UNIVERSITY OF SOUTHAMPTON

The Rate of Cosmic Thermonuclear Explosions in the Local Universe

by

Christopher M Frohmaier

A thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy

in the
Faculty of Physical Sciences and Engineering
Department of Physics and Astronomy

November 2017

UNIVERSITY OF SOUTHAMPTON

Abstract

Department of Physics and Astronomy

Doctor of Philosophy

The Rate of Cosmic Thermonuclear Explosions in the Local Universe

by Christopher M Frohmaier

This thesis investigates the volumetric rates of thermonuclear supernovae (SNe) in the Palomar Transient Factory (PTF). SN rates are a measure of how frequently stellar explosions occur as a function of cosmological volumes, or host galaxy properties. SNe are powerful cosmological probes; understanding their rates offers insight into the progenitors to the explosion and the astrophysics of galactic chemical evolution.

The Palomar Transient Factory was an automated optical sky survey designed for transient discovery. It spectroscopically confirmed $\sim$1900 SNe during the period of 2009–2012. PTF operated with a 3–5 day cadence, and scanned more than 8,000 square degrees of the sky. I quantified the performance of PTF through large scale simulations of transient events. Firstly, $\sim$7 $\times$ 10^6 fake transient events were inserted into real observational images. These 'fakes' were designed to test the real-time transient discovery pipeline. The images were treated identically to a new PTF observation, where the fakes were either recovered or not.

Multidimensional grids were created to describe how a transient would be recovered as a function of the fake's brightness and observing conditions. I found that bright fakes ($m_R < 18.5$) were recovered with $\sim$97% efficiency. PTF was 50% complete at $m_R = 20.3$. The recovery efficiency was also strongly dependent on: the limiting magnitude, the image quality, the sky background, and the immediate environment brightness.

The second stage of quantifying the performance of PTF was transient specific. Hundreds of millions of SNe Ia light curves were simulated on an artificial night sky. The simulations shared the statistical properties of the single epoch efficiencies. Through a Monte-Carlo simulation of the SNe Ia populations, I derived recovery efficiencies as a function of SNe Ia light curve parameters.

A sample of 90 SNe Ia ($z \leq 0.09$), were compared to the simulation recovery efficiencies. This provided the probability of the SNe passing rigorous quality cuts, and was used as a weighting factor. The weighted objects were summed and the volumetric SNe Ia rate was found to be 2.43 $\left(^{+0.29}_{-0.29} \text{ stat} \right) \left(^{+0.33}_{-0.19} \text{ sys} \right) \times 10^{-5}$ SNe Ia yr^{-1} Mpc^{-3} h_{70}^3.

I fit a simple delay-time model, $\Psi = \Psi_1 t^{-\beta}$, to the data and found $\beta \sim 1$.

I applied the same methodology to a newly discovered class of SN: Ca-rich SNe. Their volumetric rate was found to be 1.62($\pm$1.07)$\times 10^{-5}$ SNe yr^{-1} Mpc^{-3} h_{70}^3. Furthermore, I used model nucleosynthetic yields for Ca-Rich, SNe Ia and CCSNe models and found that a Ca-rich rate $\sim$30 $-$ 50% of the SNe Ia rate would explain the observed Ca/Fe intra-cluster medium abundances.

Contents

List of Figures

List of Tables

Abbreviations

5DC	**5 D**ay **C**adence Experiment
ACS	**A**dvanced **C**amera for **S**urveys
AGN	**A**ctive **G**alactic **N**uclei
CCD	**C**harged **C**oupled **D**evice
CMB	**C**osmic **M**icrowave **B**ackground
CSM	**C**ircum**S**tellar **M**edium
DBSCAN	**D**ensity-**B**ased **S**patial **C**lusters of **A**pplications with **N**oise
Dec.	**Dec**lination
DEIMOS	**DE**ep **I**maging **M**ulti-**O**bject **S**pectrograph
DES	**D**ark **E**nergy **S**urvey
DTD	**D**elay-**T**ime **D**istribution
DyC	**Dy**namic **C**adence Experiment
ESO	**E**uropean **S**outhern **O**bservatory
FoV	**F**ield **O**f **V**iew
HPC	**H**igh **P**erformance **C**omputing
HST	**H**ubble **S**pace **T**elescope
ICM	**I**ntra-**C**luster **M**edium
IME	**I**ntermediate **M**ass **E**lement
iPTF	**i**ntermediate **P**alomar **T**ransient **F**actory
ISIS	**I**ntermediate dispersion **S**pectrograph and **I**mage **S**ystem
JD	**J**ulian **D**ate
JLA	**J**oint **L**ight curve **A**nalysis
KBOs	**K**uiper **B**elt **O**bjects
ΛCDM	Λ (Cosmological constant) **C**old **D**ark **M**atter
LCOGT	**L**as **C**umbres **O**bservatory **G**lobal **T**elescope network

LOSS	**L**ick **O**bservatory **S**upernova **S**earch
LT	**L**iverpool **T**elescope
MCMC	**M**arkov **C**hain **M**onte **C**arlo
MW	**M**ilky **W**ay
NERSC	**N**ational **E**nergy **R**esearch **S**cientific **C**omputing Centre
P48	**P**alomar Samuel Oschin **48** inch telescope
P60	**P**alomar **60** inch telescope
PostGIS	**Post G**eographic **I**nformation **S**ystem
PostgreSQL	**Post S**tructured **Q**uery **L**anguage
PSF	**P**oint **S**pread **F**unction
PTF	**P**alomar **T**ransient **F**actory
R.A.	**R**ight **A**scension
RB	**R**eal-**B**ogus (score)
SALT	**S**pectral **A**daptive **L**ight Curve **T**emplate
SDSS	**S**loan **D**igital **S**ky **S**urvey
SFH	**S**tar **F**ormation **H**istory
SLSNe	**S**uper**L**uminous **S**uper**N**ova**e**
SN(e)	**S**uper**N**ova**e**
SNLS	**S**uper**N**ova **L**egacy **S**urvey
SQL	**S**tructured **Q**uery **L**anguage
TDE(s)	**T**idal **D**isruption **E**vent(s)
ToO	**T**arget **o**f **O**pportunity
USNO	**U**nited **S**tates **N**aval **O**bservatory
WCS	**W**orld **C**oordinate **S**ystem
WD	**W**hite **D**warf
WHT	**W**illiam **H**erschel **T**elescope
WISeREP	**W**eizmann **I**nteractive **S**upernova data **Rep**ository
VLT	**V**ery **L**arge **T**elescope
ZTF	**Z**wicky **T**ransient **F**actory

Declaration of Authorship

I, CHRISTOPHER FROHMAIER, declare that this thesis titled, 'THE RATE OF COSMIC THERMONUCLEAR EXPLOSIONS IN THE LOCAL UNIVERSE' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.

- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

- Parts of this thesis have been published in Frohmaier et al. (2017).

Signed:

_____

Date:

_____

Acknowledgements

I will be forever grateful to my supervisor, Mark Sullivan, for the opportunities he has given me to live out my dreams. I have always left our meetings more optimistic, more enthusiastic, and more confident about the road ahead. We chose to work on a difficult project together and, with your guidance, I always knew we'd achieve our goal. I will always be proud to say I was your student.

To the Southampton Supernova Group. You have been the best of friends and the best of colleagues. To Szymon, Regis, Chris, Cosimo, Natasha, Claudia, Miika, Charlotte, Florence, and Mike, thank you for the science chat, the beers, the laughs, the observing, the football and the friendship. We are a great team and I hope we will work together for years to come. Also, thank you to Peter Nugent for hosting me at UC Berkeley, and for all your support too.

To Mat, you were my second supervisor in all but the paperwork. I am a better scientist because of you, and I owe you so much. Rob and Georgios, we have climbed mountains together, looked at exploding stars together, and scored some cracking goals together. Without you two accompanying me throughout my PhD, it wouldn't have been as special.

To Rebecca, you took a special time in my life and made it better. The experience of running down a sand dune, kayaking across the seas, and simply having a lazy morning, is so much better when you have someone to share it with. Writing a thesis is not easy, but you have always been there to support and care for me.

To my officemates who made it a pleasure to come into work everyday. James, Juan, Stew and Judith, I will also consider us a team. "Fake it until you make it"; we did.

I want to say a special thank you to my Grandads. I never said it, but I always tried to live up to the examples you set. I hope that I can display the same respect for knowledge and pursuit of learning as you did. I hope that I can be as caring, gentle and happy-go-lucky as the role-model you are to me. If in any way I can be called a success, it's because I learned from the best of both worlds.

Finally, to my Mum, Dad, Emma, and Nicola, thank you for everything these past 27 years. I know it hasn't always been easy, especially trying to take me to school some days. But, in a way, I've finished the ultimate school, so you never have to worry again! Wherever life takes me, I always know I'll have a loving home to come back to.

To whom it may concern...

Chapter 1

Introduction

"They say it all started with a big bang. But, what I wonder is, was it a big bang or did it just seem big because there was no other noise to drown it out at the time?"

—Karl Pilkington

Supernovae (SNe) are extremely bright stellar explosions, with luminosities capable of temporarily outshining their entire host galaxy. This allows them to be seen out to vast cosmic distances, and far back into time. The term "super-nova" was first coined by Baade and Zwicky in 1934 when new Cepheid distances to their host galaxies enabled them to fully marvel at the luminosity of these transient events (Baade & Zwicky, 1934). The modern view of supernovae is filled with diverse classes, capturing a broad range of extreme explosion physics. Not only are they the ultimate environment to study these physical regimes, but certain types can also be used as cosmological tools to probe the universe's expansion.

Supernovae are fleeting in nature, and are typically only visible for a few weeks. Their discoveries were previously extremely rare, with decades passing between naked-eye-visible events. It was, once again, Zwicky who led the field, and established "SN Patrols" at the Palomar Observatory in California. The team regularly observed local galaxies from the 18- and eventually the 48-inch Schmidt telescope and the rate of SN discovery increased dramatically. Zwicky himself discovered more than 120 SNe; a record that stood for many years. It was with the advent of robotic telescopes, and soon after the adoption of CCDs, that really accelerated the discovery rate of SNe and our knowledge of the universe.

The last decade has seen a revolution in the study of the optical sky in the time domain. Several large-area 'rolling searches' – for example, Pan-STARRS 1 (Kaiser et al.,

2010), the Catalina Real-Time Transient Survey (Drake et al., 2009), the La Silla Quest Variability Survey (Baltay et al., 2013), and the Palomar Transient Factory (PTF[1]; Rau et al., 2009) – have repeatedly surveyed the sky on time-scales from minutes to hours, days and years. These surveys, together with dedicated spectroscopic follow-up programs (e.g., Smartt et al., 2015), have discovered thousands of galactic and extra-galactic astrophysical transients each year, filling in new and previously unexplored regions of the time-domain phase space (Figure 1.1).

It was with the increased statistical significance of high-redshift samples that SNe really changed our cosmic perspective. A special class of thermonuclear explosion, called a Type Ia supernova (SN Ia), were noted to all display remarkably similar features. The light curve shape and colour of these SNe allowed their intrinsic brightness to be calibrated with high precision. They became known as 'standard candles' and were used as cosmological probes. Famously, the High-Z Supernova Search Team and the Supernova Cosmology Project utilised their standardisable nature, and inferred the existence of "dark energy" (Riess et al., 1998; Perlmutter et al., 1999); a force that opposes gravity and causes the accelerating expansion of the universe. SNe Ia still remain our best tool to measure cosmological parameters.

1.1 Supernova Classification

The diversity of objects held under the catch-all term 'supernova' is large. Indeed, the explosion mechanism, timescale, and luminosity can differ drastically from object-to-object and thus requires sub-categories of SNe. As a first distinction, SNe can be split into two main groups: thermonuclear SNe and core-collapse SNe (CCSNe). A spectrum reveals features in the SN light that is used to categorise the transient into a sub-class (a review of SN optical spectra can be found in Filippenko, 1997). Figure 1.2 shows a decision tree for determining the SN type. If the spectrum does not show hydrogen lines then the SN is categorised as a Type I, and if it does, it's a Type II.

1.1.1 Type I Supernovae

The SNe Type I can be further broken down based on the presence of SiII λ6347,6371 (6355) absorption, which is blue-shifted from the expansion of the explosion ejecta, colloquially referred to as the '6150Å' feature. If this is present, then the supernova is classified as a SNe Ia. If it is not then subsequent spectral identification is made on the presence of (or lack of) optical helium lines (Harkness et al., 1987). Helium is seen in

---

[1] http://www.ptf.caltech.edu/
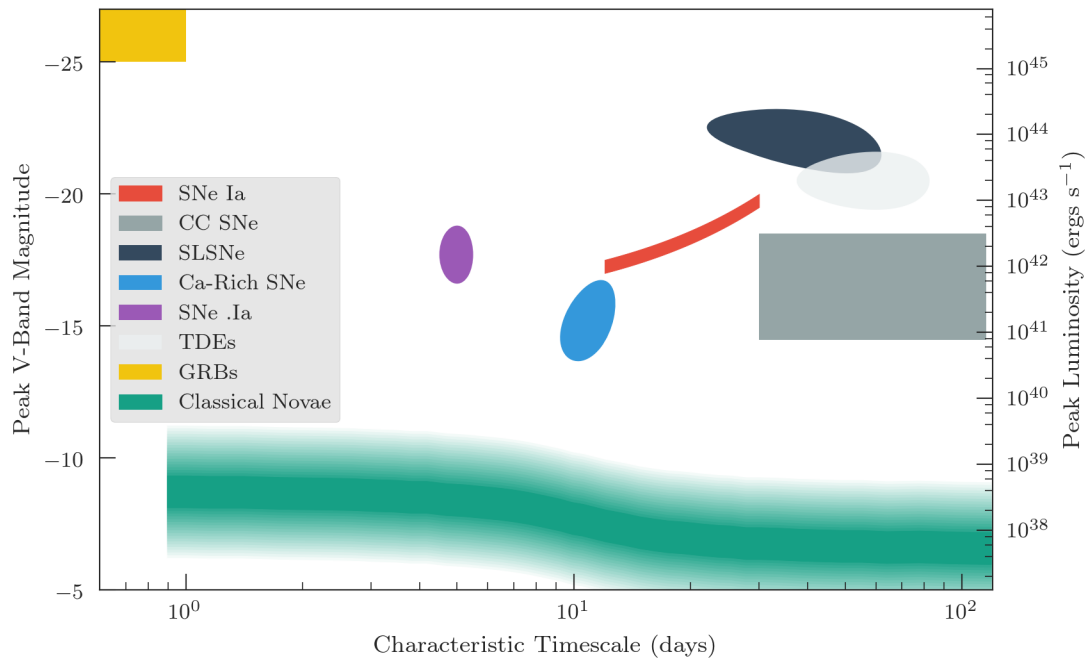
FIGURE 1.1: An illustration of the Luminosity - Timescale phase space adapted from similar representations by Kasliwal 2012, Nugent et al. 2015 and Cenko 2017. The characteristic timescale of a transient is defined as the approximate number of days for the object to fade from peak by 2 magnitudes. In this example we use the V-band magnitude but a representation in R of predicted transient classes can be seen in Rau et al. 2009. PTF's quest was to fill in the domain of this phase space with new discoveries of transients between the luminosity of classical novae and supernovae, a difference of $\sim 10^3$ in luminosity, a tasked dubbed as 'Bridging-the-Gap'. The regions depicted are not defining boundaries of a transient class, they simply represent where the bulk of the population are expected to fall. The low scatter and correlation of peak luminosity with timescale for the SNe Ia population demonstrates their standardisable nature. The GRB population extends out to much brighter luminosities (especially if the afterglow is viewed on-axis) and to faster timescales. Finally, the Classical Novae class are more diverse (Shara et al., 2017) than the standard approximation I have shown.

the SN spectra of Ib, but not Ic, with both SN sub-types believed to have Wolf-Rayet progenitors (Begelman & Sarazin, 1986). SNe Ib and Ic are often difficult to distinguish from one another due to potentially a transitional population of 'weak-helium' SNe Ib (Modjaz et al., 2014; Liu et al., 2016). For this reason they are often grouped together as SNe Ibc.

1.1.2 Type II Supernovae

Type II supernovae are the final explosive acts of stars $\geq 8 M_\odot$. Their spectra show hydrogen because they retained a significant fraction of their hydrogen envelopes prior to the explosion. Phenomenological properties of their light curves subdivide this class
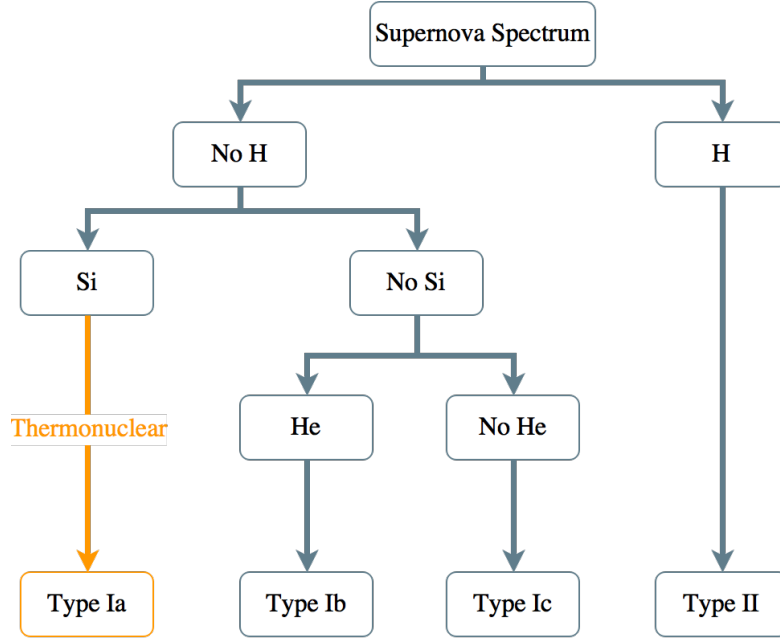
FIGURE 1.2: A decision tree for determining the classification of a SN from the emission/absorption features seen in the spectrum.

into IIP and IIL. SNe IIP show a plateau in their light curve as the thick hydrogen envelope around the progenitor is energised by the explosion (Faran et al., 2014). The plateau lasts upwards of 100 days and, once all the hydrogen has recombined, the light curve rapidly declines. SNe IIL on the other hand show a rapid linear decline in their light curves (Barbon et al., 1979). They are thought to originate from progenitors that have undergone a significant period of mass-loss.

Further to splitting the Type II populations by their light curves, additional sub-divisions can be made on their spectral features. For example SNe IIn show interaction with a circumstellar medium (CSM) via narrow Balmer emission. Indeed, the 'n' doesn't relate to a distinct explosion mechanism. Any CCSN can appear as a IIn (or Ibn (Moriya & Maeda, 2016)) because the phenomena of CSM interaction is external. Delving deeper into the classification quagmire, Type IIb spectra can initially appear identical to a normal Type II, but the H emission weakens and disappears in simultaneity with the occurrence of He. An overview of the progenitors to CCSNe can be found in Smartt (2009).

Li et al. (2011c) used data from the Lick Observatory Supernova Search (LOSS) to calculate the population fractions of the different SN sub-types. These population fractions are shown in Figure 1.3 and are derived from the observation of 175 SNe within 60 Mpc. For a volume-limited sample the largest fraction are the Type II at 57%. SNe Ia constitute just under a quarter of SNe, and the rest are the SNe Ibc. Within the sub-classes
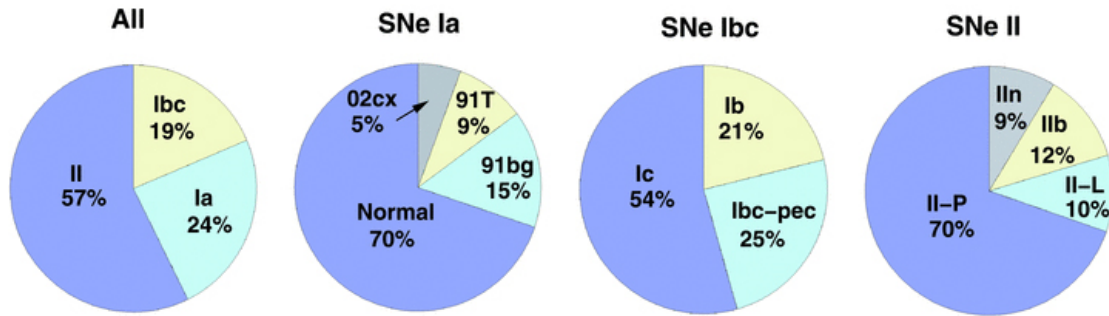
FIGURE 1.3: Pie-charts showing the population fraction of the different SN sub-classes. This figure was taken from (Li et al., 2011c).

of SN Ibc, a recent reanalysis of the LOSS data (Shivvers et al., 2017) has cast doubt on the population numbers. Owing to misclassification of the objects, they claim Ib are actually ∼1.7× more common than Ic.

The peak absolute magnitudes across the sub-types also vary considerably. Richardson et al. (2014) collected available data for as many SNe as they could find, predominantly from the Asiago Supernova Catalog (Barbon et al., 1989). They calculated the B-band absolute magnitudes (M_B) at peak brightness. For their volume limited sample they found the SNe Ia were ∼0.7 mag brighter than the brightest CCSN type and at least 1.3 mag brighter for non-interacting CCSNe. Table 1.1 shows their results for the absolute B-band magnitude of the different supernova types.

SN	M_B	N
Ia	-19.26	171
Ib	-17.54	18
Ic	-17.67	36
IIb	-17.03	15
IIL	-17.98	17
IIP	-16.80	74
IIn	-18.62	21

TABLE 1.1: The absolute B-band magnitudes for the different SN sub-type determined by Richardson et al. (2014). The column 'N' is the number of SNe in each class.

1.1.3 Superluminous Supernovae

A recent addition to the supernova family are Superluminous Supernovae (SLSNe). They are also split into the Type I and II categories (Gal-Yam, 2012). As their name suggests, they are extremely bright, with the SLSNe-I objects having an absolute B-band magnitude ≤-21 (Quimby et al., 2011). The current favoured energy source is the spin-down of a highly magnetic compact object - a magnetar (Kasen & Bildsten, 2010; Woosley, 2010). Spectrally, the objects appear as a blue continuum at early-times with a

black-body temperature $\gtrsim$15000K. As the object cools, the spectrum begins to resemble an SN Ic (Inserra et al., 2013). These events are rare; $\sim$1/1,000th the core-collapse rate (Prajs et al., 2017).

1.2 Thermonuclear Supernovae

The term 'thermonuclear supernova' is given to the class of SNe from which the explosive energy is derived from exothermic thermonuclear runaway of a white dwarf (WD). In this process, the 'fuel' is a lower binding energy species (e.g. ^{12}C, ^{16}O) that transmutes into a higher binding energy 'ash' (e.g. ^{28}Si, ^{56}Ni), under conditions of both high temperature and density. The energy released from this fusion process is sufficient to unbind the WD and hence produce an explosion (Hillebrandt & Niemeyer, 2000). This is in contrast to CCSNe, which are predominantly powered by the liberation of gravitational binding energy and driven by a colossal release of neutrinos (Woosley & Janka, 2005).

The observational results from searches of SNe largely supported claims that only SNe Ia were thermonuclear in origin. However, the recent growth of SN surveys that probed new timescales and amassed large populations of transients has identified new objects potentially at the extremes of the thermonuclear regime (Kasliwal, 2012).

I will introduce two thermonuclear families that enter into the analyses of this work. The first are the well-observed SNe Ia, and the other are newly discovered, so-called, Calcium-Rich SNe.

1.2.1 SNe Ia

SNe Ia result from the explosion of a carbon-oxygen white dwarf (CO WD) on the approach to the Chandrasekhar Mass, $M_{ch} \sim 1.4 M_{\odot}$. This mass limit is the point at which electron degeneracy pressure can no longer support the compact object. The density in the core of the WD asymptotes to infinity on this approach, and carbon is ignited. This sets off the thermonuclear runaway and, thus, the explosion.

In addition to the standard Type I classification features, SNe Ia are characterised by lines of Si, Fe, and Ca at maximum light. An annotated spectrum of a SNe Ia is shown in Figure 1.4. The expansion velocities of this ejecta, typically 10^4 km s^{-1} (Benetti et al., 2005), cause these lines to appear both blue-shifted and Doppler broadened. The mean rise-time of the light curve (shown in Figure 1.5) to maximum light is 18.98 days (Firth et al., 2015). However, the light curve itself is not powered by the explosion energy. Rather, it is the light emitted from an optically thick ejecta, reprocessed via inverse-β
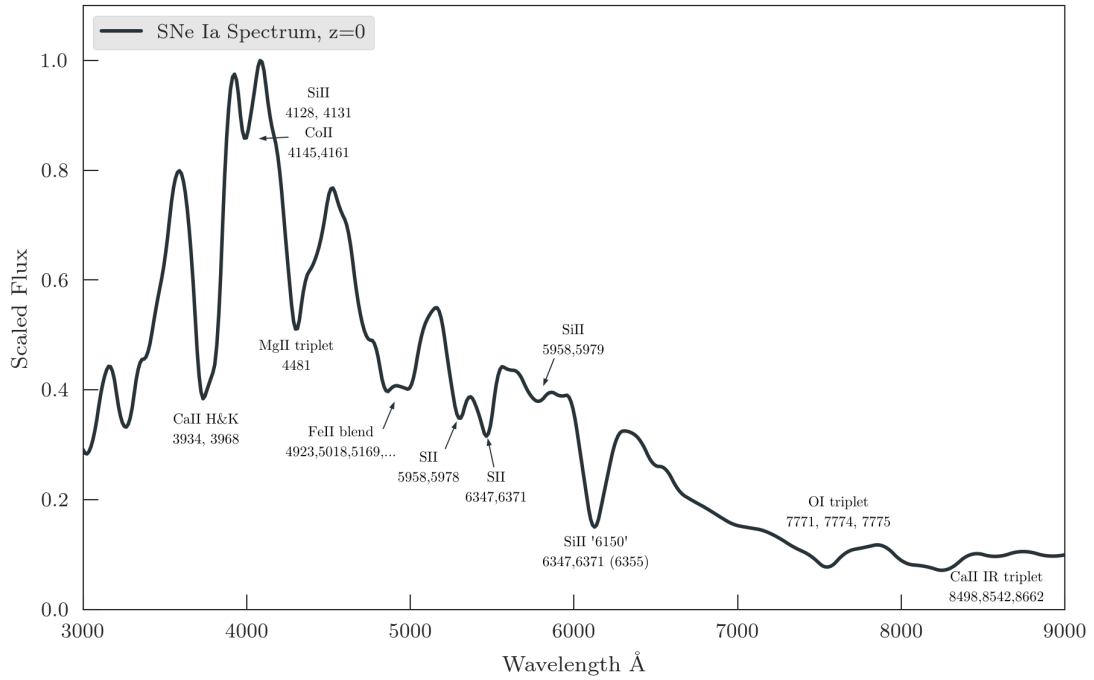
FIGURE 1.4: The spectrum of a SNe Ia at maximum brightness with the most distinctive lines labelled. The expansion velocities of the ejecta cause a blue-shifting of the absorption features.

and electron-capture of γ-rays and positrons (Colgate & McKee, 1969). This energy deposition is provided by the radioactive decay chain of ^{56}Ni

$$^{56}\mathrm{Ni} \xrightarrow{\mathrm{t_{1/2}} = 6.08 \text{ d}} {}^{56}\mathrm{Co} \xrightarrow{\mathrm{t_{1/2}} = 77.2 \text{ d}} {}^{56}\mathrm{Fe} \tag{1.1}$$
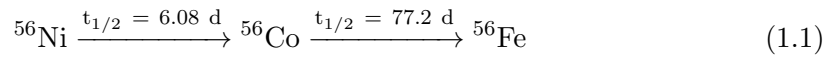
Therefore, the peak bolometric luminosity of SN Ia is governed by the mass of ^{56}Ni synthesised in the explosion. This is known as Arnett's Rule (Arnett, 1979, 1982).

