 8th edition.

- Kolmogorov, A. N. (1941). Dissipation of Energy in Locally Isotropic Turbulence. *Akademiia Nauk SSSR Doklady*, 32:16.
- Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., et al. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192(2):18.
- Kurucz, R. L. (2005). ATLAS12, SYNTHE, ATLAS9, WIDTH9, et cetera. *Memorie della Societa Astronomica Italiana Supplementi*, 8:14.
- Lada, C. J. and Lada, E. A. (2003). Embedded Clusters in Molecular Clouds. *ARA& A*, 41:57–115.
- Lallement, R., Bertin, P., Chassefiere, E., and Scott, N. (1993). Correction of spectra for telluric absorption lines with the help of a molecular data bank and high resolution forward modelling: H₂O lines around the sodium doublet at 589.5 nm. *A&A*, 271:734.
- Lang, P., Schreiber, N. M. F., Genzel, R., Wuyts, S., Wisnioski, E., Beifiori, A., et al. (2017). Falling Outer Rotation Curves of Star-forming Galaxies at $0.6 \lesssim z \lesssim 2.6$ Probed with KMOS3D and SINS/zC-SINF. *The Astrophysical Journal*, 840(2):92.
- Law, D. R., Steidel, C. C., Erb, D. K., Larkin, J. E., Pettini, M., Shapley, A. E., and Wright, S. A. (2007). Integral Field Spectroscopy of High-Redshift Star-forming Galaxies with Laser-guided Adaptive Optics: Evidence for Dispersion-dominated Kinematics. *ApJ*, 669(2):929–946.
- Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics*, 2:164–168.
- Lilienfield, J. E. (1926). Method and apparatus for controlling electric currents. U.S, Patent No. 1,745,175.
- Litzén, U. (1970). Improved Experimental Values for Some Hydrogen-like Levels of He I. *Physica Scripta*, 2(3):103–105.

- Livermore, R. C., Jones, T., Richard, J., Bower, R. G., Ellis, R. S., Swinbank, A. M., et al. (2012). Hubble Space Telescope H α imaging of star-forming galaxies at $z \simeq 1\text{--}1.5$: evolution in the size and luminosity of giant H II regions. *MNRAS*, 427(1):688–702.
- Livermore, R. C., Jones, T. A., Richard, J., Bower, R. G., Swinbank, A. M., Yuan, T.-T., et al. (2015). Resolved spectroscopy of gravitationally lensed galaxies: global dynamics and star-forming clumps on ~ 100 pc scales at $1 < z < 4$. *Monthly Notices of the Royal Astronomical Society*, 450(2):1812–1835.
- MacDougall, M. H., Bai, Y., Loose, M., and Beletic, J. W. (2006). Overview of Rockwell Scientific Imaging Technologies. In Beletic, J. E., Beletic, J. W., and Amico, P., editors, *Scientific detectors for astronomy 2005*, pages 403–410, Dordrecht. Springer Netherlands.
- Madau, P. and Dickinson, M. (2014). Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics*, 52(1):415–486.
- Markwardt, C. B. (2009). Non-linear Least-squares Fitting in IDL with MPFIT. In Bohlender, D. A., Durand, D., and Dowler, P., editors, *Astronomical Data Analysis Software and Systems XVIII*, volume 411 of *Astronomical Society of the Pacific Conference Series*, page 251.
- Marquardt, D. W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441.
- Martinsson, T. P. K., Verheijen, M. A. W., Westfall, K. B., Bershadsky, M. A., Andersen, D. R., and Swaters, R. A. (2013). The DiskMass Survey. VII. The distribution of luminous and dark matter in spiral galaxies. *A&A*, 557:A131.
- Mathis, J. S., Rumpl, W., and Nordsieck, K. H. (1977). The size distribution of interstellar grains. *ApJ*, 217:425–433.

