

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/156469>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



Hunting for the Electro-Magnetic Counterpart To Gravitational Waves

by

Ry Cutter

Thesis

Submitted to the University of Warwick

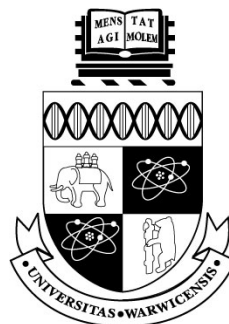
for the degree of

Doctor of Philosophy

Astronomy and Astrophysics

Submitted November 2020

Revised April 2021



*To my wonderful family Sherie, Philip, Molly, and Jake
Without whom, I'd be nothing more than a
Boring astronomer*

*“ We find ourselves in a bewildering world. We want to make sense of what we see
around us and to ask; what is the nature of the universe?”*

— Stephen Hawking, *A Brief History of Time*

Contents

List of Tables	v
List of Figures	vii
Declarations	xvi
Acknowledgements	xvii
Abstract	xviii
Regularly Used Abbreviations	xix
Chapter 1 Introduction	1
1.1 A Brief History	1
1.1.1 The Beginning of Multi-Messenger Astronomy	1
1.1.2 General Relativity and Gravitational Waves	3
1.1.3 Compact Objects as Gravitational Wave Progenitors	6
1.1.4 Initial Observations of Gravitational Waves	12
1.1.5 Gravitational Wave Detectors	13
1.1.6 Localising Gravitational-Wave Signals	18
1.1.7 Making a Counterpart	19
1.2 Science with Multi-Messenger Astronomy	21
1.3 Summary of Introduction	26
1.4 Thesis Content	26
Chapter 2 Methodology	28
2.1 The Hunt Begins - Optical Surveys	28
2.2 The Gravitational-Wave Optical Transient Observer	31
2.2.1 The Challenge	37
2.2.2 Search Strategy	37

2.2.3	Multi-waveband and Spectroscopic follow-up	40
2.3	CCD Astronomy	42
2.3.1	Standard Astronomical Practices	42
2.3.2	Making Sense of the Data	49
2.3.3	Hunting for a Gravitational Wave Counterpart	54
2.4	Summary of Methodology	55
Chapter 3	Getting Focused	56
3.1	Background	57
3.2	A New Approach to Obtaining Focus	60
3.2.1	General Requirements	60
3.2.2	Accounting for Telescope Personalities	61
3.2.3	Exploiting the Astigmatism	67
3.2.4	Initial Performance	68
3.3	An Automatic Implementation	73
3.3.1	Threading the proverbial needle	73
3.3.2	Time Spent Getting Focus is Not Time Wasted	74
3.4	Script Performance	79
3.4.1	The Current Method	79
3.4.2	Future Improvements	83
3.5	Conclusions	83
Chapter 4	Building Reference Images	86
4.1	Your Subtraction is Only as Good as Your Reference	86
4.2	Image Registration	87
4.2.1	World Co-ordinate System based methods	87
4.2.2	Affine Transforms	89
4.2.3	The Problem With Wide Fields	89
4.2.4	Spali2 - The Affine Transform with a Spline	90
4.2.5	From the Line-up, Which is Best?	91
4.2.6	Flux Conservation with Spali2	94
4.2.7	Image Registration Discussion	97
4.2.8	Image Registration Summary	97
4.3	Image Co-Addition	98
4.3.1	Proper Co-Addition	99
4.3.2	Comparing different Stacking methods	100
4.3.3	Stacking GOTO images	103
4.4	Chapter Conclusion	105

Chapter 5	Image Subtraction	107
5.1	What is Image Subtraction?	107
5.1.1	Introducing ZOGY	110
5.1.2	ZOGY in Parallel	111
5.2	Modelling the Point Spread Function	114
5.2.1	A Simple Bivariate Gaussian	115
5.2.2	Zernike Moments as PSF models	120
5.2.3	Modified PSFex	123
5.3	Comparing Image Subtraction Software	125
5.3.1	Creating Test Data	125
5.3.2	Transient tests	132
5.3.3	Speed tests	137
5.3.4	Flux Conservation	140
5.4	Using ZiP on real data	140
5.4.1	NGTS Flare Pipeline	140
5.4.2	Monitoring M31	142
5.4.3	Finding HOPE, GOTO's Favourite Transient	143
5.5	Chapter Conclusion	146
Chapter 6	Filtering out False Positives with Machine Learning	147
6.1	Introduction	147
6.2	Building a Real-Bogus Classifier	148
6.2.1	General Machine Learners used in Astronomy	148
6.2.2	Artificial Neural Networks	152
6.2.3	A Novel DNN	155
6.3	How to Train your Computer	157
6.3.1	Testing the DNN	162
6.4	Machine Learning Summary	169
6.5	A Future Classifier	170
Chapter 7	The First Follow-Up Campaign with GOTO	171
7.1	Introduction	171
7.2	GOTO follow-up	172
7.2.1	Collecting Data	172
7.3	Follow-up Results	175
7.3.1	Delving into the Samples	182
7.3.2	Variables that were vetted by a human	185
7.4	Discussion	186

7.4.1	GOTO is Doing its Job! (Gravitational-Wave Follow-Up) . .	186
7.4.2	Finding Transients	187
7.4.3	Variables and Other Science with GOTO	188
7.5	Conclusion	190
Chapter 8	What Happens Next?	191
8.1	Thesis Summary	191
8.2	The Future of GOTO	192
8.2.1	GOTO South	192
8.2.2	Transients, they don't stop coming	193
8.2.3	Variable Stars, they never stop coming	194
8.2.4	Getting Ready for O4, hit the ground running	194
Appendix A	Derivation of the ZOGY Equations	195
Appendix B	Derivation of Proper Co-Addition	199
Bibliography		201

List of Tables

3.1	Focus trial performance metrics. This shows that the most confident (and closest) method is the Gaussian combination of all three methods from figure 3.10. See figure 3.11 for more details.	71
4.1	Image registration tests, supplementary table to figures 4.2, 4.3, and 4.4. The table shows that each method is suitable for well behaved fields, however WCS methods perform notably worse in GOTO images.	91
4.2	PSF statistics for different alignment methods. Highlighting that proper co-addition conserves the PSF when stacking images.	102
5.1	Qualitative checklist of Software performance. Showing ZiP meets all the basic requirements for real time image subtraction.	114
5.2	Image subtraction test statistics. Contamination and Completeness are defined in equations 5.20 and 5.21 respectively. $Comp_{lim}$ is the completeness at the limiting magnitude of the telescope.	132
5.3	Mean subtraction time taken over 100 subtractions.	137
5.4	Mean subtraction time taken over 100 ZiP subtractions for the different slice patterns (see figure 5.20). This shows that the different slice patterns can affect the subtraction quality and time.	138
5.5	Supplementary table for the NGTS flare pipeline (figure 5.22). . . .	142
6.1	Machine Learning Performance for different methods. This reiterates earlier analysis, showing the GNB and KNN methods are weaker. It also highlights the DNN and random forest have similar performances.	167

7.1	Complete Collection of GOTO GW follow-up metrics. Event is the automatic event name and the GCN is where the announcement was first published. Region describes how large the 90% credibility localisation region is in square degrees. δt_{alert} and δt_{event} shows the response time of GOTO getting on sky after the alert was sent and when event went off, respectively. A is the area covered by GOTO in the search. PA is the probability covered over the search. PA_v is the probability covered by the region exclusively in GOTO's visible range. PA_d is the probability covered by images that have difference images. MAP describes the tool used to build the probability localisation. Time Obs, is the amount of time dedicated to following-up the event. LVC run describes which half of O3 the detection was made. m_{lim} is the average limiting magnitude of the search and err_{lim} is the error on this average. This table shows that GOTO is covering most of the probability region during its searches to a depth that can find transients similar to GW170817 (Gompertz et al., 2020b)	178
7.2	Results of the transient extraction. Showing that of 74317748 only 49 passed my pipeline filter, and 17 passed human inspection.	179

List of Figures

1.1	The two Gravitational-Wave polarisations in the xy plane and the phases of their oscillations. These represent the two ways a detector will stretch and contract when exposed to a gravitational-wave. . . .	5
1.2	Binary evolution to create a massive compact binary (Postnov and Yungelson, 2014). This is only one of many binary channels; however, the visual acts as a useful representation of how close-in compact binaries can form.	11
1.3	Comparing the orbital decay of an observed neutron star binary B1913+16 to that predicted by GR (Weisberg and Taylor, 2005). This shows how the observations of the binary orbital period matches the predicted evolution from GR almost exactly. Acting as indirect evidence of Gravitational-Radiation.	12
1.4	Visualisation of a squeezed LIGO signal. The uncertainty of the phase is squeezed below the quantum noise limit, but is traded for a higher uncertainty in the amplitude.	15
1.5	The initial LIGO-O2 noise characterisation (Fritschel, 2003). The central black line represents the total noise limit, which is the sum of all the other noise curves (excluding the initial LIGO curve at the top).	16
1.6	First Gravitational-Wave signal (Abbott et al., 2016b). Note how the signal timescale is a fraction of a second, the chirp (merger) occurs at ~ 0.4 seconds.	17
1.7	The initial LIGO map for GW150914 from Singer et al. (2014). The total localisation region spans 630 sq. degrees.	19
1.8	Cartoon of a kilonova from the axisymmetric model. The black centre shows the hypermassive neutron star engine, the green indicated the GRB jets, the red (lanthanide rich) material is shown in red, similarly for the blue (lanthanide poor) component, and finally the accretion disk is shown in purple.	23

1.9	Model kilonova lightcurve from Kasen and Barnes (2019). Highlighting the expected peak luminosity and decay timescales are on the order of days.	26
2.1	GOTO's first light, the Hercules Globular Cluster (June 11th 2017) .	30
2.2	GOTO using its 8 telescope configuration	32
2.3	GOTO's expected field of view compared to other telescopes (Steeghs et al., in prep). GOTO's final configuration will have 32 telescopes, 16 on both hemisphere's, giving it an instantaneous field of view of 80 square degrees. Making GOTO one of the widest survey instruments in astronomy.	33
2.4	GOTO's footprint for GOTO-4 (left) and GOTO-8 (right)	34
2.5	GOTO tiling pattern for GOTO-4 (left) and GOTO-8 (right)	34
2.6	Tiling pattern when hunting for a Fermi GRB trigger (Fermi 626581870). Each square on the map represents one footprint from figure 2.4. . .	35
2.7	Filter response function over given wavelengths from Steeghs et al. (in prep)	36
2.8	Showing GOTO's boom components for axial movement.	36
2.9	The spread of source shapes across a single GOTO-prototype exposure. Each frame is equivalent to the position of where the source is in the image. Highlighting the optical collimation errors.	38
2.10	Nightly report, highlighting the GOTO tiles and probability covered for S190426c ((alias?)).	39
2.11	Example of a GOTO master bias image. This is made from the stack of multiple zero second exposures and is used to account for read-out noise.	44
2.12	Example of a GOTO master dark image. This is made by stacking multiple exposures taken with the lens closed to model the thermal noise of the image.	45
2.13	Example of a GOTO flat field. This is made by stacking multiple flat fields. Flat fields are the normalised response function of each CCD pixel, so given a field of uniform brightness the flat field will correct all the pixels so the exposure is uniform.	46
2.14	An example of background estimation on a GOTO field	51
2.15	Caption for LOF	53

3.1	Difference in AutoFocus optimal solution, to best focus found manually across multiple nights. Highlighting that optimal focus was routinely missed with the original focus algorithm.	56
3.2	Example of FWHM as a function of focuser position. Optimal focus corresponds to the minimum in FWHM, which is shown to be close to the point where sources appear point-like.	57
3.3	Example of using two linear fits to obtain optimal focus with GOTO data (from equation 3.1). The blue lines show the linear fits, the orange line is their intercept and the location of the predicted optimal focus (described by equation 3.1).	59
3.4	Focus evolution over the course of continuous observation (different nights are overlayed on top of one another). There is a shared response across all of the UTs ~ 5 hours, where all the HFD increases. This means there is an environmental influence acting on all the scopes.	62
3.5	Focus response to temperature. The red points indicate the lowest 10% HFD points in that temperature bin. The red points show a flat line, this indicates temperature does not have a significant affect on HFD	63
3.6	Focus response to wind speeds. The red points indicate the lowest 10% HFD points in that wind speed bin. The red points show a step at $\sim 20\text{km/h}$, this indicates wind speed has an affect on the HFD . .	64
3.7	Focus response to wind speed and airmass. The HFD has no real dependence on airmass, dismissing the concern that telescope position may be responsible for HFD variation. This plot also reaffirms the correlation to HFD and wind speed.	66
3.8	Visualisation of Elongation and source rotation	67
3.9	Relationship of Theta and Elongation as a function of focuser position. The inverse of the elongation traces a negative parabola, with the peak at optimal focus. Source rotation demonstrating a step function, where that step occurs is also considered optimal focus. . .	68

3.10	Finding focus three different ways. The left most plot shows finding focus using source rotation, where the red line shows one of the sigmoid fits. The middle plot shows elongation finding focus, the red line here is the negative parabola fit. The plot on the right shows optimal focus found by the intersection of two fits. The green and orange line represent the linear fits either side of the uncertainty line (indicated by the grey bar). Optimal focus on each plot is indicated by the vertical blue line and the uncertainty by the transparent blue region.	70
3.11	Using Gaussian probability combination to improve optimal focus estimation. The combination of all the methods (shown as the purple dotted line) has the highest confidence and has the closest prediction to the focus found manually (the vertical black line).	72
3.12	Flowchart of the automatic focus script	75
3.13	Flowchart of the automatic focus script, where exposure time is shorter than processing time	76
3.14	HFD using different exposure times. Note: the outliers, particularly in 5s and 15s, are due to the regular sampling in these exposure times. Meaning there are more data points to sample. This shows that focus can be determined with one second exposure times, meaning the longer exposures originally used when finding focus are unnecessary.	77
3.15	Theoretical time taken to obtain focus if the exposure time (15 seconds) is greater than the processing time (7.5 seconds). The dotted blue line shows the original time it took to get focus, the orange dashed line shows the time the new algorithm takes to find focus. The green line shows time saved between the methods and the solid purple line shows the number of additional exposures that can be taken with the amount of time saved (assuming saved time has no additional benefit).	78
3.16	Theoretical time taken to obtain focus if the exposure time (2 seconds) is less than the processing time (7.5 seconds). The colours are the same from figure 3.16.	79
3.17	Finding focus with the HFD in practice. The colour scheme is the same from the right most plot in figure 3.10. This is shown to be a reliable method when both sides of the uncertainty line are properly sampled. However, UTs 5,6, and 8 failed to find focus due to under sampling of either side of focus.	80

3.18	Finding focus using elongation in practice. The colour scheme is the same from the centre plot in figure 3.10. Again, UTs 5 and 6 are under sampled. However, a prediction was able to made for UT8. . .	81
3.19	Find focus using source rotation in practice. The colour scheme is the same from the left most plot in figure 3.10. The orange line shows the most distant fit from the red fit. The method is unstable which is shown by UT2, 7, and 8.	82
4.1	The result of subtraction with poor alignment. Note the residuals leave a shadow imprint leaving behind 'butterfly' artefacts.	88
4.2	Physical time to compute each alignment on a GOTO field. (It should be noted Alipy, and Spali2 use 2 cores while SWarp and reproject use all the cores available.) Each line has a linear relation, meaning the transform time from each method can be modelled easily using the mean of each method. SWarp and reproject do not account for the time it takes to solve the WCS, this can have an additional 10-20 seconds computing time to these solutions.	92
4.3	Histograms showing separation of sources in pixels from the original image in x and y directions. A : WCS based alignment of a small field, B : Alipy alignment of a small field, C : Spali2 alignment of a small field	93
4.4	Histograms showing separation of sources in pixels from the original image in x and y directions. D : WCS based alignment of a wide field, E : Alipy alignment of a wide field, F : Spali2 alignment of a wide field	93
4.5	Flux conservation of the different alignment methods as a function of flux. SWarp (purple) has the most inconsistent performance. Reproject (green) and Spali2 (blue) have similar performances with Reproject showing a little more scatter at higher fluxes.	95
4.6	Flux conservation of spali2 as a function of position for ten different alignments. There is no positional dependence on the Spali2 transform quality. The edges of the original image are often out of frame in reference. This means the transform removes them from the aligned frame, this means sources in the original image have no comparison in the transformed frame, hence the large edge uncertainties.	96
4.7	Initial comparison of how each stacking method handles artefacts. All the co-addition methods work well, apart from the mean which shows how some statistics are not suitable for co-addition.	100

4.8	Number of 5σ detections for a given number of NGTS stacks for each method. This is compared to a scaled \sqrt{N} graph, showing the depth increase matches what is predicted by equation 4.1.	101
4.9	Time taken to stack NGTS images, using multiple cores. The Swarp and Median combine methods scale linearly with number of images and are much faster than proper co-addition which is significantly slower, but its speed is roughly independent of the number of images used.	102
4.10	Limiting magnitude for different Stacking Strategies of GOTO data. The dark stack gets the deepest reference. Using a random assortment of fields is not beneficial and shows that some curation is necessary to optimise ones reference image.	104
4.11	Comparison of PSF conservation Left: Median Stack. Right: Proper Co-Addition.	105
5.1	HOTPANTS subtraction with complex PSF. The convoluted template should match the target PSF. This is not the case, resulting in a prominent residual being left in the subtraction image.	109
5.2	Flowchart for the ZiP proper image subtraction tool.	113
5.3	The Full-Width Half Maximum as a function of X and Y (the CCD axes). This highlights that for telescopes like GOTO, the FWHM can vary by a factor 3 across the image.	116
5.4	Visual of the Bivariate Gaussian method for recreating PSFs observed in real GOTO data. The method can recreate simple PSF geometries quickly, but cannot solve non-Gaussian shape functions.	118
5.5	GOTO image of M51 (The Whirlpool galaxy). M51 acts as a contaminant when calculating the FWHM across an image which will break the Bivariate Gaussian method.	119
5.6	Using Zernike moments to reproduce arbitrarily complex 2D-surfaces	120
5.7	PSF estimation of different PSF complexities using Zernike moments. This figure highlights the power Zernike moments have when recreating PSF shapes.	121
5.8	Time taken for different degrees of Zernike polynomial to model a PSF. This shows that for smaller thumbnails (well contained PSFs), a model PSF can made on sub-minute timescales. However, for large complicated PSFs, the time to calculate a model is too long for real time image subtraction	122

5.9	PSFex point source selection (Bertin, 2013). Any sources found in the red box are considered appropriate for estimating the PSF. . . .	124
5.10	Comparison of each method at recreating a complicated PSF. The Bivariate Gaussian method failed here.	124
5.11	Comparison of the working methods using ZiP subtraction	125
5.12	Example of source injection using LazyPSF . Centre: Injected source, Left: Real Source. This figure shows LazyPSF can inject realistic sources into GOTO data.	127
5.13	Distribution of injected sources based on measured flux distribution	129
5.14	Flat user-defined distribution for source injection	129
5.15	Visualisation of the subtraction methods. Using a galaxy with sources injected using IRAF . Qualitatively, the $ZOGY_D$ image is the lowest quality (i.e. will flag the highest number of false positives). HOTPANTS is the second, but the cleanest subtraction looks to be $ZOGY_{Scorr}$	131
5.16	Injected source recovery as a function of magnitude. The top figure (red) shows HOTPANTS performance, second (green) is the original ZOGY script, and finally the ZiP raw subtraction (orange) and ZiP noise correlated image (blue). As expected the original ZOGY and ZiP have similar transient extraction performances.	133
5.17	Probability of recovery as a function of magnitude. The blue vertical line indicates GOTO's average limiting magnitude.	134
5.18	ZiP image subtraction tests with iPTF data	135
5.19	ZiP image subtraction tests with NGTS data	136
5.20	Slice patterns for table 5.4	137
5.21	Median absolute difference in flux as a function of flux. Calculated from injecting transients, subtracting a reference, and calculating the flux of the injected transient in the subtracted image. Note, the error closely corresponds to the background flux of the reference image. . .	139
5.22	NGTS flare pipeline. The black line represents the flux of the source. Each blue vertical line indicates when a new exposure has been taken, with the number on the right corresponding to the exposure number. The red stars are when alerts are sent for follow-up.	141
5.23	Example of vetting subtraction candidate	143
5.24	Using image subtraction to find and monitor a transient in M31. Showing ZiP can locate faint transients in bright and crowded environments.	144

5.25	GOTO2019hope being subtracted with ZiP.	145
5.26	Lightcurve of GOTO2019hope (AKA: SN 2019pju) a type Ia supernova. Showing GOTO can find early time signals of Supernovae. . .	145
6.1	A basic ANN schematic for a binary classifier. Input layer (red) represents the input parameters, the hidden layer (blue) applies an unknown functions and sends it to the output layer (purple). The output layer is a score between 0 and 1 representing each class in the binary.	153
6.2	DNN architecture for an automatic rapid transient hunter. Using the same colour representation from the ANN, this figure shows how the input layer is sent through multiple hidden layers. Each hidden layer has half the number of nodes before being sent to the output layer. .	156
6.3	Confusion matrix of the GNB, KNN, and Random forest methods. Not including the GNB method, as it appears to be calling a majority of sources real, the KNN and random forest method support the strength of the new feature selection.	158
6.4	Weights of each parameter when determining classification for the Random Forrest method.	159
6.5	Loss over training epochs. Training epoch refers to the number of times the trainer has iterated over the training data. In the first few epochs the curve can be treated as parallel to the loss axis.	160
6.6	Accuracy over training epochs. The peak of this graph selects the machine learner for this training iteration.	160
6.7	DNN Training confusion matrix, highlighting that 98% of transients are appropriately flagged and only 2% of the false positives are incorrectly flagged as real.	161
6.8	Proportion of injected sources (on top of galaxies) correctly flagged at the extraction and classification stages.	162
6.9	Confusion matrix for 50 percent confidence	163
6.10	Confusion matrix for 95 percent confidence	163
6.11	Recovery and classification of injected sources for each UT. The grey bar indicates the number of injected sources. The blue bar is the number of sources extracted from the subtraction step. Orange represents transients classified with >50% confidence. Purple represents transients classified with a confidence of over >95%.	164

6.12	Weights of each parameter when determining classification for the Random Forrest method. This figure uses the training sample from the galaxy only injected training sample.	165
6.13	ROC Curve for different machine learning methods on each UT. Each plot is based on the classification metrics for each UT. The black dashed line represents the ‘worst’ classifier case. The blue line is the DNN classifier, the orange line is the random forest, red is the KNN, and green GNB. There is a strong performance correlation of all the methods with UT, most likely linked with image quality. There is also a performance difference in classifier based on the method used, as seen in table 6.1	168
7.1	Visualisation GOTO’s coverage of S190923y as seen in (Gompertz et al., 2020b). GOTO covered over 720 ^{o2} totalling 58% of the observable probability	174
7.2	Example of misalignment and edge-case transient false positive . . .	175
7.3	First set of transients from GOTO’s GW follow-up from table 7.2. .	180
7.4	Second set of transients from GOTO’s GW follow-up from table 7.2.	181
7.5	RR Lyr star that passed the automatic filter but not a human. . . .	185
7.6	Dwarf Nova that passed the automatic filter but not a human	186
7.7	Example selection of false positives from the GW follow-up	189

Declarations

I declare that, unless otherwise stated, the work presented in this thesis is my own.

The first half of Table 7.1 and Figure 7.1 in Chapter 7 covering the follow-up observations of O3a appears in Gompertz et al. (2020b).

Acknowledgements

There are a lot of people who were necessary for the completion of this thesis, shoulders of giants and all that. Firstly, this work would not have been possible without the encouragement of my supervisor and PI of the colossus that is GOTO, Danny Steeghs. Duncan Galloway, GOTO's Southern node PI, also deserves recognition for his part in the birth of GOTO. The GOTO collaboration has been a support structure and the contribution of each member has been invaluable: Thomas Killestein, Krzysztof Ulaczyk, Joe Lyman, Klass Wiersema, Ben Gompertz, Elizabeth Stanway, Paul Chote, James McCormac, Tom Marsh, Don Pollacco, Richard West, Kendall Ackley, Yik Lun Mong, Evert Rol, Eric Thrane, Martin Dyer, Lydia Makrygianni, Umar Burhanudin, Vik Dhillon, James Mullaney, Stuart Littlefair, Ed Daw, Justyn Maund, Rob Eyles-Ferris, Spencer Tooke, Rhaana Starling, Paul O'Brien, Gavin Ramsay, Christopher Duffy, Micheal Brown, Saran Poshyachinda, Suparerk Aukkaravittayapun, Supachai Awiphan, Puji Irawati, David Mkrtichian, Utane Sawangwit, Rubina Kotak, Teppo Heikkilä, Seppo Mattila, Laura Nuttall, Rene Breton, Mark Kennedy, Daniel Mata-Sánchez, Andrew Levan, and Enric Pallé.

Special thanks should be given to the people who put up with me over the last four years, my office mates and friends: Odette Toloza, Elena Cukanovaite, Tim Cunningham, James Jackman, Matthew Green, Louise Wang, Ashley Chrimes, Matthew Battley, and Christopher Manser.

I am indebted to Boris Gänsicke and Iain Steele for their helpful feedback and insights to improve the thesis; and also very grateful for an enjoyable viva experience.

Finally an enormous thanks to Miriam Hogg, her patience, love, and support have been a beacon of sanity while the world fell apart. Without her light, this thesis would never have been finished. Thank you.

Abstract

The detection of GW170817, the first known binary neutron star merger, ushered in the era of multi-messenger astronomy. The Gravitational-Wave Optical Transient Observer (GOTO) is a new multi-camera instrument designed to cover large gravitational-wave localisation regions quickly, with the aim of identifying gravitational-wave counterparts quickly. This thesis covers the deployment of GOTO and the technical development of an automatic focus script, a real-time image subtraction pipeline, and a machine learner all with the aims of finding and announcing transients in real-time. The methods developed here can be used in other high-cadence optical surveys. The thesis is motivated in the introduction, summarising the history of gravitational-wave astronomy and the importance of finding counterparts. GOTO is introduced properly in the methodology section. Here, I explain how GOTO is built and optimised for rapid transient discovery. From there, I show the development of an automatic focus script that exploits source geometry to quickly achieve focus. The following two chapters detail the the development of a new image-subtraction pipeline, which proves to be faster and better quality than the techniques currently used in the literature. Finally, I conclude this work using GOTO's first Gravitational-Wave follow-up campaign in compliment with the techniques developed in this thesis to find transients coincident with Gravitational-Wave detections. Showing GOTO is indeed capable and primed to find transients associated with gravitational-waves quickly.

Regularly Used Abbreviations

Artificial Neural Nets –	ANN
Binary Black Hole –	BBH
Binary Neutron Star –	BNS
Charged Couple Device –	CCD
Compact Binary Coalescence –	CBC
Deep Neural Nets –	DNN
ElectroMagnetic –	EM
european Laser Interferometer Space Antenna –	eLISA
Full Width Half Maximum –	FWHM
General Relativity –	GR
Gravitational-Wave –	GW
Gravitational-wave Optial Transient Observer –	GOTO
Gamma-Ray Burst –	GRB
Half Flux Diameter –	HFD
Intermediate Mass Black Hole –	IMBH
Laser Interferometer Gravitational-Wave Observatory–	LIGO
LIGO VIRGO Collaboration –	LVC
Minor Planet Checker –	MPC
Neutron Star - Black Hole –	NSBH
Point Spread Function –	PSF
Probability Density Function –	PDF
Super Massive Black Hole –	SMBH
Unit Telescope –	UT
World Coordinate System –	WCS
ZOGY in Parallel –	ZiP

Chapter 1

Introduction

“ Aristotle said a bunch of stuff that was wrong. Galileo and Newton fixed things up. Then Einstein broke everything again. Now, we’ve basically got it all worked out, except for small stuff, big stuff, hot stuff, cold stuff, fast stuff, heavy stuff, dark stuff, turbulence, and the concept of time ”

— Zach Weinersmith, *Science Abridged Beyond the Point Of Usefulness*

1.1 A Brief History

1.1.1 The Beginning of Multi-Messenger Astronomy

I was in the Kjell-Henriksen observatory with six other masters students huddled around a laptop when the biggest physics announcement of the decade was made. Little did I know, this press release was going to be the foundation of my PhD. February 11th 2016 marked the official announcement of the first direct Gravitational Wave detection, GW150914; the first binary black hole merger (Abbott et al., 2016a). This first major event spurred on a slurry of searches, each looking for an electromagnetic signal corresponding to the detection of GW150914 (Evans et al., 2016; Morokuma et al., 2016; Lipunov et al., 2016; Ackermann et al., 2016; Tavani et al., 2016; Herner et al., 2017). Even with all these resources looking for a counterpart, nothing was found. This is, in part, due to the astonishingly wide field of localisation, taking up 600 square degrees in sky. With all the searches combined, roughly 61% of the certainty region was covered (Abbott et al., 2016a). Thus, if there was a counterpart, there is a good chance that it simply was not observed. The other factor is that black hole mergers are not expected to have a luminous, if any, counterpart (Berger, 2014). Multi-messenger astronomy had found its begin-

ning, but it would take another year before the world would see the true benefits of combining classical astronomy with the detection of gravitational waves.

On August 17th 2017, LIGO detected what was later confirmed to be the first neutron star merger. This moment points to a landmark in astronomical history. Almost two seconds later, FERMI and INTEGRAL announced a high confidence detection of a gamma-ray burst. This kind of temporal coincidence had been seen before (Perna et al., 2016), but the localisation of the events aligned as if by design. The position of the source in the sky could be further deduced by the fact VIRGO had a null detection of the event (Abbott et al., 2017b). Within minutes of the announcement, almost every available telescope on Earth and in space was looking for a counterpart to the LIGO discovery. It did not take long for astronomers to start suggesting potential candidates, but one source stood out (Coulter et al., 2017). This particular candidate appeared in the predicted localisation region (Abbott et al., 2017c). Confidence increased as the source was atop a galaxy with a known redshift, and the distance matched that expected from the Gravitational-Wave signal! As the evidence was stacking, more telescopes started to observe this source. It quickly became apparent that this source was like nothing that had been observed prior; but of all the astrophysical models that existed, the spectrum most closely resembled a kilonova, the ejecta expected to accompany a neutron star merger. In fact, the kilonova models matched surprisingly well (Rosswog et al., 2000; Tanvir et al., 2013; Metzger and Fernández, 2014). Theories surrounding the event were made in days, from the stellar environment (Levan et al., 2017; Blanchard et al., 2017) to the original source properties (Metzger et al., 2018). Possibly one of the most publicised facts from this event, is the attribution to kilonovae as the origin to most of the Universes heavy elements (Kasen et al., 2017). Even after a year, the event was still being studied (Lamb et al., 2019). As the only known occurrence of its kind, GW170817 offered a wealth of knowledge to astronomers and astrophysicists and is one of the most studied events in astronomical literature to date (Abbott et al., 2017b,c,a; Alexander et al., 2017; Andreoni et al., 2017; Blanchard et al., 2017; Coulter et al., 2017; Covino et al., 2017; Creminelli and Vernizzi, 2017; Drout et al., 2017; Evans et al., 2017; Ezquiaga and Zumalacárregui, 2017; Kasen et al., 2017; Levan et al., 2017; Savchenko et al., 2017; Abbott et al., 2018a; Alexander et al., 2018; Corsi et al., 2018; D’avanzo et al., 2018; Dobie et al., 2018; Lyman et al., 2018; Nakar and Piran, 2018; Abbott et al., 2019; Lamb et al., 2019; Montana et al., 2019; Hotokezaka et al., 2019; Lü et al., 2019; Piro et al., 2019; Wu and MacFadyen, 2019; Ai et al., 2020; Dhawan et al., 2020; Essick et al., 2020; Hajela et al., 2020a,b; Hamidani et al., 2020, This list is not remotely close to being exhaustive).

The detection of GW170817 marks the true beginning of the multi-messenger astronomy era and frames the motivation for this thesis, but the story of gravitational wave follow-up starts much earlier.

1.1.2 General Relativity and Gravitational Waves

A little over a century ago, Einstein first proposed the theory of General Relativity (GR). The revolutionary concept that describes gravity as a geometric representation of space-time (Einstein, 1915, 1995). Since then GR has been used to model the nature of black holes, predict Mercury’s orbit, and even peer into the moments our Universe began (Hawking, 1988). Until relatively recently, evidence of GR’s final prediction remained undiscovered (Einstein and Rosen, 1937). Before detailing how to detect them, one should first discuss the fundamental theory surrounding curved space-time and the generation of these evasive phenomena, gravitational waves!

GR was developed as a sequel to special relativity. Special relativity introduced the concept of a maximum speed limit to the Universe, the speed of light (Einstein, 1907, 1995). The theory implies the speed of light is the same in all directions and to every observer regardless of their relative motion, with the caveat that all the frames of reference be non-inertial (i.e. non-accelerating reference frames). This is what made the case special. After ten years of development, Einstein (with some help) produced the general theory of relativity. This generalised the special case and presented a solution to the physics involving non-inertial frames. The theory tackles gravity geometrically, igniting the famous phrase *“Spacetime tells matter how to move; matter tells spacetime how to curve”*. The neatest representation of the Einsteins field equations are as follows (assuming the speed of light and the gravitational constant are unity):

$$G_{\mu\nu} = 8\pi T_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \quad (1.1)$$

where $G_{\mu\nu}$ is the Einstein tensor, which is used to express the curvature of space time. It is defined by the Ricci tensor ($R_{\mu\nu}$), a measure of the difference a space-time metric ($g_{\mu\nu}$) has from a Euclidean geometry, and the Ricci Scalar (R), which is the contraction of the Ricci Tensor. Finally $T_{\mu\nu}$ is the stress-energy tensor. Simply put, $T_{\mu\nu}$ expresses the flow in the μ direction of momentum in the ν direction.

The existence of gravitational waves can be derived by adding a small perturbation ($h_{\mu\nu}$) in the flat space-time metric ($\eta_{\mu\nu}$).

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \quad (1.2)$$

Substituting equation 1.2 into 1.1 and solving the Riemann tensor, Ricci tensor and scalar, and then choosing the Lorenz gauge, one can express the modified metric as:

$$\square \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}, \quad (1.3)$$

where \square is the D'Alembertian operator:

$$\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}. \quad (1.4)$$

In a perfect vacuum $G_{\mu\nu} = 0$. This necessarily means $T_{\mu\nu} = 0$ in a vacuum also. Following this logic, if there is a perturbation of a flat space-time in space, a place considered an almost perfect vacuum, equation 1.3 becomes:

$$\square \bar{h}_{\mu\nu} = 0. \quad (1.5)$$

This is a concise way to represent ten wave equations. Applying the transverse traceless gauge (making all the non-zero spatial components transverse to the direction of the wave, while the matrix has a trace equal to 0) removes the remaining gauge freedoms to the above and presents a solution to find the amplitude of the wave equation. This results in two polarisations; plus and cross (see figure 1.1 and equation 1.6). These can be described as transverse plane waves that travel in the z direction at the speed of light and oscillate in the xy plane. Also referred to as Gravitational Waves (GWs).

$$\mathbf{h} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (1.6)$$

The generation of GWs can theoretically have many physical sources (Langlois et al., 2000; Horowitz and Kadau, 2009; Ott, 2009), the one discussed here and currently the only observed physical source of GWs, Compact Binary Coalescence (Abbott et al., 2009, CBC). CBC describes non-spherical masses with asymmetric rotation accelerating around one another, distorting space-time until they coalesce

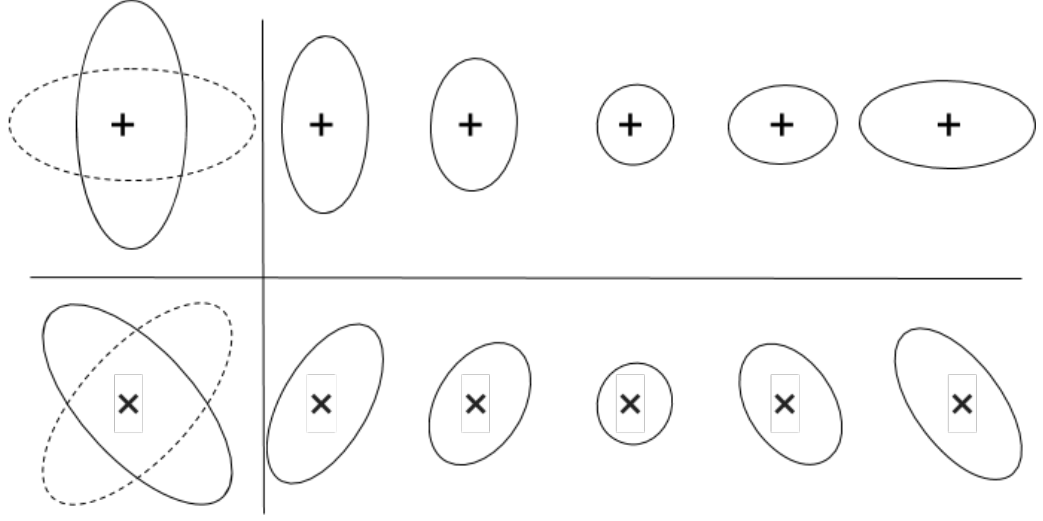


Figure 1.1: The two Gravitational-Wave polarisations in the xy plane and the phases of their oscillations. These represent the two ways a detector will stretch and contract when exposed to a gravitational-wave.

generating GWs. This refers to two compact objects, either black holes, neutron stars, or a combination of the two (and when the detectors are sensitive enough, white dwarfs (Tutukov and Yungelson, 1994)) spiralling in to one another. The decay of their orbit is due to loss of angular momentum from gravitational radiation and is emitted in the form of GWs. As the two bodies get closer their orbital period decreases, increasing the amplitude and frequency of the resulting GWs.

The above theory predicts physical stresses caused by ripples in space-time. The energy to make these ripples is generated through the decay of the emitting system's orbit. Using the formalism in Peters (1964), it is shown that the separation (a) of the two stars evolves:

$$\frac{da}{dt} = -\frac{64}{5} \frac{m_1 m_2 (m_1 + m_2)}{a^3} < 0. \quad (1.7)$$

This has a physical effect of reducing the binary separation, in turn decreasing the orbital period. Eventually the orbit will shrink sufficiently that the two bodies collide in a catastrophic event, a compact merger. One apparent physical outcome from equation 1.7 is the effect of mass on the merger timescales. The heavier the binary system, the faster the merger. Evidence of this is seen in the LIGO signals. Black hole mergers are seen to coalesce within a fraction of a second, whereas the

neutron star inspiral phase was seen in the data for over hundred seconds (Abbott et al., 2016a, 2017b). Moreover, the upper limit of the GW frequency is set by the size of the objects. This means neutron stars have higher GW frequencies than black holes and white dwarfs which are physically much larger. The separation of the binary also plays a role on the merger timescales. For smaller objects like neutron stars and stellar mass black holes, wide separations will result in merger timescales greater than a Hubble time. However, for super massive black holes, their mass allows them to overcome this separation problem. Because of their mass their inspiral phase is short, but it is also predicted that the entire merger of two super massive black holes can occur on mega-year timescales (Khan et al., 2016). The term compact in CBC comes from the nature of the GW progenitors, the dense remains of dead stars. Given their incredible densities (i.e. they are in a strong field regime), compact remnants make for valuable GW laboratories.

1.1.3 Compact Objects as Gravitational Wave Progenitors

Compact objects come in three varieties, from heaviest to lightest: Black Holes, Neutron Stars, and White Dwarfs. All of which share a common theme of being a compact stellar remnant.

Black Holes

Black holes are objects so dense, light cannot escape their gravitational pull. The radius of no escape is referred to as the event horizon. The formation of stellar mass black holes comes from the collapse of the most massive stars. The concept of a gravitational field strong enough to prevent physical escape was first proposed by John Mitchell in 1799, Schaffer (1979) and Montgomery et al. (2009) present an interesting historical account of this. In 1916, a year after Einstein published the theory of GR, Schwarzschild developed the solution to the Einstein-Field Equations (eq 1.1) for a non-rotating spherical body with a given mass. This metric predicted a finite event horizon, named the Schwarzschild radius; the physical significance of which was debated for decades. It was only in 1958 when the Schwarzschild radius was interpreted as a point where matter can only pass in one direction (Finkelstein, 1958). These theoretical objects were coined “black holes” in 1967. Then in 1971, over half a century after GR was first postulated, the first indirect evidence of a black hole was presented (Bolton, 1972; Webster and Murdin, 1972). Cygnus X-1 is a binary system; by assessing the Doppler shifts of the companion, the orbit of the companion and thus the mass of the heavier star can be determined. Further

insights from the accretion disc solidified the presence of a black hole (Rothschild et al., 1974; Tananbaum et al., 1972). Most recently, astronomers were able to map the event horizon of a Super Massive Black Hole (SMBH) using an array of telescopes across the globe (Akiyama et al., 2019). Marking the closest image to a direct detection in astronomical history.

SMBHs are one class of black hole, the heaviest black holes in the universe, reaching masses of $10^9 M_\odot$. They are found in the centre of galaxies and are predicted to be a source of GWs for future space-based GW missions (Hughes, 2000, 2001). Looking into the distant future, the big bang observatory has been proposed as a way to find GW signals from primordial black holes (Crowder and Cornish, 2005). Primordial black holes are theoretical objects made in the early universe from the gravitational collapse from the initial chaotic expansion of the universe (Zel'dovich and Novikov, 1967; Hawking, 1971). Primordial black holes are theorised to be significantly lighter than their younger, more common siblings, stellar mass black holes. Recently, stellar mass black hole mergers have been attributed as the primary mechanism to make Intermediate Mass Black Holes (IMBHs), black holes between $100-100\,000 M_\odot$ (Abbott et al., 2020c). This is just one of the many questions Gravitational-Wave astronomy has answered. Prior to this, there was no stellar evolutionary channel that could create this class of object, but some theories existed (Clesse and Garcia-Bellido, 2015; Chilingarian et al., 2018).

IMBHs are comprised of two Stellar Mass Black Holes, black holes with mass $< 100 M_\odot$. Stellar Mass Black Holes are formed from the core collapse of heavy stars, the same as neutron stars. Some theories exist modelling the collapse of neutron stars into stellar mass black-holes via accretion, but this phenomena has yet to be observed (Giacomazzo and Perna, 2012). Other models for neutron star collapse have been made thanks to multi-messenger astronomy, which I will talk about later.

Neutron Stars

Stars roughly 8 times the mass of our Sun or heavier will take the evolutionary track that results in them becoming neutron stars (or in extreme situations black holes). Unlike solar-mass stars, larger stars have the internal temperatures sufficient to fuse heavier elements (Oxygen, Carbon, and Silicon). Eventually the massive star will begin to produce iron-56 and nickel-56, which cannot be fused (i.e. they take more energy to fuse than the reaction produces). This begins the gravitational collapse process. As the star collapses, the stellar material is bombarded with high energy gamma photons, removing the protons from the atomic nuclei in a process called

photodisintegration (Chadwick and Goldhaber, 1934). As the collapse continues further, temperature and density increase monumentally leading to protons combining with electrons, dubbed electron capture (Alvarez, 1937). Once the star has collapsed, the majority of the material left is degenerate neutron matter (hence neutron star). The typical mass of a neutron star is $1.4 M_{\odot}$ (Miller, 2004; Seeds and Backman, 2009); this is also the Chandrasekhar limit, the point at which electron degeneracy pressure can no longer prevent gravitational collapse. Assuming an ideal Fermi gas, one can use the same logic applied by Chandrasekhar to find the upper limit of a neutron star ($5.76 M_{\odot}$); however, in GR, gravity is stronger and must be accounted for when determining the upper limit of a neutron star (Srinivasan, 2002). Originally, the expected neutron degeneracy pressure accounting for relativistic hydrostatic equilibrium was predicted to only sustain a star of about $0.7 M_{\odot}$ (Oppenheimer and Volkoff, 1939). Given the average mass of a neutron star is double this, something about the prediction is unreliable. The error lies in the concept that neutrons could be treated as an ideal Fermi gas, a more accurate equation of state was needed. Using an updated estimate of neutron star material, which is partly based on the observed neutron star mass, a new upper estimate of $2.9 M_{\odot}$ was made 60 years later (Bombaci, 1996; Kalogera and Baym, 1996). This is called the Tolman-Oppenheimer-Volkoff limit. GW follow-up and its role in understanding the neutron star equation of state is discussed later in the introduction.

Neutron stars are about 15-20km in diameter, they have a surface gravity of $\sim 10^{12} ms^{-2}$, some are magnetic; in fact, magnetars are seen to exhibit the strongest magnetic fields in the Universe at the order of 10^{14} Gauss (Kaspi and Beloborodov, 2017, for reference the Sun has a field strength of 1 Gauss at its surface). These magnetic fields are predicted to play a role in the physics observed in neutron star related phenomena, such as: superluminous-supernovae, Gamma-Ray Bursts, Fast Radio Bursts, and kilonovae (Dessart et al., 2012; Rowlinson et al., 2013; Beloborodov, 2017; Metzger et al., 2018). As a star collapses into a neutron star, all the material is condensed to a much smaller volume; angular momentum is conserved, so they spin quite fast (Hessels et al., 2006, up to 716 rotations per second). This was an important characteristic for finding the first indirect evidence of gravitational radiation (Hulse and Taylor, 1975). One point of interest is that neutron stars can theoretically spin a lot faster than observations account for, the two prevailing concepts as to why this is not the case are magnetic breaking through dipole radiation or GW emission from *mountains* on the neutron star surface (also known as r-mode instabilities) (Staff et al., 2012; Haskell and Patruno, 2017; Gittins and Andersson, 2019). Later, I present the results of a GW project that utilises

r-modes to constrain the size of the aforementioned *mountains*.

There is a final type of neutron star, the hyper massive neutron star. These are hypothetical objects (in that we have no observations of them) that exist for milliseconds before collapsing into black holes. These stars are supported against collapse by their immense spin (Shibata et al., 2006). Hyper massive neutron stars are expected to form from the merger of two neutron stars and are expected to play a key role in the physics of an EM counterpart (Hotokezaka et al., 2013; Metzger and Fernández, 2014).

White Dwarfs

These stars are destined to be the most common celestial objects in the Milky Way. Unlike the evolutionary track of the neutron star, white dwarfs form when a low mass star (less than $8 M_{\odot}$) uses up its fuel, but does not have the mass sufficient to fuse elements heavier than carbon. When a star uses up most of its hydrogen, the remaining hydrogen and helium begin to fuse in layered shells creating a red giant. Eventually the core will be comprised predominantly of carbon and oxygen. As fusion cannot be sustained in the core, it begins to contract until hydrostatic equilibrium is supported by electron degeneracy pressure. The surface layers will decouple from the star from stellar winds creating a planetary nebula revealing a white dwarf.

White dwarfs cores are composed of electron degenerate matter. They are about the mass of the Sun condensed into the size of Earth. Compact binaries containing a white dwarf are not detectable in current GW observatories due to their GW frequency being too low, outside of the GW observatory detection range; however, future missions such as eLISA, the TianQin project, and the Einstein telescope expect to hear them (Babak et al., 2008; Regimbau et al., 2014; Huang et al., 2020).

White dwarf binary mergers have yet to be detected in GW, but given their local commonality, white dwarf mergers should be relatively frequent. Most white dwarfs are expected to have exceptionally long merging timescales (Segretain et al., 1997; Badenes and Maoz, 2012) with observations showing some white dwarfs even drift away from their companions (Pala et al., 2016). White dwarf binaries are however, good sources of continuous signal for eLISA. As eLISA orbits the Sun it will change its orientation to the galaxy. Because of the shift in orientation, the signal from the white dwarf binaries will change. The change in signal can inform astronomers of the distribution of white dwarfs binaries in the Galaxy; allowing a 3D-map of binaries to be built. Using these signals further, astronomers can

probe the nature of the binary dynamics, comparing observation to the current understanding (Edlund et al., 2005).

Making Compact Binaries

The chances of a star having a binary partner is incredibly high, with upper estimates showing binarity to occur in sun like stars 50% of the time, 70% in A-type stars, and a predicted maximum of 100% of O-types (Duquennoy and Mayor, 1991; Raghavan et al., 2010; Duchêne and Kraus, 2013). It should be noted, not all of these systems are interacting binaries. As the frequency of binaries is so high, especially in heavier stars, there should be a reasonably high count of compact binaries too. The expectation is that $\sim 1\text{-}5\%$ of stellar mass black holes and neutron stars should appear in binaries (Tauris and Van Den Heuvel, 2006; Lamberts et al., 2018).

The general thinking around the topic of heavy compact binary formation is based on the evolution of the two comprising stars (Dominik et al., 2012). As they evolve through their lifetime, the heavier star will expand first. This creates an accretion phase. This is where the expansion of the star fills the Roche lobe and accretes to its companion. The host will explode and collapse, leaving either a black hole or neutron star remnant. The companion will eventually expand as well and accrete to the compact host star. Eventually helium burning will start, the expansion creating a common envelope. In this phase the drag force will draw the stellar cores closer together. Either the cores merge in the envelope, creating a neutron star/black hole core, or the envelope evaporates leaving a detached binary. The companion then goes supernova, if the kick is strong enough the system is unbound leaving two single compact objects. If the binary survives the explosion, what is left is a compact binary, the main GW sources. This is a surface level detailing of just one way compact binaries can form, figure 1.2 shows a cartoon of the channel discussed here. In reality, there are an infinite number of ways two stellar remnants can form a close binary. Postnov and Yungelson (2014) details multiple channels, highlighting that the channel for white dwarf mergers has nuanced differences compared to the channel for two heavier stars. The physical processes and differences between the different merger types are explained further in Section 1.1.7.

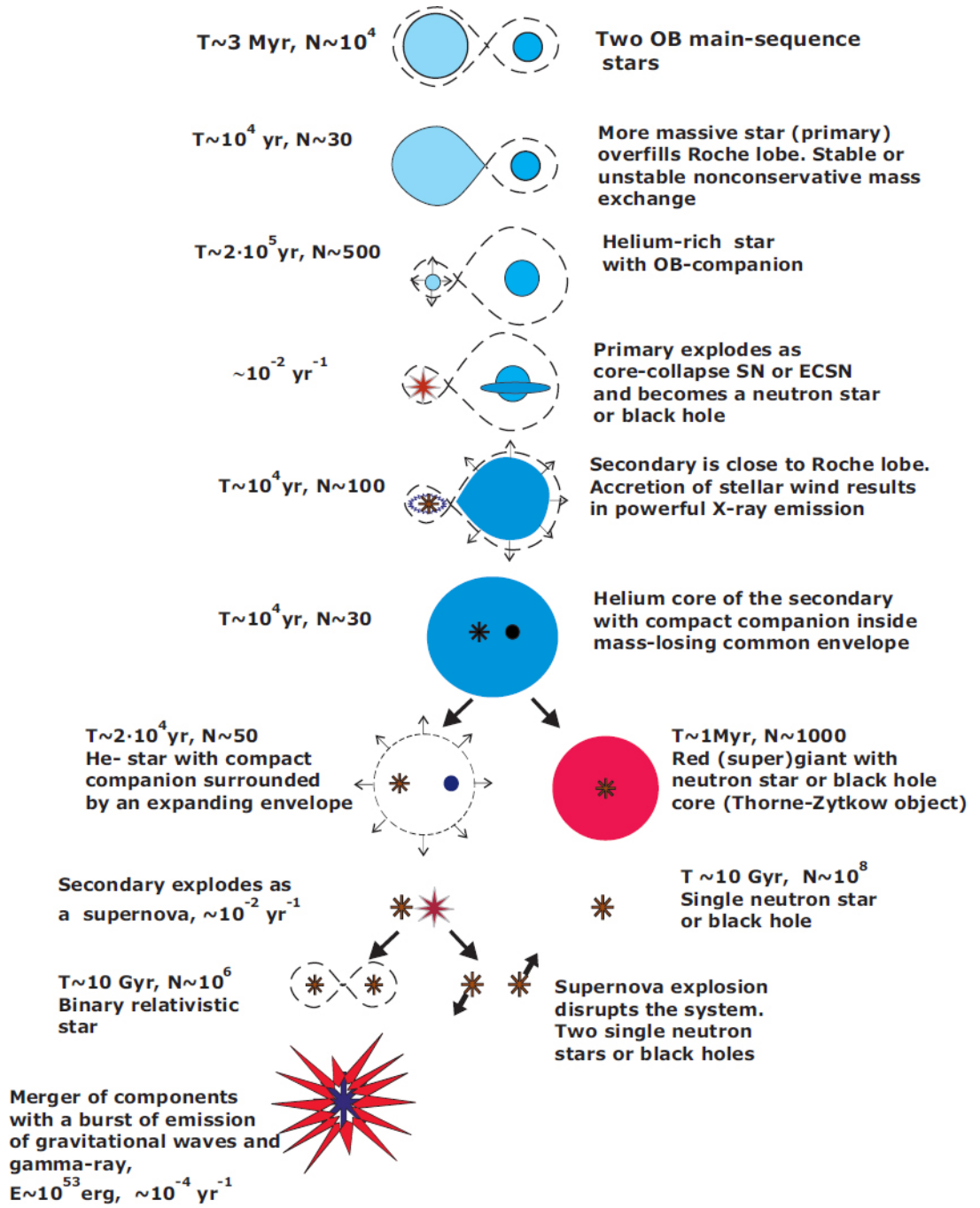


Figure 1.2: Binary evolution to create a massive compact binary (Postnov and Yungelson, 2014). This is only one of many binary channels; however, the visual acts as a useful representation of how close-in compact binaries can form.

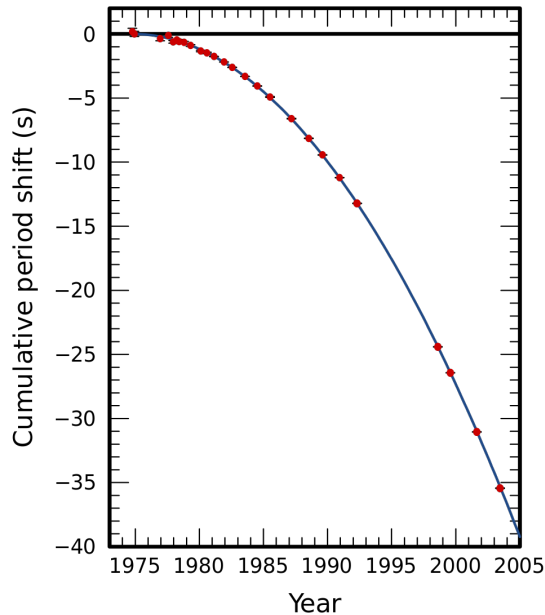


Figure 1.3: Comparing the orbital decay of an observed neutron star binary B1913+16 to that predicted by GR (Weisberg and Taylor, 2005). This shows how the observations of the binary orbital period matches the predicted evolution from GR almost exactly. Acting as indirect evidence of Gravitational-Radiation.

1.1.4 Initial Observations of Gravitational Waves

The search for GWs has been ongoing since the late 1960s. The initial searches were conducted using giant blocks of aluminum with piezoelectric sensors, called a Weber Bar (Weber, 1968), with the hope that the stretching and contraction by a Gravitational Wave would create a noticeable voltage that could measure the strain. Although optimistic (Weber, 1969), high confidence detections were not found. Unfortunately, this method was only sensitive to 10^{-16} cm changes. It is a well publicised fact, detection of GWs require sensitivities to sub-proton scales (10^{-21} cm) (Castelvecchi, 2019).

Astronomers found the first indirect evidence in support of gravitational radiation in 1974 using pulsar timing from a binary neutron star system (Hulse and Taylor, 1975). Pulsars are rapidly spinning neutron stars that emit “pulsed” radio signals. Pulsars were first discovered in 1968 (Hewish et al., 1968) and have played an important role in many areas of astrophysics. The discovery pulsars was awarded a Nobel prize in 1974, credit should also be attributed to Dame Jocelyn Bell Burnell who was sadly overlooked when the prize was awarded. An important feature of pulsars is that the “pulsed” signals happen consistently, taking the same amount of

time, like clockwork. This has dubbed pulsars the “lighthouses” or “timekeepers” of the Universe (Stairs et al., 2000; Lorimer, 1996). If a pulsar has inconsistent timing, it has something perturbing it. This fundamental notion played a key role in the first discovery of exoplanets (Wolszczan and Frail, 1992). However, it was a binary neutron star pair that was the focus of Hulse and Taylor (1975). First came the discovery of the pulsar being in a binary orbit. This was due to systematic shifts in the arrival time of the pulses. Using the delay time, one can determine the size of the binary separation. In addition, the frequency of the delay shifts can be used to figure out the orbital period. Which can finally be used to determine the mass ratio of the two stars. As shown in equation 1.7, over time the orbit is expected to shrink. By taking follow up observations of the system, astronomers were able to match the shift in orbital periastron (the point at which the stars are closest together) to the decay rate predicted by GR (Taylor and Weisberg, 1989), shown in figure 1.3. The Nobel prize in 1993 was attributed to this discovery, emphasising its importance.

Millisecond pulsars, a type of pulsar with a rotation period on the order of 10 milliseconds, have been given another role in the detection of GWs. By looking at a large number of millisecond pulsars one can use correlations in the perturbations of the pulse time of arrivals to model a galaxy-wide perturber; in other words, GWs. There are two main pulsar timing array projects, North American Nanohertz Observatory for Gravitational Waves (NANOGrav) and the European Pulsar Timing Array (EPTA), neither of which has found any significance evidence for GWs (Almoznino et al., 1993; Demorest et al., 2012; Kramer and Champion, 2013; Taylor et al., 2017).

GWs were first detected directly in 2015. Rocketing the era of GW and multi-messenger astronomy. The reason GWs remained elusive for so long is due to the sensitivity required to detect them. Thus introducing the GW Observatories.

1.1.5 Gravitational Wave Detectors

The standard GW detector works using the simple property of interferometry. The concept that when two electromagnetic waves meet they interact with one another producing an interference pattern dependent on what phase they meet. LIGO operates by using a high powered laser. Simply, the laser is fed through to a beam splitter which separates the beam into two perpendicular directions. These are the detector arms. Each arm is 4km long and at the end of each arm is a mirror reflecting the beam back to the splitter. The two perpendicular beams are recombined to make the interference pattern (Martynov et al., 2016). When a GW interacts with a detector the elliptical nature of the wave (shown in figure 1.1) is exploited. As

one arm of the detector is contracted by a GW, the other will extend. This is a physical change in the length of the arms which will alter the distance the photons travel; resulting in the photons meeting out of phase at the detector, creating an interference pattern. This interference pattern can be used to find the strain ($h_{+, \times}$ from equation 1.6) of the incoming GW.

GW detectors need to be incredibly sensitive, this sensitivity is achieved through noise mitigation with hardware. The main source of low frequency noise is terrestrial (i.e. generated on Earth). This is usually through physical vibrations like cars passing by or earthquakes. To reduce these influences the test masses are suspended on a pulley system designed to mitigate thermal noise and low frequency vibrations (Harry et al., 2010).

A key feature for detecting GWs is arm length. As the photon journey time is dependent on how much the arm is stretched, a longer arm will provide a higher resolution of this stretching. Unfortunately, the curvature of the Earth and finite resources prevent the LIGO arms from reaching 100km distances. Instead, a Fabry-Perot cavity is placed in the 4km arms. This done by adding a partially reflective one-way mirror at the beginning of the arm. The laser passes through the non-reflective side and is reflected back down the arm where it is then reflected back up the arm (effectively trapping the photons). There are approximately 280 reflections up and down the arm before the light passes through the cavity again. This increase the arm length by a factor of 280, resulting in an effective length of 1120km; giving the detectors the sensitivity required to observe GWs (Rakhmanov et al., 2004).

Power recycling is a technique used to lower shot noise by enhancing laser power (Braginsky et al., 2002). Signal recycling was first introduced in Meers (1988) and was originally implemented in GEO600. Specifically, the purpose of these power recycling mirrors is to increase the output power from the laser at the sensor from roughly 10kW to 750kW. Basically, power recycling consists of a one way reflective mirror located by the photodetector. The photons are reflected back similarly to the Fabry-Perot cavity. Because the photons are repurposed, they are said to be 'recycled', this increases the power of the laser by up to a factor of 3750. Resulting in a beam powerful enough to resolve a GW signal.

Light squeezing is also used to overcome shot noise (Abadie et al., 2011). This is done by exploiting the Heisenberg uncertainty principal. Put briefly, by extending the uncertainty of the unused signal in one direction, the orthogonal used signal's quality can be improved; with the caveat being the overall uncertainty remains less than $\hbar/4\pi$ (\hbar being Planck's constant). For LIGO, this squeeze happens in amplitude and phase, where the phase is squeezed below the quantum limit (figure

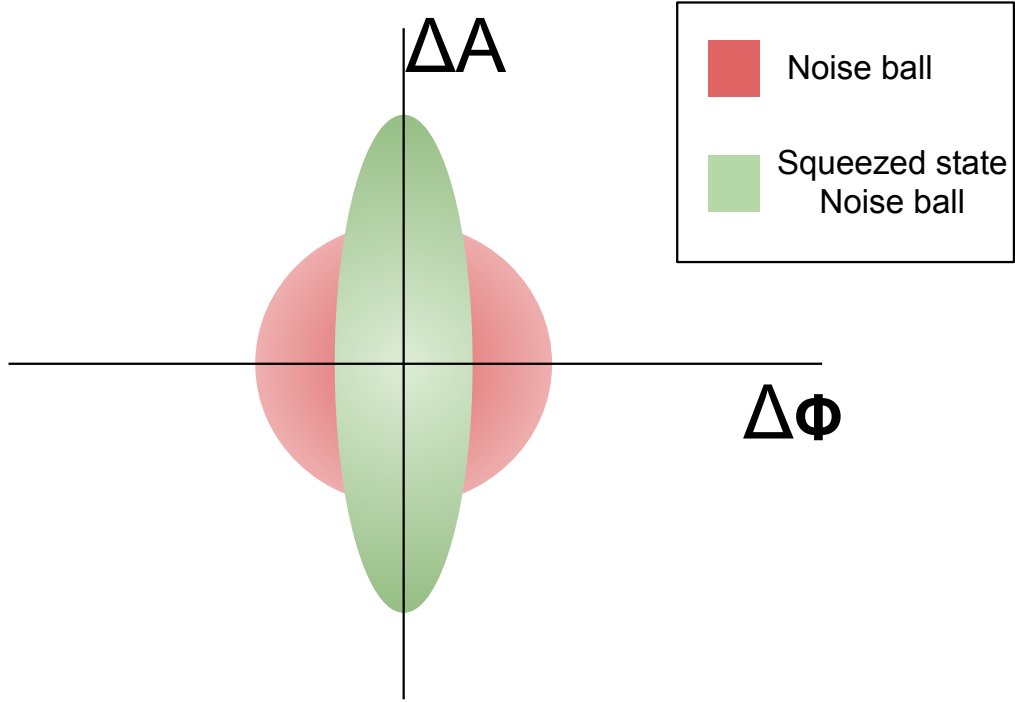


Figure 1.4: Visualisation of a squeezed LIGO signal. The uncertainty of the phase is squeezed below the quantum noise limit, but is traded for a higher uncertainty in the amplitude.

1.4). This is because the amplitude information is less important for GW detection than phase information of the incoming light source. Cryogenics are utilised to lower thermal noise further (Uchiyama et al., 2012) and will be a pivotal feature introduced to KAGRA (Somiya, 2012).

All of these features can be used to model the noise characteristics of the GW detectors (shown in figure 1.5). Fringes in the real noise curve are due to harmonics in the system, the model does not account for this which is why it looks smooth. Figure 1.5 shows the noise profile of the first LIGO detectors (Fritschel, 2003). The frequency dependence was originally optimised for finding Neutron star binary mergers, as they were predicted to be the most common source at the time. This can be verified by looking at the most sensitive region at $\approx 1\text{kHz}$, which matches the expected Neutron star binary frequency (Faber et al., 2002).

With further scientific investigation and hardware developments, GW detectors can only become more sensitive. Two new observatories will be added by 2025, and two additional missions are being planned for the near future. Ultimately

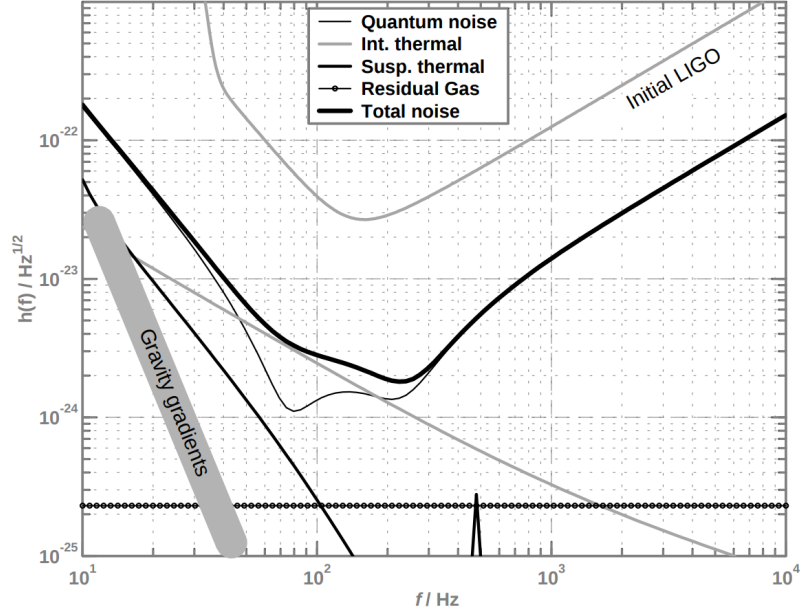


Figure 1.5: The initial LIGO-O2 noise characterisation (Fritschel, 2003). The central black line represents the total noise limit, which is the sum of all the other noise curves (excluding the initial LIGO curve at the top).

providing astronomers and astrophysicists with a surplus of GW sources (Caprini et al., 2016). The new observatories are KAGRA and LIGO-India (IndiGO) (LVC, 2019a). KAGRA will be a cryogenic GW observatory passing through the Kamioka mountains. It is optimistically aiming to see down to 140 Mpc for neutron merger events (Abbott et al., 2018b). The intention is for IndiGO to be a replica of the advance LIGO systems. These plan to see neutron star mergers as deep as 330Mpc (LVC, 2019a). For reference, current detectors reportedly see down to approximately 120Mpc (Abbott et al., 2018b). There are multiple advantages to having numerous detectors. Firstly, by having multiple observatories not all of them have to be online at once. This means longer or even continuous observing runs without having to schedule a break to upgrade observatories. More importantly, multiple detectors means better localisation. As GWs travel at the speed of light, one can use the time of the detection at different locations to triangulate the source position in the sky. The more detectors there are online, the higher the localisation information.

The other two GW missions in the future are the Einstein Telescope (Punturo et al., 2010) and eLISA (Danzmann et al., 2011). The Einstein telescope is an array of three arms with 60 degree separation, they are planned to be placed underground to mitigate terrestrial noise. Having the arms with 60 degree separation presents the opportunity to get polarisation information from the incoming GW. This offers

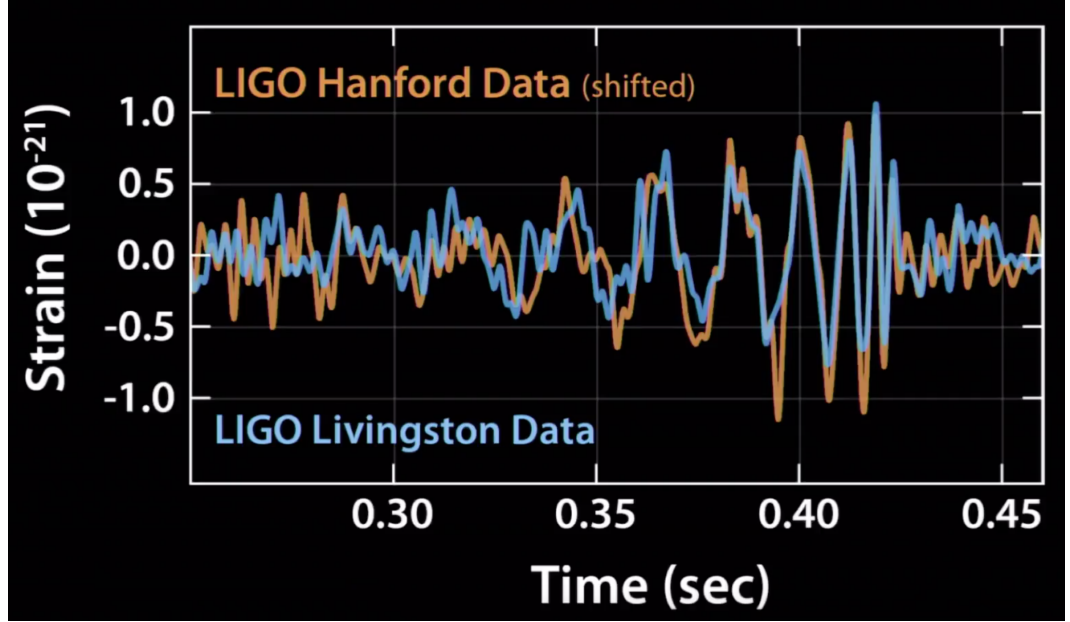


Figure 1.6: First Gravitational-Wave signal (Abbott et al., 2016b). Note how the signal timescale is a fraction of a second, the chirp (merger) occurs at ~ 0.4 seconds.

astrophysical information about the merger and, with multiple detectors, another method of localising the event.

The European Laser Interferometer Space Antenna (eLISA) is a proposed space based mission. It will be sensitive to much lighter compact binaries, with plans for it to model the distribution of double-degenerate binary white dwarfs (Amaro-Seoane et al., 2012b). Among other scientific endeavours, eLISA will provide further localisation information for astronomers and polarisation information (Amaro-Seoane et al., 2012a). This is fundamentally important as it will inform astronomers of the viewing angle, the inclination of the merger plane to Earth's line of sight. In combination to the fact eLISA will discover more neutron star mergers, having polarisation information will mean significant improvement can be made to kilonovae models. A similar design to eLISA is being proposed for another project called TianQin (Huang et al., 2020). Both eLISA and TianQin are proposed to launch at some point after 2030 and will contain three space crafts, that form a large space-based interferometer.

LIGO's 2015 detection was awarded the Nobel prize in 2017. This detection confirmed Einsteins theory of GR and showed the first ever GW signal (Figure 1.6). Properties of a GW signal can be exploited to infer parameters of the source and merger remnant. Parameters such as mass, spin, and inclination. However,

for astronomers one of the most urgent characteristics is sky localisation. Prompt sky localisation is needed for early Electro-Magnetic (EM) follow up. This is done using a rapid Bayesian approach (Singer and Price, 2016). By taking the signal and making some basic assumptions on the physical characteristics, a sky map (like the one shown in figure 1.7) can be produced. For GW150914, the localisation covered $630^{\circ 2}$ of the sky. This localisation process is important for the motivations of this thesis, therefore it is vital to understand how the skymap is built to effectively utilise it in follow-up surveys.