Given the narrow mass range of ^{56}Ni production, there is consequently a narrow range of SNe Ia luminosities. Furthermore, it was first noticed by Phillips (1993) that the shape of the light curve is strongly correlated to the peak luminosity of the SN Ia. This gave rise to the 'width-luminosity' relation and was initially described through Δm_{15}, a measure of how many magnitudes the objects has faded by in the 15 days since peak (Phillips, 1993; Phillips et al., 1999). An example of a multi-colour light curve is shown in Figure 1.5.

Additional methods of describing the light curve shape have been parametrised through a 'stretch' (e.g. Perlmutter et al., 1997b; Conley et al., 2008) or an 'x$_1$' (Guy et al., 2007). Both the stretch and x$_1$ parameters consider the shape of the light curve on either side of peak, whereas, Δm_{15} is only measured on the decline. There also exists a colour
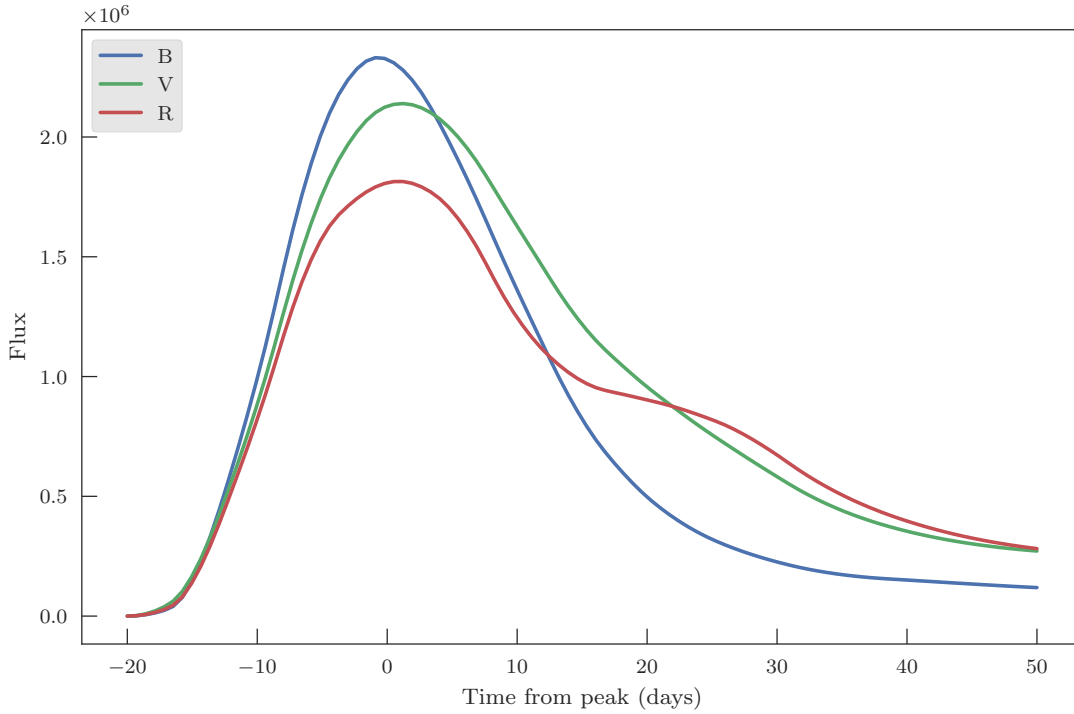
FIGURE 1.5: A multi-colour light curve of a template SNe Ia.

dependence to the luminosity (Riess et al., 1996). Therefore the total width-colour-luminosity relation is described as: brighter-broader, fainter-narrower, brighter-bluer, fainter-redder. There are sub-types within the SN Ia category that do not follow the 'normal' population and are named after the prototypical examples. For example, 91T-like SNe are over-luminous, and 91bg-like SNe are under-luminous.

Using correction techniques has allowed for precise calculations of the peak magnitudes of 'normal' SNe Ia, hence their description as 'standardisable candles'. Knowing the intrinsic brightness of an object, and comparing it to the apparent brightness allows for the distance to the object to be calculated. Both the standardisable nature of SNe Ia and their tremendous luminosities turn them into incredibly powerful tools to measure vast cosmic distances, and hence probe the fundamental nature of the universe (discussed in Section 1.3).

1.2.2 Ca-Rich SNe

Classification features in the spectra of SNe were traditionally defined at peak brightness. As both the sample sizes and follow up programs improved the coverage on any single object, the late-time diversity in SNe grew clearer. One such class of newly discovered objects are 'Ca-rich' transients (Perets et al., 2010; Kasliwal et al., 2012; Valenti

et al., 2014; Lunnan et al., 2017). From early observations they display a He-rich but H-free spectrum, which, following convention, would lead to a SN Ib classification. However, their evolution is rapid, with rise-times of 10–15 days. The spectra turns nebular extremely quickly and is dominated by [Ca II] $\lambda\lambda7291,7324$, Ca NIR triplet and [O I] $\lambda\lambda6300,6364$ emission (Figure 1.6). Indeed it is the large ratio of [Ca II] to [O I] that defines a Ca-rich object, as [O I] is much stronger in SNe Ib/c.
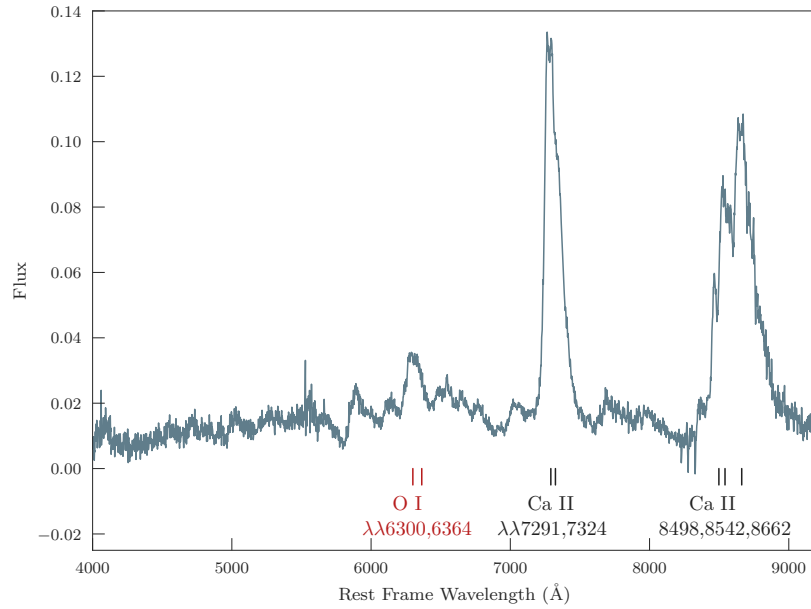
FIGURE 1.6: The nebular spectrum of a Ca-rich SN. The dominant features of the spectrum are the strong emission lines of [O I] and [Ca II]. This spectrum is of 2005E (Perets et al., 2010), an archetypal Ca-rich transient

Curiously, the literature sample of Ca-rich objects are only observed in the outskirts of E/S0 galaxies at $\sim$8–80 kpc from the centres, see Figure 1.7, perhaps even showing a preference for cluster environments (Lunnan et al., 2017). Such metal-poor, and potentially old, environments can satisfactorily rule out a CCSNe origin, although, Foley (2015) suggests the progenitor may have been 'kicked' from the host galaxy. It has also been speculated that the host may be a globular cluster (Yuan et al., 2013), but Lyman et al. (2014, 2016) have placed deep limits against this through observations from VLT and HST.

Ruling out a massive star origin leaves thermonuclear runaway as the remaining explosion mechanism. Collisionally triggered explosions in a binary system between a He donor and a CO or ONe WD have been speculated (García-Berro et al., 2017). Another mechanism involves a helium-shell detonation on the surface of a CO WD that does not trigger a core explosion itself (Waldman et al., 2011). Alternatively, the explosion may result from the tidally disrupted CO WD around an intermediate-mass black hole which
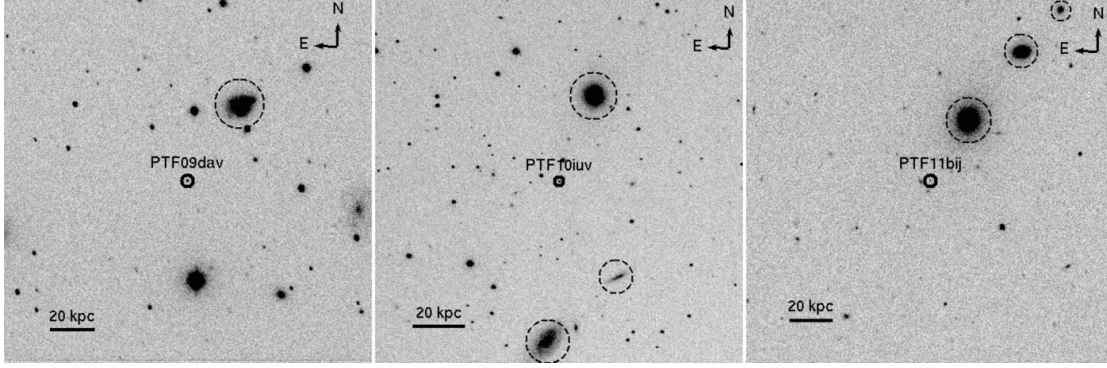
FIGURE 1.7: The remote locations of 3 Ca-rich transients (Kasliwal et al., 2012). It is currently unknown why these transients show a preference for remote locations.

would promote the globular cluster host conjecture (Rosswog et al., 2009; Metzger, 2012; Sell et al., 2015).

It is observed that the Ca to Fe ratio in the intra-cluster medium (de Plaa et al., 2007; Mernier et al., 2016) exhibits an over-abundance when contributions from the known rates of SNe Ia and CCSNe are considered. Furthermore, it has been suggested that faint thermonuclear SNe produce ^{44}Ti. This decays to ^{44}Ca, and may well be the source of Galactic positrons (Crocker et al., 2017). Ca-rich SNe are certainly faint and, as the name suggests, a rich source of Ca. Their observed locations ideally place them to spread Ca into the intra-cluster medium, as the ejecta do not need to overcome the gravitational potential of a host galaxy (Zaritsky et al., 2004). Mulchaey et al. (2014) showed that the observed Ca over-abundance could be explained by a contribution from these Ca-rich supernovae. However, there remain two points of contention: the rates of Ca-rich SNe are not known, and the remote locations may actually be a result of current surveys being unable to detect these faint-and-fast transient if they occur in the centres of galaxies.

1.3 SNe Ia as cosmological probes

Evidence for an accelerating universe was first provided by observations of high redshift SNe Ia in the late 20th century (Riess et al., 1998; Perlmutter et al., 1999). Presently, they still provide the tightest constraints on cosmological parameters (e.g. Astier et al., 2006; Sullivan et al., 2011a; Betoule et al., 2014). Combined with Cosmic Microwave Background (CMB) anisotropies from WMAP (Hinshaw et al., 2013) and Planck (Planck Collaboration et al., 2016), along with baryon acoustics peaks from large scale structure (e.g. Eisenstein et al., 2005), the overwhelming evidence points to dark energy dominating the current mass-energy density of the universe.

'Normal' SNe Ia (see Betoule et al., 2014, or Chapter 4.1 for details) have their luminosities corrected through both the width-luminosity and colour-luminosity relationships. One such method utilises the empirically derived model SALT/SALT2 (Guy et al., 2005, 2007) to perform the corrections to the absolute magnitude. The distance modulus, μ, is the difference between the observed apparent magnitude, m_{obs}, and the corrected absolute magnitude, and is found via

$$\mu = m_{obs} - (\mathcal{M} - \alpha x_1 + \beta C) \tag{1.2}$$

where $\mathcal{M}$, α and β are nuisance parameters found through fitting a cosmological sample of supernovae to minimise the scatter in the peak absolute magnitudes.
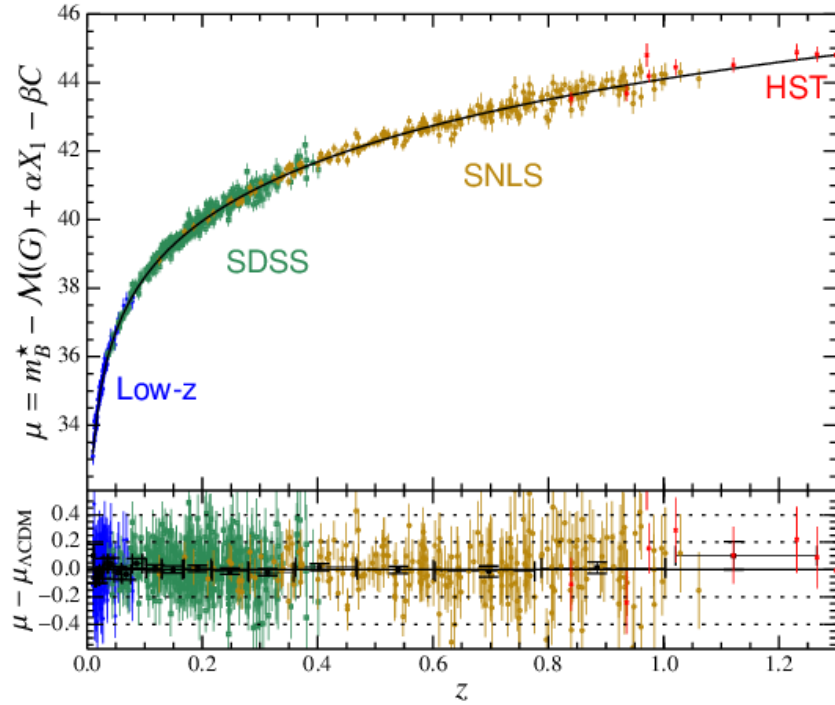
FIGURE 1.8: A Hubble diagram from Betoule et al. (2014). The cosmological model that best describes the data is shown by the black line. The Hubble residuals are shown in the bottom panel.

The distance modulus is compared to the redshift of the supernovae to produce a 'Hubble Diagram;, an example is shown in Figure 1.8. A theoretical distance modulus, μ_{theory}, constructed from a cosmological model, can be compared to the observed μ. The difference between them is called the 'Hubble residual'. Cosmological parameters are derived by minimising the Hubble residuals.

To minimise the Hubble residuals further, corrections to the SN measurements can be made based on the host galaxy properties. SNe which occur in star forming hosts are,

on average, broader and brighter than populations that occur in passive hosts (e.g. Hamuy et al., 1995; Sullivan et al., 2006; D'Andrea et al., 2011). Furthermore, Sullivan et al. (2010) found that SNe Ia of the same stretch and colour in the most massive (M> $10^{10}M_\odot$) and the lowest specific star formation rate (sSFR) galaxies were, on average, brighter by 0.08mag. This could possibly be explained by different progenitor scenarios.

Recently, Betoule et al. (2014) used SNe Ia from the SDSS-II SN Survey (Sako et al., 2014) and SNLS (Conley et al., 2011) and combined them under a join light curve analysis (JLA). For a flat ΛCDM cosmology, they found the matter density parameter to be $\Omega_m = 0.295 \pm 0.034$. Furthermore, using a combination of the JLA sample, CMB anisotropies, and BAO measurements, Betoule et al. (2014) were able to provide improved constrains on the wCDM cosmological model. This model assumes a constant dark energy equation of state parameter, w, to describe the pressure to energy density ratio of the universe. Their result of $w = -1.027 \pm 0.055$ is consistent with a description of the universe requiring a cosmological constant.

The Hubble parameter, H_0, is set by the current expansion rate of the universe and was first observed by Hubble (1929). Present day measurements of H_0 show a greater than 3σ tension between results from the CMB measurements of Planck Collaboration et al. (2016) ($H_0 = 67.3 \pm 1.2$ km s^{-1} Mpc^{-3}) to the SNe Ia sample of Riess et al. (2011, R11) ($H_0 = 73.8 \pm 2.4$ km s^{-1} Mpc^{-3}) and Riess et al. (2016) ($H_0 = 73.24 \pm 1.74$ km s^{-1} Mpc^{-3}). Even a blinded reanalysis of the R11 data by Zhang et al. (2017) could not resolve the tension ($H_0 = 72.5 \pm 3.1$(stat) ± 0.77(sys) km s^{-1} Mpc^{-3}). This is an unresolved problem in modern cosmology.

1.4   The Progenitors to SNe Ia

The primary task of any SNe Ia progenitor model is to grow the mass of a CO WD to approach or exceed M_{Ch}, and thus trigger carbon burning and a thermonuclear runaway. This is achieved through two proposed mechanisms. The first is the single degenerate (SD) scenario (Whelan & Iben, 1973) where the WD accretes material from a non-degenerate companion. The second method involves two WDs and is therefore dubbed the double degenerate (DD) scenario (Tutukov & Yungelson, 1981; Iben & Tutukov, 1984; Webbink, 1984).

SNe Ia are seen in all types of galaxies (e.g. Childress et al., 2013), suggesting that either a single progenitor scenario must act on a wide range of timescales, or the SN Ia population comes from multiple progenitors. Furthermore, SNe Ia light curve properties, such as

the width and expansion velocity, are seen to correlate with the galaxy morphology (Filippenko, 1989; Branch & van den Bergh, 1993) and mass or metallicity (Howell et al., 2009). It is observed that the brighter-broader SNe Ia prefer the spiral/irregular (late type) galaxies, whereas the fainter-narrower SNe Ia in the E/S0 (early type) galaxies. The evidence presented is highly suggestive that the age, or possibly the metallicity, of the progenitor is driving these differences.

Without direct observation of a progenitor, the details of the explosion mechanism may well remain a mystery. The somewhat uneasy reality that the precise origin of our best cosmological probe is unknown, is called 'The Progenitor Problem'. I will discuss the single- and double- degenerate scenarios here.

1.4.1   Single-degenerate

There are two principal methods to accrete material from a non-degenerate companion. Firstly, the companion can swell-up to fill its Roche-lobe, whereby material can be transferred into the gravitational potential of the WD. Secondly, a stellar wind from the companion can be blown towards, and retained by, the WD (Li & van den Heuvel, 1997). In both cases it is assumed that the accretion of material will increase the WD mass up to M_{Ch}, and cause the explosion. However, the largest theoretical challenge behind this is deciphering whether or not the WD will actually grow in mass.

The accretion rates necessary to increase the mass of a WD must occur within a narrow range, $\sim 10^{-8}$–$10^{-7} M_\odot \mathrm{yr}^{-1}$. As the WD accretes, it is possible for a thin layer of hydrogen to form on the surface of the WD. Stable burning of the hydrogen to helium can take place (Nomoto, 1982) and is observed as a super-soft X-ray source (van den Heuvel et al., 1992).

If the accretion rate is too low then the build-up of hydrogen will eventually erupt in its own thermonuclear runaway called a 'nova' (Starrfield et al., 1972). The energy liberated from hydrogen burning is greater than the energy required to escape the gravitational potential of the WD. Resultantly, more material is blown away than originally accreted.

In a regime where the accretion is too high, an optically thick wind can occur and stem the flow of further material (Hachisu et al., 1996). However, it has been suggested that this wind could self-regulate such that the rate of accreted material is maintained to a steady-state for stable nuclear burning.

The SD scenario naturally explains the homogeneity of SNe Ia luminosities. The growth of the WD must approach M_{Ch}, and thus the amount of radioactive fusion ash is approximately the same for each event. Hydrogen is expected to originate from an interaction

with the companion. From an observational perspective, this hydrogen would be seen either in very early observations, as the SN ejecta hits the companion (Kasen, 2010), or at very late times as a result of hydrogen swept-up from the companion. In a late-time study of 11 objects, Maguire et al. (2016) found tentative evidence of Hα emission in just 1 event. This finding casts doubt on the dominant contribution of SNe Ia originating from the SD scenario.

1.4.2 Double-degenerate

In the DD scenario, the WD binary system loses angular momentum via the radiation of gravitational waves. This causes an in-spiral and merger, with the combined mass approaching M_{Ch}. This leads to core ignition of carbon and an explosion.

In the case of an unequal mass binary, the more massive WD tidally disrupts and accretes the lower-mass WD. Although the accretion rate for this mechanism is sufficiently high to grow the WD mass, Nomoto & Iben (1985) speculate that an accretion rate that is too high will trigger an off centre carbon ignition. This doesn't explode completely, but produces oxygen and neon, any further accretion could result in a collapse to a neutron star (Saio & Nomoto, 1998; Shen et al., 2012). This accretion-induced-collapse may well lead to an observable transient event itself (Saio & Nomoto, 1998). As an alternative to WD coalescence, it is possible for the two degenerate objects to smash into each other. This is the so-called 'violent merger' scenario (Röpke et al., 2012; Pakmor et al., 2012; Ruiter et al., 2013; Pakmor et al., 2013). It is a further possibility that the combined mass of the binary is $>M_{Ch}$, resulting in a super-Chandrasekhar explosion (e.g. Howell et al., 2006).

From an observational perspective, the DD mechanism leaves few clues to its origin. Probably the most well constrained progenitor for a SNe Ia comes from observations of SN 2011fe (Nugent et al., 2011). Deep imaging at the location of the SN has ruled out Roche-lobe overflow from a red giant companion (Li et al., 2011b; Nugent et al., 2011; Horesh et al., 2012). The ESO SN Ia progenitor survey (SPY) (Napiwotzki et al., 2001), undertook a search of galactic WD binaries for systems that would merge in a Hubble time. Ultimately, they found no unambiguous candidates. However, Santander-Garcia et al. (2015) found a WD binary with an orbital period of $\sim$4 hours and a combined mass of $\sim$1.76$M_\odot$. This exceeds M_{Ch} and is predicted to merge within 700 million years and trigger a SNe Ia - confirming the DD hypothesis.

1.5 The SNe Ia Delay Time Distribution

The delay-time distribution (DTD) is a theoretical measure of the SN rate as a function of time that elapses between a hypothetical δ-function burst of star formation to the respective supernovae. Naturally, different progenitor scenarios require different DTDs. For SNe Ia, a bottle-neck common to all progenitors is the production rate of WDs. This process is determined by the time taken for an $8M_\odot$ star to evolve into a WD - stars more massive will explode as CCSNe and stars less massive have longer main-sequence life times. This sets the shortest delay at $\gtrsim 40$Myr.

Since the DTD is the SN response to the star formation, a theoretical rate of SNe Ia can be calculated by convolving a model delay-time with a star formation history (SFH).

$$\text{SNR}_{Ia}(t) = \int_0^t \text{SFH}(t - \tau)\Psi(\tau)d\tau \tag{1.3}$$

The SFH can either be as a function of cosmic time (producing theoretical volumetric rates), or for individual galaxies. Using the measured cosmic SFH (Li, 2008; Yüksel et al., 2008), it is possible to construct model delay-times, $\Psi(\tau)$, and compare the results to observed supernova rates.

1.5.1 Double-Degenerate Delay Times
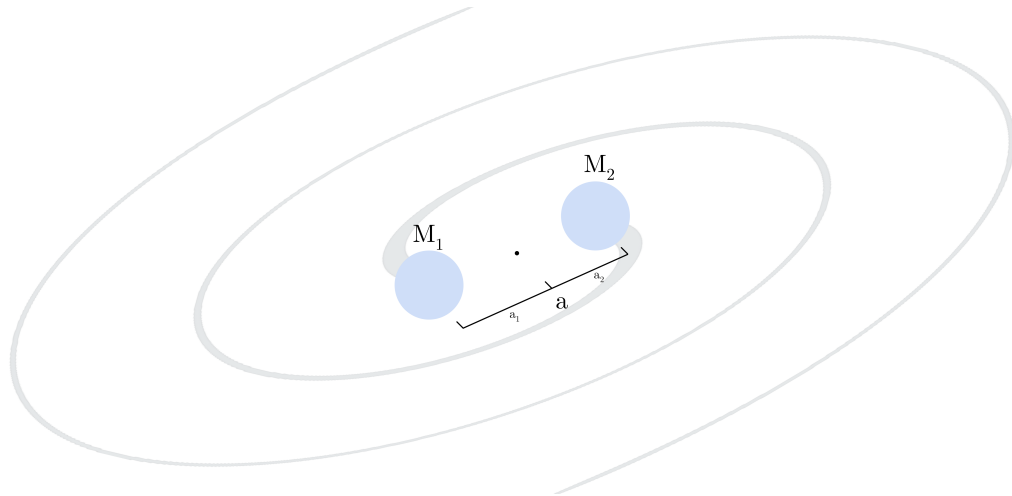
FIGURE 1.9: A cartoon for the double-degenerate progenitor system. The two WDs in the binary system are shown, separated by distance a. The binary loses angular momentum via the radiation of gravitational waves, shown by the grey spirals emanating from the system.