- McKee, C. F. and Ostriker, E. C. (2007). Theory of star formation. *Annual Review of Astronomy and Astrophysics*, 45(1):565–687.
- McLeod, B. A. and Smith, R. (2016). Mitigation of H2RG persistence with image illumination. In Holland, A. D. and Beletic, J., editors, *High Energy, Optical, and Infrared Detectors for Astronomy VII*, volume 9915, pages 159 – 165. International Society for Optics and Photonics, SPIE.
- Meadows, V. S. and Crisp, D. (1996). Ground-based near-infrared observations of the Venus nightside: The thermal structure and water abundance near the surface. *Journal of Geophysical Research*, 101(E2):4595–4622.
- Menduiña Fernández, A. (2021). *Measuring and calibrating non-common path aberrations in adaptive optics assisted image-slicer based spectrographs*. PhD thesis, University of Oxford.
- Merritt, D., Graham, A. W., Moore, B., Diemand, J., and Terzić, B. (2006). Empirical Models for Dark Matter Halos. I. Nonparametric Construction of Density Profiles and Comparison with Parametric Models. *The Astronomical Journal*, 132(6):2685–2700.
- Meyer, M. R., Edwards, S., Hinkle, K. H., and Strom, S. E. (1998). Near-Infrared Classification Spectroscopy: H-Band Spectra of Fundamental MK Standards. *ApJ*, 508(1):397–409.
- Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In Watson, G. A., editor, *Numerical Analysis*, pages 105–116, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Mortazavi, S. A. and Lotz, J. M. (2019). Dynamics and shocks from H α emission of nearby galaxy mergers. *MNRAS*, 487(2):1551–1569.
- Moseley, S. H., Arendt, R. G., Fixsen, D. J., Lindler, D., Loose, M., and Rauscher, B. J. (2010). Reducing the read noise of H2RG detector arrays: eliminating correlated noise

- with efficient use of reference signals. In Holland, A. D. and Dorn, D. A., editors, *High Energy, Optical, and Infrared Detectors for Astronomy IV*, volume 7742 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 77421B.
- Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996). The Structure of Cold Dark Matter Halos. *The Astrophysical Journal*, 462:563.
- Neeser, M. (2021). Avoiding, Controlling, and Correcting Detector Persistence in ESO Operations. ESO Presentation.
- Neichel, B., Fusco, T., and Conan, J.-M. (2008). Tomographic reconstruction for wide-field adaptive optics systems: Fourier domain analysis and fundamental limitations. *Journal of the Optical Society of America A*, 26(1):219.
- Neichel, B., Fusco, T., Sauvage, J. F., Correia, C., Dohlen, K., El-Hadi, K., et al. (2016). The adaptive optics modes for HARMONI: from Classical to Laser Assisted Tomographic AO. In Marchetti, E., Close, L. M., and Véran, J.-P., editors, *Adaptive Optics Systems V*, volume 9909 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 990909.
- Noll, R. J. (1976). Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America (1917-1983)*, 66:207–211.
- Noll, S., Kausch, W., Barden, M., Jones, A. M., Szyszka, C., Kimeswenger, S., and Vinther, J. (2012). An atmospheric radiation model for Cerro Paranal. I. The optical spectral range. *A&A*, 543:A92.
- Nyquist, H. (1928). Thermal Agitation of Electric Charge in Conductors. *Physical Review*, 32(1):110–113.
- Osterbrock, D. E. and Ferland, G. J. (2006). *Astrophysics of gaseous nebulae and active galactic nuclei*. University Science Books, California.