1.1.6 Localising Gravitational-Wave Signals

There are two techniques used when localising events. **BAYESTAR** (Singer and Price, 2016) and **LALInference** (Veitch et al., 2015). For astronomers, **BAYESTAR** is more beneficial as it has a much faster output, necessary for prompt follow-up of GW counterparts. The basis of **BAYESTAR** comes from a seminal paper on parameter estimation of merging compact binaries, Cutler and Flanagan (1994). The original **LALInference** method works by using a full parameter model based on the observations with a Markov-Chain Monte Carlo (MCMC) estimation. Depending on the event this can take on the order of days to complete, as kilonovae fade dramatically on these timescales a faster approach is necessary. Using some clever statistics, the authors were able to remove unnecessary parameters to their localisation estimates. Mainly the mass of the system. In addition to lowering the parameter space, the authors also used the matched-filter pipeline output as opposed to analysing the raw data. They found the pipeline data were able to reproduce a likelihood function similar to the raw data, this removed the need for reprocessing the data. Instead of needing to use a MCMC technique, the authors were able to implement a Bayesian estimator. This requires less computation and can be done in parallel, which MCMC methods cannot. All of these improvements combined means viable localisation skymaps (like the one seen in Figure 1.7) can be built in seconds and distributed to astronomers for the rapid identification of the EM counterpart.

These distributed skymaps are vital for astronomers as they constrain where we need to look to find a counterpart. One fundamental issue, flagged by GW150914, is the uncertainty in the localisation. Survey telescopes, at most, cover a few square degrees of the sky per exposure. The uncertainty regions presented by LIGO can span over 10,000 degrees squared. Finding a counterpart associated with these events would entirely depend on looking at the right spot; in other words, luck! To maximise the chance of finding counterpart quickly, a large survey area is needed accompanied with a high cadence. This is to cover as much of the localisation region

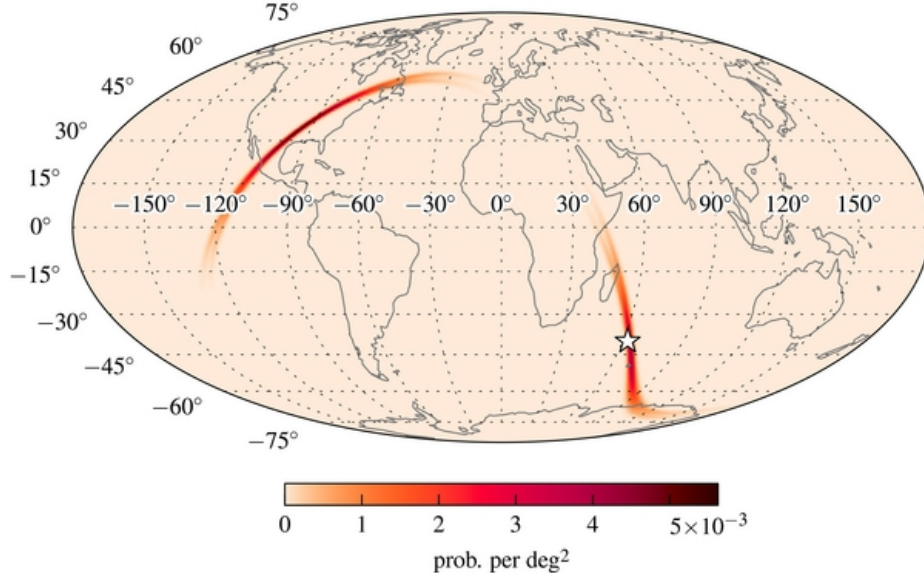


Figure 1.7: The initial LIGO map for GW150914 from Singer et al. (2014). The total localisation region spans 630 sq. degrees.

as possible. As more GW observatories are added and hardware is improved these uncertainty regions should shrink; if this happens in practice is another question. Until then, the solution for practical multi-messenger astronomy is wide-field, high cadence follow-up surveys.

The above dictates the motivation of this thesis. The current challenge is to survey the broad localisation quickly and to a particular depth to rapidly identify fast evolving transient counterparts associated with GW triggers.

1.1.7 Making a Counterpart

Binary Black Holes

Binary Black Holes (BBHs) are the primary demographic of GW detections. BBH mergers are mostly anticipated to be EM silent, a counterpart is not expected (Abbott et al., 2016a). This is because any circumbinary material is expected to be ejected through angular momentum injection and circumstellar material around the individual black holes will be engulfed by the resultant post-merger black hole. However, many theories exist that predict some emission from BBH mergers (Palenzuela et al., 2010; Moesta et al., 2012; Noble et al., 2012; Paschalidis et al., 2013; Connaughton et al., 2016; Loeb, 2016; Liu et al., 2016; Liebling and Palenzuela, 2016; Li et al., 2016; Murase et al., 2016; Woosley, 2016; Zhang, 2016; Bartos et al., 2017;

Dai et al., 2017; De Mink and King, 2017; Ioka et al., 2017; Janiuk et al., 2017; Stone et al., 2017; Frascchetti, 2018; Kimura et al., 2017a,b; McKernan et al., 2018; Perna et al., 2018). The commonality of all the theories suggest that material is accreted by the resultant post-merger black hole. The matter outside of the event horizon either comes from fallback of material with injected angular momentum from the coalescing black holes, kicks from the merger itself, or through collection of material as the black hole moves through the inter stellar medium.

Recently a detection of a flare coincident with the localisation of a potential BBH merger GW candidate (24621, GW190521g) has drawn the attention of astronomers (Graham et al., 2020). The current concept is that a merger has occurred while orbiting an active galactic nuclei (AGN), similar to Bartos et al. (2017), the surrounding AGN disc material began accreting on to the resultant black hole creating the flare. The flare is expected to repeat which would confirm the presence of a black hole. This is not equivalent to prompt emission however, therefore fast targeted follow-up campaigns should not expect to find these kinds of signals.

Targeting BBH merger events is important, not only to constrain the models and confirm if there is no expected counterpart (at least in terms of prompt emission), but also to optimise the search strategy when trying to hunt counterparts from neutron star GW triggers.

NS-NS

Merging Neutron stars are the principal source of expected EM emission. As neither component contains a black hole, none of the circumstellar material is expected to be trapped within the event horizon. As the two stars merge, they are disrupted ejecting neutron rich material into the stellar region. The result of the merger leaves a compact object in the centre surrounded by in-falling, neutron rich, material (Rosswog et al., 2000; Lippuner et al., 2017).

The resultant remnant of a neutron star merger is a rapidly rotating Hyper Massive Neutron Star (HMNS). A HMNS is an unstable, fast spinning, highly magnetic neutron star. The reason the star is unstable is because the rapid spin is preventing the star from collapsing into a black hole. The magnetic field is expected to slow the rotation of HMNS, called magnetic breaking. As the star loses angular momentum the spin cannot sustain the mass of the star and it collapses, creating a Gamma-Ray Burst (GRB). This is the first EM emission signifying a counterpart (Shibata et al., 2006). The material that gets thrown out of the merger acts to create the other counterpart to the associated GW, a kilonova. As the HMNS does not collapse immediately, the surrounding material is heated creating the kilonova. It

has been shown the collapse of the HMNS is linked to how the kilonova will appear (Metzger and Fernández, 2014). This collapse has been connected to the equation of state, but also has shown that there is another GW signal possible from this collapse (Hotokezaka et al., 2013). Highlighting the value multi-messenger astronomy has on determining the properties of neutron stars.

NS-BH

Neutron star - Black Hole (NSBH) mergers represent the intermediary branch of BBH and BNS mergers. Some theorists suggest that the neutron star can be swallowed hole, similar to expected BBH scenario, meaning no expected counterpart. However, other NSBH models show potential for a counterpart, often overlapping with the BNS models (Lamb and Kobayashi, 2017).

Once the counterpart has been localised, follow-up observations are made through a range of different wavelengths and observation styles. This is the point where models can be made and dismissed, the science with multi-messenger astronomy.

1.2 Science with Multi-Messenger Astronomy

For any astronomer following up a GW trigger, the most tantalising prospect is the discovery of a kilonova.

Kilonovae were introduced in Metzger et al. (2010) and first observed in Tanvir et al. (2013). Kilonovae are thought to be the result of colliding neutron stars or neutron star - black hole mergers. They are predicted to be roughly 1000 times brighter than nova, and hence, are referred to as kilo-nova. There has also been a term dubbed macronova (Tanaka, 2016) or r-process supernova for these events; however, kilonova has become the accepted name and will be the adopted nomenclature for the rest of this work. Kilonovae are predicted to be preceded by short GRBs, these are collimated jets of high energy photons. GRBs have a distinct signature of a bright high energy prompt emission, followed shortly by an afterglow due to synchrotron radiation (Piran et al., 1998). Both the study of kilonova dynamics and multi-messenger astronomy will further our understanding of how γ -rays are produced (Burns et al., 2019) and how the central engine of a merger can produce these jet structures (Rezzolla et al., 2011).

A fundamental characteristic of kilonovae is their red and blue components, this feature comes from the r-process; or rapid neutron capture (Woosley et al.,

1994). r-process was first proposed as a solution to the dilemma of where most of the heavy elements in the Universe originated (Burbidge et al., 1957). However, at the time r-process was only attributed to core collapse supernovae. Mergers were considered exceptionally uncommon. The r-process can only occur in catastrophic events with a high neutron density and core collapse supernovae were the only known events that met both criteria. As time passed, the community began to doubt that core collapse supernovae would be able to populate the Universe with the amount of heavy elements that we observe (Argast et al., 2004; van de Voort et al., 2014). From GW170817 alone, astronomers were able to predict that a significant fraction of the Universe's heavy element make-up comes from neutron star mergers (Kasen et al., 2017). r-process physics is still unclear, meaning continued follow-up of kilonovae will give astronomers better estimates of the universe's elemental composition and help unveil the mysteries surrounding the physics of the r-process. Recent work suggests that collapsars, the collapse of rapidly rotating massive stars, could be responsible for a portion of the Universe's heavy elements (Siegel et al., 2019; Kobayashi et al., 2020; Brauer et al., 2020). Chemical Evolution of the Universe is vital in understanding stellar populations, star formation rate, initial mass function, and galaxy composition (Tinsley, 1968). All of which are used to discuss the evolution of the universe (Eldridge and Stanway, 2016; Zapartas et al., 2019). As discussed above, kilonovae are suspected to be responsible for some of the Universe's heavy elements. However, it is suspected that current rate estimates for neutron star mergers are too low to explain the metallicity in the observed Universe (Argast et al., 2004). Some studies try to refute the low neutron star merger rate; claiming that there is significant injection of heavy material (Komiya and Shigeyama, 2016; Ojima et al., 2018), but with the current observed BNS merger rate seeming incomplete, it is impossible to conclude if there are enough mergers to fuel the observed universal composition. The physics of kilonovae have still proven to be non-negligible for a galaxy's elemental evolution (Kasen et al., 2015; Ji et al., 2016).

The kilonova model has shown to have some predicted structure (Metzger and Fernández, 2014). There are two main components, a red central part in the plane fuelled by heavy elements some of which are lanthanides from the r-process. At higher latitudes, the explosion turns blue from lighter elements. This blue part is predicted to come from a 'punch' generated by the jets, this gives favour to axisymmetric models when simulating kilonova explosions (Perego et al., 2017; Yi et al., 2018; Kóbori et al., 2020). Figure 1.8 is a cartoon of the axisymmetric model. What determines the colour of the kilonova is the electron fraction (Y_e) in that part of the ejecta. Regions with $Y_e < 0.25$ (neutron rich) will undergo r-process nuclear

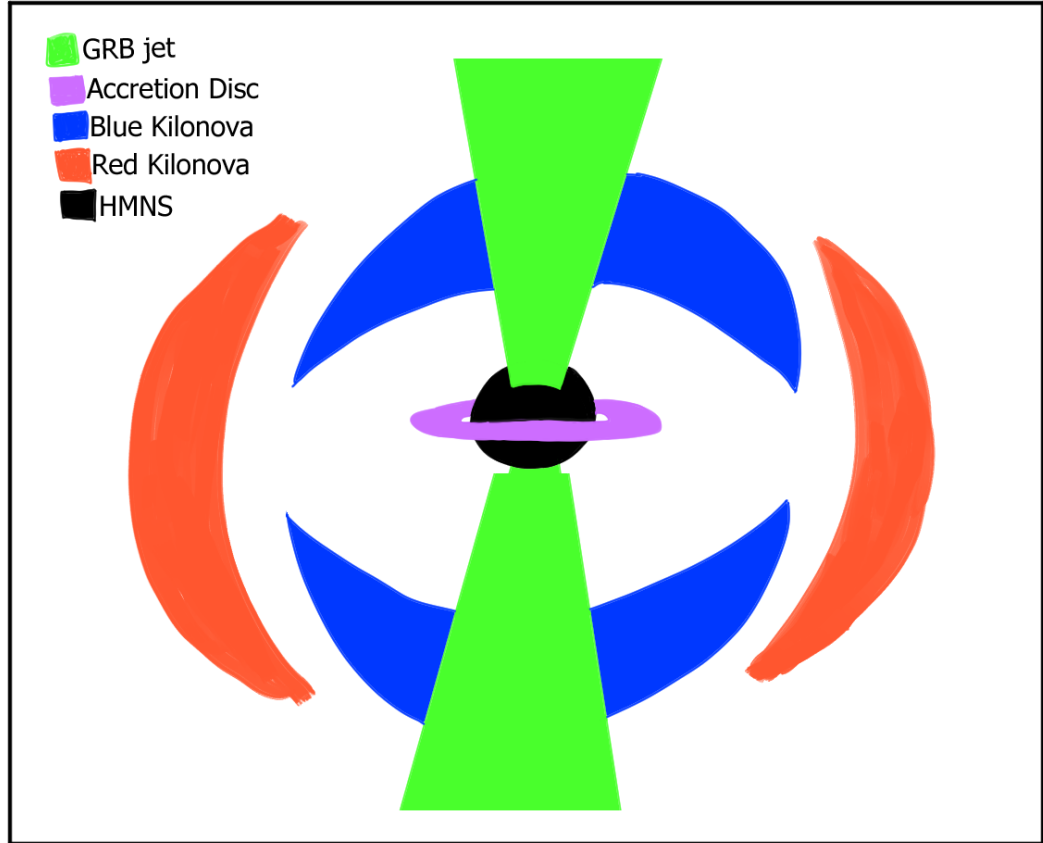


Figure 1.8: Cartoon of a kilonova from the axisymmetric model. The black centre shows the hypermassive neutron star engine, the green indicated the GRB jets, the red (lanthanide rich) material is shown in red, similarly for the blue (lanthanide poor) component, and finally the accretion disk is shown in purple.

synthesis, resulting in high opacity red material. This acts to hide the electron rich material, often referred to as a ‘Lanthanide curtain’ (Kasen et al., 2015; Nativi et al., 2020). This emphasises the viewing angle dependence when looking at the colour information of kilonovae.

Due to the distances, it is impossible to resolve kilonovae geometry with photometry or spectroscopy. Instead, polarimetry provides the best alternative to determining the kilonova’s structure (Bulla et al., 2018). Polarimetry is the observation of light passing through a polarisation modulator. The modulator will cycle through different states (usually rotating through phase angles) which causes a change in detected light intensity, this gives astronomers polarisation information of a source. If a light-source is not polarised, the intensity of the light remains static over all modulation states. Using Stoke’s parameters, one can make estimates of structure of stellar explosions. Due to lack of events, the current literature leaves the geometry of kilonovae a mystery (Covino et al., 2017; Bulla et al., 2019), but the working theory is the cocoon model (Gottlieb et al., 2018). This is the idea that the GRB jet produced by the collapsing central star will heat the ejected material, causing it to inflate into a cocoon-like structure (Murguía-Berthier et al., 2014; Nagakura et al., 2014; Nakar and Piran, 2016).

GWs offer advancements in other areas of science. This can be in the form of numerical methods (Lange et al., 2017). Numerical relativity can be used to predict signals from less common systems, or model received signals to determine the most likely astrophysical source to a particular GW detection. Furthermore, numerical relativity can even be used to model the equation of state for neutron stars (Kiuchi et al., 2020).

Merging compact objects are now being used to find cosmological parameters. The Hubble constant has been readdressed in the framework of GW signal analysis (Collaboration et al., 2017). It is hoped that with more observations of GWs and their counterparts, the estimated Hubble constant will help resolve the current dichotomy of the constant’s prediction from type 1a supernovae and Cosmic Microwave Background (CMB) observations (Chen et al., 2018). The CMB acts as a shield for astronomers; the CMB is the beginning of the observable universe, at least in terms of EM observation. Due to the fact that no earlier light can pass through the CMB, EM observations can only probe to the first 380,000 years of the Universe (Ryden, 2017). We know that the Universe must be older due to current age estimates from Universal expansion and the physics necessary to produce the Universe we see today. GWs are not blocked by the CMB. Meaning, primordial GWs present the best opportunity to probe deeper into the early universe (Mag-

giore, 2000). There are plans for a potential space observatory dedicated to finding these early GW emissions (Crowder and Cornish, 2005).

The lack of a GW detection can also have a scientific impact. As mentioned earlier, the spin of neutron stars are an important and enigmatic property of the stellar remnants. One characteristic expected is that if there are any imperfections (*mountains*) on the surface of a millisecond pulsar there would be a GW signal matching the frequency of the pulses. Researchers have constrained the *smoothness* of multiple pulsars from the null detection of GWs in the expected frequency bins (Abbott et al., 2020a). The study concludes that neutron stars are smooth to within a hair’s breadth, but more detectors with higher sensitivities are required to accurately measure the spin-down property of pulsars due to structural deformities.

One interesting feature from observing kilonovae is constraining the neutron star equation of state (EoS) (Bauswein and Janka, 2012). The neutron star EoS is vital in understanding multiple phenomena. The mass-radius relation of neutron stars and cooling age are two features that are dependent on the EoS. There is also an intrinsic connection between the EoS and when the neutron star collapses into a black hole. Furthermore, the EoS is a probe into understanding the elementary particle physics under extreme gravitational pressure. It has been shown that detailed observations of GWs and kilonovae are gold mines when it comes to constraining the EoS (Radice et al., 2018). By looking at the GW counterparts, researchers have estimated the upper limits of a neutron star to be $2.3\text{--}2.9M_{\odot}$ (Gao et al., 2016; Margalit and Metzger, 2017; Shibata et al., 2017; Ruiz et al., 2018; Rezzolla et al., 2018; Shibata et al., 2019). More follow up campaigns will lead to a better understanding and constraints.

Kilonovae are expected to evolve on the timescale of days. Figure 1.9 shows the current model lightcurve expected of a kilonova. Finding them early offers the chance to see the first stages of the kilonova progression, this will expose the initial dynamics shortly after the merger and help physicists understand the environment of extreme systems like merging neutron stars. Due to the nature of LIGO and VIRGO, localisation will be a problem for astronomers for the foreseeable future (Abbott et al., 2018b). With large scale sky localisations, counterpart searches with small fields of view would take on the order of weeks. This cadence is well below the needed time-span of hours/days needed to find a kilonova. One solution to speed up this search is to use a high cadence wide field astronomical sky survey and monitor changes in the sky that coincide with GW alerts. The Gravitational-wave Optical Transient Observer (GOTO) is a wide field optical telescope on LaPalma, whose primary purpose is to rapidly follow up optical counterparts to GW alerts

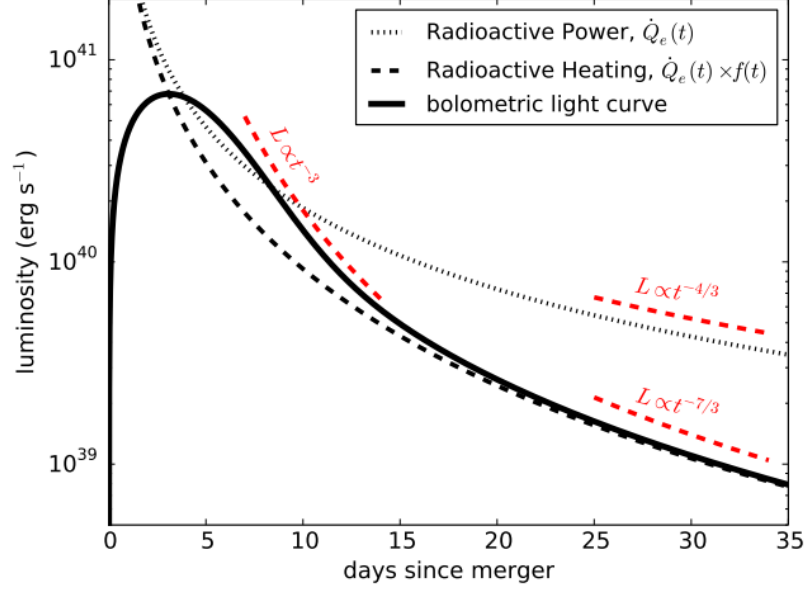


Figure 1.9: Model kilonova lightcurve from Kasen and Barnes (2019). Highlighting the expected peak luminosity and decay timescales are on the order of days.

(Gompertz et al., 2020b; Steeghs et al., in prep). GOTO is the main driver behind this project, with the focus surrounding fast processing and optimisation of the GOTO data.

1.3 Summary of Introduction

This chapter has detailed the physics and detection of GWs and the subsequent process of follow up observations for an EM counterpart, and has shown scientific motivation for finding kilonovae quickly. The challenges of this endeavour originate from the wide localisation regions. As the expected counterparts to GWs are predicted to be rapidly evolving, scouring the large areas of sky quickly maximises ones chances of locating the counterpart. This thesis will discuss some of the underlying techniques used in the Gravitational-wave Optical Transient Observer, a project specifically designed to face the challenges presented.

1.4 Thesis Content

This following is split into seven chapters:

- In the next chapter, I will discuss previous missions to find counterparts to

GWs. Looking at the pitfalls and strengths of older projects, I introduce the Gravitational-wave Optical Transient Observer as a mission designed for the rapid identification of GW counterparts. Then present the astronomical methods used in the context of CCD astronomy.

- Chapter 3 introduces GOTO's fast optics. The chapter outlines the importance of focus and how GOTO achieves optimal focus quickly and automatically.
- Co-Addition is a process used to remove artefacts and reduce noise in the data. Chapter 4 presents different co-addition techniques as methods for building reference images for image subtraction. In addition, a new source alignment tool is described, emphasising the tool's application for image subtraction.
- In this chapter I will detail the uses of image subtraction and introduces new methods designed specifically for wide field of view images. Moreover, this method is then compared to those used in the field currently. Chapter 5 concludes with real-world application of the software.
- A machine learning method is developed for the image subtraction pipeline described in Chapters 4 and 5. The classifier is trained using a novel injection method, which is finally applied to the Gravitational-wave Optical Transient Observer's first GW follow-up campaign.
- Using GOTO's follow-up of the LVC O3 GW observing run, I show transients can be found in real-time coincident with GW triggers. Chapter 7 concludes with the GOTO prototype performance across O3.
- The final chapter summarises the work in this thesis, drawing attention to future developments and what the Gravitational-wave Optical Transient Observer has to offer during the LIGO downtime.

Chapter 2

Methodology

“Some humans would do anything to see if it was possible to do it. If you put a large switch in some cave somewhere, with a sign on it saying ‘End-of-the-World Switch. PLEASE DO NOT TOUCH’, the paint wouldn’t even have time to dry.”

— Terry Pratchett, *Thief of Time*

When a GW skymap is sent out, astronomers point their telescopes at the highest likelihood regions and iterate from there. Here, I am going to discuss the original searches for the first GW triggers, highlighting their weaknesses when looking for a counterpart.

2.1 The Hunt Begins - Optical Surveys

As stated in the introduction, astronomers are dependent on the skymaps built from the GW detection to find electromagnetic counterparts. Skymaps are a co-ordinate grid in right ascension and declination, where each pixel represents a region of the sky. The value of each pixel is a probability of the detected merger being there. The skymap contains the 90% confidence region; if an observer were to search the entire localisation region, there would be a 10% chance the observer would miss the counterpart. The 90% credibility region can span from a few 100 square degrees up to the 10,000s (Abbott et al., 2018b, 2020b; Gompertz et al., 2020b). It is the enormity of these credibility regions that makes finding a counterpart so challenging. The variability of these skymaps arises for a few reasons. Firstly, the number of detectors; at least two detectors are necessary to build a skymap. Using the time difference between the incoming signal at the different observatories, the source location can be triangulated on the sky. With just two observatories, the timing

method produces the classic arcs across the sky. As more detectors are added, the signal location can be better constrained. By making assumptions about the source, the sky location can be reduced further (Singer et al., 2014; Pankow et al., 2018). The other reason is signal to noise, the detectors are incredibly sensitive and therefore have a lot of noise sources. This noise can act to obscure localisation information.

Because of the large localisation regions, wide optical surveys present the best chance of finding a counterpart quickly. An example collaborative effort was the Japanese collaboration for Gravitational wave ElectroMagnetic follow-up (J-GEM). This search was primarily carried out with the Kiso Wide Field Camera (KWFC) (Morokuma et al., 2016). This is a 4 deg^2 field of view survey attached to the 1.05m Schmidt telescope on Mt. Ontake. Its primary focus is a supernova survey (Sako et al., 2012; Morokuma et al., 2014). The KWFC takes a single 3 minute exposure per pointing, with a limiting magnitude of ~ 18.9 . When searching for a counterpart, this survey strategy covers tens of degrees per event.

The Master survey boasts an assortment of survey telescopes in seven different observatories (Lipunov et al., 2016). Each scope uses 180 second exposures while searching for a counterpart, their wide-field of view, covering thousands of square degrees, reportedly seeing down to 19th magnitude; though the paper never flags a transient deeper than 18.2.

The Dark Energy Survey (DES) and the Dark Energy Camera (DECam) on the 4m Blanco Telescope at the Cerro Tololo Inter-American Observatory are also participants in GW follow-up (Herner et al., 2017, 2020). Using a 2.2 deg^2 field of view, DECam can see down to 20th magnitude with a 90 second exposure. However, like KWFC, DECam can only survey a small section of the probability area.

ATLAS (Tonry et al., 2018) is a wide field, all sky survey that can see down to ~ 19 th magnitude. Because of its high cadence and depth it is able to hunt for counterparts to Gravitational-Waves and transients quite effectively. One of the biggest drawbacks is its pixel-scale. Large pixels leaves the survey wanting for sky resolution.

ASAS-SN (pronounced Assassin) is another wide-field transient hunting survey (Kochanek et al., 2017). ASAS-SN can cover the entire sky in one night with a relatively small pixel scale. The sacrifice here is depth; ASAS-SN can only see down to 17th magnitude.

The intermediate Palomar Transient Factory(iPTF) (Kasliwal et al., 2016) upgraded to Zwickey Transient Facility (ZTF) (Ghosh et al., 2017), has the highest potential for transient identification. It has a wide-field of view, seeing $\sim 1000 \text{ deg}^2$

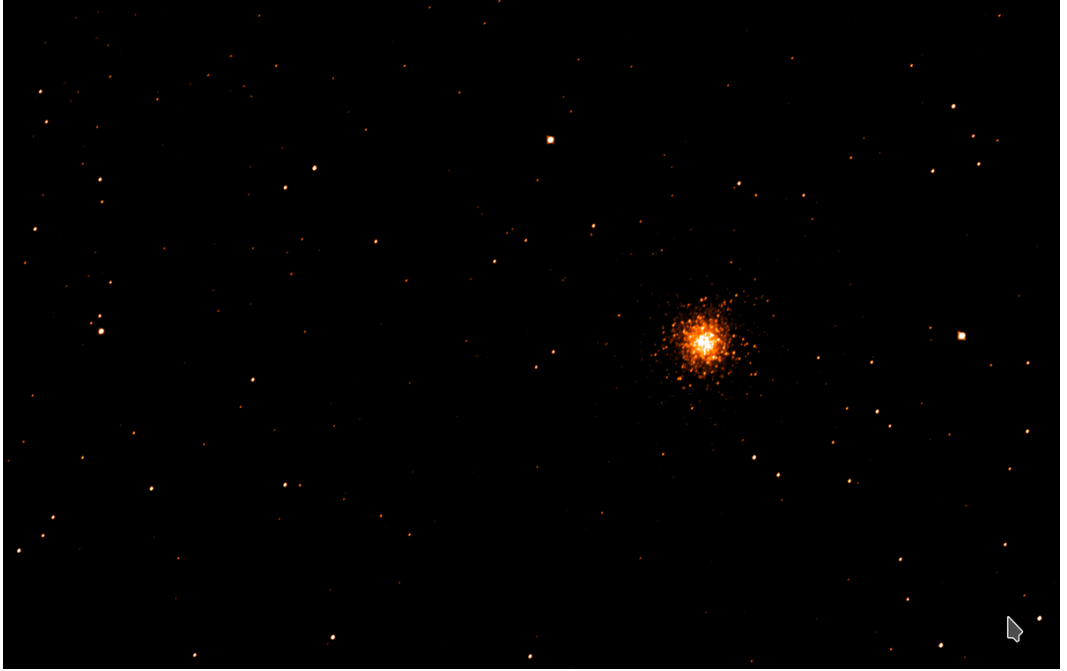


Figure 2.1: GOTO’s first light, the Hercules Globular Cluster (June 11th 2017)

per night, to a depth of 21st magnitude (Kasliwal et al., 2020). The trade here is cost. To build and run ZTF required \$ 24 million (£18.96 million as of Dec 2017).

All of the above projects were built and designed for research outside of GW follow-up. This work introduces a project dedicated specifically to the hunt for EM counterparts to GW events. Combining what these other searches have shown; to find a counterpart to a GW effectively, your survey must see deep (down to $\sim 20^{th}$ magnitude (White, 2014)), span a wide-field, and have a sky cadence on the order of a kilonova timescale (days). An additional challenge is that this is all done with a tight budget. This thesis works on the development of such a survey telescope. The Gravitational-wave Optical Transient Observer (GOTO).

GOTO is an array of multiple wide field of view optical telescopes equipped with large format charge coupled devices (CCDs). GOTO first received support in the beginning of 2014, hosting a collaboration of 5 institutions (University of Warwick, Monash University, University of Sheffield, University of Leicester, and the Armagh Observatory). Construction on La Palma swiftly began in 2015 and on June 11th 2017, a GOTO prototype achieved first light (Fig 2.1). Since then, GOTO has demonstrated its niche in the field of transient astronomy and the consortium has expanded, boasting collaborators from over 12 institutions worldwide

(the original 5 plus, The National Institute of Astronomy in Thailand, Portsmouth University, Instituto de Astrofísica de Canarias, University of Manchester, University of Turku, University of Edinburgh, and Radboud University).

The rest of this chapter will detail the specifications of GOTO and the methodology implemented both in setting-up and taking scientific observations of the night sky. I will then discuss how the features of GOTO are used to optimise hunting for a counterpart to GWs.

2.2 The Gravitational-Wave Optical Transient Observer

GOTO was conceptualised well before the detection of any GWs (White, 2014). Using similar principles as SuperWASP to finding transiting exoplanets (Pollacco et al., 2006), GOTO uses a wide-field multi-camera array to quickly survey localisation regions and identify transients. Figure 2.3 demonstrates GOTO’s expected coverage compared to other missions. GOTO was proposed as a fast transient hunter at low cost, this meant the two main challenges were finding cost appropriate sensors and determining the optimal aperture size for them. Simulations showed that the best compromise for coverage time, depth, and cost was 40 centimetre scopes with a focal-ratio $f/2.5$, this gives the 8304x6220 detector a coverage of ~ 206 arcseconds per millimetre; with the $6\ \mu\text{m}$ pixels this gives GOTO a field of view of 1.24 arcseconds per pixel, also known as the pixel scale (Steehns et al., in prep). One final design hurdle for GOTO was the mount’s boom arm. Unlike other multiplex systems that have unique mounts for each scope, GOTO was designed to have multiple UTs on one mount. This makes GOTO scale-able and easily upgrade-able too. In the first instance, GOTO had 4 telescopes on the mount.

GOTO exists as two entities across this thesis; the prototype, 4-telescope, phase (GOTO-4 see Steehns et al. (in prep)) and the prototype 8-telescope phase (GOTO-8). GOTO-4 consisted of 4 Unit Telescopes (UTs); these are Wynne-Riccardi astrographs (a type of Newtonian telescope with a Wynne corrector), each with a CCD containing a grid of 8000 by 6000 $6\mu\text{m}$ pixels. GOTO-8 has an additional 4 RASA telescopes (UTs 5-8). Figure 2.2 highlights the GOTO-8 configuration. The four additional astrographs increased GOTO’s field of view from ~ 19 degrees squared per pointing, to ~ 40 as shown in Figures 2.3 and 2.4. The addition of UTs 5-8 were as a test phase. These scopes were temporary to help path-find GOTO’s strategy for when 8 second generation scopes are added. Because of this, the footprint (the layout of the four cameras) was sub-optimal as demonstrated on



Figure 2.2: GOTO using its 8 telescope configuration

the right hand plot of figure 2.4.

It is important to note, GOTO’s working configuration has since replaced the original four UTs with second generation ASA-H400 (hereafter ASA) telescopes and future plans are to have a uniform UT set where all the UTs are ASA telescopes. This thesis is concerned with the deployment of the GOTO prototype and only considers the prototype GOTO-4 and GOTO-8 set-ups.

GOTO’s field of view is crucial to its performance. The idea being, GOTO will survey the localisation regions quickly. Having a wide field of view allows for more of the skymap to be searched in fewer pointings. Furthermore, a larger coverage also means a better all-sky cadence. Part of transient hunting is having an up to date reference of the patch of sky being searched. GOTO’s coverage gives an all-sky cadence between 7 (with 8 astrographs) and 14 (with 4 astrographs) days. In GOTO’s full form it will have 16 telescopes in both hemispheres, giving the system an all sky coverage of ~ 4 days; making GOTO one of the fastest optical survey telescopes in the world.

To achieve a relevant reference map, a tiling pattern is used across the sky (Dyer et al., 2018). Each pointing has an associated tilename. When looking for transients, GOTO can query the tilename to find the most relevant reference. By

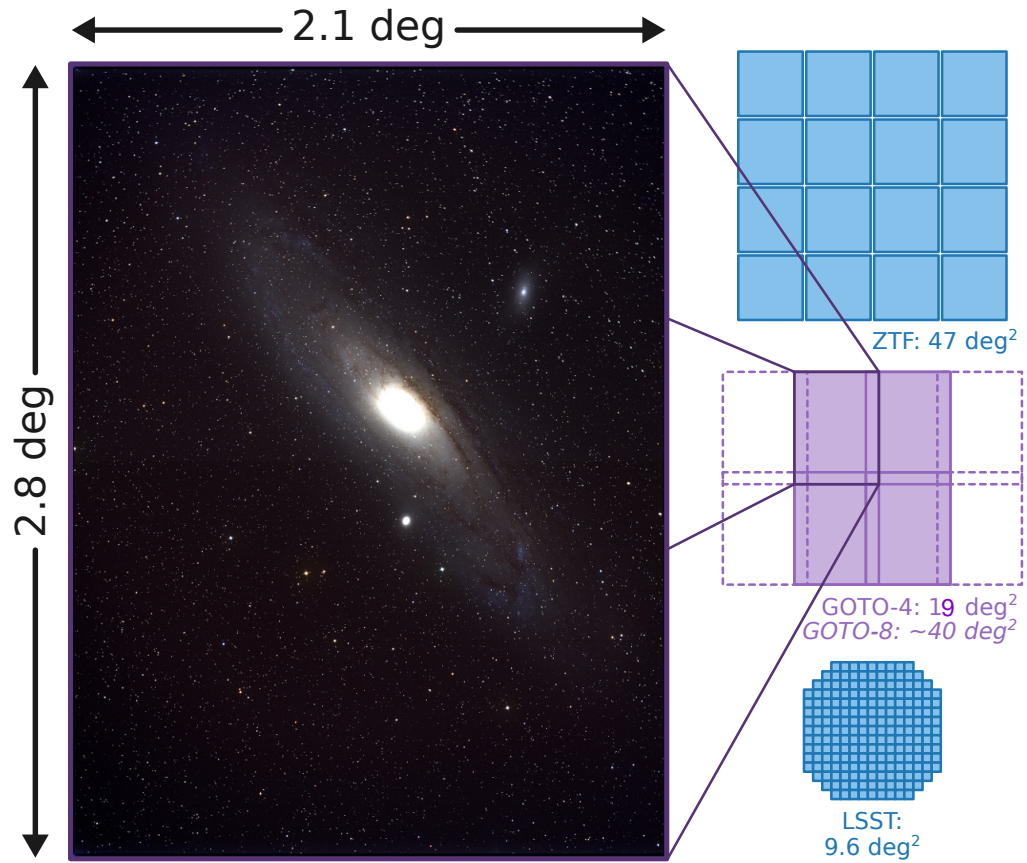


Figure 2.3: GOTO's expected field of view compared to other telescopes (Steehls et al., in prep). GOTO's final configuration will have 32 telescopes, 16 on both hemisphere's, giving it an instantaneous field of view of 80 square degrees. Making GOTO one of the widest survey instruments in astronomy.

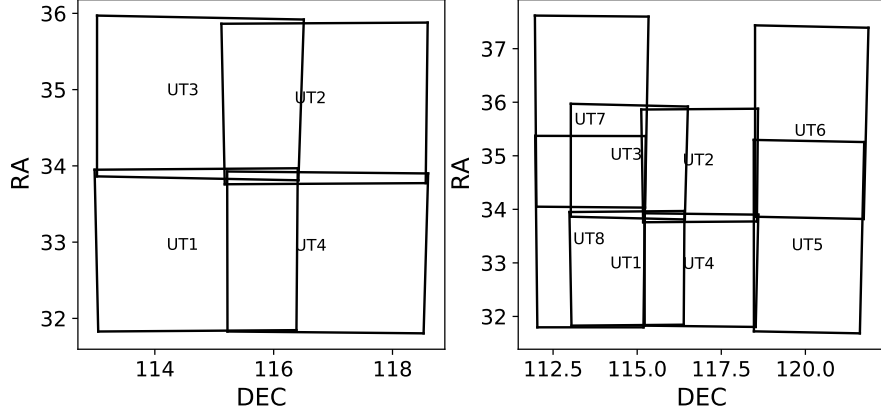


Figure 2.4: GOTO's footprint for GOTO-4 (left) and GOTO-8 (right)

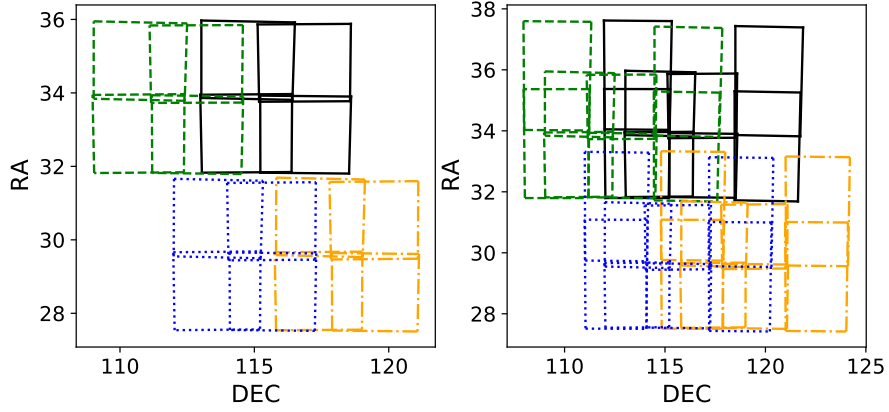


Figure 2.5: GOTO tiling pattern for GOTO-4 (left) and GOTO-8 (right)

having a static reference grid (see figure 2.6), the GOTO logic can also assess the age of the reference. GOTO's aim is to find new transients associated with GW triggers; so, the more recent a reference image the more constrained the age of a transient can be. In other words, by having an up to date sky grid, the age of new transients can be limited to the time between the reference epoch and the updated science frame. This dictates the need for high cadence all-sky reference surveys in addition to the need for high cadence counterpart searches.

The four original telescopes have a filter wheel attached. The filters are Red, Blue, Green, *L-band*, and clear. *L-band* can be described as the combination of the Red, Green, and Blue filters (as seen in figure 2.7) and is similar to the combined SDSS *g* and *r* bands. While GOTO-4 has the capacity to do wavelength dependent

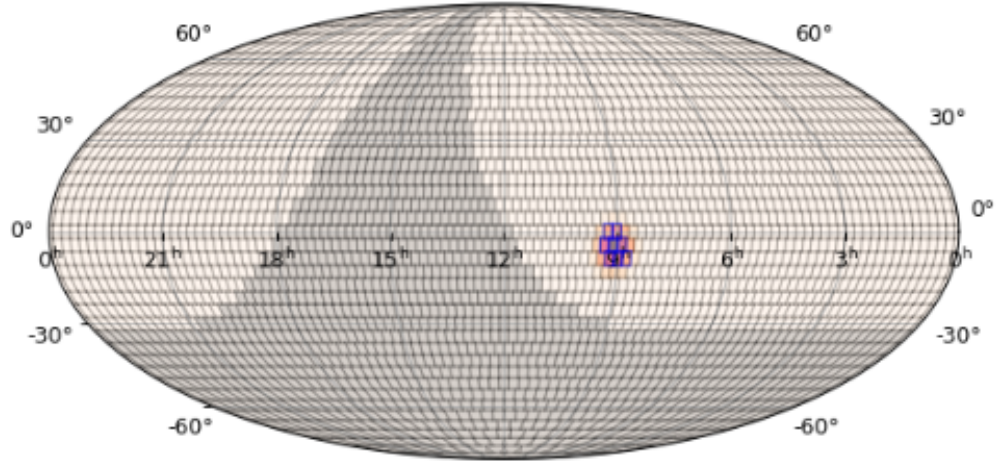


Figure 2.6: Tiling pattern when hunting for a Fermi GRB trigger (Fermi 626581870). Each square on the map represents one footprint from figure 2.4.

photometry, this feature is not utilised here and only the *L-band* is considered.

Each pair of UTs are connected to an intel Next Unit of Computing (NUC). These are used to control the camera operations, described further in (Dyer, 2020). All the UTs are fixed to the boom arm, which is driven by two axial motors, one fixed to the right ascension, the other declination. These are the drivers in moving the telescopes quickly to cover the wide LIGO uncertainty regions, figure 2.8 shows this.

GOTO is a robotic telescope. This means everything should be done automatically; from scheduling observations to assessing image quality. Crucially, if a GW trigger is sent in, GOTO can formulate a search strategy and execute it with no human involvement. This is just one way in which GOTO is one of the fastest follow-up facilities on the planet.

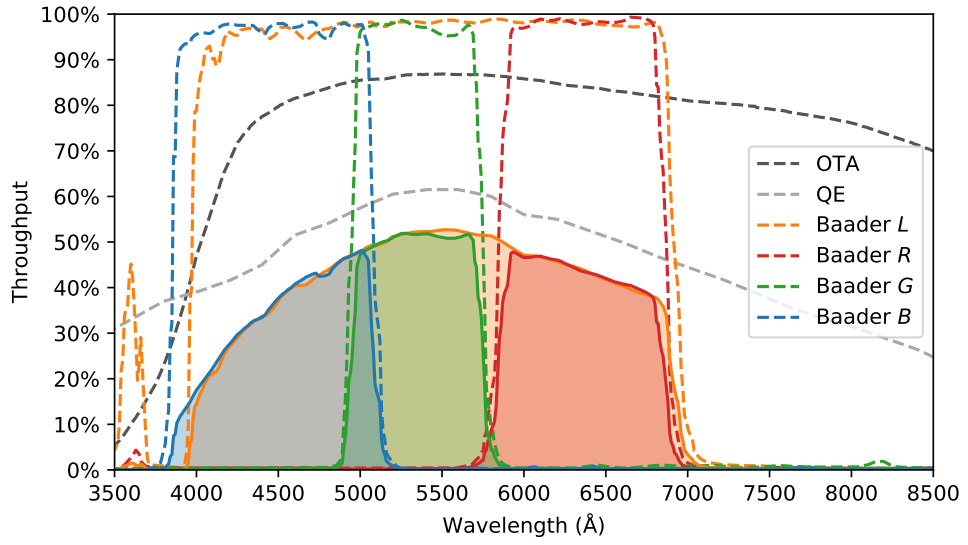


Figure 2.7: Filter response function over given wavelengths from Steeghs et al. (in prep)

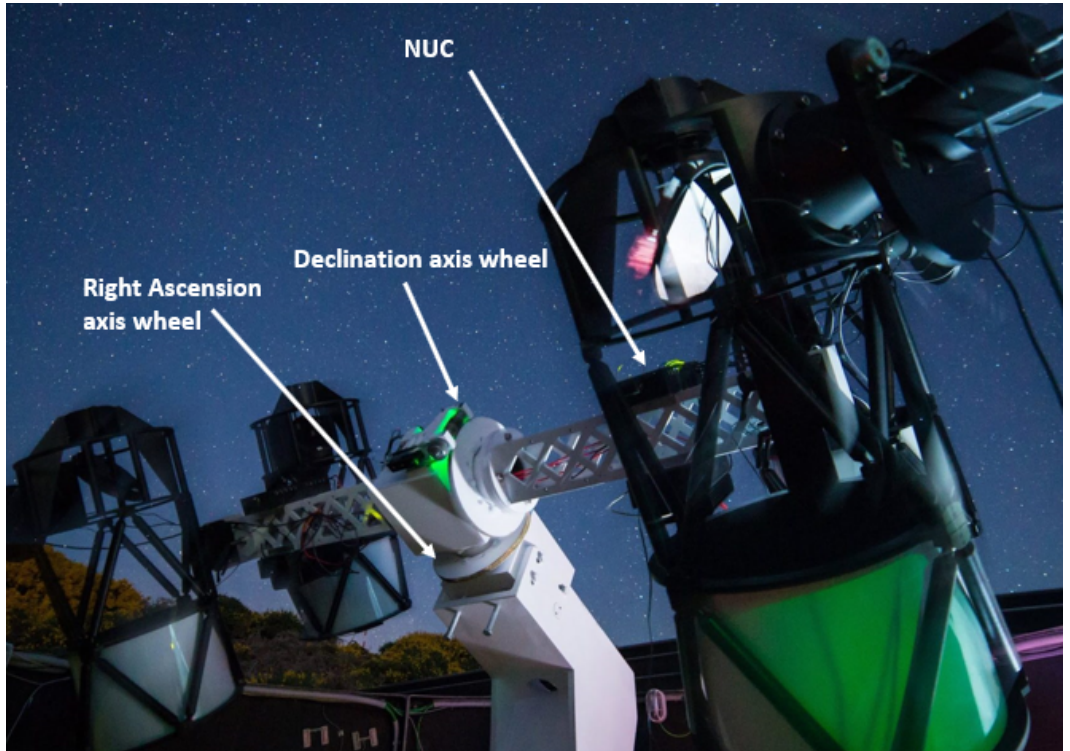


Figure 2.8: Showing GOTO's boom components for axial movement.

2.2.1 The Challenge

GOTO is a unique system. The combination of its fast optics, wide-field, and robotic nature presents a challenge for fast data processing and quality analysis. GOTO takes on the order of Terabytes of data a night, robust methods need to be in place to ensure GOTO is consistently taking science grade observations in real-time. Furthermore, this thesis works on the foundations of the prototype configurations. This adds the additional obstacle of the optical distortions on top of the already challenging wide-fields (see figure 2.9). This thesis will detail the development of GOTO’s data processing stages that are designed to optimise the rapid extraction of transients associated with gravitational waves.

2.2.2 Search Strategy

Most scopes when doing their searches would aim for the highest probability regions within their observable range. Opting for the strategy of looking in the most likely places first (Morokuma et al., 2016; Lipunov et al., 2016). This may be the best option for telescopes that have a narrow field of view. For GOTO, its wide field means more careful strategies can be put into motion. This is done by creating an all sky reference out of repeated fields (the sky grid can be seen in figure 2.6). We do this using GOTO TILE (Dyer et al., 2018). By having an exact sky reference for each part of the sky, we can revisit those tiles to see what has changed without the compromise of using a different part of the telescope to see the same sources. Because the field of view is so large, GOTO can prioritise localisation regions that will fall below the horizon earlier in the evening and still target higher probability regions in the same night. This search method is not feasible with smaller fields of view.

Additionally, by having a fixed tiling grid, GOTO can quickly subtract images using the fixed references. This allows for prompt identification of transients for follow-up telescopes to target (Steeghs et al., 2019a,b,c). The fixed grid also allows fast output metrics of the search; outlining the size of the probable region covered, the likelihood covered, and any interesting candidates flagged during a GW alert. A graphic of this report (figure 2.10) is made, displaying this information. Currently GOTO utilises a greedy tiling method similar to the one discussed in Coughlin et al. (2018); however, given GOTO’s wide field of view and fast optics, this may not be the optimal strategy.

Compared to other telescopes, GOTO handles LIGO’s large localisation regions effectively. Most telescopes can manage a maximum of $\approx 2\%$ (10s of square de-

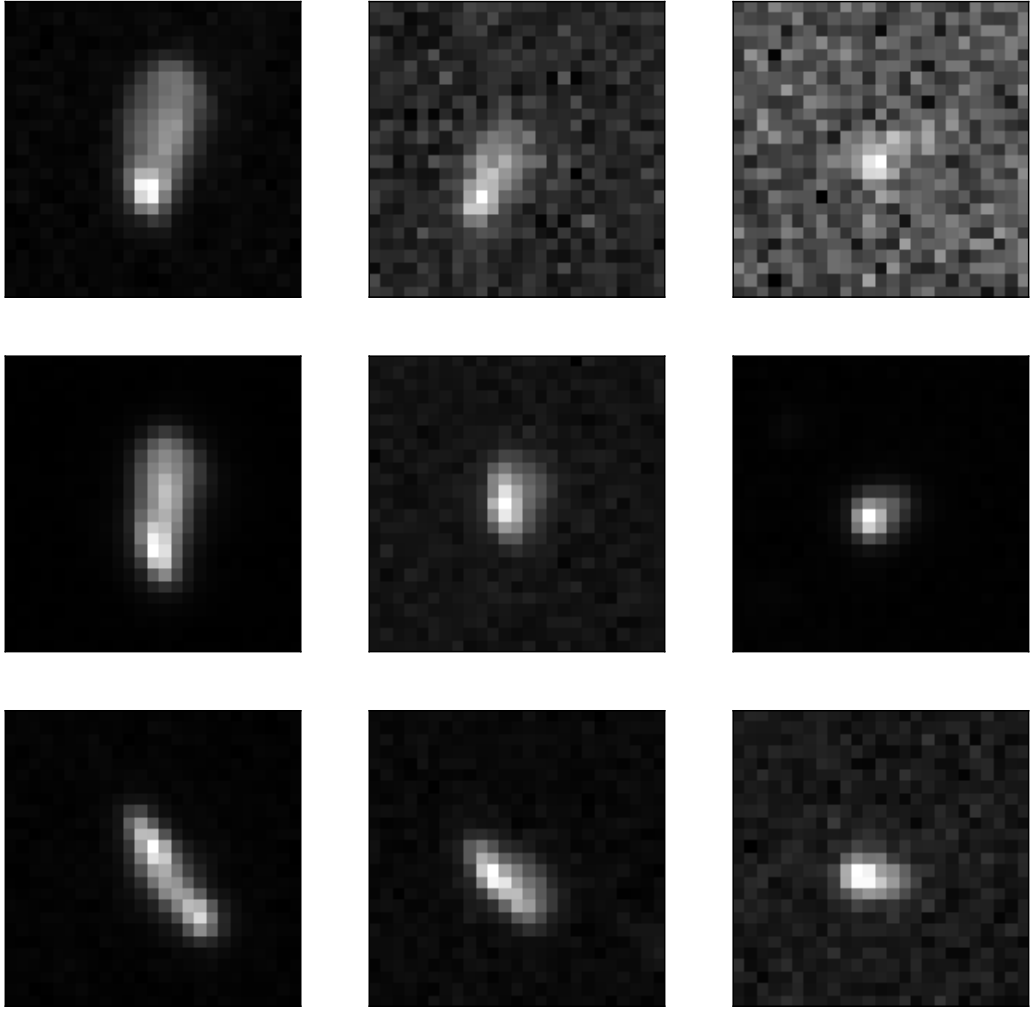


Figure 2.9: The spread of source shapes across a single GOTO-prototype exposure. Each frame is equivalent to the position of where the source is in the image. Highlighting the optical collimation errors.

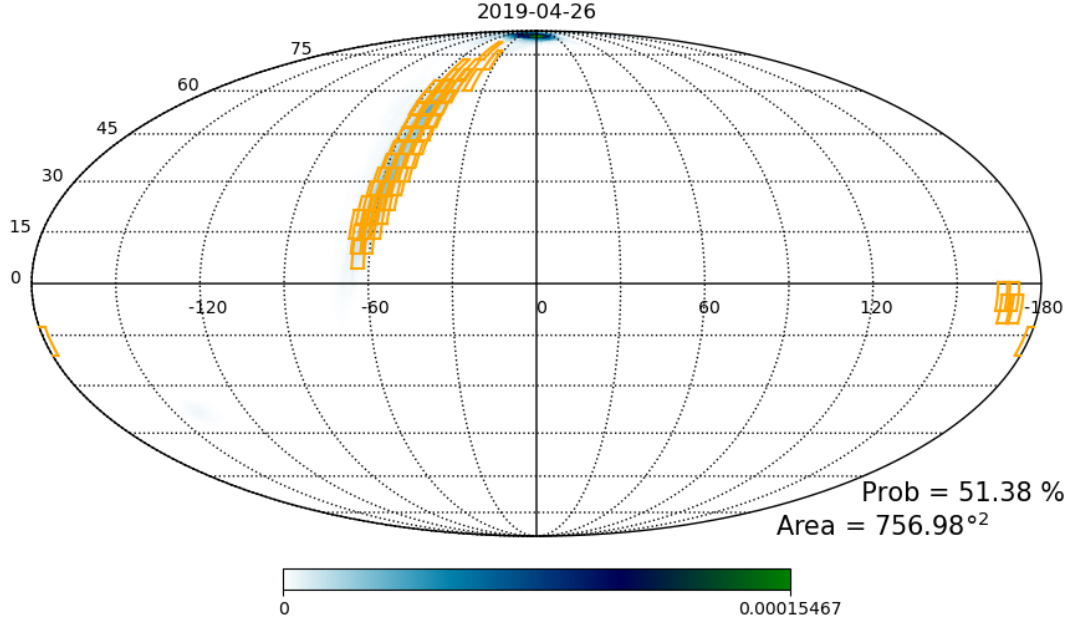


Figure 2.10: Nightly report, highlighting the GOTO tiles and probability covered for S190426c (24237).

degrees) of the error box per night. GOTO has demonstrated multiple times throughout the most recent LIGO campaign that it can survey more than 50% (over 2000 degrees squared) of these regions within the space of two evenings; presenting the potential to pinpoint sky position of GW events much faster than the competition (Gompertz et al., 2020b; ENGRAVE et al., 2020).

If the error region is small enough, sophisticated searches can be done. This is exactly what happened with GW170817. Because the event was only detected in two of the three available detectors, researchers were able to reduce the error region by taking VIRGO’s blind-spot into account. The sky position was homed in further by a detection of γ -rays in both SWIFT and Integral. The resultant 90% confidence region contained only 40 galaxies. By assuming that the most likely place for a transient to occur is atop a galaxy, one can target these high-profile galaxies first. In only four pointings the kilonova was found by the Swope Supernova Survey (SSS) (Coulter et al., 2017). Given the serendipity of a small localisation, smaller scopes with a strong search strategy can compete in finding EM counterparts. The problem, highlighted at the end of 1.1.5, is that usual localisation regions span hundreds of square degrees and this has proven to be an optimistic estimate; with some Neutron Star mergers from the O3 run having confidence regions up to tens of thousands of

degrees (LVC, 2019b).

Filtering processes

A lot of candidate transients are found through the photometric survey process. Many of these will either be false positives or unrelated astrophysical sources (Minor planets, Supernova, Stellar flares, Variable Stars, etc.). Different surveys use different methods to filter out these events.

There are many ways to filter out interlopers. One way that gained traction around a decade ago is citizen science (Lintott and Reed, 2013). This was a useful method when dealing with large amounts of data that could not be vetted by a single person. However, with surveys now finding sources in the millions, even citizen science lacks the resources to filter the data (Bloom et al., 2012). GOTO currently filters out known variables, minor planets, and implements a basic machine learning algorithm to ignore blatant false positives.

As the quantity of data is now beyond the capacity of human processing, the need for machine learning has become necessary to cope with the demands. Many surveys already implement some form of machine learning (du Buisson et al., 2015; Wright et al., 2015; Morii et al., 2016). The basic premise of all the implementations is flagging some features; either of the subtracted image or from the original science frame, and determining how likely the residual is to be a real signal. The intermediate Palomar Transient Factory (iPTF) used multiple cuts to test not only if the source is real, but if the features are indicative of a particular type of transient (Kasliwal et al., 2016) (i.e. occurring on top of a galaxy, and galaxy distance). Chapter 6 explores this further.

Once a transient has been localised, follow-up observations in other wavelengths and with spectroscopy are needed to classify the transient.

2.2.3 Multi-waveband and Spectroscopic follow-up

After the initial detection of a high confidence potential counterpart, more telescopes will join the hunt to obtain detailed light curves and spectral profiles of the event.

Gamma-rays

Gamma rays are made of high energy photons and define the prompt emissions of many transients. These are the photons made at the crescendo of a cataclysmic event. Therefore, gamma-rays signify the birth of a transient and usually offer a glimpse of the most extreme physics (Ackermann et al., 2016; Tavani et al., 2016;

Savchenko et al., 2016). Observations of GW170817 have tied the production of short GRBs to BNS mergers (Wu and MacFadyen, 2019). Gamma-rays often form in collimated jets (they are beamed), this means observing them is down to viewing angle. Therefore, many kilonovae (and other transients) are expected to occur without an observable GRB precursor.

X-rays and Ultraviolet

X-rays and UV rays are wavelengths shorter than visible light (higher energy). A particular challenge with these wavelengths (and gamma-rays) is that they are absorbed by the Earth’s atmosphere before they can be observed. This requires space based observatories. Like gamma-rays, X-rays can be part of the prompt emission phase. They are linked to extreme physics and high temperatures. However, for GW170817 there was a delayed monotonic brightening in the X-rays seen in tandem with the radio observations; this most likely means they have the same physical cause. This is where the primary support for an off-axis jet model and the cocoon model come from (Evans et al., 2016; Haggard et al., 2017).

While not initially detected in the X-rays, GW170817 did have early UV emission. This can inform astronomers of a variety of physical features of the explosion, from viewing angle to ejecta composition. It has been stated that while the observations were useful in testing some models, faster follow-up is needed in both the optical and ultraviolet wavelengths (Arcavi, 2018).

Radio

Radio observations are used predominantly for late time follow-up (Broderick et al., 2020). However and especially for GW counterparts, broadband characterisation is necessary to have a comprehensive understanding of the physics happening (Palliyaguru et al., 2016). Radio observations happen between 3 kilohertz and about 900 gigahertz (100 kilometre - 0.1 millimetre wavelengths). As stated in the X-ray section, radio emissions are just as crucial in understanding the physics as the high energy radiation.

Spectroscopy

Taking spectra of stars and explosions is a foundation of modern astronomy. It is used to classify every source in the night sky. This is how astronomers can determine redshift and composition of stars. For GW170817 like events, spectral

information can be used to determine elemental abundances and kinematics of the merger explosions (Shappee et al., 2017).

Host galaxy properties are also an important part of follow-up. Using the multi-band spectra, one can estimate the age, star formation rate, mass, metallicity, and stellar properties of a galaxy. This in combination with where in the galaxy a transient occurs can help uncover the conditions required to make the different types of transients (Chrimes et al., 2018; Artale et al., 2019).

For optical telescopes like GOTO, CCD astronomy is the standard. The practice of scientific astronomy with CCD telescopes is well established and has a wealth of literature behind it (Berry and Burnell, 2000), and GOTO presents many opportunities to exploit this rich history.

2.3 CCD Astronomy

2.3.1 Standard Astronomical Practices

When astronomical data is recorded, it is commonly saved to a Flexible Image Transport System (FITS) file (Wells and Greisen, 1979). These files exist to compactly store and move image data easily. The structure of these files contains a primary header and some accompanying primary data. The primary header should contain information regarding the size of the image array, when it was taken, and by which telescope. Additional information can be stored in the headers. This can be anything, here are some examples: sky position, object of interest, exposure time, or weather conditions. The primary data works by turning each CCD pixel into a component of an output array. That is, the first pixel on the fifth row will have the array position 1,5 and its value will correspond to the readout value of the CCD pixel. The astronomy community has built a python package, called *astropy*, specifically designed to aid astronomers. One particular package `astropy.io.fits` is used to read these fits arrays and headers so they can be handled easily (Robitaille et al., 2013). *Astropy* is used frequently throughout this thesis.

One of the main standards in astronomy is image reduction (Howell, 2006; Wall and Jenkins, 2012). This is the concept that all images contain noise characteristics that can be modelled or predicted in some way. The main sources of noise for GOTO are:

- Photon (shot) noise
- Readout noise

- Thermal noise (also called dark current)
- CCD response (Quantum Efficiency)

Considering the above, one can estimate the signal to noise ratio of a pixel. Shot noise is Poissonian, meaning the noise generated by the signal is exactly the root of the signal. $\delta_{shot} = \sqrt{\chi_e D_p}$, where χ_e is the quantum efficiency and D_p is the number of incoming photons.

Dark current count rate (D_{rate}) is measured in electrons per pixel per second. To find the noise, one must multiply the rate by the exposure time (t). Following that, the noise can be found as the root of the Dark current signal: $\delta_{dark} = \sqrt{D_{rate} t}$. Unlike the other values readout noise is pseudo-static i.e. does not vary with exposure time and therefore just has the standard value δ_r .

Noise (N) is considered the Euclidean distance of all the noise combined: $N = \sqrt{\delta_{shot}^2 + \delta_{dark}^2 + \delta_r^2}$. The signal to noise can be found by taking the signal and dividing by this noise term. Giving the final result:

$$\frac{S}{N} = \frac{\chi_e D_p}{\sqrt{\chi_e D_p + D_{rate} t + \delta_r^2}} \quad (2.1)$$

By using image processing techniques, one can minimise the influence of these noise sources on the raw data.

Readout noise

When a photon hits a CCD pixel an electron is displaced and recorded as a voltage. To record this, the voltage is passed across the CCD to a readout monitor. This is done with additional voltage, which influences the readout voltage. Fortunately, this noise is relatively static for each pixel and independent of photon count, meaning the noise can be modelled. This is done with a zero second exposure (with a closed shutter). The resultant image is called a Bias image, and contains the readout noise value per pixel (typical readout noise for GOTO is ~ 12 e⁻ with a bias level of ~ 1000 ADU). GOTO takes 10 of these and median combines them to create a Master Bias image (figure 2.11). It should be noted, each GOTO sensor has two readout channels, which results in a slight discrepancy in the readout noise on either half of the image.

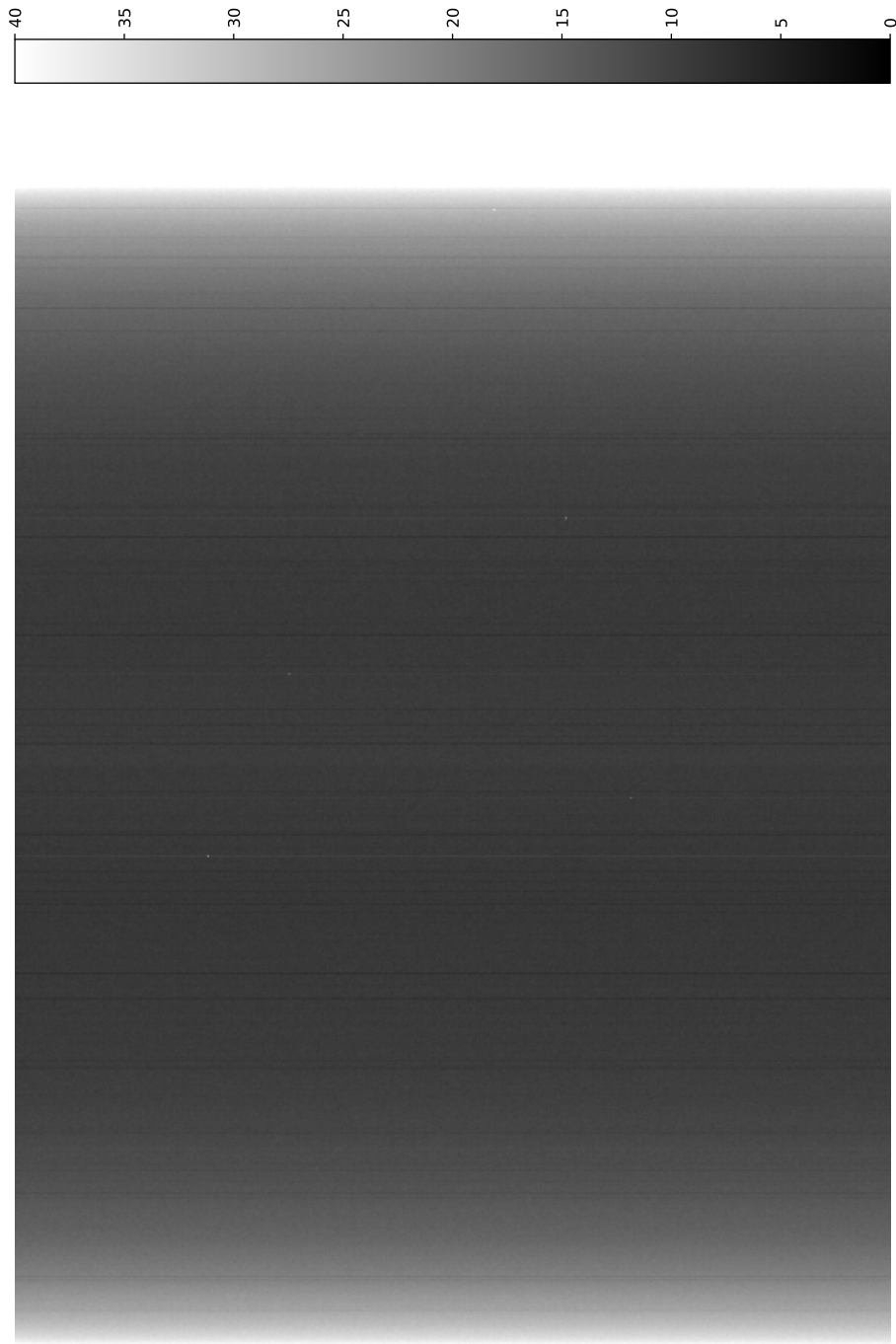


Figure 2.11: Example of a GOTO master bias image. This is made from the stack of multiple zero second exposures and is used to account for read-out noise.

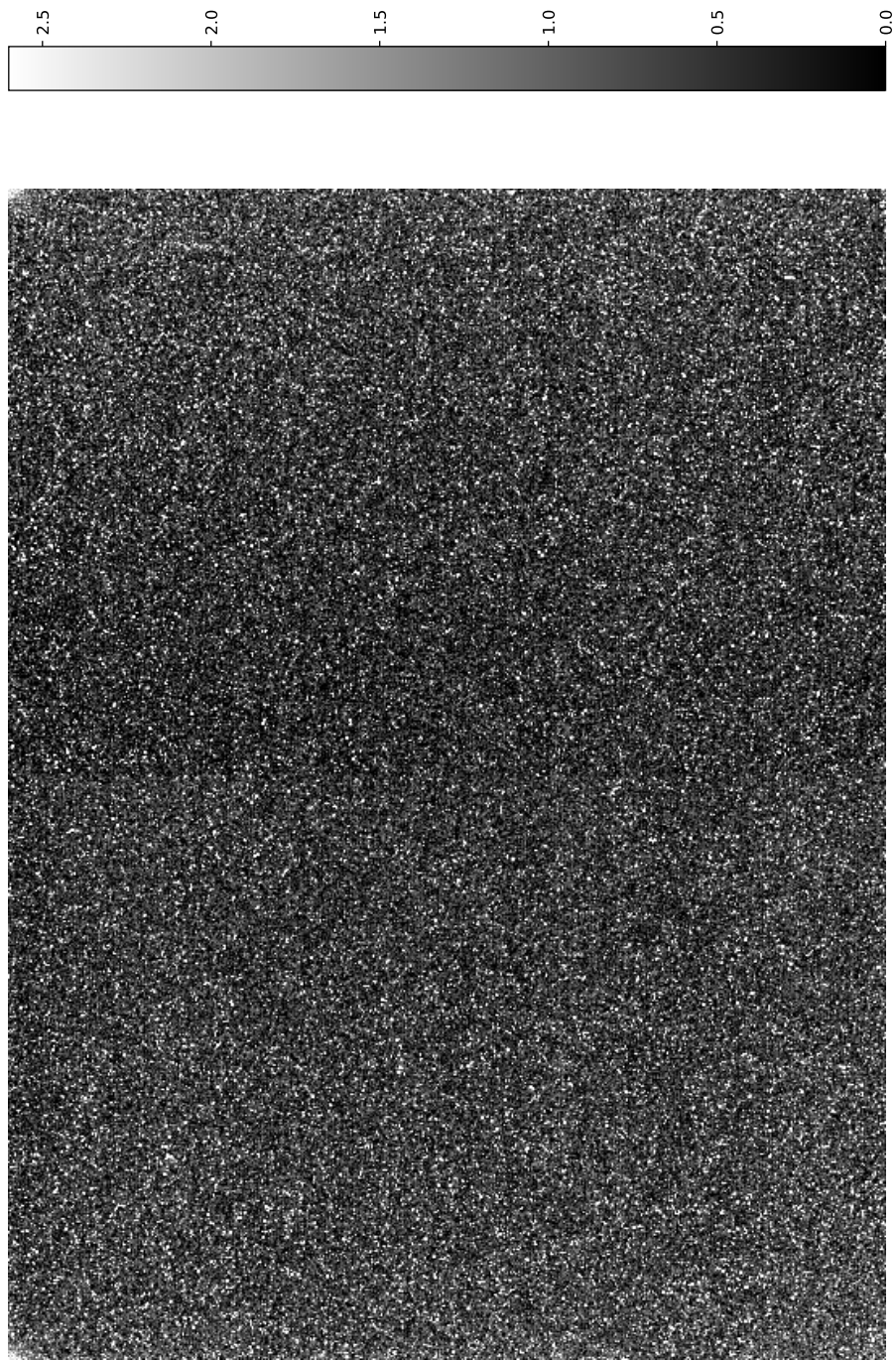


Figure 2.12: Example of a GOTO master dark image. This is made by stacking multiple exposures taken with the lens closed to model the thermal noise of the image.