For the DD model the timescale is driven by the loss of angular momentum through the radiation of gravitational waves. The coalescence timescale for a WD binary can be described through Keplerian mechanics and a schematic of the system is shown in Figure 1.9. The binary system consists of two WDs of mass M_1 and M_2, orbiting around the centre of mass and separated by distance, $a = a_1 + a_2$. The reduced mass of the system is expressed as

$$\mu = \frac{M_1 M_2}{M}; \quad M = M_1 + M_2 \tag{1.4}$$

From Kepler's 3rd law, the angular orbital frequency, Ω, of the binary is related to the orbital separation of the system via

$$\Omega^2 = \frac{GM}{a^3} \tag{1.5}$$

where G is the universal gravitational constant. The total energy of the system (kinetic + gravitational potential), E, is

$$E = \frac{\Omega^2 \left(M_1 a_1^2 + M_2 a_2^2 \right)}{2} - \frac{GM_1 M_2}{a} \quad = \quad \frac{-G\mu M}{2a} \tag{1.6}$$

where Equation 1.5 was used to reduce the expression. As the binary system radiates gravitational energy it loses angular momentum. To conserve this momentum the orbital separation must decrease accordingly. The rate at which this energy is lost corresponds to the luminosity of the gravitational waves

$$\frac{dE}{dt} = \frac{dE}{da} \frac{da}{dt} \tag{1.7}$$

The gravitational wave luminosity, L_{GW} is found through the third order time derivative of the binary quadrupole tensor (see e.g. Kokkotas, 2002, for the derivation), and for a circular orbit is

$$L_{GW} = \frac{-dE}{dt} = \frac{32G^4 \mu^2 M^3}{5c^5 a^5} \tag{1.8}$$

The derivative of Equation 1.6 w.r.t a and Equation 1.8 can be combined to find the rate of change of the orbital separation.

$$\frac{da}{dt} = \left(\frac{dE}{da} \right)^{-1} \frac{dE}{dt} = \frac{-64G^3 \mu M^2}{5c^5 a^3} \tag{1.9}$$

The system coalesces from an initial separation a_0 to $a=0$ in time, τ, which is found by integrating Equation 1.9

$$\tau = \int_{t_0}^{t_f} dt = \int_{a_0}^{0} \frac{-5c^5 a^3}{64G^3 \mu M^2} = \frac{5c^5 a_0^4}{256G^3 \mu M^2} \tag{1.10}$$

$$\therefore \tau \propto a^4 \tag{1.11}$$

Now that the coalescence time as a function of the initial separation has been established, it is possible to calculate a theoretical SNe Ia rate if the orbital separation distribution of the progenitors is known. It was first observed by (Öpik, 1924) that the distribution of main-sequence binary orbital separations follows a power-law with index ϵ.

$$\frac{dN_{MS}}{da} \sim a^\epsilon \tag{1.12}$$

Under the assumption that after the stellar evolutionary and common envelope phases, the WD orbital separation follows the same power-law, then the time derivative of the number of SNe Ia is

$$\frac{dN_{Ia}}{dt} = \frac{dN_{MS}}{da} \frac{da}{dt} \quad \sim a^\epsilon \; a^{-3} \tag{1.13}$$

where the two contributing derivatives were taken from Equation 1.10 and Equation 1.9 respectively. Observations of the orbital separation of binary stars (Abt, 1983; Poveda et al., 2007) have found an $\epsilon \approx -1$. Indeed, widely used Binary Population Synthesis codes (e.g. Ruiter et al., 2009) also draw from a -1 power-law index for their simulations. Finally, we can use the observed value of ϵ and the relationship between coalescence time, t, and separation, a, (Equation 1.11) to calculate the DTD for a population of SNe Ia from a gravitational wave driven progenitor system to be

$$\frac{dN_{Ia}}{dt} \sim t^{\frac{\epsilon-3}{4}} \quad \sim t^{-1} \tag{1.14}$$

Following on from this result, the variable of interest is the power-law index. Multiple comparisons of volumetric SN Ia rates, and a $\Psi(t) \propto t^{-\beta}$ DTD, have been explored and found to be consistent with $\beta = 1$ (Dilday et al., 2010; Graur et al., 2011; Perrett et al., 2012; Graur & Maoz, 2013). This is an encouraging result for the DD scenario as the theory appears to match observations, although, Maoz & Mannucci (2012) show that some SD models can produce a $\sim t^{-1}$ DTD too.

Other DTD models propose a two channel progenitor population of SNe Ia with contributions from a 'prompt' and 'tardy' mechanism. For the prompt channel, it is speculated that the rates will follow proportionally to the CCSNe rate (Mannucci et al., 2005) or to the star formation rate (Sullivan et al., 2006). This would result in a typical delay-time of ∼100–500 Myr. The tardy channel would follow the stellar mass with delay-times around 5 Gyrs. However, the interpretation of stellar mass is open to some debate. Neill et al. (2006) argue that this is the total mass ever formed. Whereas, others consider corrections to the mass to take account of the stars that have died, which is more easily linked to current galaxy observables (Sullivan et al., 2006).

Of course, a theoretical DTD is only as strong as the observations that support it. The past two decades have seen the rapid development of CCD technology and computational infrastructure. This has enabled large synoptic surveys to observe huge volumes of the universe and discover thousands of transients. Detailed studies of SN rates provide powerful anchors to constrain SN delay times and help investigate progenitor scenarios.

1.6 Supernova rates and surveys
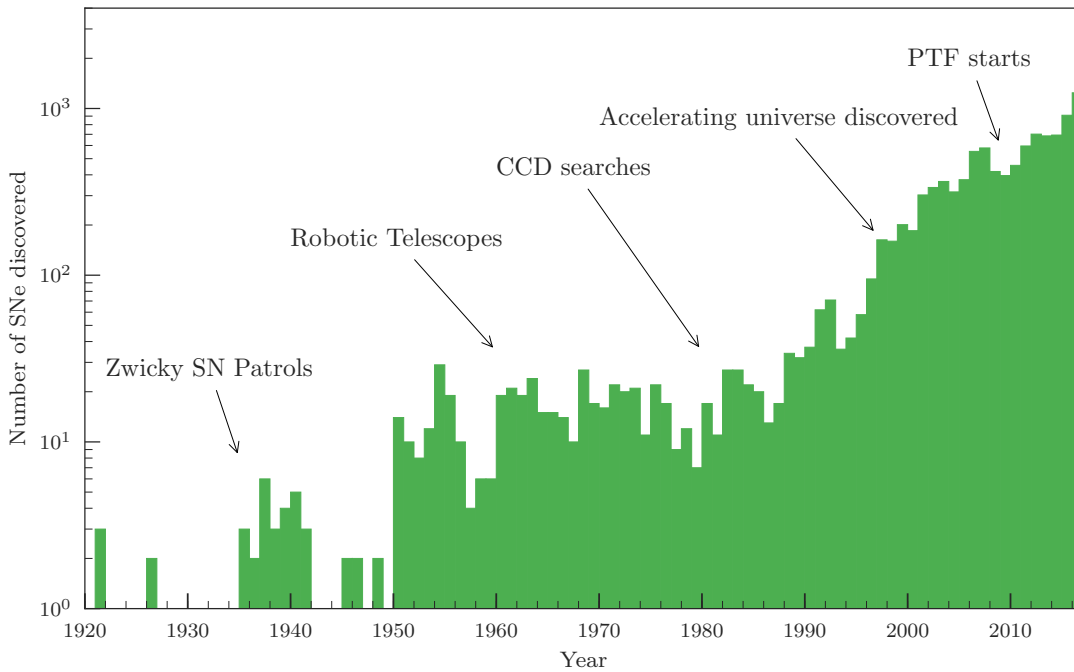
FIGURE 1.10: SN discoveries from 1920 to present day. Key events in the history of transient hunting are labelled. This data was compiled from on-line sources of SNe discoveries[2]

Supernova rates describe the number of supernovae that occur in a given timespan, either as a function of galaxy properties or within a comoving volume. Rate calculations are difficult: they require large samples of supernovae and a substantial survey infrastructure to produce reliable and contaminate free results. Low-redshift surveys have typically been galaxy-targeted, and hence biased in the SNe populations they find. Additionally, their rates are calculated as a function of the galaxy properties (i.e. mass, luminosity etc) and converted to volumetric rates via an assumed galaxy luminosity function. This potentially leaves the results open to uncertainties caused by preferentially sampling from the brighter end of the distribution and, again, biasing the type of supernovae being observed.

Historically, supernovae were discovered through a visual assessment of galaxies, either with the naked-eye in real-time (Evans et al., 1989), or with photographic plates. Consequently, the number of discoveries was small; of order a few SNe each year. Indeed, Cappellaro et al. (1997) combined five independent SN searches (4 photographic plate, 1 visual) from decades (1959-1994) of SN hunting efforts. Their rates sample consisted of just 110 suitable SNe across all types, and it's no surprise they claimed:

"Very few groups of professional astronomers have had the perseverance and force to carry out a SN search".

Times have changed, the rise of robotic telescopes, the wide adoption of CCDs, and the development of sophisticated image analysis software has vastly increased the discovery rate of supernovae. Figure 1.10 shows the number of SNe discovered each year from the start of the 20th century to present day.

The long-term (1998 March – 2008 December) SN hunt from the Lick Observatory Supernova Search (Leaman et al., 2011) was the first to take advantage of new CCD and image subtraction technological developments, and dedicate it to a SN survey. Over their 11 year survey they made 2.3 million observations of 14,882 galaxies and discovered 1036 SNe. This was the largest sample of SNe collected from a single survey, establishing the modern trend of time-domain astronomy.

1.6.1 Rate Calculation Methods

In this section I will discuss the two main methods of calculating volumetric rates. At the heart of all methodologies is a need to account for incompleteness in the survey. Ideally, the rate can be expressed as

---

[2]A history of supernova discovery (https://ishivvers.github.io/maps/sne.html),
IAU supernova list (http://www.cbat.eps.harvard.edu/lists/Supernovae.html),
the Transient Name Server (https://wis-tns.weizmann.ac.il),
and the Rochester Astronomy archives (http://www.rochesterastronomy.org/snimages/sndateall.html).

$$r(V) = \frac{N}{V\Delta T} \tag{1.15}$$

where V is the total surveyed volume, ΔT is the duration of the survey and N is the total number of SNe that exploded during the survey. Accounting for incompleteness is the trickiest part of any rate calculation, and is usually captured in a parameter known as the 'efficiency'. Understanding the efficiency with which surveys operate and detect objects is of paramount importance in understanding the astrophysics of the transient populations that they uncover. For every transient that is detected, it is important to know how many events, with the same properties, were not detected during the survey period. There are many reasons why transients can be missed or not detected by surveys, beyond simple Malmquist bias effects. For example, the observational cadence of the survey may be too long to detect rapidly evolving events; gaps in observing as a result of poor weather, seeing, or technical problems may occur; some parts of the survey area may be inaccessible due to saturated foreground stars, gaps between CCDs, or bad pixels; the detection sensitivity may change as a function of the lunar cycle or other variables; inefficiencies in the complex data reduction and transient detection pipelines may result in transients of any brightness being lost. All surveys will therefore make an inevitably incomplete sampling of the transient population, which will consequently impact the determination of transient volumetric rates, luminosity functions, the dependence of the transient on the underlying stellar populations, and, in the case of cosmological studies using supernovae, the measured cosmological parameters.

1.6.1.1   The Control-Time Method

The control-time method was first created by Zwicky (1942), improved by van den Bergh (1991), Cappellaro et al. (1993, 1997) and further refined by Leaman et al. (2011); Li et al. (2011c). The execution of this method is complex but detailed thoroughly in the literature (Cappellaro et al., 1997; Leaman et al., 2011). However, I will briefly discuss a qualitative overview of the principles.

The control-time method is attractive because it is relatively inexpensive to compute and doesn't require large scale simulations. In a galaxy targeted survey, the selected fields are returned to after a period of time to check if a SN had recently occurred. The distance to each galaxy is known and so light curves for each type of supernovae are constructed from assumed luminosity functions and translated into their apparent magnitudes. Observational effects, such as extinction, are often included on a galaxy-by-galaxy basis. The control-time is the time interval for each SN model, for which it is brighter than the limiting magnitude of each observation. For example, if your limiting

magnitude is 19 mag, and your SN model asserts anything that exploded within the past 12 days should be brighter than 19 mag, then your *effective* survey duration from one observation is 12 days.

This control-time is calculated for each observation, unique to each SN model type. The final rate is then

$$r_{\mathrm{SN}} = \frac{N_{SN}}{\sum_{i=1}^{N_{\mathrm{G}}} T_i} \qquad (1.16)$$

where N_{SN} is the number of SNe in the sample, N_G is the number of galaxies and T_i is the control-time for the ith galaxy. This produces rates as a function of the galaxy properties and are often transferred into volumetric rates.

1.6.1.2 The Monte-Carlo Method

The prevalence of high-performance computing (HPC) resources has lead to the adoption of a Monte-Carlo method of rate calculations. Studies that attempt this require large-scale simulations that can be computationally very expensive. They invariably work via the insertion of 'fake' transients into a survey imaging data stream, passing the adjusted data through the same survey detection pipeline as used to find real transients, and assessing the degree to which the fake transients can then be recovered. This can be done either 'offline' once a survey has been completed (e.g., Pain et al., 2002; Perrett et al., 2010), or in real-time while the survey is operational and the data being collected (e.g., Sako et al., 2008; Kessler et al., 2015). The fake events are usually designed to replicate the properties of the entire range of transients that might be detected, from their apparent magnitude to their host galaxy environment and local surface brightness.

Real transients that enter into the analysis are compared to simulated transients that share similar properties. The results of the simulations then provide a weighting to each real object. This accounts for objects with properties that make them difficult to find and hence are mostly missed. The rate equation is therefore

$$r_{\mathrm{V}}(z) = \frac{1}{\mathrm{V}\Delta\mathrm{T}} \sum_{i=1}^{N} \frac{1+z_i}{\epsilon_i} \qquad (1.17)$$

where $\Delta\mathrm{T}$ is the duration of the survey with a fixed comoving volume, V. ϵ_i and z_i are the efficiency of detecting and redshift of the ith SN respectively. Many of the recent and future sky surveys are untargeted, as they systematically scan the sky. This makes the control-time method inordinately complex when variable cadences, on wide-area

searches, are adopted. The Monte-Carlo method naturally handles these subtitles and is the method adopted in this thesis.

A novel approach to the Monte-Carlo method was performed by Prajs et al. (2017) to calculate the rate of SLSNe from SNLS. Their method adopted the relevant detection efficiencies of Perret et al. (2010), but didn't use them to re-weight the discovered objects. Instead, they assumed an underlying SLSNe rate and simulated light curves with the detection efficiencies applied accordingly. Their real SLSN sample consisted of 3 objects, therefore the assumed rates were checked to see how frequently 3 simulated objects were successfully 'observed'. This method naturally built up a probability density function for the underlying rate. This comes with the requirement to simulate orders of magnitude more light curves than the other Monte-Carlo methods, and was only achievable due to the intrinsically low rate of SLSNe and small survey area of SNLS - 4 square degrees, compared to PTF's $\sim 8,000$ square degrees.

1.6.2 The current state of SNe Ia Rates

SNe Ia rates have been calculated over a wide range of redshifts, thanks to their large luminosities. In Figure 1.11, the current best volumetric rates from the literature are shown. This is a non-exhaustive list and, in many cases, new SNe were added to older samples and the rates re-analysed years later. In such a circumstance, only the latest analysis from the survey is included.
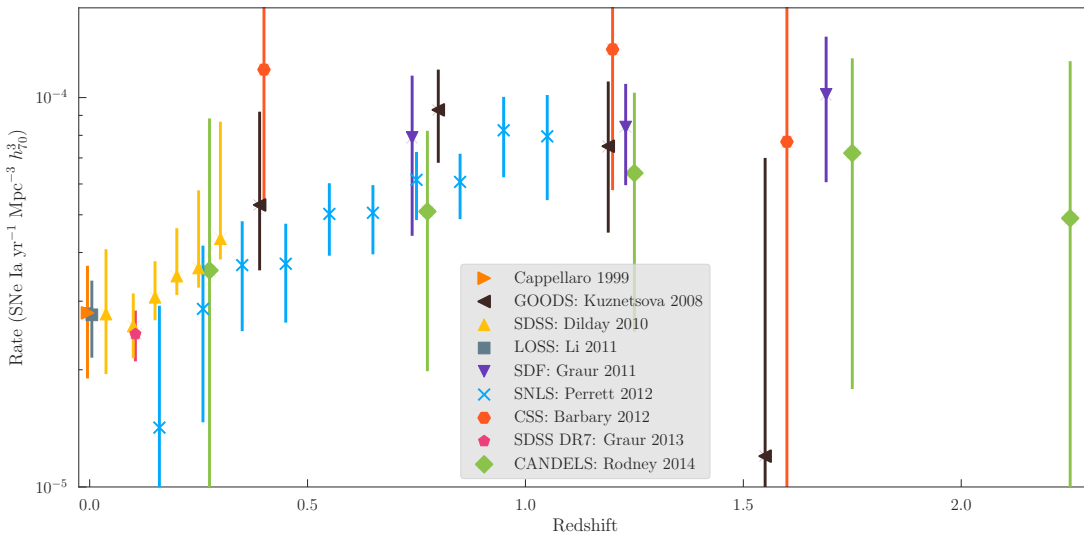
FIGURE 1.11: Literature rate calculations from several different surveys. Shown are Cappellaro et al. (1999); Kuznetsova et al. (2008); Dilday et al. (2010); Li et al. (2011a); Graur et al. (2011); Perret et al. (2012); Barbary et al. (2012); Graur & Maoz (2013); Rodney et al. (2014)

These calculations are invariably limited by the statistical samples of SNe. The low-redshift rates require long periods of time and large survey areas to build up significant numbers of objects in their catalogues. This is both expensive to maintain and, if the survey is un-targeted, produces huge volumes of data. This is why only galaxies were targeted in older surveys, so that the number of SNe discovered was maximised with the fewest observations necessary. The advantage of discovering low-redshift SNe Ia is that they are comparatively bright, which means spectroscopic classification is cheap.

Conversely, the high redshift surveys require long exposures to reach the necessary depths to even detect SNe. Therefore, the telescopes are large, or in orbit, and operated at huge expense. Furthermore, the SNe Ia discovered at higher redshifts are most often photometrically classified, which poses higher risks of contamination from non-SNe Ia. For example, Barbary et al. (2012) found an object consistent with both a SN Ibc at z~0.9 and a SN Ia at z~1.3. With fewer than 6 objects in that redshift bin, this degeneracy represents a significant source of uncertainty. However, time-dilatation increases the duration of high-redshift transients, meaning fewer observations are needed for a good rest-frame coverage.

The first high redshift rates (z>1) were calculated from data acquired by the Advanced Camera for Surveys (ACS) instrument on the Hubble Space Telescope (HST) (Strolger et al., 2004; Dahlen et al., 2004). Additional analysis and data was added to this Great Observatories Origins Deep Survey (GOODS) SNe Ia sample by Dahlen et al. (2007, 2008). It was claimed that the SNe Ia rate was consistent with being flat at z>0.5 and decreasing beyond z~1.2. If true, it is suggestive of a long delay-time. This was also supported by a different Bayesian-based analysis technique (Kuznetsova et al., 2008) on the same data. However, the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS) (Grogin et al., 2011; Koekemoer et al., 2011), also an HST survey, found rates a factor of 2 lower than similar HST rates at z=1.25. Furthermore, their results also show a steadily increasing rate to z~1.2. However, beyond this the rate is either flat or slowly declining (Rodney et al., 2014). All HST measurements of the rate used the control-time method.

These HST high-redshift rates overlap with the study of Poznanski et al. (2007) (later re-analysed by Graur et al., 2011) from the ground-based Subaru Deep Field (SDF) survey. Within the uncertainties of the measurement, the SDF rates are consistent with the HST results. However, the design of the SDF survey left it open to large biases and uncertainties, as each observation was acquired with an approximately one year cadence. Additionally, only one object received a spectroscopic confirmation (out of 28 SNe Ia), the rest were inferred from the light curve. The host redshifts were mostly photometric with ~15% spectroscopically observed. SDF offered an improvement over

HST rates as the samples sizes were significantly larger. The highest redshift SDF bin (z>1.5) contained 10 SNe Ia compared to ∼3 in the Dahlen et al. (2008) sample and ∼1 in the Barbary et al. (2012) analysis. The SDF rates are consistent with a flattening around z∼1 but are not consistent with the decline advocated by Dahlen et al. (2008); Kuznetsova et al. (2008).

The Supernova Legacy Survey (SNLS) reveals the trend of rates increasing from z=0 to z∼1 (Perrett et al., 2012). Contrary to the claims of Dahlen et al. (2008), there is no evidence that the rates flatten at z>0.5. The SNLS rate was supported by a large sample of SNe Ia (691 events), with one-third spectroscopically confirmed. The survey was also characterised through the Monte-Carlo efficiency method by simulating 2.5×10^6 artificial SNe in the observational fields. Due to the large extent of the redshift covered, 0.1≤z≤1.1, and relatively small uncertainties, DTD models were constrained more confidently. Combining their data with other surveys they found a simple power-law $\Psi(t) \propto t^{-\beta}$ with $\beta \sim 1$ can explain the entire SNe Ia rates sample. However, they also found a 'prompt' and 'tardy' model is able to explain the data equally well. To separate the two model better rates data is needed.

Graur & Maoz (2013) performed a rates analysis unlike any of the others reviewed here. Their data were not images, rather, contaminated SNe Ia light was searched for in the spectra of galaxies as part of SDSS DR7. They found 90 SNe Ia in the spectra of ∼700,000 galaxies. They adopted the control time method to derive SNe Ia volumetric rates.

The SDSS-II SN Survey (Frieman et al., 2008) had SNe Ia rates calculated by Dilday et al. (2008, 2010). These rates (z≤0.3) bridged the intermediate redshift rates of SNLS to the local universe surveys of Cappellaro et al. (1999) and Li et al. (2011a). Their detection efficiencies were calculated by Dilday et al. (2008), and ∼17,000 light curve Monte-Carlo simulations were performed. Their sample of 516 SNe Ia was an order of magnitude improvement over the rates within a similar redshift range that preceded them. However, only 11 objects were in their lowest redshift bin.

The local universe supernova rate is currently described by targeted surveys and results dominated by statistical uncertainties. To understand the SNe Ia progenitor channel a DTD needs a low-redshift anchor. The Palomar Transient Factory was the revolutionary survey needed to build up large samples of low-redshift objects. It is the data from this survey that forms the foundation of this thesis.

1.7   Thesis Outline

In this thesis, I use the substantial data products from PTF to calculate the most well constrained SNe Ia volumetric rate in the local universe. This process necessarily requires detailed simulations of fake transients. The methods laid out over the course of this thesis can easily be applied to other transients if they have a known light curve evolution. Primarily, this thesis is concerned with the analysis of SNe Ia. An outline of this work is as follows.

In Chapter 2, I introduce the Palomar Transient Factory (PTF). I discuss how PTF finds transients in observations, uses machine-learning to determine whether they are astrophysically real and then follows-up the transient from additional facilities. I also described data engineering steps I took to transform the PTF data into an optimised structure for later simulation needs.

In Chapter 3, I begin the analysis of the real-time recovery efficiencies. I describe the method of inserting fake point sources into real observational data. I replicate the real-time transient hunting techniques of PTF to recover the controlled population of fakes. From this I present the real-time recovery efficiencies of the PTF transient detection pipeline.

In Chapter 4, I perform the SNe Ia specific light curve modelling in a simulation that replicates the operation of PTF. I detail my choice of SNe Ia model and how I reconstruct the PTF operation. I also justify where and when on the night sky I simulate SNe. This process returns the recovery efficiency of PTF as a function of SNe Ia specific properties.

In Chapter 5, I examine the real SNe Ia that PTF discovered. This includes the spectroscopically confirmed objects, the photometrically identified objects, and the missed objects. The final two categories require an analysis of the entire PTF transient database. To all objects in this sample I fit light curves and derive parameters that allow them to be compared to the simulated SNe Ia.

In Chapter 6, I calculate the rate of SNe Ia. I compare the real SNe to the results of the earlier simulations to determine population weights. I also examine a simple power-law DTD on a combined sample of literature rates, including my new result.

In Chapter 7, I examine the rate of a newly discovered transient class called Calcium-Rich Supernovae. I take the PTF discoveries and construct model light curves. The framework for the SNe Ia rate is adopted, but the Ca-Rich model is swapped in and millions of light curves were simulated. From these results I calculate the first ever rate of Ca-Rich transients.

In Chapter 8, I provide a summary of my work, and a perspective on the future applications.

Chapter 2

The Palomar Transient Factory and Data Structures

In this Chapter I discuss the operation of the Palomar Transient Factory (PTF) and present results on its performance, such as number of supernovae discovered. I also describe some of the important follow-up facilities that aided PTF and provided classifications for later study in this work. The final section of this chapter describes an improvement I made to the PTF database structure to significantly increase the computational performance for later applications in this work.

2.1 The Palomar Transient Factory

The Palomar Transient Factory (Rau et al., 2009; Law et al., 2009) was the largest untargeted rolling sky survey of its time. It discovered more than 50,000 non-moving transients and spectroscopically confirmed almost 1,900 supernovae during its 3.5 year lifetime. PTF operated under an automated transient detection system within a tightly integrated collaboration where observation, to data reduction, to transient detection, occurred through a totally automated pipeline. The stated aim of PTF was to systematically explore the optical sky and fill in unknown and sparsely populated regions of the transient phase-space diagram, shown in Figure 1.1. The observational infrastructure of PTF consisted of the Samuel Oschin 48 inch telescope (P48) at the Palomar Observatory in California. The sight itself experiences a typical seeing of $1.1''$ and is populated by a variety of telescopes offering follow-up capabilities. PTF saw 'first light' on 2008 December 13, entering a commissioning phase with the survey itself commencing on 2009 March 1.

2.1.1 The Detector and Filters

The detector on the P48 is the CFH12K which was previously mounted on the Canada-France-Hawaii Telescope at Mauna Kea. It consisted of 12 CCDs mounted in an arrangement of 2 rows by 6. PTF primarily operated in the Mould R filter (R_{P48} or, simply, R) and the g' filter (or g_{P48}), the filter responses can be seen in Figure 2.1. The vast majority, 83%, of PTF observations were taken in R so that the survey could operate throughout an entire lunation. During dark time the g-band filter was used to take advantage of the increased sensitivity to the blue nature of the supernova transients PTF was designed to catch.

Early into the operation of PTF, CCD03 malfunctioned. It was decided that the expense and the risk of possible damage to the other CCDs during any repairs was too high. Therefore, the rest of the survey was carried out on 11 CCDs. Each CCD had a size of 2048×4096 pixels and, when combined, the instrument's field of view (FoV) was 7.26 deg^2 providing a plate scale of $1.01''$pixel^{-1}. The gap between each of the CCDs varies between $33''$ and $46''$, minimising both the lost area for each observation and the need for large dithering. The PTF photometric system (Ofek et al., 2012a) was calibrated against sources in SDSS with r∼16 mag on a field-by-field basis, achieving a precision of 0.02–0.04 mag in both R and g$'$. Fields which weren't in SDSS were calibrated against other PTF fields. A photometric catalogue from PTF of 2.1×10^7 static sources < 19 mag in R and g$'$ can be found in Ofek et al. (2012b).