- Peebles, M. S., Werk, J. K., Tumlinson, J., Oppenheimer, B. D., Prochaska, J. X., Katz, N., and Weinberg, D. H. (2014). A budget and accounting of metals at $z \sim 0$: Results from the COS-Halos survey. *The Astrophysical Journal*, 786(1):54.
- Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., et al. (2020). Planck 2018 results. VI. Cosmological parameters. *A&A*, 641:A6.
- Powell, L. C., Slyz, A., and Devriendt, J. (2011). The impact of supernova-driven winds on stream-fed protogalaxies. *MNRAS*, 414(4):3671–3689.
- Price, S. H., Kriek, M., Shapley, A. E., Reddy, N. A., Freeman, W. R., Coil, A. L., et al. (2016). The MOSDEF survey: Dynamical and baryonic masses and kinematic structures of star-forming galaxies at $1.4 \leq z \leq 2.6$. *The Astrophysical Journal*, 819(1):80.
- Ragazzoni, R. (1996). Pupil plane wavefront sensing with an oscillating prism. *Journal of Modern Optics*, 43(2):289–293.
- Rasera, Y. and Teyssier, R. (2006). The history of the baryon budget. Cosmic logistics in a hierarchical universe. *A&A*, 445(1):1–27.
- Rauscher, B. J. (2015). Teledyne H1RG, H2RG, and H4RG Noise Generator. *PASP*, 127(957):1144.
- Rauscher, B. J., Arendt, R. G., Fixsen, D. J., Lander, M., Lindler, D., Loose, M., et al. (2012). Reducing the read noise of HAWAII-2RG detector systems with improved reference sampling and subtraction (IRS²). In Holland, A. D. and Beletic, J. W., editors, *High Energy, Optical, and Infrared Detectors for Astronomy V*, volume 8453 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 84531F.
- Ren, D. and Allington-Smith, J. (2002). On the Application of Integral Field Unit Design Theory for Imaging Spectroscopy. *Publications of the Astronomical Society of the Pacific*, 114(798):866–878.

- Richardson, M. L. A., Routledge, L., Thatte, N., Tecza, M., Houghton, R. C. W., Pereira-Santaella, M., and Rigopoulou, D. (2020). ramses2hsim: RAMSES output to 3D data cube for HSIM.
- Richardson, M. L. A., Routledge, L., Thatte, N., Tecza, M., Houghton, R. C. W., Pereira-Santaella, M., and Rigopoulou, D. (2020). Simulating gas kinematic studies of high-redshift galaxies with the HARMONI integral field spectrograph. *Monthly Notices of the Royal Astronomical Society*, 498(2):1891–1904.
- Rieke, G. (2002). *Detection of Light: From the Ultraviolet to the Submillimeter*. Cambridge University Press, 2 edition.
- Rieke, G. (2007). Infrared Detector Arrays for Astronomy. *Annual Review of Astronomy and Astrophysics*, 45:77–115.
- Roberts-Borsani, G. W., Saintonge, A., Masters, K. L., and Stark, D. V. (2020). Outflows in star-forming galaxies: Stacking analyses of resolved winds and the relation to their hosts’ properties. *MNRAS*, 493(3):3081–3097.
- Roddier, F. (1981). The effects of atmospheric turbulence in optical astronomy. *Progress in Optics*, 19:281–376.
- Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., et al. (2009). The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(9-10):533–572.
- Routledge, L., Pereira Santaella, M., Thatte, N., Fusco, T., and Neichel, B. (2020). Improving the ELT/HARMONI Science Simulation Pipeline, HSIM. In Pizzo, R., Deul, E. R., Mol, J. D., de Plaa, J., and Verkouter, H., editors, *Astronomical Society of the Pacific Conference Series*, volume 527 of *Astronomical Society of the Pacific Conference Series*, page 607.