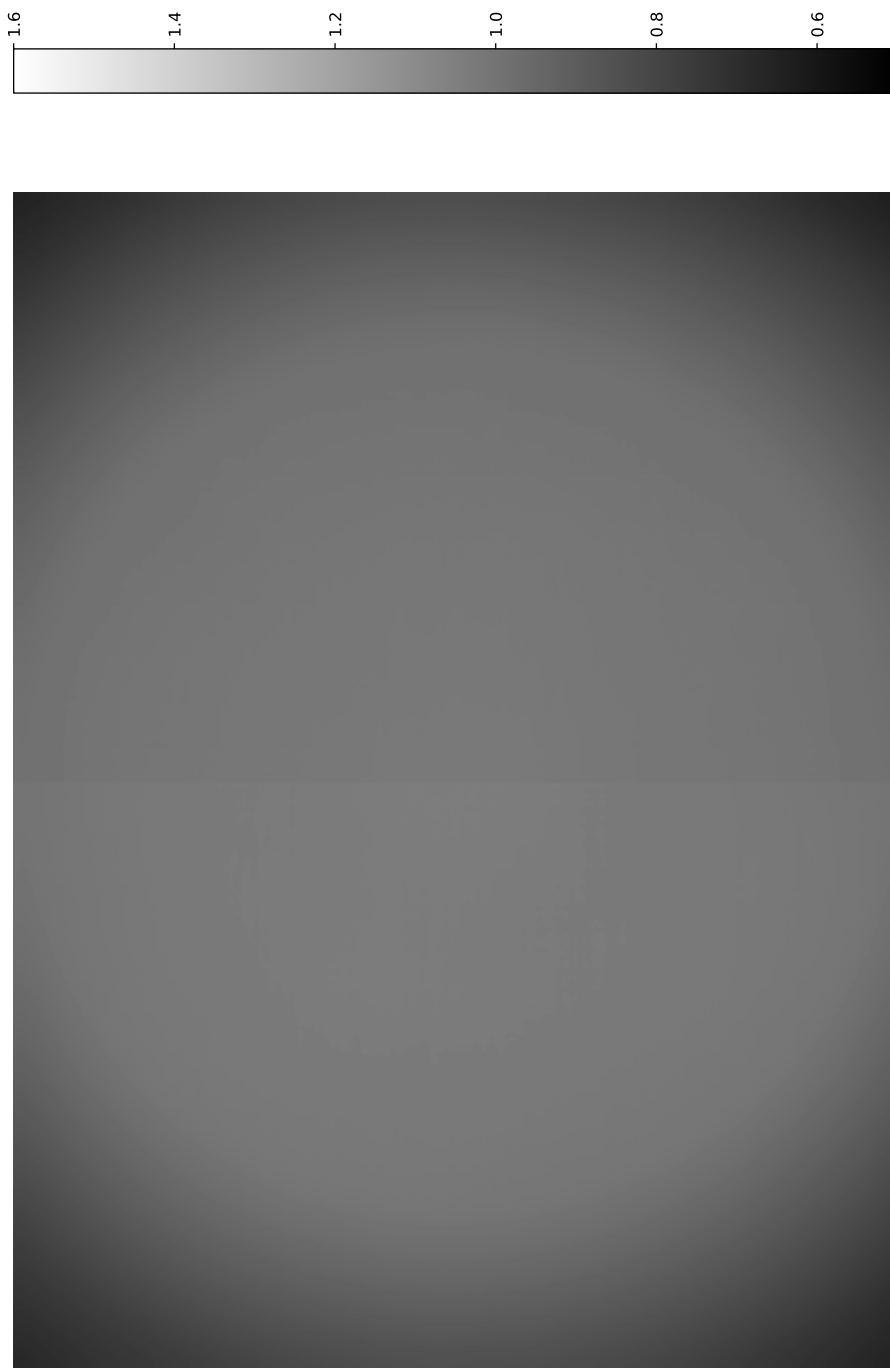


Figure 2.13: Example of a GOTO flat field. This is made by stacking multiple flat fields. Flat fields are the normalised response function of each CCD pixel, so given a field of uniform brightness the flat field will correct all the pixels so the exposure is uniform.

Thermal noise

Thermal noise is the result of a stored current due to thermal fluctuations that occur in the CCD pixel during an exposure. Similar to readout noise, thermal noise can also be modelled. This is done by taking an exposure that matches the exposure time of the images, but the lens cap is kept on. In GOTO's case the shutters are kept closed and the dome is shut. This means the only recorded signal received will be due to readout noise and Dark Current, and the readout noise has already been determined. These images are referred to as Darks. Again, a Master Dark, shown in figure 2.12, is created by median combining 10 Darks. To minimise the influence of dark current, the GOTO cameras are fan-cooled to -20°C . Typical dark current values are roughly $0.0015 \text{ e}^{-}\text{s}^{-1}$ (see Dyer (2020) for more details).

CCD response and Vignetting

CCDs are not perfect, in that they do not record every photon that lands on them. This efficiency can be represented as a percentage of photons that are recorded divided by the total number of photons. However, it is not effective to measure the quantum efficiency of every pixel. Instead it is easier to compare each pixel to every other pixel in the image and normalise them. This is done using a flat field (figure 2.13). A flat field is a bright uniform field. The field needs to be bright so every pixel has a high number of counts (i.e. the field is not dominated by other noise factors), uniformity means that every pixel should have the same value. GOTO takes flats by taking an exposure of the twilight sky. Twilight offers the best flat field for a wide field of view telescope, but a sky gradient needs to be accounted for (Chromey and Hasselbacher, 1996). The image is then normalised by dividing every pixel by the brightest pixel. Each pixel now contains the fraction of photons that will be detected compared to the most sensitive pixel. Dividing the flat field by the normalised image (called the Master Flat) should necessarily return an image with uniform pixels, this correction by division is the same used when reducing data.

The design of the Newtonian telescope allows the the secondary mirror to fit the beam produced by an on-axis source. However, a fraction of the beam from sources off-axis will not hit the secondary mirror. Therefore off-axis sources will appear dimmer, this is accounted for in the flat field seen by the vignetting in figure 2.13.

Image Reduction

Now all the calibration images are in place, the raw science image can be reduced. Each pixel can be found using:

$$Science = \frac{(Raw - MasterDark - MasterBias)}{MasterFlat} \quad (2.2)$$

This only produces a single reduced frame. While science can be done on this frame there are going to be artifacts and other imperfections in the image. To remove these, the combination of multiple science frames can be used. This not only removes artifacts, but increases the depth of the image. These benefits are discussed in more detail later in chapter 4.1

Finding the Apparent Magnitudes

Astronomy is often regarded as the oldest science, this is reflected in the standard unit of measurement, magnitude. The magnitude system is most commonly credited to Hipparchus of Nicaea, a Greek scholar, where the brightest stars were given a magnitude of 1 and the faintest a magnitude of 6 (Toomer et al., 1984). This ancient practice was formalised in 1856 by astronomer Norman Robert Pogson (Pogson, 1856), making each magnitude a 2.512 times brighter than the next. Albeit a not entirely intuitive system, this archaic way of measuring stellar brightness has persisted.

Most, if not all telescopes have a different flux sensitivity. This means, unless calibrated to one another, the results from separate telescopes cannot be combined or compared. The best way to calibrate a GOTO image is to compare bright, low variable sources in an image to those found in recent data-releases. GOTO originally used 2MASS and Gaia DR2 as reference catalogues (Cutri et al., 2003; Gaia-Collaboration et al., 2018); however, the most useful reference catalogue has been REFCAT2 (Tonry et al., 2018). To achieve calibration, sources with matching sky positions in both catalogues are compared on the recorded apparent magnitude and measured flux. Using the following, a zero-point for the data can be determined:

$$m = -2.5 \log_{10}(F) + Z \quad (2.3)$$

$$\therefore Z = 2.5 \log_{10}(F_{zero}) \quad (2.4)$$

Where m is the apparent magnitude in the reference catalogue (for a true calibration source $m = 0$) and F is the flux of the same source from the science data. This can

be used to estimate the apparent magnitudes of new sources.

$$m_{new} = 2.5 \log_{10} \left(\frac{F_{zero}}{F_{new}} \right) \quad (2.5)$$

This however, requires two pieces of information currently undefined. Sky position of each source in the field and their fluxes. Finding these manually for a large field would take a tremendous amount of time. Data processing methods have been introduced to speed up and reduce labour when mining these values from the data.

2.3.2 Making Sense of the Data

Source Extraction

Source extraction is a critical tool in survey astronomy. Accurately isolating sources in an image is important not only for science, but as I will demonstrate later, for diagnosing and improving telescope parameters. There have been multiple source extraction tools developed (Sun and Zhao, 2013; Makovoz and Marleau, 2005; Savage and Oliver, 2007), however the utilities most regularly implemented for GOTO is SEXTRACTOR (Bertin and Arnouts, 1996) and its python implementation SEP (Barbary, 2016). Source analysis is dependent on the source shape. Point sources are detected using peak finding algorithms, this is looking for a Gaussian based distribution of flux and selecting the brightest pixel within that distribution as the source centre. The second method is for low surface brightness objects, primarily galaxies. Thresholding, is the concept of defining a background limit; anything above that limit is considered a detection. Once SEXTRACTOR has completed extraction, a catalogue is produced detailing the source number found, its xy position on the CCD, and any other parameters the user has asked to be collected; for example, the source flux. Source flux is calculated by assigning the source an aperture, usually calculated by the Gaussian profile of the source, placing the aperture over the source centroid, and summing the value of the pixels inside the aperture. The aperture is usually circular, however they can be any shape depending on the Point Spread Function of the source. Determining the flux count alone is often not enough when looking at sources in the image, understanding the background is vital too.

Background Estimation

A significant feature of source detection is background estimation, especially for thresholding. There are a multitude of approaches when tackling background esti-

mation (Almoznino et al., 1993; Bijaoui, 1980; Blanton et al., 2011; Newell, 1983). However, the method most commonly used and regarded the most stable is the one discussed in Da Costa (1992) and implemented in SEXTRACTOR. This details an iterative process called κ - σ clipping and mode estimation. The κ - σ clipping method is used to ignore outlying pixels. Using three terms: the number of iterations, the standard deviation (σ) and the multiplier (κ). The pixel value is calculated and compared to κ multiplied by σ . Any pixel with a magnitude higher than $\kappa \cdot \sigma$ is considered to deviate too strongly and is rejected from the background estimation. This process is repeated until either the estimate reaches 3σ certainty or the σ deviates by less than 20%. Mode estimation is then used on the selected pixels from the above method. This is done by finding the mean and the median to estimate the mode, hence mode estimation. Da Costa (1992) demonstrates that with κ - σ clipping, this estimation is a slightly altered form of the usual mode estimation:

$$mode = 2.5 \times median - 1.5 \times mean \quad (2.6)$$

This offers a positional background estimate of an image. Finally, a bilinear interpolation is used to smooth the gradient of the background, creating a background map. Aside from thresholding, background estimation can be used to remove the background from images. Which is seen in chapter 5. Typically, background flux for GOTO varies dramatically (between a flux of 200 to 500) as seen in figure 2.14.

Forced photometry is used when trying to get the flux of a specific position in an image, usually of a known source from another survey; or to find the flux of an interesting source at an earlier time or in a different wavelength filter. To obtain the flux an aperture is used. The aperture is defined as a circle with a given radius of pixels that contains the flux of a source, assuming a point source. Anything contained within the aperture is summed to give the flux, similar to automatic aperture flux method from earlier. However, when looking at a single source the background can be estimated by taking the flux of the same source using a wider aperture and subtracting the flux from the original aperture. This will give a count of the local background flux around the source.



Figure 2.14: An example of background estimation on a GOTO field

Astrometry

The night sky can be treated as a celestial sphere, which means it can have a co-ordinate system (latitude and longitude or right-ascension and declination). This map is especially vital to astronomers as it tells us which part of the sky we need to look at to target specific stars. Astrometry is the process of finding source positions in the sky precisely and predicting their motions. The most common way of assigning images sky co-ordinates is through the relation between source positions in the image (Lang et al., 2010). The first step is to determine source positions in image co-ordinates, this is computed using a source extraction step like the ones discussed in the previous section. The second part is to group stars together in sets of four, these sets are referred to as 'quads'. Trial parameters of the quads are taken; parameters such as separation, separation angles, and relative positions to other quads. These parameters are then compared to a reference grid containing similar quads.

A Bayesian analysis is presented, verifying if a hypothesis grid matches the input parameters. If the hypothesis grid passes the Bayesian test, a World Coordinate System (WCS) is applied to the input image. A WCS is a set of instructions in the image header detailing the orientation of the image relative to the sky grid, and how many arcseconds each pixel is worth in Right Ascension (RA) and Declination (DEC) (Greisen and Calabretta, 2002). More sophisticated models can account for curvature of the optics in the image. The WCS is used to map sources in an image to their sky position in a global co-ordinate system, the data now becomes useful for other astronomers. There are further uses to the WCS that will be discussed later in the thesis.

Image registration

Image registration is similar to astrometry in the sense that one needs to locate sources across multiple images. Instead of needing to relate the stars to a co-ordinate system, image registration is to make sure the same sources fall on the same pixel in the image co-ordinates. There are many ways images can be aligned to one another. These methods are discussed later in chapter 4.2. The primary purpose for image registration in this work is for image subtraction, which is used for finding changes in the field. Image registration is not to be mistaken for mosaicing which requires image registration, but combines images to create a larger image. Figure 2.15 shows an example of a GOTO mosaic.



Figure 2.15: Example GOTO mosaic, the horsehead nebula ¹

Co-Addition

Image Co-Addition, or stacking, is the process of combining multiple frames of the same patch of sky to lower uncorrelated noise and remove artefacts. This works by using a statistic, usually the median, of three or more frames. The benefits and uses are discussed in more detail in chapter 4.3

Science from the Data

The data have now been processed and can be used to do science (Izzo et al., 2019; Borisov et al., 2018; Abbott et al., 2017c). However, having usable data alone is not enough to achieve the desired science goals. Section 2.3.3 is going to detail how GOTO conducts its search to find a counterpart and how this compares to its competitors.

2.3.3 Hunting for a Gravitational Wave Counterpart

GOTO’s prototype consists of four 40cm telescopes, each with ≈ 5 square degrees field of view. In optimal conditions these scopes can see down to 21st magnitude, with 120 second exposures.

A main feature of GOTO is its fast optics, which is how it achieves its wide field of view. Furthermore, GOTO takes short exposures with a relatively deep limiting magnitude. Included with this is a short slew time, going from one position in the sky to another; boasting speeds of ~ 4 degrees per second. Hardware with these specifications can have a bespoke search strategy that exploits these features. GOTO is designed to fulfil the rapid localisation needs discussed in at the beginning of this chapter.

Finding counterparts quickly requires more than just fast hardware. A fast image processing pipeline is needed to reduce the data and extract potential transient candidates quickly.

Image Subtraction Pipelines

Image subtraction is a tool used when hunting for changes in the sky. By subtracting all the stars in the field with a reference image, the residuals left will be sources that either were not there previously or changed in brightness. By looking exclusively for new sources, one can quickly locate transients. Detailed discussion of the methodology and implementation of image subtraction is discussed later in

¹goto-observatory.org/photos/photo/horsehead-nebula/

chapter 5. By having an automated pipeline images can be reduced and subtracted quickly; then, using a source extractor tool on the subtracted image one can find the residuals in the image.

By implementing all of the standard practices discussed above, GOTO is able to achieve real-time follow-up of GW alerts.

2.4 Summary of Methodology

This chapter has introduced GOTO, a wide-field astronomical survey for the rapid identification of GW counterparts. I explained the hardware specifications and why they are appropriate for GOTO's needs. This section also presents the challenge of GOTO's collimation errors, which will make a regular occurrence throughout the thesis. I then detailed the standard practices used in scientific survey astronomy and related these to GOTO. The following chapters detail the automation of some of these practices, all culminating to rapid and automatic discovery of transients in the GOTO GW follow-up survey.

Chapter 3

Getting Focused

“ A policewoman comes up and tells us to move along. ‘Nothing to see here’, she says. ‘Not even that stuff you’re looking at.’ ”

— Lincoln Michel, *Upright Beasts*

Historically, at the start of each night telescopes will find optimal focus and iterate over the night as focus shifts, usually due to temperature changes. GOTO is not much different. Due to the fast nature of the GOTO optics, focuser position is critical to data quality and overall performance. Unfortunately, each UT has its own characteristics; meaning one cannot just use the same base values for each telescope. When GOTO first began commissioning data, the issue of obtaining focus automatically became an apparent problem. Figure 3.1 highlights this, showing optimal focus was routinely missed by the initial autofocus implementation and focus needed to be found manually.

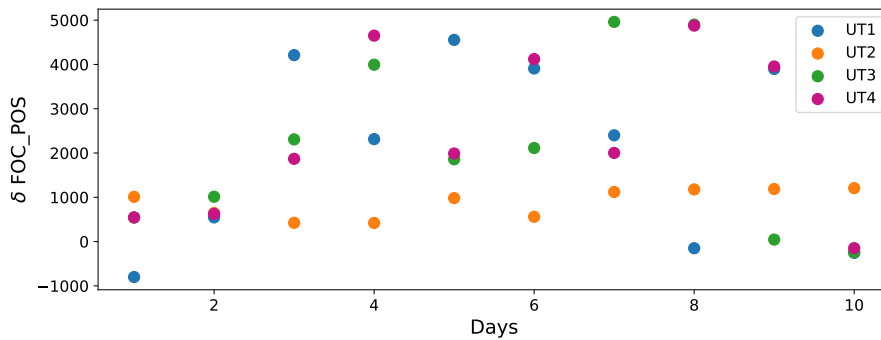


Figure 3.1: Difference in **AutoFocus** optimal solution, to best focus found manually across multiple nights. Highlighting that optimal focus was routinely missed with the original focus algorithm.

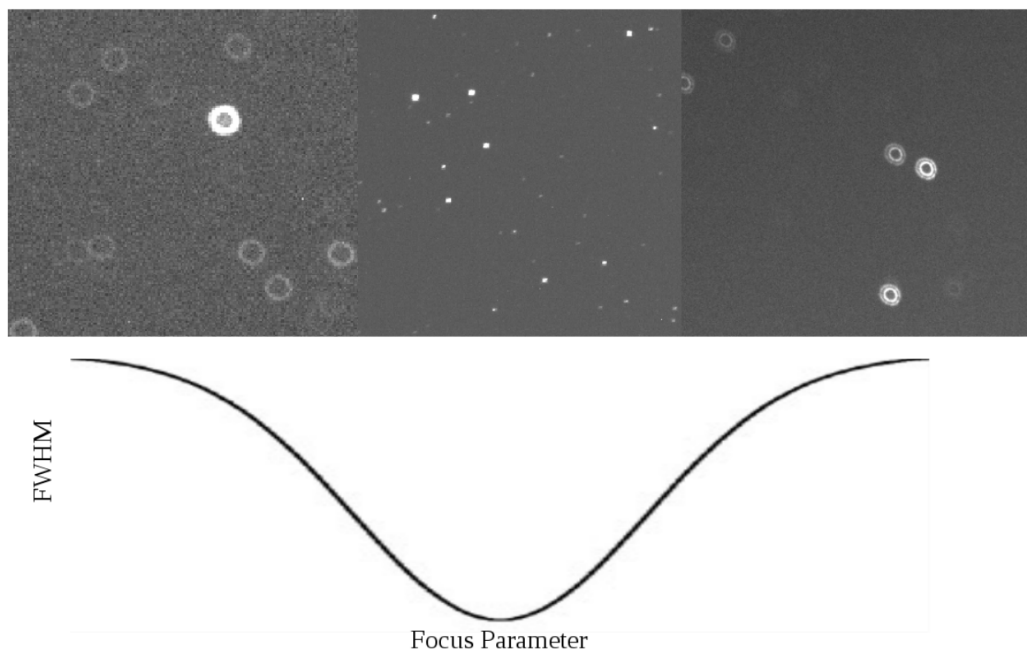


Figure 3.2: Example of FWHM as a function of focuser position. Optimal focus corresponds to the minimum in FWHM, which is shown to be close to the point where sources appear point-like.

This chapter is going to describe the importance of getting focus for telescopes, then highlights the pitfalls of GOTO’s original methodology when finding focus. By exploiting those pitfalls, I describe a new way of obtaining focus automatically and describe future improvements for the, current, working implementation.

3.1 Background

Obtaining focus is a critical part of astronomy, but is often overlooked as most telescopes have an astronomer at hand to monitor the image quality. However, GOTO and many other robotic telescopes are designed so that they will operate independently, without oversight. This means the subtle art of finding focus by eye has to be expressed mathematically, using measurable parameters. When determining focus, astronomers are adjusting a focuser looking for the moment point sources look most ‘point-like’ (see fig 3.2). This is essentially a non-parametric way of estimating the point spread function (PSF). The reason astronomers choose the most ‘point-like’ PSF is because this will contain the highest signal to noise ratio. Note, this chapter works on the premise that a circular point PSF has the highest signal

to noise; however, there can be solutions where elliptical PSFs contain the highest signal to noise.

Focusers operate by incrementally tuning the optics of the telescope, usually the primary mirror. For GOTO however, the focuser adjusts the detector position until the optics align at the point of optimal focus. To understand the scale of the problem, one is trying to find the point (within $\pm 15\mu\text{m}$) where a meter long beam meets focus. The reason refocusing is required is that the tubes will flex and expand depending on the conditions. The frequency of changing focus is telescope dependent; some set-ups require multiple adjustments per night, others need to be set weekly.

Obtaining focus automatically is a relatively under published area. Most scopes will iterate focus position and use the observed minimum Full Width Half Maximum (FWHM) as the point of optimal focus (fig 3.2). Weber and Brady (2001, hereafter WB1) found that the FWHM was an unreliable metric when obtaining focus (because the PSF is not Gaussian when out of focus) and instead opt for the Half-Flux Diameter (HFD). This describes the point at which the inner circle, with a given diameter, contains exactly half the flux of a source. WB1 conclude that HFD can be modelled either side of focus with a linear fit against focuser position, and that one can use the intercept of these two fits to find focus. The fits are made by modelling the HFD of one bright source and iterating across different focuses.

GOTO's initial implementation used the fundamental model presented in WB1, but made a few additional assumptions to lower time dedicated to achieving focus. Once the HFD has been modelled as a function of focuser position, a linear fit is made ($y = mx + c$). If one assumes no astigmatism (which is not unreasonable), only a single point past the minimum needs to be found as the gradient of one line will be the inverse equivalent of the other (i.e. $m_1 = -m_2$). Using this relation, one can employ the intersect of the two linear fits to find the optimal focus parameter:

$$F_{\text{optimal}} = \frac{c_2 - c_1}{2m_1} \quad (3.1)$$

Looking at 3.3, the two linear models are used to predict the optimal focus position. Further variations of this altered method have been trialled. One of these is looking at the minimum HFD for each focus position. This works on the assumption that the smallest HFD is the most pristine; in other words, has been tampered the least by seeing conditions and wind-shake. This is dangerous as the seeing conditions (especially at the beginning of the night) can fluctuate dramatically, and the minimum of just three exposures will most likely result in a poor fit. This technique is still better than GOTO's original method of taking a single exposure

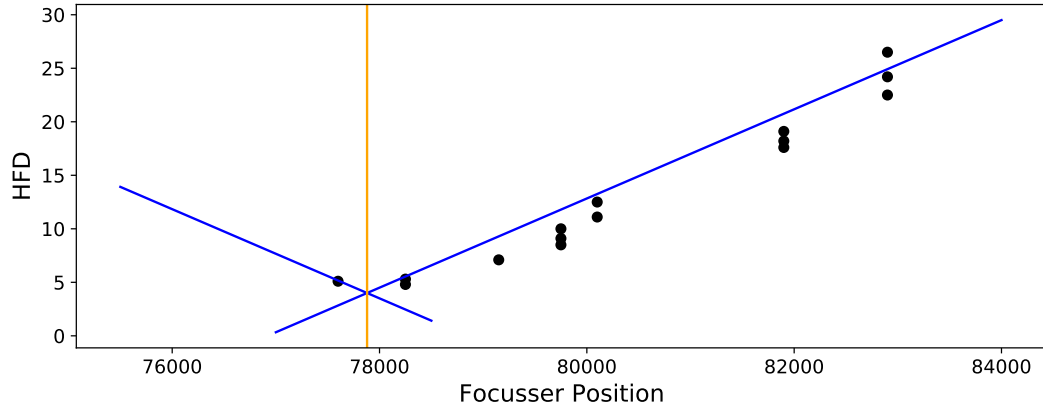


Figure 3.3: Example of using two linear fits to obtain optimal focus with GOTO data (from equation 3.1). The blue lines show the linear fits, the orange line is their intercept and the location of the predicted optimal focus (described by equation 3.1).

per focuser position.

Additionally, there have been projects highlighting the behaviour of the HFD near focus (La Cour, 2012). These results are flagged in the GOTO data also, showing that as one approaches focus the linear approximation breaks down. The near focus issue raises a few concerns:

- Firstly, sampling frequency. The original (and current) method of finding focus uses fixed step sizes when moving the focuser. This is not an inherent problem away from focus, but under-sampling near focus will result in the non-linear regime contaminating the linear fit.
- Secondly, for scopes like the ones in GOTO, the optimal focuser position can be obscured by noise in the HFD. That is, the resolution of the system is dominated by other errors (such as wind-shake and early evening seeing fluctuations). This means instead of having a single best focus, there are a range of focuser positions that provide the same level of accuracy. Optimistically, one would assume that optimal focus would lie in the mid-point of this range; however, given that the GOTO-prototype optics are astigmatic there is a strong chance this is not the case.

Astigmatism is the result of the optical best focus not aligning precisely in both of the optical planes. This results in a point source being stretched in the axis of the out of focus plane. A great visualisation tool can be found in ¹. The astigmatism

¹www.olympus-lifescience.com/en/microscope-resource/primer/java/aberrations/astigmatism/

is unique for each scope and its effect varies across the image. More importantly for this discussion, the HFD's gradient is not symmetric either side of focus. This means the logic applied in equation 3.1 is fundamentally flawed for this system of telescopes.

Given the nature of the GOTO optics, a different strategy on obtaining focus needs to be developed. The following details the ideas and subsequent implementation of the new method developed.

3.2 A New Approach to Obtaining Focus

3.2.1 General Requirements

When developing a script to find focus, a few considerations should be made to keep the routines practical and rigorous. Obviously, a script to find focus should reliably find focus; however, if the script, for any reason, cannot obtain focus there needs to be a safety net which informs observers that optimal focus could not be found and returns the script to a pre-defined working focus. Furthermore, a comprehensive log to provide a way of debugging where the malfunction occurred must be available.

GOTO is an ever evolving array of telescopes. Having more and different telescopes added to the system means the script needs to be flexible to these new extensions. Thus, if telescopes are replaced or a set of optics have considerably different focus parameters, the script will not need to be redeveloped from the ground up.

Another issue that has not been addressed to this point is exposure time. Originally, the exposure time when doing a focus run was 30 seconds. Ideally, the exposure time needs to be as short as possible. This will avoid added noise from wind-shake and seeing fluctuations. However, one must take care to avoid too short an exposure time as the source count needs to be filled adequately to determine source shape on the chip (30 seconds is too long). Later in this chapter I will demonstrate how to find the optimal exposure time for GOTO when conducting a focus run. Furthermore, longer exposure times for a focus run eats deeper into observing time. Shortening the exposures to what is absolutely necessary increases the amount of science that can be accomplished. Another way to increase the efficiency of the focus step is to process the data while it is being taken. At the moment, **AutoFocus** does not utilise the multiple cores GOTO's data acquisition machine has available, this can add up to an additional 20 seconds per image. Threading will be introduced in section 3.3.

3.2.2 Accounting for Telescope Personalities

Each telescope has its own imperfections and characteristics. How they respond to different sources of noise is therefore dependent on the individual and a universal focus model cannot be assumed. Figure 3.4, shows the nightly focus evolution of each UT. There is a clear shared response in the focus quality across the images; however, finding what that dependence is will require some investigation. Understanding what is disrupting the HFD is important as this could affect a routine finding optimal focus.

The biggest influence on focus stability, for most telescopes, is temperature. Many telescopes are cooled to stabilise temperature fluctuations. GOTO has active cooling on the CCDs to keep the cameras at a constant temperature, but does not condition the rest of the optics. Therefore, it is vital to see how the focus characteristics are effected by temperature changes. Figure 3.5 shows the focus response for each UT as a function of temperature. Again, the scopes show a shared behaviour, but it appears the noise introduced by temperature fluctuations is smaller than other sources of noise.

Wind shake is another notable source of uncertainty. This happens when the mount is blown by winds. Because of the UTs cylindrical structure, they are able to catch the force of the wind, which adds a torque on the mount arm. This actively increases tracking error. Because the scopes share a common mount and all have similar noise properties, there is a strong likelihood the mount is involved in the shared uncertainties. Tracking errors act to spread a point source across multiple pixels. An exaggerated example of this, usually done for artistic effect, is source streaking. Having a still camera expose on the moving sky. This will necessarily cause the HFD to extend due to the source being spread across more pixels. In Figure 3.6 I show there is a slight correlation in wind speed and HFD.

There was also concern that the seeing conditions are responsible for focus performance. Cloud opacity is the measure of percentage light scattered and absorbed so it cannot be observed. It would be expected that higher cloud coverage would result in more scattering, resulting in a larger HFD. After finding no correlation between cloud cover and HFD, one can conclude this factor does not have a substantial effect on focus. One interesting consideration was Moon effects. Moon brightness increases sky background, this could make faint sources appear smaller than their PSF dictates. As GOTO uses long exposure times and avoids looking directly at the moon, this is not a big issue when observing. It should be noted when using shorter exposure times, like when finding focus, read-out noise or bias could become a more influential factor due to low counts.

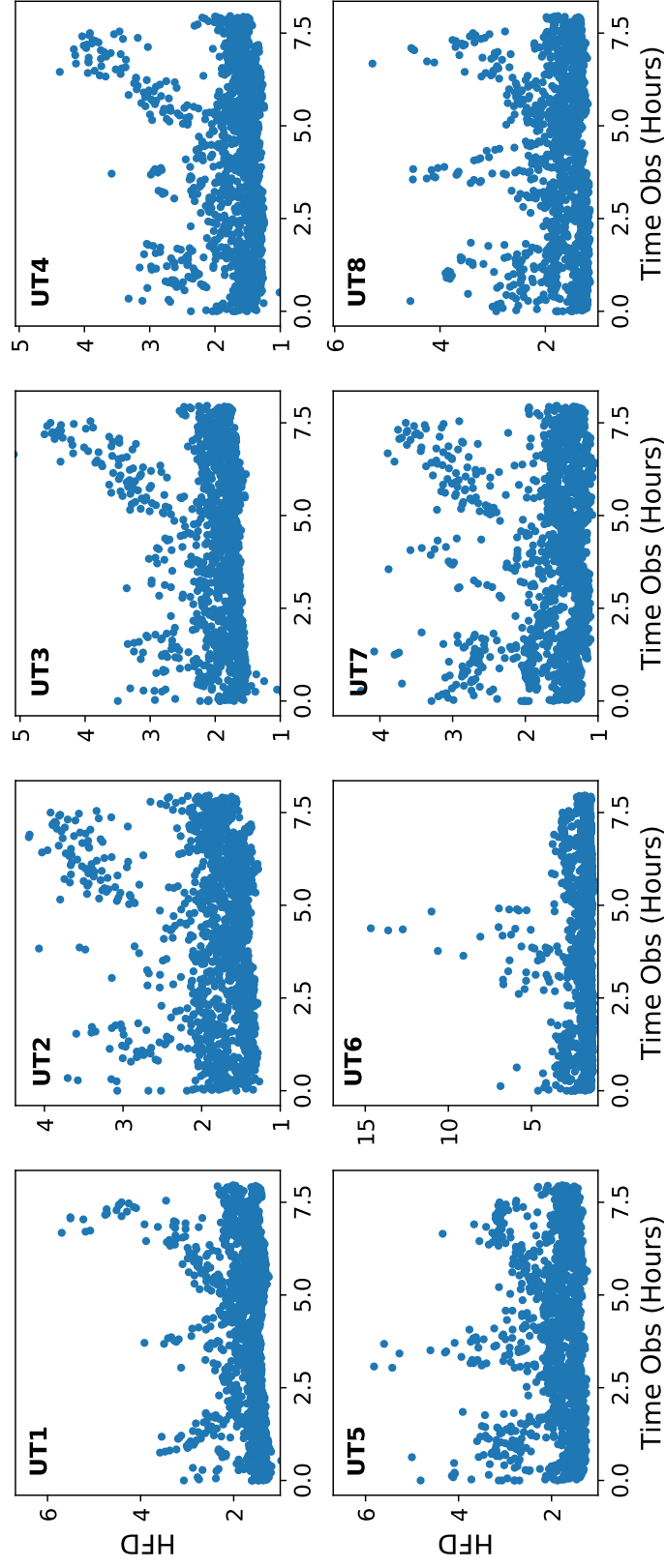


Figure 3.4: Focus evolution over the course of continuous observation (different nights are overlaid on top of one another). There is a shared response across all of the UTs ~ 5 hours, where all the HFD increases. This means there is an environmental influence acting on all the scopes.

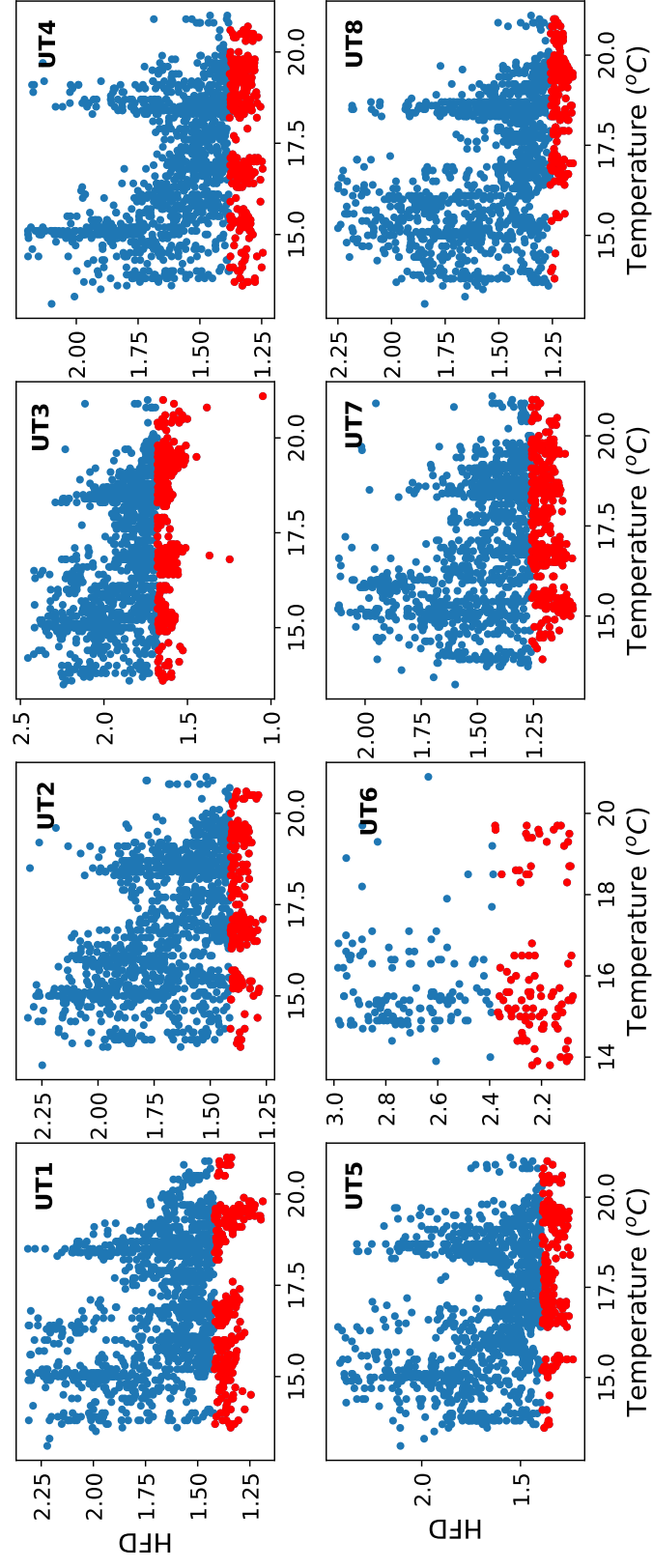


Figure 3.5: Focus response to temperature. The red points indicate the lowest 10% HFD points in that temperature bin. The red points show a flat line, this indicates temperature does not have a significant affect on HFD

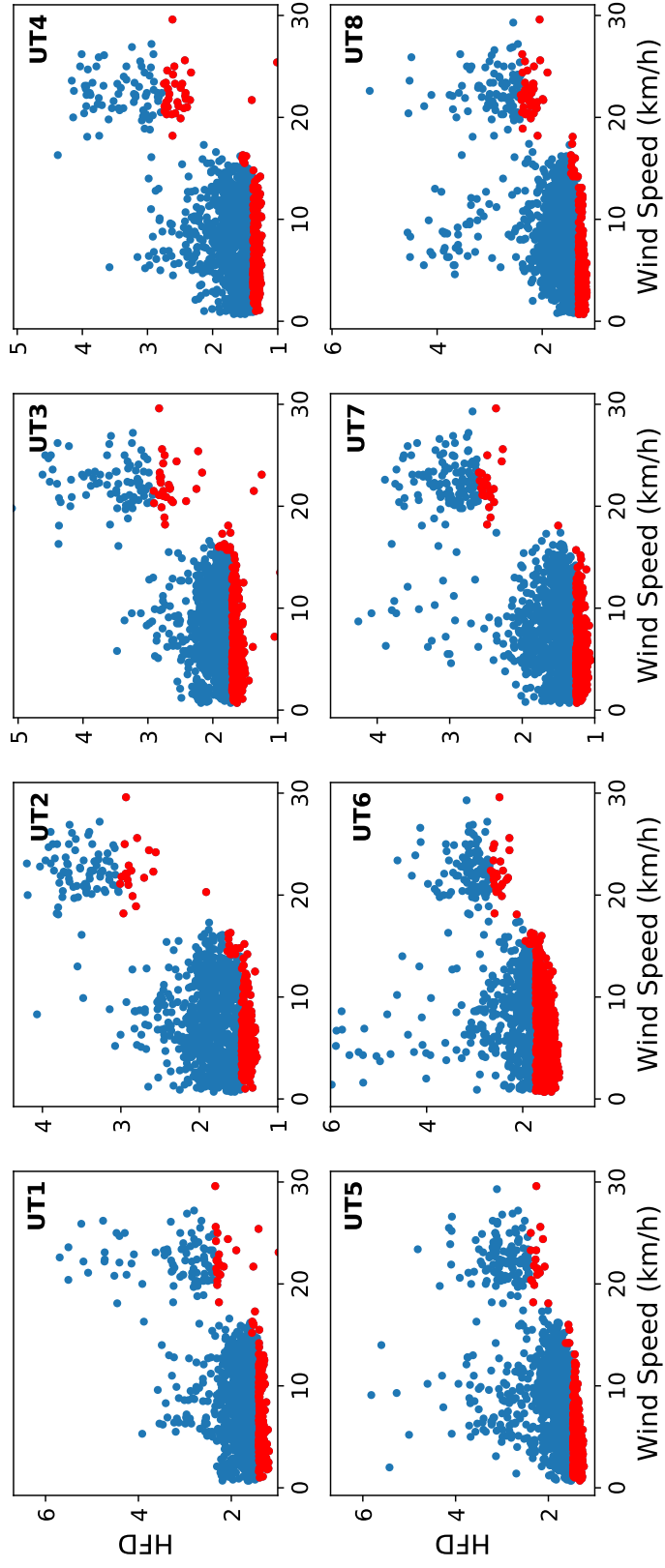


Figure 3.6: Focus response to wind speeds. The red points indicate the lowest 10% HFD points in that wind speed bin. The red points show a step at $\sim 20\text{km/h}$, this indicates wind speed has an affect on the HFD

Airmass is another kind of seeing condition; this is how much air light has to pass to reach the observer. The thinnest airmass is directly up from the observer, Zenith. Initially, there appears to be some HFD response to airmass; however, this behaviour is not consistent with what was expected. This is likely due to most observations happening at relatively low and high airmasses, leaving a small sample of mid-range airmasses that appear to have a smaller PSF. Airmass is also an indirect measure of mount position. This could mean the origin of the HFD behaviour is tied to physical components as opposed to environmental conditions.

It was suggested from the relationship of the HFD with airmass that focus was not based on what was obstructing view, but a mechanical fault. Immediate disquiet was met with this speculation, as this most likely meant a component of the optics was not fixed in place and was moving at particular mount angles. More measurements were taken, showing there is no HFD relationship with mount position.

Given that airmass and wind were the only two variables that seemed connected to focus performance, the next step was to compare the two. Figure 3.7 highlights, almost damningly, that wind speed is the culprit for poor focus. Figure 3.7 exhibits the same profile as was seen between the HFD and airmass; furthermore, above 30 kilometres per hour, every scope has a larger HFD. This means when getting focus and observing, it is vital to monitor wind speeds to assure data quality.

This section has shown that modelling optimal focus parameters based on environmental variables is incredibly difficult. This is primarily because wind shake acts as the dominant noise source and drowns out the noise introduced from temperature and the other conditions. I have also shown that wind shake will be the biggest obstacle when finding focus. Therefore, the optimal focus strategy will need to minimise the amount of time spent exposing to minimise the influence of wind shake. Furthermore, as scope position has a negligible effect on focus, it makes sense to face zenith to minimise seeing errors. The last point to consider is background noise. Wind shake requires the exposure time be as short as possible, whereas background requires a minimum exposure time. Thus, one challenge will be to find the minimum exposure time necessary to overcome background noise.

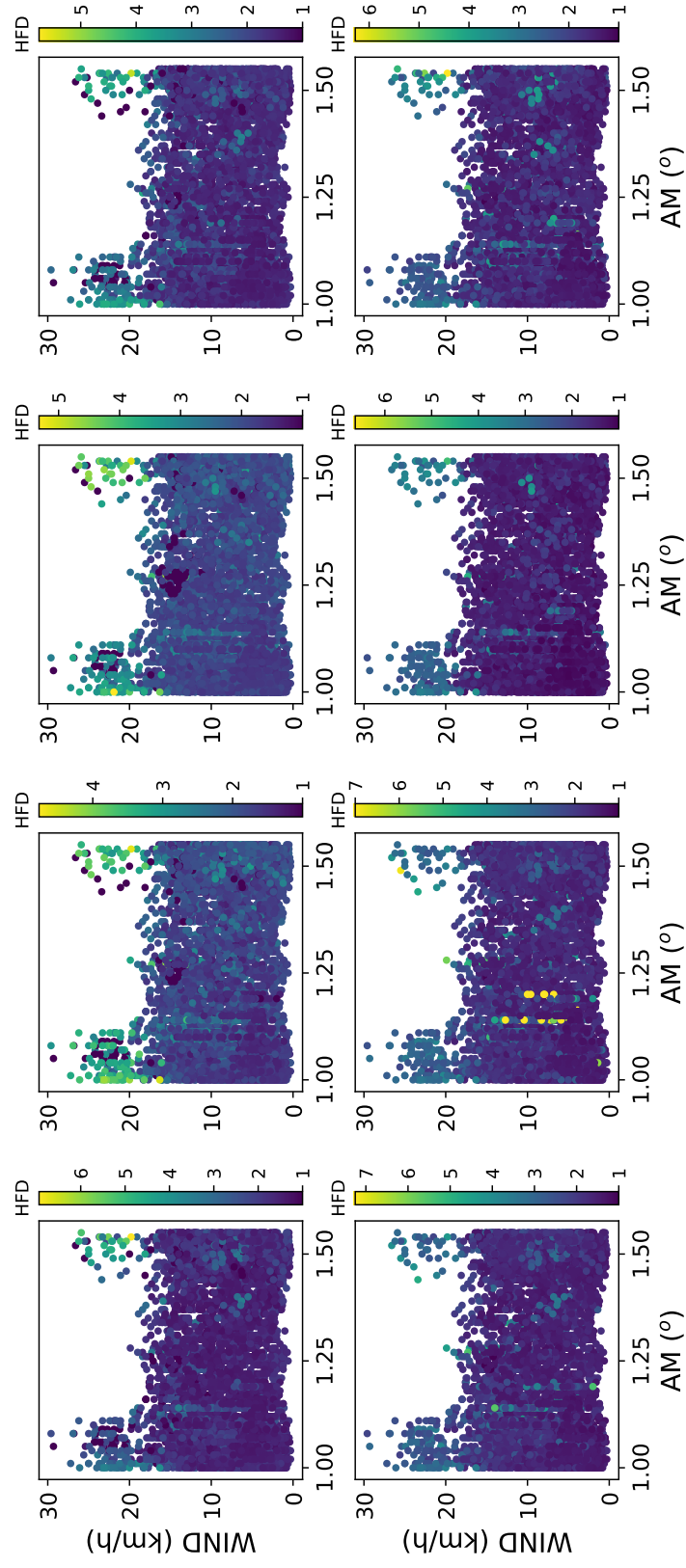


Figure 3.7: Focus response to wind speed and airmass. The HFD has no real dependence on airmass, dismissing the concern that telescope position may be responsible for HFD variation. This plot also reaffirms the correlation to HFD and wind speed.

3.2.3 Exploiting the Astigmatism

Luna et al. (2007) demonstrates that when out of focus, astigmatic point sources express an exaggerated geometry and a preferential rotation. So in addition to using the HFD, geometric properties of the out of focus sources can be used to determine focus.

The two metrics used here to determine source geometry are the rotation (θ) and elongation (ϵ). Where θ is the angle of the sources semi-major axis (A) from the primary axis in the image, and elongation is defined:

$$\epsilon = \frac{A}{B} \quad (3.2)$$

where B is the semi-minor axis of the source (See fig 3.8).

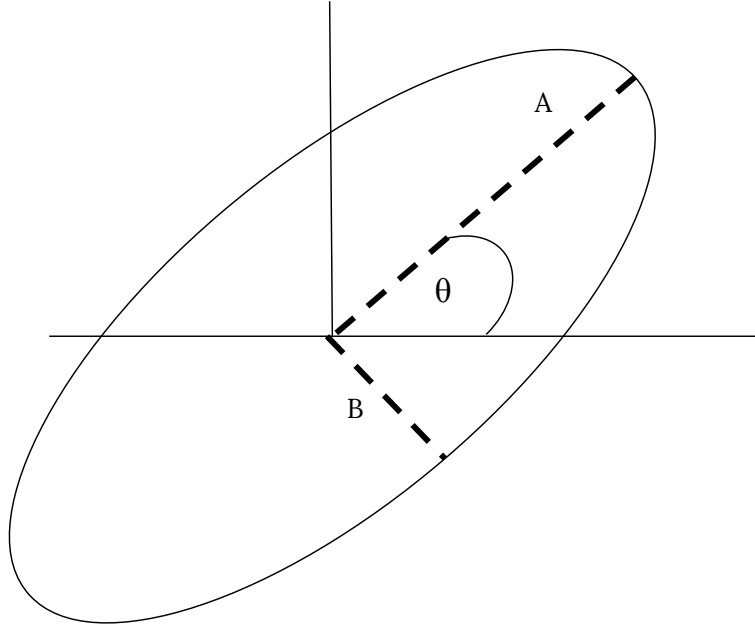


Figure 3.8: Visualisation of Elongation and source rotation

As the optics approach optimal focus, ϵ^{-1} tends to 1. As unity is the maximum value of one over elongation, the function of this parameter should trace a reverse parabola as one sweeps through focus. The expected function for θ is a

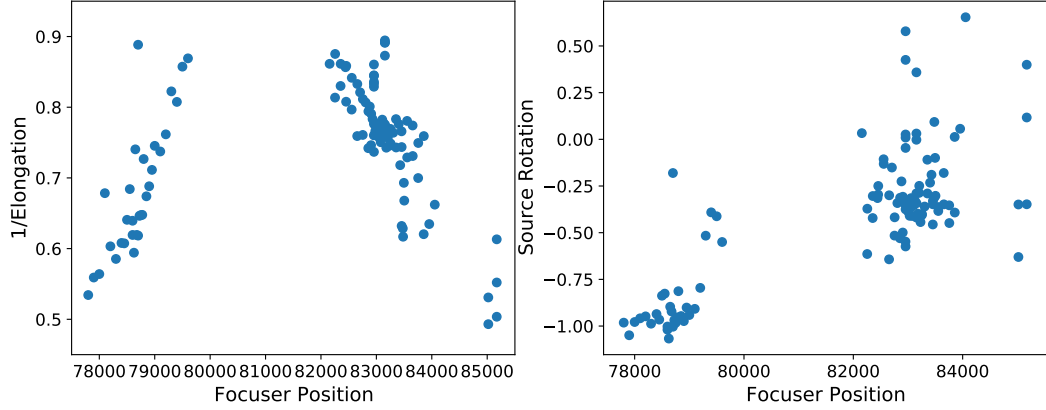


Figure 3.9: Relationship of Theta and Elongation as a function of focuser position. The inverse of the elongation traces a negative parabola, with the peak at optimal focus. Source rotation demonstrating a step function, where that step occurs is also considered optimal focus.

little more complex. Theta is determined through the asymmetry of the source; if the source is not circular theta has a magnitude, otherwise the rotation equals 0. This effectively makes source rotation a step-function of focus. Where focus is the point at which the step occurs. Figure 3.9 highlights initial findings, showing these relationships, albeit tentatively. In the following section, these relationships will be utilised to identify optimal focus.

3.2.4 Initial Performance

Using the HFD effectively

Earlier in this chapter I described the method used in WB1 and the way it was originally applied to GOTO. The two primary conflicts were the astigmatism not being accounted for, and the non-linear behaviour of the HFD near focus.

To force the WB1 method to consider astigmatism, one simply has to model both linear components either side of optimal focus. The rest is the same as the original, using the intersect of the two linear fits:

$$F_{optimal} = \frac{c_1 - c_2}{m_2 - m_1} \quad (3.3)$$

Navigating the non-linear regime requires a little more inspection. The linearity is broken due to the noise becoming a dominant factor in determining the HFD. By ignoring the noise dominated samples, the linear regime is conserved and

the method does not need to increase the sampling frequency unnecessarily. The challenge here is determining at what point noise influence becomes non-negligible. For now, the upper limit of this influence is considered double the median standard deviation of the HFD; though more statistical rigour is required for absolute precision. Any points in the distribution below the upper limit are masked when modelling the fit. The error on this measurement can be determined from the intersect of the uncertainties from the two models.

Elongation

Figure 3.9 shows the expected relationship between the inverse elongation and the focus position. This can be fitted using a simple second order polynomial. The maximum of this fit is the point of optimal focus. As with any model, there will be an associated uncertainty. I obtain the uncertainty in the maximum by taking the residual of the fit and summing it with the fit. The bounds are defined as the focus position where the sum of the fit and residual are equal to or smaller than that of the fit's maximum.

Rotation

As theta is predicted to behave as a step function, the first action is to find the closest representation of a step function to model the data. This is a sigmoid function:

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (3.4)$$

Instead of modelling the step function directly, I aim to predict the region in which the step occurs. Conveniently, this can also be modelled with a sigmoid fit. Using a coefficient sweep method, I model the function multiple times and taking the intercept of these models I can predict best focus. The maximum and minimum difference from these models are used to determine the uncertainty range of the model's prediction.

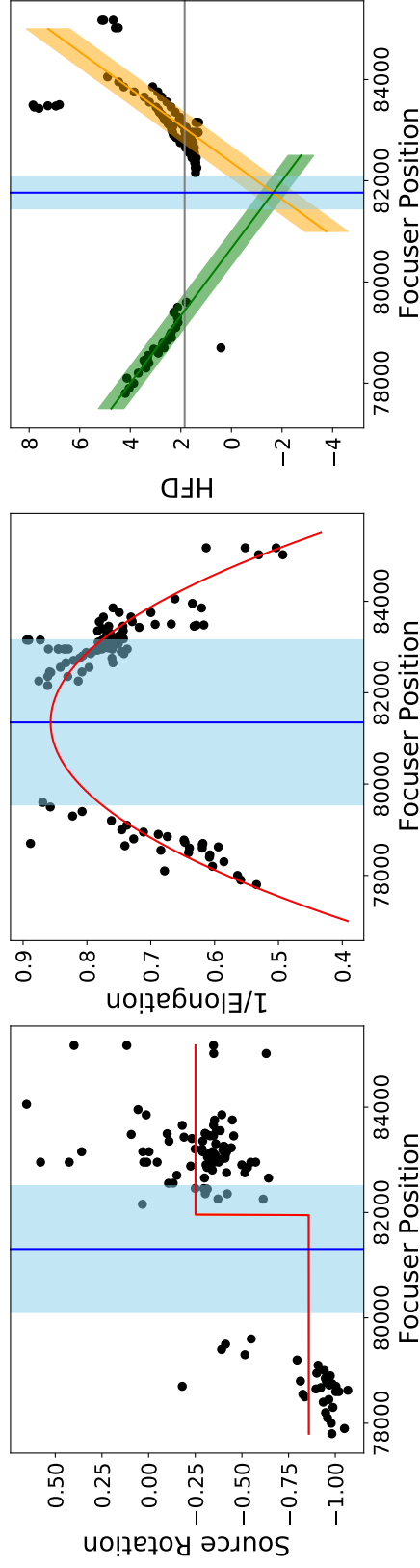


Figure 3.10: Finding focus three different ways. The left most plot shows finding focus using source rotation, where the red line shows one of the sigmoid fits. The middle plot shows elongation finding focus, the red line here is the negative parabola fit. The plot on the right shows optimal focus found by the intersection of two fits. The green and orange line represent the linear fits either side of the uncertainty line (indicated by the grey bar). Optimal focus on each plot is indicated by the vertical blue line and the uncertainty by the transparent blue region.

Initial Tests

Figure 3.10 shows the above three determinations in practice. The key take away is that the HFD method has the lowest uncertainty by quite a large margin. However, this uncertainty range still spans more than 500 focuser units (the average astronomer for GOTO has a can find optimal focus within $\sim \pm 250$ units). Part of the challenge presented in WB1 was to automatically obtain focus of equal or better quality than manual acquisition, that goal should be the same here. By having three predictions of where best focus should be, I can combine them to make a joint prediction. To do this, I model the uncertainty region as a Gaussian probability density function (PDF) and using the product of multiple Gaussian distributions (Smith, 2011) I create a cumulative prediction. There is a slight oversimplification in this step as I am not accounting for correlation in the three predictions, but this should have a much smaller influence than the anticipated level of precision. Figure 3.11 shows the different predictions using the PDF combination technique. For the example given, the true, manually obtained, best focus value is known. Table 3.1 contains the performance metrics of each method.

Method	Optimal Focus	Uncertainty	Distance
Manual	81600	± 250	0
HFD	81765	± 318	165
Source Rotation	81268	± 708	332
Elongation	81351	± 902	249
Gaussian Combination	81651	± 276	51
Astigma Combination	81299	± 557	301

Table 3.1: Focus trial performance metrics. This shows that the most confident (and closest) method is the Gaussian combination of all three methods from figure 3.10. See figure 3.11 for more details.

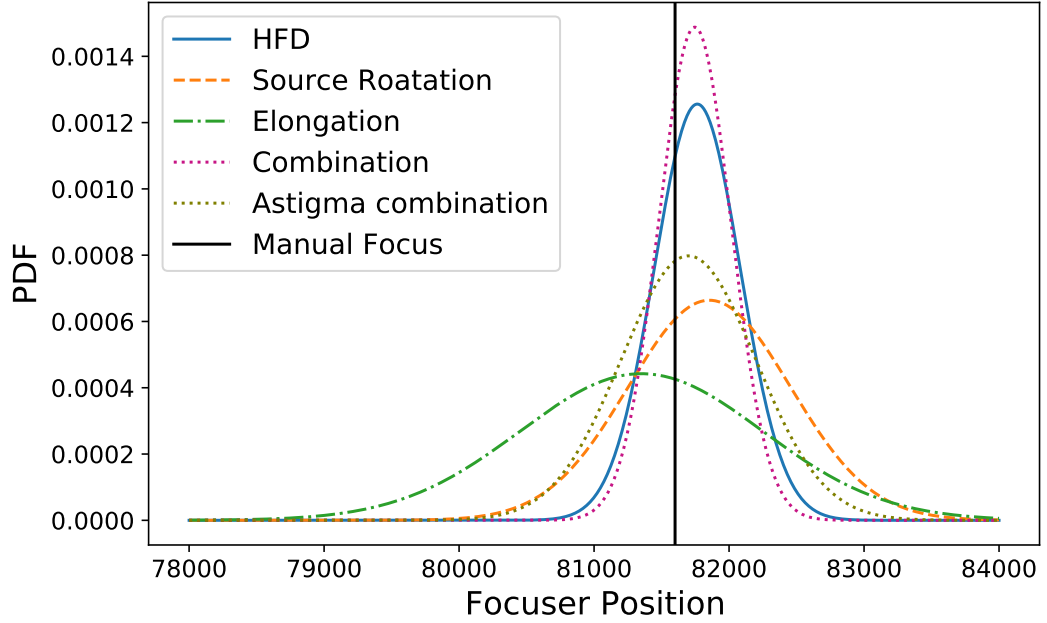


Figure 3.11: Using Gaussian probability combination to improve optimal focus estimation. The combination of all the methods (shown as the purple dotted line) has the highest confidence and has the closest prediction to the focus found manually (the vertical black line).

In this situation, the combination of the three predictions improved the overall performance of the autofocus. Unfortunately, the prediction is still not as certain as a human finding focus. The combination of just the methods exploiting the astigmatism was also made to see how well the new techniques performed. Interestingly, the accuracy of the elongation method deteriorated when merged with the source rotation method; however, they both gained significant improvement in precision; though the astigmatism methods alone are still a over factor 2 from being as precise as a human observer. Given that the data used to make these predictions were taken with sub-optimal exposure parameters, one could expect significant improvement in both accuracy and precision with a more methodical approach. Working under the assumption that a combination method can provide at least equivalent performance to a human, the next step is to implement it automatically.

3.3 An Automatic Implementation

3.3.1 Threading the proverbial needle

The previous section detailed the requirements and methodology of obtaining focus. While theoretically automatic, in practice implementing the above without oversight requires communication with the telescopes. This step will also address the goals of making the script scale-able and fast. The technique implemented here makes use of the multiple cores the GOTO computers have available.

To begin the process, first one must decide the number of focus positions and how many exposures are required, this method requires an initial estimate of where the user thinks optimal focus is. Given that focus is found on a nightly basis, it is safe to assume the focus found the previous night is sufficient. From this point no user input is required. The above can be programmed to begin using a cron job, this means predefined inputs can be used; so no human involvement is necessary.

Once the script is initiated, the mount will be put into ‘tracking mode’, this is to prevent sources from streaking. Nothing of importance happens during the first set of exposures as there is no data to process. When the exposure finishes, the program changes the focus position for the next exposure and starts the loop. As the second exposure is started, a thread is used on the latent core (it will be inactive waiting for the camera to respond). This is done for each UT, this is the point that allows for scaling as the implementation can work for up to as many telescopes as there are cores. The statistics of the previous image are taken. Recording the HFD, rotation, Elogation, Focuser Position, and a few other parameters specific to each telescope. This process repeats on a loop until the last exposure.

After the designated set of exposures are completed, the csv files are read by another script which processes the data for each telescope using the method discussed in 3.2.4 to find optimal focus. This will produce one of two outputs. Either optimal focus could not be found, this will occur because either the focus sweep did not cross optimal focus. Or the data quality was not good enough to determine optimal focus. Plots can be made and automatically sent to slack to help diagnose the issue. Optionally, when an optimal focus is found it is tested against the hypothesised (initial) best focus. If the uncertainty of the optimal focus overlaps with the hypothesised best focus, the optimal focus is selected. If the calculated focus certainty does not land on the hypothesis, exposures are taken to compare the two positions. If the optimal focus produces a HFD equal to or better than the hypothesised focus, the optimal focus is selected. Otherwise, the focus returns to the initial focus position.

This whole process is given by the flowchart in figure 3.12.

The previous implementation assumes that the exposure time required to get focus is less than the time taken to analyse the images. This implementation is the *fastest* implementation using the minimum computing resources. A secondary implementation has also been designed for the situation where the exposure time is shorter than processing time. Later in this chapter, I show this is the routine most suitable for GOTO. However, this routine is less computationally efficient than the one shown in Figure 3.12.

The second method works by looping over all the focus exposures and then processing all the images simultaneously, the flowchart for this version is given in 3.13. This will have the same speed improvement as the first method, but, as stated earlier, will use more computer resources.

3.3.2 Time Spent Getting Focus is Not Time Wasted

Exposure time is a clear player in the performance of a focus script. Not only does it effect the quality of the focus images, but also decides how long it will take to achieve focus. Assuming optimal focus is achieved first try the original focus scripts will take:

$$t_{ori} = N(t_{exp} + t_{pro}) \quad (3.5)$$

With the threading component, the new script time becomes:

$$t_{new} = Nt_{exp} + t_{pro} \quad (3.6)$$

Where t_{exp} is exposure time and t_{pro} is the time taken to process one GOTO image. N is the number of exposures. This means the improvement time, δt , can be shown:

$$\delta t = t_{ori} - t_{new} = t_{pro}(N - 1) \quad (3.7)$$

Changing the exposure time will not improve one method over the other, instead it is the processing time and number of exposures.

Regardless of performance improvement, it is still vital to minimise the exposure time. Both for script speed and to minimise the introduction of noise. As discussed earlier, a short exposure time is needed to prevent wind-shake and seeing fluctuations from dominating the focus parameters. However, too short an exposure time and the sources may not express their geometry correctly. A fast and painless way to figure out what exposure time is optimal is to find focus manually and sweep the exposure times and look at the HFD. Using Figure 3.14, the HFD is shown to

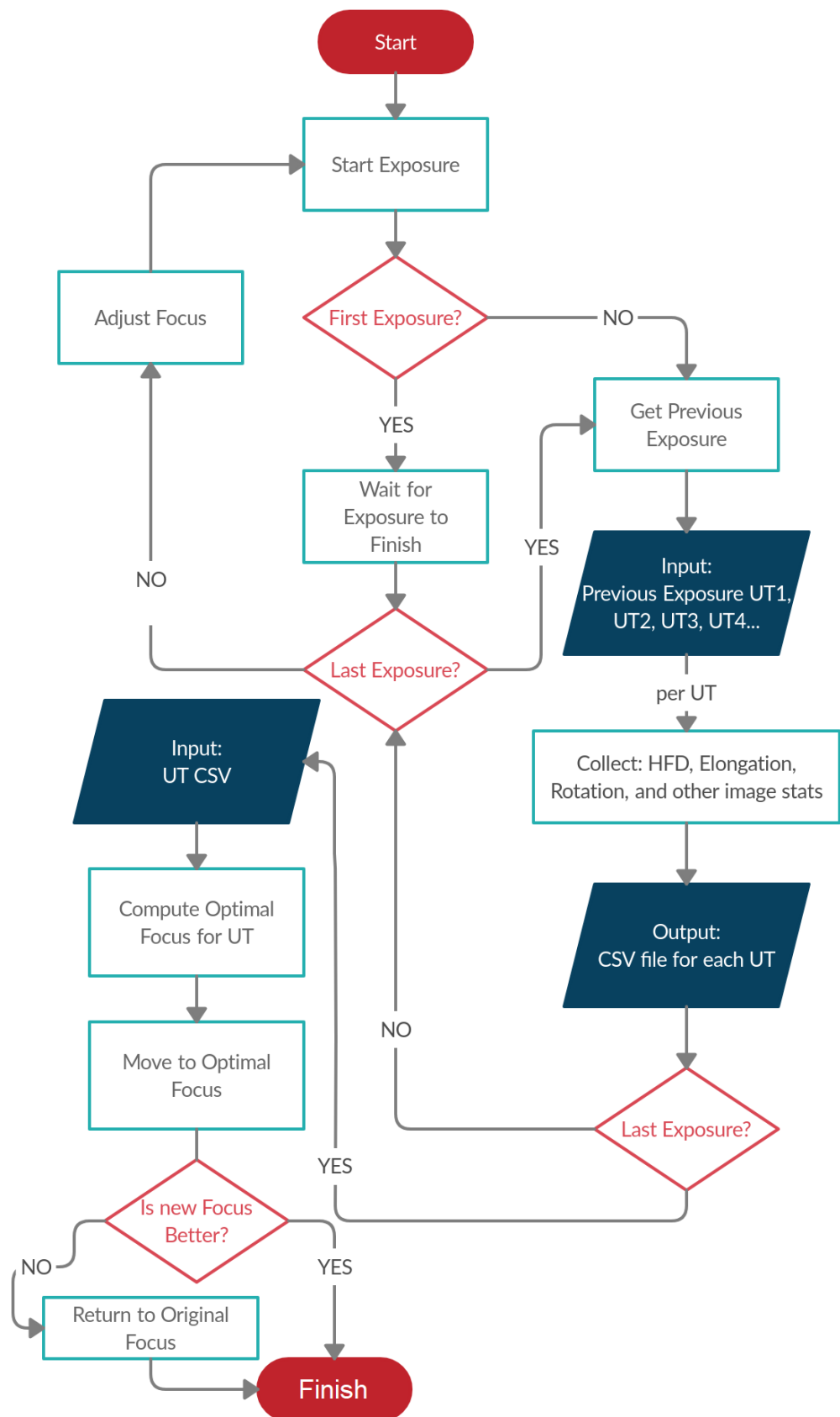


Figure 3.12: Flowchart of the automatic focus script

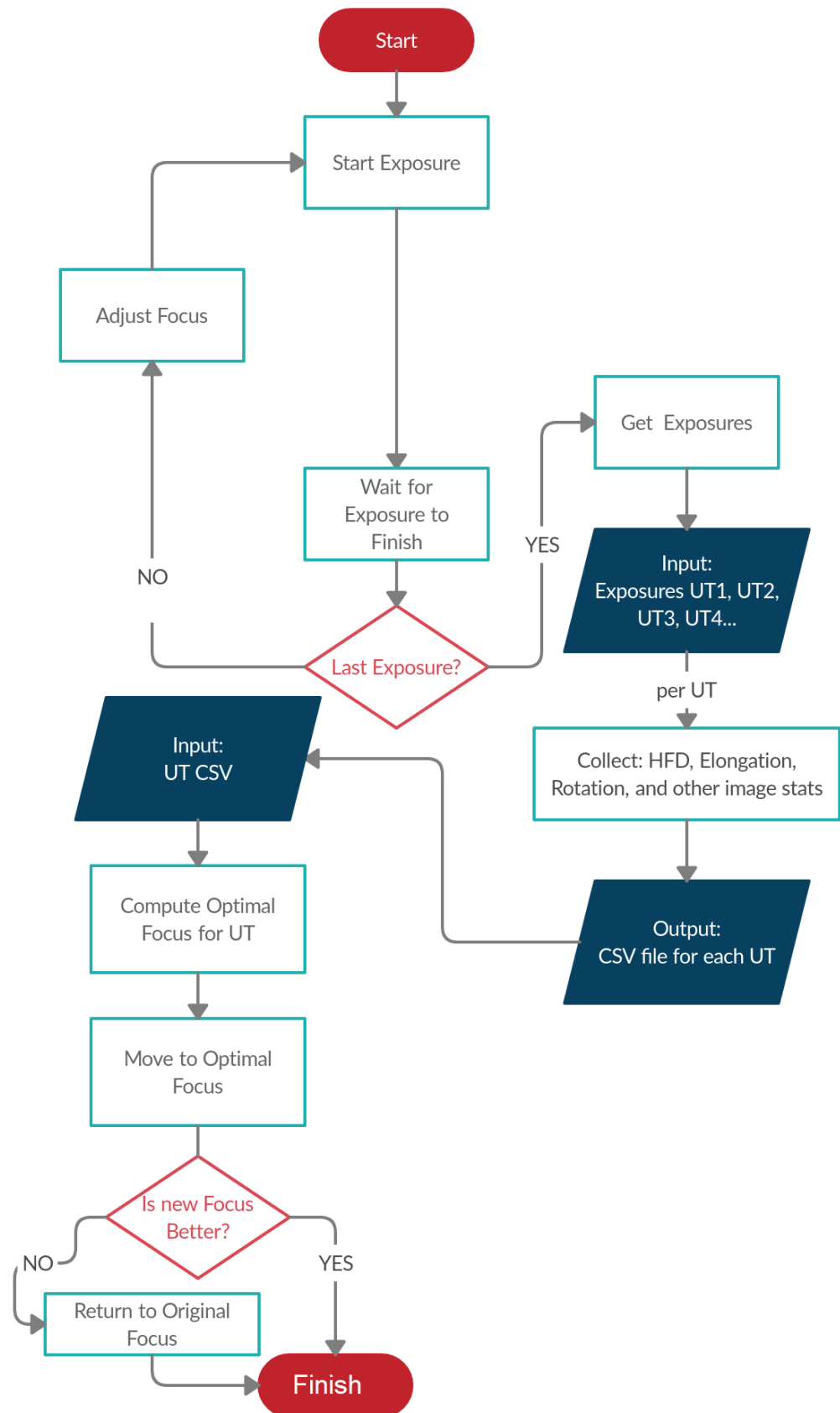


Figure 3.13: Flowchart of the automatic focus script, where exposure time is shorter then processing time

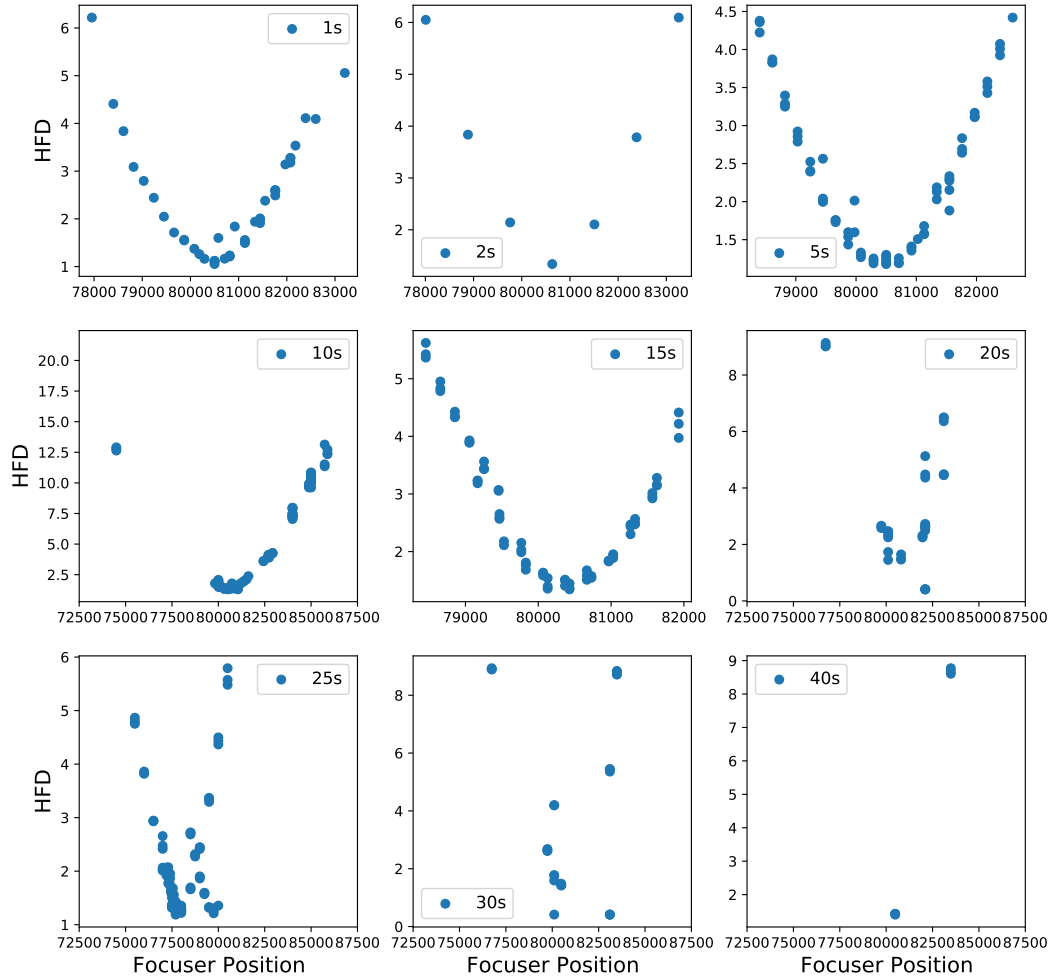


Figure 3.14: HFD using different exposure times. Note: the outliers, particularly in 5s and 15s, are due to the regular sampling in these exposure times. Meaning there are more data points to sample. This shows that focus can be determined with one second exposure times, meaning the longer exposures originally used when finding focus are unnecessary.