2.1.2 Cadence Experiments

Maximising the number of supernovae discovered in a sky survey is a balance between depth of each observation and cadence. The cadence of a survey is the mean frequency with which a field is returned to for another observation. Deep, high-redshift surveys, such as Dark Energy Survey (DES) (Kessler et al., 2015) or the Supernova Legacy Survey (SNLS) (Perrett et al., 2010) can take advantage of time-dilation effects to survey large volumes with a slower cadence, but still able achieve a dense rest-frame coverage. Of course, this comes at the expense of local events being missed either through saturation or fast light curve evolution. The étendu[1] of the PTF optical system favours a relatively local universe search; PTF therefore focussed on a high cadence operation. Several cadence 'experiments' were designed to not only build large sample numbers of known transient classes, but to also discover previously unseen events.

---

[1] The product of the telescope's primary mirror area (in m^2) and the FoV (in degree^2)
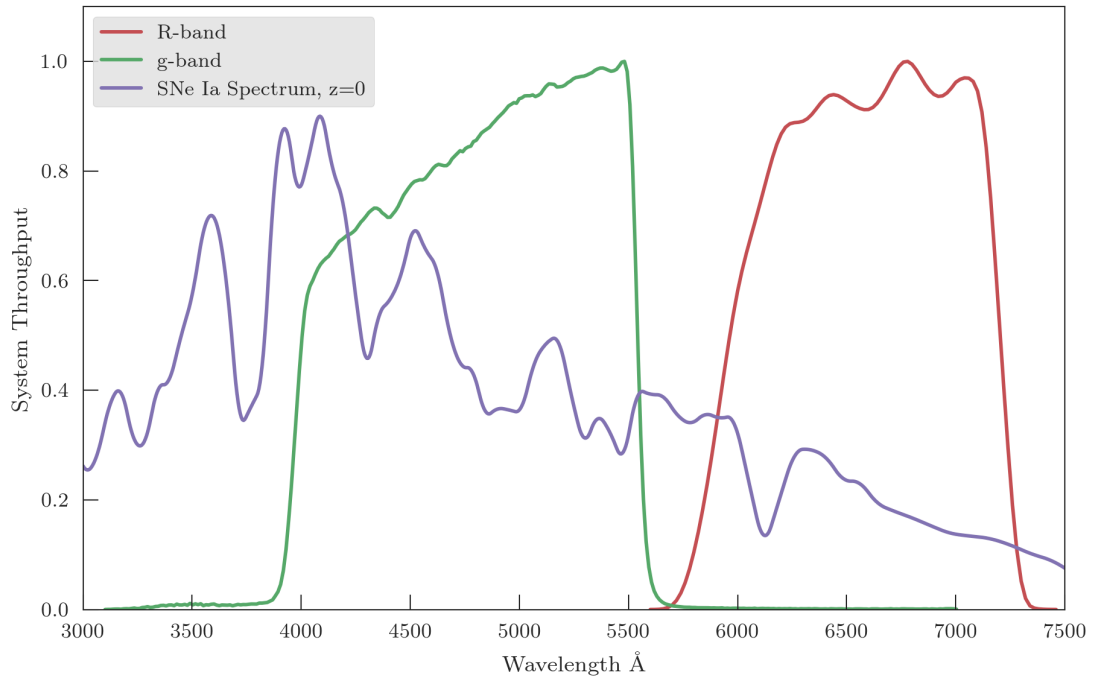
FIGURE 2.1: System throughput through the two main filters used during the PTF survey are shown and a template rest frame SN Ia spectrum is plotted for comparison. The system throughput describes what fraction of the incident light, as a function of wavelength, onto the filter continues along the optical path to the detector which, ultimately, is registered as a charge in the CCD. The system throughput is a combination of the filter transmission function and the CCD quantum efficiency of the CFH12K, measured by Cuillandre et al. (2000). The choice of R-band as the predominate filter was to maximise the on-sky time during periods when the moon is bright and, secondly, R was already on the CFH12k detector. Perhaps most importantly, Mould-R is very similar to SDSS r′, allowing for easy comparisons of PTF to existing catalogues and follow-up instruments.

Survey	Mirror (m)	FoV (deg^2)	Étendu (m^2 deg^2)	Notes
ASAS-SN	(4×) 0.14	73	1.1	Shappee et al. (2014)
CRTS	(2×) 0.7	8	3.1	Drake et al. (2009)
LSQ	1.0	8.7	6.8	Baltay et al. (2013)
PTF	1.2	7.3	8.2	Rau et al. (2009)
Skymapper	1.3	5.7	7.5	Scalzo et al. (2017)
DES	4.3	3.0	37.7	Kessler et al. (2015)
ZTF	1.2	47	53.1	2017+; Bellm (2014)
LSST	6.7[†]	9.6	319.5	2022+; Ivezic et al. (2009)

TABLE 2.1: A comparison of the telescope étendu for similar sky surveys to PTF with a supernova discovery program. The number of objects a survey can detect scales with the telescopes étendu and is therefore a useful figure-of-merit. [†]The mirror on LSST is 8.4m in diameter but has an effective size of 6.7m.

2.1.2.1 5 Day Cadence

The first experiment, and the one which provides the bulk of the data for this thesis, is the '5 Day Cadence Experiment' (5DC); imaginatively named to characterise its goal of surveying, in total, an $\sim$8000 deg^2 footprint on the sky every 3-5 days. The experiment ran from 1 March to 31 October each year and was allocated $\sim$80% of the P48 time, achieving >50% open shutter time in good observing conditions. The 5DC aimed to discover large samples of supernovae, both of the thermonuclear and core-collapse variety. Inevitably, the scale and ambition of the experiment resulted in secondary science achievements, namely the study of active galactic nuclei (AGN) variability (Charisi et al., 2016), cataclysmic variables (Levitan et al., 2013) and tidal disruption events (TDEs) (Arcavi et al., 2014). During this time the P48 was actively surveying at least 2700 deg^2 throughout all phases of the Moon and would adapt this footprint as new areas began to rise or set during the year. In both the R and g$'$ filters an exposure time of 60s was used and the field was returned to after 45-60 minutes for a second exposure. If this secondary observation had not been performed then detections of transients would be heavily contaminated with small solar system bodies and cosmic-rays.

2.1.2.2 Dynamic Cadence

The 'Dynamic Cadence' (DyC) experiment was designed to survey the sky on time scales of seconds up to 3 days in the hunt for rapidly evolving transients. With this in mind PTF chose to focus their observations on regions rich in galaxies less than 200Mpc away, capturing clusters such as Virgo and Perseus. This cadence evolved throughout the survey in its quest to find rare and new transients. Ultimately, this experiment proved extremely successful, identifying new sub-classes known as 'gap-transients' (Kasliwal, 2012), so-called for sitting in the luminosity gap between novae and supernovae in Figure 1.1. Excitingly, for this work on thermonuclear rates, Ca-rich gap transients such as PTF 09dav (Sullivan et al., 2011b), PTF10hcw (Lunnan et al., 2017), PTF 10iuv, PTF 11bij, PTF 11kmb (Foley, 2015; Lunnan et al., 2017), PTF12bho (Lunnan et al., 2017) were discovered in greater numbers than previously seen, allowing questions on their environments to be answered and rates to be calculated with more certainty.

2.1.2.3 Other Experiments

To maximise on-sky time, each lunation when sky brightness increases around full Moon, PTF switched to a narrow-band Hα filter for 3 days to perform a 'Deep Hα Sky Survey'.

Additionally, a single field in the constellation of Orion was observed in R for an entire night with 30s exposures to search for Jupiter sized planets orbiting young stars.

2.1.3 Computational Support

Support beyond telescope and observational hardware was provided by the National Energy Research Scientific Computing Centre (NERSC), located at University of California Berkeley (Nugent et al., 2015; Cao et al., 2016). NERSC allocated PTF over 150khr of computing resources to cope with the more than 100GB of data that was generated each night. In total PTF has used more than 200TB of data storage on spinning disks and performed daily archiving on a 60PB magnetic tape. In 2008 the PTF collaboration wrote the entire real-time data reduction pipeline, described in Section 2.2, to take advantage of time allocated on the Carver supercomputer at NERSC. All of the image manipulation and data processing I performed in Chapter 3 made use of the Carver facility. On 2015 September 30 the Carver supercomputer was retired from service.

The strategy of PTF can be summarised through the following work flow: Observe the sky, find the transients, trigger high quality follow-up and, finally, publish the science (Nugent et al., 2015). The goal was to minimise the time on the first 3 of these stages so that the evolution of the transient could be studied in as much detail as possible. The effectiveness of PTF in achieving this is most beautifully captured in the discovery of the SNe Ia, PTF11kly, later designated SN2011fe (Nugent et al., 2011). This object was first detected by PTF on 2011 August 24 03:59UT. Within a matter of hours the PTF consortium was alerted and multi-wavelength follow-up was immediately triggered. Such was the rapid response of the observations that an OI feature was discovered in the spectrum that had never been observed at that epoch in a SNe Ia before. This feature is seen to substantially evolve in the 8 hours to the second observation, slowing from 18,000km s^{-1} to 14,000km s^{-1}. The early-time observations of 2011fe we were able to constrain the explosion epoch to within 20 minutes as 2011 August 23, 16:29UT. Whilst it is true that discovering supernovae within $\sim$6Mpc so soon after explosion requires a substantial amount of luck, both for a local supernova to occur and for it to be caught, without all stages of the survey operating harmoniously then this enigmatic object may not have become the rich avenue of research it is today.

2.1.4 Past, Present and Future of Palomar Observatory

The original PTF survey ended in 2012 December 21 and transformed into the intermediate Palomar Transient Factory, iPTF, which officially started its survey 2013 February 13 (Kulkarni, 2013). Building upon the success of PTF, iPTF maintained the hardware of
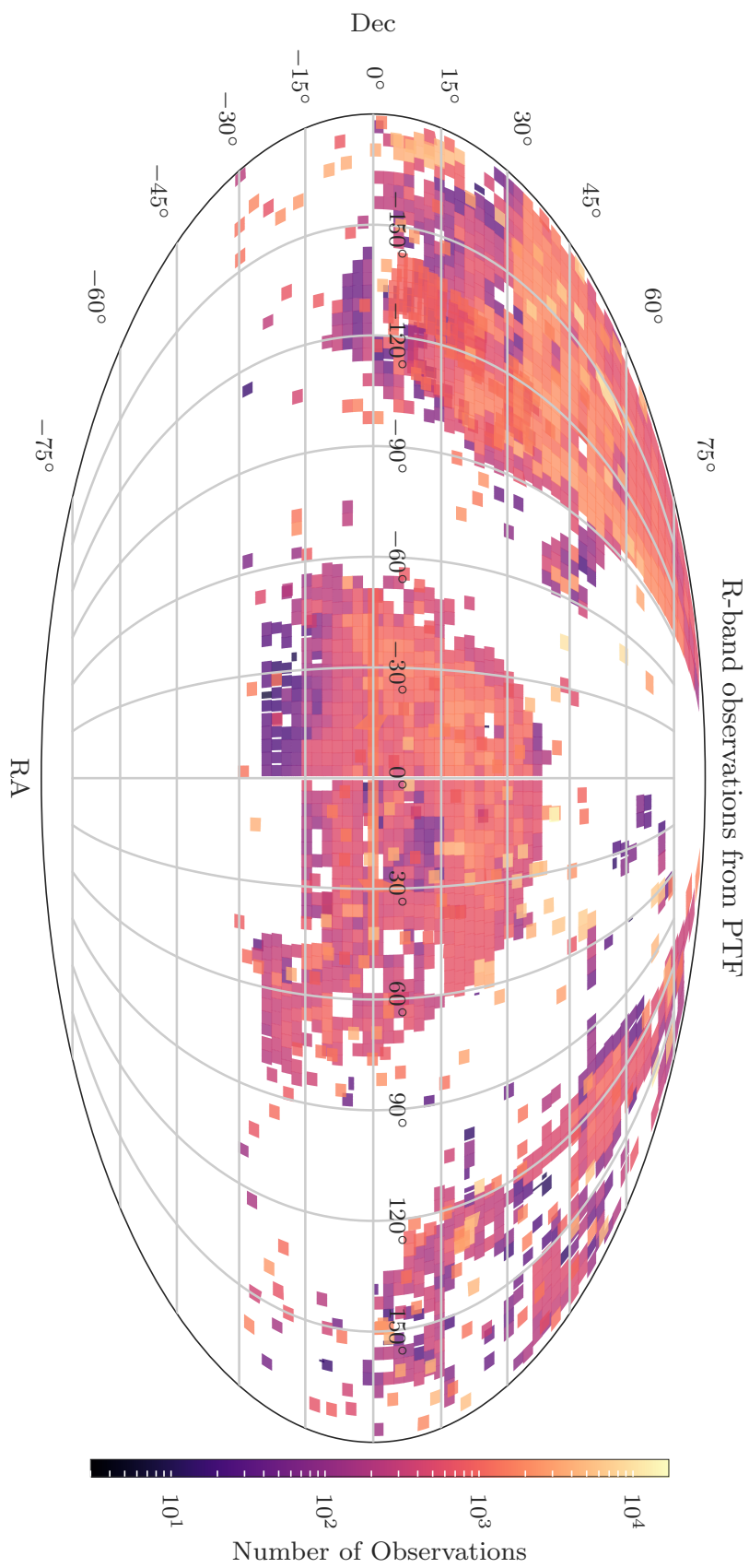
FIGURE 2.2: The total observed sky throughout the survey is shown. Only the R-band observations have been used and data from both the 5DC and DyC are used.

its predecessor but offered improvements to the transient discovery infrastructure. This was achieved through a rewritten photometric pipeline that aided in transient detection and also in the study of variable sources. Additionally, the data taken by PTF was used to produce high-quality reference images so that iPTF could operate immediately. As I demonstrate in Chapter 3.5.2.2, the first few months of PTF's transient detection pipeline were affected partially by poor quality reference images producing subtractions with unreliable source detection. iPTF saw no such period of inefficient operation as it was able to utilise the PTF reference images. Finally, iPTF acted as an experimental platform for the future of time-domain astronomy by accepting user suggested cadence patterns to optimise specific transient sub-class discovery. As we enter a new era of large scale and high cadence sky surveys, the data on transients with rapid evolution could provide new insights into thermonuclear and core-collapse explosion physics.

Indeed, at the time of writing, iPTF has just finished survey operations and is in the process of being upgraded to the Zwicky Transient Factory (ZTF) (Bellm, 2014; Smith et al., 2014). ZTF will herald a revolution in transient astronomy, the CFH12k detector is being replaced by 16 e2v 6kx6k CCDs. Such a vast improvement in design takes full advantage of the P48 focal plane area of $\sim$47 deg^2, all whilst maintaining its $1''\mathrm{pixel}^{-1}$ resolution. ZTF will cover 3760 deg^2 hr^{-1} and visit each field, on average, 300 times per year. Personally, my favourite statistic from ZTF is that in the high cadence survey mode it will detect one SN, within 24 hours of explosion, *every night*!

Whilst both PTF and iPTF have ceased operations, this thesis only contains data from the PTF survey. In Chapter 5.3.3, I use an iPTF catalogue to determine whether sources detected in PTF are known AGN or variable stars so they can be discarded from the supernova search.

2.2 Real-time Transient Hunting

In this section I will detail how transient discoveries were made from the initial stage of observation to extraction and then their release to the collaboration. Additionally, I will describe the way in which PTF handles and stores the data products. In many circumstances it was necessary for me to replicate the PTF database locally so that it could interface with the University of Southampton High Performance Computing Facility, Iridis4. Where relevant I will discuss the improvements I made to the way the data are handled with justifications for doing so.

2.2.1 Observations

Each night the P48 observations are sent, in real-time, to NERSC. Standard corrections are applied to each each image, such as bias subtraction and flat field corrections. Each CCD from the field was processed in parallel beyond this stage. From here on in an 'image' refers to a CCD's single observation of 2048x4096 pixels and not a 'field' which is the tiling of the 11 functional CCDs. The first stage of the pipeline used the Astromatic code for source detection, SExtractor[2] (Bertin & Arnouts, 1996) to create a catalogue of all the sources in the image. These sources were then sent to astrometry.net[3] for comparisons to known catalogues of stars to astrometrically solve the image. The astrometry.net algorithms uses the well defined positions of its own catalogues to determine accurate astrometric solutions for each PTF image. Finally, the source detection software SExtractor (Bertin & Arnouts, 1996), determined fluxes of each of the sources in the image which were then compared to the United States Naval Observatory (USNO)-B1 catalogues (Monet et al., 2003). From there, image zero-point, 3σ limiting-magnitude and seeing were calculated. During this initial stage every modification made to or calculations made on the image were stored in the FITS header. At the end of this process the entire header is written out to a database and forms one of many tables I will refer to throughout this thesis as the 'observing log'. In this instance the metadata are stored in the `proc_image` table. The source catalogues produced by SExtractor are not saved in a database as, over the duration of the survey, this would have become too numerous for the available computing resources.

2.2.2 Image subtraction

The following stages of the pipeline are focussed primarily on the detections of transients. PTF finds astronomical transients through a process of image subtraction. The principles behind this technique are relatively simple but the practicalities are often complex and computationally expensive. The process involves a new 'science image' taken on a given night and subtracted from it is a 'reference image'. This reference acts as a template of a static sky, ideally at a higher quality - in terms of both depth and PSF width - than the science image. PTF achieved this by median stacking several good observations of the same field throughout the duration of the survey. Since the science images that go into the creation of the references may well contain transients themselves, the choice to median stack reduces the probability that contaminant flux would produce negative subtractions and confusion in the transient detection stage.

---

[2]http://www.astromatic.net
[3]http://astrometry.net

The subtraction package HOTPANTS[4] was used to create a subtraction image from the new and reference images. To remove the unchanging sources in the science image, the reference image is convolved to the quality of the science image. This is achieved by sampling the point spread function (PSF) of the science image and then running a kernel over the reference image to match them. Additionally, the reference image is astrometrically aligned to the science image so that, spatially, like-for-like are being subtracted. The reference image is then flux scaled to the photometric system of the science image and subtracted from it, leaving only objects which are variable in brightness or are moving (e.g an asteroid). Additional artefacts are also present in a subtraction, caused by events such as cosmic rays, image misalignment or CCD defects. At regular intervals during the survey operations, reference images were created and updated from previous observations of each field.
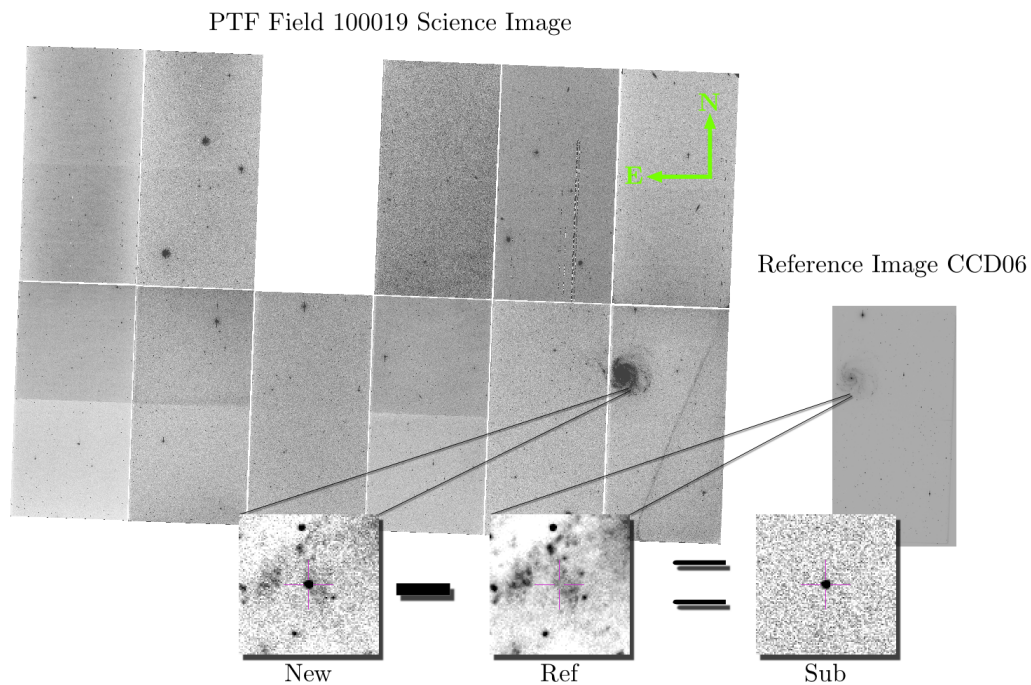
FIGURE 2.3: PTF finds transients through an image subtraction technique. The PTF Field 100019 hosts the bright, nearby galaxy, M101 and is shown as the tiling of the 11 functional CCDs that make up the array. The science image is taken each night and a reference image, composed of a stack of previous images is subtracted from it. Any unchanging sources are removed from the subtraction, leaving only sources which have changed. The stamps show zoomed in regions of the the SN Ia, PTF11kly aka 2011fe, in the Science image 'New', the reference image 'Ref' and the subtraction image 'Sub'. The subtraction image is passed through a source detection algorithm and a machine learning code determines whether the detections are 'Real' or 'Bogus'.

---

[4]http://www.astro.washington.edu/users/becker/v2.0/hotpants.html

Object detection on the resultant subtraction image was performed using, once again, SEXTRACTOR. Anything detected in a subtraction image is called, under PTF nomenclature, a 'candidate'. The subtraction image metadata was stored in the `subtraction` table and the candidates detected in those subtractions were stored in the `candidates` table. Metadata on the reference images were stored in the `deep_ref` table, joined to the subtraction table via a common ID as a single reference image could be used multiple times before being updated with additional information.

On any given night PTF found, in total, between 1-1.5 million objects across all subtraction images (Bloom et al., 2012). Clearly this is an overwhelming task for a human to scan through in order to find the objects of interest. Indeed, only 1 in 1,000 detections in a PTF subtraction could be considered astrophysically real (Bloom et al., 2012; Brink et al., 2013). Therefore PTF utilised an automated process to classify all the candidates, with a probability of the detection being real or 'bogus'

2.2.3 Real or Bogus?

In this section I describe the machine learning process of the PTF pipeline. The algorithm of Bloom et al. (2012) was developed specifically for the PTF pipeline with the remit to make a statistically supported assertion as to whether a candidate found in a subtraction is astrophysically real or 'bogus'. The algorithm was trained on the assessments of human scanners who evaluated a set of 3,750 candidates constructed from the first spectroscopically confirmed SNe during the commissioning stage of the survey. 15% of these candidates were real and 85% were chosen from false positives which closely matched the real supernovae. These scanners were asked to assess cut-out images of candidates from the image subtractions, and to assign a score to each candidate from 0 (bogus) to 1 (real) as a so-called Real-Bogus (RB) score. Additional candidates were constructed iteratively from the human scanner results to ensure the algorithm was well trained. Ultimately from this process, a set of 'features' were determined from the SEXTRACTOR output catalogues which were used to assign an RB score to a candidate so that it best replicated the results of the human scanners. A full list of all 28 features can be found in Table 1 of Bloom et al. (2012). These features needed to be available at the point of detection and, at this stage, could not be derived from external catalogues. This requirement was borne from the desire for rapid follow-up as comparing 1-1.5 million candidates a night to known AGN, variable star or minor body catalogues would have hindered the follow-up process. The only feature that had an 'awarenesses' of its surroundings was the `good_cand_denisty` parameter. This was informative to use as it was found that in poor subtractions the image artefacts were seen to cluster spatially, therefore an atypically high density of candidates would suggest bogus detections.

Since the goal of PTF was to find as many SNe as possible, especially rapid transients, coupled with a concern of missing faint, nearby and young SNe, the decision was made to join candidates which were co-located, with an RB score ≥0.07, to form a light curve. This was seen as a worthwhile compromise to maximise the efficiency of the survey at the expense of a low purity candidate sample. In other words, many supernovae would be found at the expense of a database abundant with spurious candidates.

As previously mentioned, all candidates were entered into the `candidate` table, along with their RB score. Each new entry, with RB≥0.07, was cross-matched spatially to the other candidates already in the database. Anything co-located within $2.0''$and temporally existent between 45 minutes and 6 days of another single candidate is joined and said to be a 'preliminary discovery'. Since the typical cadence of the SN search is around 3-5 days it would appear that the 6 day upper limit for cross-matching is well suited, however, as I show in Chapter 5.3, this requirement can, very occasionally, cause a real object to be missed.

So far, these machine learning algorithms have only considered events based on the information made available in real-time. I make extensive use of this machine learning stage in Chapter 3.4, and I am only concerned with the real-time candidate discovery for the work of this thesis. However, there are additional steps to ensure that the follow-up process acts as efficiently as possibly. It is extremely undesirable to waste telescope time, which is assigned for supernova science, on objects such as M-dwarf flares, AGN or even noise! Therefore, a secondary machine classification stage was employed which made a contextual assessment of the preliminary discoveries. Whilst I do not consider this stage in later assessments of PTF's performance, I will present a brief summary of the process as it influences the spectroscopic follow-up program in PTF.

2.2.3.1 Contextual Assessment

There is much to be learned from the immediate environment, colour and external database comparisons for each potential discovery. For example, a transient event discovered in an 'old, red and dead' galaxy is most likely to be a SN Ia. PTF therefore created a secondary classification stage named OARICAL (Bloom et al., 2012). Queries to external sources such as SDSS DR7, USNO-B1 and SIMBAD were made and a database of the features, such as distance to a known source and source type, was appended to. This contextual information, plus any limited time-series information on the transient was used to assign the event one of four classes: `SN/Nova`, `AGN-cnSN-TDE`, `VarStar` or `rock` which are Supernovae/Novae, circumnuclear event, variable star or asteroid respectively. A confidence score was also assigned to each classification and a continual

- Image Ingest
- Image Analysis
- Realbogus generation

Candidate & Subtraction Database
LBNL/NERSC

subtraction completed

Web query interface to Oarical

Query for real candidates in a date range or specified position, with other candidate matches

Create contextualized scored of possible sources based on nearby candidates; build light curve of source

SIMBAD

USNO B1.0

SDSS

Determine context features by querying webservices

Store discovered source in internal Oarical database; determine time-domain features

Save & annotate high-scoring discoveries

Determine PTF type, robotclass, and classification. Store in internal database.