- Rubin, V. C., Ford, W. K., J., and Thonnard, N. (1978). Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa $\rightarrow$ Sc. *ApJL*, 225:L107–L111.
- Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *ApJ*, 121:161.
- Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., van de Ven, G., Vílchez, J. M., Wisotzki, L., et al. (2012). CALIFA, the Calar Alto Legacy Integral Field Area survey. I. Survey presentation. *A&A*, 538:A8.
- Schmidt, M. (1959). The Rate of Star Formation. *ApJ*, 129:243.
- Schneising, O., Buchwitz, M., Burrows, J. P., Bovensmann, H., Bergamaschi, P., and Peters, W. (2009). Three years of greenhouse gas column-averaged dry air mole fractions retrieved from satellite – Part 2: Methane. *Atmospheric Chemistry and Physics*, 9(2):443–465.
- Scott, N. A. (1974). A direct method of computation of the transmission function of an inhomogeneous gaseous medium - I: Description of the method. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 14(8):691–704.
- Seifahrt, A., Käufl, H. U., Zängl, G., Bean, J. L., Richter, M. J., and Siebenmorgen, R. (2010). Synthesising, using, and correcting for telluric features in high-resolution astronomical spectra . A near-infrared case study using CRIRES. *A&A*, 524:A11.
- Sellwood, J. A. and Spekkens, K. (2015). DiskFit: a code to fit simple non-axisymmetric galaxy models either to photometric images or to kinematic maps. *arXiv e-prints*, page arXiv:1509.07120.
- Sellwood, J. A., Spekkens, K., and Eckel, C. S. (2021). Uncertainties in galaxy rotation curves. *MNRAS*, 502(3):3843–3854.
- Sharples, R., Bender, R., Agudo Berbel, A., Bezawada, N., Castillo, R., Cirasuolo, M.,

- et al. (2013). First Light for the KMOS Multi-Object Integral-Field Spectrometer. *The Messenger*, 151:21–23.
- Shatokhina, I., Hutterer, V., and Ramlau, R. (2020). Review on methods for wavefront reconstruction from pyramid wavefront sensor data. *Journal of Astronomical Telescopes, Instruments, and Systems*, 6(1):1 – 39.
- Smette, A. (2020). Molecfit: a general tool for telluric absorption correction. ESO Presentation.
- Smette, A., Sana, H., Noll, S., Horst, H., Kausch, W., Kimeswenger, S., et al. (2015). Molecfit: A general tool for telluric absorption correction. I. Method and application to ESO instruments. *A&A*, 576:A77.
- Smith, R. M., Zavodny, M., Rahmer, G., and Bonati, M. (2008). A theory for image persistence in HgCdTe photodiodes. In Dorn, D. A. and Holland, A. D., editors, *High Energy, Optical, and Infrared Detectors for Astronomy III*, volume 7021 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 70210J.
- Smith, R. M., Zavodny, M., Rahmer, G., and Bonati, M. (2008). Calibration of image persistence in HgCdTe photodiodes. In Dorn, D. A. and Holland, A. D., editors, *High Energy, Optical, and Infrared Detectors for Astronomy III*, volume 7021, pages 192 – 203. International Society for Optics and Photonics, SPIE.
- Sobral, D., Best, P. N., Matsuda, Y., Smail, I., Geach, J. E., and Cirasuolo, M. (2012). Star formation at $z=1.47$ from HiZELS: an $H\alpha$ + $[O\ II]$ double-blind study. *MNRAS*, 420(3):1926–1945.
- Spekkens, K. and Sellwood, J. A. (2007). Modeling Noncircular Motions in Disk Galaxies: Application to NGC 2976. *ApJ*, 664(1):204–214.
- Tatarski, V. I., Silverman, R. A., and Chako, N. (1961). Wave Propagation in a Turbulent Medium. *Physics Today*, 14(12):46.