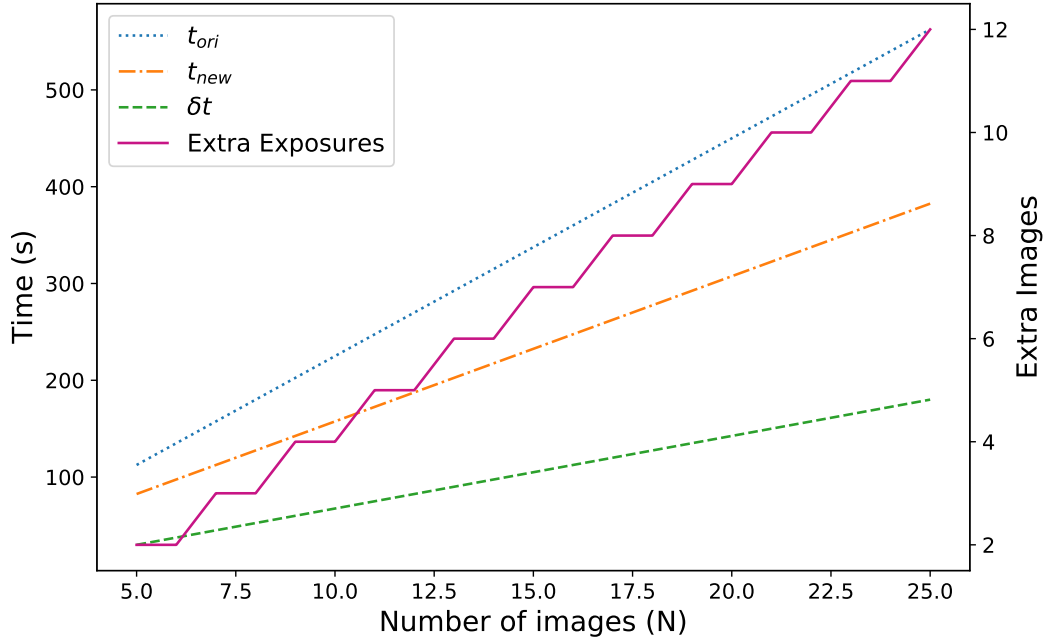


Figure 3.15: Theoretical time taken to obtain focus if the exposure time (15 seconds) is greater than the processing time (7.5 seconds). The dotted blue line shows the original time it took to get focus, the orange dashed line shows the time the new algorithm takes to find focus. The green line shows time saved between the methods and the solid purple line shows the number of additional exposures that can be taken with the amount of time saved (assuming saved time has no additional benefit).

manifest with just a single second exposure. Processing time on the images ranges from 5-10 seconds, 7.5 seconds is used for the theoretical performance in figures 3.15 and 3.16.

Assuming one is taking between 18 and 25 exposures per focus run, this new method should save between 100–150 seconds. This spare time can be used to take pressure off the nightly start-up. But if the additional time is not needed, 10–100 additional exposures can be dedicated to finding focus. As shown above, the exposure time required is much shorter than the exposure time currently being used. Figure 3.16 shows that focus can be obtained in a matter of minutes, or shorter, saving approximately 75% of the time presently dedicated to achieving focus.

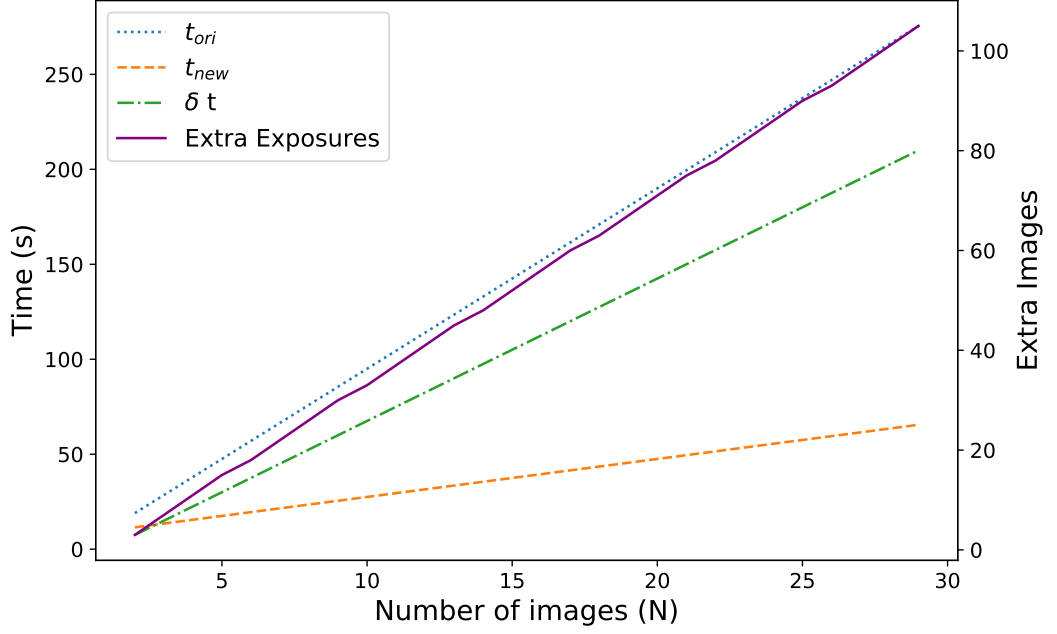


Figure 3.16: Theoretical time taken to obtain focus if the exposure time (2 seconds) is less than the processing time (7.5 seconds). The colours are the same from figure 3.16.

3.4 Script Performance

3.4.1 The Current Method

Currently, the focus run has been upgraded to sweep past optimal focus using the masked HFD method, but does not include the other two methods. This has shown to be robust and functional for the most part. Figures 3.17, 3.18, and 3.19 give an example run of finding optimal focus with each method. Figure 3.19, shows how unreliable the source rotation method is. The step function is mostly distinct for each UT, however the fitting method and predicted uncertainty for each telescope is completely wrong. All three methods break for the same reason in UTs 5 and 6, the sweep does not pass optimal focus.

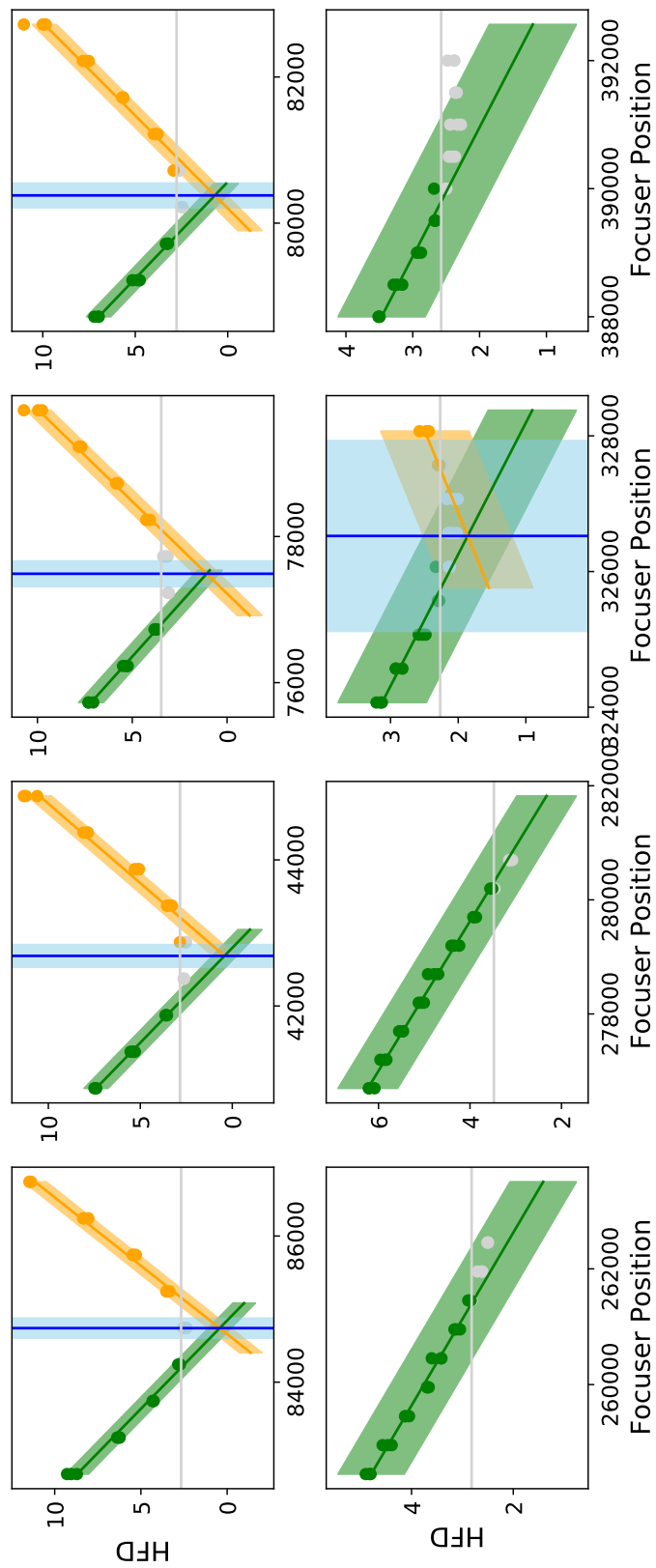


Figure 3.17: Finding focus with the HFD in practice. The colour scheme is the same from the right most plot in figure 3.10. This is shown to be a reliable method when both sides of the uncertainty line are properly sampled. However, UTs 5,6, and 8 failed to find focus due to under sampling of either side of focus.

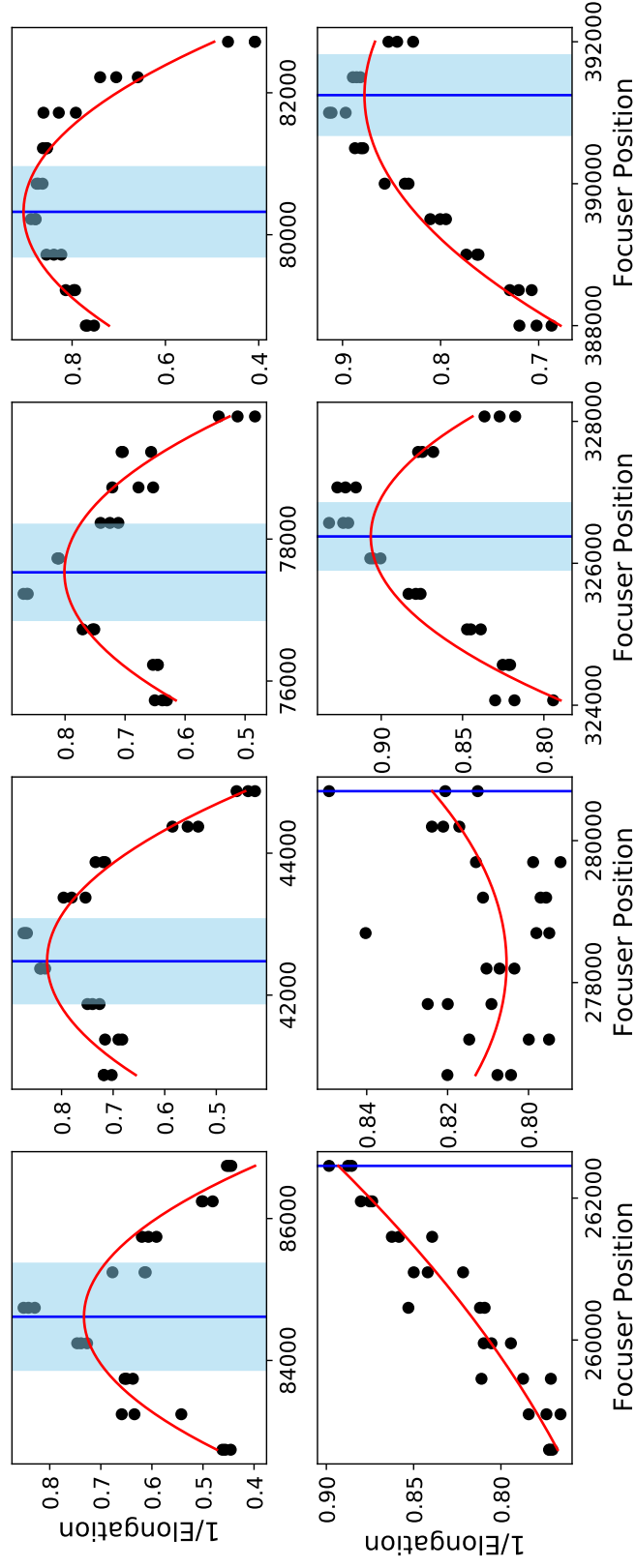


Figure 3.18: Finding focus using elongation in practice. The colour scheme is the same from the centre plot in figure 3.10. Again, UTs 5 and 6 are under sampled. However, a prediction was able to made for UT8.

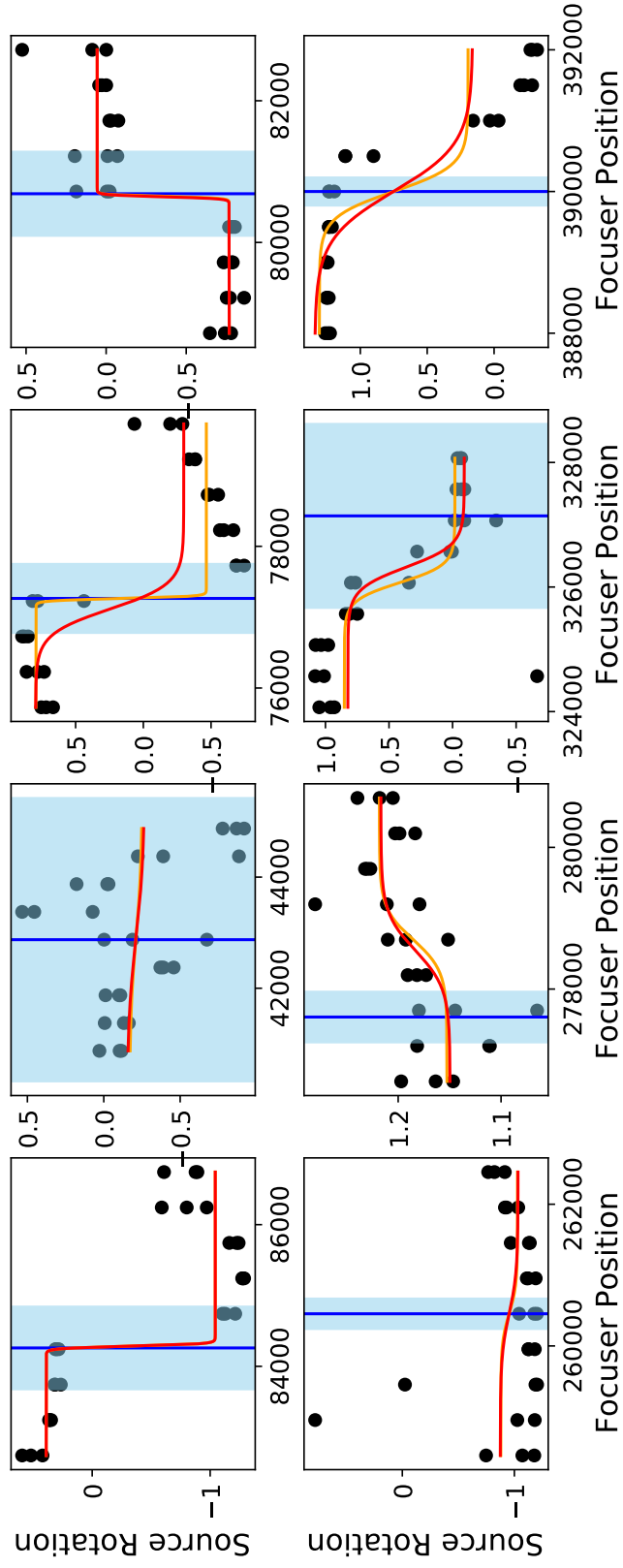


Figure 3.19: Find focus using source rotation in practice. The colour scheme is the same from the left most plot in figure 3.10. The orange line shows the most distant fit from the red fit. The method is unstable which is shown by UT2, 7, and 8.

3.4.2 Future Improvements

Addressing the primary concern first, using fixed predefined steps is problematic. If the sweep does not pass optimal focus or, to be more accurate, does not cross either side of the uncertainty boundary, optimal focus cannot be predicted. Figures 3.17, 3.18, and 3.19 show this for UTs 5 and 6. Interestingly, UT8 does not reach optimal focus with the HFD method, but does reach a prediction with elongation and source rotation. Instead of achieving focus through a sweep pivoting around a given focus, I suggest implementing a logic that bases step size on a given time limit. This would work by presenting a hypothetical best focus still; but instead of providing a set of focus steps, offer a time limit. This means the logic can do the first four steps in the sequence normally, this should be enough to provide the linear model of one side of focus. After the fourth set of exposures, the step size is increased with the aims of trying to cross the uncertainty region. The step size should then be modelled based on the remaining time available. With more drastic measures being taken as time runs out. This means reducing the number of exposures per focus position and larger step sizes. While this will result in lower precision, having an estimate available will save more time than not having a prediction at all and needing to rerun the script.

Source rotation, by inspection, offers useful information about where optimal focus lies. However, analysing this metric automatically is difficult. UTs 1,2, and 4 show it is possible to predict focus using the method developed in this work, but it is unstable. I maintain that source rotation is a valuable metric, though I can conclude that this method would not be suitable for automatic estimation of focus. As a final note, much of this chapter can be considered surface level machine leaning, computers making predictions from the given data. It would be an interesting study to use the data with different machine learners to predict optimal focus. Most importantly, the source rotation relationship has two distinct populations. This would make this parameter a prime target for unsupervised machine learning techniques (Cai et al., 2010). Developing a machine learner for optimal focus would take time; however, with correct method selection could potentially reduce the time cost of achieving focus on a nightly basis.

3.5 Conclusions

Finding focus is an important task in astronomy. This responsibility is usually assumed by the observing astronomer. GOTO's independence means no human is expected to carry out this duty; instead, GOTO is to find focus on its own. The

challenge of finding focus automatically and GOTO’s faults when trying to achieve that goal were articulated in this chapter.

I demonstrated, for GOTO, that the main error on the HFD comes from wind shake. This showed that modelling focus from other noise sources will be challenging, but also these noise errors do not need to be considered when obtaining optimal focus at the start of the evening. I close the discussion by showing that a new focus script must be scale-able and able to run on a short as possible timescale.

In addition to improving the original implementation, other metrics were taken to build novel methods to help find focus. The two new metrics were chosen from their dependence on the astigmatism, the primary weakness in the original method. Those metrics are source rotation and elongation. Furthermore, a new threading technique is introduced to minimise the amount of time dedicated to finding focus which saves $\sim 30\text{--}75\%$ (exposure time depending) of the time spent reaching focus.

One of the biggest weaknesses with the new autofocus routine was the source rotation step estimate. This parameter shows promise for finding locating optimal focus; however, the automated method used here was unstable. To correct for this, I believe an advanced machine learning technique would be of significant benefit. This may also be a beneficial parameter when doing an initial low resolution sweep to find the bounds of where optimal focus lies as source rotation is less affected by seeing conditions.

Elongation was also shown to be a useful parameter. It was able to find optimal focus in situations where using the HFD alone would have failed. This is, however, at the expense of having a lower certainty in the estimates. One way of increasing the prediction certainty is by combining the probability distribution of multiple methods. A future study should look at the correlation in the different parameters so a more statistically rigorous prediction can be made. For now, a more immediate improvement would be to increase the number of focus samples to better constrain the focus models.

Part of the new method has been introduced to the GOTO system, but improvements are still warranted. The current implementation has no safety net if focus is not met, I propose a way to make sure focus is achieved within a given time range as opposed to within a number of exposures. Finally, I suggested an unsupervised machine learning tool to predict optimal focus using the source rotation parameter.

Future improvements to finding focus automatically are going to be vital with the advancement of large robotic survey telescopes. Humans will not have the

resources available to run these telescopes manually, this means robust and reliable logic must be in place to provide high quality data without observer involvement. This is just one way GOTO acts as a pathfinder for future astronomy missions.

Chapter 4

Building Reference Images

“It’s so difficult, isn’t it? To see what’s going on when you’re in the absolute middle of something? It’s only with hindsight we can see things for what they are.”

— S.J. Watson, *Before I Go to Sleep*

4.1 Your Subtraction is Only as Good as Your Reference

As mentioned in section 2.3.3, GOTO revisits particular tiles of the sky. Using these tiles we can build reference images of the quiescent sky for image subtraction. The following chapter is going to detail the optimal way to build reference images for GOTO and assess new tools for both image registration and co-addition.

In this chapter I will address two key features needed for science grade image subtraction:

- The sources in both images must be aligned so their centroids fall on the same pixels.
- The reference must be of equal or better quality than the science frame to make the best use of data.

The following chapter is split into two parts solving both of these issues. Due to GOTO’s wide-fast optics and tracking errors, the data must undergo more complex transformations to be used for image subtraction. Because image co-addition requires the data to be aligned also, this chapter will start by discussing and solving the registration issue first.

4.2 Image Registration

Image registration is the process of taking all the source centroids and matching them to the source centroids of another image; then using transforms to align the sources in the same pixel in both images. Figure 4.1 is a demonstration of what happens when alignment is not achieved prior to the image subtraction step. This is a big problem when hunting for transients, if every source leaves a residual then the subtraction loses purpose. Similar errors appear with image co-addition. To remedy this, one image must be transformed so that the same sources are guaranteed to land on the same pixels in both images before further manipulation.

Image alignment can be done in one of two ways. The first is solving the World Co-ordinate System (WCS). This is a grid defined from a set of instructions in the header which dictates what part of the sky an image is looking at (Greisen and Calabretta, 2002). The images are then aligned so that both grids overlap.

The second is to build geometries based on source position in the images. Each image will have similar shapes (arbitrary constellations in the image), then using mathematical transforms align the shapes based on source position.

I will explain these methods in more detail and show why large fields of view are more complicated, which necessitates higher order transforms than are currently available in the literature.

4.2.1 World Co-ordinate System based methods

Described above is the basic idea of a WCS. The fits file will have multiple headers detailing instructions of how to project the image onto a global grid. GOTO uses a tool called **Astrometry.net** (Lang et al., 2010). In summary, this works by using a source detection routine, similar to the one mentioned in chapter 2, to find star positions in image co-ordinates. Stars are grouped into sets of four, called quads, and compared to a grid of known stars. Using Bayesian analysis, a test hypothesis for grid instructions is compared to the null hypothesis. If the hypothesis passes the instructions are written to the header, otherwise another hypothesis is tested. This routine is repeated until either a solution is found, the wall time is exceeded, or the code runs out of viable hypotheses.

Once the WCS of the two images is known, one can use this information to align the sources on specific pixels. The main tool used to do this is **SWarp** (Bertin, 2010); by using a specific set of input instructions with **SWarp**, one can align images without co-adding them. This is a relatively fast process (see section 4.2.5); however, one concern for this method is its flux conservation. The other

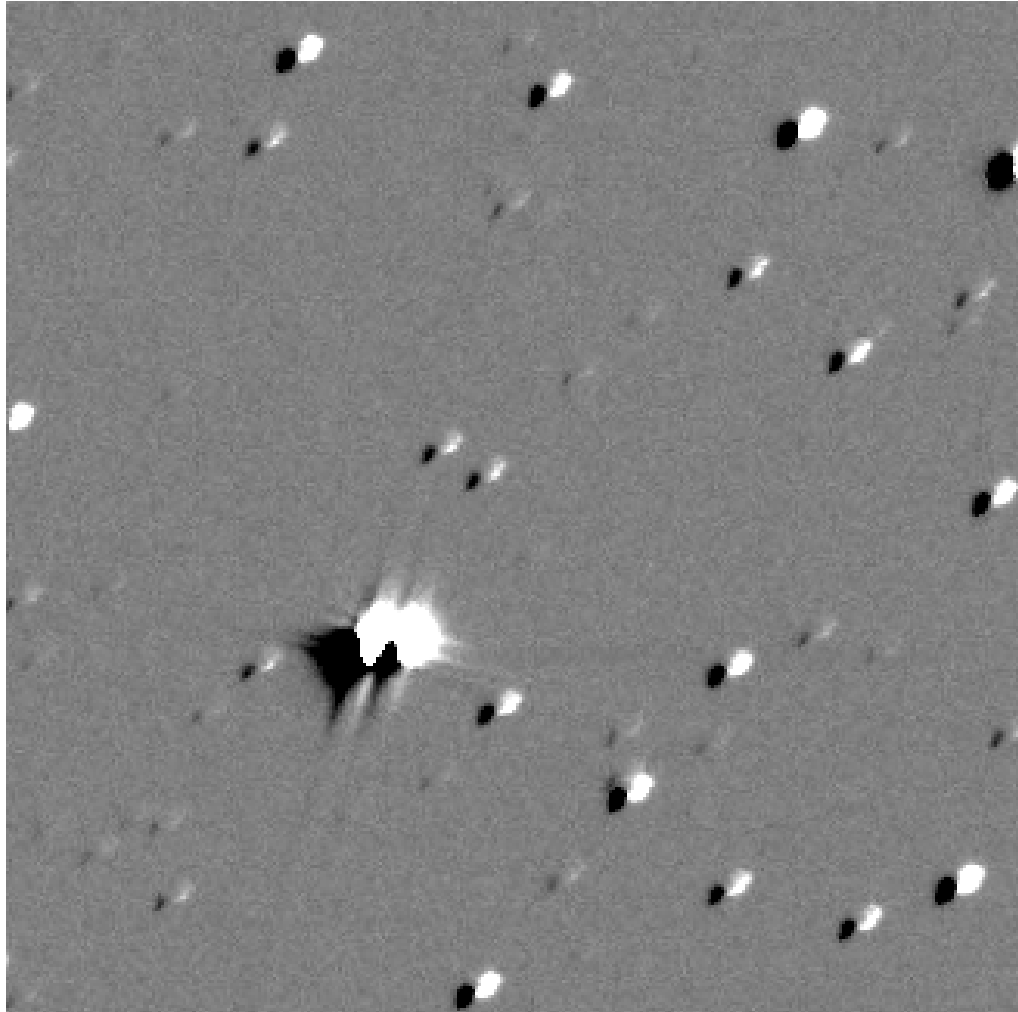


Figure 4.1: The result of subtraction with poor alignment. Note the residuals leave a shadow imprint leaving behind 'butterfly' artefacts.

main concern is the simplicity of the image transforms. As mentioned in chapter 5, GOTO fields are much larger, both issues are exaggerated through the simplicity of the image transforms on GOTO’s wider fields.

In addition to Swarp, other WCS based alignment algorithms exist. One of particular note is reproject (Robitaille et al., 2018). Reproject uses an updated version of the **Montage** mosaicing algorithm (Berriman et al., 2008). It takes particular care to conserve the flux of the image when doing WCS based transforms (Jacob et al., 2010). This software will be used as the benchmark for flux conservation later in the chapter. The performance of these transforms are subject to the quality of the WCS. When compared to other methods later in this chapter, the weaknesses in the WCS become apparent for GOTO data.

The WCS model is made for both images, this means the sources in both images are not actually matched to each other, but a third intermediary grid. This makes the WCS a weaker method for techniques like image subtraction that need two images to be aligned to pixel precision because the potential for an error being made is more than doubled.

4.2.2 Affine Transforms

Instead of using an intermediary grid, another way to align images is to match the stars in both images directly. **Alipy** does this in a similar way to **Astrometry.net** (Tewes, 2013). Quads are built in both images, but instead of being compared to a pre-built grid, they are compared to each other. The difference in the source positions are taken of *matching* stars and used to build an affine transform.

Affine transforms are sets of mathematical instructions for Euclidean space that preserve the collinearity (in other words, all points lying on a line prior to a transform remain on said line after transformation) and distance ratios of collinear points (for example, the midpoint of a cross-section remains the midpoint after transformation). This is usually done in an array format, and uses a list of linear transforms to scale and rotate points. The transform is built by taking the matched quads and building the transform from the difference in their source positions.

The software used in this thesis to compute the affine transform is from **scipy** Virtanen et al. (2020).

4.2.3 The Problem With Wide Fields

GOTO has the case of having fast, wide field optics and non-negligible tracking errors. Unlike other, smaller, telescopes, this results in wide field distortions at the

edges of the image and source shifting. When aligning, the first order transforms discussed above are insufficient at dealing with these edge effects. This results in poor alignment residuals shown in figure 4.1. Because of these errors, a new alignment technique is needed to register GOTO images.

Alignment is a process in the image subtraction pipeline that must occur on relatively short timescales. As the images need to be aligned in real-time, the methodology should reflect that.

To summarise the introduction of this section, GOTO fields are too difficult to manage with the first order transforms discussed above. A higher order transform method is necessary to remedy this. The problem described here has added difficulty, because not only do the images need more involved transforms, there are time constraints on the alignment period. The following was developed as a solution to these issues.

4.2.4 Spali2 - The Affine Transform with a Spline

Spali2 is the successor to the code Spalipy (Lyman, 2018). Spali2 has been designed specifically to be integrated into the ZOGY in Parallel (ZiP) framework (this package is introduced in the next chapter), additionally it has features designed to make it more user friendly¹. Moreover, it also has parallel steps introduced that saves between 10 and 20 seconds on alignment time, depending on the source density of the images.

Spali2 accounts for the non-linear edges of the GOTO-optics. It starts by using the base transform from Alipy. This involves building quads and computing the affine transform. From here, the difference from the x and y positions of the transformed sources and template are calculated. These are then fed into the spline step, which calculates the higher order transforms. The final transform is then applied to the original image.

In addition to registering the image, the key headers of the original are saved to the transformed image, but, if available, the WCS headers of the template are used in the newly aligned image instead. This is a sensible solution as it prevents confusion later with the co-ordinates. If the WCS of the template is not available, the WCS is simply removed from the aligned image. Additional comment headers are included to indicate that the image has been aligned using Spali2 for posterity.

Spali2 has been implemented into the main GOTO workflow already. However, the performance of the script has not been measured or compared to the other

¹<https://zogy-in-parallel.readthedocs.io/en/latest/Spali2.html>

alignment algorithms currently used in astronomy. The next section will test and compare the performance of the different registration methods.

4.2.5 From the Line-up, Which is Best?

The following test has been designed to measure two components, alignment time and precision. To do this a group of seven random GOTO and Next Generation Transit Survey (NGTS) fields were selected. Each field is identified by a tile name. From there, a set of 10 images are made from each field. This means a total of 70 images for both telescopes. A group is chosen blindly and two images in that set are selected at random to be aligned, this is repeated 100 times for both telescopes. Figure 4.2 shows the alignment time for each alignment method over a variety of fields. The precision tests are split into two comparisons; the handling of the smaller easy (NGTS) fields, where the images were taken hours apart, and then another alignment of the large GOTO fields that were taken days apart. The stats are shown in table 4.1, figures 4.3 and 4.4 show the accuracy of one alignment for each telescope.

Method	Test	Number of sources	Within a pixel difference
WCS based	A	97110	88701
Alipy	B	97312	88911
Spali2	C	98817	89257
WCS based	D	1173306	548743
Alipy	E	1177715	952558
Spali2	F	1177417	987346

Table 4.1: Image registration tests, supplementary table to figures 4.2, 4.3, and 4.4. The table shows that each method is suitable for well behaved fields, however WCS methods perform notably worse in GOTO images.

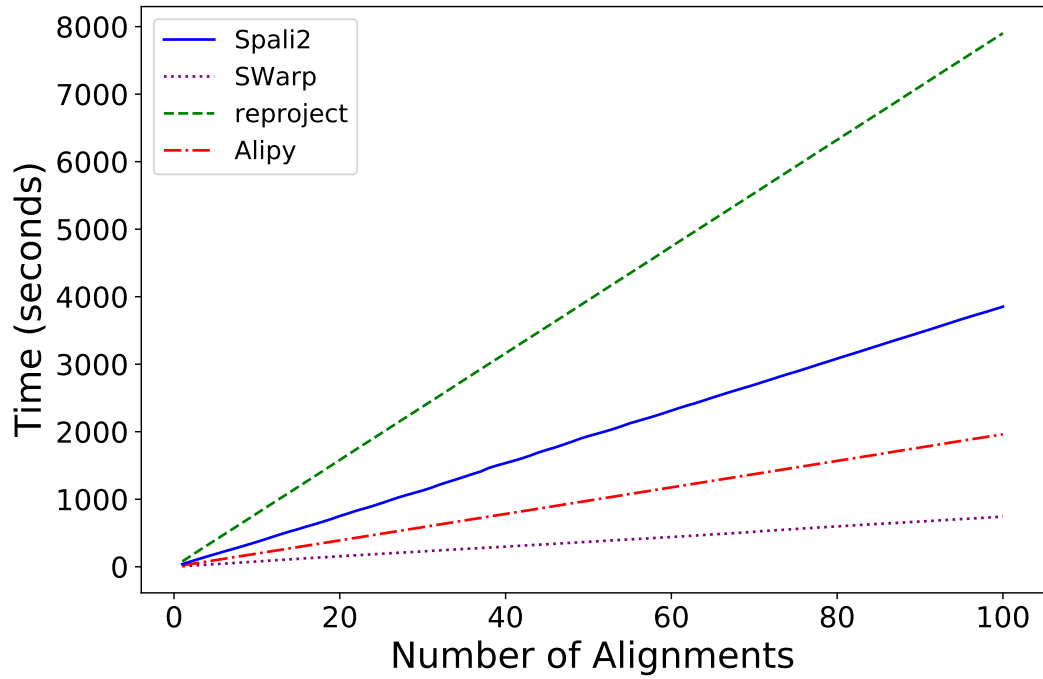


Figure 4.2: Physical time to compute each alignment on a GOTO field. (It should be noted Alipy, and Spali2 use 2 cores while SWarp and reproject use all the cores available.) Each line has a linear relation, meaning the transform time from each method can be modelled easily using the mean of each method. SWarp and reproject do not account for the time it takes to solve the WCS, this can have an additional 10-20 seconds computing time to these solutions.

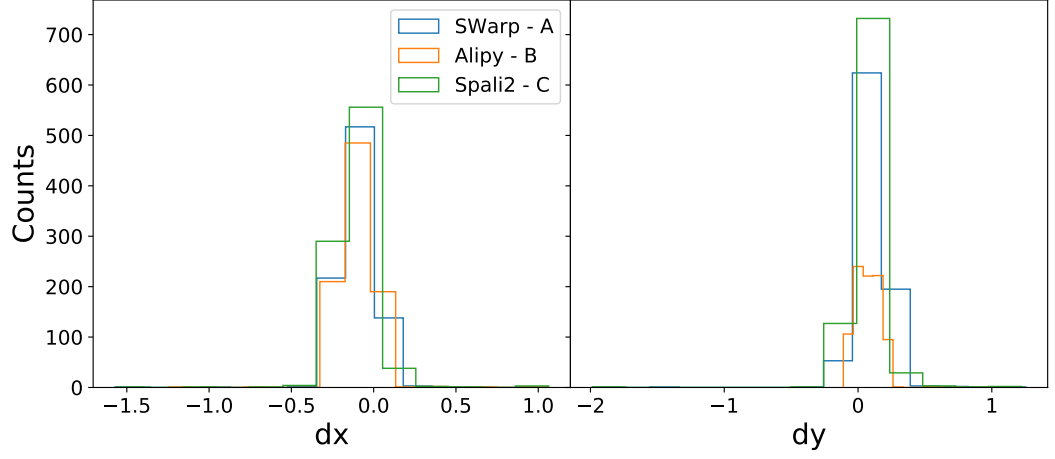


Figure 4.3: Histograms showing separation of sources in pixels from the original image in x and y directions. **A**: WCS based alignment of a small field, **B**: Alipy alignment of a small field, **C**: Spali2 alignment of a small field

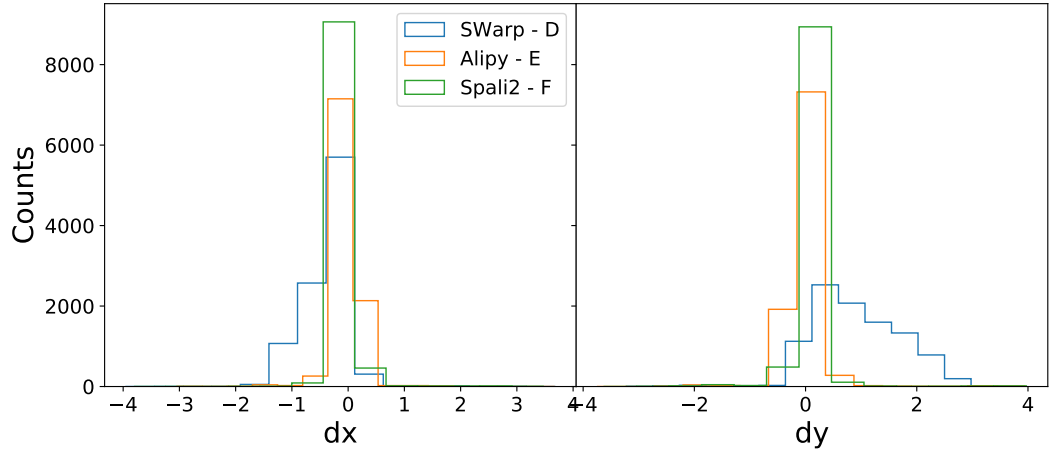


Figure 4.4: Histograms showing separation of sources in pixels from the original image in x and y directions. **D**: WCS based alignment of a wide field, **E**: Alipy alignment of a wide field, **F**: Spali2 alignment of a wide field

4.2.6 Flux Conservation with Spali2

When transforming an image, it is important sources do not lose information. The main feature that needs to be conserved is flux. That means that when an image is transformed all the sources in the image should contain the same number of counts after the transform. The following is a systematic test to show that Spali2 does not introduce error into the flux of sources when transforming the image.

To measure the conservation of flux, the WCS is needed. There needs to be a set of rules in place to decide which sources are the same in the pre and post transformed image. For each source in the original field the WCS is taken with the aperture flux. The image is then aligned and the new transformed image assumes the WCS of the image it was registered to. The flux and source position are recorded again in the new image. As shown in figure 4.4, the WCS is not entirely reliable, to account for this error a few criteria have been put in place. The flux is compared on WCS location, anything more than 20 arc-seconds away is considered a different source. Additionally, as transients and variables can skew the results, any source flux beyond 50% difference is considered an outlier and is ignored. The latter criterion was relaxed and showed no major impact on the results. The results are plotted in figure 4.5.

Furthermore, Spali2 uses a positional based spline step. This means there may be a spatially dependent error on the transform that is masked by looking at the flux as a whole. Figure 4.6 is the spatial dependence on flux conservation.

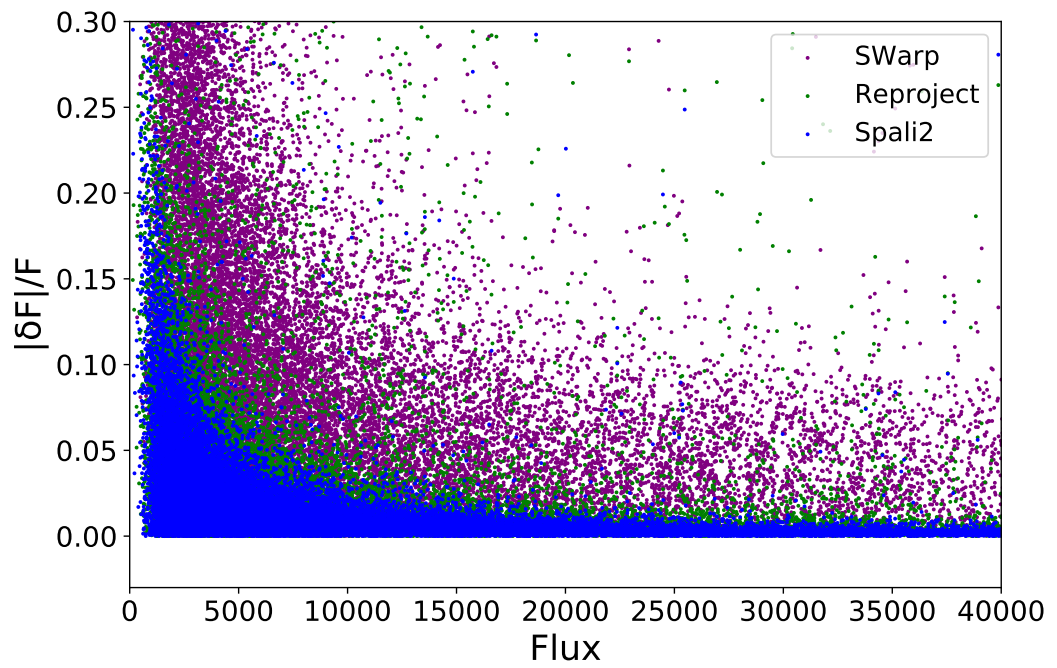


Figure 4.5: Flux conservation of the different alignment methods as a function of flux. SWarp (purple) has the most inconsistent performance. Reproject (green) and Spali2 (blue) have similar performances with Reproject showing a little more scatter at higher fluxes.

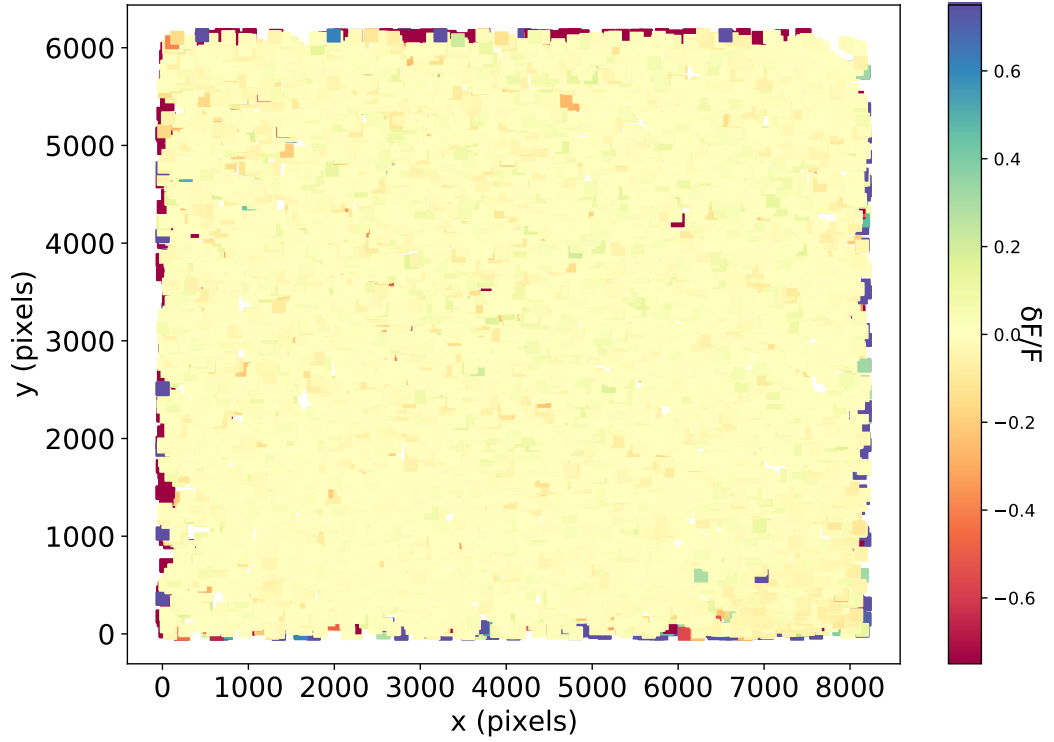


Figure 4.6: Flux conservation of spali2 as a function of position for ten different alignments. There is no positional dependence on the Spali2 transform quality. The edges of the original image are often out of frame in reference. This means the transform removes them from the aligned frame, this means sources in the original image have no comparison in the transformed frame, hence the large edge uncertainties.

4.2.7 Image Registration Discussion

Figure 4.2 shows that each method has a linear scaling with time, independent of field complexity. The **SWarp** algorithm takes the shortest amount of time (7.5 seconds per image). This is a little misleading as it does not account for the time it takes to build the WCS, which can add up to 20 seconds computing time per image. This makes the fastest algorithm **alipy**, at 19.8 seconds per image. **Spali2** shows that it can complete alignment on the same order timescales, taking 39.9 seconds per image. **Reproject** takes the longest, this is most likely due to the effort dedicated to maintaining flux conservation.

Using the NGTS alignment tests in fig 4.3. It can be seen that all the methods perform equally. This may be due to the fact the fields are almost aligned already as they were taken only a few hours apart and NGTS has good mount guiding, but also because the field is relatively small and the PSF does not vary dramatically across the CCD. However, looking at table 4.1 and figure 4.4 there is a performance difference for the GOTO data. The first order transforms have a harder time correcting for the distortions of the wide field. This problem is exasperated in the WCS alignment case as errors are made in both images to the grid as apposed to a single set of errors from a direct comparison of the image. Because of the higher order correction, **Spali2** is shown to perform the best in both alignment precision tests. However, this improvement is unnecessary for the smaller fields. This means, the best alignment method is dependent on the field size being used. Because of this, the spline step was made optional in **Spali2** to keep run times optimal.

The flux conservation tests have shown that there is almost no uncertainty introduced in the aligned images. This is flagged notably in figure 4.5, **Spali2** shows a clear preservation of flux compared to other methods used in astronomy and even out performs **reproject**. Positional uncertainty does not appear to be a concern, however, this is image quality dependent. Exposures with more sever warping will require more involved transforms. At a certain point the spline over compensates, the flux and source shape become unusable. To account for this, a basic quality control is implemented in **Spali2** that requires a certain level of confidence in the number of high quality matched sources. This safeguards the algorithm from outputting an image distorted by the spline step.

4.2.8 Image Registration Summary

The above has shown that large images with wide fields of view are prone to optical distortions. This makes aligning images difficult. By introducing a higher order

transform, one can effectively register larger images in preparation for co-addition and subtraction. Spali2 has been shown to be more reliable for all image types and has shown to conserve flux indiscriminate of source flux or source position. This makes Spali2 optimal for wide field fast optical systems like GOTO.

4.3 Image Co-Addition

Looking at the latter criterion for image subtraction, if the reference is of lower quality than the science image then transients will be obscured in the subtraction, lost in the noise. The most stable remedy is to ensure the reference has better signal to noise than the science image. The main way to guarantee science images have better references is image stacking, also referred to as image co-addition.

GOTO's current image subtraction pipeline utilises a basic median stacking algorithm. Comprised of three reduced images for both the science and reference images. This is most likely the optimal way to co-add a science frame in real time as it is fast and uses all the available data. It is a different case for reference images. By building references this way, the reference quality is going to be of the same order as the science frame. There is a 50% chance the reference will be at a poorer standard than the science which violates the last rule of image subtraction. As references can be prepared offline (i.e. build references when the sun is up), more care can be taken when constructing them. This means the issue of reference quality can be solved in two ways. By selecting multiple high quality epochs in the stack, and by using more detailed algorithms. In the following, I will show a new co-addition tool I have built and compare it to other techniques used in astronomy and show the best way to construct reference images with GOTO.

Image co-addition is the process of taking three or more images of the same patch of sky and combining them in such a way that the resultant image is a pixel by pixel average of the combined images. The goal of doing this with science images is to remove any one shot artefacts, like cosmic rays, from cluttering the final subtracted image. The other benefit, and the one more applicable for reference images, is increasing signal to noise. This works because the signal is spatially correlated, but the noise is randomly distributed for each image; meaning, when the images are stacked, the signal (the real sources) will add constructively, but the noise will cancel out. The standard expectation for any image stacking method is that the limiting depth will increase as such:

$$m_{lim} \propto \log \left(\frac{N}{\sqrt{N}} \right) \quad (4.1)$$

where N is the number of stacked images.

Median combination is simply that, each pixel is combined and the median of the pixels is taken for the stack. GOTO uses a background-scaled median. This is because in some sets of images the background can vary drastically over the collection of exposures. The variation in background is accounted for and consequently equates the flux of the sources in the images. Some statistics are not useful when building reference images. For example, the standard mean, as shown in figure 4.7, is poor choice as bright cosmic rays are not removed from the image.

A preexisting tool that has image co-addition functionality that is sometimes used when stacking images, called **SWarp**, co-adds images by giving them predefined weights and adding them in an optimal way (Bertin, 2010) . Another method also claims to co-add multiple frames in a statistically optimal way (Zackay and Ofek, 2015a,b). This latter method has been dubbed *proper co-addition* and I have developed an algorithm around this method.

4.3.1 Proper Co-Addition

Proper Co-Addition has the following formalism:

$$\hat{R} = \frac{\sum_j \frac{F_j}{\sigma_j} \overline{\hat{P}_j} \hat{M}_j}{\sqrt{\sum_j \frac{F_j^2}{\sigma_j^2} |\hat{P}_j|}} \quad (4.2)$$

Where a bar represents a complex conjugate and a hat the 2D Fourier transform. M_j is the background subtracted j^{th} image and P_j is the corresponding PSF. The derivation of the above can be found in appendix B. The algorithm was built in python and works with optimal speed. This works by solving all the sum terms simultaneously, computing the two sums, and finally completing the division. One of the key principals of this method is that it preserves the PSF of the combined images. This is to stop close sources getting blurred together and retains spatial information. To date, this claim has not been tested in the literature and so will be trialled in the following section.

\hat{R} symbolises the 2D-Fourier transform of the reference image R . The co-addition uses a PSF estimation convolved with the original image (M) which is variance scaled such that the signal to noise is maximised, assuming a background dominated image (see appendix B). The sum of the Fourier transforms results in correlated spatial frequencies adding constructively and uncorrelated frequencies destructively. This means sources and artefacts that only appear in one of the co-added set are removed in this process.

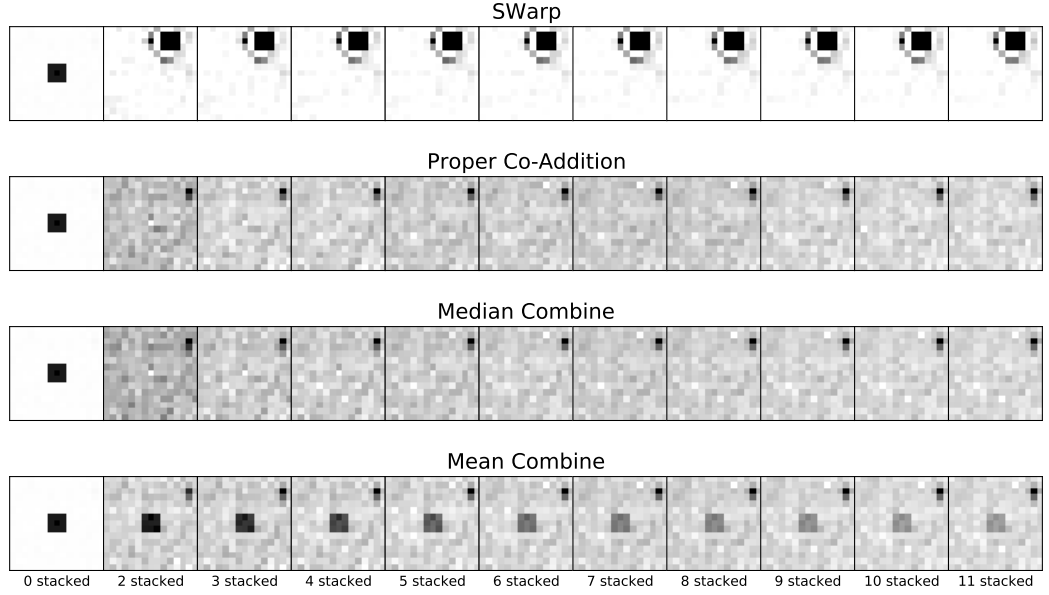


Figure 4.7: Initial comparison of how each stacking method handles artefacts. All the co-addition methods work well, apart from the mean which shows how some statistics are not suitable for co-addition.

The PSF is estimated using a modified version of PSFex (Bertin, 2013). This modified version is a python wrapper around the base code that works on sections of the image in parallel to solve the PSF, accounting for spatial variation, across multiple of an exposure. The algorithm is discussed in further detail in section 5.2.3.

4.3.2 Comparing different Stacking methods

The following will discuss each methods performance when co-adding images and use that performance to make assessments about how both references and science images should be built.

Firstly, each method is compared on how they handle bright cosmic rays. Figure 4.7 details this, showing that each method will remove artefacts within the first three stacked images. The mean is included to highlight why certain statistics are poor choices with image co-addition.

Knowing that each of the methods can handle inconsistencies, one can move on to measure output image quality. This is done in two ways, The first is to look at the limiting depth of the image. The number of high confidence sources detected in the image can be used as a proxy for this if zero-point information is not available, which is done in figure 4.8 (Mong et al., 2020). The second is to check the stability and size of the PSF across the co-added images (table 4.2).

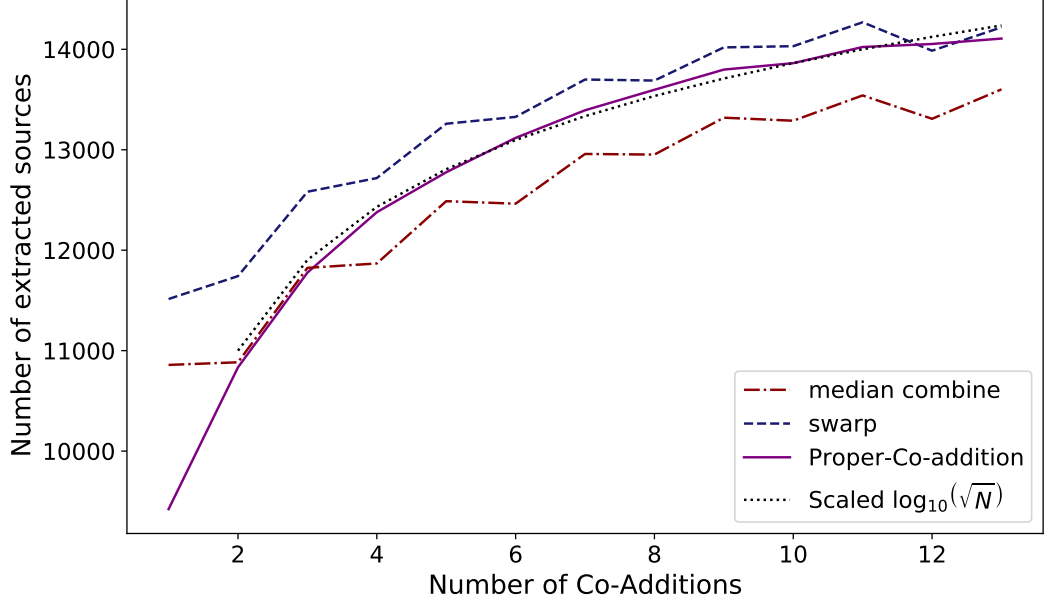


Figure 4.8: Number of 5σ detections for a given number of NGTS stacks for each method. This is compared to a scaled \sqrt{N} graph, showing the depth increase matches what is predicted by equation 4.1.

It can be seen from figure 4.8 that for all the methods the source count stops increasing at roughly 12 stacked images for all methods, all following a \sqrt{N} relation predicted by equation 4.1. This limit is a ceiling introduced through diminishing returns, where the performance of the stack no longer improves significantly with additional images. Median combine is demonstrably the worst performer of the methods. However, both the **SWarp** and Median combine methods show dependence on image quality. When some stacks are added the overall performance is diminished, this is never the case for proper co-addition. **SWarp** is seen to be the best method for low image count, but for high image count **SWarp** and proper co-addition are equal contenders, with proper co-addition demonstrating a greater stability.

Table 4.2 highlights the FWHM statistics as a proxy for PSF quality. The stats were taken using **SExtractor** parameters. The FWHM of every σ_5 source was taken and averaged, ignoring any saturated sources. The first thing to note is that median-combine and Proper Co-addition have notably smaller PSFs than **SWarp** (this can also be seen qualitatively in figure 4.7). This is because **SWarp** uses a simple convolution on each image to ensure they have the same source shapes. This however, results in a loss of spatial resolution that median combine and proper co-addition keep. Furthermore, while slight, the variance in PSF from proper-coaddition

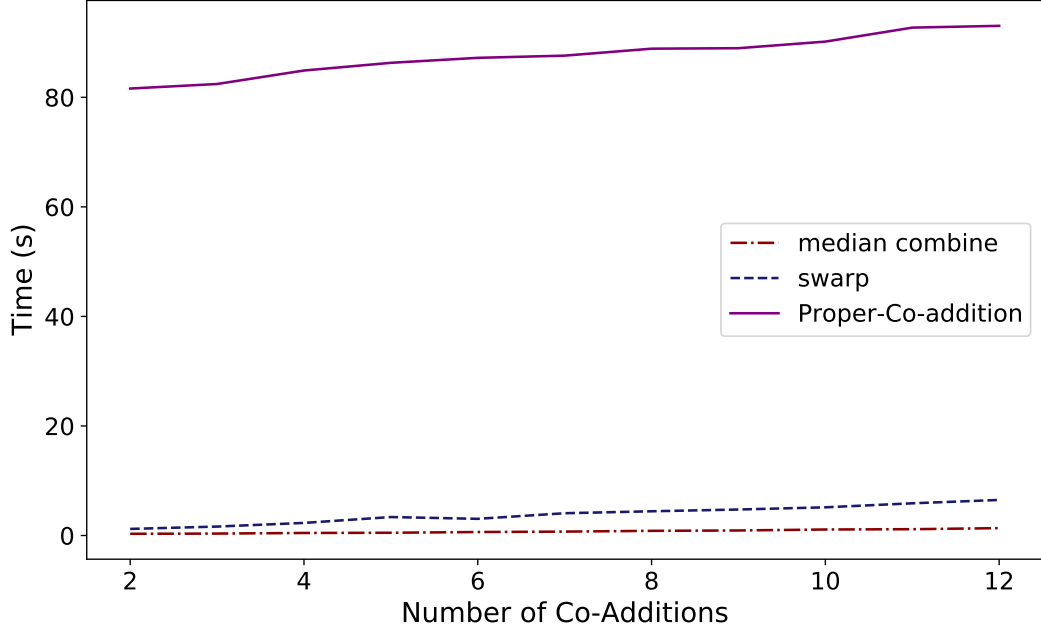


Figure 4.9: Time taken to stack NGTS images, using multiple cores. The Swarp and Median combine methods scale linearly with number of images and are much faster than proper co-addition which is significantly slower, but its speed is roughly independent of the number of images used.

is lower. Highlighting the PSF stability of the method.

Method	Mean FWHM (pixels)		Median FWHM (pixels)
Swarp	1.95	± 0.016	1.63
Median Combine	1.76	± 0.016	1.42
Proper Co-addition	1.77	± 0.010	1.42

Table 4.2: PSF statistics for different alignment methods. Highlighting that proper co-addition conserves the PSF when stacking images.

Finally, in the case of needing to stack reference images in real-time a time test is computed. This tests the amount of time each method takes to stack a number of images.

Looking at figure 4.9 it is apparent that Proper co-addition is not suitable for real-time image building. 80 seconds to co-add images is too long for real time searches. This may be a valuable tool to look at legacy data when there are no time constraints however.

In summary, proper co-addition offers better signal to noise than median combine and presents a better PSF stability than **SWarp**. All of these benefits come at the expense of a longer run time, which when utilised in a non time pressured scenario is not a concern; so, suitable for building reference images.

The next section will work under the assumption that GOTO's references will be made in preparation to a survey and not at the same time the data is being taken.

4.3.3 Stacking GOTO images

Currently, as discussed earlier in this chapter, GOTO builds reference the same way it stacks sets of science images; by combining a set of 3 (sometimes 4) exposures. Plans for the future are to increase this limit; however, as building an entire reference grid for the sky takes priority, most tiles in GOTO's catalogue have only been visited once and therefore do not have multiple frames to stack. This is a problem when doing image subtraction as it means the science image has a strong likelihood of seeing deeper than the reference. This section will detail two different image stacking routines using proper co-addition and highlight what method is optimal for future GOTO campaigns.

Exclusively Dark Time

This method is based on using only the highest quality data. Data that are taken when the moon is down. Dark time data will have a lower sky background, making the signal to noise higher. By also excluding poor seeing data, there is better PSF containment and expected higher gains per image stacked. The disadvantage of this method is how strict it is. Using only dark time data may result in an overall better quality reference, but this requires a much longer survey and discards a large fraction of potentially usable data.

The Random Selection Method

Every available image from a random selection is stacked together. Looking back at figure 4.8, this should be a viable way to stack images without input discrimination. As higher weighting is naturally given to the better quality data (i.e. images with higher signal to noise relative to the background), the overall output image should always be of optimal quality. Whether this is consistently true in practice yet to be seen. There is also a case to be made if this is computationally the most sustainable method. There are bound to be useless input images past a certain point and

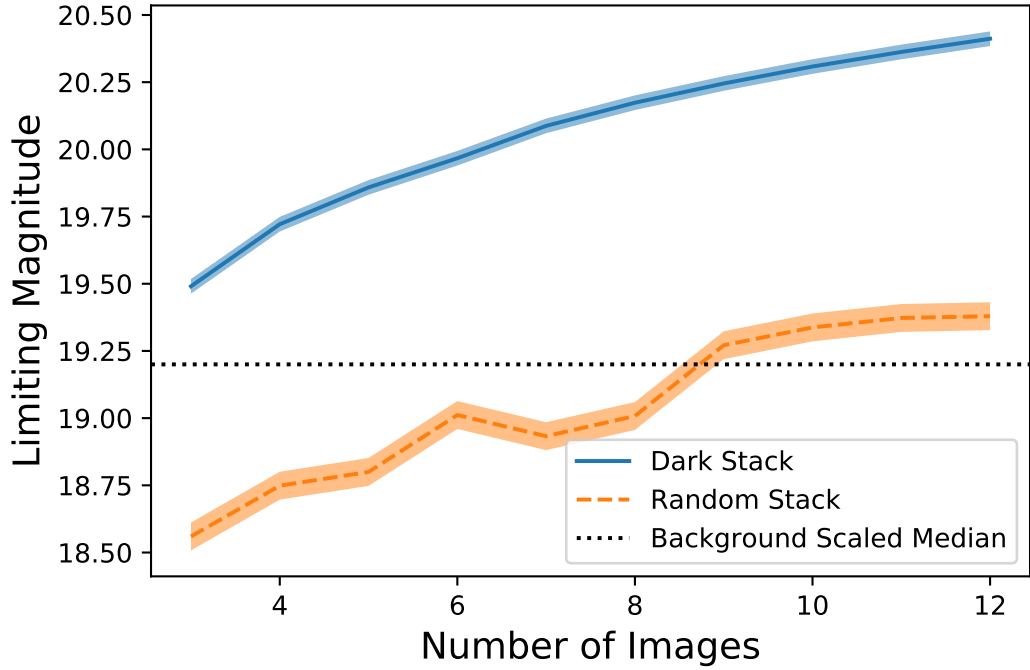


Figure 4.10: Limiting magnitude for different Stacking Strategies of GOTO data. The dark stack gets the deepest reference. Using a random assortment of fields is not beneficial and shows that some curation is necessary to optimise ones reference image.

processing them all is costly in more than just a temporal sense (Broekema et al., 2020). Cores use energy and have an expected lifetime, unnecessary overuse has an environmental impact and is a problem in computational astronomy.

Results

Looking at figure 4.10, the limiting magnitude of the co-added image is primarily dependent on the quality of the input images. This shows, when making reference images, the best option is to select exclusively good seeing images taken during dark time. This can be done over a long range of epochs, which should also present the best reference of the quiescent sky (as short timescale variations are removed). Compared to the average background scaled median (A GOTO science exposure), the optimal references will always see deeper. However, given that dark time templates are not always available, the best insurance is to have a stack of 12 images (see figure 4.10).

Another point to address is the overall performance. It appears that as long as the background is accounted for, the chosen stacking algorithm does not

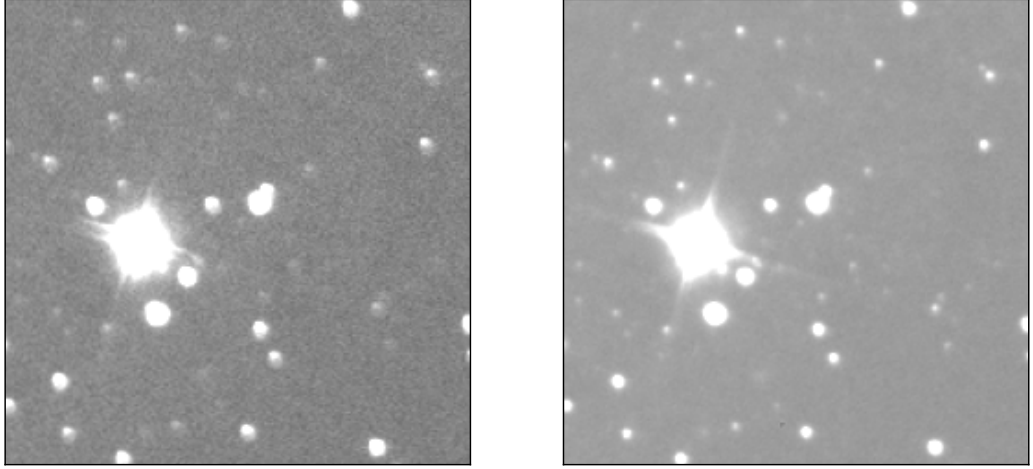


Figure 4.11: Comparison of PSF conservation Left: Median Stack. Right: Proper Co-Addition.

play a significant role in limiting depth of an image. Although, PSF resolution is an important feature to conserve, especially for image subtraction. Figure 4.11 presents the difference in PSF quality for a stack of (the same) 12 images using a median stack and proper co-addition. Proper co-addition, at least qualitatively, has a more symmetric PSF.

4.4 Chapter Conclusion

In this chapter I discussed the key techniques necessary to build reference images for image subtraction. The first being image registration, the act of aligning sources in one image to sources in another. I detailed the shortcomings of the current methods used in the literature and highlighted how for GOTO fields these issues are exaggerated. To solve the issues faced by the simple transforms, I developed a new higher order transform that uses a spline step to correct source placement at the edges of the image. This accounts for the warping caused by the wide-field optics. I show that this implementation runs on the required timescales needed for image subtraction and readily aligns the most sources in the image compared to its competitors. Furthermore, I demonstrated that this new technique conserves the flux of the sources, meaning no information is lost when transforming images using Spali2. It is noted that, for smaller fields, it is faster and more sensible to use standard first order transforms, as the benefits of Spali2 are negligible compared to the time lost doing a simple transform.

The second half of the chapter covers image co-addition. The method of combining multiple images of the same patch of sky to create a deeper reference image. The issue here is that GOTO's references see to the same depth as the science images. This means there is a strong chance that when doing image subtraction, transients will be lost. Firstly, a new algorithm is developed using a novel statistical approach to image subtraction. This technique shows to compete with the algorithms currently used in the literature. However, proper co-addition takes significantly longer to compute, meaning on the fly co-addition is not possible for real-time processing with proper co-addition. This is seen not to be a huge dilemma with GOTO as the telescope uses a tiling system; so, reference images do not need to be made in real-time and instead can be made in preparation. Secondly, real references are built using this new, proper co-addition, technique and compared to the scaled median method currently used by GOTO. Two selection methods are used here, a random assortment and exclusively dark time images. The random assortment of images is worse than the curated scaled median images, this means that one must be selective when deciding what images to use when building a reference. The dark time exclusive selection sees deeper than scaled median, however this comes at the cost of only using dark time images when making a reference grid. I show that the benefits of using one method over another is negligible and that reference image standard is mostly dependant on number of images being stacked and quality of said images. Going forward GOTO will need to decide which method is more advantageous, fewer dark time images, or more lower quality images.

Having now developed a way to build reference images, the next stage is to test image subtraction. The following chapter will introduce image subtraction using the results from this section to help find transients in real time.

Chapter 5

Image Subtraction

“When you’re looking for a needle in a haystack, don’t be afraid to burn the haystack to save yourself from spending half your life picking through strands of straw.”

— A.J. Darkholme, *Rise of the Morningstar*

Image subtraction was briefly mentioned in 2.3.3. This chapter will take a comprehensive look at image subtraction and its uses in astronomy. I will then detail the specific motivations for developing a new image subtraction tool for GOTO and demonstrate the advantages of this tool for subtractions on large images with wide fields of view in real time.

5.1 What is Image Subtraction?

Even with a modest telescope, thousands of stars are captured with every exposure. Finding new or varying signals is near impossible to accomplish by eye and time consuming to cross check every source to a database. This problem is only exasperated when wide fields are introduced. Wide field surveys (e.g. GOTO, LSST, Black-GEM, ZTF, Pan-STARRS) are becoming common in the transient hunting game, all of them will see millions of stars per night. Finding changes in these fields on appreciable timescales is a current challenge. Image subtraction is a computationally frugal way to quickly spot changes in the dynamic sky. Image subtraction (or difference imaging) is the process of taking an image of the night sky (called a science image) and subtracting out a reference (or template) of that same sky position from an earlier epoch, using the residuals in the subtracted image to find changes in the frame.