PTF Marshal
Caltech

Generate web summary pages for human scanners & science-specific project groups (Fig. 9)

Push email summaries to PTF collaboration

FIGURE 2.4: The flow chart describing the stages of the machine learning process to discover transients. This image was taken from the work of Bloom et al. (2012). The initial stage of the pipeline titled 'Candidate & Subtraction Database' is the primary concern of this thesis as this is where the candidates are discovered in real-time.

update to this classification was made as new information became available. Along with the Oarical assessment, any candidate with an RB score >0.3 was automatically gilded with a 'PTF Name' and added to the PTF Marshal (Section 2.2.4) for human assessment and follow-up allocation. Members of the collaboration with different science goals could create custom alerts for objects with different Oarical confidence, for example, a SN team member could customise an alert for anything Oarical deems to be a `SN/Nova` with more than 25% probability but less than a 1% chance of being a `rock` (Law et al., 2009). This alert allowed for a rapid response and the user can access the PTF Marshal to visually assess the object.

2.2.4   The PTF Marshal

The PTF Marshal is a database with a web based GUI front-end containing information on all objects assigned a PTF name. Members of the PTF collaboration were able to login and view a cut-out stamp of each object's discovery image, co-located position in a reference image, the resultant subtraction stamp and a colour stamp of the SDSS region. The real-time light curve from the P48 is also displayed along with additional observations from other facilities. From this available information the collaboration were able to discuss the object, provide classification opinions and cross-reference to other transients. As spectra were taken of the objects they were also added to the Marshal in addition to the Weizmann Interactive Supernova data Repository (WISeREP[5]) (Yaron & Gal-Yam, 2012). Within the Marshal the users can interactively manipulate the data to identify lines in the spectra, measure expansion velocities and classify candidates. By the end of PTF operations the Marshal contained over 4,400 spectra across all types and followed more than 1,800 spectroscopically confirmed SNe.

2.3   Transient Follow-up

The way in which spectroscopic follow-up was prioritised in PTF and assigned time at partner telescopes was a somewhat complex process with substantial human input. Photometric follow-up from the Palomar 60 inch telescope (P60) (Cenko et al., 2006) was allocated when candidates were assigned high confidence scores in Oarical and met special criteria, such as in nearby galaxies. The collaboration also benefits from a global workforce which takes advantage of the 8-10 hour time difference between the Palomar Observatory and Europe. As candidates could be entered into the PTF Marshal within 45 minutes of the observations being made, then an update to the Marshal coincides with

---

[5]http://wiserep.weizmann.ac.il/

a European collaboration member's normal office hours. This allows those members of the collaboration to vet the most promising candidates and any objects potentially of special interest were assigned spectroscopic follow-up. The multi-wavelength follow up strategy is detailed in Gal-Yam et al. (2011). External facilities such as the Intermediate dispersion Spectrograph and Image System (ISIS) on the William Herschel Telescope (WHT), the Low Resolution Imaging Spectrometer (LRIS) (Oke et al., 1995) on Keck-I, the DEep Imaging Multi-Object Spectrograph (DEIMOS) (Faber et al., 2003) were utilised most often as part of a standard observing run to obtain quick spectra on as many targets as possible. In Table 2.2 I have listed the different telescopes and associated instruments which have contributed spectra to the PTF Marshal. The ambitious goal of this extensive program was to produce the most spectroscopically complete and unbiased[6] sample of supernovae, to-date, in the local universe (z<0.1). Additionally, a Target of Opportunity (ToO) program was established with facilities such as the Liverpool Telescope (LT), the Las Cumbres Observatory Global Telescope network (LCOGT) and the KAIT 0.75m telescope to achieve, rapid, same-day classification and improved photometric coverage gaining colour information. All objects which had spectra taken as part of a PTF program were uploaded to the WISeREP archive where members of the collaboration could manipulate and download both the 1D and 2D spectra.

In Figure 2.5 the redshift distributions are shown for the spectroscopically confirmed supernovae belonging to the different sub-classes. For low redshifts (z≤0.05) it can be seen that the Type II events are more common than the SNe Type Ia. This is expected as the Ia rate is approximately 25-40% that of the CC rate in the smallest ($\sim 10^9 M_\odot$) to largest ($\sim 10^{11} M_\odot$) galaxies respectively (Li et al., 2011a). However, the total number of SNe Ia spectroscopically classified far exceeds that of any other sub-group. This is mostly due to their brightness allowing them to be seen to greater distances, d, which increases the search volume, $\propto d^3$. At the brightest and rarest end of the distribution, SLSNe are also found by PTF, whilst only 24 were identified in real-time (8 more were found after a reassessment of public light curves (Perley et al., 2016)) it still represents one of the largest samples of SLSNe in a survey catalogue.

2.4 Data Engineering

Much of the PTF data was stored in a relational database, PostgreSQL[7] (colloquially called postgres), based on the Structured Querying Language (SQL). This allowed for

---

[6]Amusingly, at a conference I was asked about the simulations I perform of supernovae in PTF and if I can account for the 'Silverman bias'. It was later explained to me that after the high priority objects were complete, if this particular PhD student was at the telescope then the 'funny' named objects were observed. He proudly exclaimed that PTF12ass, a 91bg-like SNe, was his discovery.

[7]https://www.postgresql.org

Telescope	Instrument	Spectra	Classifications
Keck I 10m	LRIS	1166	583
Palomar 5.1m Hale	DBSP	1164	759
WHT 4.2m	ISIS	482	278
Lick 3-m	KAST	465	164
Kitt Peak 4m	RC Spec	320	230
WHT 4.2m	ACAM	219	172
Keck II 10m	DEIMOS	187	74
University of Hawaii 88inch	SNIFS	139	115
Gemini North	GMOS-N	89	58
Apache Point 3.5m	DIS	56	46
Telescopio Nazionale Galileo	DOLORES	39	23
Hobby-Eberly Telescope	LRS	23	11
Gemini South	GMOS-S	14	12
Wise 1m	FOSC	10	0
Nordic Optical Telescope	ALFOSC	8	5
VLT	X-Shooter	7	7
Magellan Clay	LDSS3	6	5
Gran Telescopio Canarias 10.4m	OSIRIS	5	4
MDM 2.4m Hiltner	Mk III	2	2
Palomar 1.2m Oschin	CFH12k	2	0
Palomar 5.1m Hale	TSPEC	2	0
Faulkes Telescope North	FLOYDS	1	0
Liverpool Telescope	Frodospec	1	1
Palomar 1.5m	P60 Camera	1	1
VLT	FORS2	1	1

TABLE 2.2: A table of the telescopes and instruments utilised by PTF and which supplied at least one spectrum uploaded to the PTF Marshall. An analysis of a large spectroscopic SNe Ia sample can be found in Maguire et al. (2014).

quick and efficient querying of billions of data entries from a host of tables which were joined via mutual relationships and from which the entire observing history and transient discovery system could be retraced. Whilst the use of postgres was the most appropriate solution for entering modest amounts of data each night, which of course added up to vast quantities over 3 years ($> 10^9$ candidate entries), for my own simulations there were limitations to this system. I needed to rerun the survey hundreds of millions of times in my own Monte Carlo simulations of transient populations. In early development I discovered a bottleneck when performing queries with an RA and Dec. dependence. To alleviate this bottleneck I adopted an extension to postgres which has been extensively developed for Geo-Physicsts called PostGIS (Post Geographic Information System).

2.4.1 Spatial Querying

Simple coordinate systems in Cartesian space require a transformation when projected on a spherical surface, such as the celestial sphere, in order to preserve the geometry
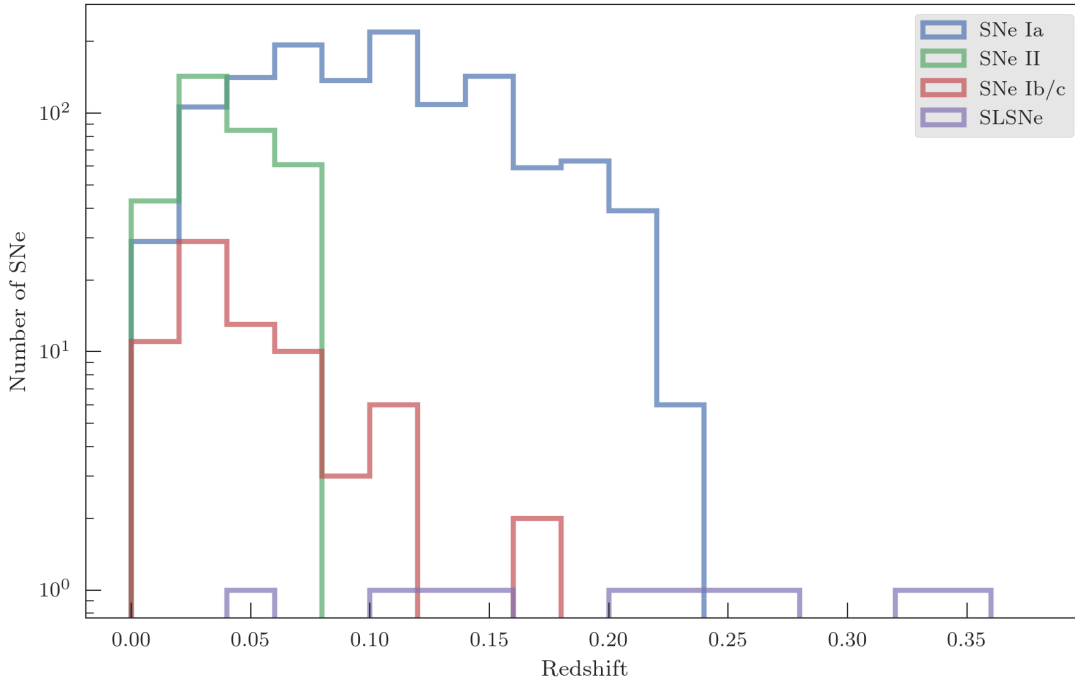
FIGURE 2.5: The four major classes of the supernova family with the number of spectroscopically confirmed objects in each redshift bin shown. Whilst at very low redshift Type II supernovae are the most common sub-type, their comparatively low luminosity limits the distance to which they can be seen. The redshift as which spectroscopic incompleteness occurs can be estimated by the maximum of each distribution. The relatively new class of supernovae, the Superluminous Supernovae (SLSNe), are rare and extremely bright, less than 1/1,000 the Core-Collapse rate (Quimby et al., 2011) but at $M_B \sim$-21., are still found by PTF.

from which they originated. Standard data structures for storing information like dates, object names and ID can easily be indexed through some hierarchical sorting method such as, in the case of postgres, a B-tree. Numerical and alphabetical sorting is both intuitive on the user end and fast on the server side. Consider a randomly ordered column of data with N entries and a query to return all matches; a check must be performed on all N rows to see if it satisfies the conditions of the query. Such an execution would scale as $\sim$O(N) and hence undesirably slow for large N. Indexing a database creates a copy of the desired column and sorts it so that a query does not have to search all records for suitable matches. Most indexing procedures reduce the execution time to that of $\sim$O(log(N)), a significant overall reduction for large tables, such as those in PTF.

However, there are other types of information one might wish to store and index on with particular applications to astronomy. What is the logical way to index millions of spherically projected polygons like a CCD or field footprint? How are differently-dimensioned data efficiently handled and returned only when they intersect, such as 1D lines (asteroid orbital paths) and 2D polygons (CCD footprints)? And, what is the best

Type	Number Classified
SN Ia	1241
SN II	220
AGN	119
SN IIn	87
Galaxy	69
SN IIP	56
SN Ic	40
SN IIb	35
SN Ib	29
SN Ic-BL	17
SLSN-I	12
SLSN-II	9
SN Ia 02ic-like	9
SN Ib/c	9
SN I-faint	7
SN Ia 02cx-like	6
SLSN-R	3
SN Ibn	3
TDE	3
SN IIL	2
Gap I	1

TABLE 2.3: This table shows the number of objects classified belonging to different transient types. These data were taken from the `ptftrans` database which is a collation of the collaboration comments on objects in the Marshal. The classification were assigned based on the Marshal comments. Any re-analysis of the data, which has resulted in a classification change, is not represented in this table. In the case of Type II it was common for the generic 'II' label to be assigned as a catch-all for these objects.

way of storing coordinates (RA, Dec.) so that the nearest N neighbours within a given number of arcseconds can be clustered into a group? These are all examples of data handling problems experienced in the completion of this work and hence novel solutions were used. To index a spatial database the smallest rectangle that encloses a shape, called the bounding box, is drawn and stored. Subsequent sub-, sub-sub- and sub-sub-sub- etc. boxes are constructed and information on the overlaps are stored. Queries are executed and conditions are checked to see which boxes overlap, this follows down an unbroken chain of smaller boxes that meet the condition down to a desired resolution. This type of sorting algorithm is called an R-tree and is utilised by PostGIS. Further speed improvements to the spatial data are achieved by 'clustering' the data structure so that objects which sit close together in physical space are rewritten on the disk to be near each other in 'silicon-space'.
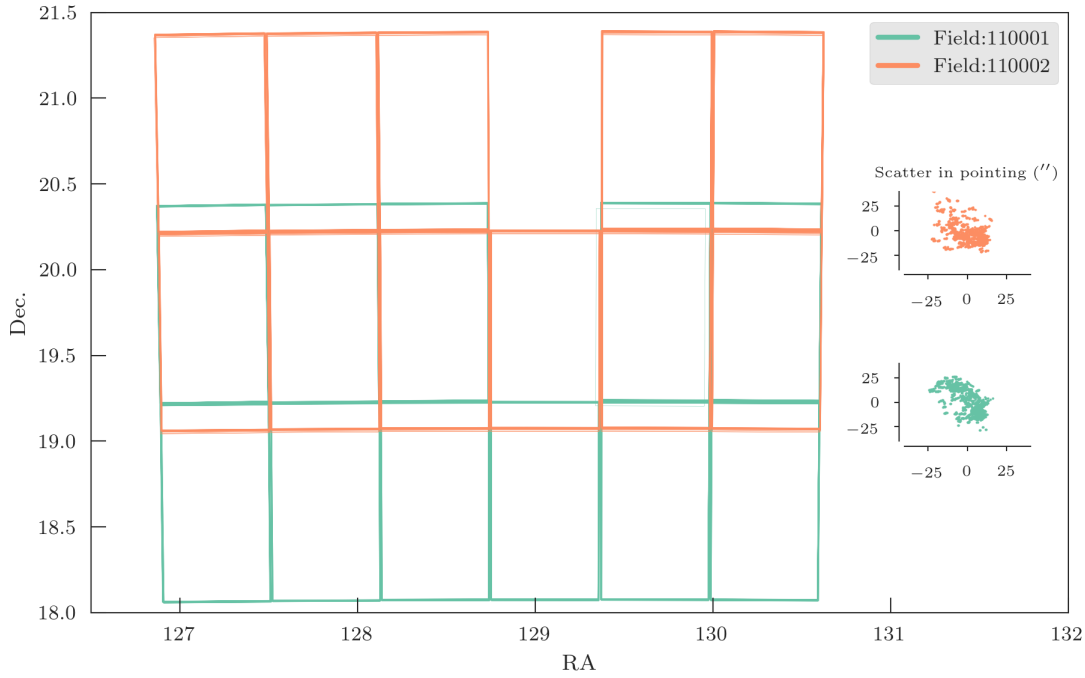
FIGURE 2.6: Two fields from the DyC experiment, 110001 and 110002 are shown with each individual pointing of a CCD drawn. The fields have 1699 and 1416 separate images - 11 per pointing. Occasionally the astrometry fails under poor conditions, explaining why there are not integer numbers of subtractions per field. The dither in the pointing is small but the scatter of the centre away from the mean is shown in the inset plots for the two fields. Both the scatter and the overlap of the fields demonstrate the need to consider the sky as a union of all PTF pointings and not individual fields as PTF fields do not always represent unique areas of the sky.

2.4.1.1 CCD Footprints

PTF stores the RA and Dec. of the corner pixels for every science, reference and subtraction image in the postgres relational database. Each PTF observation has a small dither, meaning that pointings to the same field are not identically aligned, in many cases shifting by as much as several tens of arcseconds. Therefore a transient that was detected in one CCD one night, may not necessarily remain on the same CCD for all subsequent observations. The naive assumption of consistent CCD coverage would have allowed a query to return all possible observations with only the PTF field, CCD number and a date range with 'classical' indexes on the data. Additionally, unique fields did not always observe unique regions of the sky resulting in fields overlapping. Again, simply assuming the discovery CCD remains constant in time would result in missed observations. The scatter on single field pointing and the overlap of fields is illustrated with two example fields in Figure 2.6. The inset plots show the scatter away from the mean central pointing for each observation of the field. Whilst these example fields represent an extreme case of field overlap from the DyC experiment, it none-the-less

highlights the importance of considering PTF's observations as one large area on the sky, rather than lots of individual fields like other sky surveys, e.g. the 4 low Galactic extinction fields of SNLS (Guy et al., 2010).
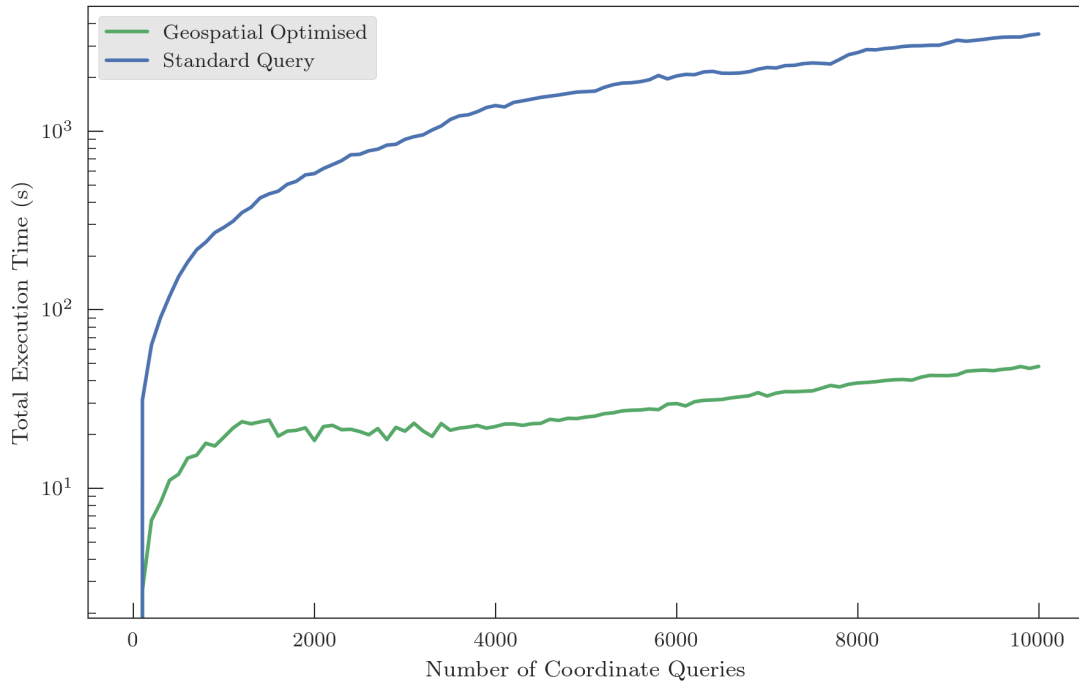
FIGURE 2.7: I compare two different methods of querying the PTF database. N randomly generated coordinates and data ranges were used to query the database to find all the separate observations of that location. The time taken to complete the query was recorded. The blue line represents a standard looping querying, on an unoptimised data structure, checking each CCD in turn to see if it encloses the desired point. The green line represents the query time on an optimised, geo-spatially constructed, table. Immediately it is clear that just a few tens of coordinates display a noticeable and significant difference in execution time.

Motivated by this troublesome property of the data, I created a polygon shape on a spherical surface in PostGIS for each unique observation and subtraction image. For the subtractions alone this was 1,628,115 polygons and a clone of the `subtraction` table was made but spatially reconstructed. I was still able to index the other values in that table, such as observation JD, so that fast spatial queries with constraints on observation dates could be performed. In Figure 2.7 I demonstrate the improvements to the data structure by representing CCD footprints as polygons in PostGIS. To perform this experiment I randomly generated N RA and Dec. coordinates uniform in spherical density. Additionally I generated N JDs as start and final observation dates. The database was queried to find all CCDs that would have observed those coordinates between the dates supplied. The same random variables were passed to each of the query experiments. The green line represents the geo-spatial query syntax (Section

2.4.1.2) following a 'point-in-polygon' algorithm, whereas the 'standard' query (Section 2.4.1.3), in blue, sequentially checks each CCD to see if the desired RA and Dec. falls within the boundary of the CCD. These queries are representative of the type used in later Chapters where hundreds of millions of queries are required and performed in parallel to the database. As the number of queries grows to modest values of just a few hundred coordinate searches, it becomes clear of the dramatic advantage geo-spatial libraries provide. Approximately two orders of magnitude in time separate the different techniques, which for the case of this thesis is the difference between completing the necessary simulations and not! Further improvements to the 'standard' method could have been made by indexing the relevant columns. However, this was an experiment in comparing the untouched database I cloned from PTF to the spatially modified version. Not recorded in Figure 2.7 is the overhead associated with downloading the data once the query has been executed. This, presumably, is identical for both methods as the returned results are identical.

2.4.1.2 Spatial Query Example

```
1  SELECT geo_sub.ujd, geo_sub.ptffield, geo_sub.ccdid,
2  ra_ur, ra_lr, dec_ur, dec_lr, ra_ul, ra_ll, dec_ul, dec_ll
3  FROM geo_sub
4  WHERE st_contains(geo_sub.geo_ccd, ST_GeometryFromText('Point(ra_P dec_P)',4326))=True
5  and geo_sub.ujd>JD_Start and geo_sub.ujd<JD_Stop
6  and filter='R'
7  ORDER BY geo_sub.ujd ASC
```

2.4.1.3 Standard Query Example

```
1  SELECT subtraction.ujd, subtraction.ptffield, subtraction.ccdid ,
2  ra_ur, ra_lr, dec_ur, dec_lr, ra_ul, ra_ll, dec_ul, dec_ll
3  FROM subtraction
4  WHERE ra_lr< ra_P and ra_ur> ra_P and dec_lr< dec_P
5      and dec_ur< dec_P and ra_ll< ra_P and ra_ul> ra_P
6      and dec_ll> dec_P and dec_ul> dec_P
7      and subtraction.ujd> JD_Start and subtraction.ujd< JD_Stop
8      and filter='R'
9  ORDER BY subtraction.ujd ASC;
```

Chapter 3

Artificial point sources and the recovery efficiency

We now turn our attention towards the first step in fully quantifying the survey efficiency; the addition of artificial point sources (fakes) into real observation data. This forms the foundational element of any study that wishes to calculate rates, luminosity functions, or perform any population statistics. In this Chapter I discuss the method and results associated with simulating fake single-epoch events in PTF observational data. This initial stage is free of any transient evolution as it aims to determine the probability of a single epoch event being recovered. However, in later Chapters I build upon this by combining the statistics of single epoch recovery with whole light curve evolution for model transients. The results of this Chapter have been published in Frohmaier et al. (2017) and the data catalogues used to generate the efficiency grids are publicly available[1].

3.1 The Efficiency

It is important to understand the efficiency with which a survey operates as, ultimately, a study based on a population of events is biased towards the objects easiest to find. For every transient that is detected, it is important to know how many events with the same properties were not detected during the survey period. There are many reasons why transients can be missed or not detected by surveys, beyond simple Malmquist bias effects. For example, the observational cadence of the survey may be too long to detect rapidly evolving events; gaps in observing as a result of poor weather, seeing, or technical problems may occur; some parts of the survey area may be inaccessible due

---

[1] http://doi.org/10.5258/SOTON/D0030

to saturated foreground stars, gaps between CCDs, or bad pixels; the detection sensitivity may change as a function of the lunar cycle or other variables; inefficiencies in the complex data reduction and transient detection pipelines may result in transients of any brightness being lost. All surveys will therefore make an inevitably incomplete sampling of the transient population, which will consequently impact the determination of transient volumetric rates, luminosity functions, the dependence of the transient on the underlying stellar populations, and, in the case of cosmological studies using supernovae, the measured cosmological parameters.

These effects and losses can be corrected for, if the efficiency of a survey can be determined. Studies that attempt this require large-scale simulations that can be computationally very expensive. They invariably work via the insertion of 'fake' transients into a survey imaging data stream, passing the adjusted data through the same survey detection pipeline as used to find real transients, and assessing the degree to which the fake transients can then be recovered. This can be done either 'offline' once a survey has been completed (e.g., Pain et al., 2002; Perrett et al., 2010), or in real-time while the survey is operational and the data being collected (e.g., Sako et al., 2008; Kessler et al., 2015). The fake events are usually designed to replicate the properties of the entire range of transients that might be detected, from their apparent magnitude to their host galaxy environment and local surface brightness.

Determining the efficiency of PTF in order to fully exploit these samples for population studies is challenging. Surveys focused on the detection and study of high-redshift type Ia supernovae (SNe Ia), e.g., the Dark Energy Survey (Kessler et al., 2015) and the Supernova Legacy Survey (Perrett et al., 2010), often use a Monte Carlo approach to determining detection efficiencies, synthesizing the light curves of thousands of supernovae over a particular observing season, and inserting fake point sources into each image with the correct photometric properties following the evolution of the synthesized events. This allows the simultaneous determination of both the efficiency on any given epoch, and the recovery efficiency of the underlying SN Ia population. While this is practical for surveys that observe a limited number of fixed fields with a primary interest in one particular supernova type, it does not translate effectively into a survey such as PTF, where we wish to study the populations of any supernova-like transient that PTF could detect.

Indeed, PTF presents its own unique challenges. As discussed in Chapter 2, PTF covered a large area of sky (approximately 8000 deg^2 in the 3-5 day cadence experiment) and operated 9 months per year for four years. During this period, $\geq 2.2 \times 10^6$ images were taken and processed generating just over 1PB of total data in the pipeline including reference, subtraction and noise images, as well as a nearly 1TB database storing the

metadata from every image and all candidate transient detections. It is thus impractical to insert fakes into all of these images in sufficient numbers to study the recovery efficiency on a per-field basis.

The approach I adopted for this work is a two-step process. In the first step, detailed in this Chapter, I choose a single representative field in PTF observed hundreds of times over the four years, with observing conditions that sample the full range that PTF experienced. I inserted millions of fake point sources ('fakes' or 'fake SNe') into every image of this single area, passed them through the detection pipeline, and constructed a recovery efficiency grid as a function of variables such as the transient brightness, image photometric zeropoint, and seeing. The second stage is transient specific and adopts light curve models to build on the single-epoch recovery efficiencies to construct an efficiency grid for the modelled transient type.