- Teyssier, R. (2002). Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. *A&A*, 385:337–364.
- Thatte, N., Tecza, M., Clarke, F., Davies, R. L., Remillieux, A., Bacon, R., et al. (2010). HARMONI: a single-field wide-band integral-field spectrograph for the European ELT. In McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735, pages 1000 – 1010. International Society for Optics and Photonics, SPIE.
- Thatte, N. A., Bryson, I., Clarke, F., Ferraro-Wood, V., Fusco, T., Mignant, D. L., et al. (2020). HARMONI: first light spectroscopy for the ELT: instrument final design and quantitative performance predictions. In Evans, C. J., Bryant, J. J., and Motohara, K., editors, *Ground-based and Airborne Instrumentation for Astronomy VIII*, volume 11447, pages 415 – 425. International Society for Optics and Photonics, SPIE.
- Thoning, K. W., Tans, P. P., and Komhyr, W. D. (1989). Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974-1985. *Journal of Geophysical Research*, 94(D6):8549–8565.
- Toledo, D., Irwin, P. G. J., Rannou, P., Fletcher, L. N., Teanby, N. A., Wong, M. H., and Orton, G. S. (2020). Constraints on Neptune’s haze structure and formation from VLT observations in the H-band. *ICARUS*, 350:113808.
- Toledo, D., Irwin, P. G. J., Rannou, P., Teanby, N. A., Simon, A. A., Wong, M. H., and Orton, G. S. (2019). Constraints on Uranus’s haze structure, formation and transport. *ICARUS*, 333:1–11.
- Tulloch, S. (2018). Persistence Characterisation of Teledyne H2RG detectors. *arXiv e-prints*, page arXiv:1807.05217.
- Tulloch, S., George, E., and ESO Detector Systems Group (2019). Predictive model of persistence in H2RG detectors. *Journal of Astronomical Telescopes, Instruments, and Systems*, 5:036004.

- Übler, H., Schreiber, N. M. F., Genzel, R., Wisnioski, E., Wuyts, S., Lang, P., et al. (2017). The Evolution of the Tully–Fisher Relation between $z \sim 2.3$ and $z \sim 0.9$ with KMOS^{3D}. *The Astrophysical Journal*, 842(2):121.
- Uttenthaler, S., Pontoppidan, K. M., Seifahrt, A., Kendrew, S., Blommaert, J. A. D. L., Pantin, E. J., et al. (2010). Correcting METIS spectra for telluric absorption to maximize spectral fidelity. In McLean, I. S., Ramsay, S. K., and Takami, H., editors, *Ground-based and Airborne Instrumentation for Astronomy III*, volume 7735 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 77357A.
- Vanderriest, C. and Lemonnier, J. P. (1988). Silfid - a Versatile Fiber-Optics Spectrograph for Faint Objects. In *Instrumentation for Ground-Based Optical Astronomy*, page 304.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.
- Walborn, N. R. and Fitzpatrick, E. L. (1990). Contemporary optical spectral classification of the OB stars - a digital atlas. *Publications of the Astronomical Society of the Pacific*, 102:379.
- Walsh, J. R. and Roy, J. R. (1990). Area Spectroscopy and Correction for Differential Atmospheric Refraction. In *European Southern Observatory Conference and Workshop Proceedings*, volume 34 of *European Southern Observatory Conference and Workshop Proceedings*, page 95.
- Wechsler, R. H. and Tinker, J. L. (2018). The Connection Between Galaxies and Their Dark Matter Halos. *Annual Review of Astronomy and Astrophysics*, 56(1):435–487.
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. (1996). 3D: The next generation near-infrared imaging spectrometer. *AAS*, 119:531–546.

- Welford, B. P. (1962). Note on a Method for Calculating Corrected Sums of Squares and Products. *Technometrics*, 4(3):419–420.
- Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., Wuyts, E., Bandara, K., Wilman, D., et al. (2015). The KMOS^{3D} Survey: Design, First Results, and the Evolution of Galaxy Kinematics from $0.7 \leq z \leq 2.7$. *ApJ*, 799(2):209.
- Wright, E. L. (2006). A Cosmology Calculator for the World Wide Web. *PASP*, 118(850):1711–1715.
- Wuyts, S., Förster Schreiber, N. M., Wisnioski, E., Genzel, R., Burkert, A., Bandara, K., et al. (2016). KMOS3D: Dynamical Constraints on the Mass Budget in Early Star-forming Disks. *ApJ*, 831(2):149.
- Zieleniewski, S., Thatte, N., Kendrew, S., Houghton, R. C. W., Swinbank, A. M., Tecza, M., et al. (2015). HSIM: a simulation pipeline for the HARMONI integral field spectrograph on the European ELT. *Monthly Notices of the Royal Astronomical Society*, 453(4):3754–3765.
- Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127.