Image subtraction has many uses in astronomy, from monitoring asteroids

(Gary and Healy, 2006), to fast identification of variable stars (Olech et al., 1999). For GOTO, difference imaging is intended to find transients associated with GWs. Similar strategies have been used for other GW follow-up missions (Kasliwal et al., 2016; Morokuma et al., 2016); however, as I highlight later, the methods used here do not make optimal use of the data and are not designed to handle wide fields.

Future surveys, like the LSST, intend on having a real-time transient alert system (Kantor, 2014). This would require finding changes in the field on sub-minute timescales. With the current software in the literature, these timescales are not achievable. The LSST images will be too large for current difference imaging tools to complete the subtraction on the necessary timescales. Methodologies should be in place before LSST comes online.

The first attempts of difference imaging were conducted in Tomaney and Crotts (1996). In this paper, the authors describe using a simple algorithm to monitor gravitational lensing events. This technique involves using a ratio of Fourier transforms from the two frames and a Gaussian profile estimate to model the wings of the sources in the image. This in essence captures the primary challenge of difference imaging, Point Spread Function (PSF) estimation. Because Tomaney and Crotts (1996) was an initial study into image subtraction, the process was closer to an improvised method which made results sub-optimal and inconsistent. Alard and Lupton (1998) became the first study to address the idea of a rigorous PSF estimation technique for image subtraction. The formulation of the technique has the following structure:

$$D = I - (T \otimes K) \quad (5.1)$$

Where the initial image (I), is subtracted from T (the template) convolved with the PSF kernel (K). The paper has additional discussion about PSF variations across the image and accounting for background noise. However, the latter consideration plays a less substantial role in the quality of the subtraction at this stage.

In Alard and Lupton (1998), the PSF kernel is derived using a linear regression method combined with a shape function based on a simple 2-dimensional Gaussian with n components:

$$K(x, y) = \sum_n \sum_{d_n^x} \sum_{d_n^y} a_n \exp\left(-\frac{x^2 + y^2}{2\sigma}\right) x^{d_n^x} y^{d_n^y} \quad (5.2)$$

Depending on how this is implemented, this style of estimation is unintuitive and therefore relatively quick to compute. However, as shown later, it is not viable for highly varying PSFs. The first open source image subtraction tool of note was

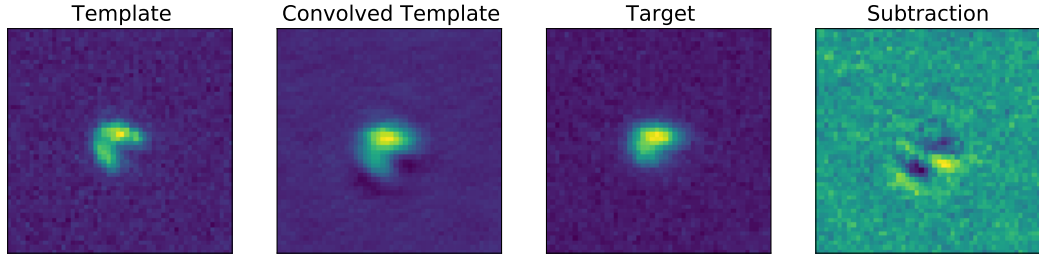


Figure 5.1: HOTPANTS subtraction with complex PSF. The convolved template should match the target PSF. This is not the case, resulting in a prominent residual being left in the subtraction image.

based on the above and has gained traction when searching for transients with image subtraction pipelines (Kasliwal et al., 2016; Morokuma et al., 2016). This utility is called **HOTPANTS** (Becker, 2015) and is an almost direct implementation of the Alard and Lupton (1998) method, with a slight twist on coping with PSF variation. In Alard and Lupton (1998), the variation is handled by sampling the PSF for almost as many times as there are sources, this is computationally intensive and would take inordinate amounts of time to solve. **HOTPANTS** by default will split the image into stamps of 100×100 pixels and use that to model a spatially varying kernel.

HOTPANTS was first implemented in Becker et al. (2004), here the benefits of using image subtraction to quickly locate transients and variable stars is articulated. Soon after, the source code was made openly available. Once the kernel has been calculated, the subtraction from equation 5.1 is computed. There are 3 fundamental issue with this method:

1. The PSF model is not robust enough to handle GOTO’s complicated source geometries (see figure 5.1)
2. The subtraction is too slow for sub-minute subtractions on large images
3. The convolution results in an underestimation of the noise

Point one is a clear concern for GOTO. An incorrect PSF estimation can leave residuals in the field that clutters the difference image with false positives. Point three was only recently addressed in Zackay et al. (2016), which has created a resurgence in image subtraction algorithms. Because the noise is not accounted for correctly in **HOTPANTS**, the difference images do not see as deep as the science images. This results in a lot of wasted data. Even with the package’s short comings, **HOTPANTS** is still the most widely used subtraction software in the literature; however, the new algorithm described in Zackay et al. (2016) has shown to make

optimal use of the data and is beginning to overtake the competition (Zackay et al., 2016; Sánchez et al., 2018; Masci et al., 2018).

5.1.1 Introducing ZOGY

Proper image subtraction (from here on called ZOGY, named after the three authors of the original paper Zackay, Ofek, and Gal-Yam (2016)) is a statistical approach to image subtraction. The three main equations are as follows:

$$\hat{D} = \frac{(F_r \hat{P}_r \hat{N} - F_n \hat{P}_n \hat{R})}{\sqrt{\sigma_n^2 F_r^2 |\hat{P}_r|^2 - \sigma_r^2 F_n^2 |\hat{P}_n|^2}} \quad (5.3)$$

This is referred to as the *proper difference image*. D here is equivalent to D in equation 5.1. A hat indicates a 2-dimensional Fourier transform. Subscript r denotes the reference (or template) image, and subscript n is the new (science) image. P is the normalised PSF (the PSF sum to unity), F is the flux-based zero-point, and σ is the standard deviation of background. Equation 5.3 creates two images in the numerator, these are the flux normalised science image matched with the reference image PSF in Fourier space ($F_r \hat{P}_r \hat{N}$) subtracted by the flux normalised reference image matched with the science image PSF in Fourier space ($F_n \hat{P}_n \hat{R}$). The subtraction creates a flux normalised difference image in Fourier space which is scaled using the denominator.

$$S = F_d D \otimes \overleftarrow{P_D} \quad (5.4)$$

S is then determined from parameters defined by the D image. $\overleftarrow{P_D}$ is defined as the coordinate reversed PSF, i.e. $\overleftarrow{P_D}(x, y) = P_D(-x, -y)$. S is a statistic mapping to the original science image, where the value of each pixel is the likelihood there is a residual in that space.

$$S_{corr} = \frac{S}{\sqrt{V(N_s) + V(R_s) + V_{ast}(N_s) + V_{ast}(R_s)}} \quad (5.5)$$

Finally, S_{corr} is the noise correlated statistic. By using noise correlation, one can see deeper; however, this term also results in a higher rate of false positives. V indicates the flux variance of the sources in each input image and V_{ast} is the astrometric variance of the sources. The derivation of the above can be found in appendix A.

Each equation produces a type of difference image. D being the equivalent of

a raw subtraction means the difference in flux can be taken from it. For transients this is very useful, as it means flux photometry can be done on the subtracted image. This is not the case for variables and caution should be taken when using subtracted images to find the flux of variable sources.

ZOGY has already shown to be better than HOTPANTS in subtraction performance (Sánchez et al., 2018; Masci et al., 2018), this result is reiterated in chapter 5.3. However, there is a drastic difference in computing time. The ZOGY algorithm at best takes approximately 4 times longer than HOTPANTS (Ofek, 2014; Vreeswijk, 2017; Sánchez et al., 2018). This is not suitable for real time transient hunting. Additionally, the current algorithms use a global PSF estimate, the model is made from the entire image as a whole. This way of modelling the PSF is insufficient for wide fields and perpetuates the problem highlighted by figure 5.1.

Now that I have covered the background of this chapter, the following section details the development and implementation of ZOGY in Parallel (ZiP), an algorithm I designed to complete optimal subtraction on exposure time equivalent timescales, with detailed PSF estimation. The software has been professionally packaged^{1,2}, is openly available, and has all the documentation available online³.

5.1.2 ZOGY in Parallel

Achieving image subtraction on exposure length timescales is important for a few reasons. Scientifically, this kind of speed allows for follow-up of short lived events like stellar flares that are only visible on the timescale of minutes (Doyle et al., 2018), or kilonovae whose spectra evolve on an hourly basis (Shappee et al., 2017). Logistically, by calculating subtractions quickly, one frees up computing resources. Allowing more processing in a shorter time frame. This prevents pipelines getting bottle-necked. Finding high profile transients fast is also vital in terms of competition, other surveys are also gunning for the event and given the nature of transient hunting, finding counterparts first can be beneficial.

The timescales of the current ZOGY algorithm for GOTO images (6000 by 8000 pixels) is ~ 4 minutes. This is more than the exposure time for an average set of GOTO images (3 minutes). Using this unaltered subtraction method will result in a backlog of processing, meaning most transients will not be highlighted the same evening they are detected. While HOTPANTS can perform the subtractions within a 1.3 minute timescale, the PSF and noise issues result in the data being poorly

¹GitHub - <https://github.com/GOTO-OBS/ZiP>

²PyPi - <https://pypi.org/project/zogyp/>

³Documentation - <https://zogyp-in-parallel.readthedocs.io/en/latest/ZOGY.html>

utilised.

ZiP is a parallelised version of the ZOGY algorithm. Using functions defined in Ofek (2014) and using the base infrastructure from Vreeswijk (2017) I built a new program that can see optimally deep, process in real time, and has comprehensive PSF estimation. This is done in two segments. Firstly, both the reference and science images are split into sub-images. This feature has many benefits. Wide fields of view can have high varying PSFs, by splitting the image into sub-images this makes estimating a PSF model simpler and faster. Splitting the images lowers memory constraint as the data arrays are smaller, this is useful for smaller machines that cannot handle large data arrays and lowers computation time when doing image transforms. The benefit with the most impact is that the sub-images can be processed in parallel. Here the PSF for each sub-image is modelled and their backgrounds are estimated and subtracted simultaneously. The second speed improvement comes from convolving the variable PSF across the sub-images in segments. The PSF convolution and subtraction are computed in parts, but simultaneously, to provide the output images. This entire process is described in the flowchart 5.2.

Caution should be taken when analysing the subtracted images from this method as the background subtraction and PSF model will not be congruent across the sub-images. This is why the output sub-images are returned individually and not as one bulk image. To account for edge artefacts, each sub-image has some overlap, this prevents loss of transients that occur at the edge of a sub-image.

Additional adjustments to the algorithm were made, eq 5.5 was found to over-estimate the variance in source centroid alignment. It was found a better estimate of alignment accuracy was the FWHM. Giving the following:

$$S_{corr} = \frac{S}{\sqrt{V(N_s) + V(R_s) + V_{fwhm}(N_s) + V_{fwhm}(R_s)}} \quad (5.6)$$

Another feature added to ZiP, more for practical convenience, is masking parts of the science image that do not overlap with the reference. This feature is not included in other methods and has two benefits. The first is removal of false positives due to subtraction from nothing. The residuals left in the field where the images do not overlap are present because there is no data to subtract. Building automated software (like machine learners) to flag changes is made considerably harder when there are many false but realistic residuals in the subtracted image. The second reason is that the subtraction convolution has a smoother gradient.

To summarise, ZiP presents a solution to real time image subtraction at

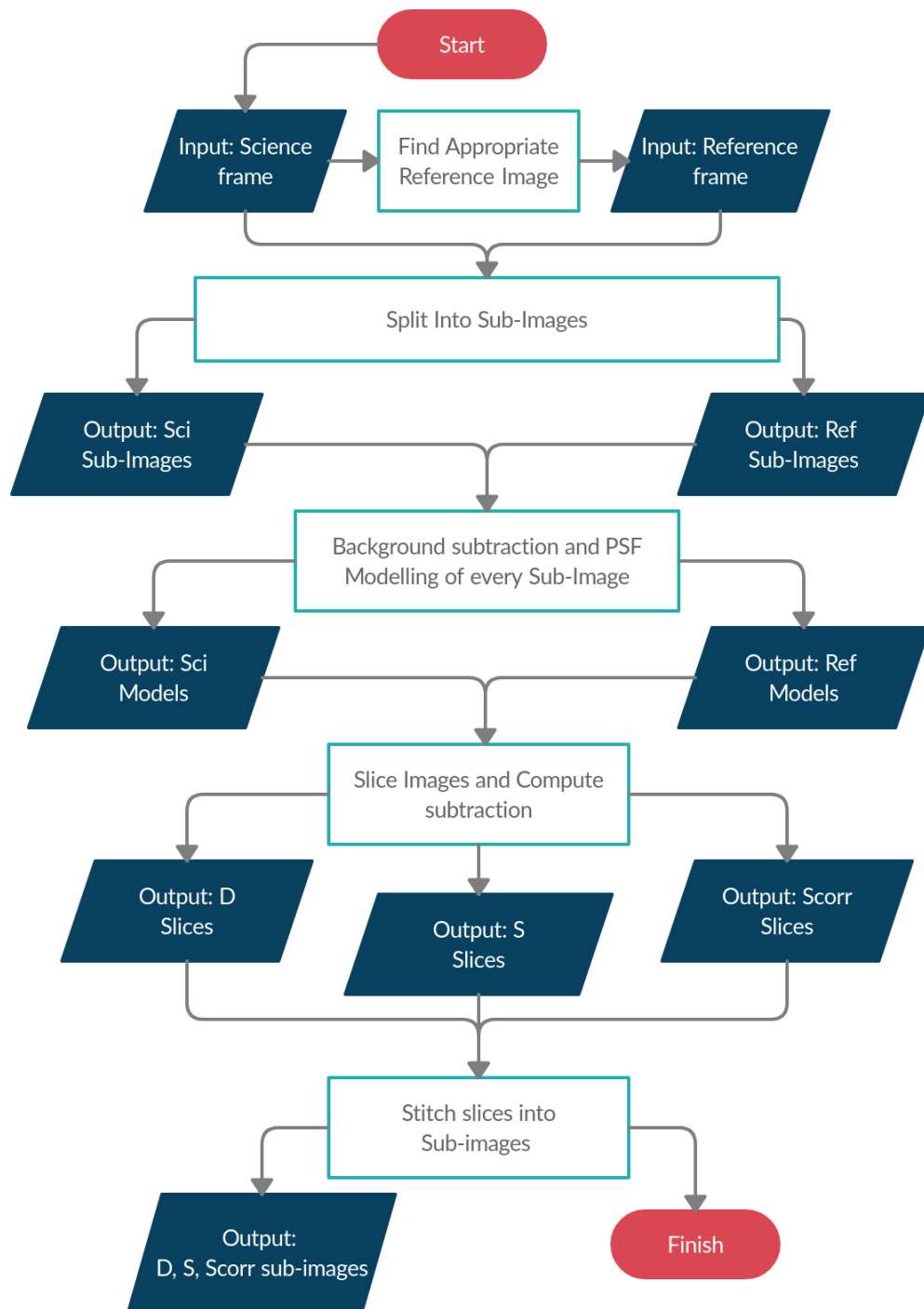


Figure 5.2: Flowchart for the ZIP proper image subtraction tool.

	HOTPANTS	ZOGY	ZiP
Real time subtractions	Y	N	Y
Optimal Depth	N	Y	Y
Comprehensive PSF fitting	N	N	Y
Handles Overlap	N	N	Y

Table 5.1: Qualitative checklist of Software performance. Showing ZiP meets all the basic requirements for real time image subtraction.

optimal depth for wide fields. Table 5.1 offers an overview of the benefits of the ZiP algorithm.

For image subtraction to work as intended, regardless of method, three conditions have to be met:

1. The PSF of both images must be modelled accurately
2. The sources in both images must be aligned so their centroids fall on the same pixels
3. The reference must be of equal or better quality than the science frame to make the best use of data.

The remainder of this chapter will work under the assumption that conditions 2 and 3 are already satisfied using the techniques from chapter 4.

One of the primary time consumers is PSF estimation. Given the vitality of this modelling process to the subtraction quality, it is valuable to test different methods for their viability; both in model accuracy and calculation time. The following discussion shows the different methods that I have developed and tested, then highlights the most applicable method which is then integrated into the ZiP framework. Once a PSF modelling technique is decided, section 5.3.2 will compare the capability of the new ZiP algorithm.

5.2 Modelling the Point Spread Function

PSF estimation is a vital component to image subtraction. The reason for this importance is so subtractions can be done cleanly. The PSF describes how a point source's flux is distributed across CCD pixels. A perfect point source would be expected to illuminate a single pixel, however due to a myriad of environmental and systematic distortions it is not realistic to assume this is the case. The PSF is

a response function, by convolving the expected image with the PSF the observed image is produced.

By modelling the PSF of one or both images, the source shapes can be matched when doing the subtraction. This mitigates false residuals in the image due to the source shapes not being equivalent. However, this fitting has to be more detailed with the GOTO prototype optics because of their significant distortions across the image. Equation 5.2 implemented in **HOTPANTS** is one example of modelling the PSF. This is a relatively common practice, summing ordered polynomials to build a kernel (Alard and Lupton, 1998; Bertin, 2013; Chan et al., 2013; Yin et al., 2017). Unfortunately, as shown in figure 5.1, a global PSF fit is insufficient for GOTO. This style of estimation also takes too long to compute if not done in parallel, which **HOTPANTS** does not have the functionality for.

As most PSF estimation methods are either too slow or too simple, a more versatile model needs to be in place. The following sections discuss models I have built or improved on to quickly model complex PSF functions for image subtraction in GOTO. The following have all been integrated into the ZiP workflow described in figure 5.2.

5.2.1 A Simple Bivariate Gaussian

Instead of using a polynomial based prescription and summing them to model the PSF, one can use measured parameters to quickly create a spatially varying kernel. One way of doing this is to employ a bivariate-normal distribution where X and Y are not correlated (also known as a 2D-Gaussian):

$$f(X, Y) = \frac{1}{\sqrt{2\pi}\sigma_X\sigma_Y} \exp\left(-\frac{X^2}{2\sigma_X^2} - \frac{Y^2}{2\sigma_Y^2}\right) \quad (5.7)$$

Using parameters described in section 3.2.4, a basic spatially varying kernel can be built. Using **SExtractor**, the Full Width Half Maximum (F), elongation (ε), and rotation (θ) of the sources can be catalogued along with their position in the field. The average F , ε , and θ are all estimated using the positional median turning the parameters into functions dependent on position ($\bar{F}(X, Y)$, $\bar{\varepsilon}(X, Y)$, $\bar{\theta}(X, Y)$ see figure 5.3 as an example.) Once the values have been ascertained they can be applied to recreate the source shape.

Using the known relation between the FWHM and σ :

$$\sigma_X = \frac{F}{2\sqrt{2\ln(2)}} = \frac{F}{2.355} \quad (5.8)$$

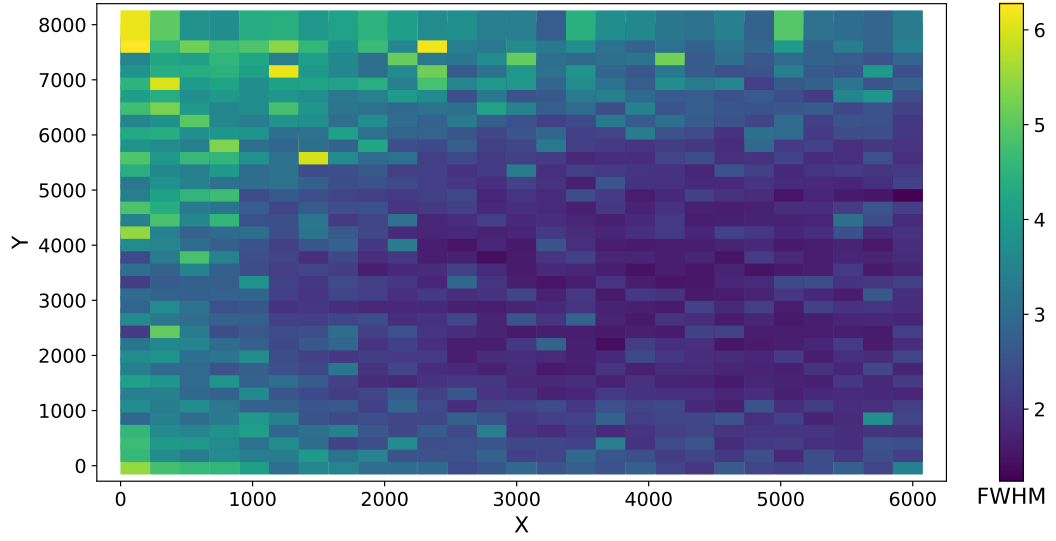


Figure 5.3: The Full-Width Half Maximum as a function of X and Y (the CCD axes). This highlights that for telescopes like GOTO, the FWHM can vary by a factor 3 across the image.

From equation 3.2, the ratio of the two lengths can be determined and applied to find the length of the longer perpendicular component of 5.8:

$$\sigma_Y = \frac{F\varepsilon}{2.355} \quad (5.9)$$

Given that the sources can rotate, (i.e. will not necessarily be aligned with the axis of the image) the final variable (θ) can be introduced for X and Y:

$$X = \sqrt{x^2 + y^2} \cos(\theta + \phi_p) \quad (5.10)$$

$$Y = \sqrt{x^2 + y^2} \sin(\theta + \phi_p) \quad (5.11)$$

Where lower case x and y indicate a PSF centred co-ordinate system and ϕ_p is the phase angle of x and y. As the co-ordinates of x and y are arbitrary, the most stable implementation is to use the *arctan2* function:

$$\phi_p = \arctan2(y, x) \quad (5.12)$$

This means for any x and y a PSF can be built. Taking equations 5.8, 5.9, 5.10, and 5.11 and substituting them into 5.7 with some simple algebraic rearrangement gives:

$$K(x, y, F, \varepsilon, \theta) = \frac{1}{\sqrt{1.133F^2\varepsilon}} \exp \left(-2.773 \frac{x^2 + y^2}{F^2} \left(\frac{\sin^2(\theta + \phi_p)}{\varepsilon^2} + \cos^2(\theta + \phi_p) \right) \right) \quad (5.13)$$

5.13 can now model PSF kernels with spatial variation. Irrespective of complexity, the estimation takes an average of 10 seconds to complete. This is because the sections of the image are small and source features can be mined in parallel.

This way of building a PSF model is quick, it is faster than any modelling technique openly available, but there are a few major weaknesses. Providing the source shapes are all Gaussian, this method will not struggle reproducing the PSF. However, like in figure 5.4, if the PSF is non-Gaussian the model cannot work because the method makes a fundamental assumption about the source shape. One additional problem, flagged by figure 5.3, is the sampling function. Given the slices are small, the PSF features can be under-sampled. This means the median source can be influenced by a small sample. For example, figure 5.5 has messier 51 in the centre. As the galaxy is a dominating source in that segment of the image, the FWHM prescribed to that slice will be an over estimation. This means the PSF estimation will be incorrect for this slice. Using a 2D-interpolation across the function will help minimise this issue, but does not remove it completely. Overall, this method would work for well behaved fields, but more advanced techniques need to be in place for complex PSFs and low source counts.

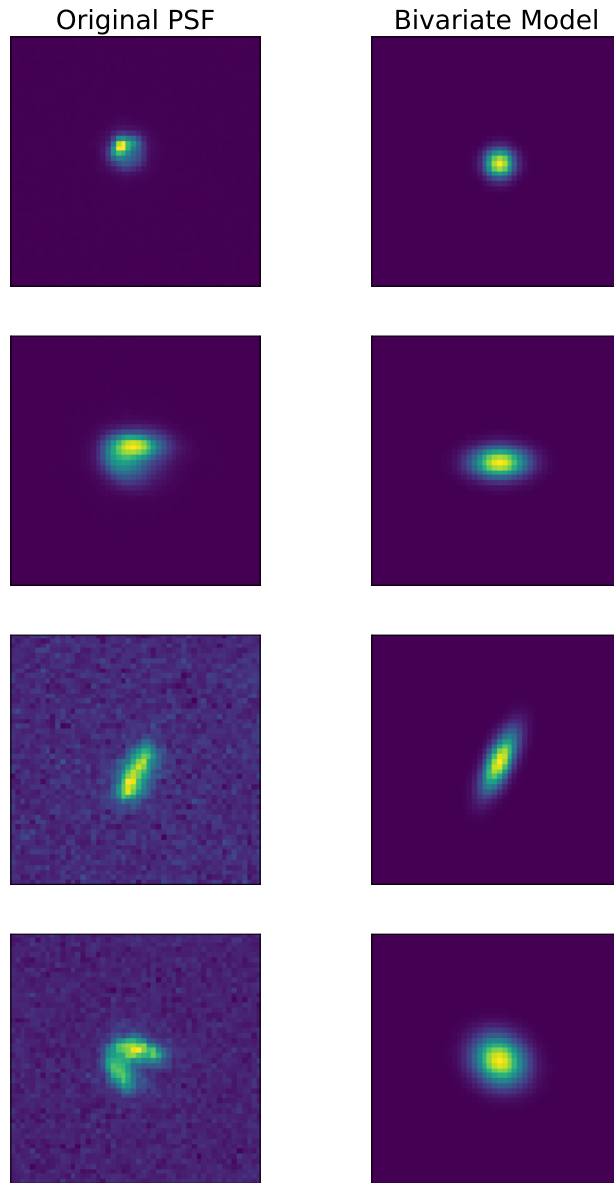


Figure 5.4: Visual of the Bivariate Gaussian method for recreating PSFs observed in real GOTO data. The method can recreate simple PSF geometries quickly, but cannot solve non-Gaussian shape functions.

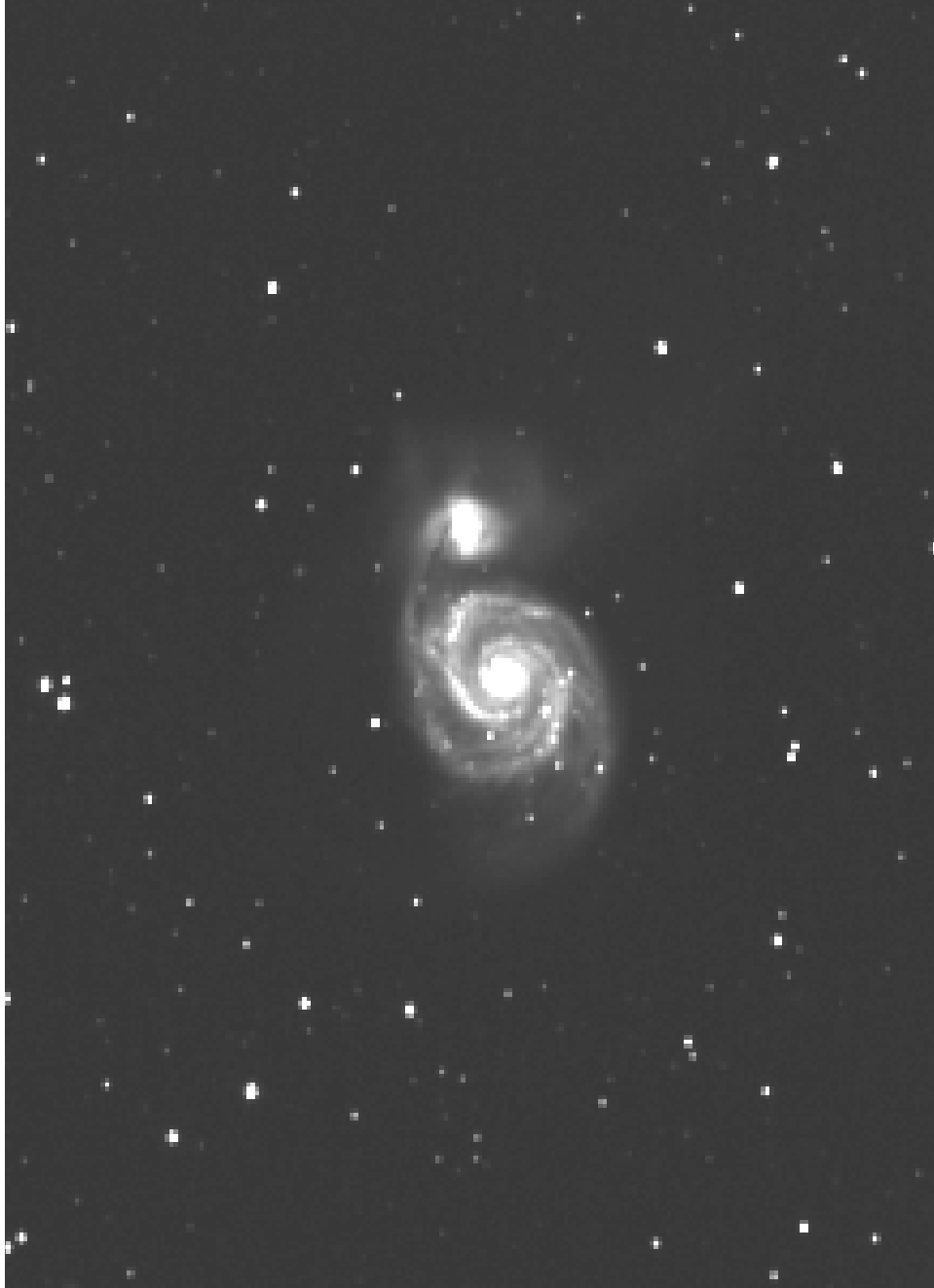


Figure 5.5: GOTO image of M51 (The Whirlpool galaxy). M51 acts as a contaminant when calculating the FWHM across an image which will break the Bivariate Gaussian method.

5.2.2 Zernike Moments as PSF models

Zernike polynomials were first developed in 1934 by Fritz Zernike (Zernike, 1934). They are used primarily when characterising optics. However, due to the nature of the polynomials having an almost orthogonal basis (no two polynomials have overlapping information) and the fact they form a complete set (meaning that they can recreate arbitrarily complex surfaces given a high enough order, see fig 5.6), Zernike polynomials have the potential to be a powerful tool when modelling source shapes. Zernike polynomials have the basic structure:

$$Z_n^m = \begin{cases} N_n^m R_n^m(\rho) \cos(\theta m) & \text{if } (m \geq 0) \\ N_n^m R_n^m(\rho) \sin(\theta m) & \text{if } (m < 0) \end{cases} \quad (5.14)$$

Where R_n^m is the Radial polynomial component and N_n^m is a normalisation factor.

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! (\frac{n+m}{2} - k) (\frac{n-m}{2} - k)!} \rho^{n-2k} \quad (5.15)$$

This works under the conditions that $|m| \leq n$ and $n - |m| = (\text{even})$. A Zernike moment is a function that maps an image onto a set of complex Zernike polynomials. Zernike moments can be used to describe image features and has been used in machine learning studies already (Rao et al., 2013). As there are no base assumptions on source shape, the PSF, like modelling surfaces, can be arbitrarily complex and entirely described by these Zernike moments.

Unfortunately, most data arrays cannot be described simply using polar co-ordinates. One has to rewrite the equation so it matches the form of the data structure. This means eq 5.14 can be expressed in Euclidean (pixel) co-ordinates as such:

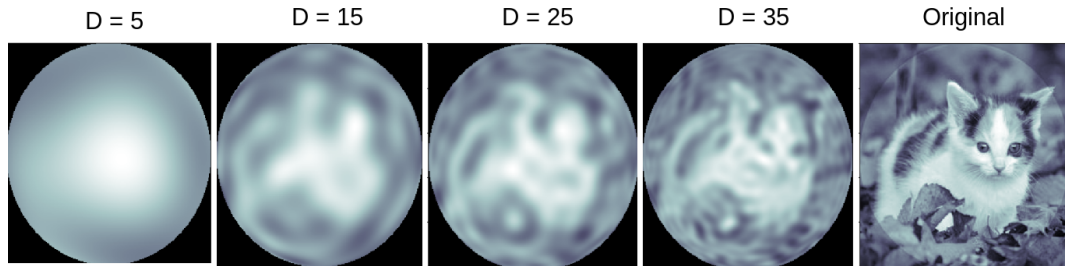


Figure 5.6: Using Zernike moments to reproduce arbitrarily complex 2D-surfaces

$$Z_n^m(x, y) = N_n^m R_n^m(1) \sqrt{x^2 + y^2}^{(n-2k)} \left(\frac{1}{\sqrt{\frac{y^2}{x^2} + 1}} + \frac{iy}{x\sqrt{\frac{y^2}{x^2} + 1}} \right) \quad (5.16)$$

Where:

x = the PSF centred pixel coordinate in the image row

y = the PSF centred pixel number in the image column

The moments are taken using a modified version of the algorithm discussed in Hwang and Kim (2006) that I wrote to reconstruct the PSF kernel:

$$K(x, y) = \sum_{n=0}^D Z_n^m(x, y) \quad (5.17)$$

Where D describes the maximum polynomial order (or degree). Below are some examples of how well this method performs when estimating PSF shapes.

The sources in figure 5.7 are normalised so they sum to one. The thumbnails are colour-scaled to the brightest pixel, this is why the fringes of the image look brighter in some of the examples.

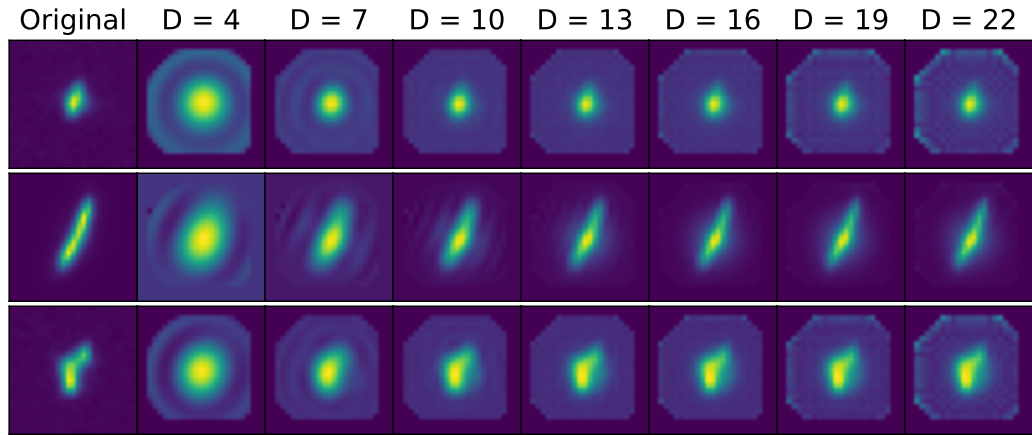


Figure 5.7: PSF estimation of different PSF complexities using Zernike moments. This figure highlights the power Zernike moments have when recreating PSF shapes.

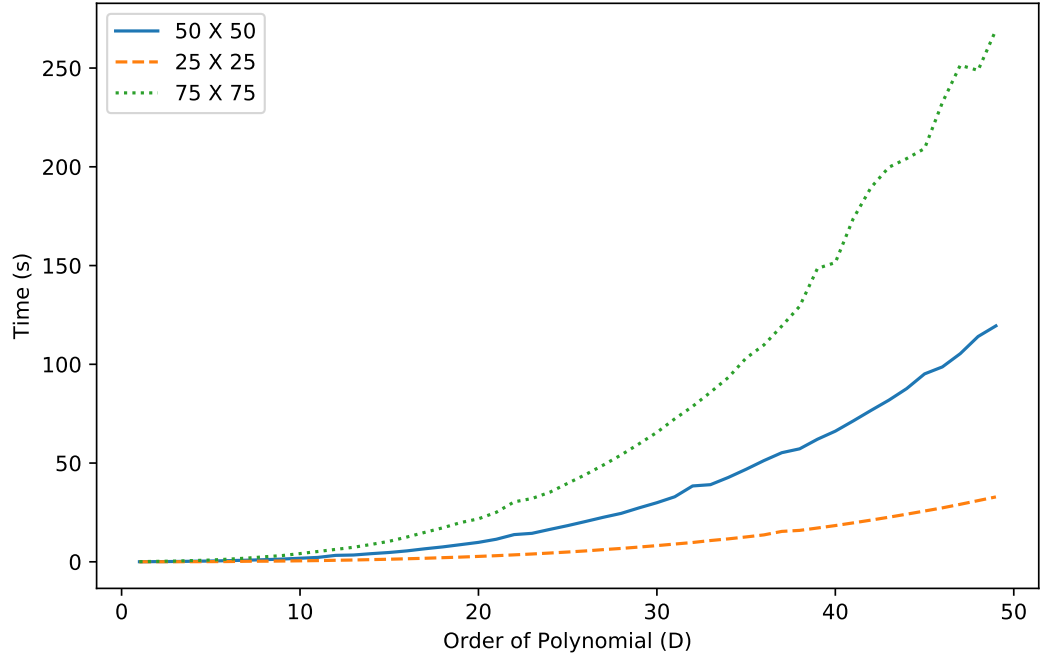


Figure 5.8: Time taken for different degrees of Zernike polynomial to model a PSF. This shows that for smaller thumbnails (well contained PSFs), a model PSF can be made on sub-minute timescales. However, for large complicated PSFs, the time to calculate a model is too long for real time image subtraction

Zernike moments offer a powerful PSF model, showing it is capable at recreating complicated shapes. There are still a few major issues. Over fitting becomes a quick problem. PSF complexity defines at what order is best; however, as D gets larger, the model becomes more unstable. This means an optimal way to determine what order the model should use needs to be developed. In addition to finding an automatic way to determine the optimal order, there is a temporal issue. The larger the PSF kernel the more stable the model at higher orders. However, larger kernels take exponentially longer computing times. This means the most suitable kernel is one that can be stable at large orders, but does not take too long to compute. There is scope to implement a parallel step to speed up the process, but this does not solve the stability issue. Because of the current instabilities and time consumption, it was deemed Zernike moments would take too long to optimise on the timescale of this thesis.

This means a method that can recreate non-Gaussian PSFs quickly is still needed. As PSF estimation tools already exist, the next option was to modify a current implementation into the ZiP framework.

5.2.3 Modified PSFex

PSFex is a PSF estimation tool (Bertin, 2013). It utilises **SExtractor** output parameters to model the source shapes across the image. PSFex has two components that I wanted to exploit. The first is that it works on a Gauss-Laguerre basis, also known as polar shapelets (Massey and Refregier, 2005). These were presented as a solution to the issue of characterising source shapes in a Cartesian basis function. The circular symmetry and orthonormal nature of the shapelet functions made them natural choices for modelling PSF shapes, not unlike the Zernike moments mentioned above. However, the circular symmetry introduced from the Gaussian assumption allows the profiles to be modelled quickly to a reasonable level of complication.

The second feature that made the code appealing is the capability of source selection and filtration when modelling the PSF. Figure 5.9 shows how PSFex uses relations between flux and source radius to automatically highlight the most appropriate sources for PSF modelling. This way of selecting sources can be reappropriated to help solve the biases for the other two methods discussed in 5.2.1 and 5.2.2.

Currently, this is all done with serial processing by building continuous polynomials across the image. With the rapid shape change across the GOTO images, building a smooth function across an entire frame is almost impossible. To alleviate this, I have integrated PSFex into a python wrapper and enhanced it so it can make multiple PSF models across the image in parallel. This not only increases PSF model performance, but lowers the amount of computing time. The models performance is enhanced as it is creating a smooth polynomial over a smaller area, there are fewer discontinuities to account for across the function. Additionally, the time decrease comes from solving multiple smaller models at the same time. This updated version was merged into the ZiP architecture.

This method of PSF estimation is suitable, both in handling complexity (see figure 5.10) and in amount of time taken to compute. As this is the most suitable method for now, this is the one that has been used for the rest of this work and used as the default for ZiP.

While testing ability to replicate the PSF is necessary, comparing how these methods cope with image subtraction is also vital. A good PSF estimation can qualitatively look adequate; however, the subtraction does a better job of testing suitability. Figure 5.11 demonstrates the same subtraction described in fig 5.1 using the Bivariate Gaussian and modified PSFex methods. While there is arguably a residual left in the subtracted image for both methods, they still demonstrate a notable improvement to the residual left from the HOTPANTS subtraction.

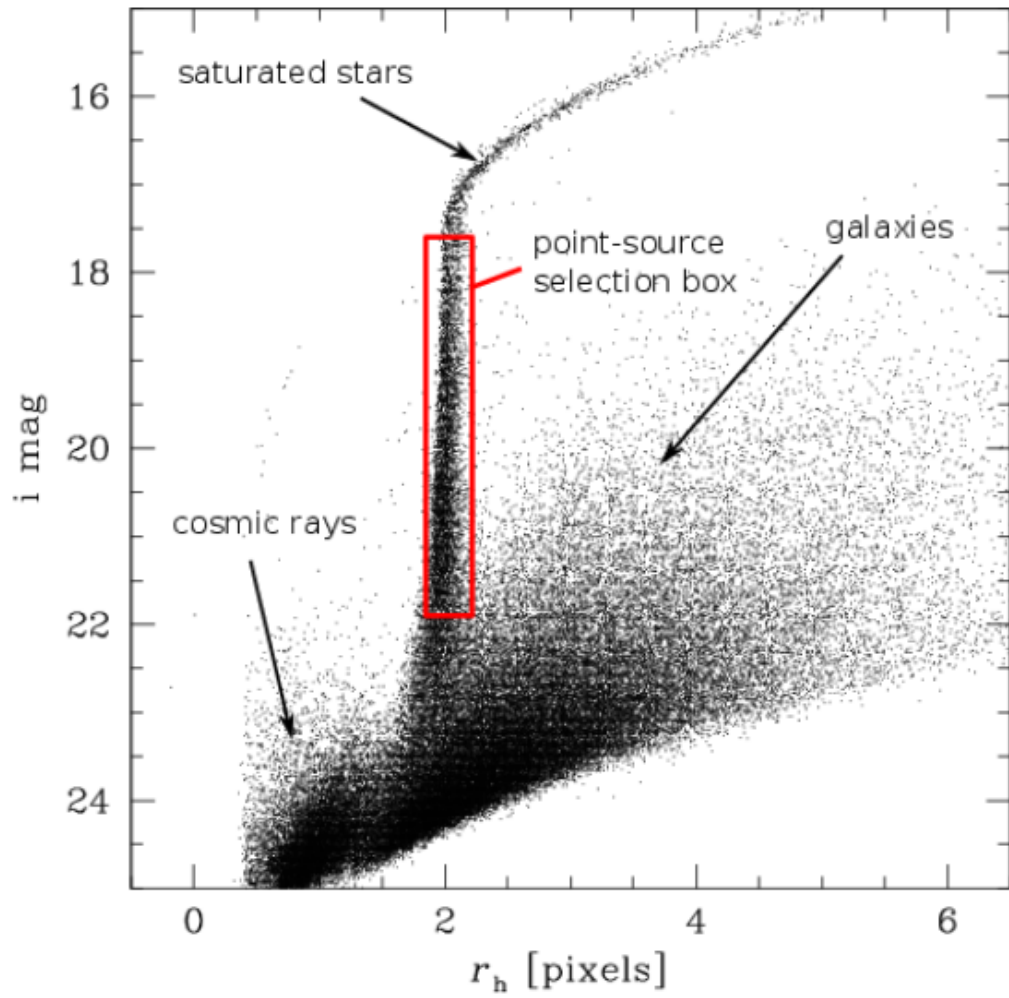


Figure 5.9: PSFex point source selection (Bertin, 2013). Any sources found in the red box are considered appropriate for estimating the PSF.

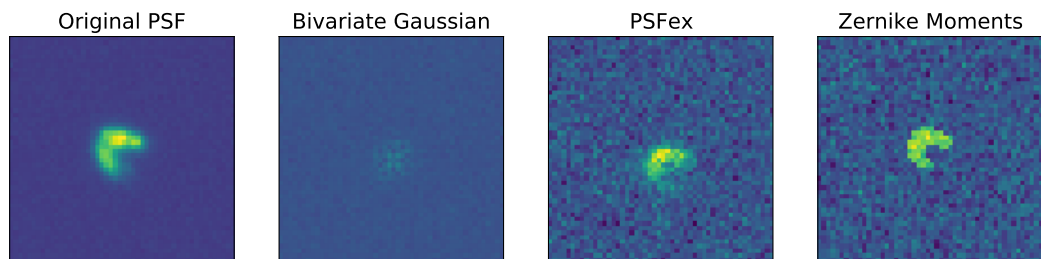


Figure 5.10: Comparison of each method at recreating a complicated PSF. The Bivariate Gaussian method failed here.

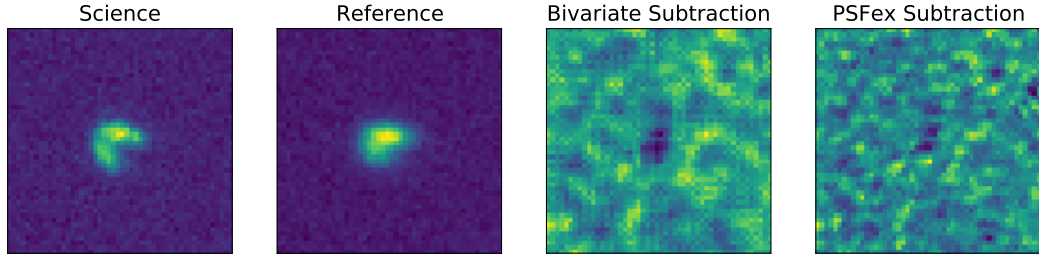


Figure 5.11: Comparison of the working methods using ZiP subtraction

5.3 Comparing Image Subtraction Software

This section will highlight the three main software options discussed throughout this chapter. That is, `HOTPANTS` (Becker, 2015), `ZOGY` (Vreeswijk, 2017)⁴, and `ZiP`. The implementation of `ZOGY` from Vreeswijk (2017) was not stable with the wide-field GOTO data. This may be in part due to how it estimates the PSF, but is most likely due to how images are processed and aligned by the algorithm. These PSF and registration features of `ZOGY` have been overridden to allow for direct comparison of subtraction techniques. Chapter 4 concentrated on the issues of registration and reference building.

Comparing the performance of the subtraction methods empirically is difficult as transients do not occur frequently enough to provide concrete analysis of the different methods. This means injection tests are required to adequately compare the different subtraction software.

5.3.1 Creating Test Data

A routine approach to injection tests and simulating data

Injection tests are a technical component to an already technical analysis, the literature is sparse for these kinds of tests and currently lacks a clear methodical approach. The aim of this subsection is to articulate the benefits of a package I developed, called `LazyPSF` (Cutter, 2019).

This tool was initially designed to simulate data expected from a telescope of given parameters. That is to answer the question, for a telescope with a given field of view, exposure time, pixel size, and noise parameters, what do the expected images look like? Looking at the literature, many examples of image simulation exist; yet,

⁴The `ZOGY` algorithm from Sánchez et al. (2018) was also trialled; however, the base code seemed to be broken. Additionally, when forced, the subtracted images were sub-standard for GOTO and NGTS data.

the methods were either unsuited for astronomy (Farrell et al., 2003), for specific telescopes (Dobke et al., 2010), or underdeveloped for the science needs (Bertin, 2009; Carbillet et al., 2010). This left me with no option, but to build my own image simulations. Image simulation draws many parallels with image subtraction, in that the PSF and noise characteristics must be modelled carefully. The noise can be modelled in multiple fashions. The first is by building an expected noise function. This is a user defined spatio-temporal function based on CCD location and exposure time. This is used to model dark current, and bias. Shot noise is defined by injected source flux and background signal (moonlight for example) is not accounted for as it can be considered part of an incoming signal. This was done to model the expected GOTO data performance before the telescopes were online. Once GOTO started commissioning data the software no longer needed to model noise statistics as these characteristics were being measured.

LazyPSF got its name as I was trying to model more realistic data performance, this required understanding GOTO’s PSF across the CCD. In the beginning, before GOTO’s optics were properly collimated, the PSF variation was extreme. This made it exceptionally difficult to provide a global PSF model when injecting sources into the simulated data. I resorted to building a simplified version (or precursor) of the modified PSFex technique discussed in section 5.2.3. This was an effortless way to model the variable PSF quickly and hence, **LazyPSF**. There are two features that make this tool especially noteworthy, the first is that it is written in python. No other open-source image simulation tools currently run using python. Python is the most commonly used language in astronomy (Momcheva and Tollerud, 2015), having easy to use tools readily available is beneficial for the community as a whole. More importantly, the tool is broad enough to be applied to any photometric optical telescope, highlighted later in 5.3.2.

As **LazyPSF** already exhibited the necessary infrastructure of injecting fake sources with realistic PSFs (figure 5.12), it was intuitive to refurbish the script for testing image subtraction.

Predicting magnitudes without a zero-point

One fundamental issue with the GOTO data, particularly when it first began commissioning data, was zero-point estimation. Due to the highly varying PSF and the infancy of the pipeline, the calculated zero-points of the images were regularly incorrectly estimated. Because the zero-points were unreliable, injecting fake sources based on this information often resulted in too faint sources. They were not re-

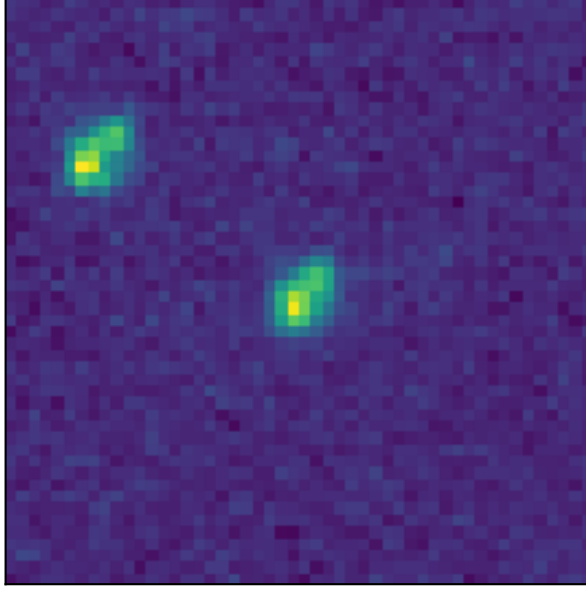


Figure 5.12: Example of source injection using **LazyPSF**. **Centre:** Injected source, **Left:** Real Source. This figure shows **LazyPSF** can inject realistic sources into GOTO data.

solvable in the subtracted image. This led to the development of a flux injection estimation method for given magnitudes. (It should be noted this work was done as an initial study testing GOTO's transient recovery as a function of magnitude. For the purposes of subtraction method comparison, relative fluxes would suffice.)

Assuming a point source, the injected flux can be computed using the exposure time (t), pixel scale (P), and calibration flux rate (C).

$$m - m_0 = -2.5 \log_{10} \left(\frac{F_{inj}}{F_0} \right) \quad (5.18)$$

Where $m_0 = 0$ and $F_0 = PtC$

$$F_{inj} = 10^{-m/2.5} PtC \quad (5.19)$$

For GOTO's L -band filter, C was found to be $1.51372 \times 10^{13} \text{ m}^{-2}\text{s}^{-1}$. This was obtained by taking the flux of a static source and using its magnitude to find F_0 .

From here, the point source is convolved with the estimated PSF to give the resultant injected source (see fig 5.12).

As the image noise is already included in the image, the injected source does not need to account for specific noise properties. However, shot noise does need to

be accounted for. This is done by adding to the injected flux the square root of the injected flux \pm a random integer between 0.01% and 5% of the square root of the injected flux to simulate some variation in the noise (in retrospect 5% may have been too high).

This method provides an upper limit for the injected flux. This is because **LazyPSF** cannot account for foreground seeing conditions such as cloud cover. This will result in a slight over estimation in magnitude limits from the subtraction tests. This method also breaks down for extended sources. If a source can be resolved in more than one pixel the base assumption of the convolution breaks, resulting in a poorly modelled PSF and drastic underestimation of the flux.

LazyPSF for image subtraction

To test the transient detection from image subtraction, one must already have some record of the transients they are looking for. **LazyPSF** does this by building a catalogue linked to the image file containing a list of injected sources. Detailed are the sources locations in x and y pixel coordinates, the flux, and optionally the apparent magnitude. Fake stars can be inserted in many ways and many configurations. One fundamental test is placing the sources in a fixed grid. This can be used to verify PSF model performance and allows an eyeballing run for any obvious errors in either the injection step or subtraction process before doing any arduous data analysis. The second method injects stars stochastically, this is also done blindly so human-bias does not influence the success of detection (I do not know where the transients are located). In addition to source position, the input flux is also generated. This can be as a static flux (every injected source has the same brightness), as a function of the source flux distribution (figure 5.13), or as a flat distribution between two user defined limits (figure 5.14). The latter should be between the bright-limit and magnitude-limit of the telescope. Here the optimal test was to use a randomly distributed set of source positions, with a flat distribution of magnitudes.

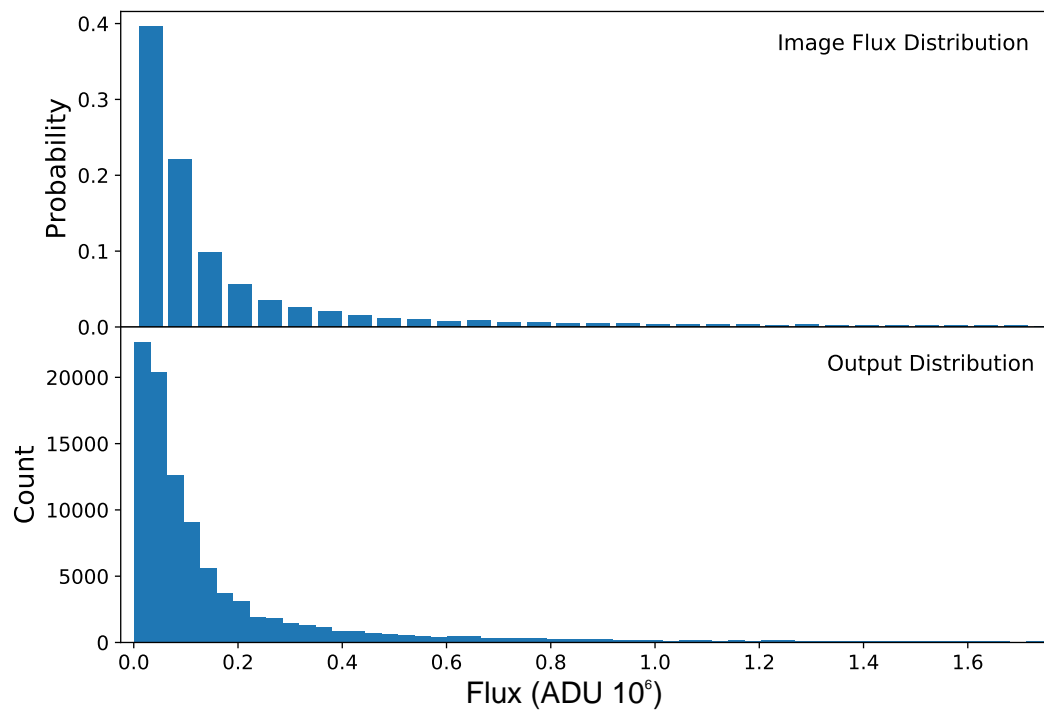


Figure 5.13: Distribution of injected sources based on measured flux distribution

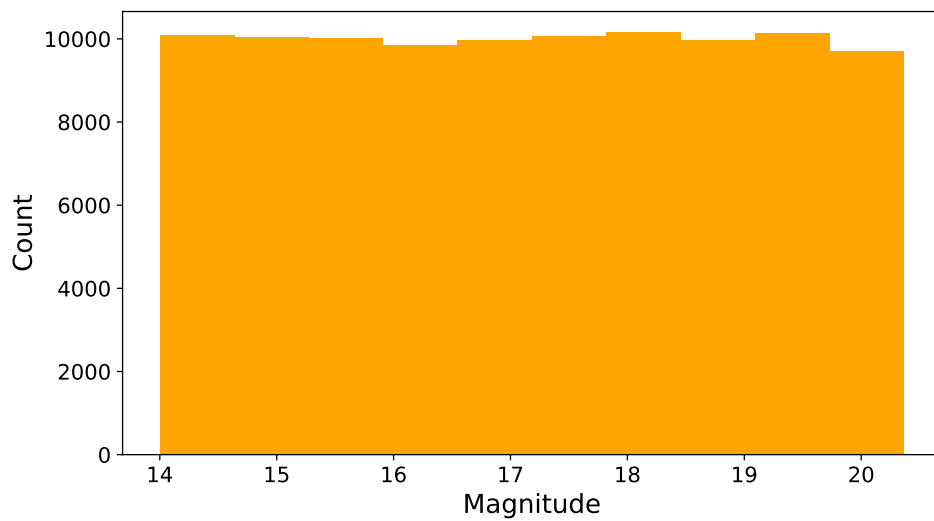


Figure 5.14: Flat user-defined distribution for source injection

Building test data sets

When building test data, other practical considerations need to be made. Source density in the field is one dependence of time consumption, busier fields take longer to process and subtract. Source density also plays a role in PSF estimation and background estimation. Therefore the test data should have a mix of densely packed fields and sparse fields. Furthermore, many transients occur atop galaxies, having fields where a source is likely to land on a galaxy is also important.

The three fields selected for the test set are:

- M31 - Andromeda, a highly dense field
- M51 - Whirlpool galaxy, a relatively dense field
- The VIRGO galaxy cluster - a sparse field with a lot of galaxies.

These fields are of particular value as GOTO has revisited them over several epochs, providing a surplus of viewing conditions in addition to the different levels of source crowding.

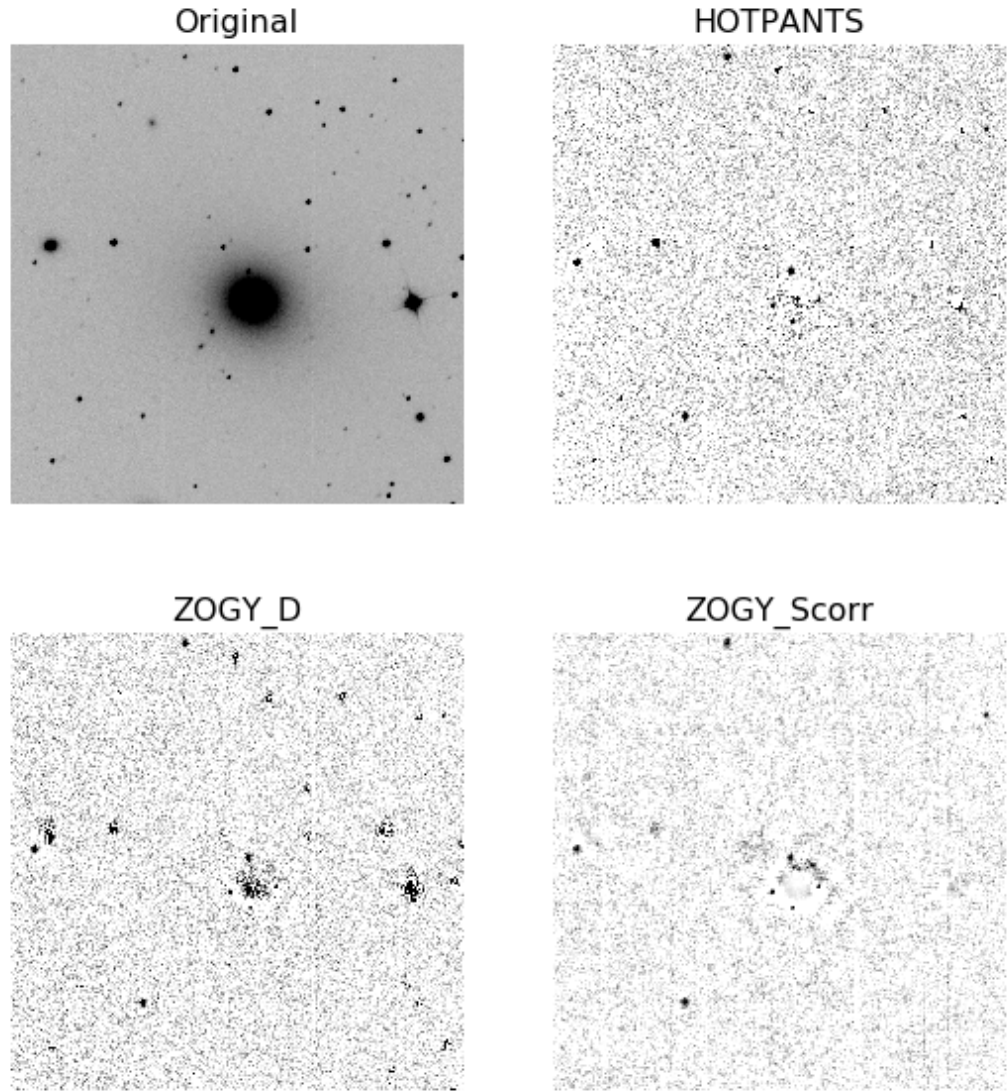


Figure 5.15: Visualisation of the subtraction methods. Using a galaxy with sources injected using IRAF. Qualitatively, the $ZOGY_D$ image is the lowest quality (i.e. will flag the highest number of false positives). HOTPANTS is the second, but the cleanest subtraction looks to be $ZOGY_{Scorr}$.

5.3.2 Transient tests

Results

Using the test data set described above, image subtraction is computed on each field. Each method is then compared on their ability to detect 3σ transients with **SExtractor**. The criterion for a successful detection is that the extracted centroid be within 2.5 pixels of the injected source. Anything wider than this is treated as a false positive. It is expected that GOTO’s magnitude range for this test 14 – 19.5; however, the range was increased beyond this to see how quickly transient detection decays as sources get fainter. The results are in figure 5.17

Two other statistics are useful here, Contamination and Completeness:

$$Contamination = \frac{False\ Extractions}{Total\ Extractions} \quad (5.20)$$

$$Completeness = \frac{Real\ Extractions}{Total\ Injections} \quad (5.21)$$

These are regularly used in machine learning to measure the performance of a binary categoriser (Ivezić et al., 2014). These are useful here as Contamination is the ratio of false positives compared to correct predictions and Completeness is the fraction of sources correctly identified. In other words, Contamination informs us of the percentage of false extractions one should expect with each method and Completeness is the percentage of transients in your field one expects to find.

It should be noted, the statistics in 5.2 do not consider the fact that the magnitude range goes beyond GOTO’s limits (resulting in a lower completeness) and does not account for real transients in the field (resulting in a marginally higher contamination). $Comp_{lim}$ is included to show the completeness of each method at

Software	False Positives	Contamination	Completeness	$Comp_{lim}$
HOTPANTS	291389	0.9626	0.7498	0.8148
ZOGY _{Scorr}	812336	0.9846	0.8411	0.9185
ZiP _D	307414	0.9657	0.7235	0.7915
ZiP _{Scorr}	837004	0.9851	0.8415	0.9192

Table 5.2: Image subtraction test statistics. Contamination and Completeness are defined in equations 5.20 and 5.21 respectively. $Comp_{lim}$ is the completeness at the limiting magnitude of the telescope.

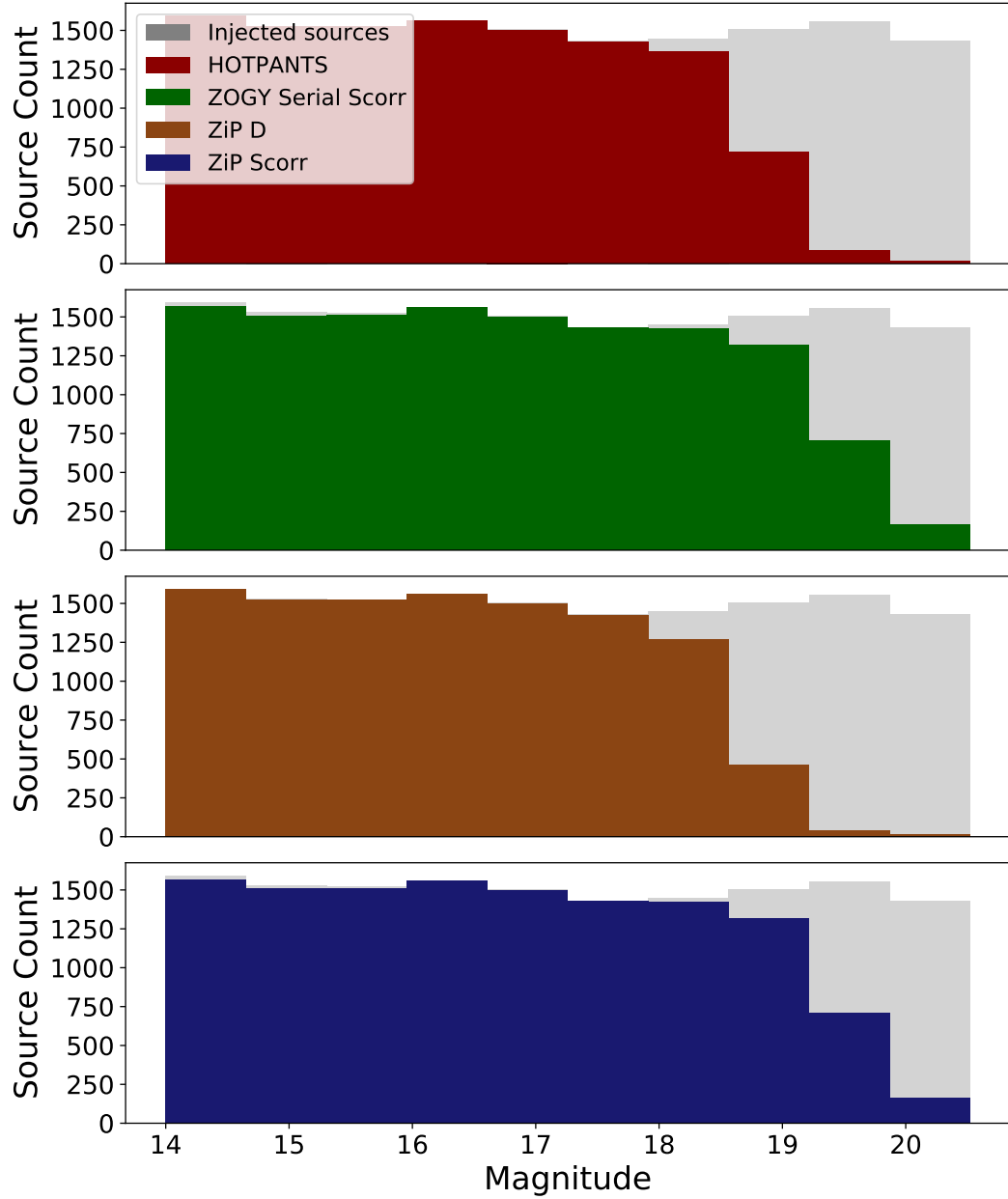


Figure 5.16: Injected source recovery as a function of magnitude. The top figure (red) shows HOTPANTS performance, second (green) is the original ZOGY script, and finally the ZiP raw subtraction (orange) and ZiP noise correlated image (blue). As expected the original ZOGY and ZiP have similar transient extraction performances.

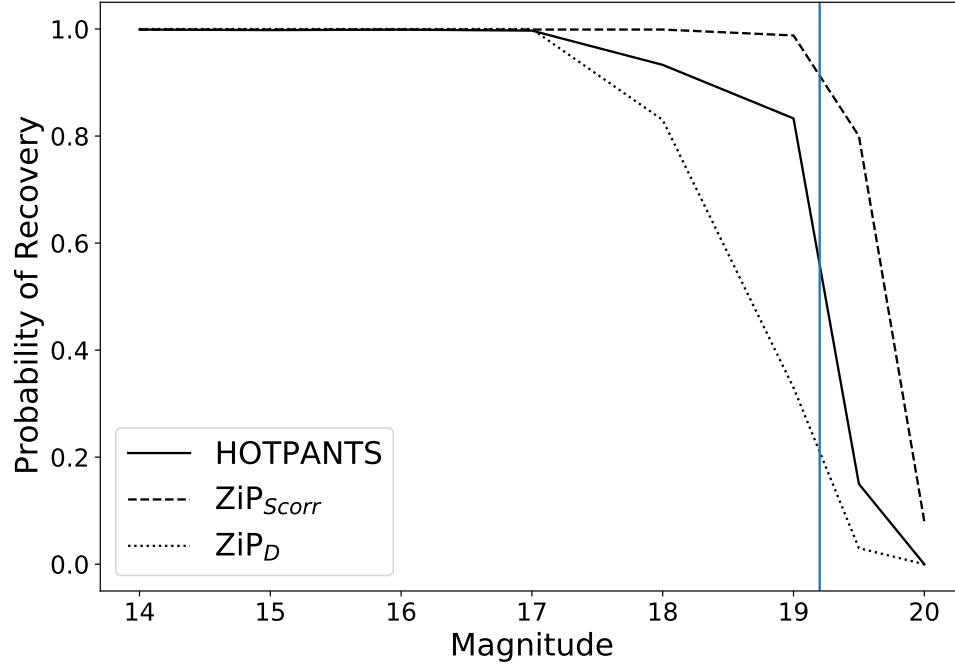


Figure 5.17: Probability of recovery as a function of magnitude. The blue vertical line indicates GOTO’s average limiting magnitude.

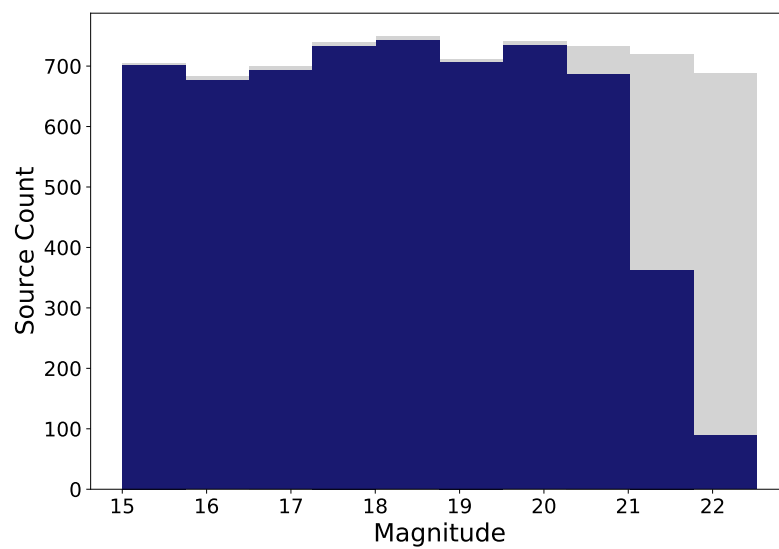
GOTO’s limiting magnitude.