3.2 Recovery efficiencies in PTF

PTF, like many other sky surveys, finds astronomical transients through a process of image subtraction (see Chapter 2 for in depth details). In this process, a new 'science' image taken on a given night is astrometrically and photometrically aligned to a 'reference' template image constructed from a median stack of several images taken previously in good conditions. The PSF of the two images is then matched, and the reference image subtracted from the new science image. This leaves an image containing only astrophysical transients that have changed in brightness or position between the two images, as well as subtraction artefacts due to imperfections in the image subtraction process, and other artefacts such as cosmic-rays. Different astrophysical transients can be characterized by a different spatial and temporal evolution: as a trivial example, asteroids move quickly across a field, whereas supernovae are static but change in brightness.

A significant amount of additional metadata are generated by the real-time pipeline describing the context and properties of each CCD image, characterized by over 90 variables, and I make extended use of these image metadata throughout this Chapter. In particular, these data describe the effect of the observing conditions on the images. The metadata, stored for every CCD and most relevant for this study, include:

1. The 3σ limiting apparent magnitude on each unsubtracted image in the R_{P48} filter (m_R^{lim}),

2. The zeropoint to calibrate instrumental magnitudes to the USNO-B1 photometric system (m_R^{zp}),

3. The Full Width at Half Maximum (FWHM) of the image PSF (hereafter referred to as the image quality, IQ). Additionally, the ratio of the IQ in the science image to the IQ of the reference image Φ_{IQ} is stored,

4. The median sky level in counts (F_{sky}),

5. The airmass of the observations,

6. The mean ellipticity of sources in the image,

7. The moon illumination fraction, with 0 denoting new moon, and -1 or 1 denoting full moon.

3.3 Simulations

My simulations were designed to test the performance of the real-time PTF pipeline, therefore the data products I generated from this study are only applicable to the real-time outputs. Any additional image calibration, external to the real-time pipeline, would change the results I find for the transient detection pipeline. For a given set of transient properties and observing conditions, the 'recovery efficiency' ϵ is defined as the ratio of the number of transients found by a survey, to the total number of similar transients that occurred within a fixed sky area. That is, the probability that an astrophysical event with a given set of properties is recovered on a given epoch. I regularly refer to this as the 'single epoch' recovery efficiency, and it is a complex multi-dimensional function of transient properties (e.g., the transient apparent magnitude m_R), astrophysical environmental properties (e.g., local host galaxy surface brightness), and observing conditions (e.g., IQ, m_R^{lim}, etc.). Although some surveys monitor such a recovery efficiency in near real-time by inserting artificial point sources into the data as it is taken each night (e.g., the DES SN program; Kessler et al., 2015), this approach was not used in PTF due to the heavy computational demand of doing this on a near-continuous data stream.

This analysis was performed on PTF data taken between 2009 and 2012 when the survey was fully operational. I evaluated the recovery efficiency by inserting a population of artificial point sources ('fakes') into the PTF imaging data. The resultant images were then treated identically to a new observation, and processed through the same transient detection pipeline as used during the survey (Chapter 2.2), including the machine learning classification. A comparison between the input fake population and the population recovered by the pipeline then provides information on the recovery efficiency on any epoch as a multi-dimensional function of the fake's properties and observing parameters that describe the data.

The computational load of this process – inserting fakes and running the detection pipeline on the resulting image – is high, taking around 7.7 s per PTF exposure (running the 11 CCDs of each exposure in parallel). Thus to analyse every image used by PTF in the image subtraction pipeline once, would require >150 days of supercomputer time. In reality, many additional iterations on each image would be required in order to accumulate the necessary statistics on each epoch, further increasing the required computing time.

Instead, I performed this analysis on a single PTF field, but one that sampled a representative range of observing conditions experienced by the survey. This field was PTF field 100019, observed 1290 times over the survey duration. This field contains the galaxy M101 that hosted the SN Ia SN 2011fe[2] (Nugent et al., 2011), and was observed with an almost daily cadence as part of the 'dynamic cadence' PTF program (Law et al., 2009) in order to study novae and 'fast and faint' transients (e.g., Kasliwal, 2012).

Figure 3.1 shows how the image metadata and observing conditions of field 100019 compare to that experienced by the PTF survey as whole. While identical distributions are not required, it is important that the full range of conditions is sampled by field 100019, and that the distributions are similar, so that the computational resources are used efficiently. It is clear in Figure 3.1 that there is a good agreement between our chosen field and that of PTF as a whole.

3.3.1 Selecting point sources

The fakes are sampled from real point-sources located in each image. I used SEXTRACTOR to locate the 20 brightest, unsaturated, and isolated point sources (i.e., 'stars'), ensuring each was > 50 pixels from the CCD edge. The selection was based both on the SEXTRACTOR neural network CLASS_STAR classifier, which assigns every object a value from 0 (not star-like) to 1 (star-like), and the relevant PSF measurements to meet the aforementioned quality cuts. These cuts removed galaxies and cosmic rays from the fakes catalogue, which I confirmed by visual inspection from a random sample of 1084 candidate stars. Unlike the actual PTF pipeline, I retain all the SEXTRACTOR catalogues and store this information in a postgres table, this table contains more than 350,000,000 source detections. In Figure 3.2 (left) I show the distribution of the FWHM for both the fakes and a random selection of 10 million other sources identified by SEXTRACTOR

---

[2]Although the typical exposure time in PTF is 60 s, due to the brightness of SN 2011fe (reaching $m_R \sim 10$ mag), the exposure time for observations of field 100019 were shortened during the period that SN 2011fe was bright, to avoid saturation of the SN. These shorter exposures, which make up 15% of the field 100019 observations, we discarded from my analysis as they were not representative of PTF as a whole.
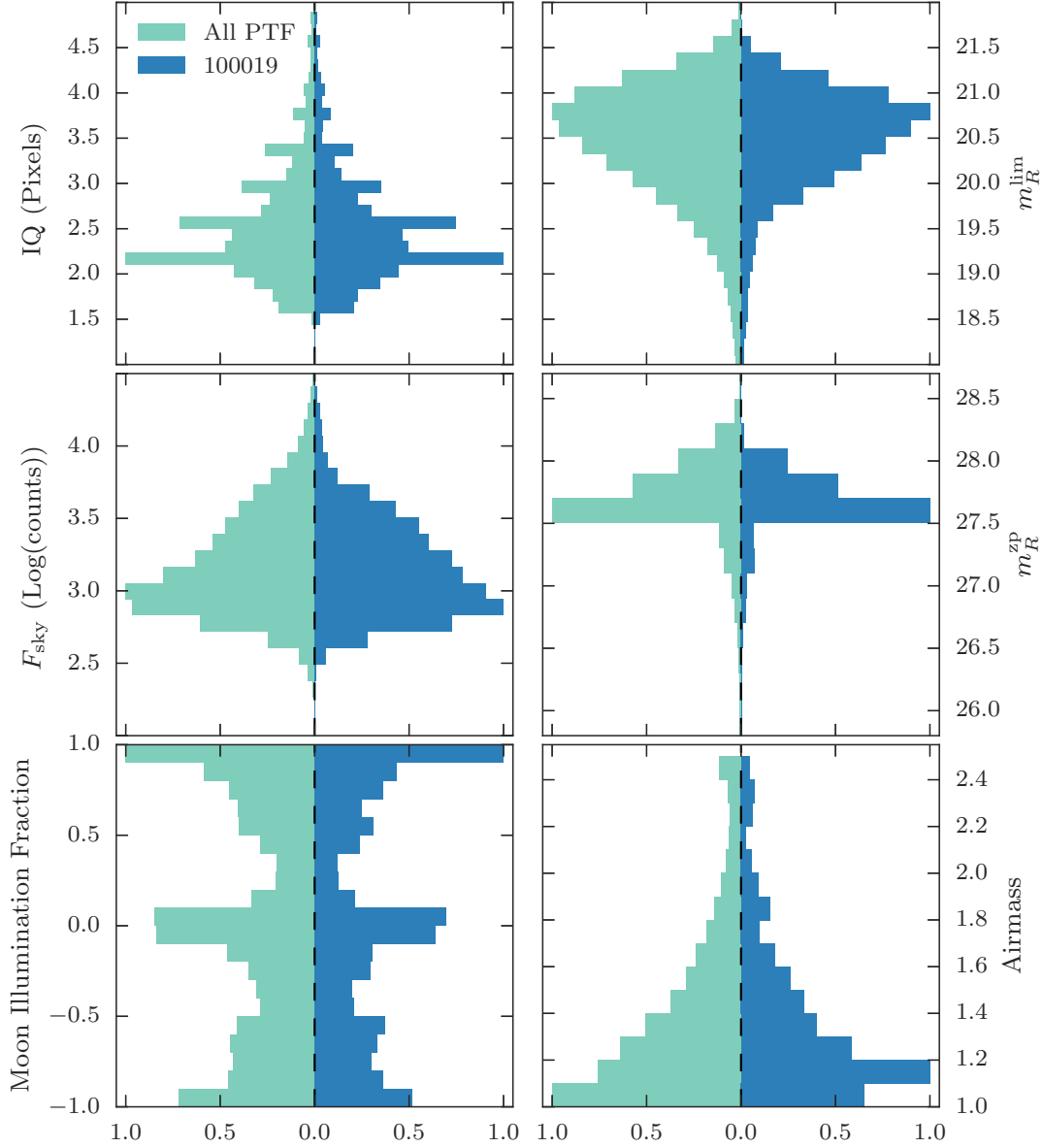
FIGURE 3.1: The (renormalised) distributions of the image metadata and observing conditions across the entire PTF survey (left, light-shaded histograms) compared to those of the PTF field 100019 used in our recovery efficiency simulations (right, dark-shaded histograms). The top left panel shows the comparison for the image quality (IQ), the top right panel the limiting magnitude (m_R^{lim}), the center left panel the median sky counts (F_{sky}), the center right panel the photometric zeropoint (m_R^{zp}), the lower left panel the moon illumination fraction (0 = new moon, -1,1 = full moon), and the lower right the airmass of the observation.
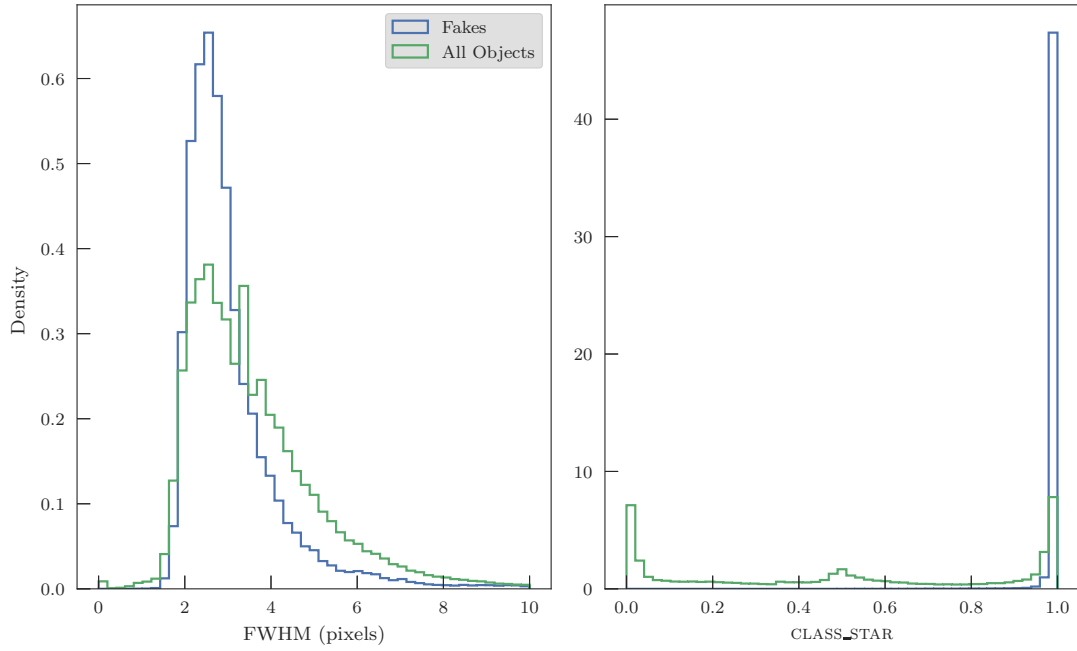
FIGURE 3.2: The distribution of the FWHM of the source stars used to construct fake point sources (left) and the distribution of the CLASS_STAR score for each source star. In both cases a random selection of 1×10^7 other sources in the PTF images were chosen to demonstrate the purity of the sample.

catalogues. Clearly the fakes are sampled preferentially from the peak of the distribution, where most of the stars will lie, the tail of the distribution is a mixture of galaxies and stars observed in poor conditions producing more 'smeared' out PSFs. In Figure 3.2 (right) I show the distribution of the CLASS_STAR scores for the selected stars sources, 99.5% of the objects had a CLASS_STAR score >0.92, further increasing my confidence in the quality of the sample.

It must be noted that there are contaminating sources in the source star sample ($< 1\%$) and I have shown examples of blooming (Figure 3.3(a,c)), blended sources (e), artefacts (b) and an obvious non-star (d). This contamination is difficult to filter out as these 'stars' still receive a high CLASS_STAR value in SEXTRACTOR ($\gtrsim 0.9$), are not considered blended or elongated and have zero quality control flags. It was only through visual inspection that these contaminants could be identified. Whilst it may seem remarkable that these objects were both used and, in some cases, later detected in Chapter 3.5, it is a problem of the SEXTRACTOR algorithm not an inherent problem with my methodology. Identical versions and parameter files for SEXTRACTOR were used for this process as were actually used during PTF's operation. Therefore, if they remain relatively unchanged in later image subtraction processes then they should be detectable.
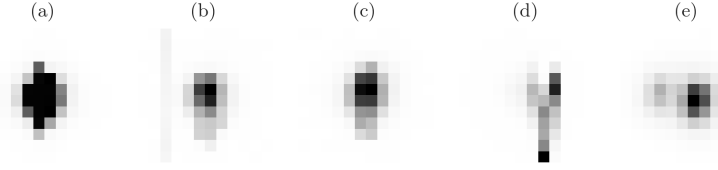
FIGURE 3.3: An example of 5 poor quality source stars that were used to construct fake events. All of these objects passed the rigorous selection criteria in Section 3.3.1 and can only be identified as spurious through visual inspection. Objects like these make up $< 1\%$ of the total sample of objects.

3.3.2 Clone-Stamping

My fake source are constructed by 'clone-stamping' stars in the source sample: I take a box of 9 pixels on a side that encloses the PSF, subtract the local SEXTRACTOR background, and re-scale the star to the desired fake apparent magnitude (m_R). This method ensures that the fakes have a PSF that is both representative of real objects in the image, but also carry the intrinsic variation of the PSF (the object-to-object variation) within the simulation. I generated fakes with a uniform magnitude distribution from m_R=15–22 mag. In addition I enforce the condition that each fake must be a least one magnitude fainter than the original star from which it was generated. In Figure 3.4 I have provided an illustration of the methodology for adding point sources into the images.

3.3.2.1 Inserting fakes into the data

A key consideration when inserting the fakes into the PTF data is that the presence of these 'extra' sources does not distort the machine learning classification process. One of the 28 metrics (Bloom et al., 2012) that goes into the RB score is the spatial density of good candidates, defined as the 'ratio of the number of candidates in that subtraction to the total usable area on that array'. Thus, saturating an image with an artificially high density of fakes may lead to unrepresentative RB scores. A secondary effect is that adding too many fakes into an image could affect the astrometric alignment of the science image to the reference, and thus cause an increased number of subtraction artefacts.

I therefore investigated, using a random sample of 281 images made available to me for pipeline development from the tape archives, how the addition of fakes changed the RB scores of real candidates in the images. In Figure 3.5, I compare our baseline RB scores of real candidates (when there are no fakes in an image) with the RB scores of the same candidates but with an increasing number of fakes added. I found that even a small
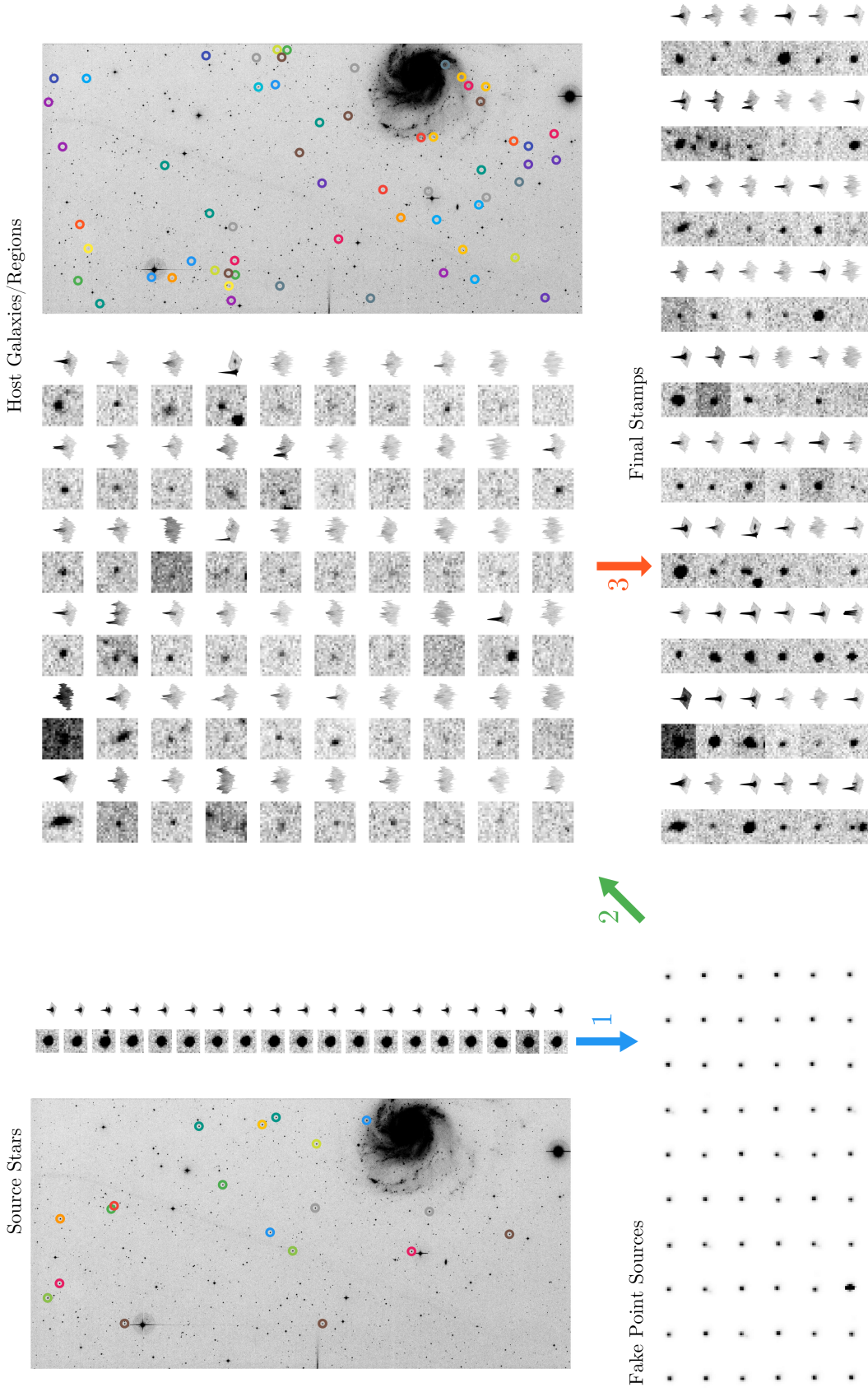
FIGURE 3.4: An illustration of the 'clone-stamp' methodology. 20 source stars are chosen from good quality point source objects. In step 1 those 20 stars are cloned and randomly scaled to 60 different magnitudes (fakes). In parallel, 54 host galaxies and 6 empty regions are selected and in step 2 the fakes are paired to the host locations. In step 3 each fake is stamped into the image to simulate a transient event at that location.
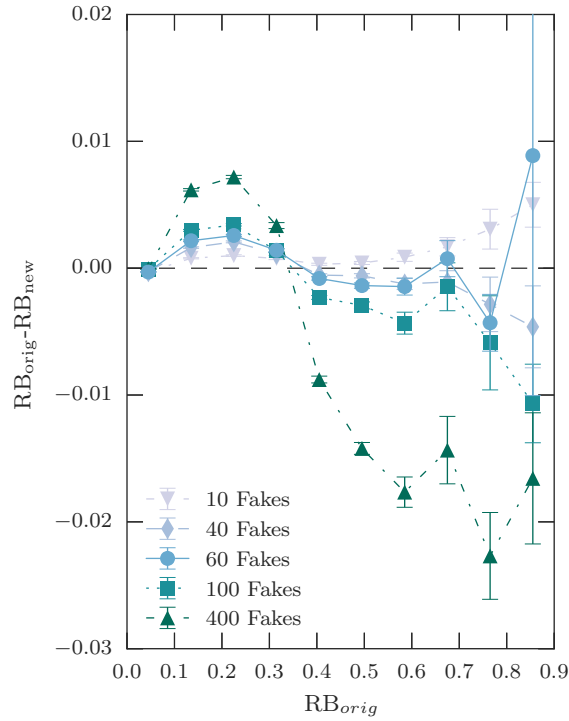
FIGURE 3.5: The original Real-Bogus (RB; Chapter 2.2.3) scores ($\mathrm{RB_{orig}}$) of the real candidate objects in our sample images, compared to the RB scores of the same objects with different numbers of fakes added to the same images ($\mathrm{RB_{new}}$). The figure shows cases where 10, 40, 60, 100 and 400 fakes have been added to the images.

number of fake objects slightly distorts the RB scores; however these effects remain negligible when of order tens of fakes are added, only becoming important with >100 fakes. I considered 60 fake objects per image to be a satisfactory compromise between maximizing the computational efficiency and distorting the RB scores. I also note that even with 400 fakes per image, the astrometric alignment to the reference image was not changed.

3.3.2.2 Fake Point Source locations

Most real astrophysical transient events occur within an associated host galaxy. However, if my fakes were added to random locations on the sky, then the majority would instead be placed in host-less regions, and consequently would provide poor statistics on the recovery efficiency as a function of host galaxy parameters, such as local surface brightness. This would require many more fake point-source simulations in order to adequately map this parameter space.

I therefore chose to bias the locations of the fakes to ensure that 90% of them are placed within a detected galaxy. To select a host for these fake point sources, the SEXTRACTOR

catalogues were used to randomly choose galaxies in each image, with the galaxy pixel positions given by $(x_{\mathrm{gal}}, y_{\mathrm{gal}})$. A fake was added at a pixel position $(x_{\mathrm{SN}}, y_{\mathrm{SN}})$ at an elliptical radius R within the isophotal limit of each galaxy. The elliptical shape parameters are measured by SEXTRACTOR, defined by the semi-major (r_A) axis, the semi-minor (r_B) axis, and the position angle (θ), with R given by

$$R^2 = C_{xx}(x_{\mathrm{SN}} - x_{\mathrm{gal}})^2 + C_{yy}(y_{\mathrm{SN}} - y_{\mathrm{gal}})^2 + C_{xy}(x_{\mathrm{SN}} - x_{\mathrm{gal}})(y_{\mathrm{SN}} - y_{\mathrm{gal}}) \qquad (3.1)$$

where $C_{xx} = \cos^2(\theta)/r_A^2 + \sin^2(\theta)/r_B^2$, $C_{yy} = \sin^2(\theta)/r_A^2 + \cos^2(\theta)/r_B^2$, and $C_{xy} = 2\cos(\theta)\sin(\theta)(1/r_A^2 - 1/r_B^2)$. A value of $R \sim 3$ corresponds to the isophotal limit of each object. The location of each fake was not refined further, for example to follow a galaxy surface brightness profile. The remaining 10% of the fakes were added into blank regions of the sky. I also ensured that a fake is not within 40 pixels of another fake, regardless of whether it is in a galaxy or not.

3.3.3 Application of the Transient Detection Pipeline

The simulation method described above was applied 10 times to all observations of the PTF field 100019 taken over 2009–2012, generating a sample of $\approx 7 \times 10^6$ fakes in the data. To perform image subtraction on the fakes data I utilised the Carver supercomputer at NERSC. Careful consideration was taken to preserve the exact process adopted by PTF during original survey operations and as described in Chapter 2.2. By default the pipeline uses the latest reference image for the observation it is given. Of course, the latest reference images have updates in 2012 and using those on my data would produce a subtraction image that were not representative of the pipeline at the time. Therefore, each observational image I created was linked to the reference used on the original observation date. Additionally, the pipeline was modified for my needs so that the output would not write to the original PTF database, instead, the relevant schema was cloned and new tables were created to accept the pipeline results. The images which needed processing were pooled on Carver and processed in parallel, taking just under 2 weeks.

The product of the simulation and processing were two PostgreSQL database tables. The first stores a complete description of the parameters describing each fake, the spatial location and any host galaxy information, the fake magnitude, and the observing conditions metadata. The second table stores the output from the real-time detection pipeline run on the images containing the fakes, including the machine learning RB
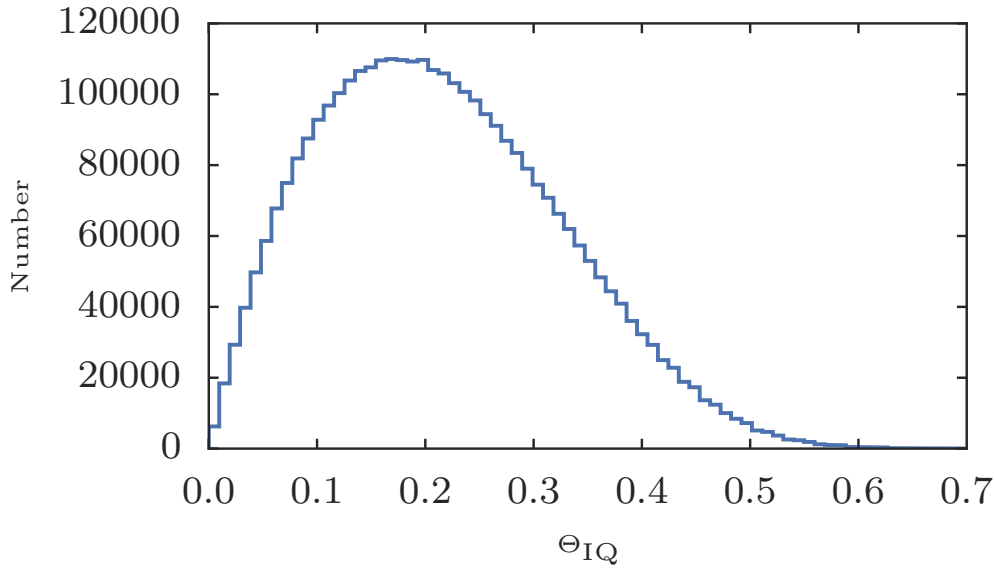
FIGURE 3.6: The distribution of Θ_{IQ}, where Θ_{IQ} is defined as the separation between the position of the input fake and the position of the nearest recovered candidate, divided by the IQ of the image. I require that $\Theta_{IQ} < 0.6$ in order to consider a fake to be recovered by the pipeline.

scores; i.e., it contains information on which fakes were recovered by the pipeline (as well as all the real astrophysical transients and false-positives).

3.4 Fake supernova recovery

To determine whether a fake was recovered by the pipeline, I performed a geo-spatial matching of the two databases (fake positions versus recovered positions), and required that any matched fake must have a RB score ≥ 0.07, the same as during the PTF survey operation (Bloom et al., 2012). The matching radius between a fake and a recovered candidate varies with the IQ (seeing), and to remove spurious associations I define a new variable Θ_{IQ} as the ratio of the separation of a fake and the nearest recovered candidate, to the IQ. The histogram of all Θ_{IQ} is shown in Figure 3.6, and I enforced $\Theta_{IQ} < 0.6$ in order to consider a fake to be recovered. Any fake without a detection satisfying RB ≥ 0.07 and $\Theta_{IQ} < 0.6$ is considered not recovered.

3.4.1 Recovered fake point source properties

I next compared the recovered fake's magnitude to that input into the pipeline (Figure 3.7). Although this is not a critical part of the analysis, as I do not use the recovered
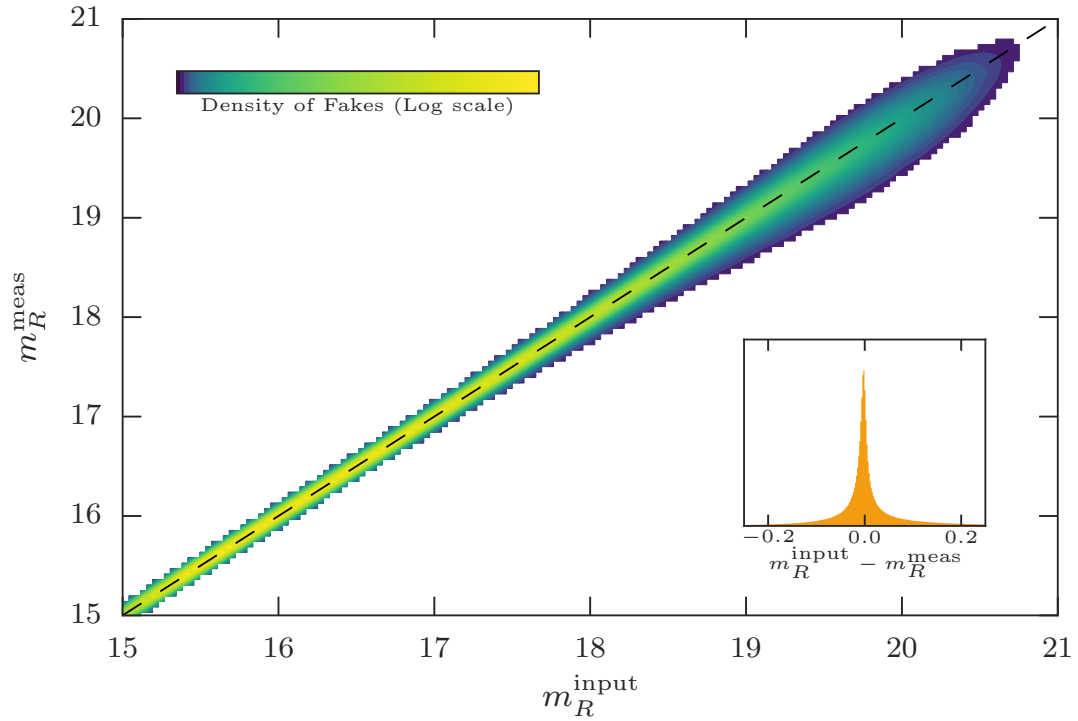
FIGURE 3.7: The SEXTRACTOR measured magnitudes (m_R^{meas}) for the recovered fakes compared to their input magnitudes (m_R^{input}). The main panel shows the overall comparison with the dashed line denoting a 1:1 agreement, and the inset panel shows the distribution of $m_R^{\mathrm{meas}} - m_R^{\mathrm{input}}$ which is sharply peaked at 0 showing no systematic offset.

photometry in any analysis, this test acts as a useful sanity check that the efficiency pipeline is working as expected, and that the PTF real-time pipeline itself can recover reasonably accurate photometry. The agreement is generally good, and as expected, the fainter fake SNe show a larger scatter between their input and recovered magnitudes as the signal-to-noise (S/N) decreases; however the overall comparison shows a good agreement with no systematic offset. I find that 92% of the recovered fake magnitudes are within 0.2 mag of their input magnitudes, and splitting the fakes into bright objects ($m_R \leq 18.5\,\mathrm{mag}$) and fainter objects ($m_R > 18.5\,\mathrm{mag}$) I find 98% and 77% of the magnitudes are recovered within 0.2 mag. Thus, the PTF real-time search pipeline accurately recovers the input magnitudes of the fakes.

3.5 Recovery statistics

I now study the performance of the pipeline in recovering fakes under different observing conditions, and as a function of the input fake's properties and location. This analysis will later motivate the construction of a multi-dimensional recovery efficiency grid as a

function of the smallest number of parameters that affect the recovery of a fake. In later Chapters this multi-dimensional grid can be used together with Monte Carlo simulations to calculate the recovery efficiency of any real transient event.

3.5.1 Recovery efficiencies

The data are binned based on the input fake properties and observing conditions with bin widths and number driven by the precision with which the data are measured. For example, the m_R^{zp} values are determined by the real-time pipeline to an accuracy of 0.1 mag, and so fewer, larger, bins are required compared to m_R^{lim}, which is measured to a higher precision. The same binning is applied to the equivalent data associated with the fakes that are recovered by the PTF pipeline. I then define, in each bin i, the recovery efficiency ϵ_i to be the ratio of the number of fake objects recovered in each bin (k_i), to the total number of fakes originally created in that bin (n_i) i.e., $\epsilon_i = k_i/n_i$. One-dimensional recovery efficiencies for each variable are shown in Figure 3.8, in each case marginalized over the other variables.

3.5.1.1 Uncertainties

An important question is the calculation of uncertainties for each ϵ_i. In each bin there are k successes (detections) out of n independent trials (fakes). To describe the probability of detecting an object from simulations, the binomial distribution was adopted. Through the use of Bayes Theorem[3], we are able to both define our uncertainties for small number cases and still tend towards the Poisson distribution as $n \to \infty$. Additionally, we can set a prior on ϵ to keep it within physical boundaries.

In each bin the probability of success, k, given n trials is a binomially distributed variable expressed by

$$p(k|\epsilon, n) = \frac{n!}{k!(n-k)!}\epsilon^k(1-\epsilon)^{n-k} \tag{3.2}$$

where ϵ is the probability of success for each trial. Adopting Bayes Theorem, we can find the probability of the efficiency given the data as

$$p(\epsilon|k, n) = \frac{p(k|\epsilon, n)p(\epsilon|n)}{\mathcal{Z}} \tag{3.3}$$

where $\mathcal{Z}$ is a normalisation such that $\int_{-\infty}^{+\infty} p(\epsilon|k, n) = 1$ and our choice $p(\epsilon|n)$ is set as a uniform prior such that $p(\epsilon|n) = 1$ when $0 \leq \epsilon \leq 1$ and 0 otherwise. We are finally left with our posterior probability function

---

[3] $P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$,

$$p(\epsilon|k,n) = \frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)}\epsilon^k(1-\epsilon)^{n-k} \qquad (3.4)$$

where Γ is the Euler gamma function; see Paterno (2004) for an in depth discussion and derivation of Equation 3.4. To find the uncertainties I numerically integrate the function to find the shortest interval containing 68.3% of the probability.

3.5.2 Single Parameter Recovery Efficiencies

In addition to the the fake's magnitude, I chose to explore 8 metadata parameters and a 9th of my own construction. The recovery efficiencies for all these parameters can be seen in Figure 3.8. Several clear (and expected) trends are apparent; for example fake objects are more difficult to recover when fainter. However, even when the fake is bright ($m_R < 18.5\,\text{mag}$), we note that a consistent $\approx 3\%$ of objects are not recovered, implying that some small fraction of objects are missed no matter what the brightness. Fake objects are also more difficult to recover as the IQ of the science image becomes poorer relative to that of the reference image; as the limiting magnitude becomes brighter; and as the photometric zeropoint becomes brighter (i.e., the data have more attenuation, presumably from clouds). The recovery fraction is also a strong function of median sky counts (a brighter sky makes the fake harder to detect), a weak function of the moon illumination fraction (objects are harder to recover with a bright moon), and a weak function of airmass (objects are marginally more difficult to recover at high airmass). There is no measurable trend with image ellipticity, indicating the image subtraction works well across most PTF data.

3.5.2.1 Host galaxy surface brightness

As the fakes were inserted (see Chapter 3.3.2.2), I recorded the total integrated R-band apparent magnitude of any host galaxy (m_R^{host}) from the SEXTRACTOR catalogue, as well as the local surface brightness at the position of the fake. I denote this latter parameter 'F_{box}', defined as the background-subtracted sum of the pixel counts at the fake position over different configurations of pixels. I recorded this metric in box sizes from 1×1 to 11×11 pixels, however the default for all F_{box} measures is to use the integrated counts in a 3×3 box as this is close in size to the PSF of a typical fake. This metric provides local environment information for an object's recovery efficiency, i.e., the transient detection pipeline's ability to discover sources against a bright background. The F_{box} metric is the only parameter I discuss that was not output from the from the real-time pipeline during survey operations between 2009-2012. Thus any study based on the results of
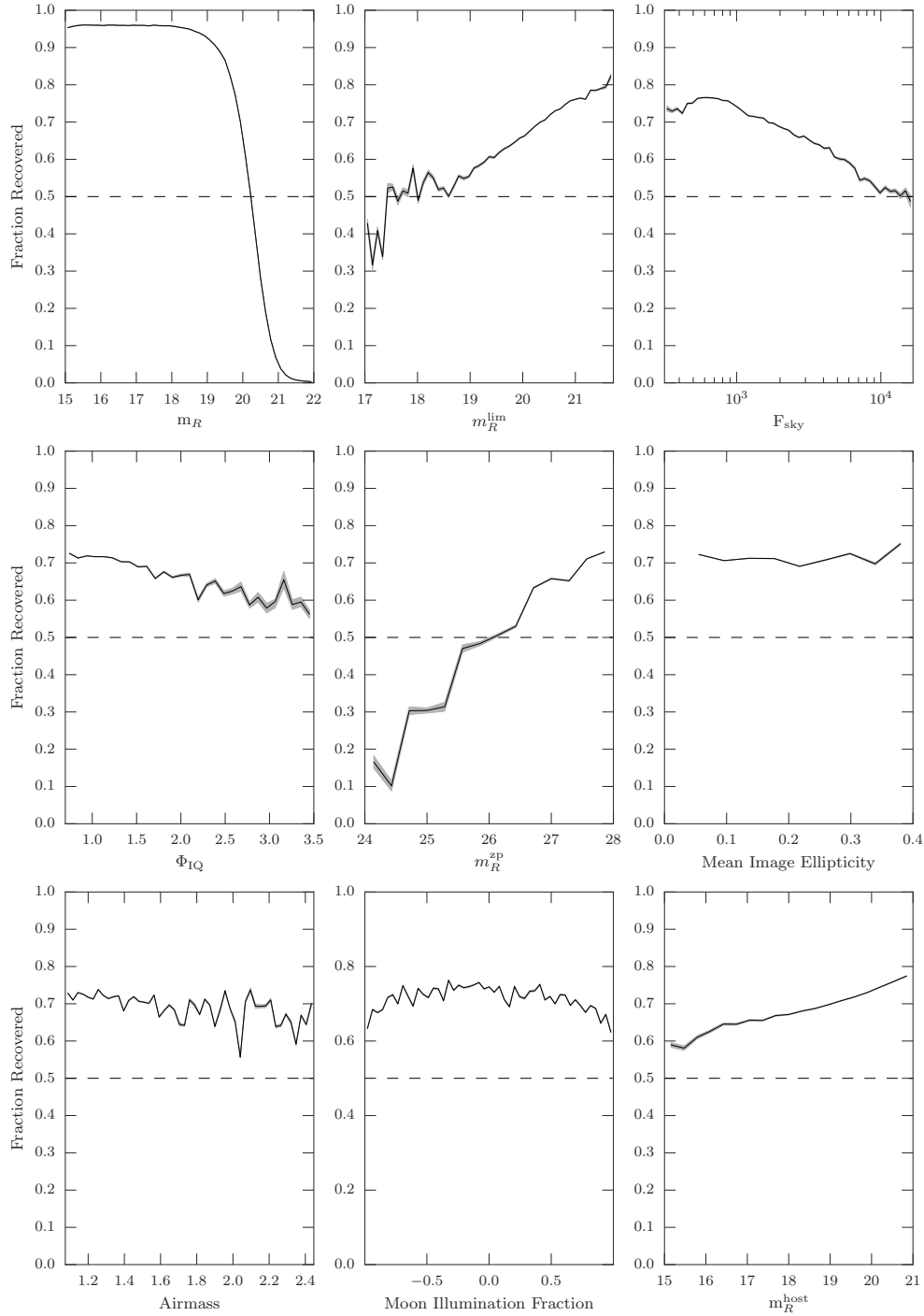
FIGURE 3.8: The fake recovery efficiency ϵ as a function of 8 of the variables we consider and the fake's magnitude. In each individual panel, the recovery efficiencies are marginalized over the other parameters. These are (top row) the fake magnitude m_R, the limiting magnitude of the image m_R^{lim}, the median sky counts in the image F_{sky}, the ratio of the seeing in the science and reference images Φ_{IQ} the image photometric zeropoint m_R^{zp}, the mean image ellipticity, the airmass of the observation, the moon illumination fraction at the observing epoch and the host galaxy magnitude for the fake m_R^{host}. The dashed lines represent the points at which 50% of the fakes are recovered and the shaded regions contain 68.3% of the probability.

these efficiencies that explicitly require the use of F_{box}, will need to measure F_{box} for the transient objects in that sample so that they are directly comparable to my fake simulations. The real-time pipeline did measure a fixed aperture flux of 5 pixels in both the subtraction and the reference, referred to as the *flux-ratio* in Bloom et al. (2012). However, while useful for computing the real-bogus score, it was insufficient for my needs as it was in general too large compared to the typical PSF. Furthermore, the pipeline would only record that metric if it found an object, whereas F_{box} was recorded for all objects regardless of whether they were detected or not. Using the pipeline metric would require me to perform forced photometry at the location of each undiscovered fake whereas constructing my own metric and recording it at the point of fake creation was a more efficient use of resources.

Figure 3.8 shows that fakes become more difficult to recover in brighter galaxies. However, m_R^{host} is a poor choice of metric shown only for information. It is not applicable to all real transient events (where the host association may be uncertain; e.g., Sullivan et al., 2006; Gupta et al., 2016), and can be mis-leading if, say, a transient is well-separated from a bright host galaxy. Instead, the information is more usefully encapsulated by the F_{box} metric. In Figure 3.9 (left) I inspect the recovery efficiency as a function of F_{box} split into bins of fake magnitude, and an expected trend where fakes in regions of higher surface brightness are less likely to be recovered can be seen. I also extended this analysis to a new parameter, $\vartheta_{\mathrm{ratio}}$: the ratio F_{box} to the flux from the fake. This new parameter, when considered alone, provides an insight into how cleanly the image subtraction has been performed, which can particularly affect the fainter fakes on bright galaxies. I note that $\vartheta_{\mathrm{ratio}}$ has a degeneracy with m_R (as both include the counts from the fake) and in Chapter 3.6 I do not use m_R in conjunction with $\vartheta_{\mathrm{ratio}}$ for this reason.

In Figure 3.9 (right) I examine the recovered fraction of fakes as a function of $\vartheta_{\mathrm{ratio}}$. I find the expected trend where fakes that are located in an environment of high surface brightness relative to the object itself are less likely to have been detected by the pipeline. The pipeline maintains a consistently high ability to discover the fakes whilst the fakes are $\approx 10\times$ brighter than F_{box}. The recovered fraction rapidly drops off after this point, with 50% recovered at $\vartheta_{\mathrm{ratio}} \approx 0.7$

3.5.2.2   Efficiencies as a function of time

Due to the improvement and updating of the reference images during the survey (Chapter 2.2), it is expected for the recovery efficiencies to show a time dependence. I therefore analysed the recovery efficiencies as a function of m_R for each year of the survey (Figure 3.10), and find that 2009 has a significantly lower recovery efficiency than the
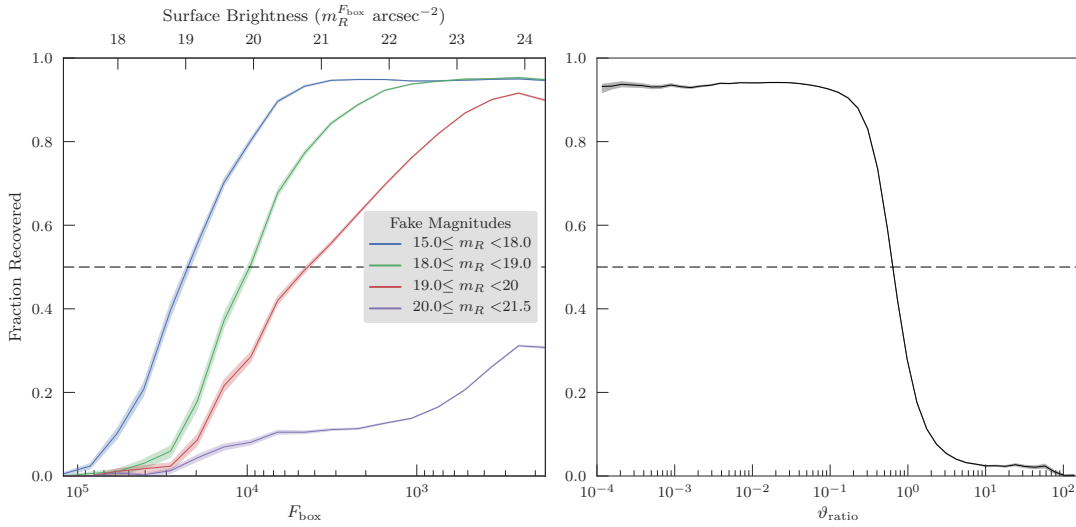
FIGURE 3.9: Left: The recovery efficiency as a function of the F_{box} parameter on the bottom axis and $m_R^{F_{\text{box}}}$ surface brightness on the top axis. I see the expected trend of fewer fakes recovered when they are situated in bright regions (analogous to bright host galaxies). Right: The recovery efficiency as a function of the ratio of counts in F_{box} to the counts of the fake point source (ϑ_{ratio}). I find that if the local surface brightness is $\lesssim 0.1\times$ than the fake, the pipeline finds it $\approx 95\%$ of the time. Beyond this the recovered fraction sharply falls off and the 50% recovery fraction is at $\vartheta_{\text{ratio}} \approx 10^{-0.2}$

subsequent years. The later years – 2010, 2011, 2012 – all show consistent trends. Given the large discrepancy between 2009 and the later years, I exclude 2009 from this study. While the effect in 2009 is partly explainable due to the likely lower quality of the references during 2009 (both in terms of depth and IQ), I also note that the data from 2009 suffered from a 'fogging' problem on the PTF camera window (described in detail in Ofek et al., 2012a). This likely dramatically decreased the efficiency of the survey in the parts of the image affected by the fogging during that period.

3.5.2.3   50% recovery efficiencies

The 50% recovery magnitude m_R^{50} – the magnitude at which PTF finds the same number of transients as it misses – is another useful way of parametrising the survey efficiencies. Taken over all observing conditions, $m_R^{50} \approx 20.3\,\text{mag}$ (Figure 3.8). However, m_R^{50} depends strongly on the observing conditions and galaxy surface brightness. I analyse m_R^{50} as a function of m_R^{lim}, F_{sky}, Φ_{IQ} m_R^{zp}, airmass and moon illumination fraction parameters in Figure 3.11; the trends are as expected.
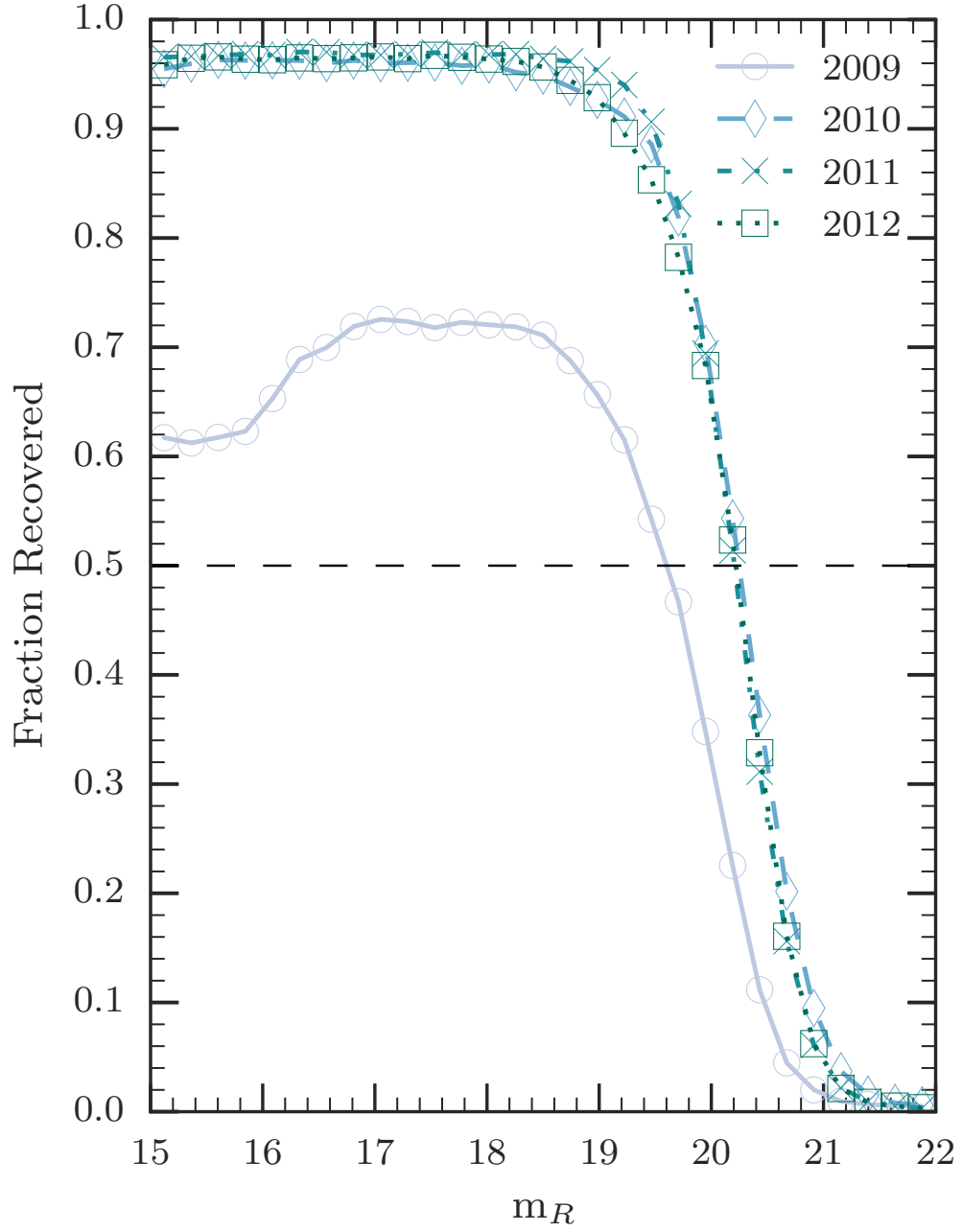
FIGURE 3.10: The recovery efficiency as a function of fake apparent magnitude (m_R) for each year of the PTF survey (averaged over all observing conditions). The years 2010–2012 are consistent, but the year 2009 (the first year of the PTF survey) shows a large discrepancy (see discussion in section 3.5.1). We exclude 2009 from our analysis.
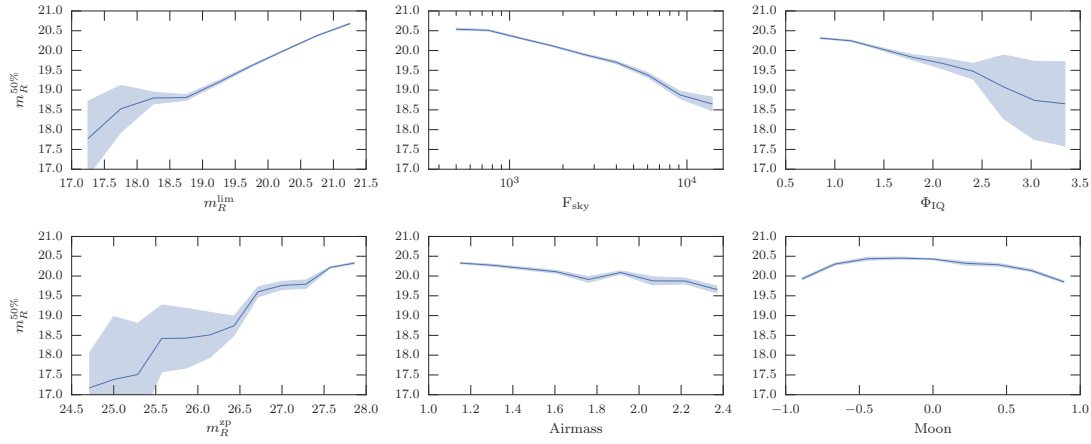
FIGURE 3.11: The 50% recovery magnitude, m_R^{50}, as a function of various observing conditions. The results are plotted on the same axis scale to highlight the negligible dependence of these parameters in shifting the recovery efficiencies which is explained further in Chapter 3.6. The shaded areas indicate the statistical uncertainties (containing 68.3% of the probability).