Further tests were conducted for ZiP using the Next Generation Transit Survey (NGTS) and intermediate Palomar Transient Facility (iPTF) (Wheatley et al., 2017; Kulkarni, 2013). NGTS has much larger pixels and a smaller field of view, this presents notably different photometric challenges for ZiP. iPTF has smaller pixels than GOTO and, similarly to NGTS, a smaller FoV. These telescopes are representatives for general photometric surveys, figures 5.18 and 5.19 highlight that ZiP is still a strong software choice for smaller FoV time-domain survey astronomy.

Discussion

Here I have shown that the HOTPANTS method cannot perform subtractions to the same depth as the ZOGY method. This result has been corroborated in Sánchez et al. (2018). Showing ZOGY makes optimal use of the data by finding residuals to the magnitude limits of the telescopes.

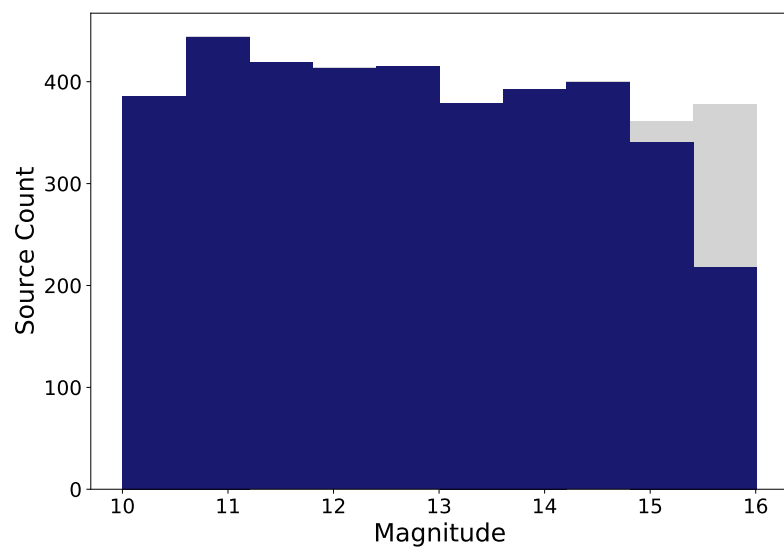
The primary take away from this section is that ZiP has an almost identical performance to the regular ZOGY method. This is not an entirely surprising result given that both methods employ the same numerical process. However, it is also important to note that while subtraction performance is the same, the time scales



(a) Transient recovery for iPTF data

Statistic	
Completeness	0.8541
Contamination	0.9443
False Positives	103787
Limiting Magnitude	>20.5

Figure 5.18: ZiP image subtraction tests with iPTF data



(a) Transient recovery for NGTS data

Statistic	
Completeness	0.9541
Contamination	0.5088
False Positives	3944
Limiting Magnitude	16

Figure 5.19: ZiP image subtraction tests with NGTS data

are notably different.

5.3.3 Speed tests

One of the primary aspects of this process was to create a tool that could produce subtracted images on a real-time basis. Below is a table of three methods.

Method	Time taken (s)	\pm (s)
HOTPANTS	73.8	1.2
ZOGY	260.5	4.8
ZiP	28.4	0.4

Table 5.3: Mean subtraction time taken over 100 subtractions.

Here, the issue of ZOGY’s subtraction duration being too long is cemented. Taking over four minutes per subtraction, this is too long to complete real-time subtraction on an entire nights worth of data. HOTPANTS takes a little less time than the exposures for GOTO. ZiP is notably faster than both methods, while the subtraction step still takes up a fraction of the exposure time; by preventing a back-up in the pipeline, there is now room for other steps.

There are additional parameters to test with the ZiP method. Each sub-image is processed in parallel, this means different slice patterns will yield different results, both in speed and subtraction quality. Using the slice patterns from figure 5.20, the results in table 5.4 are obtained. It should be noted, pattern E in figure 5.20 was the test used for ZiP in Table 5.3.

Table 5.4 highlights the biggest drawback for the ZiP method. Using too

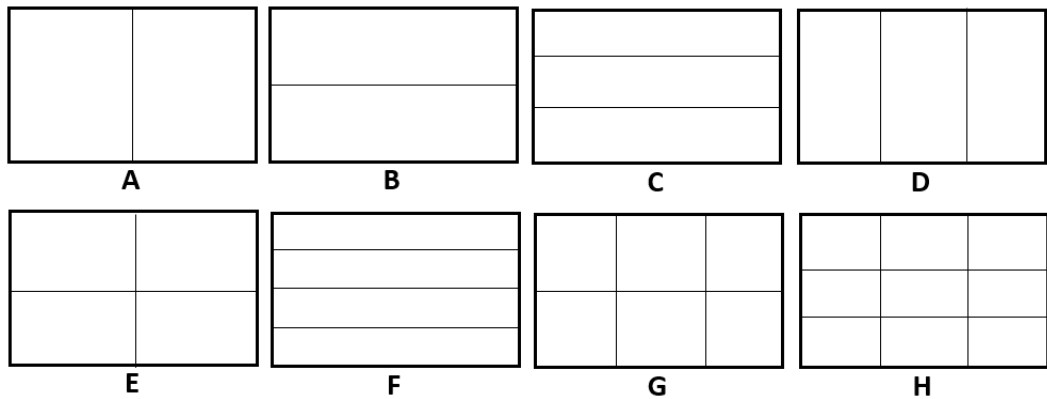


Figure 5.20: Slice patterns for table 5.4

Pattern (fig 5.20)	Time taken (s)	Contamination	Completeness
A	34.9	0.9726	0.8414
B	41.2	0.9746	0.8413
C	28.9	0.9867	0.8401
D	27.0	0.9854	0.8264
E	28.4	0.9851	0.8416
F	30.7	0.9820	0.8253
G	35.2	0.9967	0.6470
H	N/A	N/A	N/A

Table 5.4: Mean subtraction time taken over 100 ZiP subtractions for the different slice patterns (see figure 5.20). This shows that the different slice patterns can affect the subtraction quality and time.

many slices can result in an under-sampling of the PSF, this makes the subtraction unstable (in the case of pattern G) or unable to be completed (pattern H). There is a balance between number of sub-images for optimal subtraction and computing time, this will vary depending on the telescope.

Mock LSST speed tests

One of the challenges discussed at the beginning of this chapter was getting image subtraction to work on timescales equivalent to those expected from the LSST operation. Currently, the literature shows no contending algorithm that would be able to do this subtraction optimally and within the given time limit of ~ 30 seconds (Kantor, 2014). Using `LazyPSF` one can build synthetic data of LSST. From Gilmore et al. (2008) and Bradshaw et al. (2018) the camera specifications and CCD response functions can be modelled. One important feature of the LSST is that each CCD is comprised of $4K \times 4K$ pixels. Depending on how the systems ultimately processes these files will determine the optimal way to use ZiP. For the tests here, I have assumed each CCD is processed individually. Every exposure is expected to find roughly 25 000 sources. The injected sources were given a basic Gaussian profile, with a small amount of elongation as the sources approach the edges of the CCD for added modelling complexity. As the tests here are for computation time, noise accuracy and subtraction quality have been forgone.

If the LSST has viable computing facilities, each CCD can be subtracted in ~ 20 seconds. As seen in section 5.4, it is possible to stream multiple subtractions

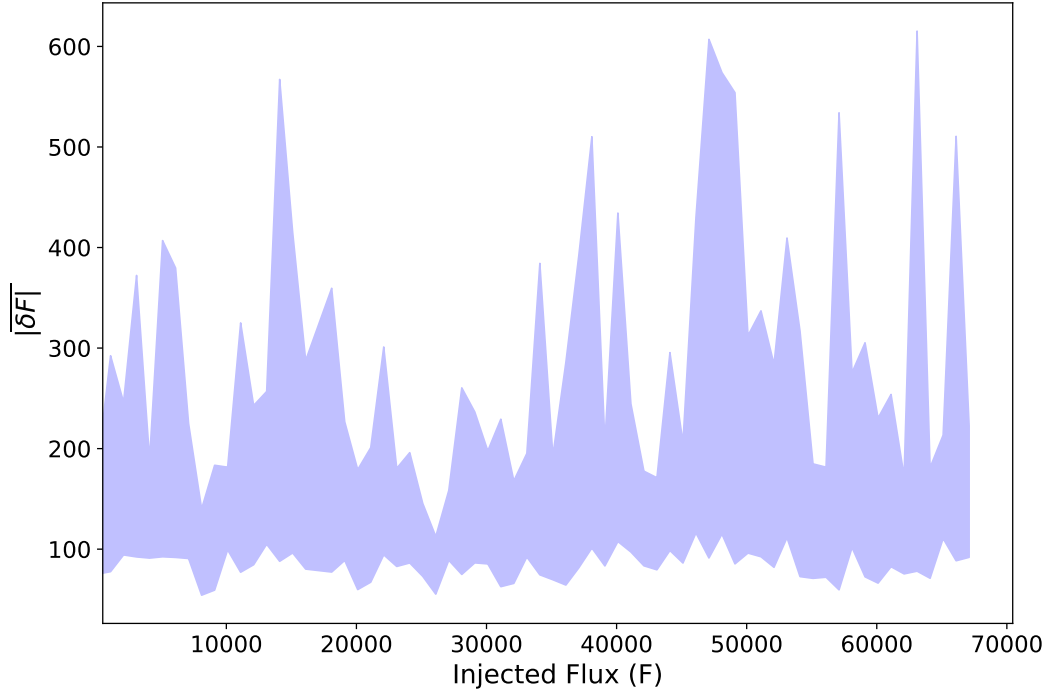


Figure 5.21: Median absolute difference in flux as a function of flux. Calculated from injecting transients, subtracting a reference, and calculating the flux of the injected transient in the subtracted image. Note, the error closely corresponds to the background flux of the reference image.

simultaneously. This means each focal plane array should take a minimum of ~ 20 seconds. This makes ZiP a powerful transient hunting tool for the LSST as alerts can be streamed in real-time to other astronomers.

Time test summary

This subsection has demonstrated the ZiP difference imaging algorithm completes subtraction in sub-minute timescales on GOTO data. This is important for pipeline purposes as it means not only can transients be flagged the same night they are observed, but can further improve the search strategy by allowing for same night survey updates when searching for counterparts to GWs. This section also acts as a pathfinder for LSST transient hunting algorithms, showing it is feasible to return high confidence transients in real-time with mock LSST data.

5.3.4 Flux Conservation

Even though ZiP has shown to be of comparable quality to methods used in the literature, one still needs to show that the D image conserves flux when looking at transients in the subtracted image. This test is important, as it shows there is no noise being introduced or information lost from the subtraction process. This was done by injecting transients with a given flux into a selection of science images and subtracting them. The aperture flux from the same position is taken in the D image and compared to the input flux.

Figure 5.21 highlights that the difference in flux obtained from the subtracted image is independent of the input flux. This means any change in output flux is most likely dependent on the image noise and any other transforms on the image.

5.4 Using ZiP on real data

Up until this point, the data used has been in some way modified or simulated. This section will detail the discovery of real events on untouched data using ZiP.

5.4.1 NGTS Flare Pipeline

NGTS is an exoplanet transit hunting survey. Transits of planets cause dips in stellar light curves, which is one way exoplanets get discovered. However, because of the high cadence and short exposures, NGTS has a few auxiliary science capabilities; the most major of which is searching for short lived transients like stellar flares (Jackman et al., 2018). Stellar flares are short lived events, lasting only minutes. The flare is associated with magnetic reconnection from the star's field (Benz and Güdel, 2010). Following flares is vital to understanding their properties. This in part is to test habitability of exoplanets around flaring stars (Segura et al., 2010). Fast identification is needed for rapid follow-up; which is necessary as these events last for such a short period.

NGTS in survey mode focuses on one patch of sky for a period of time repeatedly taking images. Each exposure lasts 11 seconds and has ~ 1 second upload time. Using a ZiP based architecture, I developed a flare hunting pipeline that alerts the user when a high likelihood candidate gets detected. Figure 5.22 shows how the pipeline works. Blue lines represent the incoming images. The black line highlights the flux of the source in question. The red stars indicate when an alert for the event is sent. Table 5.5 shows the supplementary information to the graph.

Each subtraction takes on average a little under 17 seconds. The pipeline was

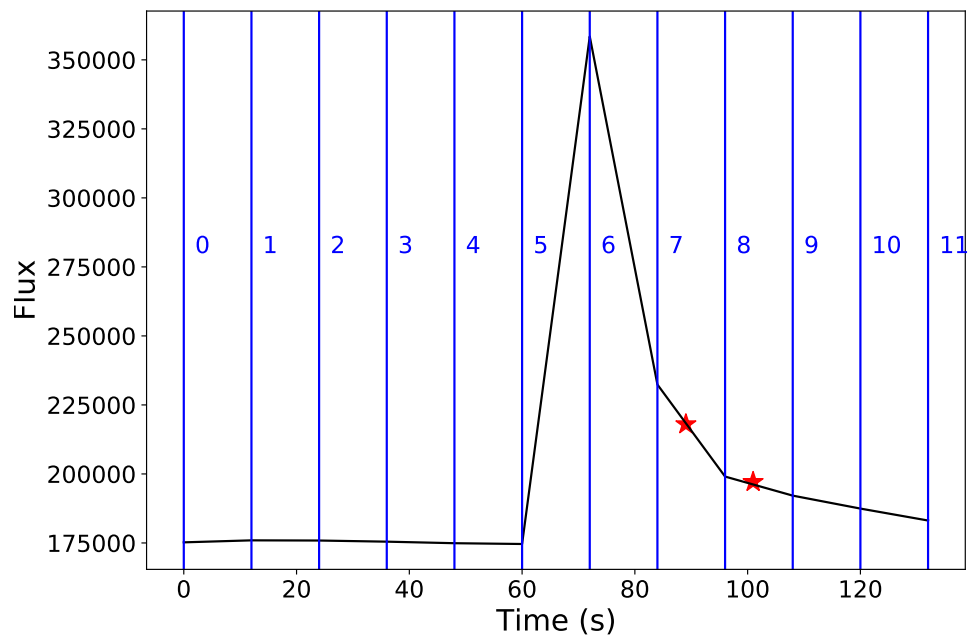


Figure 5.22: NGTS flare pipeline. The black line represents the flux of the source. Each blue vertical line indicates when a new exposure has been taken, with the number on the right corresponding to the exposure number. The red stars are when alerts are sent for follow-up.

Exposure #	Time taken (s)	False Extractions	Flare Detected
0	18.05	1	No Flare to Detect
1	15.99	2	No Flare to Detect
2	15.44	1	No Flare to Detect
3	15.60	1	No Flare to Detect
4	16.92	3	No Flare to Detect
5	16.37	4	No Flare to Detect
6	17.05	2	Yes
7	16.97	3	Yes
8	16.63	18	No
9	15.84	7	No
10	15.53	31	No
11	15.67	26	No

Table 5.5: Supplementary table for the NGTS flare pipeline (figure 5.22).

written in such a way that multiple subtractions can be computed simultaneously, one subtraction does not have to wait for another to end. This prevents bottlenecks. This method is only sustainable if the cores are available.

By being able to compute subtraction in real time on NGTS data, ZiP has shown that same night follow-up can be achieved on some of the shortest lived transients on any kind of telescope, not just for wide-field optics. Furthermore, this also demonstrates the versatility of survey patterns used to find transients. Some telescopes will cover multiple areas of the sky. Here, NGTS streams just one patch. ZiP can handle multiple survey patterns and find transients just as effectively.

5.4.2 Monitoring M31

Image subtraction is primarily conducted on relatively empty fields with moderately low varying sky background. One of the challenges when finding high energy transients, such as GRBs and different novae, is that they are usually located on top of a host galaxy. Being able to remove a host galaxy cleanly and without removing the source transient can be difficult. Andromeda offers a particularly unorthodox challenge in this regard. With a non-uniform background and crowded field, M31 delivers the same hurdles as a normal host galaxy with increased difficulty.

GOTO conducted a survey where it revisits M31 on a nightly basis. Subtractions are computed and changes in the field are extracted. One way of assessing if an extracted transient is real is through human vetting. This is a plausible method when extraction count is low or if there is an additional filter removing blatant

false positives (this will be addressed in chapter 7). A particularly useful output for potential candidate vetting is shown in figure 5.23.

Figure 5.23 is a real example of using human vetting to verify a potential nova. Upon finding the source, prior visits to the field can be analysed to find when the nova started. The transient can additionally be monitored further shown in figure 5.24. Increased confidence in this detection is offered through independent announcements (Hornoch and Kucakova, 2018).

By finding a faint transient in the centre of Andromeda, a busy field with a complex background, ZiP has shown that it will not struggle finding transients that occur on top of galaxies.

5.4.3 Finding HOPE, GOTO’s Favourite Transient

GOTO’s primary mission is to find counterparts to gravitational-wave events. One such trigger, S190901ap (25606), was GOTO’s first event with a transient spatially and temporally coincident with a GW alert. This transient, formally known as AT2019pju, was observed by GOTO and identified prior to the first announcement (Lipunov et al., 2019) (discovered in the MASTER survey). This transient was internally named GOTO2019hope (Ackley et al., 2019).

Given how vital the GOTO2019hope detection was to GOTO and the fact it is a clear demonstration of GOTO finding transients coincident with GWs, it is imperative that ZiP is able to find transients of this level of importance too.

Figure 5.25 shows the subtraction of GOTO2019hope. This demonstrates that ZiP is capable of finding transients in GOTO as faint as 19.6 ± 0.1 magnitude. After finding GOTO2019hope, GOTO continued observing this field for for some time and was able to build a lightcurve of the transient, which was later classified as a type 1a supernova. The lightcurve is shown in figure 5.26.

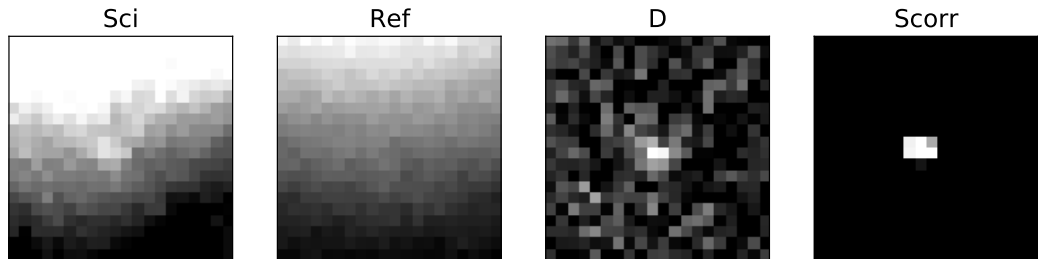


Figure 5.23: Example of vetting subtraction candidate

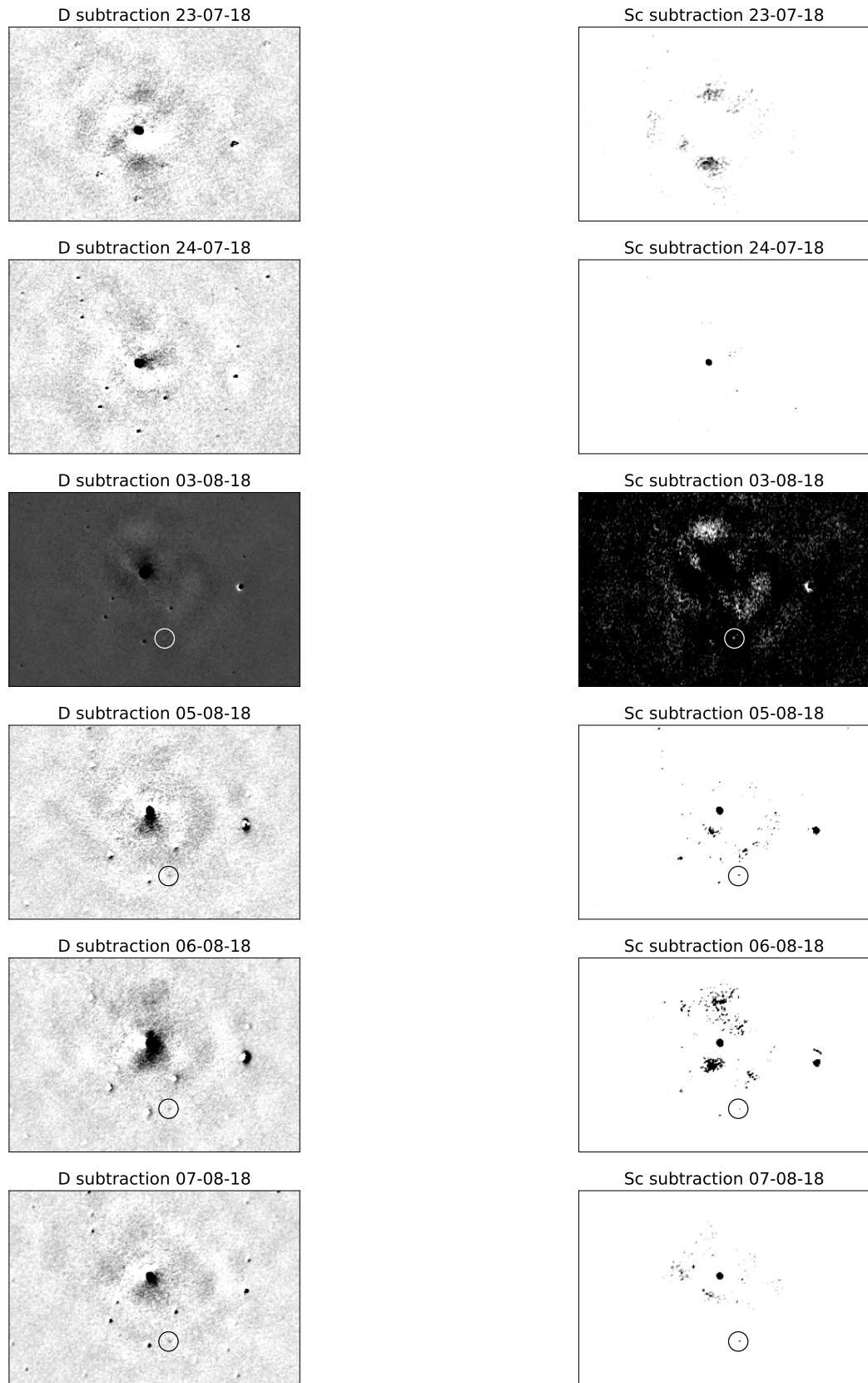


Figure 5.24: Using image subtraction to find and monitor a transient in M31. Showing ZiP can locate faint transients in bright and crowded environments.

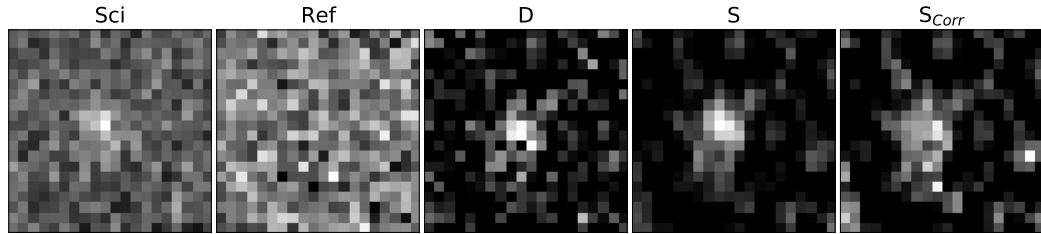


Figure 5.25: GOTO2019hope being subtracted with ZiP.

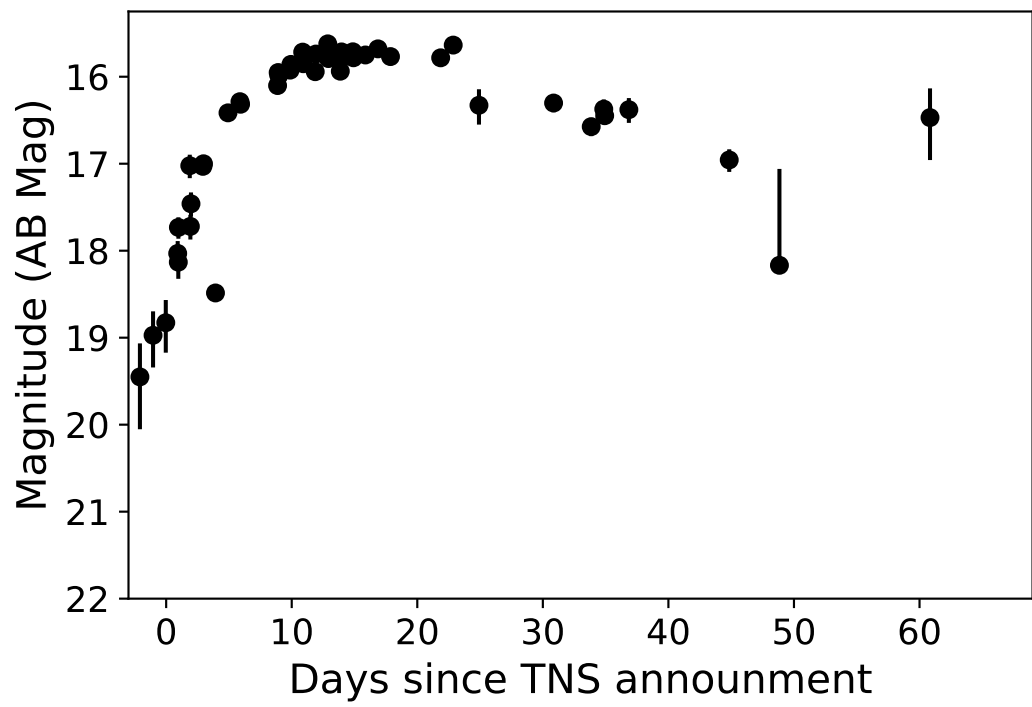


Figure 5.26: Lightcurve of GOTO2019hope (AKA: SN 2019pju) a type Ia supernova. Showing GOTO can find early time signals of Supernovae.

5.5 Chapter Conclusion

Image subtraction is a vital component both in hunting for and the serendipitous discovery of transients in survey data. Surveys are seeing deeper into the universe and their images are becoming much larger. Current algorithms in the literature are insufficient in either optimising the data or finishing the subtractions quickly enough for real time searches. In the above chapter, I have detailed a new algorithm I designed to solve these two issues.

In addition, different PSF modelling methods were trialled to find the most suitable. Two of which are entirely new methods; given their novelty, improvements can still be made to them. The selected PSF estimation technique was a modified version of PSFex. Future improvements for the two new PSF estimation methods include a more stable source selector. Currently, both methods use a positional average to determine what the PSF is at a given CCD coordinate. Instead a more selective method, like the one implemented in PSFex, should provide a more comprehensive PSF estimate. Additionally, the Zernike moments method has shown to be powerful when building PSF models, however it can take some time at higher orders. By parallelising the algorithm and using some optimisation techniques, Zernike moments have the potential to build non-Gaussian positional dependent PSF models on the order of ten seconds.

I have shown that this new image subtraction algorithm, ZiP, is not only the fastest on the open-source market, but competes with leading image subtraction quality as well. ZiP has the potential to compute wide-field subtractions in real time. This makes it optimal for future surveys like the LSST. More importantly, it can be used in the GOTO pipeline for real-time transient identification when hunting for the counterpart to Gravitational Waves, as demonstrated when finding GOTO2019hope. Examples of finding real transients were given, highlighting ZiP's power when looking for transients in wide field survey data. A version of ZiP now needs to be integrated into GOTO's next transient pipeline. This will be a pipeline optimised to GOTO's specifications and so ZiP will need to be modified as to streamline transient discovery and make the process as efficient as possible. The ZiP slicing pattern may also need to be revisited when GOTO is using next generation astrographs, as the new hardware will most likely have a different optimal slicing pattern.

In the final chapter of this thesis I will combine GOTO's first GW survey with ZiP and highlight all the high confidence GW coincident transients found using this algorithm. But first, an algorithm must be in place to filter out the false positives.

Chapter 6

Filtering out False Positives with Machine Learning

“It’s still magic, even if you know how it’s done ”

— Terry Pratchett, *A Hat Full of Sky*

6.1 Introduction

One issue surrounding difference imaging not addressed thus far in this thesis is the handling of false positives. As the contamination is on the order of 99% (irrespective of subtraction method, see table 5.2), vetting all the flagged candidates by eye is an impossible task. To address this, a machine learner has been developed for fast classification of transient candidates.

A long standing problem in wide-field optical transient surveys is the filtering of false-positives, or bogus detections (Romano et al., 2006; Bailey et al., 2008; Bloom et al., 2008, 2012). The main challenge being, filter out as many bogus detections as possible while preserving all of the real sources. This chapter builds on the work already done on classifying sources in difference images to build a novel classifier for potential transient candidates found in ZiP subtracted data.

Studies using GOTO data have made progress for machine learning classification already (Mong et al., 2020; Killestein et al., in prep). Here, I will discuss the development of a novel Deep Neural Network (DNN) for the rapid identification of transients in ZiP subtracted images.

Most of the machine learners in transient astronomy use exclusively data from the subtraction image (Bloom et al., 2012; Sánchez et al., 2018; Mong et al., 2020).

While the subtraction image alone contains a significant amount of information for a human, the reduced features of the same image can obscure the ‘*realness*’ of a candidate. Here, I develop an original real-bogus classifier using a bespoke Deep Neural Network utilising features from the reference, science, and ZiP subtracted images.

I will start by discussing the standard machine learning tools used in astronomy. I then apply my novel feature selection to these methods, showing the parameters are powerful when classifying potential transients. After this, I will describe neural networks in preparation for the development of a novel deep neural net specific for the feature selection described above.

6.2 Building a Real-Bogus Classifier

6.2.1 General Machine Learners used in Astronomy

Most, if not all, binary classifiers function by giving weights to a selection of features based on some training data. The same features (without a known classification) can then be fed through the classifier which will output a number between 0 and 1 (where 0 is a real transient and 1 is a false positive). While the output of all these methods is in the same format, how they operate under the hood is quite different. The following section is going to explain the standard machine learning tools.

K-neighbours

Also known as K-Nearest Neighbours (KNN), this method requires no training period. Instead, the input parameters are compared to its closest matching neighbours. That is, the distance of the inputs in parameter space are calculated and the closest K sources determine the class. This distance is usually calculated using the Euclidean:

$$D_q = \sqrt{\sum_{i=0}^{N_{params}} W^i \left(P_{input}^i - P_q^i \right)^2} \quad (6.1)$$

P^i here represents the i^{th} parameter, q is the q^{th} neighbour in the sample, and W^i is a weight term for the given parameter. How these distances are analysed will determine how the classification is made. For this scenario, a list with all the distances is made with the corresponding class:

$$LIST = [(D_0, y_0), (D_1, y_1), (D_2, y_2), \dots (D_N, y_N)] \quad (6.2)$$

Where N is the number of samples from the training data. First, K smallest distances are taken from the list. Next, a class weight is given by summing the terms of each class. Smaller distances are given a higher weighting by using one over the distance. As this is a binary classification problem, the algorithm can be sped up. Instead of doing multiple summations for different classes, do the sum once and multiply 1 or -1 depending on the class:

$$y_{pred} = \frac{1}{K} \sum_{i=0}^K (2y_i - 1) \frac{1}{D_i} \quad (6.3)$$

This means if y_{pred} is less than zero, it is of class zero. Otherwise it is class 1. Nearest neighbour classifiers are often considered one of the more powerful supervised learning methods (Cunningham and Delany, 2020). However, when training sets become too large, the method slows down drastically. This was certainly the case with the training set used here. In this regard, I propose a parallel ensemble bagging method to make multiple predictions that are then averaged:

$$y_{final} = \frac{1}{N_{preds}} \sum_{pred=0}^{N_{preds}} y_{pred} \quad (6.4)$$

Where the predictions (y_{pred}), are made on a batch size smaller than the training sample in parallel. This is significantly faster, and potentially more accurate than the standard KNN method (AlBaghdadi and Alkoot, 2005).

Naive Bayes

The Naive Bayes (NB) method is a probabilistic classifier. By measuring the probability that a given set of inputs belong in a particular class (Rish et al., 2001). Bayes probability takes the form:

$$P(y_i | p_0, p_1, p_2, \dots, p_n) = \frac{P(\mathbf{p} | y_i) P(y_i)}{P(\mathbf{p})} \quad (6.5)$$

As the evidence, $P(\mathbf{p})$, can be considered a constant normalising factor, it can be ignored for the purposes of classification (i.e. $P(\mathbf{p}) = 1$). The above equation can be expanded to look like:

$$P(y_i | \mathbf{p}) = P(p_0 | p_1 \dots p_n, y_i) P(p_1 | p_2 \dots p_n, y_i) \dots P(p_{n-1} | p_n, y_i) P(p_n | y_i) P(y_i) \quad (6.6)$$

The naive assumption here is that all the parameters (p_i) are independent on determining class (y_i), meaning equation 6.6 can be expressed:

$$P(y_i|\mathbf{p}) = P(p_0|y_i) P(p_1|y_i) \dots P(p_n|y_i) P(y_i) \quad (6.7)$$

This expression can be made neater still:

$$P(y_i|\mathbf{p}) = P(y_i) \prod_{j=0}^n P(p_j|y_i) \quad (6.8)$$

The largest $P(y_i|\mathbf{p})$ is taken for a given input and y_i becomes the predicted class.

Gaussian Naive Bayes

The problem with using the NB method is that it still requires one to understand the probability of getting a parameter p_k with a class y_i ($P(p_k|y_i)$). Providing the input parameters are non-discrete, an assumption can be made that the distribution of said parameters is normal. This refers to Gaussian Naive Bayes (GNB). That is:

$$P(p_k|y_i) = \frac{1}{\sqrt{2\pi\sigma_{pk}^2}} \exp\left(-\frac{p_k - \bar{p}_k^{y_i}}{2\sigma_{pk}^2}\right) \quad (6.9)$$

Where $\bar{p}_k^{y_i}$ represents the mean of the parameter p_k distributed within the class label y_i (John and Langley, 2013). Training in this situation will refer to finding the mean and standard deviation, $\bar{p}_k^{y_i}$ and σ_{pk} respectively. This method is incredibly fast; however, certain assumptions may be detrimental to classifier in this scenario. Firstly, the naive assumption of independence is erroneous. The subtraction image is correlated to both the science frame and reference frame. Secondly, the Gaussian assumption is flawed also. Some of the input parameters do not have a normal distribution, but rather a flat one. It has been shown that these assumptions do not have to be entirely accurate to provide meaningful classification (Hand and Yu, 2001); though with more accurate methods available for this problem, GNB will most likely falter. The GNB method has already been packaged in the `sciKit` module and is the tool used here (Pedregosa et al., 2011).

Random Forrest

To understand a random forest, one must first detail how a decision tree works. A decision tree is a cascade of conditional statements to draw a conclusion, either in a classification problem, or when regressing to predict a value (Wu et al., 2008).

A decision tree is said to ‘grow’ as more conditional branches are added. It has been routinely stated that single trees with arbitrary complexity are inconsistent and lacking in accuracy (Breiman, 2001; Hastie et al., 2009). Instead of having an individual overgrown tree (i.e. an overfitting classifier), one can use an ensemble of randomly grown trees (Ho, 1995; Breiman, 2001). This is a *wisdom of crowds* logic (Golub and Jackson, 2010). That is, every tree in the forest will draw a conclusion on what classes the input data are and a consensus is drawn on the what the majority of trees decide. Similarities of this method to the KNN method are drawn in Lin and Jeon (2006); the random forest method is often considered (and even shown to be) the stronger of the two however (Scornet, 2016). This is because the trees can be considered independent, the low correlation between each tree means each *vote* contains new information. The errors in the classification are minimised, leading the ensemble of trees to the most likely answer even if a minority draw the wrong conclusion. The most common way of training a random forest is by “bagging”, also known as bootstrapping. This is where the training data contains N samples and N samples are given to the classifier, but some of the inputs are omitted and replaced with copies of other samples in the data. This is done randomly for each tree. An additional feature implemented as standard now is subspace bagging (Ho, 2002; Bryll et al., 2003). This is where a random selection of parameters are used in each tree to help prevent correlation in decision making.

A particularly useful characteristic in both decision trees and random forests is that a metric of feature importance can be extracted. How one calculates feature importance is still in hot debate (Strobl and Zeileis, 2008; Parr et al., 2018); but the one used here is the default in `SciKit`. It is called mean decrease in impurity or *Gini-importance* and is based off the Gini-Impurity-Index (GII) (Breiman, 1996). Simply put, the GII is a measure of how effective a decision split is at separating two or more classes:

$$G = 1 - \sum_{c=0}^{n_c} p_c^2 \quad (6.10)$$

Where p_c is the ratio of inputs with class n_c over the total number of inputs. The importance of the individual branch is then calculated:

$$I = G_{parent} - G_{child1} - G_{child2} \quad (6.11)$$

Finally, the importance is found by averaging over all the splits in the forest that

use parameter (q).

$$I_q = \frac{1}{N_{splits}} \sum_{i=0}^{N_{splits}} I_i \quad (6.12)$$

Both the random forest and the feature importance methods discussed here are built in to `sciKit` (Pedregosa et al., 2011). These tools are used later in the chapter.

6.2.2 Artificial Neural Networks

All of the methods above have deterministic properties; in other words, one can trace the methodology to understand and predict the output. In recent years, neural nets have become the leading machine classifiers. However, there is some contention of their use within the scientific community as neural nets are stochastic or 'a black box' and cannot be back traced. Here, I explain how neural nets work and, hopefully, articulate how powerful these methods are for rapid classification.

This section is split into multiple subsections clarifying important components and terms surrounding neural nets.

Artificial Neural Networks (ANNs) got their name from their likeness to biological neural-nets; brains (McCulloch and Pitts, 1943). Figure 6.1 is a demonstration of an ANN, highlighting the main components of a binary ANN classifier (Beale et al., 1996). The neurons (symbolised by white circles) take an input then using an internal equation, called an activation function, the neurons will produce an output. These outputs are weighted and sent to the Output layer where a score is given between 0 and 1, where each integer represents a class (0=real, 1=bogus).

The input layer neurons represent the number of input parameters. Hyper-parameters are parameters that are set before learning begins. This includes number of neurons, layers, batch size, etc. One of the most important hyper-parameters, and widely considered to be the reason neural-nets have become so powerful, is dropout rate. Dropout is a form of regularisation, which is a way to stop a classifier overfitting and only training specific neurons (called Co-adaptation Baldi and Sadowski (2013)). First introduced in Hinton et al. (2012b), dropout is the process of ignoring a number of neurons during training to more accurately work out the weights of each input neuron. It has been shown the maximum regularisation occurs with a dropout rate of 0.5, so that is the value used in the rest of this work. ANNs take advantage of stochastic gradient descent to be able to make accurate predictions.

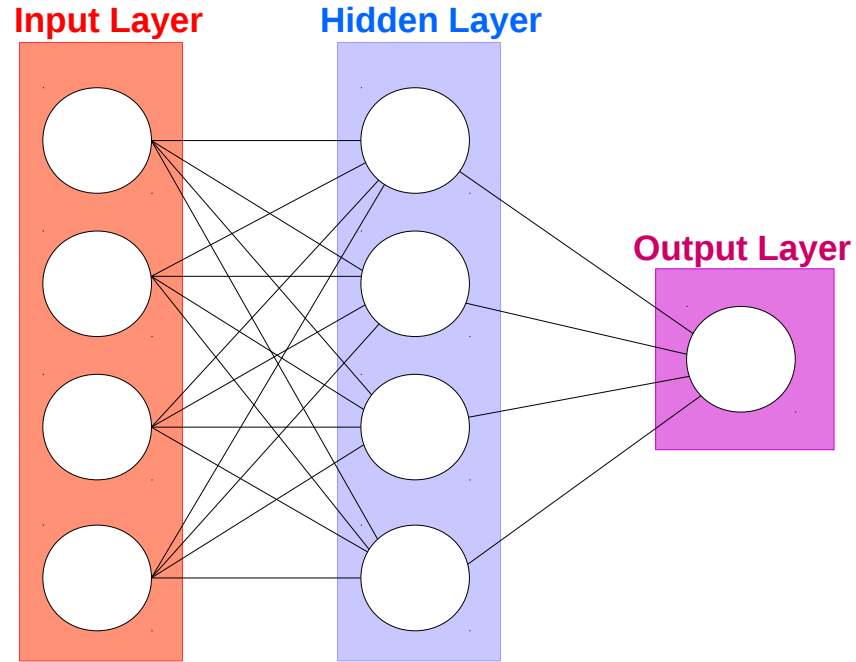


Figure 6.1: A basic ANN schematic for a binary classifier. Input layer (red) represents the input parameters, the hidden layer (blue) applies an unknown functions and sends it to the output layer (purple). The output layer is a score between 0 and 1 representing each class in the binary.

Gradient descent algorithms

A gradient descent algorithm is an iterative process to find the minima of the loss function (also know as a cost function (Ruder, 2016)). The loss function is an estimate of the errors on a set of predictions. Gradient descent works by taking the output of the loss function (J) and making a weighted step down the gradient to find the minimum with given input parameters (θ):

$$\theta^{i+1} = \theta^i - \eta \Delta J(\theta^i) \quad (6.13)$$

This method is referred to as batch gradient descent (BGD) as it involves minimising the loss function at each step for the entire training data set. As the training data in our situation are so large, BGD is not a viable way to train the data.

Stochastic gradient descent

Stochastic Gradient Descent (SGD) takes BGD to the other extreme. Instead of operating on the whole training data, it will use a single randomly selected sample from that data (x) with a label (y). This often results in a noisier and longer path to the minimum (i.e. takes more iterations); however, the time taken to get to the minimum is much shorter as training is not done on the whole sample.

For this work, a combination of the two methods, SGD and BGD, is used. This is called mini-batch gradient descent (MBGD). This method works by taking n random samples (also known as the batch size) of the training data to minimise the cost function:

$$\theta^{i+1} = \theta^i - \eta \Delta J(\theta^i, x^n, y^n) \quad (6.14)$$

The Cost Function

The classifier's task here is to determine if the source detected in the difference image is real or not. This means the classes can be characterised as a binary. Because the classifier is working in a binary regime, a binary specific loss function can be used. Binary Cross Entropy uses the following formula (Hagenauer et al., 1996):

$$H_p = \frac{1}{n} \sum_{i=1}^n y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - (p(y_i))) \quad (6.15)$$

In short, binary cross entropy is the mean entropy a batch of predictions are from their true value. The smaller the cost function, the closer the predictions are to the expected class.

Momentum

Momentum is a natural step forward to increase learning rate. Given that initial classification predictions are essentially blind, the gradient of the learning curve can be considered parallel to the loss function (see fig 6.5 as an example). Taking uniformly small steps down the steep gradient (ΔJ) is effectively the same as taking one large step, as the direction remains the same (Qian, 1999). The most prominent, though not necessarily the first, implementation of this line of thinking is found in Rumelhart et al. (1986). Here the difference in the entropy's gradient is used to model a weight dictating the step size. For steep gradients, the step size is large. For shallow gradients, the step size is smaller, this averages out small oscillations when approaching the minimum in the learning curve. One of the most recent

developments in learning rate optimisation is Adaptive Moment Estimation, Adam. Note, Adam is not an acronym which is why it is not written ADAM (Kingma and Ba, 2014). Adam makes the primary assumption that the learning curve is convex, this has shown to cause some issues and multiple attempts to remedy this exist (Dozat, 2016; Zhang et al., 2017; Loshchilov and Hutter, 2017); however, given the simplicity of the input data for this task, it can be assumed the learning curve is convex. Adam is considered an extension to the SGD method. Unlike base SGD, Adam uses an adaptive learning rate for each parameter, these subset of procedures are commonly referred to as Adaptive Gradient Descent algorithms, AdaGrads (Duchi et al., 2011). In particular, Adam is considered an improvement of RMSProp (Hinton et al., 2012a). Adam works by taking the first and second moments of the gradient (the mean (m_1) and un-centred variance (m_2)) and using them with a debiasing factor to predict the parameter's next step:

$$\theta^{i+1} = \theta^i - \frac{\eta}{\epsilon + \sqrt{m_2^i}} m_1^i \quad (6.16)$$

Ruder (2016) presents an excellent summary of the above, relating abstract computational concepts to physical scenarios.

Activation Functions

Activation functions are deceptively simple, but appear as the culmination of all of the above. This is a threshold to decide if a neuron should fire. In other words, the weighted sum of the inputs should be greater than a given activation threshold to decide if the neuron should pass information to the next layer. The most commonly used functions for classification problems are ReLu and Sigmoid, but studies show that there is sparse difference in classifier performance based on activation function (Karlik and Olgac, 2011; Sibi et al., 2013). For this classifier a Sigmoid activation function was selected as it showed to converge to the learning curve minimum faster.

6.2.3 A Novel DNN

DNN stands for Deep Neural Network. DNNs are ANNs with multiple hidden layers. The architecture used here takes the parameters and feeds them to 512 neurons. Each hidden layer has half the number of neurons as its predecessor. This is shown in figure 6.2.

Multiple models are made (i.e. the classifier is trained numerous times), then the mean prediction from the different models is taken. This results in a

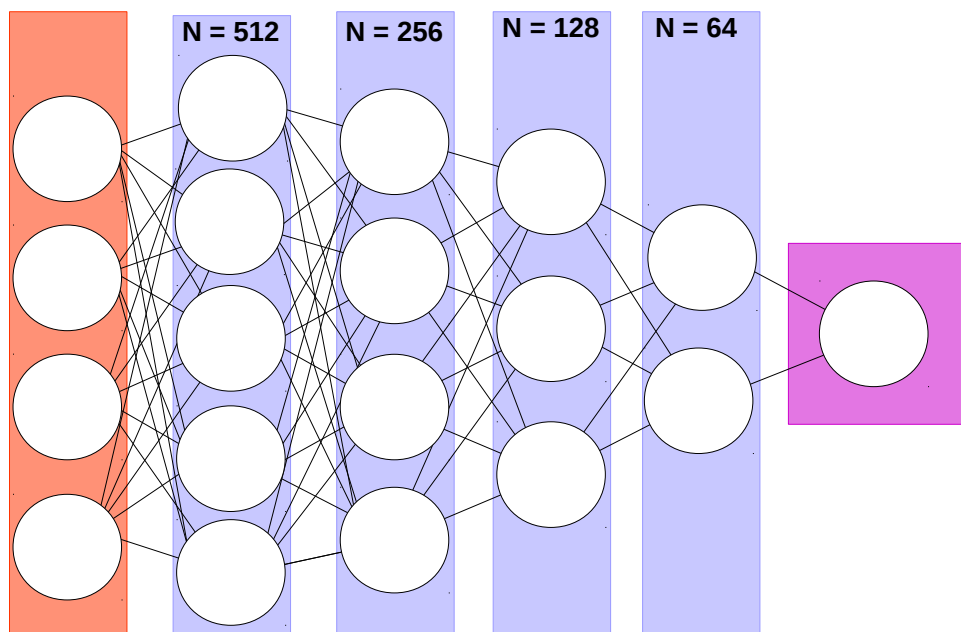


Figure 6.2: DNN architecture for an automatic rapid transient hunter. Using the same colour representation from the ANN, this figure shows how the input layer is sent through multiple hidden layers. Each hidden layer has half the number of nodes before being sent to the output layer.

more accurate prediction, similar to ensemble methods (Tumer and Ghosh, 1996; Opitz and Maclin, 1999; Rokach, 2010). Some methods using machine learning on difference images already exist in the literature (Morii et al., 2016; Wright et al., 2015; Sánchez et al., 2018), but most train using HOTPANTS. In particular, Sánchez et al. (2018) concludes that the best combination for a classifier is a random forest on HOTPANTS data, despite finding that the ZOGY algorithm can see to at least the same depth as the other methods. This is a worrying conclusion, and reiterates the problem framed in section 6.1, the classifiers performance is entirely dependent on the feature selection of the difference images. To remedy this, features from the science and reference frames are also taken.

Source selection is initiated by extracting high confidence sources in the S_{Corr} image. The flux-statistic is recorded from the S_{Corr} image, as is the difference flux in the D image. Forced photometry is done in the same position in the reference and science images. The Half-flux radius, flux, background count, and source position are also given to the classifier. I will show the importance of the feature selection in the following section.

6.3 How to Train your Computer

The same as when I did the subtraction tests in chapter 5, I need to inject fake sources into the image as the number of real transients is too low to reliably train a classifier. Fake transients are made by injecting sources into the training data using LazyPSF. An important feature of this method is that labelling can be automatic. As the number of real transients is much smaller than the amount of bogus detections, it is a safe assumption that all the non-injected sources can be labelled bogus. This will result in some incorrectly labelled true transients; however, this will have a negligible impact on the bogus training sample. Conversely, anything injected can be considered ‘real’. A similar method is used in Mong et al. (2020) and shows to be a valuable way to train a classifier. The data is subtracted and the sources are extracted using `sep` (Barbary, 2016). From here, the source locations are taken and feature analysis is done on the reference image and science image.

Initial classification tests are done with the KNN, GNB, and Random Forrest methods. These methods show, especially the random forest, that the parameter selection is informative when deciding whether a source is real or not. This can be seen in the confusion matrices in figure 6.3. As expected, the GNB classifier has the weakest performance, and the Random forest has the best of the three. Using the GINI importance metric, I have found the feature importance plotted in figure 6.4.

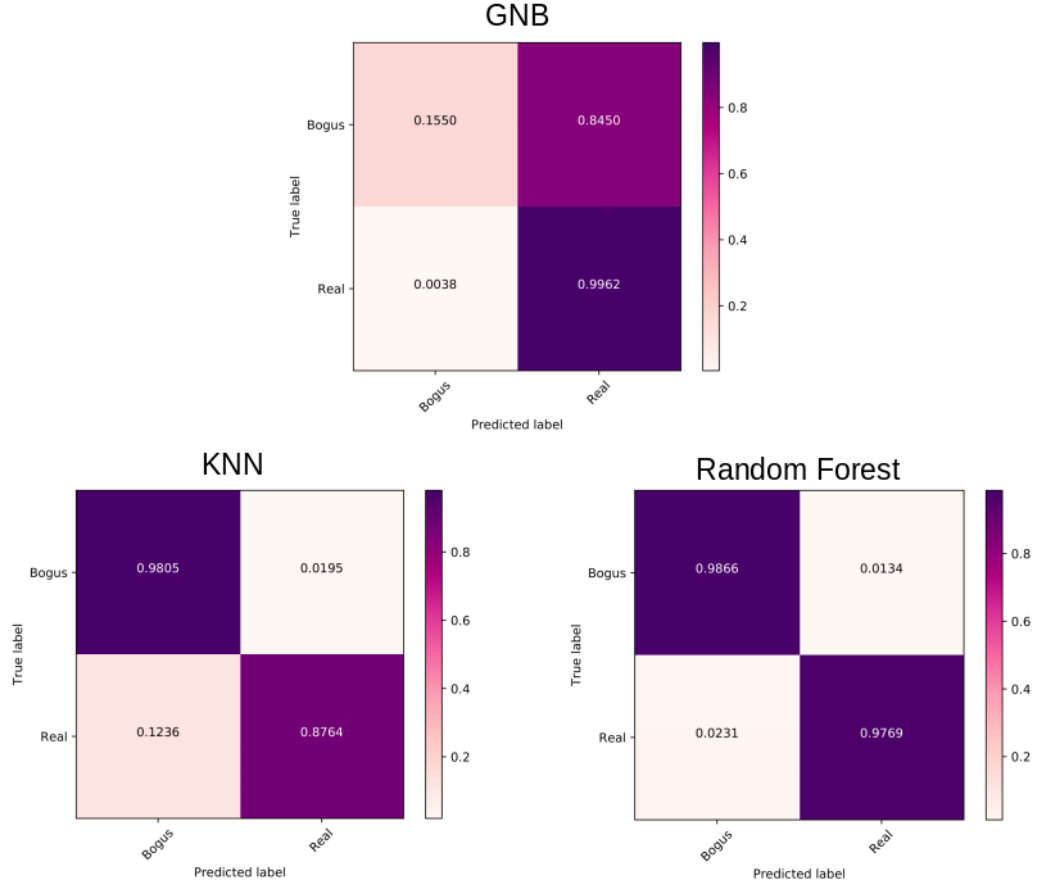


Figure 6.3: Confusion matrix of the GNB, KNN, and Random forest methods. Not including the GNB method, as it appears to be calling a majority of sources real, the KNN and random forest method support the strength of the new feature selection.

Once the feature selection was shown to be suitable on the basic classifiers, I then applied the same parameter space to the DNN classifier.

Training is done by minimising the loss function over a set of epochs (see figure 6.5). The best model is then selected as the epoch with the highest validation accuracy, figure 6.6.

Confusion matrices are a good way to highlight the predictive power of the ensemble DNN method. The top right panel of figure 6.7 represents the number of false positives. The bottom left is missed transients. This shows, that of the training sample, 2% of the injected transients are lost.

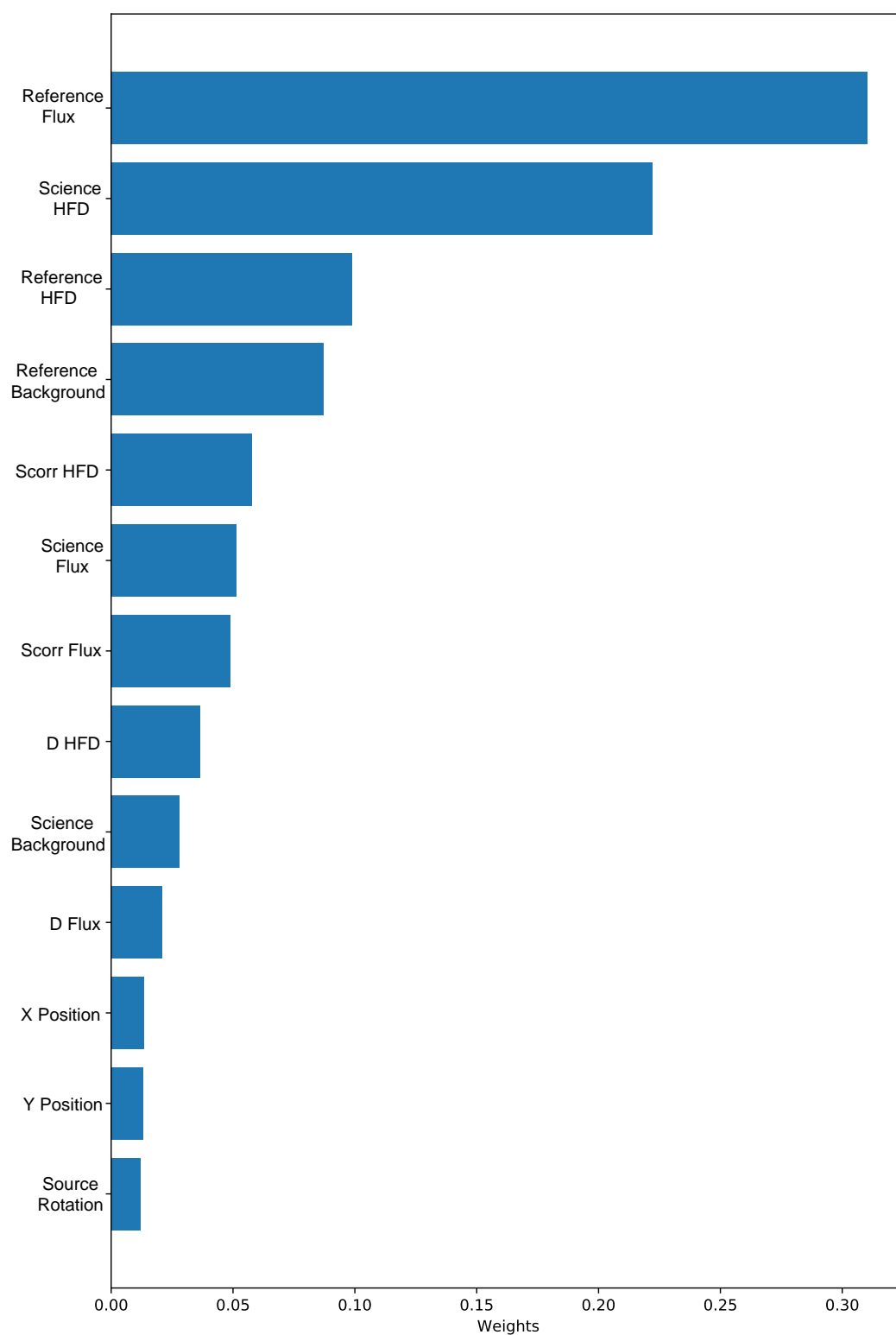


Figure 6.4: Weights of each parameter when determining classification for the Random Forrest method.

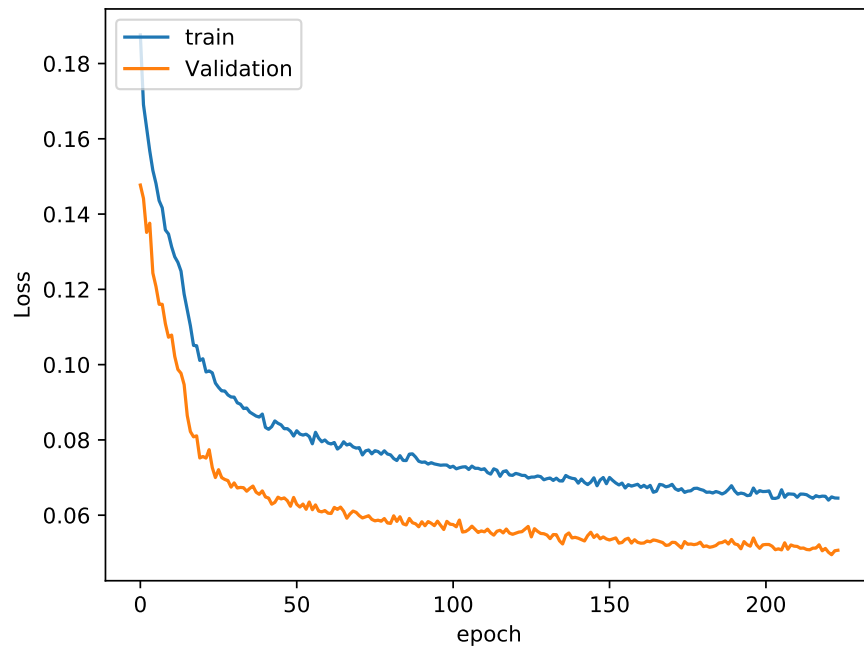


Figure 6.5: Loss over training epochs. Training epoch refers to the number of times the trainer has iterated over the training data. In the first few epochs the curve can be treated as parallel to the loss axis.

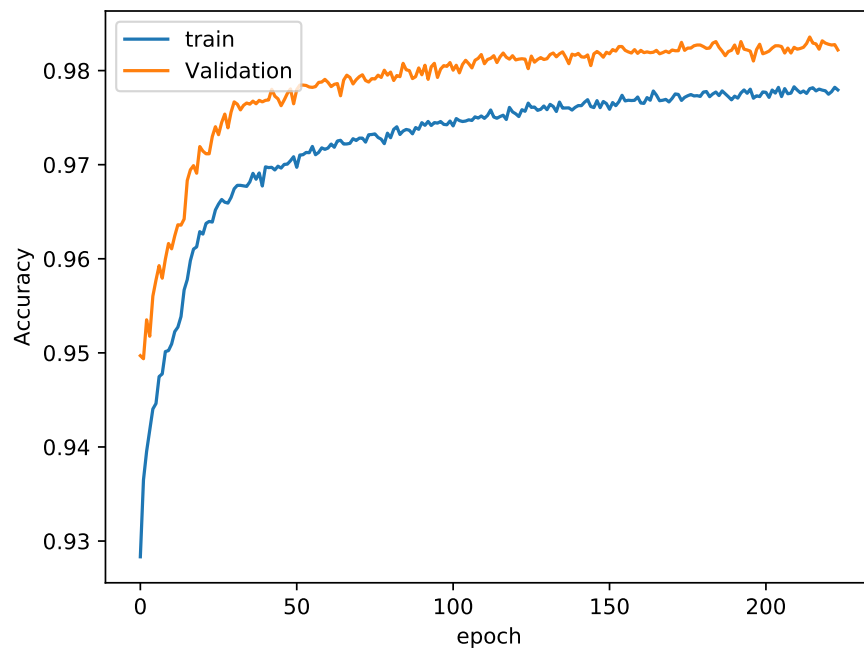


Figure 6.6: Accuracy over training epochs. The peak of this graph selects the machine learner for this training iteration.

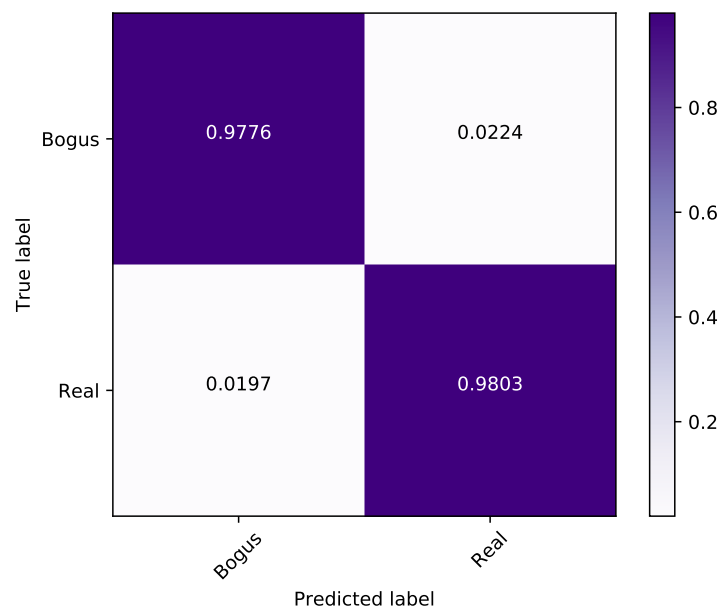


Figure 6.7: DNN Training confusion matrix, highlighting that 98% of transients are appropriately flagged and only 2% of the false positives are incorrectly flagged as real.

6.3.1 Testing the DNN

Finding injected transients on Galaxies

As transients usually occur on top of galaxies, it is important to test transients are actually being detected on top of galaxies. This was done by selecting a number of fields taken by from UT4 with galaxies present and injecting sources on to the galaxy position, with a systematic offset from the centre of 10-20 pixels from the galactic centre. I then subtract the data using ZiP and finally classify the extractions using the trained models. Figure 6.8 highlights the results of this process.

The blue bars represent the recovery of transients from the extraction step, this is analogous to the recovery stats from chapter 5. The orange bar is the number of correctly identified transients, the purple bar represent number of correctly identified candidates with classification confidence greater than 95%.

Figures 6.9 and 6.10 show the confusion matrix at the different thresholds. The reason for a higher threshold is to more strictly exclude false positives, but this is at the expense of more false negatives. It appears the classifier does not struggle finding injected transients found atop galaxies. The random forest method was also used on the galaxy training data. Another feature importance plot as made, shown in figure 6.12. The plot highlights the same conclusions from figure 6.4, that the input images hold more classification potential than the subtraction images.

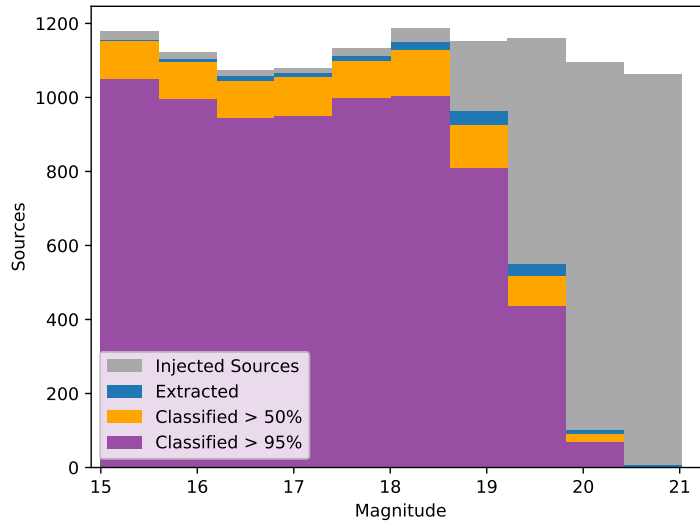


Figure 6.8: Proportion of injected sources (on top of galaxies) correctly flagged at the extraction and classification stages.

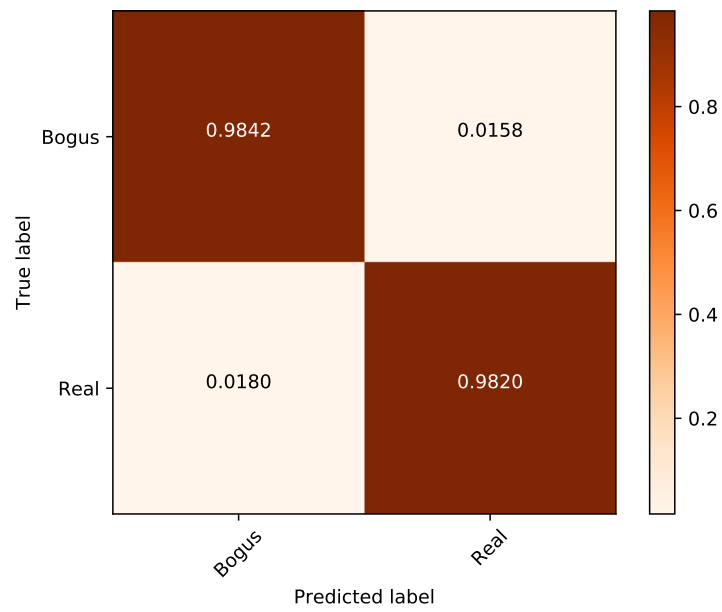


Figure 6.9: Confusion matrix for 50 percent confidence

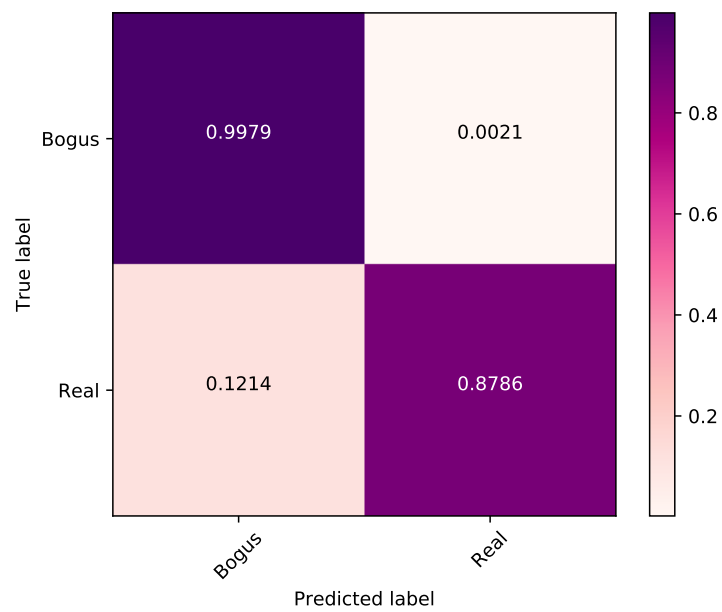


Figure 6.10: Confusion matrix for 95 percent confidence

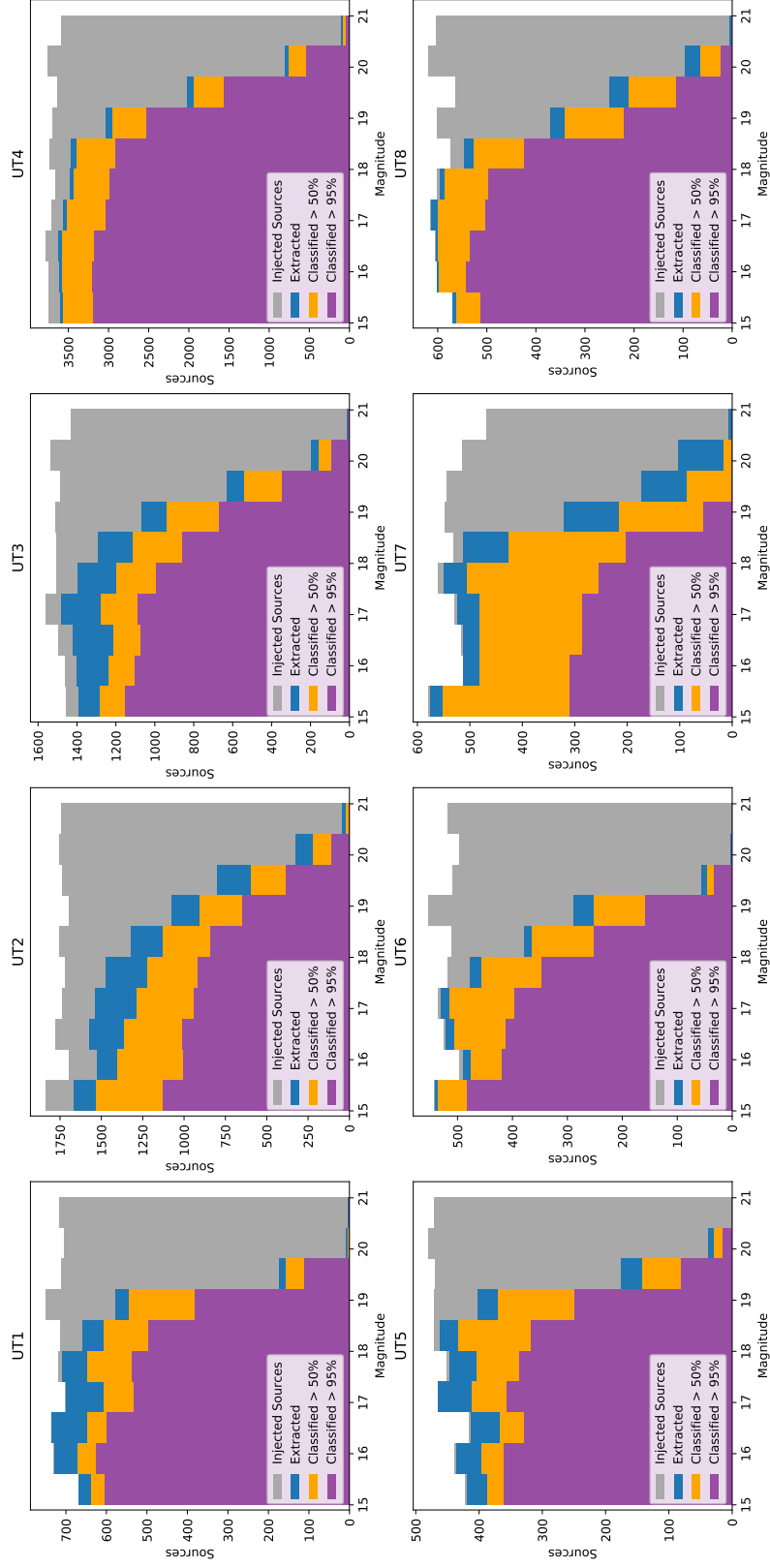


Figure 6.11: Recovery and classification of injected sources for each UT. The grey bar indicates the number of injected sources. The blue bar is the number of sources extracted from the subtraction step. Orange represents transients classified with $>50\%$ confidence. Purple represents transients classified with a confidence of over $>95\%$.

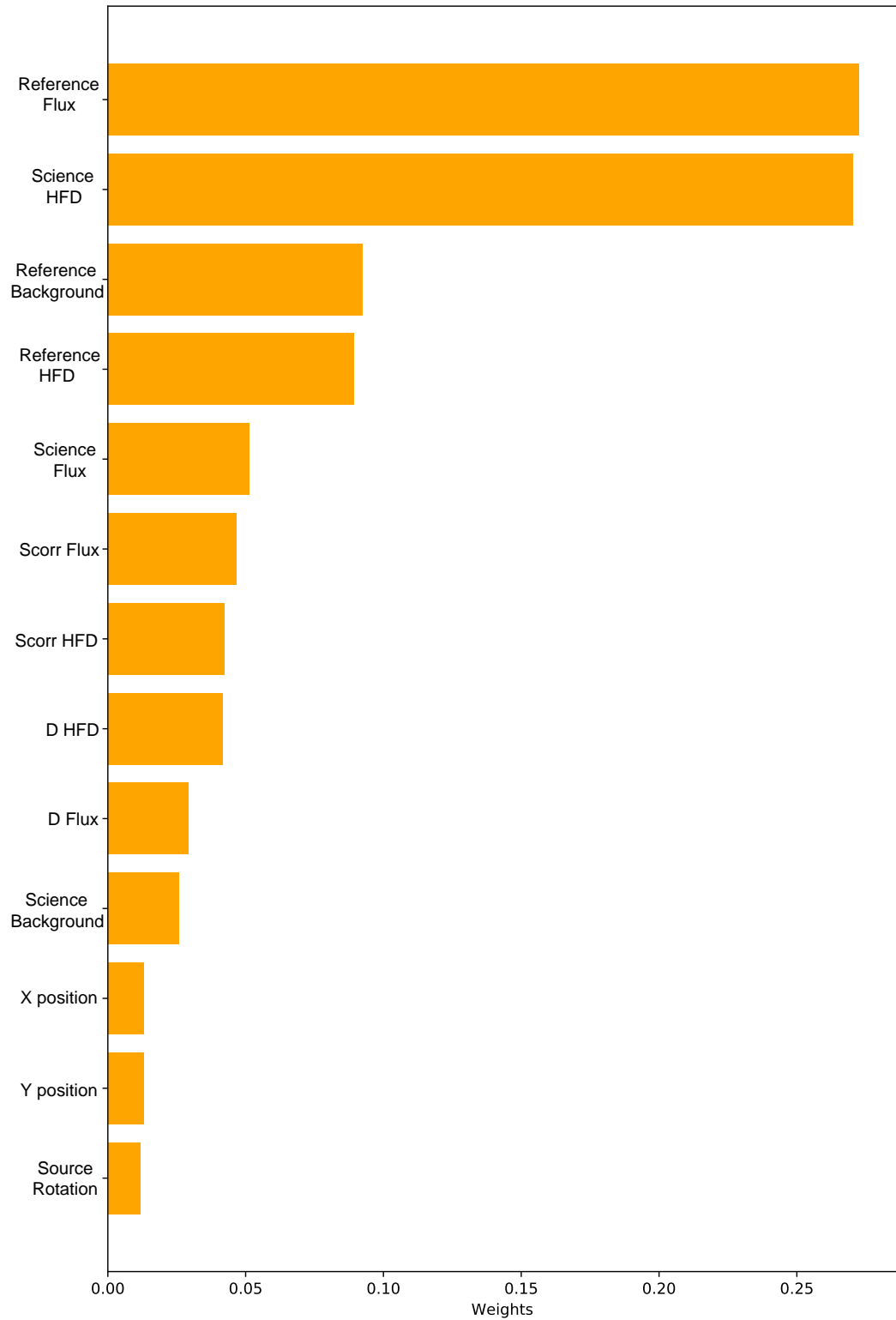


Figure 6.12: Weights of each parameter when determining classification for the Random Forrest method. This figure uses the training sample from the galaxy only injected training sample.

Finding injected transients

The data were trained on the GOTO astrographs. However, the performance of each UT varies; therefore, to test recovery of the machine learner each validation set has been split into component UTs. The colours here correspond to the same metrics as in figure 6.8. Figure 6.11 indicates the recovery metrics for each UT. The transients were injected randomly on a random selection of fields. Therefore, figure 6.11 highlights classifier strength for each UT. Figure 6.11 shows that the training data favours the original astrographs, UTs1-4 (see figure 2.4). This is most likely due to the surplus of data and performance optimisation for the original UTs. At the time of making the training data, the new UTs had only been on GOTO for a few months. This meant they had not been properly collimated and did not have the quality reference templates that the original UTs had. Figure 6.11 also shows that there is some benefit to training the UTs individually instead of having a global classifier.

Comparison

Receiver Operating Characteristics (ROC) curves are a diagnostic tool, predominantly for binary classifiers (Spackman, 1989; Fawcett, 2004). The ROC curve is built by plotting the true positive rate against the false positive rate:

$$TPR = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (6.17)$$

$$FPR = \frac{False\ Positive}{True\ Negative + False\ Positive} \quad (6.18)$$

Once the curve has been graphed, the Area Under the Curve (AUC) is ascertained. If the AUC is equal to 0.5, the classifier can be considered 'guessing'. If the AUC is > 0.5 the classifier is more often predicting correctly, where an AUC of 1 highlights a perfect classifier. Interestingly, if the AUC < 0.5 the classifier is more likely to be wrong, where an AUC of 0 is an anti-classifier, it is 100% incorrect every-time. This means, almost paradoxically, the closer to 0.5 the AUC, the worse the classifier. Figure 6.13 shows the ROC curves for every method for the individual UTs.

Method	Training Time (s)	Classification Time (s) per 100 000 sources	Mean AUC
KNN	N/A	35.8	0.755
GNB	0.31	0.013	0.616
Random Forrest	210	5.2	0.893
DNN	14400	8.5	0.904

Table 6.1: Machine Learning Performance for different methods. This reiterates earlier analysis, showing the GNB and KNN methods are weaker. It also highlights the DNN and random forest have similar performances.

Finally, training time and classification time are measured. Training time is less vital, nevertheless, it presents insight into practical classifiers. Table 6.1 shows the results of the different classification methods.

By combining the results from figures 6.10, 6.11, and 6.13 and table 6.1, I have shown that it is possible to find transients and filter them in real time using a combination ZiP and a novel machine learner. Furthermore, because my DNN method has a score attributed with it, not only can I highlight high confidence transients automatically; I can also prioritise which transients should be followed-up. This all has implications for future GW searches. Firstly, GOTO can now highlight new transients coincident with GW triggers in real-time. This can mean a new search strategy that has live updates to constrain or even classify the transient. Secondly, and more importantly, if multiple transients are found during a single follow-up campaign, the score of the transient and its probability in the localisation region can be used to inform which candidate is most valuable to follow-up. Finally, if GOTO ever receives a follow-up facility (such as upgrading the Warwick 1 metre telescope to get rapid low resolution spectra) transients found in this pipeline can be fed to such an instrument for automatic classification.

In summary, the workflow from subtraction to classification can be computed within the time it takes for one GOTO exposure. This stops images getting backlogged for processing and allows for real-time same night identification and classification of transients. The low number of false positives that are let through prevent the system from getting cluttered making this pipeline a strong aid in finding transients associated with GWs quickly.

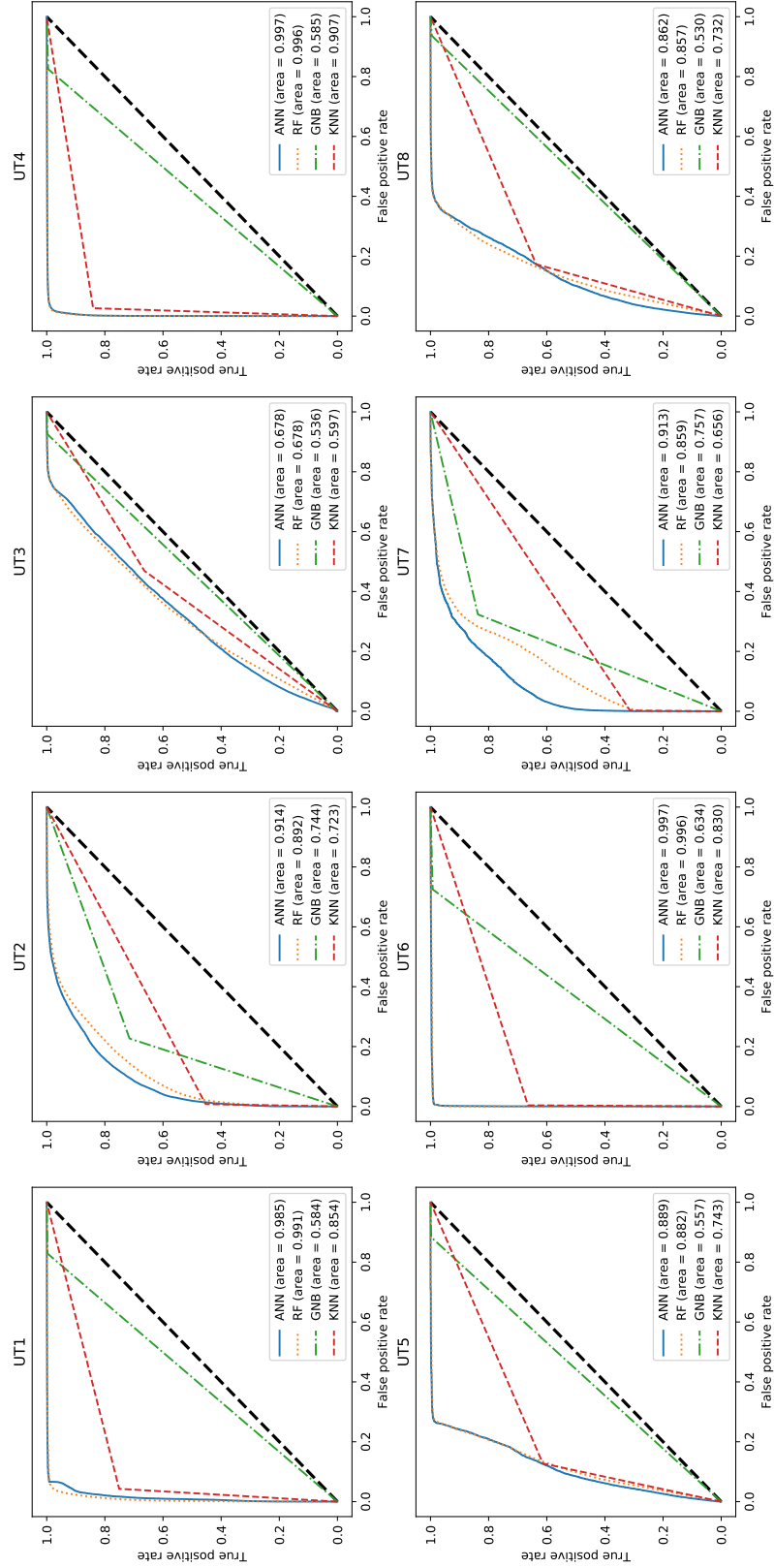


Figure 6.13: ROC Curve for different machine learning methods on each UT. Each plot is based on the classification metrics for each UT. The black dashed line represents the ‘worst’ classifier, the orange line is the DNN classifier, the orange line is the random forest, red is the KNN, and green GNB. There is a strong performance correlation of all the methods with UT, most likely linked with image quality. There is also a performance difference in classifier based on the method used, as seen in table 6.1

6.4 Machine Learning Summary

The success of the subtraction step, flagging potential transient locations, in combination with the fast and effective removal of false positive by the DNN described in this section creates a strong transient-hunting pipeline for real time follow-up of gravitational-wave follow-up (and other localisation missions).

For the given parameter space, the DNN has shown to be the most reliable classifier. However, the Random Forest also has a relatively strong performance. This is indicative of a good parameter selection. This is a novel input as it makes use of the reference, science, raw subtraction, and Scorr images. Neural nets are disadvantaged in that they take much longer to train. However, for GOTO’s needs the classifier only needs to be trained once and does not need to be trained on each image as they come in. Meaning, there is no cost to a long training time and only classification time has impact on GOTO’s strategy. This makes the DNN an appropriate fast real-time classifier, classifying entire images on ~ 2 second timescales. Given more time, it would have been beneficial to train a classifier for each UT. It appears that the different UT characteristics make it difficult to build a universal classifier. This is represented in all the machine learning methods, where certain UTs have much lower classifier performance. One future improvement would be to design a classifier for each UT instead of having a global classifier for the entire array.

Convolutional Neural Nets (CNNs) and other computer vision techniques may offer more comprehensive classifications as they *look* at the images directly. Nonetheless, CNNs will take orders of magnitude longer training time (order of 10 hours) and have limited training space. The DNN was able to use a training sample of $>$ a million sources, whereas CNNs max out at $\sim 400,000$. Training time is not vital compared to classification time, CNNs have shown to be able to classify entire images on the order of seconds (Killestein et al., in prep). Given the nature of image classification, CNNs offer the most predictive power (Duev et al., 2019; Jia et al., 2020).

As a final note, training will be more reliable on real data. By using actual astrophysical sources from ZOGY subtracted data vetted by humans, the classifier will not be biased to PSF-matched injected transients. This allows for one final improvement. By using real transients the modelled PSF of the image could also be used as an independent measure of source realness. Real sources will match the PSF of their respective images and artefacts will not. This would be an optimal study for a CNN classifier.

Now that the classifier has been trained, it can be used on GOTO's GW follow-up survey to find transients coincident with GW triggers.

6.5 A Future Classifier

A particular point of interest arose from figures 6.4 and 6.12. Over 50% of the classification is based exclusively on the input images. This means a significant amount of time can be saved simply by classifying sources on their science and reference frames. Disregarding the subtraction phase entirely. A preliminary study has been done to look into this method and has shown to hold a lot of promise, the only drawback is the extraction of faint transients. Developing an '*image subtraction-less*' pipeline would require a comprehensive way of flagging faint sources on top of galaxies without using a subtraction step. There are multiple solutions and workarounds to this problem and future investigation should yield promising results for real-time transient identification.

Chapter 7

The First Follow-Up Campaign with GOTO

“I like the stars. It’s the illusion of permanence, I think. I mean, they’re always flaring up and caving in and going out. But from here, I can pretend...I can pretend that things last. I can pretend that lives last longer than moments. Gods come, and gods go. Mortals flicker and flash and fade. Worlds don’t last; and stars and galaxies are transient, fleeting things that twinkle like fireflies and vanish into cold and dust. But I can pretend...”

— Neil Gaiman, *The Sandman: Brief Lives*

7.1 Introduction

On Friday March 27th 2020, the LIGO and VIRGO Scientific Collaboration (LVC) ended the 3rd GW observing run (hereafter O3)¹, the results of which are compiled in Abbott et al. (2020b). Throughout the entirety of O3, GOTO followed-up every viable GW trigger. That is, any alert where a part of the localisation was in GOTO’s observable horizon. In this chapter, I will show GOTO’s first follow-up campaign over O3 using the ZiP subtraction algorithm to find transients.

Across the year-long observing run, the LVC announced 82 GW superevents, 26 of which were retracted. GOTO targeted 46 of the non-retracted events. Superevents are events detected in multiple LVC processing pipelines, and therefore have a much higher confidence of being real. Throughout O3, GOTO was indiscriminate of the type of events it followed-up (BBH, BNS, Mass-gap, etc.). This was in

¹This was 1 month premature due to the COVID-19 pandemic

part to test GOTO’s rapid response and sky coverage capabilities. However, there is also scientific merit in looking for counterparts to BBH mergers. As detailed in the introduction of this thesis, some models predict that BBH mergers can have an associated counterpart (Loeb, 2016; De Mink and King, 2017; McKernan et al., 2018). Following-up BBH merger events can place limits on the models, constraining the physics of these extreme systems.

One of the most important events for O3 was S190814bv (25324), the first NSBH merger. No counterpart was found this time (ENGRAVE et al., 2020; Gompertz et al., 2020b; Steeghs et al., in prep); however, the discovery shows that the pathway for binary evolution to make a NSBH system exists. The expectation is that the black hole engulfed the neutron star, but models exist emphasising that it is possible for these collisions to have a kilonova-like counterpart (Troja et al., 2008; Gompertz et al., 2020a; Kawaguchi et al., 2020). Finding counterparts to GW events quickly is crucial given the timescales it takes for these transients to evolve and fade. Using the ZiP subtraction pipeline built throughout this thesis, I will show it is possible to find high confidence transients, across large areas of the sky, coincident with GW triggers in real-time, with minimal human involvement.

7.2 GOTO follow-up

In chapters 1 and 2 I covered the methodology of searching for a counterpart and the scientific importance of doing so. In this section, I will highlight the culmination of GOTO’s efforts in the hunt for a GW counterpart. The following will discuss the follow-up of 46 GW triggers with GOTO, highlighting GOTO’s performance, demonstrating GOTO’s hold of the sky coverage monopoly. Finally, I will combine the above machine learning section to GOTO’s search to flag transients coincident with GW detection.

7.2.1 Collecting Data

LVC O3 run was split into two parts, O3a (which ran from the 1st of April to the 31st of September 2019) and O3b (November 1st 2019 to March 27th 2020). (For most, there was no significant change between the two phases.) For GOTO this was the upgrade from GOTO-4 to GOTO-8. This effectively doubled GOTO’s field of view, figures 2.4 and 2.5 shows this increase.

All of the data collected here are using either the *L*-band filter (UTs 1-4), or the clear filter (UTs 5-8). The pass bands can be considered equivalent for the purposes of this work. As discussed in chapter 4, GOTO uses a fixed grid of

tiles. This means each UT will always point at the same patch of sky for every tile and references can be built using the tile-number for each UT. Most of the tiles only contain one reference epoch at the time observation (this is usually a set of three exposures), therefore a considerable number of the subtractions in this chapter will be sub-optimal. On rare occasions only one exposure of the set could be used as a reference, the issues relating to this will be highlighted later. For this work, the sky co-ordinates are taken from the reference images which are solved using `atrometry.net` (Lang et al., 2010). However, recent work on the GOTO pipeline has improved the astrometric solutions. In the initial study on the O3a data, Gompertz et al. (2020b) used `HOTPANTS` as the primary subtraction method and subsequently fed the extracted sources to a simple Random Forest Classifier. This describes GOTO’s (old) real-time pipeline. I do not compare the results of the pipelines directly here because the criteria for transient detection in my new pipeline is much stricter. In other words, this new pipeline is exclusively looking for transients that occur on top of galaxies (from one catalogue) that have a 95% chance of being real, I cannot compare this directly to a classifier that looks at sources across the sky indiscriminately with a $>50\%$ chance of being real. The old pipeline has the luxury to vet this many candidates because it has the man-power to filter out lots of false positives. More spurious candidates could be sent for human vetting in the old pipeline.

The data used here are archival, but are treated through a real-time processing pipeline; that is `ZiP` subtraction whose products are fed to the machine learner discussed in Chapter 6. The data are collected with the image and follow-up statistics. Using these data combined with the published skymaps for each event², I can determine the probability region covered. To do this, I first find the area enclosed by each image in the survey. These calculated areas are then overlaid on the `HEALpix` skymap. Each pixel in the skymap represents an area of the sky and the pixel value contains the probability of the astrophysical source being in that region. If two or more GOTO exposures overlap, the `HEALpix` pixels contained within the overlap are only counted once. The sum of all the collected pixels are taken to find the total area covered and the percent localisation coverage³. The data in Gompertz et al. (2020b), for O3a, were also used to determine volumetric coverage. As this is not the purpose of this study and to avoid replicating previous work, GOTO’s effective depth is not considered here.

Once the image statistics are taken, the focus turns to transient detection.

²Skymaps can be found at <https://gracedb.ligo.org/>

³This routine was turned into a handy package and was used in Gompertz et al. (2020b) and ENGRAVE et al. (2020)

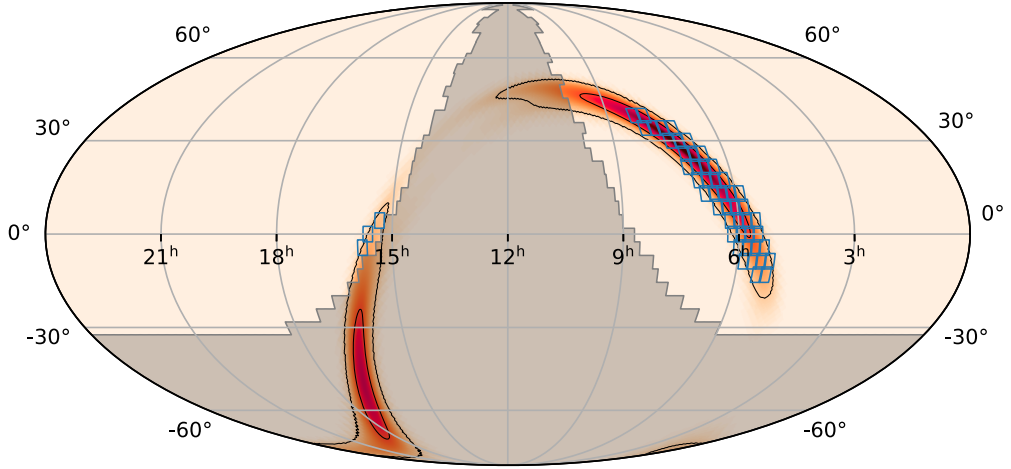


Figure 7.1: Visualisation GOTO’s coverage of S190923y as seen in (Gompertz et al., 2020b). GOTO covered over $720^{\circ 2}$ totalling 58% of the observable probability

Every source detected in the S_{Corr} images are fed through the DNN classifier. They are then filtered on the machine learning class score. Anything that passes the 95% confidence threshold is then given to the minor planet checker (MPC). As minor planets make up the primary astrophysical contaminant when searching for transients, filtering them out is a natural step. The MPC uses the predicted sky-positions of known minor planets with the catalogue from the minor planet centre (Marsden, 1986). If a flagged source falls within 10 arcseconds of an expected minor planet it is ignored. Finally, transient candidates are selected based on their proximity to a galaxy. Any sources closer than 10 arcseconds to a galaxy in the GLADE catalogue are treated as potential extra-galactic transients (Dályá et al., 2018). It should be noted most galaxy catalogues are relatively incomplete at fainter magnitudes, meaning there will be a subset of host-less transients and transients on top of unresolved galaxies that this method will miss (Fremling et al., 2020). Additional quality checks are based on number of extracted sources in the difference image. If there are over 50,000 sources in the difference image, subtraction quality is considered too low. If there are more than 50 high confidence transients in a single subtraction, the science and reference frames are probably misaligned. Finally any sources found on the border of the CCD, the first and last 50 rows and columns of pixels, are disregarded as they are most often artefacts left from the convolution step of the subtraction process (see Figure 7.2)

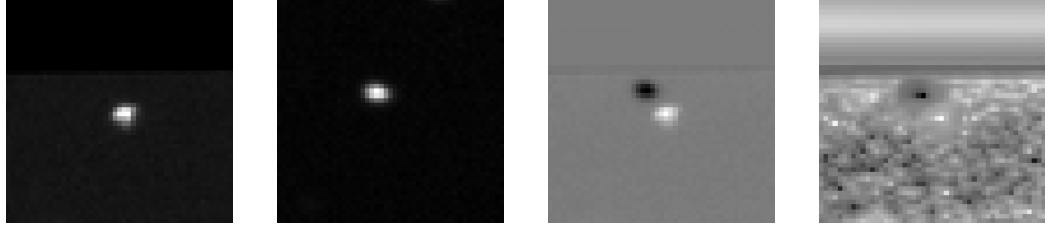


Figure 7.2: Example of misalignment and edge-case transient false positive

The results of all of this are shown in the next section and then discussed in 7.4.

7.3 Follow-up Results

Table 7.1 contains all the metrics regarding GOTO’s O3 GW follow-up. Each event has their 90% localisation regions, followed by GOTO’s response time to the alert (δt_{alert}) and the on-sky time between the detection of the event (δt_{event}). The area covered (A) and probability area covered, in total (PA), within GOTO’s visible horizon (PA_v), and the region covered exclusively with tiles that have difference imaging (PA_d) are provided. PA_v is calculated by excluding the probability regions behind the sun and below GOTO’s horizon and normalising the new skymap. The time dedicated observing each event and the average limiting magnitudes are given also.

The second table, Table 7.2, contains all the transient detection numbers for each GW trigger. Figures 7.3 and 7.4 display all the high confidence human vetted transients, their reference frames, and ZiP subtractions. The transients are presented in the order they appear in Table 7.2. Figure 7.7 provides some example false positives for later discussion.

Event ⁴	GCN	Region ⁵ deg ²	δt_{alert} Hours	δt_{event} Hours	A deg ²	PA %	PA _v %	PA _d %	MAP	Time Obs Hours	LVC run	m_{lim} mag	err _{lim} ±
S190408an	24096	387	10.81	11.38	156.14	20.24	22.49	13.22	BAYESTAR	0.85	O3a	19.34	0.36
S190412m	24098	156	14.0	15.0	295.23	94.44	94.72	30.26	BAYESTAR	1.80	O3a	19.34	0.33
S190421ar	24141	1444	29.05	48.3	114.3	8.88	36.99	<0.01	LALinference	1.85	O3a	19.31	0.28
S190425z	24168	7461	9.50	12.35	2667.12	22.02	54.01	17.42	LALinference	22.30	O3a	19.98	0.48
S190426c	24237	1131	5.00	5.3	772.69	54.13	66.69	9.68	LALinference	4.85	O3a	19.72	0.32
S190510g	24442	1166	0.40	1.42	116.09	0.21	1.06	0.19	LALinference	0.35	O3a	19.88	0.21
S190512at	24503	252	2.50	2.78	315.14	87.1	92.15	48.07	LALinference	8.95	O3a	19.55	2.58
S190513bm	24522	691	0.05	0.55	116.25	28.55	57.25	1.47	BAYESTAR	1.70	O3a	19.46	0.03
S190517h	24570	939	15.18	15.87	112.71	14.77	51.64	6.72	BAYESTAR	0.45	O3a	17.94	0.04
S190519bj	24598	967	4.35	5.35	664.77	76.80	83.65	5.26	BAYESTAR	6.95	O3a	18.73	0.02
S190521g	24621	1163	0.05	0.13	393.25	43.68	86.70	41.82	LALinference	2.90	O3a	19.40	0.04
S190521r	24632	488	15.12	15.2	720.66	91.91	92.87	66.0	BAYESTAR	4.50	O3a	19.04	0.02
S190630ag	24922	1172	2.40	2.40	1170.33	60.92	82.65	10.8	LALinference	6.80	O3a	19.62	0.5
S190706ai	24998	826	0.03	0.33	543.86	36.67	88.11	26.33	BAYESTAR	3.50	O3a	19.76	0.63
S190707q	25012	1375	11.73	12.38	722.95	34.38	76.07	19.57	BAYESTAR	3.70	O3a	19.8	0.38
S190718y	25087	7246	6.10	6.58	242.48	61.16	73.89	59.59	BAYESTAR	6.35	O3a	19.19	0.48
S190720a	25115	443	0.04	0.08	1358.26	62.12	73.30	45.7	LALinference	8.60	O3a	19.38	0.71

⁴On Tuesday March 10 2020, GOTO started hardware improvements, shortly after La Palma was closed. This resulted in S200311bg and S200316bj getting no follow-up

⁵As found on <https://gracedb.ligo.org/>

S190727h	25164	1357	14.90	15.03	714.66	42.34	93.37	30.11	LALinference	3.65	O3a	19.63	0.47
S190728q	25187	104	14.53	14.77	146.95	89.54	94.01	89.5	LALinference	1.75	O3a	19.63	0.24
S190814bv	25324	23	1.50	1.83	717.93	94.14	99.12	93.81	LALinference	7.05	O3a	17.67	0.92
S190828j	25497	228	15.81	16.08	442.19	9.11	23.66	7.9	LALinference	2.35	O3a	19.96	0.26
S190828l	25503	359	16.5	16.9	453.63	1.94	6.34	1.73	LALinference	2.80	O3a	20.01	0.15
S190901ap	25606	13613	0.04	0.12	2523.55	29.74	42.51	16.81	BAYESTAR	18.30	O3a	19.69	0.28
S190910d	25695	2482	0.03	0.13	1675.02	41.18	85.11	24.99	LALinference	10.75	O3a	18.77	0.58
S190915ak	25753	318	29.80	29.92	18.22	0.08	5.64	<0.01	LALinference	0.05	O3a	15.23	0.63
S190923y	25814	2107	13.69	13.83	723.68	39.38	58.88	35.7	BAYESTAR	2.35	O3a	19.43	0.18
S190924h	25829	303	2.90	2.97	281.27	70.24	70.59	67.21	LALinference	2.70	O3a	19.38	0.26
S190930s	25871	1748	6.20	6.28	2139.94	92.16	92.17	87.97	LALinference	14.25	O3a	19.77	0.21
S190930t	25876	24220	12.71	12.82	918.19	6.84	9.54	6.14	BAYESTAR	5.65	O3a	19.36	0.18
S191129u	26303	852	12.55	13.90	425.47	18.86	51.47	18.13	LALinference	1.30	O3b	19.5	0.48
S191205ah	26386	6378	45.28	45.33	2224.13	61.83	68.13	54.67	BAYESTAR	5.95	O3b	19.0	0.43
S191213g	26399	1393	21.02	21.38	1537.56	34.1	59.18	28.38	LALinference	12.25	O3b	18.07	1.93
S191215w	26441	361	44.96	45.08	891.41	84.39	98.67	81.47	LALinference	2.85	O3b	19.3	0.93
S191216ap	26454	253	21.43	21.68	613.06	82.77	82.77	76.03	LALinference	2.25	O3b	19.43	0.86
S191222n	26543	1850	0.02	0.10	2238.08	43.61	85.78	38.93	LALinference	12.65	O3b	19.57	0.95
S200105ae	26688	7373	26.40	27.92	2339.74	41.53	50.76	32.88	LALinference	7.40	O3b	18.47	0.6
S200112r	26715	4004	3.53	3.53	1890.21	37.51	55.06	32.54	LALinference	6.40	O3b	19.0	0.46
S200114f	26734	403	0.32	0.35	598.84	90.86	93.30	86.93	CWB	5.10	O3b	19.41	0.41
S200115j	26759	765	0.15	0.18	1143.38	48.34	70.47	41.38	LALinference	12.20	O3b	19.63	0.39
S200128d	26906	2293	0.09	0.12	2534.51	54.01	91.15	51.73	LALinference	13.25	O3b	19.51	0.86
S200208q	27014	28	6.79	6.82	1287.79	25.52	97.87	0.2	LALinference	2.90	O3b	17.88	0.75

S200213t	27042	2326	0.05	0.08	1892.32	65.83	76.86	58.52	LALinference	13.45	O3b	18.59	1.44
S200219ac	27130	781	10.90	11.97	803.11	34.30	43.63	33.35	LALinference	2.45	O3b	19.75	0.41
S200224ca	27184	72	25.68	25.72	151.86	87.25	87.25	86.51	LALinference	0.35	O3b	19.5	0.18
S200225q	27193	22	15.67	15.70	925.51	1.05	10.45	1.03	LALinference	2.10	O3b	19.69	0.19
S200302c	277278	5656	18.24	18.27	1196.73	23.1	42.68	19.24	LALinference	3.65	O3b	19.22	0.58
Mean		2390	10.85	11.83	911.55	46.73	64.80	34.91		5.75		19.44	
Median		953	10.16	11.68	719.30	41.94	71.95	30.19		3.68		19.45	

Table 7.1: Complete Collection of GOTO GW follow-up metrics. Event is the automatic event name and the GCN is where the announcement was first published. Region describes how large the 90% credibility localisation region is in square degrees. δt_{alert} and δt_{event} shows the response time of GOTO getting on sky after the alert was sent and when event went off, respectively. A is the area covered by GOTO in the search. PA is the probability covered over the search. PA_v is the probability covered by the region exclusively in GOTO's visible range. PA_d is the probability covered by images that have difference images. MAP describes the tool used to build the probability localisation. Time Obs, is the amount of time dedicated to following-up the event. LVC run describes which half of O3 the detection was made. m_{lim} is the average limiting magnitude of the search and err_{lim} is the error on this average. This table shows that GOTO is covering most of the probability region during its searches to a depth that can find transients similar to GW170817 (Gompertz et al., 2020b)

Event	Extractions	50% Confidence	MPC + 95% Confidence	Galaxy Coincident	Human Vetted
S190408an	129760	3797	5	0	0
S190412m	102825	1361	130	10	5
S190421ar	1330611	45438	0	0	0
S190425z	4310088	2149	51	4	0
S190426c	3389718	886	82	3	1
S190510g	35740	506	42	1	1
S190513bm	1282862	278	11	1	0
S190517h	62663	454	45	1	0
S190519bj	1302089	368	49	1	1
S190521g	223227	526	11	0	0
S190521r	1782291	196	35	0	0
S190630ag	2792717	1728	55	5	0
S190706ai	458744	1038	23	2	0
S190707q	3691333	262	1	0	0
S190718y	3534203	29293	0	0	0
S190720a	1847589	33804	0	0	0
S190727h	3037178	2913	5	0	0
S190728q	241952	25629	0	0	0
S190814bv	315541	107	37	1	0
S190828j	2143937	1676	3	1	0
S190828l	730317	302	26	0	0
S190901ap	1605211	12332	22	1	1*
S190910d	1101370	34479	0	0	0
S190923y	966056	14531	8	0	0
S190924h	457078	494	20	0	0
S190930s	7230830	148	3	0	0
S190930t	538085	122	13	0	0
S191129u	149269	393	62	2	2
S191205ah	787597	16554	46	3	1
S191213g	3080622	30349	0	0	0
S191215w	1157137	90	9	0	0
S191216ap	1451279	33153	1	0	0
S191222n	3816181	1528	20	0	0
S200105ae	1123799	520	32	0	0
S200112r	2270999	295	39	2	1
S200114f	3972069	14222	0	0	0
S200115j	3019999	344	31	0	0
S200128d	4327510	258	31	1	1
S200208q	118274	233	32	0	0
S200213t	1202412	5603	38	0	0
S200219ac	218380	471	63	5	0
S200224ca	50877	317	106	5	3
S200225q	1000850	21002	3	0	0
S200302c	1926480	21223	12	0	0
Total	74317748	361371	1202	49	17

Table 7.2: Results of the transient extraction. Showing that of 74317748 only 49 passed my pipeline filter, and 17 passed human inspection.

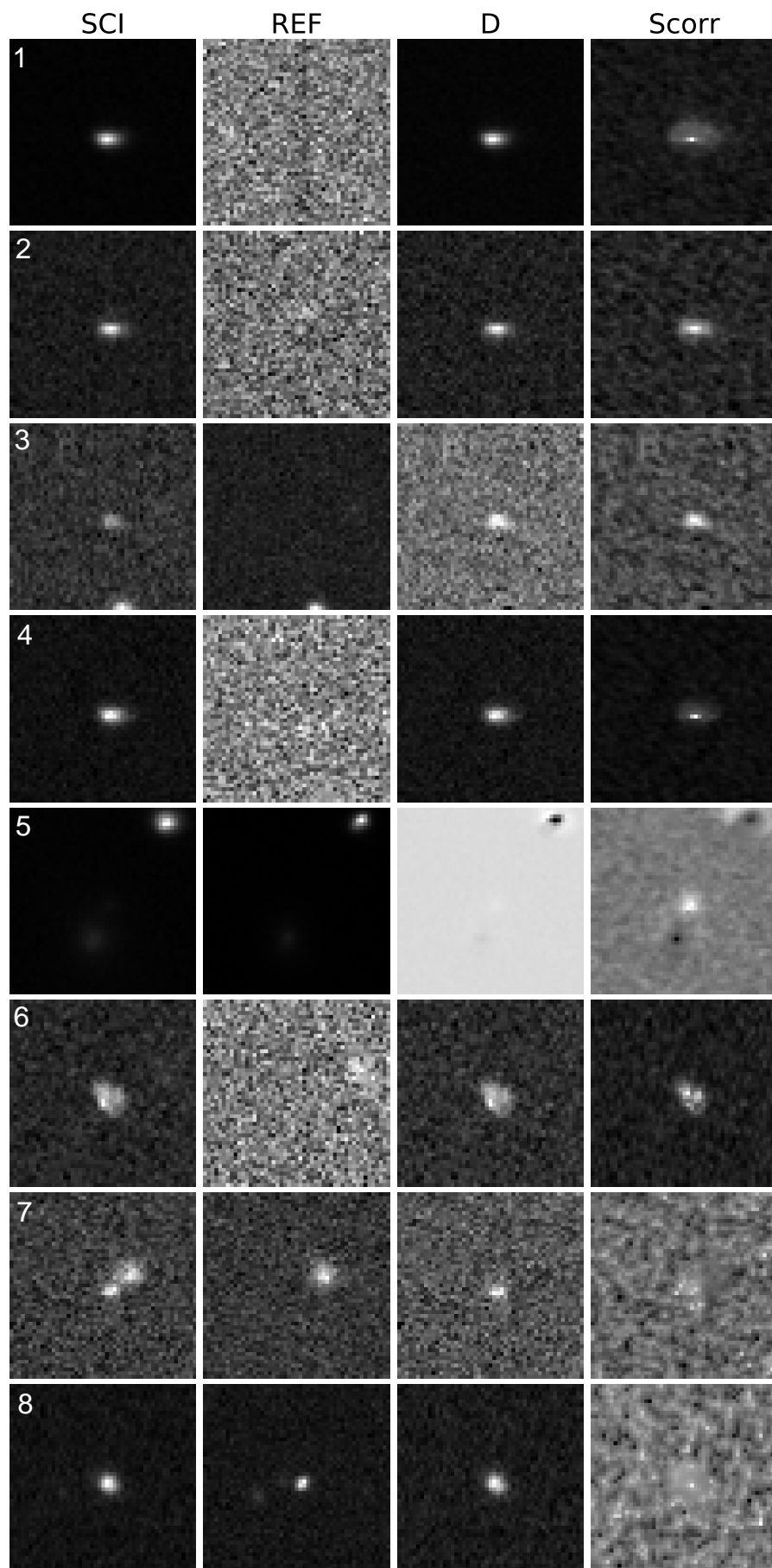


Figure 7.3: First set of transients from GOTO's GW follow-up from table 7.2.

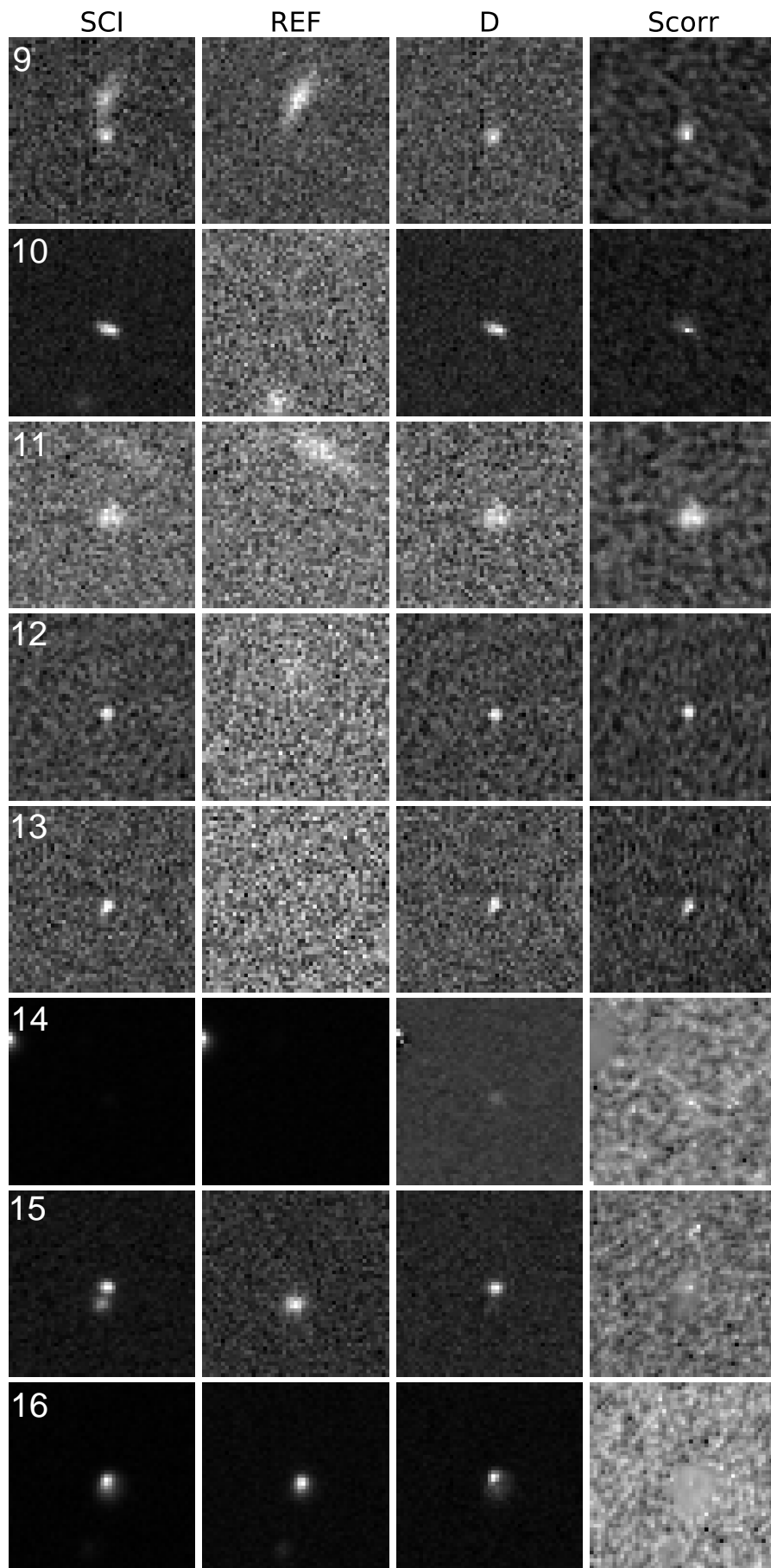


Figure 7.4: Second set of transients from GOTO's GW follow-up from table 7.2.

7.3.1 Delving into the Samples

Here transients are labelled with the number they are listed with in figures 7.3 and 7.4 and their event title followed by a letter listed alphabetically corresponding to what order they appear. All of the transients found here were first detected in GOTO after the GW announcement; in other words, the source had not been observed prior to the GW trigger and could therefore be attributed to the GW. The following distances were calculated assuming a Hubble constant (H_0) of 69.8, but the true value of this constant is still in debate (Hotokezaka et al., 2019; Freedman et al., 2019; Sokol, 2019; Yuan et al., 2019).

1 S190412m-A

This source has no announcement on the Transient Name Server (TNS) or GCN. This source is coincident with GAMA-86278 a galaxy with a redshift $z = 0.0244$ and has a magnitude of 16.5. The host galaxy is 110Mpc away, which does not overlap with the expected 812Mpc of the trigger. Therefore, this transient is not a counterpart. No follow-up of this transient exists, so has not been classified. The absolute magnitude of the source (M) is -18.7, which most likely makes this a type-1a supernova.

2 S190412m-B

This source has no announcement on the TNS or GCN. This source is coincident with GAMA-100195 (AKA 2dFGRS TGN386Z037) a galaxy with a redshift $z = 0.1593$. The host galaxy is at a distance of 740Mpc. Given the magnitude of the source is 15, the transient is most likely not attributed to the galaxy it is coincident with. The transient would have an absolute magnitude of -24.3, which is 10^5 times brighter than average peak luminosity of the average supernova. However, on the chance that it is, this galaxy is at the lower limit of the GW predicted distance.

3 S190412m-C

This source has no announcement on the TNS or GCN. This source is coincident with GAMA-24450 a galaxy with a redshift $z = 0.3288$. The coincident galaxy is too distant for the transient to be associated with this galaxy.

4 S190412m-D

This source is closest to GMAJ122238.63+004344.9; however, this galaxy does not have a known redshift and grazes just outside the 10 arcsecond minor planet coincidence. This source is most likely a minor planet that has diverged from the predicted orbit in the MPC, or the WCS solution on the image is erroneous.

5 S190412m-E

This is a type-1a supernova. On the TNS it has been dubbed SN 2019cdc and was detected at 18th Magnitude at a redshift of 0.02777. This is too close to be the associated counterpart. Using its light-curve, we can see this supernova exploded before the GW trigger, offering more evidence this transient is not a counterpart.

6 S190426c-A

This detection has no associated announcement. It is coincident with galaxy 3263608 found exclusively in Dályá et al. (2018) at $z = 0.02927$. This event is too close to be associated with S190426. No additional follow-up of this transient exists. The transient has an apparent magnitude of ~ 17 giving it an absolute magnitude of -18.5 .

7 S190510g-A

This source occurred directly on top of SDSS J144312.23+100239.0 a galaxy at $z = 0.0920$. There is no associated transient announcement. It falls in the correct distance regime for S190510g; however, does not coincide with the localisation of the updated skymap. This transient appears in GOTO with a magnitude of 17.5, this gives it an absolute magnitude of -20.3 . This would make it a super-luminous supernova.

8 S190519bj-A

Is a Quasar, QSO J1246-2547. In quiescence it is 17th magnitude. The flare occurs directly on the galactic centre. Meaning this is most likely nuclear variability.

***S190901ap-A (GOTO2019hope)**

This transient is GOTO2019hope (Ackley et al., 2019) and was detailed in chapter 5.4.3. As it was specifically used to test transient detection with ZiP it has been excluded from the follow up statistics. The source was not found blindly as

the pipeline was tested on this source and therefore should not be included in the performance metrics for finding sources blindly.

9 S191129u-A

This is supernova SN 2019tvv which occurred at a brightness of 16th magnitude in a galaxy with a redshift $z = 0.026$. This is too close to be the associated counterpart. Additional evidence shows detonation was well before the GW trigger.

10 S191129u-B

This source is closest to 2dFGRS TGN296Z256 with a redshift of 0.115 and grazes just outside the 10 arcsecond coincidence for a minor planet. This source is most likely a minor planet that has diverged from the predicted orbit in the MPC or the image has a poor WCS solution.

11 S191205ah-A

Coincident with LEDA 1101926, no known redshift. No additional follow-up of this source exists. This transient has a magnitude of ~ 17.5 .

12 S200112r-A

This source lands on SDSS J121927.66+002740.1, which is a Galaxy cluster at a redshift of 0.45. The transient has a brightness of 16th magnitude; so it is likely not associated with the galaxy. There is no transient announcement and no follow-up.

13 S200128d-A

SDSSJ025706.68+032451.6 is the closest galaxy to the source, no redshift available. Using the image WCS, this source also lands pretty close to an eclipsing binary, CRTS J025648.1+032438. Though this is probably not the same source, the eclipsing binary's minima is also 17th magnitude, so should be able to see the minima in the reference image.

14 S200224ca-A

This source is coincident with 2dFGRS TGN240Z118 a galaxy at $z=0.2535$. This falls within the expected error region for the GW event, but given the distance, it is unlikely the source is associated with this galaxy. No follow-up of this transient exists.

15 S200224ca-B

The coincident galaxy is 2dFGRS TGN235Z198 $z=0.117$, which is too close to be associated with the GW event. This source was seen at 18.5th mag in GOTO, which would give it an absolute magnitude of -19.6. This has no transient announcement, and therefore has not been followed-up.

16 S200224ca-C

This is most likely a known dwarf nova (V* TW Vir). It is aligned with 2dFGRS TGN174Z271 a galaxy at $z = 0.1203$. This is too close to be associated with the GW event.

7.3.2 Variables that were vetted by a human

There were other astrophysical sources that were identified by the machine learner, but these failed human inspection.

LINEAR 8377916 – RR Lyr Variable Star

This source passed the machine learning filter and all the other transient checks. This is an astrophysical source that is coincident with 2MASX J13092484+1850014. It is an RR Lyrae star, a type of star characterised by its pulsations. Figure 7.5 shows the subtraction of the source. As the source is relatively bright in both frames it was considered a false positive.

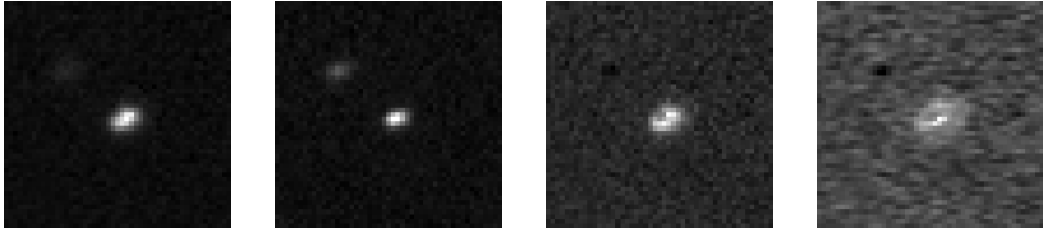


Figure 7.5: RR Lyr star that passed the automatic filter but not a human.

V* TW Vir – Dwarf Nova

This final source is another Dwarf novae. Again it passed all of the classifier checks, occurring on a galaxy labelled 2dFGRS TGN174Z271. It is pretty clear that the source exists in both frames and gets noticeably brighter in Figure 7.6. This is not a new source and therefore cannot be a transient associated with a coincident GW, this is why it was disregarded in the human vetting stage.

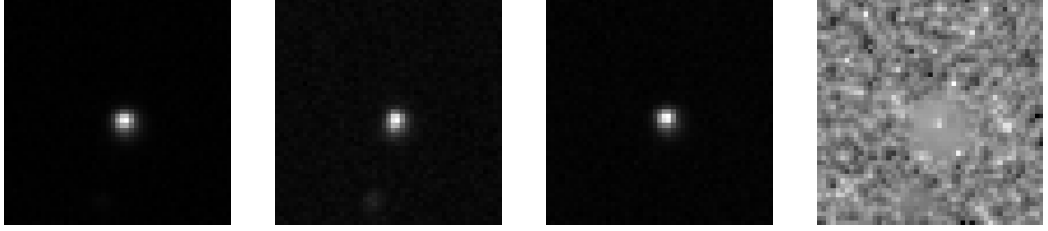


Figure 7.6: Dwarf Nova that passed the automatic filter but not a human

7.4 Discussion

7.4.1 GOTO is Doing its Job! (Gravitational-Wave Follow-Up)

Over the course of 11 months, the GOTO prototype was able to chase 46 GW triggers, Table 7.1 shows the product of these efforts. GOTO was established so a system could be developed that can search wide localisation regions quickly for multi-messenger astronomy. Even with just 4 telescopes, GOTO has the capability to cover $\sim 2000\text{deg}^2$ in just one night. This value is already setting the gold standard for high cadence transient surveys (Kasliwal et al., 2016; Ghosh et al., 2017; Cowperthwaite et al., 2018; Antier et al., 2020; ENGRAVE et al., 2020). By quadrupling its field of view on La Palma, GOTO will not struggle to cover large localisation regions on a nightly basis. Not only is GOTO capable of wide field surveying, its automatic ingestion of triggers allows for a rapid response time to GW alerts. On average, the response time is 10 hours, this is due to the GW triggers coming in while the sun is up on La Palma. GOTO is able to hit the sky less than a minute after a trigger, this is one of the fastest response times of any survey scope. This is demonstrated in ENGRAVE et al. (2020) where multiple surveys of S190814bv are used to constrain models of NSBH mergers.

With its current setup, GOTO can observe almost the entire localisation region visible to it in one to two nights of observing. A concern is GOTO only being able to cover the Northern hemisphere. This means there is $\sim 50\%$ of the sky not

being looked at. Developments are underway for a GOTO-South in Australia. This will be an additional 16 telescopes looking at the Southern hemisphere. The GOTO network will see all of the GW localisation regions, excluding anything behind the sun.

One issue, flagged by Gompertz et al. (2020b), is GOTO’s limiting depth. The average limiting magnitude is a little short of GOTO’s expected performance. More and improved GW detectors will drastically lower the localisation region. This, in combination with 8 more second generation telescopes, will mean GOTO can search entire skymaps in just a few pointings. As GOTO will have the luxury of such a wide field of view, it will be able to spend more time looking at specific areas of the sky. This should remedy the limiting depth problem.

Table 7.1 shows that GOTO can, in fact, do its job hunting counterparts to GWs. Future improvements both to GOTO and the discovery of GWs will only increase GOTO’s impact in the multi-messenger community.

7.4.2 Finding Transients

Of the 17 sources found 2 were identified by other surveys, 3 were known variables, and 2 are suspected minor planets not caught by the MPC. One of the most interesting transients is the fifth transient for S190412m, S190412m-E. This is the only sources that was exclusively detected in the S_{Corr} image. The source is 18.01 magnitude and is close to the limiting magnitude for the image it is in. Because of how faint it was, S190412m-E was not found in the HOTPANTS subtraction pipeline. This emphasises the message from Chapter 5, ZiP makes optimal use of the data and can therefore find transients other algorithms would miss.

Using the ratio of high confidence transients and accounting for the false positive rate and population of variables, we would expect to find a transient every 15 square degrees, with an upper limit of 9 and lower limit of 40. This figure is highly dependent on source and galaxy density of each field. Kochanek et al. (2017) and Ho et al. (2018) provide similar numbers for different optical surveys. The transient rates determined here are conservative estimates based on the performance of the classifier and act as a lower limit for the performance of the GOTO-prototype. With a more lenient source classifier and fewer restrictions on where transients can occur, the number of transient sources detected in GOTO will increase.

One thing to point out, noticeably fewer transient were found in UTs 5-8 in the O3b data-set. This may be due to a bias in the classifier, or because the image quality was under improvement through telescope collimation during the transient searches. Additionally, the scopes were new to the system; meaning there were

limited data available. In particular, many of the fields observed in UTs 5-8 did not have templates and therefore no subtraction could be done. When the bias in source classification is addressed and with optimal UT performance the transient rate for GOTO should improve. The classifier was also incredibly strict as to minimise false positives, with more human power the classifier can be relaxed because the workload will be more manageable.

The main source of false positives come from bad stacking, either in the reference image, science image, or both. GOTO takes a set of three exposures, if even one of these fails, for any reason, the resultant combined image will still contain imperfections from the two combined images. Chapter 4.3 shows that the combination of just two fields can be worse than using a single frame. Figure 7.7 provides some examples. The most regular contaminant appears to be bad columns. These can be masked easily enough. The other source of false positives comes from cosmic rays, hot pixels, and dead pixels. If the GOTO pipeline is unable to stack more than 2 images these artefacts are not removed, meaning the impostors get through as potential transients. There are two ways to solve this issue, the first is to make sure every image that is subtracted has both a science and reference frame that are comprised of at least 3 stacked images. (GOTO has started taking sets of 4 exposures to help mitigate this problem.) The other and more economic solution is to train the classifier to compare the PSF against the potential transient. This was discussed earlier in Section 6.4. If the source matches the PSF it is most likely real. The biggest problem is only having one reference epoch. Revisits are vital with image subtraction, having only one epoch means the reference image cannot be significantly better quality than the science frame. The most pressing improvement from this study is to have comprehensive reference grid ready for transient detection.

Assuming optimal circumstances, each of the high confidence transients would be sent for human vetting within a minute of the exposure being taken. Therefore, it is entirely reasonable to believe, that for GOTO, high confidence transients can be announced the same night they have been detected. This is the best way to announce potential counterparts to GWs. This way, spectroscopic surveys can do detailed follow-up of these transients to quickly classify them.

7.4.3 Variables and Other Science with GOTO

Due to the rapid nature of GOTO, the dynamic sky is captured in many forms. This makes GOTO useful for a range of time-domain astronomy, highlighted by the detection of multiple dwarf novae and variable stars. GOTO not only has the capabilities for transient astronomy, but has the potential for many auxiliary science

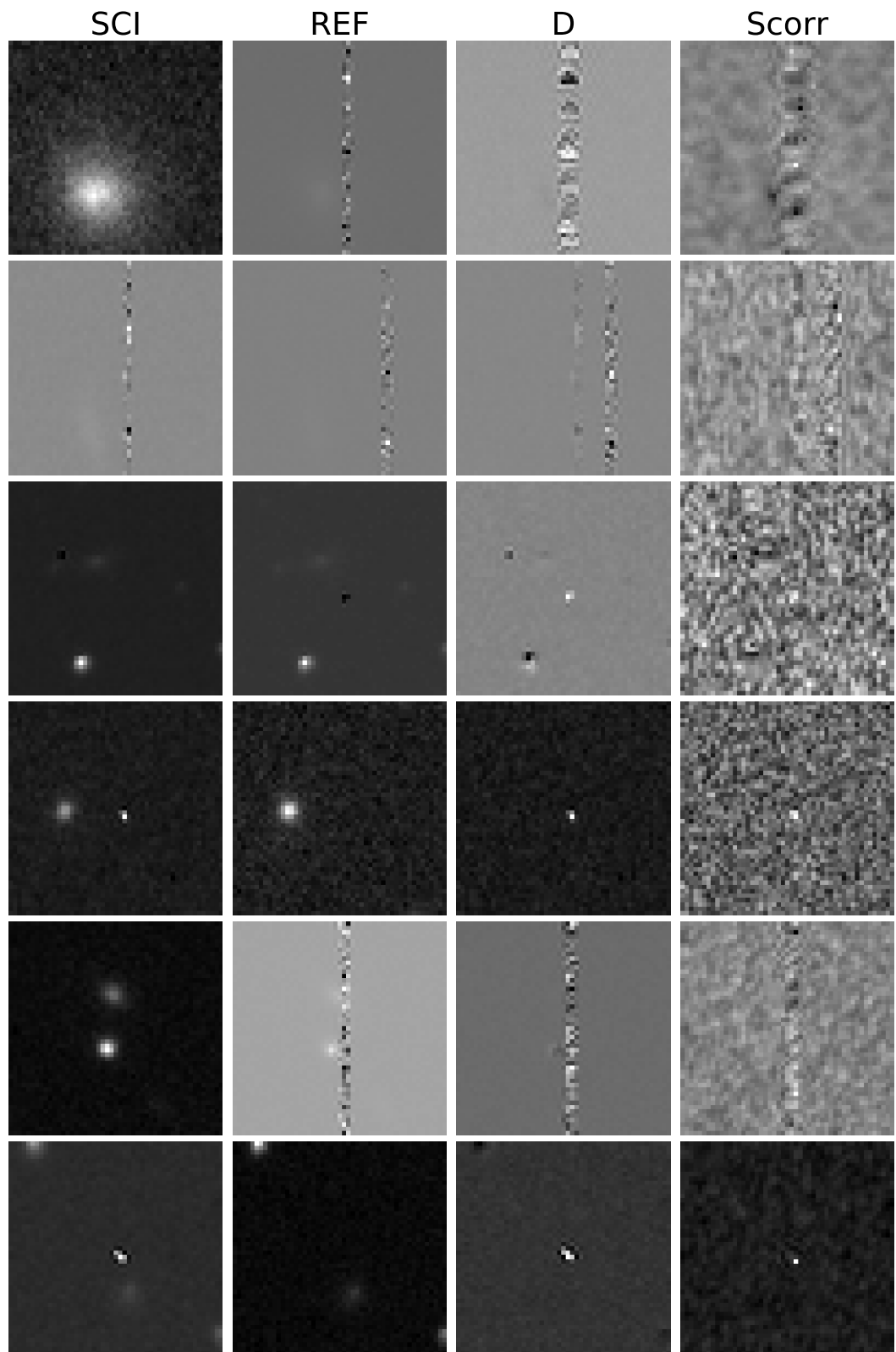


Figure 7.7: Example selection of false positives from the GW follow-up

cases. GOTO’s impact in time-domain astronomy is evidenced by the number of working group branches in the collaboration. A few examples: asteroids, Active galactic nuclei, Accreting Binaries, White Dwarfs, and Stellar Variability.

7.5 Conclusion

This chapter covered GOTO’s GW follow-up survey. The data taken throughout this campaign were passed through the ZiP image subtraction pipeline and then classified with the machine learner developed at the beginning of the chapter. I showed that GOTO was able to find a number of transients in real-time and that GOTO has the potential to redefine multi-messenger astronomy, by finding and announcing transients on sub-hour timescales of their detection. The discussion is closed with an introduction into GOTO’s capabilities for auxiliary science, evidenced with the variable sources found in the GW survey data.

Finally, I detailed the ZiP pipeline results and what they meant for the GW follow-up campaign. I used the results of the entire survey to analyse GOTO’s performance, highlighting both GOTO’s strength and shortcomings; I then offered solutions to the pitfalls through planned future improvements.

Given the future developments, GOTO has been and will continue to be a foundation for the next generation of multi-messenger astronomy programs.

Chapter 8

What Happens Next?

8.1 Thesis Summary

This thesis has detailed the conception of a new survey telescope; the Gravitational-wave Optical Transient Observer (GOTO). In the introduction, I detailed the science behind GW detection and the motivation behind fast identification and rapid follow-up of the explosive events. In Chapter 2, I covered the methodology of searching for a counterpart with optical CCDs. I paid special attention to the short comings of narrow field survey telescopes and how they lacked the capacity to cover the LVC localisation regions in any meaningful time-frame. This emphasised the necessity for deep, wide-field, high-cadence surveys like GOTO.

Chapter 3 opened up the first challenges faced when working with wide-field fast-optics data. Showing that obtaining focus automatically is a non-trivial task and requires a robust and flexible algorithm. I presented a novel solution to this challenge, highlighting that one can actually exploit astigmatisms in the optics to aid achieving focus quickly.

Chapter 4 covers two key preparation features required for image subtraction. The first is co-addition, drawing attention to a new method developed from Zackay and Ofek (2015a,b). I found that this algorithm was much slower than competing methods, however provided better quality output in both depth and resolution. This algorithm was then used on GOTO data and had a slight improvement to limiting magnitude depth, but not significantly enough to warrant changing the algorithm. The second half of the chapter detailed a new image registration technique. This new technique was shown to compute on the same timescales as leading methods, with a more accurate alignment result. This meant GOTO images could be co-added and aligned in real time in preparation for image subtraction. Furthermore, the errors

introduced using the new alignment transforms were shown to be negligible.

The new image subtraction algorithm, ZOGY, is introduced in chapter 5. It is shown to outperform the other contending algorithm, HOTPANTS, in terms of subtraction quality; however, the original ZOGY algorithm proved to be too slow. This was the motivation for a new parallel algorithm, ZiP. ZiP was shown to have the same image subtraction quality as ZOGY, but completed subtraction in only a fraction of the time. A variety of new PSF estimation techniques were also presented as a solution to coping with the source shape variability across GOTO’s wide field images. With a modified version of PSFex (Bertin, 2013) being the most reliable for GOTO’s needs. Finally I showed that ZiP works for more than just GOTO data and can be used on a variety of optical surveys. I concluded the chapter showing the new subtraction tool can find real transients in difficult data and in real-time during GW follow-up surveys.

Chapter 6 introduces machine learning as a way to filter out bogus detections from the difference images. Here, I developed a novel way to classify sources using a Deep Neural Network that looks at information from the input images. One future study from this work is looking at finding transients by just looking at the input frames without doing any difference imaging. Using the feature importance of the classifiers shows this is a realistic, and faster solution to finding transients in survey data.

Finally, chapter 7 is the culmination of chapters 4, 5, and 6; the implementation of an image subtraction pipeline using new co-addition, alignment, subtraction, and machine classification algorithms. The chapter closes affirming GOTO’s potential as a GW counterpart hunter by analysing its performance across O3. I claim GOTO is able to scan an impressive fraction of the sky and can rapidly identify transients, proving GOTO is fit for purpose and more.

8.2 The Future of GOTO

8.2.1 GOTO South

GOTO has just been awarded funding to run throughout O4, which is predicted to begin June 2022. GOTO is designed to scale, this means more scopes and more observatories. GOTO’s horizon includes adding 8 additional telescopes on the La Palma site, increasing its field of view by another $40^{\circ 2}$ (this will mean 16 telescopes on La Palma). This will add an ambidextrous functionality to GOTO’s searches. By having two independent mounts, one mount can do targeted searches while the other continues to build a reference map. This is just an example of how GOTO

can be exploited. Furthermore, new mount designs are in place with the addition of second generation ASA telescopes. The new mount will have a direct-drive motor. This means the mount will be connected directly to the motor which, in theory, will lower tracking errors from wind shake. The performance improvement in both the mount and astrographs will push GOTO to the leading edge of transient discovery.

More importantly, GOTO South, the Australian node of GOTO, is in development. GOTO will have a further 16 ASA telescopes in Australia fixed to a copy of the new direct-drive mount. This will unlock the restricted access of the Southern Hemisphere by adding a copy of the new mount down-under. GOTO will have total dominance over the entire sky, the only place GOTO will struggle to find counterparts is behind the sun.

By having 32 astrographs split evenly across both hemispheres, GOTO will have a combined field of view of $\sim 160^{\circ 2}$. GOTO will be able to survey the entire sky in a matter of days. In addition to its field of view and all sky cadence, GOTO will see down to ~ 20 -21st magnitude; making GOTO an unprecedented survey for sky-coverage and depth combined. For reference, ZTF has a coverage of 47 sq. deg. down to 21st magnitude (Kasliwal et al., 2020) and the evryscope has a sky coverage of 8150 sq. deg. but only down to 16th magnitude (Ratzloff et al., 2019).

8.2.2 Transients, they don't stop coming

Over the next 3 to 4 years, LIGO will be upgrading their GW detectors. This does not stop stars exploding or mergers from happening, but finding them may be more challenging. The LIGO hiatus will offer the unique opportunity to follow-up GRBs, and find Supernovae and other transients. This in itself will have scientific merit, but will also allow GOTO to further hone the skill in rapid identification of transients.

Furthermore, the hunt for kilonovae does not stop here. The ENGRAVE collaboration still stands, of which GOTO is a contributing member. These cosmic explosions still hold many secrets, finding more of them is important even if it is without their GW information. Recent work has shown that there is strong correlation with short GRBs and neutron star mergers (Jespersen et al., 2020), meaning chasing these high energy events offers one of the best opportunities at finding kilonovae. However, GRBs are viewing angle dependent, making blind kilonova searches an important part of the kilonova hunt.

8.2.3 Variable Stars, they never stop coming

GOTO’s cadence has ancillary science application. These changes in the dynamic sky caught by GOTO has benefit to other astronomers and astrophysicists not looking for stellar explosions. One example of this comes early in GOTO’s career. During its time commissioning data, GOTO was able to capture the rotation of a nearby asteroid (Borisov et al., 2018). This auxiliary science offers another way GOTO can contribute to astronomers while still preparing for the future of multi-messenger astronomy.

A different survey style has been proposed for GOTO during the LIGO hiatus. This would monitor specific sky tiles for long periods of time to monitor stellar variability. This is important for observing eclipsing binaries, pulsating stars, and potentially observing exoplanet transits as these stars need to be monitored on shorter cadences (minutes to hours) compared to the evolution of transients (hours to days). GOTO’s future ambidexterity will allow time-domain astronomy across multiple cadences allowing for simultaneous transient searches and short-term variability surveys.

8.2.4 Getting Ready for O4, hit the ground running

O4 will itself be a historical milestone. The expected sensitivities will detect an even broader range of GW progenitors; however, after O3, astronomers are prepared for what to expect in the upcoming mission. Current predictions expect O4 to see 10^{+52}_{-10} BNS merger events (Abbott et al., 2018b). This will mark the golden age for wide field surveys and GOTO has the potential to be one of the leaders during this time. Preparation is key, learning from the strengths and weaknesses of the O3 campaign GOTO will have the opportunity to find more GW counterparts than ever before.

The future of multi-messenger astronomy is exciting and GOTO will be in prime position to guarantee the science lives up to the hype.

Appendix A

Derivation of the ZOGY Equations

The ZOGY method (Zackay et al., 2016) works on a statistical premise. Assuming a reference image can be expressed as the following (anything subscript r refers to the reference image, anything subscript n refers to the new image, and anything subscript y applies to both):

$$R = F_r B \otimes P_r + \epsilon_r \quad (\text{A.1})$$

Note, this is a similar form as described in Alard and Lupton (1998). Where B denotes the background, the PSF kernel is given P_r , ϵ_r is the reference image additive noise, and F_r is the flux.

Using the null hypothesis (no new sources in the new frame), the new image can be seen as:

$$N_{|\mathcal{H}_0} = F_n B \otimes P_n + \epsilon_n \quad (\text{A.2})$$

If the image has a new point source not in the reference, a new hypothesis can be developed, assuming a source of flux (α) at position (q):

$$N_{|\mathcal{H}_1(\alpha, q)} = F_n B \otimes P_n + F_n \alpha \delta(q) \otimes P_n + \epsilon_n \quad (\text{A.3})$$

Where $\delta(q)$ is a 2 dimensional delta function that equals unity at q and zero everywhere else. Given that the images are properly background subtracted and that the dominant noise is from the background. One can assume all the pixels are uncorrelated, i.e. $\text{COV}(x_1, x_2) = 0$ for $[x_1 \neq x_2]$ and have uniform variance $V(\epsilon_y[x]) = \sigma_y^2$.

Applying the Neyman-Pearson lemma, that states the most powerful test of two simple hypothesis is given by the likelihood ratio of the two predictions (Neyman and Pearson, 1933). The power of a hypothesis test is the probability that the test will correctly reject the null hypothesis when the alternate hypothesis is true.

$$\mathcal{L}(q, a) = \frac{\mathcal{P}(N, R|\mathcal{H}_0)}{\mathcal{P}(N, R|\mathcal{H}_1(q, \alpha))} \quad (\text{A.4})$$

This is a ratio of probabilities. Unfortunately, one cannot make any assumptions about the background and no prior information about the background is given. This means the two probabilities in A.4 cannot be found directly. However, they can be solved using the law of conditional probabilities.

$$\mathcal{L}(q, a) = \frac{\mathcal{P}(N|R, \mathcal{H}_0)\mathcal{P}(R|\mathcal{H}_0)}{\mathcal{P}(N|R, \mathcal{H}_1(q, \alpha))\mathcal{P}(R|\mathcal{H}_1(q, \alpha))} \quad (\text{A.5})$$

Given that the probabilistic model for the reference is the same for both hypotheses $\mathcal{P}(R|\mathcal{H}_0) = \mathcal{P}(R|\mathcal{H}_1(q, \alpha))$, A.5 can be simplified to:

$$\mathcal{L}(q, a) = \frac{\mathcal{P}(N|R, \mathcal{H}_0)}{\mathcal{P}(N|R, \mathcal{H}_1(q, \alpha))} \quad (\text{A.6})$$

The next step is to find $\mathcal{P}(N|R, \mathcal{H}_0)$ and $\mathcal{P}(N|R, \mathcal{H}_1(q, \alpha))$. Assuming the images are still background dominated, the Fourier transform of all of the hypotheses is presented:

$$\widehat{R} = F_r \widehat{B} \widehat{P}_r + \widehat{\epsilon}_r \quad (\text{A.7})$$

$$\widehat{N}_{|\mathcal{H}_0} = F_n \widehat{B} \widehat{P}_n + \widehat{\epsilon}_n \quad (\text{A.8})$$

$$\widehat{N}_{|\mathcal{H}_1(q, \alpha)} = F_n (\widehat{B} + \alpha \widehat{\delta(q)}) \widehat{P}_n + \widehat{\epsilon}_n \quad (\text{A.9})$$

Where the hat denotes a Fourier transform.

Using the fact that R is a measured value, A.7 can be rearranged to predict the background.

$$\widehat{B} = \frac{\widehat{R}}{F_r \widehat{P}_r} - \frac{\widehat{\epsilon}_r}{F_r \widehat{P}_r} \quad (\text{A.10})$$

As there is now an expression for \widehat{B} , both the null and real hypotheses can be written:

$$\widehat{N}_{|\widehat{R} \mathcal{H}_0} = \left(\frac{\widehat{R}}{F_r \widehat{P}_r} - \frac{\widehat{\epsilon}_r}{F_r \widehat{P}_r} \right) F_n \widehat{P}_n + \widehat{\epsilon}_n \quad (\text{A.11})$$

$$\widehat{N}_{|\widehat{R}} \mathcal{H}_1(q, \alpha) = F_n \left(\left(\frac{\widehat{R}}{F_r \widehat{P}_r} - \frac{\widehat{\epsilon}_r}{F_r \widehat{P}_r} \right) + \alpha \widehat{\delta}(q) \right) \widehat{P}_n + \widehat{\epsilon}_n \quad (\text{A.12})$$

This means the probability to observe $N_{|\mathcal{H}_0}$:

$$\log \left(\mathcal{P} \left[\widehat{N} | \widehat{R}, \mathcal{H}_0 \right] \right) = \sum_f \frac{|\widehat{N} - \frac{F_n \widehat{P}_n \widehat{R}}{F_r \widehat{P}_r}|^2}{2V \left(\widehat{\epsilon}_n + \frac{F_n \widehat{P}_n \widehat{\epsilon}_r}{F_r \widehat{P}_r} \right)} \quad (\text{A.13})$$

Where f denotes the spatial frequencies. Using the linearity and scalar multiplication properties of the variance and simplifying, A.13 becomes:

$$\log \left(\mathcal{P} \left[\widehat{N} | \widehat{R}, \mathcal{H}_0 \right] \right) = \frac{1}{2} \sum_f \frac{\left| F_r \widehat{P}_r \widehat{N} - F_n \widehat{P}_n \widehat{R} \right|^2}{\sigma_n^2 F_r^2 |\widehat{P}_r|^2 - \sigma_r^2 F_n^2 |\widehat{P}_n|^2} \quad (\text{A.14})$$

Similarly, one can do this same process with A.12 to get:

$$\log \left(\mathcal{P} \left[\widehat{N} | \widehat{R}, \mathcal{H}_1(q, \alpha) \right] \right) = \frac{1}{2} \sum_f \frac{\left| F_r \widehat{P}_r \widehat{N} - F_n \widehat{P}_n \widehat{R} - \alpha F_n F_r \widehat{P}_n \widehat{P}_r \widehat{\delta}(q) \right|^2}{\sigma_n^2 F_r^2 |\widehat{P}_r|^2 - \sigma_r^2 F_n^2 |\widehat{P}_n|^2} \quad (\text{A.15})$$

Given that:

$$\log(\mathcal{L}(q, \alpha)) = \log \left(\mathcal{P} \left[\widehat{N} | \widehat{R}, \mathcal{H}_0 \right] \right) - \log \left(\mathcal{P} \left[\widehat{N} | \widehat{R}, \mathcal{H}_1(q, \alpha) \right] \right) \quad (\text{A.16})$$

and subbing A.14 and A.15 into A.16:

$$\log(\mathcal{L}(q, \alpha)) = \sum_f \frac{\Re \left[\left(F_r \widehat{P}_r \widehat{N} - F_n \widehat{P}_n \widehat{R} \right) \overline{\alpha F_n F_r \widehat{P}_n \widehat{P}_r \widehat{\delta}(q)} \right]}{\sigma_n^2 F_r^2 |\widehat{P}_r|^2 - \sigma_r^2 F_n^2 |\widehat{P}_n|^2} \quad (\text{A.17})$$

Where the bar represents the complex conjugate and \Re defines the real component of the term. The statistic is shown:

$$S(q) \equiv \frac{\log(\mathcal{L}(q, \alpha))}{\alpha} \quad (\text{A.18})$$

Introducing the proper image subtraction term, this is the same as 5.3:

$$\widehat{D} = \frac{\left(F_r \widehat{P}_r \widehat{N} - F_n \widehat{P}_n \widehat{R} \right)}{\sqrt{\sigma_n^2 F_r^2 |\widehat{P}_r|^2 - \sigma_r^2 F_n^2 |\widehat{P}_n|^2}} \quad (\text{A.19})$$

The PSF of D can be expressed as a function of the two modelled PSFs in A.19:

$$\widehat{P}_D = \frac{(F_r \widehat{P}_r F_n \widehat{P}_n)}{F_D \sqrt{\sigma_n^2 F_r^2 |\widehat{P}_r|^2 - \sigma_r^2 F_n^2 |\widehat{P}_n|^2}} \quad (\text{A.20})$$

where F_D is a normalisation flux:

$$\widehat{P}_D = \frac{(F_r F_n)}{\sqrt{\sigma_n^2 F_r - \sigma_r^2 F_n}} \quad (\text{A.21})$$

A.19, A.20, and A.21 can then be used to express A.18 with intuitive quantities:

$$S(q) = \left(F_d D \otimes \overleftarrow{\widehat{P}_D} \right) (q) \quad (\text{A.22})$$

This is the same expression as 5.4.

Appendix B

Derivation of Proper Co-Addition

Proper Co-Addition is described in Zackay and Ofek (2015a) and Zackay and Ofek (2015b), and is claimed to make optimal use of the data. Maximising signal to noise with the given input images. The following derivation is a statistical approach to building references from Zackay et al. (2016).

$$R_j = F_j B \otimes P_j + \epsilon_r \quad (\text{B.1})$$

$$\mathcal{L}(q, a) = \frac{\mathcal{P}(N, R_1 \dots R_j | \mathcal{H}_0)}{\mathcal{P}(N, R_1 \dots R_j | \mathcal{H}_1(q, \alpha))} \quad (\text{B.2})$$

Same as before, all references have the same likelihood. Law of conditional probability.

$$\mathcal{L}(q, a) = \frac{\mathcal{P}(N | R_1 \dots R_j, \mathcal{H}_0)}{\mathcal{P}(N | R_1 \dots R_j, \mathcal{H}_1(q, \alpha))} \quad (\text{B.3})$$

This provides the same result as eq A.10 and rearranging for \widehat{R} gives:

$$\widehat{R} = \widehat{B} F_r \widehat{P}_r - \widehat{\epsilon}_r \quad (\text{B.4})$$

Using the same analysis as employed to get eq A.13, one can find the best signal to noise statistic for B .

$$\widehat{B} = \frac{\sum_j \frac{F_j}{\sigma_j} \widehat{P}_j \widehat{M}_j}{\sqrt{\sum_j \frac{F_j^2}{\sigma_j^2} |\widehat{P}_j|}} + \widehat{\epsilon}_r \quad (\text{B.5})$$

Substituting B.5 into B.4 provides the final result:

$$\widehat{R} = \frac{\sum_j \frac{F_j}{\sigma_j} \widehat{P_j} \widehat{M_j}}{\sqrt{\sum_j \frac{F_j^2}{\sigma_j^2} |\widehat{P_j}|}} \quad (\text{B.6})$$

The same as in equation 4.2.

Bibliography

- Abadie J. et al., 2011. A gravitational wave observatory operating beyond the quantum shot-noise limit, *Nature Physics*, 7(12):962.
- Abbott B.P. et al., 2009. Search for gravitational waves from low mass compact binary coalescence in 186 days of ligos fifth science run, *Physical Review D*, 80(4):047101.
- Abbott B.P. et al., 2016a. Localization and broadband follow-up of the gravitational-wave transient gw150914, *The Astrophysical journal letters*, 826(1):L13.
- Abbott B.P. et al., 2016b. Observation of gravitational waves from a binary black hole merger, *Physical review letters*, 116(6):061102.
- Abbott B.P. et al., 2017a. Gravitational waves and gamma-rays from a binary neutron star merger: Gw170817 and grb 170817a, *The Astrophysical Journal Letters*, 848(2):L13.
- Abbott B.P. et al., 2017b. Gw170817: observation of gravitational waves from a binary neutron star inspiral, *Physical Review Letters*, 119(16):161101.
- Abbott B.P. et al., 2017c. Multi-messenger observations of a binary neutron star merger, *Astrophys. J. Lett*, 848(2):L12.
- Abbott B.P. et al., 2018a. Gw170817: Measurements of neutron star radii and equation of state, *Physical review letters*, 121(16):161101.
- Abbott B.P. et al., 2018b. Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra, *Living Reviews in Relativity*, 21:1–57.
- Abbott B.P. et al., 2019. tests of general relativity with gw170817, *Physical review letters*, 123(1):011102.

- Abbott R. et al., 2020a. Gravitational-wave constraints on the equatorial ellipticity of millisecond pulsars, *The Astrophysical journal letters*, 902(1):L21.
- Abbott R. et al., 2020b. Population properties of compact objects from the second ligo-virgo gravitational-wave transient catalog, *arXiv preprint arXiv:2010.14533*.
- Abbott R. et al., 2020c. Properties and astrophysical implications of the 150 m binary black hole merger gw190521, *The Astrophysical Journal Letters*, 900(1):L13.
- Ackermann M. et al., 2016. Fermi-lat observations of the ligo event gw150914, *The Astrophysical journal letters*, 823(1):L2.
- Ackley K. et al., 2019. LIGO/Virgo S190901ap: GOTO optical coverage and detection of AT2019pjb, *GRB Coordinates Network*, 25654:1.
- Ai S., Gao H., and Zhang B., 2020. What constraints on the neutron star maximum mass can one pose from gw170817 observations?, *The Astrophysical Journal*, 893(2):146.
- Akiyama K. et al., 2019. First m87 event horizon telescope results. iv. imaging the central supermassive black hole, *The Astrophysical Journal Letters*, 875(1):L4.
- Alard C. and Lupton R.H., 1998. A method for optimal image subtraction, *The Astrophysical Journal*, 503(1):325.
- AlBaghdadi A.J. and Alkoot F.M., 2005. Bagging knn classifiers using different expert fusion strategies, In *Proceedings of the International Workshop on pattern recognition in information systems*, 219–224.
- Alexander K. et al., 2017. The electromagnetic counterpart of the binary neutron star merger ligo/virgo gw170817. vi. radio constraints on a relativistic jet and predictions for late-time emission from the kilonova ejecta, *The Astrophysical Journal Letters*, 848(2):L21.
- Alexander K. et al., 2018. A decline in the x-ray through radio emission from gw170817 continues to support an off-axis structured jet, *The Astrophysical Journal Letters*, 863(2):L18.
- Almoznino E., Loinger F., and Brosch N., 1993. A procedure for the calculation of background in images, *Monthly Notices of the Royal Astronomical Society*, 265(3):641–648.
- Alvarez L.W., 1937. Nuclear k electron capture, *Physical Review*, 52(2):134.

- Amaro-Seoane P. et al., 2012a. elisa: Astrophysics and cosmology in the millihertz regime, *arXiv preprint arXiv:1201.3621*.
- Amaro-Seoane P. et al., 2012b. Low-frequency gravitational-wave science with elisa/ngo, *Classical and Quantum Gravity*, 29(12):124016.
- Andreoni I. et al., 2017. Follow up of gw170817 and its electromagnetic counterpart by australian-led observing programmes, *Publications of the Astronomical Society of Australia*, 34.
- Antier S. et al., 2020. The first six months of the advanced ligos and advanced virgos third observing run with grandma, *Monthly Notices of the Royal Astronomical Society*, 492(3):3904–3927.
- Arcavi I., 2018. The first hours of the gw170817 kilonova and the importance of early optical and ultraviolet observations for constraining emission models, *The Astrophysical Journal Letters*, 855(2):L23.
- Argast D. et al., 2004. Neutron star mergers versus core-collapse supernovae as dominant r-process sites in the early galaxy, *Astronomy & Astrophysics*, 416(3):997–1011.
- Artale M.C. et al., 2019. Host galaxies of merging compact objects: mass, star formation rate, metallicity, and colours, , 487(2):1675–1688.
- Babak S. et al., 2008. The mock lisa data challenges: from challenge 1b to challenge 3, *Classical and Quantum Gravity*, 25(18):184026.
- Badenes C. and Maoz D., 2012. The merger rate of binary white dwarfs in the galactic disk, *The Astrophysical Journal Letters*, 749(1):L11.
- Bailey S. et al., 2008. Object classification at the nearby supernova factory, *Astronomische Nachrichten: Astronomical Notes*, 329(3):292–294.
- Baldi P. and Sadowski P.J., 2013. Understanding dropout, In *Advances in neural information processing systems*, 2814–2822.
- Barbary K., 2016. Sep: Source extractor as a library, *The Journal of Open Source Software*, 1(6):59.
- Bartos I. et al., 2017. Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei, *The Astrophysical Journal*, 835(2):165.

- Bauswein A. and Janka H.T., 2012. Measuring neutron-star properties via gravitational waves from neutron-star mergers, *Physical review letters*, 108(1):011101.
- Beale H.D., Demuth H.B., and Hagan M., 1996. Neural network design, *Pws, Boston*.
- Becker A., 2015. Hotpants: High order transform of psf and template subtraction, *Astrophysics Source Code Library*.
- Becker A.C. et al., 2004. The deep lens survey transient search. i. short timescale and astrometric variability, *The Astrophysical Journal*, 611(1):418.
- Beloborodov A.M., 2017. A flaring magnetar in frb 121102?, *The Astrophysical Journal Letters*, 843(2):L26.
- Benz A.O. and Güdel M., 2010. Physical Processes in Magnetically Driven Flares on the Sun, Stars, and Young Stellar Objects, , 48:241–287.
- Berger E., 2014. Short-duration gamma-ray bursts, *Annual review of Astronomy and Astrophysics*, 52:43–105.
- Berriman G. et al., 2008. The montage image mosaic service: custom image mosaics on-demand.
- Berry R. and Burnell J., 2000. Astronomical image processing, *Willman-Bell, Inc.*
- Bertin E., 2009. Skymaker: astronomical image simulations made easy., *Memorie della Societa Astronomica Italiana*, 80:422.
- Bertin E., 2010. Swarp: Resampling and co-adding fits images together, *Astrophysics Source Code Library*.
- Bertin E., 2013. Psfex: Point spread function extractor, *Astrophysics Source Code Library*.
- Bertin E. and Arnouts S., 1996. Sextractor: Software for source extraction, *Astronomy and Astrophysics Supplement Series*, 117(2):393–404.
- Bijaoui A., 1980. Sky background estimation and application, *Astronomy and Astrophysics*, 84:81–84.
- Blanchard P. et al., 2017. The electromagnetic counterpart of the binary neutron star merger ligo/virgo gw170817. vii. properties of the host galaxy and constraints on the merger timescale, *The Astrophysical Journal Letters*, 848(2):L22.

- Blanton M.R. et al., 2011. Improved background subtraction for the sloan digital sky survey images, *The Astronomical Journal*, 142(1):31.
- Bloom J. et al., 2008. Towards a real-time transient classification engine, *Astronomische Nachrichten: Astronomical Notes*, 329(3):284–287.
- Bloom J. et al., 2012. Automating discovery and classification of transients and variable stars in the synoptic survey era, *Publications of the Astronomical Society of the Pacific*, 124(921):1175.
- Bolton C.T., 1972. Identification of cygnus x-1 with hde 226868, *Nature*, 235(5336):271.
- Bombaci I., 1996. The maximum mass of a neutron star., *Astronomy and Astrophysics*, 305:871.
- Borisov G. et al., 2018. Rotational variation of the linear polarization of the asteroid (3200) phaeon as evidence for inhomogeneity in its surface properties, *Monthly Notices of the Royal Astronomical Society: Letters*, 480(1):L131–L135.
- Bradshaw A.K., Lage C., and Tyson J.A., 2018. Characterization of lsst ccds using realistic images, before first light, In *High Energy, Optical, and Infrared Detectors for Astronomy VIII*, volume 10709, 107091L. International Society for Optics and Photonics.
- Braginsky V.B., Strigin S.E., and Vyatchanin S.P., 2002. Analysis of parametric oscillatory instability in power recycled ligo interferometer, *Physics Letters A*, 305(3-4):111–124.
- Brauer K. et al., 2020. Collapsar R-Process Yields Can Reproduce [Eu/Fe] Abundance Scatter in Metal-Poor Stars, *arXiv e-prints*, arXiv:2010.15837.
- Breiman L., 1996. Some properties of splitting criteria, *Machine Learning*, 24(1):41–47.
- Breiman L., 2001. Random forests, *Machine learning*, 45(1):5–32.
- Broderick J. et al., 2020. Lofar 144-mhz follow-up observations of gw170817, *Monthly Notices of the Royal Astronomical Society*, 494(4):5110–5117.
- Broekema P.C. et al., 2020. On optimising cost and value in compute systems for radio astronomy, *Astronomy and Computing*, 30:100337.

- Bryll R., Gutierrez-Osuna R., and Quek F., 2003. Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets, *Pattern recognition*, 36(6):1291–1302.
- Bulla M. et al., 2018. Shedding light on the geometry of kilonovae, *The Messenger*, 174:34–36.
- Bulla M. et al., 2019. The origin of polarization in kilonovae and the case of the gravitational-wave counterpart at 2017gfo, *Nature Astronomy*, 3(1):99.
- Burbidge E.M. et al., 1957. Synthesis of the elements in stars, *Reviews of modern physics*, 29(4):547.
- Burns E. et al., 2019. A summary of multimessenger science with neutron star mergers, *arXiv preprint arXiv:1903.03582*.
- Cai D., Zhang C., and He X., 2010. Unsupervised feature selection for multi-cluster data, In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 333–342.
- Caprini C. et al., 2016. Science with the space-based interferometer elisa. ii: Gravitational waves from cosmological phase transitions, *Journal of cosmology and astroparticle physics*, 2016(04):001.
- Carbillet M. et al., 2010. The caos problem-solving environment: recent developments, In *Adaptive Optics Systems II*, volume 7736, 773644. International Society for Optics and Photonics.
- Castelvecchi D., 2019. Gravitational-wave hunt restarts-with a quantum boost., *Nature*, 568(7750):16.
- Chadwick J. and Goldhaber M., 1934. A nuclear photo-effect: disintegration of the dipion by-rays.
- Chan R.H., Yuan X., and Zhang W., 2013. A phase model for point spread function estimation in ground-based astronomy, *Science China Mathematics*, 56(12):2701–2710.
- Chen H.Y., Fishbach M., and Holz D.E., 2018. A two per cent hubble constant measurement from standard sirens within five years, *Nature*, 562(7728):545.
- Chilingarian I.V. et al., 2018. A population of bona fide intermediate-mass black holes identified as low-luminosity active galactic nuclei, *The Astrophysical Journal*, 863(1):1.

- Chrimes A.A. et al., 2018. Investigating a population of infrared-bright gamma-ray burst host galaxies, *Monthly Notices of the Royal Astronomical Society*, 478(1):2–27.
- Chromey F.R. and Hasselbacher D.A., 1996. The flat sky: calibration and background uniformity in wide field astronomical images, *Publications of the Astronomical Society of the Pacific*, 108(728):944.
- Clesse S. and Garcia-Bellido J., 2015. Massive primordial black holes from hybrid inflation as dark matter and the seeds of galaxies, *Physical Review D*, 92(2):023524.
- Collaboration L.S. et al., 2017. A gravitational-wave standard siren measurement of the hubble constant, *Nature*, 551(7678):85–88.
- Connaughton V. et al., 2016. Fermi gbm observations of ligo gravitational-wave event gw150914, *The Astrophysical Journal Letters*, 826(1):L6.
- Corsi A. et al., 2018. An upper limit on the linear polarization fraction of the gw170817 radio continuum, *The Astrophysical Journal Letters*, 861(1):L10.
- Coughlin M.W. et al., 2018. Optimizing searches for electromagnetic counterparts of gravitational wave triggers, *Monthly Notices of the Royal Astronomical Society*, 478(1):692–702.
- Coulter D. et al., 2017. Swope supernova survey 2017a (sss17a), the optical counterpart to a gravitational wave source, *Science*, 358(6370):1556–1558.
- Covino S. et al., 2017. The unpolarized macronova associated with the gravitational wave event gw 170817, *Nature Astronomy*, 1(11):791.
- Cowperthwaite P. et al., 2018. An empirical study of contamination in deep, rapid, and wide-field optical follow-up of gravitational wave events, *The Astrophysical Journal*, 858(1):18.
- Creminelli P. and Vernizzi F., 2017. Dark energy after gw170817 and grb170817a, *Physical review letters*, 119(25):251302.
- Crowder J. and Cornish N.J., 2005. Beyond lisa: Exploring future gravitational wave missions, *Physical Review D*, 72(8):083005.
- Cunningham P. and Delany S.J., 2020. k-nearest neighbour classifiers—, *arXiv preprint arXiv:2004.04523*.

- Cutler C. and Flanagan E.E., 1994. Gravitational waves from merging compact binaries: How accurately can one extract the binarys parameters from the inspiral waveform?, *Physical Review D*, 49(6):2658.
- Cutri R. et al., 2003. Vizier online data catalog: 2mass all-sky catalog of point sources (cutri+ 2003), *yCat*, II–246.
- Cutter R., 2019. ryanc123/lazypsf: Lazypsf.
- Da Costa G., 1992. Basic photometry techniques, In *Astronomical CCD Observing and Reduction Techniques*, volume 23, 90.
- Dai L., McKinney J.C., and Miller M.C., 2017. Energetic constraints on electromagnetic signals from double black hole mergers, *Monthly Notices of the Royal Astronomical Society: Letters*, L92–L96.
- Dálya G. et al., 2018. Glade: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era, *Monthly Notices of the Royal Astronomical Society*, 479(2):2374–2381.
- Danzmann K. et al., 2011. Lisa: Unveiling a hidden universe, *Assessment Study Report ESA/SRE*, 3(2).
- D’avanzo P. et al., 2018. The evolution of the x-ray afterglow emission of gw 170817/grb 170817a in xmm-newton observations, *Astronomy & Astrophysics*, 613:L1.
- De Mink S. and King A., 2017. Electromagnetic signals following stellar-mass black hole mergers, *The Astrophysical Journal Letters*, 839(1):L7.
- Demorest P.B. et al., 2012. Limits on the stochastic gravitational wave background from the north american nanohertz observatory for gravitational waves, *The Astrophysical Journal*, 762(2):94.
- Dessart L. et al., 2012. Superluminous supernovae: 56ni power versus magnetar radiation, *Monthly Notices of the Royal Astronomical Society: Letters*, 426(1):L76–L80.
- Dhawan S. et al., 2020. Constraining the observer angle of the kilonova at2017gfo associated with gw170817: Implications for the hubble constant, *The Astrophysical Journal*, 888(2):67.

- Dobie D. et al., 2018. A turnover in the radio light curve of gw170817, *The Astrophysical Journal Letters*, 858(2):L15.
- Dobke B.M. et al., 2010. Astronomical image simulation for telescope and survey development, *Publications of the Astronomical Society of the Pacific*, 122(894):947.
- Dominik M. et al., 2012. Double compact objects. i. the significance of the common envelope on merger rates, *The Astrophysical Journal*, 759(1):52.
- Doyle J. et al., 2018. Stellar flare oscillations: evidence for oscillatory reconnection and evolution of mhd modes, *Monthly Notices of the Royal Astronomical Society*, 475(2):2842–2851.
- Dozat T., 2016. Incorporating nesterov momentum into adam.
- Drout M. et al., 2017. Light curves of the neutron star merger gw170817/sss17a: Implications for r-process nucleosynthesis, *Science*, 358(6370):1570–1574.
- du Buisson L. et al., 2015. Machine learning classification of sdss transient survey images, *Monthly Notices of the Royal Astronomical Society*, 454(2):2026–2038.
- Duchêne G. and Kraus A., 2013. Stellar multiplicity, *Annual Review of Astronomy and Astrophysics*, 51.
- Duchi J., Hazan E., and Singer Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization., *Journal of machine learning research*, 12(7).
- Duev D.A. et al., 2019. Real-bogus classification for the zwicky transient facility using deep learning, *Monthly Notices of the Royal Astronomical Society*, 489(3):3582–3590.
- Duquennoy A. and Mayor M., 1991. Multiplicity among solar-type stars in the solar neighbourhood. ii-distribution of the orbital elements in an unbiased sample, *Astronomy and Astrophysics*, 248:485–524.
- Dyer M.J., 2020. A telescope control and scheduling system for the gravitational-wave optical transient observer, *arXiv preprint arXiv:2003.06317*.
- Dyer M.J. et al., 2018. A telescope control and scheduling system for the gravitational-wave optical transient observer (goto), In *Observatory Operations: Strategies, Processes, and Systems VII*, volume 10704, 107040C. International Society for Optics and Photonics.

- Edlund J.A. et al., 2005. White-dwarf–white-dwarf galactic background in the lisa data, *Physical Review D*, 71(12):122003.
- Einstein A., 1907. Pelativitatsprinzip und die aus demselben gezogenen folgerungen, *Jahrbuch der Rdioaktivitat*, 4:411–462.
- Einstein A., 1915. Die feldgleichungen der gravitation, *SPAW*, 844–847.
- Einstein A., 1995. *Relativity: the special and the general theory*. Broadway Books. ISBN 0517884410.
- Einstein A. and Rosen N., 1937. On gravitational waves, *FrInJ*, 223:43–54.
- Eldridge J. and Stanway E., 2016. Bpass predictions for binary black hole mergers, *Monthly Notices of the Royal Astronomical Society*, 462(3):3302–3313.
- ENGRAVE et al., 2020. Observational constraints on the optical and near-infrared emission from the neutron star-black hole binary merger s190814bv, *arXiv preprint arXiv:2002.01950*.
- Essick R., Landry P., and Holz D.E., 2020. Nonparametric inference of neutron star composition, equation of state, and maximum mass with gw170817, *Physical Review D*, 101(6):063007.
- Evans P. et al., 2017. Swift and nustar observations of gw170817: detection of a blue kilonova, *Science*, 358(6370):1565–1570.
- Evans P.A. et al., 2016. Swift follow-up of the gravitational wave source gw150914, *Monthly Notices of the Royal Astronomical Society: Letters*, 460(1):L40–L44.
- Ezquiaga J.M. and Zumalacárregui M., 2017. Dark energy after gw170817: dead ends and the road ahead, *Physical review letters*, 119(25):251304.
- Faber J.A. et al., 2002. Measuring neutron-star radii with gravitational-wave detectors, *Physical review letters*, 89(23):231102.
- Farrell J.E. et al., 2003. A simulation tool for evaluating digital camera image quality, In *Image Quality and System Performance*, volume 5294, 124–131. International Society for Optics and Photonics.
- Fawcett T., 2004. Roc graphs: Notes and practical considerations for researchers, *Machine learning*, 31(1):1–38.

- Finkelstein D., 1958. Past-future asymmetry of the gravitational field of a point particle, *Physical Review*, 110(4):965.
- Fraschetti F., 2018. Possible role of magnetic reconnection in the electromagnetic counterpart of binary black hole merger, *Journal of Cosmology and Astroparticle Physics*, 2018(04):054.
- Freedman W.L. et al., 2019. The carnegie-chicago hubble program. viii. an independent determination of the hubble constant based on the tip of the red giant branch, *The Astrophysical Journal*, 882(1):34.
- Fremling C. et al., 2020. The zwicky transient facility bright transient survey. i. spectroscopic classification and the redshift completeness of local galaxy catalogs, *The Astrophysical Journal*, 895(1):32.
- Fritschel P., 2003. Second generation instruments for the laser interferometer gravitational wave observatory (ligo), In *Gravitational-Wave Detection*, volume 4856, 282–292. International Society for Optics and Photonics.
- Gaia-Collaboration et al., 2018. VizieR online data catalog: Gaia dr2 (gaia collaboration, 2018), *VizieR Online Data Catalog*, 1345.
- Gao H., Zhang B., and Lü H.J., 2016. Constraints on binary neutron star merger product from short grb observations, *Physical Review D*, 93(4):044065.
- Gary B.L. and Healy D., 2006. Image subtraction procedure for observing faint asteroids, *Minor Planet Bulletin*, 33:16–18.
- Ghosh S. et al., 2017. Hunting electromagnetic counterparts of gravitational-wave events using the zwicky transient facility, *Publications of the Astronomical Society of the Pacific*, 129(981):114503.
- Giacomazzo B. and Perna R., 2012. General relativistic simulations of accretion induced collapse of neutron stars to black holes, *The Astrophysical Journal Letters*, 758(1):L8.
- Gilmore K. et al., 2008. The lsst camera overview: design and performance, In *Ground-based and Airborne Instrumentation for Astronomy II*, volume 7014, 70140C. International Society for Optics and Photonics.
- Gittins F. and Andersson N., 2019. Population synthesis of accreting neutron stars emitting gravitational waves, *Monthly Notices of the Royal Astronomical Society*, 488(1):99–110.

- Golub B. and Jackson M.O., 2010. Naive learning in social networks and the wisdom of crowds, *American Economic Journal: Microeconomics*, 2(1):112–49.
- Gompertz B., Levan A., and Tanvir N., 2020a. A search for neutron star–black hole binary mergers in the short gamma-ray burst population, *The Astrophysical Journal*, 895(1):58.
- Gompertz B.P. et al., 2020b. Searching for Electromagnetic Counterparts to Gravitational-wave Merger Events with the Prototype Gravitational-wave Optical Transient Observer (GOTO-4), *Monthly Notices of the Royal Astronomical Society*. ISSN 0035-8711. Staa1845.
- Gottlieb O., Nakar E., and Piran T., 2018. The cocoon emission—an electromagnetic counterpart to gravitational waves from neutron star mergers, *Monthly Notices of the Royal Astronomical Society*, 473(1):576–584.
- Graham M. et al., 2020. Candidate electromagnetic counterpart to the binary black hole merger gravitational-wave event s190521g, *Physical Review Letters*, 124(25):251102.
- Greisen E.W. and Calabretta M.R., 2002. Representations of world coordinates in fits, *Astronomy & Astrophysics*, 395(3):1061–1075.
- Hagenauer J., Offer E., and Papke L., 1996. Iterative decoding of binary block and convolutional codes, *IEEE Transactions on information theory*, 42(2):429–445.
- Haggard D. et al., 2017. A deep chandra x-ray study of neutron star coalescence gw170817, *The Astrophysical Journal Letters*, 848(2):L25.
- Hajela A. et al., 2020a. Chandra observations of gw170817 at 2.5 years since merger (all epochs 4 of 4), *GCN*, 27414:1.
- Hajela A. et al., 2020b. X-ray emission from gw170817 2.5 years after the merger, *Research Notes of the AAS*, 4(5):68.
- Hamidani H., Kiuchi K., and Ioka K., 2020. Jet propagation in neutron star mergers and gw170817, *Monthly Notices of the Royal Astronomical Society*, 491(3):3192–3216.
- Hand D.J. and Yu K., 2001. Idiot’s bayesnot so stupid after all?, *International statistical review*, 69(3):385–398.

- Harry G.M., Collaboration L.S. et al., 2010. Advanced ligo: the next generation of gravitational wave detectors, *Classical and Quantum Gravity*, 27(8):084006.
- Haskell B. and Patruno A., 2017. Are gravitational waves spinning down psr j 1023+0038?, *Physical Review Letters*, 119(16):161103.
- Hastie T., Tibshirani R., and Friedman J., 2009. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Hawking S., 1971. Gravitationally collapsed objects of very low mass, *Monthly Notices of the Royal Astronomical Society*, 152(1):75–78.
- Hawking S., 1988. *A Brief History of Time*. Bantam Books. ISBN 9780553109535.
- Herner K. et al., 2017. Optical follow-up of gravitational wave triggers with decam, In *Journal of Physics. Conference Series*, volume 898. Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States).
- Herner K. et al., 2020. Optical follow-up of gravitational wave triggers with decam during the first two ligo/virgo observing runs, *arXiv preprint arXiv:2001.06551*.
- Hessels J.W. et al., 2006. A radio pulsar spinning at 716 hz, *Science*, 311(5769):1901–1904.
- Hewish A. et al., 1968. In *Pulsating Stars*, 5–9. Springer.
- Hinton G., Srivastava N., and Swersky K., 2012a. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent, *Coursera*, 14(8).
- Hinton G.E. et al., 2012b. Improving neural networks by preventing co-adaptation of feature detectors, *arXiv preprint arXiv:1207.0580*.
- Ho A.Y. et al., 2018. iptf archival search for fast optical transients, *The Astrophysical Journal Letters*, 854(1):L13.
- Ho T.K., 1995. Random decision forests, In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, 278–282. IEEE.
- Ho T.K., 2002. A data complexity analysis of comparative advantages of decision forest constructors, *Pattern Analysis & Applications*, 5(2):102–112.
- Hornoch K. and Kucakova H., 2018. Discovery of an Apparent Nova in M31, *The Astronomer’s Telegram*, 11904.

- Horowitz C. and Kadau K., 2009. Breaking strain of neutron star crust and gravitational waves, *Physical Review Letters*, 102(19):191102.
- Hotokezaka K. et al., 2013. Remnant massive neutron stars of binary neutron star mergers: Evolution process and gravitational waveform, *Physical Review D*, 88(4):044026.
- Hotokezaka K. et al., 2019. A hubble constant measurement from superluminal motion of the jet in gw170817, *Nature Astronomy*, 3(10):940–944.
- Howell S.B., 2006. *Handbook of CCD astronomy*, volume 5. Cambridge University Press.
- Huang S.J. et al., 2020. Science with the TianQin Observatory: Preliminary results on Galactic double white dwarf binaries, , 102(6):063021.
- Hughes S.A., 2000. Gravitational waves from inspiral into massive black holes, In *AIP Conference Proceedings*, volume 523, 75–81. American Institute of Physics.
- Hughes S.A., 2001. Gravitational waves from extreme mass ratio inspirals: Challenges in mapping the spacetime of massive, compact objects, *Classical and Quantum Gravity*, 18(19):4067.
- Hulse R.A. and Taylor J.H., 1975. Discovery of a pulsar in a binary system, *The Astrophysical Journal*, 195:L51–L53.
- Hwang S.K. and Kim W.Y., 2006. A novel approach to the fast computation of zernike moments, *Pattern Recognition*, 39(11):2065–2076.
- Ioka K. et al., 2017. Gw 150914-like black holes as galactic high-energy sources, *Monthly Notices of the Royal Astronomical Society*, 470(3):3332–3345.
- Ivezić Ž. et al., 2014. *Statistics, data mining, and machine learning in astronomy: a practical Python guide for the analysis of survey data*, volume 1. Princeton University Press.
- Izzo L. et al., 2019. Signatures of a jet cocoon in early spectra of a supernova associated with a γ -ray burst, *Nature*, 565(7739):324.
- Jackman J.A. et al., 2018. Ground-based detection of g star superflares with ngts, *Monthly Notices of the Royal Astronomical Society*, 477(4):4655–4664.
- Jacob J.C. et al., 2010. Montage: a grid portal and software toolkit for science-grade astronomical image mosaicking, *arXiv preprint arXiv:1005.4454*.

- Janiuk A. et al., 2017. On the possible gamma-ray burst–gravitational wave association in gw150914, *New Astronomy*, 51:7–14.
- Jespersen C.K. et al., 2020. An Unambiguous Separation of Gamma-Ray Bursts into Two Classes from Prompt Emission Alone, , 896(2):L20.
- Ji A.P. et al., 2016. R-process enrichment from a single event in an ancient dwarf galaxy, *Nature*, 531(7596):610.
- Jia P., Liu Q., and Sun Y., 2020. Detection and classification of astronomical targets with deep neural networks in wide-field small aperture telescopes, *The Astronomical Journal*, 159(5):212.
- John G.H. and Langley P., 2013. Estimating continuous distributions in bayesian classifiers, *arXiv preprint arXiv:1302.4964*.
- Kalogera V. and Baym G., 1996. The maximum mass of a neutron star, *The Astrophysical Journal Letters*, 470(1):L61.
- Kantor J., 2014. Transient alerts in lsst, In *The Third Hot-wiring the Transient Universe Workshop*, Wozniak, PR, Graham, MJ, Mahabal, AA, and Seaman, R., eds, 19–26.
- Karlik B. and Olgac A.V., 2011. Performance analysis of various activation functions in generalized mlp architectures of neural networks, *International Journal of Artificial Intelligence and Expert Systems*, 1(4):111–122.
- Kasen D. and Barnes J., 2019. Radioactive heating and late time kilonova light curves, *The Astrophysical Journal*, 876(2):128.
- Kasen D., Fernández R., and Metzger B.D., 2015. Kilonova light curves from the disc wind outflows of compact object mergers, *Monthly Notices of the Royal Astronomical Society*, 450(2):1777–1786.
- Kasen D. et al., 2017. Origin of the heavy elements in binary neutron-star mergers from a gravitational-wave event, *Nature*, 551(7678):80.
- Kasliwal M. et al., 2016. iptf search for an optical counterpart to gravitational-wave transient gw150914, *The Astrophysical Journal Letters*, 824(2):L24.
- Kasliwal M.M. et al., 2020. Kilonova luminosity function constraints based on zwicky transient facility searches for 13 neutron star mergers, *arXiv preprint arXiv:2006.11306*.

- Kaspi V.M. and Beloborodov A.M., 2017. Magnetars, *Annual Review of Astronomy and Astrophysics*, 55:261–301.
- Kawaguchi K., Shibata M., and Tanaka M., 2020. Diversity of kilonova light curves, *The Astrophysical Journal*, 889(2):171.
- Khan F.M. et al., 2016. Swift coalescence of supermassive black holes in cosmological mergers of massive galaxies, *The Astrophysical Journal*, 828(2):73.
- Killestein T. et al., in prep. Transient-optimised real-bogus classification with bayesian convolutional neural networks sifting the goto candidate stream, *0*, 0(0):0.
- Kimura S.S., Murase K., and Mészáros P., 2017a. Outflow-driven transients from the birth of binary black holes. i. tidally locked secondary supernovae, *The Astrophysical Journal*, 851(1):52.
- Kimura S.S., Murase K., and Mészáros P., 2017b. Outflow-driven transients from the birth of binary black holes. ii. primary-induced accretion transients, *The Astrophysical Journal*, 851(1):53.
- Kingma D.P. and Ba J., 2014. Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*.
- Kiuchi K. et al., 2020. Sub-radian-accuracy gravitational waves from coalescing binary neutron stars in numerical relativity. ii. systematic study on the equation of state, binary mass, and mass ratio, *Physical Review D*, 101(8):084006.
- Kobayashi C., Karakas A.I., and Lugaro M., 2020. The origin of elements from carbon to uranium, *The Astrophysical Journal*, 900(2):179.
- Kóbori J., Bagoly Z., and Balázs L.G., 2020. Kilonova rates from spherical and axisymmetrical models, , 494(3):4343–4348.
- Kochanek C. et al., 2017. The all-sky automated survey for supernovae (asas-sn) light curve server v1. 0, *Publications of the Astronomical Society of the Pacific*, 129(980):104502.
- Komiya Y. and Shigeyama T., 2016. Contribution of neutron star mergers to the r-process chemical evolution in the hierarchical galaxy formation, *The Astrophysical Journal*, 830(2):76.

- Kramer M. and Champion D.J., 2013. The european pulsar timing array and the large european array for pulsars, *Classical and Quantum Gravity*, 30(22):224009.
- Kulkarni S., 2013. The intermediate palomar transient factory (iptf) begins, *The Astronomer's Telegram*, 4807.
- La Cour M., 2012. *Statistical view on FocusMax*. http://Statistical%20view%20on%20FocusMax_00.pdf. Online Article.
- Lamb G. et al., 2019. The optical afterglow of gw170817 at one year post-merger, *The Astrophysical Journal Letters*, 870(2):L15.
- Lamb G.P. and Kobayashi S., 2017. Electromagnetic counterparts to structured jets from gravitational wave detected mergers, *Monthly Notices of the Royal Astronomical Society*, 472(4):4953–4964.
- Lamberts A. et al., 2018. Predicting the binary black hole population of the milky way with cosmological simulations, *Monthly Notices of the Royal Astronomical Society*, 480(2):2704–2718.
- Lang D. et al., 2010. Astrometry. net: Blind astrometric calibration of arbitrary astronomical images, *The astronomical journal*, 139(5):1782.
- Lange J. et al., 2017. Parameter estimation method that directly compares gravitational wave observations to numerical relativity, *Physical Review D*, 96(10):104041.
- Langlois D., Maartens R., and Wands D., 2000. Gravitational waves from inflation on the brane, *Physics Letters B*, 489(3-4):259–267.
- Levan A. et al., 2017. The environment of the binary neutron star merger gw170817, *The Astrophysical Journal Letters*, 848(2):L28.
- Li X. et al., 2016. Implications of the tentative association between gw150914 and a fermi-gbm transient, *The Astrophysical Journal Letters*, 827(1):L16.
- Liebling S.L. and Palenzuela C., 2016. Electromagnetic luminosity of the coalescence of charged black hole binaries, *Physical Review D*, 94(6):064046.
- Ligo Collaboration and VIRGO Collaboration, 2019a. LIGO/Virgo S190924h: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25829:1.

- Ligo Collaboration and VIRGO Collaboration, 2019b. LIGO/Virgo S190930s: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25871:1.
- Ligo Collaboration and VIRGO Collaboration, 2019c. LIGO/Virgo S190930t: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25876:1.
- Ligo Collaboration and VIRGO Collaboration, 2019d. LIGO/Virgo S191129u: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26303:1.
- Ligo Collaboration and VIRGO Collaboration, 2019e. LIGO/Virgo S191205ah: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26386:1.
- Ligo Collaboration and VIRGO Collaboration, 2019f. LIGO/Virgo S191213g: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26399:1.
- Ligo Collaboration and VIRGO Collaboration, 2019g. LIGO/Virgo S191215w: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26441:1.
- Ligo Collaboration and VIRGO Collaboration, 2019h. LIGO/Virgo S191216ap: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26454:1.
- Ligo Collaboration and VIRGO Collaboration, 2019i. LIGO/Virgo S191222n: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26543:1.
- Ligo Collaboration and VIRGO Collaboration, 2020a. LIGO/Virgo S200105ae: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26688:1.
- Ligo Collaboration and VIRGO Collaboration, 2020b. LIGO/Virgo S200112r: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26715:1.
- Ligo Collaboration and VIRGO Collaboration, 2020c. LIGO/Virgo S200114f: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26734:1.

- Ligo Collaboration and VIRGO Collaboration, 2020d. LIGO/Virgo S200115j: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26759:1.
- Ligo Collaboration and VIRGO Collaboration, 2020e. LIGO/Virgo S200128d: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 26906:1.
- Ligo Collaboration and VIRGO Collaboration, 2020f. LIGO/Virgo S200208q: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27014:1.
- Ligo Collaboration and VIRGO Collaboration, 2020g. LIGO/Virgo S200213t: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27042:1.
- Ligo Collaboration and VIRGO Collaboration, 2020h. LIGO/Virgo S200219ac: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27130:1.
- Ligo Collaboration and VIRGO Collaboration, 2020i. LIGO/Virgo S200219ac: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27193:1.
- Ligo Collaboration and VIRGO Collaboration, 2020j. LIGO/Virgo S200219ac: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27278:1.
- Ligo Collaboration and VIRGO Collaboration, 2020k. LIGO/Virgo S200224ca: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 27184:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019a. LIGO/Virgo S190408an: Identification of a GW binary merger candidate., *GRB Coordinates Network*, 24069:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019b. LIGO/Virgo S190412m: Identification of a GW binary merger candidate., *GRB Coordinates Network*, 24098:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019c. LIGO/Virgo S190421ar: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24141:1.

- Ligo Scientific Collaboration and VIRGO Collaboration, 2019d. LIGO/Virgo S190425z: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24168:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019e. LIGO/Virgo S190426c: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24237:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019f. LIGO/Virgo S190510g: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24442:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019g. LIGO/Virgo S190512at: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24503:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019h. LIGO/Virgo S190513bm: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24522:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019i. LIGO/Virgo S190517h: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24570:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019j. LIGO/Virgo S190519bj: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24598:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019k. LIGO/Virgo S190521g: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24621:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019l. LIGO/Virgo S190521r: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24632:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019m. LIGO/Virgo S190630ag: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24922:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019n. LIGO/Virgo S190706ai: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 24998:1.

- Ligo Scientific Collaboration and VIRGO Collaboration, 2019o. LIGO/Virgo S190707q: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25012:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019p. LIGO/Virgo S190718y: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25087:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019q. LIGO/Virgo S190720a: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25115:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019r. LIGO/Virgo S190727h: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25164:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019s. LIGO/Virgo S190728q: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25187:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019t. LIGO/Virgo S190814bv: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25324:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019u. LIGO/Virgo S190828j: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25503:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019v. LIGO/Virgo S190828l: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25497:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019w. LIGO/Virgo S190901ap: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25606:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019x. LIGO/Virgo S190910d: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25695:1.
- Ligo Scientific Collaboration and VIRGO Collaboration, 2019y. LIGO/Virgo S190915ak: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25753:1.

- Ligo Scientific Collaboration and VIRGO Collaboration, 2019z. LIGO/Virgo S190923y: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25814:1.
- Lin Y. and Jeon Y., 2006. Random forests and adaptive nearest neighbors, *Journal of the American Statistical Association*, 101(474):578–590.
- Lintott C. and Reed J., 2013. In *Handbook of human computation*, 153–162. Springer.
- Lippuner J. et al., 2017. Signatures of hypermassive neutron star lifetimes on r-process nucleosynthesis in the disc ejecta from neutron star mergers, *Monthly Notices of the Royal Astronomical Society*, 472(1):904–918.
- Lipunov V. et al., 2016. First gravitational-wave burst gw150914: Master optical follow-up observations, *Monthly Notices of the Royal Astronomical Society*, 465(3):3656–3667.
- Lipunov V. et al., 2019. LIGO/Virgo S190901ap: MASTER OT J171434.95+280725.6/AT2019pju detection, *GRB Coordinates Network*, 25649:1.
- Liu T. et al., 2016. Fast radio bursts and their gamma-ray or radio afterglows as kerr–newman black hole binaries, *The Astrophysical Journal*, 826(1):82.
- Loeb A., 2016. Electromagnetic counterparts to black hole mergers detected by ligo, *The Astrophysical Journal Letters*, 819(2):L21.
- Lorimer D., 1996. Clocks in the cosmos, *Physics World*, 9(2):25.
- Loshchilov I. and Hutter F., 2017. Decoupled weight decay regularization, *arXiv preprint arXiv:1711.05101*.
- Lü H.J. et al., 2019. Diagnosing the remnants of binary neutron star merger from gw170817/grb170817a event, *Monthly Notices of the Royal Astronomical Society*, 486(4):4479–4484.
- Luna E. et al., 2007. Geometric method to measure astigmatism aberration at astronomical telescopes, *Applied optics*, 46(17):3439–3443.
- LVC, 2019a. Gravitational wave astronomy with ligo and similar detectors in the next decade, *arXiv preprint arXiv:1904.03187*.
- LVC, 2019b. LIGO/Virgo S190910h: Identification of a GW compact binary merger candidate., *GRB Coordinates Network*, 25707:1.

- Lyman J., 2018. Spalipy. <https://github.com/Lyalpha/spalipy>.
- Lyman J. et al., 2018. The optical afterglow of the short gamma-ray burst associated with gw170817, *Nature Astronomy*, 2(9):751.
- Maggiore M., 2000. Gravitational wave experiments and early universe cosmology, *Physics Reports*, 331(6):283–367.
- Makovoz D. and Marleau F.R., 2005. Point-source extraction with mopex, *Publications of the Astronomical Society of the Pacific*, 117(836):1113.
- Margalit B. and Metzger B.D., 2017. Constraining the maximum mass of neutron stars from multi-messenger observations of gw170817, *The Astrophysical Journal Letters*, 850(2):L19.
- Marsden B.G., 1986. Catalogue of cometary orbits, *IAUC*, 4168:2.
- Martynov D. et al., 2016. Sensitivity of the advanced ligo detectors at the beginning of gravitational wave astronomy, *Physical Review D*, 93(11):112004.
- Masci F.J. et al., 2018. The zwicky transient facility: Data processing, products, and archive, *Publications of the Astronomical Society of the Pacific*, 131(995):018003.
- Massey R. and Refregier A., 2005. Polar shapelets, *Monthly Notices of the Royal Astronomical Society*, 363(1):197–210.
- McCulloch W.S. and Pitts W., 1943. A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, 5(4):115–133.
- McKernan B. et al., 2018. Constraining stellar-mass black hole mergers in agn disks detectable with ligo, *The Astrophysical Journal*, 866(1):66.
- Meers B.J., 1988. Recycling in laser-interferometric gravitational-wave detectors, *Physical Review D*, 38(8):2317.
- Metzger B. et al., 2010. Electromagnetic counterparts of compact object mergers powered by the radioactive decay of r-process nuclei, *Monthly Notices of the Royal Astronomical Society*, 406(4):2650–2662.
- Metzger B.D. and Fernández R., 2014. Red or blue? a potential kilonova imprint of the delay until black hole formation following a neutron star merger, *Monthly Notices of the Royal Astronomical Society*, 441(4):3444–3453.

- Metzger B.D., Thompson T.A., and Quataert E., 2018. A magnetar origin for the kilonova ejecta in gw170817, *The Astrophysical Journal*, 856(2):101.
- Miller M.C., 2004. Introduction to neutron stars, *University of Maryland, College Park*.
- Moesta P. et al., 2012. On the detectability of dual jets from binary black holes, *The Astrophysical Journal Letters*, 749(2):L32.
- Momcheva I. and Tollerud E., 2015. Software use in astronomy: an informal survey, *arXiv preprint arXiv:1507.03989*.
- Mong Y.L. et al., 2020. Machine learning for transient recognition in difference imaging with minimum sampling effort, *arXiv preprint arXiv:2008.10178*.
- Montana G. et al., 2019. Constraining twin stars with gw170817, *Physical Review D*, 99(10):103009.
- Montgomery C., Orchiston W., and Whittingham I., 2009. Michell, laplace and the origin of the black hole concept, *Journal of Astronomical History and Heritage*, 12:90–96.
- Morii M. et al., 2016. Machine-learning selection of optical transients in the Subaru/hyper supprime-cam survey, *Publications of the Astronomical Society of Japan*, 68(6).
- Morokuma T. et al., 2014. Kiso supernova survey (kiss): Survey strategy, *Publications of the Astronomical Society of Japan*, 66(6):114.
- Morokuma T. et al., 2016. J-gem follow-up observations to search for an optical counterpart of the first gravitational wave source gw150914, *Publications of the Astronomical Society of Japan*, 68(4).
- Murase K. et al., 2016. Ultrafast outflows from black hole mergers with a minidisk, *The Astrophysical Journal Letters*, 822(1):L9.
- Murguia-Berthier A. et al., 2014. Necessary conditions for short gamma-ray burst production in binary neutron star mergers, *The Astrophysical Journal Letters*, 788(1):L8.
- Nagakura H. et al., 2014. Jet collimation in the ejecta of double neutron star mergers: a new canonical picture of short gamma-ray bursts, *The Astrophysical Journal Letters*, 784(2):L28.

- Nakar E. and Piran T., 2016. The observable signatures of grb cocoons, *The Astrophysical Journal*, 834(1):28.
- Nakar E. and Piran T., 2018. Implications of the radio and x-ray emission that followed gw170817, *Monthly Notices of the Royal Astronomical Society*, 478(1):407–415.
- Nativi L. et al., 2020. Can jets make the radioactively powered emission from neutron star mergers bluer?, .
- Newell E., 1983. Crowded-field stellar photometry, In *Astronomical Measuring Machines Workshop*, 15–40.
- Neyman J. and Pearson E.S., 1933. IX. on the problem of the most efficient tests of statistical hypotheses, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231(694–706):289–337.
- Noble S.C. et al., 2012. Circumbinary magnetohydrodynamic accretion into inspiraling binary black holes, *The Astrophysical Journal*, 755(1):51.
- Ofek E.O., 2014. MATLAB package for astronomy and astrophysics. Astrophysics Source Code Library.
- Ojima T. et al., 2018. Stochastic chemical evolution of galactic subhalos and the origin of r-process elements, *The Astrophysical Journal*, 865(2):87.
- Olech A. et al., 1999. Variable stars in the globular cluster m5: application of the image subtraction method, *Monthly Notices of the Royal Astronomical Society*, 310(3):759–764.
- Opitz D. and Maclin R., 1999. Popular ensemble methods: An empirical study, *Journal of artificial intelligence research*, 11:169–198.
- Oppenheimer J.R. and Volkoff G.M., 1939. On massive neutron cores, *Physical Review*, 55(4):374.
- Ott C.D., 2009. The gravitational-wave signature of core-collapse supernovae, *Classical and Quantum Gravity*, 26(6):063001.
- Pala A. et al., 2016. Effective temperatures of cataclysmic-variable white dwarfs as a probe of their evolution, *Monthly Notices of the Royal Astronomical Society*, 466(3):2855–2878.

- Palenzuela C., Lehner L., and Liebling S.L., 2010. Dual jets from binary black holes, *Science*, 329(5994):927–930.
- Palliyaguru N. et al., 2016. Radio follow-up of gravitational-wave triggers during advanced ligo o1, *The Astrophysical Journal Letters*, 829(2):L28.
- Pankow C. et al., 2018. Improvements in gravitational-wave sky localization with expanded networks of interferometers, *The Astrophysical Journal Letters*, 854(2):L25.
- Parr T. et al., 2018. Beware default random forest importances, *March*, 26:2018.
- Paschalidis V., Etienne Z.B., and Shapiro S.L., 2013. General-relativistic simulations of binary black hole-neutron stars: precursor electromagnetic signals, *Physical Review D*, 88(2):021504.
- Pedregosa F. et al., 2011. Scikit-learn: Machine learning in python, *the Journal of machine Learning research*, 12:2825–2830.
- Perego A., Radice D., and Bernuzzi S., 2017. AT 2017gfo: An Anisotropic and Three-component Kilonova Counterpart of GW170817, , 850(2):L37.
- Perna R., Lazzati D., and Giacomazzo B., 2016. Short gamma-ray bursts from the merger of two black holes, *The Astrophysical Journal Letters*, 821(1):L18.
- Perna R. et al., 2018. Binary black hole mergers within the ligo horizon: statistical properties and prospects for detecting electromagnetic counterparts, *Monthly Notices of the Royal Astronomical Society*, 477(3):4228–4240.
- Peters P.C., 1964. Gravitational radiation and the motion of two point masses, *Physical Review*, 136(4B):B1224.
- Piran T., Narayan R. et al., 1998. Spectra and light curves of gamma-ray burst afterglows, *The Astrophysical Journal Letters*, 497(1):L17.
- Piro L. et al., 2019. A long-lived neutron star merger remnant in gw170817: constraints and clues from x-ray observations, *Monthly Notices of the Royal Astronomical Society*, 483(2):1912–1921.
- Pogson N., 1856. Magnitudes of thirty-six of the minor planets for the first day of each month of the year 1857, *Monthly Notices of the Royal Astronomical Society*, 17:12–15.

- Pollacco D.L. et al., 2006. The WASP Project and the SuperWASP Cameras, , 118(848):1407–1418.
- Postnov K.A. and Yungelson L.R., 2014. The evolution of compact binary star systems, *Living Reviews in Relativity*, 17(1):3.
- Punturo M. et al., 2010. The einstein telescope: a third-generation gravitational wave observatory, *Classical and Quantum Gravity*, 27(19):194002.
- Qian N., 1999. On the momentum term in gradient descent learning algorithms, *Neural networks*, 12(1):145–151.
- Radice D. et al., 2018. Gw170817: joint constraint on the neutron star equation of state from multimessenger observations, *The Astrophysical Journal Letters*, 852(2):L29.
- Raghavan D. et al., 2010. A survey of stellar families: multiplicity of solar-type stars, *The Astrophysical Journal Supplement Series*, 190(1):1.
- Rakhmanov M. et al., 2004. Characterization of the ligo 4 km fabry–perot cavities via their high-frequency dynamic responses to length and laser frequency variations, *Classical and Quantum Gravity*, 21(5):S487.
- Rao P.B., Prasad D.V., and Kumar C.P., 2013. Feature extraction using zernike moments, *International Journal of Latest Trends in Engineering and Technology*, 2(2):228–234.
- Ratzloff J.K. et al., 2019. Building the evryscope: hardware design and performance, *Publications of the Astronomical Society of the Pacific*, 131(1001):075001.
- Regimbau T., Meacher D., and Coughlin M., 2014. Second einstein telescope mock science challenge: Detection of the gravitational-wave stochastic background from compact binary coalescences, *Physical Review D*, 89(8):084046.
- Rezzolla L., Most E.R., and Weih L.R., 2018. Using gravitational-wave observations and quasi-universal relations to constrain the maximum mass of neutron stars, *The Astrophysical Journal Letters*, 852(2):L25.
- Rezzolla L. et al., 2011. The missing link: Merging neutron stars naturally produce jet-like structures and can power short gamma-ray bursts, *The Astrophysical Journal Letters*, 732(1):L6.

- Rish I. et al., 2001. An empirical study of the naive bayes classifier, In *IJCAI 2001 workshop on empirical methods in artificial intelligence*, volume 3, 41–46.
- Robitaille T., Deil C., and Ginsburg A., 2018. Image reprojection (resampling).
- Robitaille T.P. et al., 2013. Astropy: A community python package for astronomy, *Astronomy & Astrophysics*, 558:A33.
- Rokach L., 2010. Ensemble-based classifiers, *Artificial Intelligence Review*, 33(1-2):1–39.
- Romano R.A., Aragon C.R., and Ding C., 2006. Supernova recognition using support vector machines, In *2006 5th International Conference on Machine Learning and Applications (ICMLA'06)*, 77–82. IEEE.
- Rosswog S., Freiburghaus C., and Thielemann F.K., 2000. Nucleosynthesis calculations for the ejecta of neutron star coalescences, *arXiv preprint astro-ph/0012046*.
- Rothschild R. et al., 1974. Millisecond temporal structure in cygnus x-1, *The Astrophysical Journal*, 189:L13.
- Rowlinson A. et al., 2013. Signatures of magnetar central engines in short grb light curves, *Monthly Notices of the Royal Astronomical Society*, 430(2):1061–1087.
- Ruder S., 2016. An overview of gradient descent optimization algorithms, *arXiv preprint arXiv:1609.04747*.
- Ruiz M., Shapiro S.L., and Tsokaros A., 2018. Gw170817, general relativistic magnetohydrodynamic simulations, and the neutron star maximum mass, *Physical Review D*, 97(2):021501.
- Rumelhart D.E., Hinton G.E., and Williams R.J., 1986. Learning representations by back-propagating errors, *nature*, 323(6088):533–536.
- Ryden B., 2017. *Introduction to cosmology*. Cambridge University Press.
- Sako S. et al., 2012. Kwfc: four square degrees camera for the kiso schmidt telescope, In *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446, 84466L. International Society for Optics and Photonics.
- Sánchez B. et al., 2018. Machine learning on difference image analysis: A comparison of methods for transient detection, *arXiv preprint arXiv:1812.10518*.

- Savage R.S. and Oliver S., 2007. Bayesian methods of astronomical source extraction, *The Astrophysical Journal*, 661(2):1339.
- Savchenko V. et al., 2016. Integral upper limits on gamma-ray emission associated with the gravitational wave event gw150914, *The Astrophysical Journal Letters*, 820(2):L36.
- Savchenko V. et al., 2017. Integral detection of the first prompt gamma-ray signal coincident with the gravitational-wave event gw170817, *The Astrophysical Journal Letters*, 848(2):L15.
- Schaffer S., 1979. John mitchell and black holes, *Journal for the History of Astronomy*, 10:42.
- Scornet E., 2016. Random forests and kernel methods, *IEEE Transactions on Information Theory*, 62(3):1485–1500.
- Seeds M. and Backman D., 2009. *Astronomy: The Solar System and Beyond*. Cengage Learning. ISBN 9780495562030.
- Segretain L., Chabrier G., and Mochkovitch R., 1997. The fate of merging white dwarfs, *The Astrophysical Journal*, 481(1):355.
- Segura A. et al., 2010. The effect of a strong stellar flare on the atmospheric chemistry of an earth-like planet orbiting an m dwarf, *Astrobiology*, 10(7):751–771.
- Shappee B. et al., 2017. Early spectra of the gravitational wave source gw170817: Evolution of a neutron star merger, *Science*, 358(6370):1574–1578.
- Shibata M. et al., 2006. Magnetized hypermassive neutron-star collapse: a central engine for short gamma-ray bursts, *Physical Review Letters*, 96(3):031102.
- Shibata M. et al., 2017. Modeling gw170817 based on numerical relativity and its implications, *Physical Review D*, 96(12):123012.
- Shibata M. et al., 2019. Constraint on the maximum mass of neutron stars using gw170817 event, *Physical Review D*, 100(2):023015.
- Sibi P., Jones S.A., and Siddarth P., 2013. Analysis of different activation functions using back propagation neural networks, *Journal of theoretical and applied information technology*, 47(3):1264–1268.
- Siegel D.M., Barnes J., and Metzger B.D., 2019. Collapsars as a major source of r-process elements, *Nature*, 569(7755):241.

- Singer L.P. and Price L.R., 2016. Rapid bayesian position reconstruction for gravitational-wave transients, *Physical Review D*, 93(2):024013.
- Singer L.P. et al., 2014. The first two years of electromagnetic follow-up with advanced ligo and virgo, *The Astrophysical Journal*, 795(2):105.
- Smith J.O., 2011. *Spectral Audio Signal Processing*. <http://ccrma.stanford.edu/~jos/sasp/>. Online book, 2011 edition.
- Sokol J., 2019. Debate intensifies over speed of expanding universe, *Science*, 07–20.
- Somiya K., 2012. Detector configuration of kagra—the japanese cryogenic gravitational-wave detector, *Classical and Quantum Gravity*, 29(12):124007.
- Spackman K.A., 1989. Signal detection theory: Valuable tools for evaluating inductive learning, In *Proceedings of the sixth international workshop on Machine learning*, 160–163. Elsevier.
- Srinivasan G., 2002. The maximum mass of neutron stars, *The Astronomy and Astrophysics Review*, 11(1):67–96.
- Staff J.E. et al., 2012. Spindown of isolated neutron stars: gravitational waves or magnetic braking?, *The Astrophysical Journal*, 751(1):24.
- Stairs I., Lyne A., and Shemar S., 2000. Evidence for free precession in a pulsar, *Nature*, 406(6795):484.
- Steeghs D. et al., 2019a. LIGO/Virgo S190412m: GOTO optical coverage - no notable counterparts, *GRB Coordinates Network*, 24116:1–+.
- Steeghs D. et al., 2019b. LIGO/Virgo S190425z: GOTO observations, *GRB Coordinates Network*, 24224:1–+.
- Steeghs D. et al., 2019c. LIGO/Virgo S190426c: GOTO optical coverage - no notable counterparts, *GRB Coordinates Network*, 24291:1–+.
- Steeghs D. et al., in prep. The gravitational-wave optical transient observer(goto): prototype performance and prospects for transient science, *0*, 0(0):0.
- Stone N.C., Metzger B.D., and Haiman Z., 2017. Assisted inspirals of stellar mass black holes embedded in agn discs: solving the final au problem, *Monthly Notices of the Royal Astronomical Society*, 464(1):946–954.

- Strobl C. and Zeileis A., 2008. Danger: High power!—exploring the statistical properties of a test for random forest variable importance.
- Sun R.Y. and Zhao C.Y., 2013. A new source extraction algorithm for optical space debris observation, *Research in Astronomy and Astrophysics*, 13(5):604.
- Tanaka M., 2016. Kilonova/macronova emission from compact binary mergers, *Advances in Astronomy*, 2016.
- Tananbaum H. et al., 1972. Observation of a correlated x-ray transition in cygnus x-1, *The Astrophysical Journal*, 177:L5.
- Tanvir N. et al., 2013. A kilonova associated with the short-duration γ -ray burst grb 130603b, *Nature*, 500(7464):547.
- Tauris T. and Van Den Heuvel E., 2006. Formation and evolution of compact stellar x-ray sources, *csxs*, 39:623–665.
- Tavani M. et al., 2016. Agile observations of the gravitational-wave event gw150914, *The Astrophysical Journal Letters*, 825(1):L4.
- Taylor J.H. and Weisberg J.M., 1989. Further experimental tests of relativistic gravity using the binary pulsar psr 1913+ 16, *The Astrophysical Journal*, 345:434–450.
- Taylor S. et al., 2017. All correlations must die: Assessing the significance of a stochastic gravitational-wave background in pulsar timing arrays, *Physical Review D*, 95(4):042002.
- Tewes M., 2013. Alipy: a python implementation of source alignment.
- Tinsley B.M., 1968. Evolution of the stars and gas in galaxies, *The Astrophysical Journal*, 151:547.
- Tomaney A.B. and Crofts A.P., 1996. Expanding the realm of microlensing surveys with difference image photometry, *arXiv preprint astro-ph/9610066*.
- Tonry J. et al., 2018. Atlas: a high-cadence all-sky survey system, *Publications of the Astronomical Society of the Pacific*, 130(988):064505.
- Tonry J.L. et al., 2018. The ATLAS All-Sky Stellar Reference Catalog, , 867(2):105.
- Toomer G. et al., 1984. *Almagest*, volume 13. Springer-Verlag.

- Troja E. et al., 2008. Different progenitors of short hard gamma-ray bursts, *Monthly Notices of the Royal Astronomical Society: Letters*, 385(1):L10–L14.
- Tumer K. and Ghosh J., 1996. Error correlation and error reduction in ensemble classifiers, *Connection science*, 8(3-4):385–404.
- Tutukov A. and Yungelson L., 1994. Merging of binary white dwarfs, neutron stars and black holes under the influence of gravitational wave radiation, *Monthly Notices of the Royal Astronomical Society*, 268(4):871–879.
- Uchiyama T. et al., 2012. Reduction of thermal fluctuations in a cryogenic laser interferometric gravitational wave detector, *Physical review letters*, 108(14):141101.
- van de Voort F. et al., 2014. Galactic r-process enrichment by neutron star mergers in cosmological simulations of a milky way-mass galaxy, *Monthly Notices of the Royal Astronomical Society*, 447(1):140–148.
- Veitch J. et al., 2015. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library, *Physical Review D*, 91(4):042003.
- Virtanen P. et al., 2020. Scipy 1.0: fundamental algorithms for scientific computing in python, *Nature methods*, 1–12.
- Vreeswijk P., 2017. Zogy. <https://github.com/pmvreeswijk/ZOGY>.
- Wall J.V. and Jenkins C.R., 2012. *Practical statistics for astronomers*. Cambridge University Press.
- Weber J., 1968. Gravitational-wave-detector events, *Physical Review Letters*, 20(23):1307.
- Weber J., 1969. Evidence for discovery of gravitational radiation, *Physical Review Letters*, 22(24):1320.
- Weber L. and Brady S., 2001. Fast auto-focus method and software for ccd-based telescopes, In *Minor Planet Amateur/Professional Workshop*, 104–113.
- Webster B.L. and Murdin P., 1972. Cygnus x-1a spectroscopic binary with a heavy companion?, *Nature*, 235(5332):37.
- Weisberg J. and Taylor J., 2005. Binary radio pulsars, *Rasio and IH Stairs, eds., San Francisco: Astronomical Society of the Pacific*, 25–31.

- Wells D.C. and Greisen E.W., 1979. Fits-a flexible image transport system, In *Image Processing in Astronomy*, 445.
- Wheatley P.J. et al., 2017. The next generation transit survey (ngts), *Monthly Notices of the Royal Astronomical Society*, 475(4):4476–4493.
- White D., 2014. Electromagnetic follow-up of gravitational wave candidates, *PhD Thesis*.
- Wolszczan A. and Frail D.A., 1992. A planetary system around the millisecond pulsar psr1257+ 12, *Nature*, 355(6356):145.
- Woosley S., 2016. The progenitor of gw150914, *The Astrophysical Journal Letters*, 824(1):L10.
- Woosley S. et al., 1994. The r-process and neutrino-heated supernova ejecta, *The Astrophysical Journal*, 433:229–246.
- Wright D. et al., 2015. Machine learning for transient discovery in pan-starrs1 difference imaging, *Monthly Notices of the Royal Astronomical Society*, 449(1):451–466.
- Wu X. et al., 2008. Top 10 algorithms in data mining, *Knowledge and information systems*, 14(1):1–37.
- Wu Y. and MacFadyen A., 2019. Gw170817 afterglow reveals that short gamma-ray bursts are neutron star mergers, *The Astrophysical Journal Letters*, 880(2):L23.
- Yi T. et al., 2018. Compact binary merger and kilonova: outflows from remnant disc, *Monthly Notices of the Royal Astronomical Society*, 476(1):683–689.
- Yin Q. et al., 2017. Blind deconvolution for astronomical spectrum extraction from two-dimensional multifiber spectrum images, *Optics express*, 25(5):5133–5145.
- Yuan W. et al., 2019. Consistent calibration of the tip of the red giant branch in the large magellanic cloud on the hubble space telescope photometric system and a redetermination of the hubble constant, *The Astrophysical Journal*, 886(1):61.
- Zackay B. and Ofek E.O., 2015a. How to coadd images? i. optimal source detection and photometry using ensembles of images.
- Zackay B. and Ofek E.O., 2015b. How to coadd images? ii. a coaddition image that is optimal for any purpose in the background dominated noise limit.

- Zackay B., Ofek E.O., and Gal-Yam A., 2016. Proper image subtraction optimal transient detection, photometry, and hypothesis testing, *The Astrophysical Journal*, 830(1):27.
- Zapartas E. et al., 2019. The diverse lives of progenitors of hydrogen-rich core-collapse supernovae: the role of binary interaction, *arXiv preprint arXiv:1907.06687*.
- Zel'dovich Y.B. and Novikov I., 1967. The hypothesis of cores retarded during expansion and the hot cosmological model, *Soviet Astronomy*, 10:602.
- Zernike F., 1934. Diffraction theory of the knife-edge test and its improved form, the phase-contrast method, *Monthly Notices of the Royal Astronomical Society*, 94:377–384.
- Zhang B., 2016. Mergers of charged black holes: Gravitational-wave events, short gamma-ray bursts, and fast radio bursts, *The Astrophysical Journal Letters*, 827(2):L31.
- Zhang Z. et al., 2017. Normalized direction-preserving adam, *arXiv preprint arXiv:1709.04546*.