3.6 Multidimensional recovery efficiencies

I extended the analysis of the single parameter recovery fractions to study PTF's performance as a function of multiple variables – the final recovery efficiency grid. This method allows for situations to be studied which cannot be encapsulated by any single parameter, for example bright transients occurring in poor observing conditions. It is possible to create a multi-dimensional efficiency grid from all of the parameters discussed in Chapter 3.5.1 and shown in Figure 3.8; however, several of these variables are likely to encapsulate similar information, and are therefore degenerate (the correlations are given in Figure 3.12). For computational reasons, it is more efficient to construct a final recovery efficiency grid composed of the fewest dimensions possible, but which capture the great majority of the variation. In this section, I therefore examined the most important variables that make up the final efficiency grid. I stress that whilst I aim to reduce the number of dimensions to produce a final efficiency grid applicable for most purposes, there is flexibility in this method to include any number of parameters to meet the specific science goals of a study.

The first dimension of the final efficiency grid is the apparent magnitude of the fake object (m_R), a variable that is clearly essential. The second dimension is F_{box}, again containing information not captured by the other variables. The remaining dimensions are then drawn from the observing conditions. In Figure 3.12, I explore the Pearson correlation coefficients for the 6 pieces of recorded metadata listed in Chapter 3.2; I neglected the image ellipticity, as it has little impact on the efficiencies (Figure 3.8). I then constructed, in Figure 3.13, the 6-dimensional grid of efficiencies where each cell in
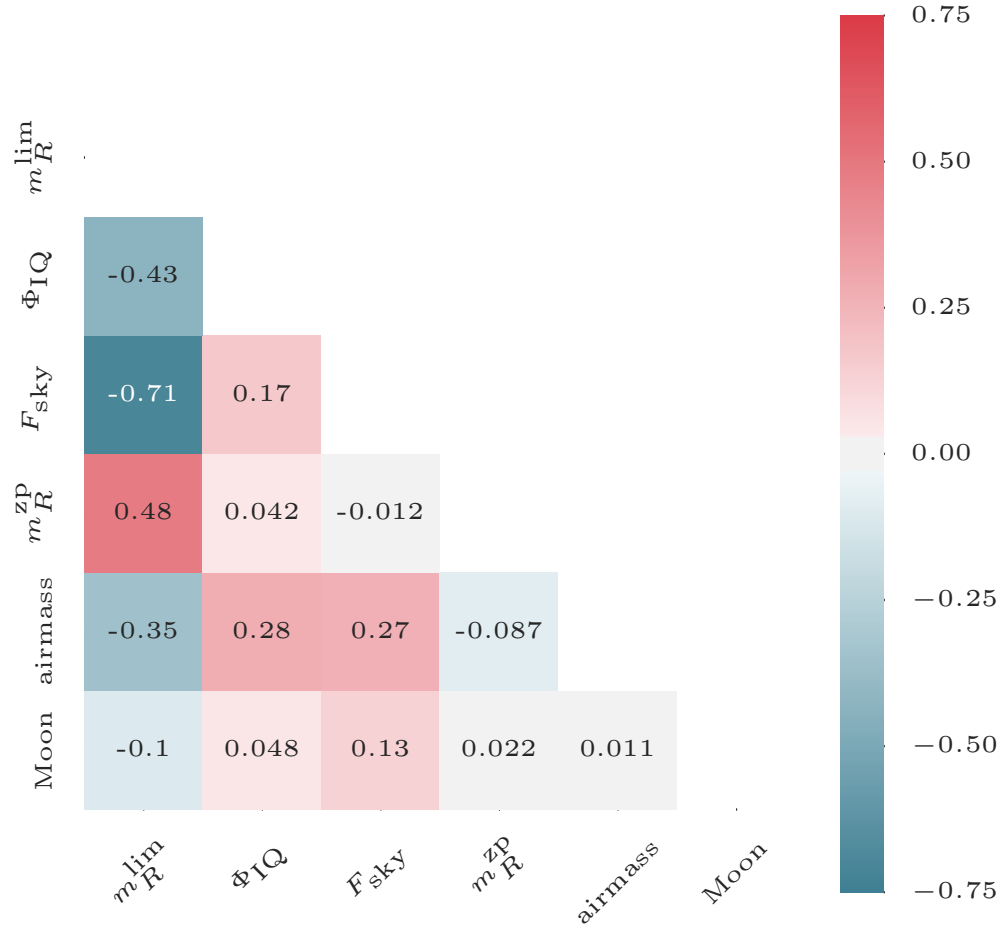
FIGURE 3.12: The correlation matrix for the observational metadata we record in our database. I explored the inter-dependencies of the parameters so that the number of dimensions in the final efficiency grid can be minimised to exclude strongly correlated parameters where no new information is gained. The values in each of the cells are the Pearson correlation coefficients.

the grid is the probability of recovering a transient as a combination of these 6 observing conditions. These parameters were binned in an identical way to the one-dimensional efficiencies as described in Chapter 3.5.1.[4]

To find the remaining dimensions with the most power, I weighted each multi-dimensional element in the grid by the inverse of the 1-dimensional detection efficiency associated with that bin for the parameter I was interested in. I then assessed the remaining 1-dimensional projections for indications of residual trends in efficiency that would indicate that there is information in that axis that was not also contained in the parameter used for the weighting. I extended this analysis to include several dimensions at once. This

---

[4]Moon illumination fraction is expressed as a value between between -1,1 (see Section 3.2 for details). For this part of the study I used the absolute value.
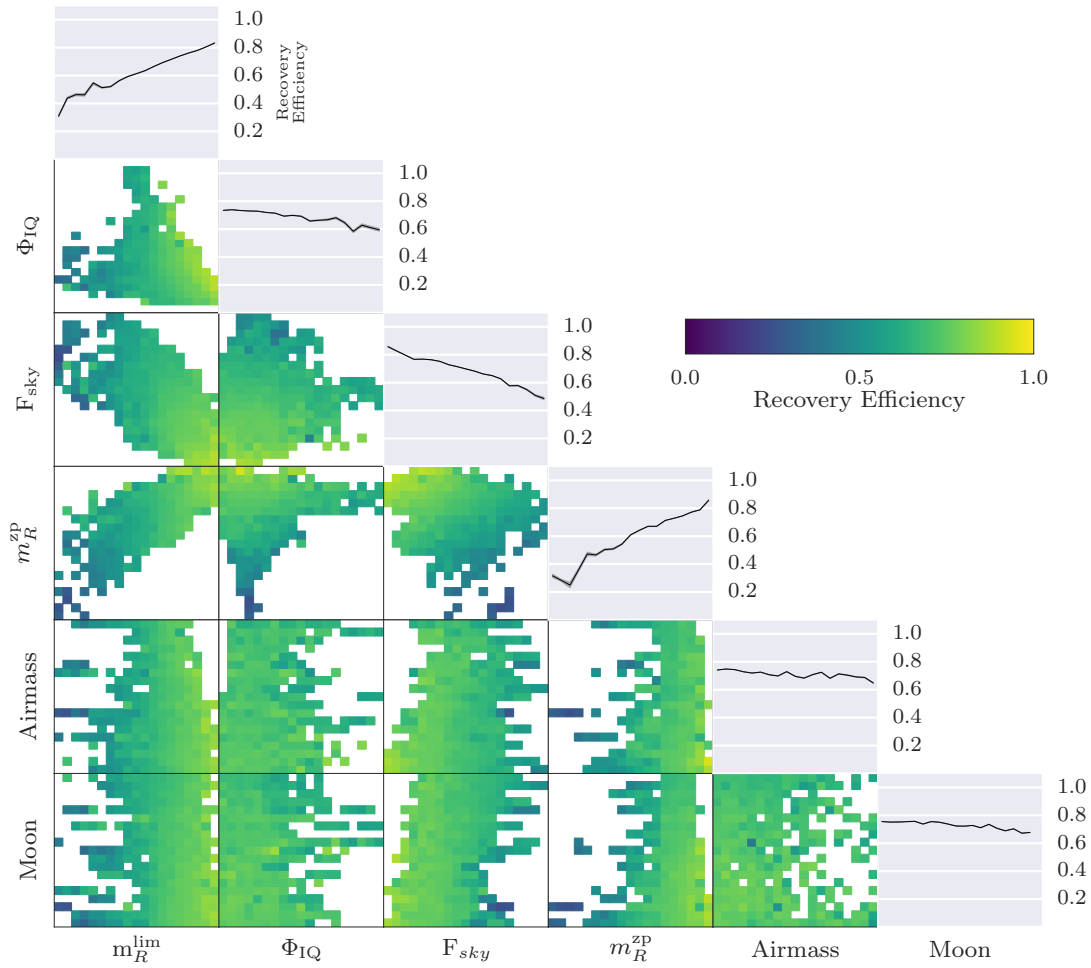
FIGURE 3.13: The 6-dimensional grid of the observing conditions metadata. The diagonal elements are the 1-dimensional recovery efficiencies ϵ, projected along the axis of interest, with the gray shading denoting the area containing 68.3% of the probability in ϵ. The off-diagonal elements represent the different combinations of all the parameters.

was performed by sequentially re-weighting different combinations of the parameters, and assessing the unweighted dimensions for residual trends. It was found, after much experimentation, that re-weighting the m_R^{lim} axis, followed by Φ_{IQ}, and then the F_{sky} axis would remove residual efficiency trends in the m_R^{zp}, airmass and moon illumination fraction parameter space. The result of this process is shown in Figure 3.14 (Note some residual trends remain with m_R^{zp}, but only at the extremes of the distribution representing poor observing conditions, presumably cloudy). Hence, an efficiency grid constructed from m_R^{lim}, Φ_{IQ}, and F_{sky} captures the vast majority of the efficiency's dependence on the observing conditions, with the fewest number of dimensions.

The final efficiency in defined by the 5 parameters; m_R, F_{box}, m_R^{lim}, Φ_{IQ}, and F_{sky} and shown in Figure 3.15. The reduced dimensionality of this final grid also allowed a finer binning of the data, increasing the resolution. The grid can then be used to estimate
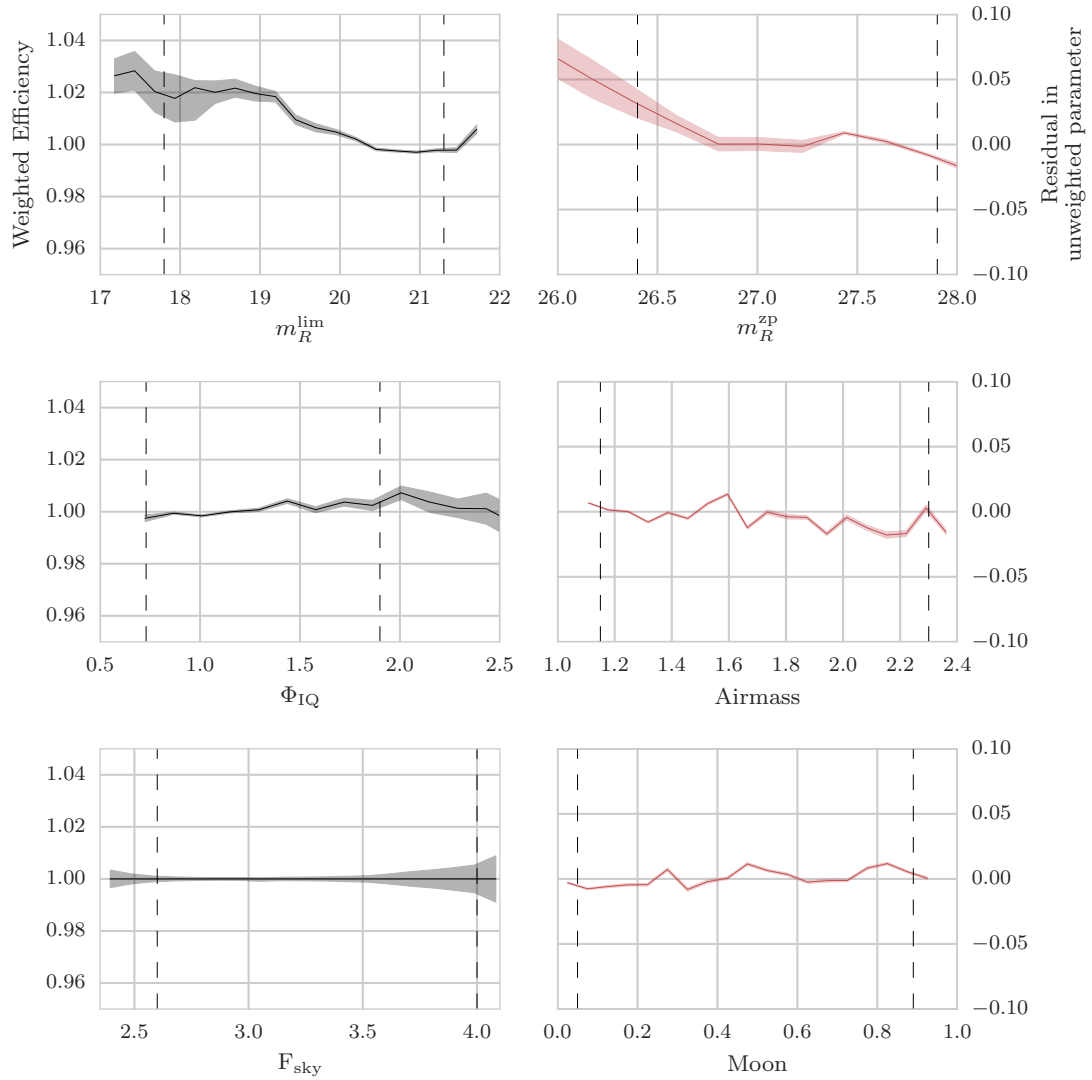
FIGURE 3.14: I explored which dimensions of the efficiency grid contain the most information. I show that by re-weighting the $m_R^{\lim}$, Φ_{IQ} and F_{sky} axis (left column), I can remove the efficiency trends in the m_R^{zp}, airmass and |moon illumination fraction| axes. I show this by plotting the residuals from a perfect recovery efficiency (right column). The dashed vertical lines represent bounds containing 99% of the data.

the recovery efficiency of a point source observed under any PTF observing conditions. This probability of a detection, given m_R, F_{box}, $m_R^{\lim}$, F_{sky}, and Φ_{IQ}, was calculated using a linear interpolation on the final efficiency grid.

3.7 Additional Applications of this method

In 2016 I was asked to adapt the code I developed for this stage of the project to accept the IPAC images of the PTF pipeline. Recently, there have been suggestions of a 9th Planet (P9) in the Solar System (Batygin & Brown, 2016) supported by numerical
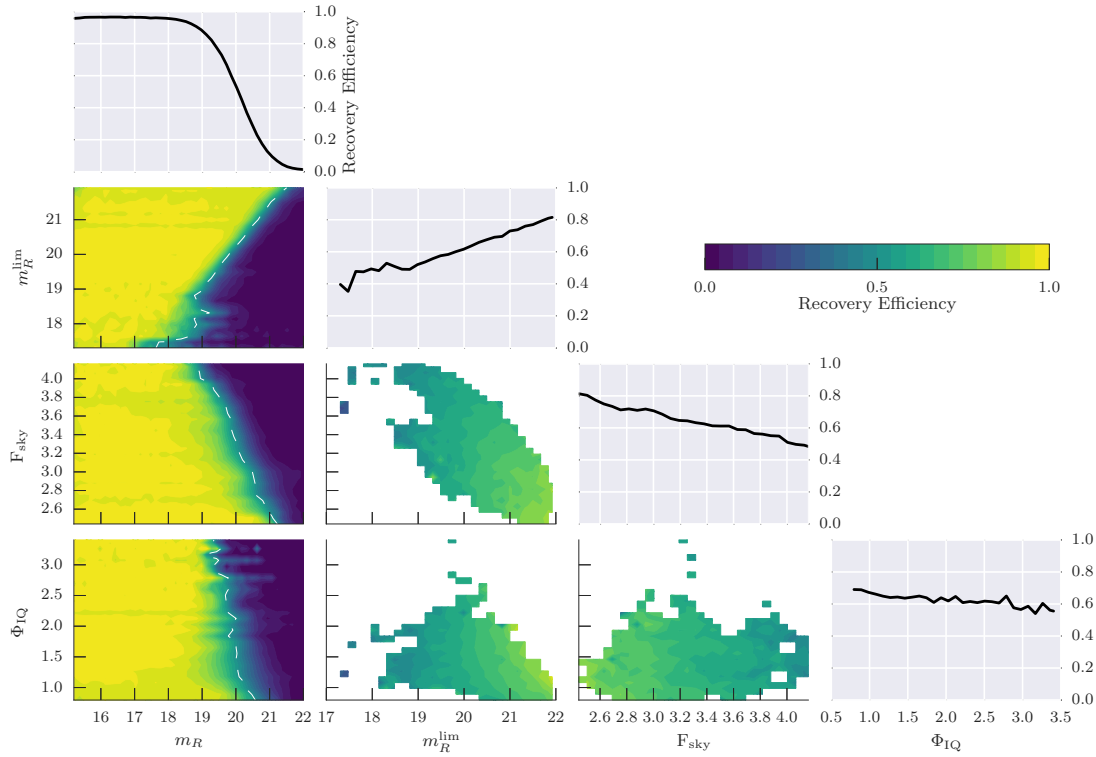
FIGURE 3.15: The final multi-dimensional efficiency grid. The off diagonal entries show the two dimensional efficiencies for combinations of the parameters. The diagonal entries show the one dimensional recovery efficiencies created by marginalising other the other grid parameters. The white dashed lines on the m_R axis denote the 50% recovery efficiency for this parameter against the other observing condition parameters.

simulations derived from observations of Kuiper Belt Objects (KBOs). The apparent perturbations of KBOs can be explained if there exists a massive body, $\geq 10 M_\oplus$ at 700AU from the sun on a highly elliptical orbit, $e = 0.6$. To-date, this theoretical body has never been observed, even recent attempts by a BBC Star Gazing Live citizen science project using data from the Skymapper project (Scalzo et al., 2017) have come up empty.

Some members of the PTF collaboration have expressed a desire to search through archival data for Planet 9, and to do so they need to understand detection limits in single and stacked observations. As per their request I modified my pipeline to generate fake point sources at a desired magnitude and to be placed in the image at a specified RA and Dec.. By stacking the images containing the fakes and running forced photometry at the location of each fake, detection limits can be set for each of the fields the P9 hunt is taking place in. Additionally, if automated pipelines are used for the search then a knowledge of the detection efficiency will also be required. Ultimately, it is unlikely PTF will discover P9, but it will be able to exclude search areas down to well constrained magnitude limits. This work is still on-going.

3.8   Summary

In this Chapter I have performed an analysis of the PTF single-epoch recovery efficiency. This work forms the foundation for all later efficiencies analysis and is an essential stage of determining the recovery efficiency for any transient class. Furthermore, I have taken this analysis beyond that performed by similar sky surveys and considered efficiencies as a multidimensional function of several observational parameters in addition to the point source properties.

To achieve such rigour in the efficiencies analysis I simulated $\sim 7 \times 10^6$ artificial point sources (fakes) into real observational data. The fakes were created from stars in the images so that they carried the intrinsic PSF variation of the image. Additionally the fakes were preferentially inserted into galaxies so variations in the immediate local environment could be studied. This local environment measurement was captured in a new metric of my own creation called F_{box} - the background subtracted sum of the counts in a 3x3 pixel box centred on the eventual position of the fake.

All the information on the fake and the observational metadata were stored in a database. The faked images were then sent through a clone of the PTF real-time transient detection pipeline which performed image subtraction and source detection. Additionally, the detected sources were analysed using the PTF machine learning algorithms to assign a 'Real-Bogus' score to assign a confidence that the object was astrophysically real. The output of the detections and machine learning were stored in a second database.

To assess the detection efficiencies the two databases were compared. The detection efficiency is defined as the ratio of the number of objects, in each bin, successfully passing the detection criterion to the number of objects simulated in that bin. I looked at 10 different 1-dimensional metrics to determine the detection efficiency and reduced these to the 5 most powerful which ultimately formed a multidimensional efficiency grid. Any combination of these parameters describes the probability of a single-epoch event passing the detection criteria and hence being entered into the PTF candidates database.

In the next Chapter I will utilise these efficiency grids in conjunction with a model of the most well studied supernova class, a Type Ia. Each epoch on a model, idealised, light curve can be thought of as existing within the efficiency grid created in this Chapter. Therefore some epochs will have a low probability of detection and others will have a high probability. A Monte-Carlo simulation of these light curves and mutually performed selection criteria between real and simulated objects transforms the single epoch efficiencies to transient class efficiencies.

Chapter 4

Simulating Type Ia Light curves

The goal of this chapter is to determine how well PTF performs at recovering SNe Ia as a function of the light curve properties during different epochs of the survey. The objective of this thesis is to quantify the true number of SNe Ia that exploded in a given volume and time period. This can only be achieved by understanding why some objects were missed. Simulating SNe light curves allows for a controlled population of transients to enter the survey. Repeating the survey, under identical conditions to the real operation, and 'observing' the simulated light curves provides insight into whether the SNe would have been found or not. Ultimately, real supernovae can then be compared to the simulations to see how likely its own discovery was. This provides the sought after weighting.

When the artificial light curves are placed into the survey, the actual night's conditions for each epoch needs to be replicated. This is where the work of Chapter 3 enters. A single-epoch on the artificial light curve is either detected or not according to statistical properties I previously derived.

To build up a good understanding of the properties contributing to the detectability of SNe, many millions of light curves were simulated. These simulations were performed across large swathes of the sky and interspersed within the duration of the 5DC experiment. The single-epoch efficiencies were applied to each epoch on a simulated light curve and used to determine whether that epoch should be kept or discarded. After this culling, the population of simulated objects were identical to a real sample that PTF would have recovered.

Additionally, the properties of the simulated objects which do not pass cuts provide an insight into the performance of PTF. Real transients are missed for a variety of reasons that are typically an interconnected result of the intrinsic supernova properties and

the survey infrastructure. For example, transient characteristic timescale and survey cadence, or transient brightness and survey mirror size both link supernova properties to the survey hardware.

Moreover, stochastic phenomena, such as the atmospheric conditions, also adversely affect performance and are just as essential to consider. Modelling light curves and observing them under these real, replicated, conditions creates a metric that is able to account for these effects and quantify the performance of the survey.

Figure 4.1 highlights these challenges, where I have demonstrated the concept behind simulating light curves with consideration of the observing conditions. The figure shows ten artificial SNe Ia I simulated at a randomly chosen spatial location in 2010. The coloured crosses show epochs when that particular metric was a 3σ outlier, which results in a low efficiency and hence an increased likelihood of the transient being missed on that epoch. The figure demonstrates the necessity of considering more than just the transient brightness, as some of the objects were poorly sampled and therefore much more dependent on the conditions they were observed in. For example, the left-most orange light curve was observed 4 times on the rise; however, one of those observations was taken in poor conditions likely resulting in a non-detection. Therefore the observing conditions can be more dominant factors in the detectability than the object brightness. Furthermore, the period around 2010-06-02 had a high cadence where a fainter object's light curve, in purple, was observed. Each faint observation has a lower probability of detection, but since the sampling density was so high the overall probability of passing quality cuts is increased. We are starting to build a picture of the importance of considering a global picture of efficiency which couples the transient properties and the survey infrastructure.

In this Chapter I discuss the simulations of the well-studied class of supernovae, Type Ia. I first provide details of the templates I use to construct model light curves, followed by a brief introduction to the Python library I make extensive use of, `sncosmo`. I then discuss my pipeline to recreate the PTF observing strategy and an artificial night sky with the statistical properties derived from the single-epoch efficiencies. Finally, from these simulations I will derive type Ia specific recovery efficiency grids so that in later chapters the real supernovae can be correctly weighted for rate calculations.

The work of this Chapter has been published in Frohmaier et al. (2017) as a demonstration application of the simulation method.
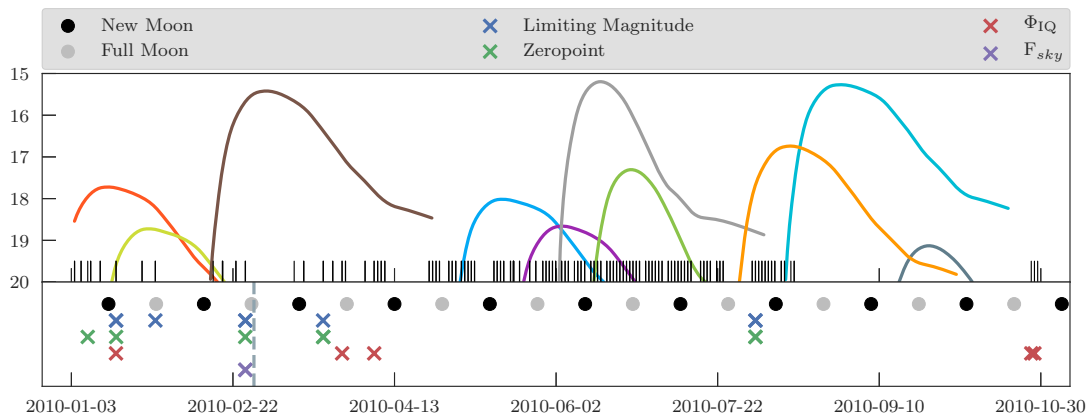
FIGURE 4.1: An example of a cadence plot from a period in 2010. The upper plot has 10 simulated SNe Ia light curves and the black dashes show epochs where PTF would have observed the field. The lower panel shows epochs where the observing conditions, denoted by the crosses, were 3σ outliers and therefore would most likely result in the event being missed (see Chapter 3.6). The figure justifies quantifying the survey efficiency by considering the transient properties, observing conditions and survey strategy as coupled elements.

4.1 Modelling Type Ia Light curves

I use the Spectral Adaptive Light Curve Template (SALT2, Guy et al., 2005, 2007) to generate light curves (I use SALT 2.4, but refer to it as SALT2 throughout this work). SALT2 is an empirically derived model which produces a spectral energy distribution (SED) time-series for SNe Ia. Integrating the SED through a filter transmission function converts the SED into magnitude- or flux-space. SALT2 can model SNe Ia spectroscopic evolution for a duration of 70 days from explosion, with around 20 days before peak brightness to 50 days after. A 40 days time-series for a typical SN Ia is shown in Figure 4.2

The SALT2 model takes several parameters to describe a light curve through its rest-frame shape, x_1 and color, C. x_1 is analogous to the light curve 'stretch'(see Perlmutter et al. (1997a); Guy et al. (2007)) and C, is effectively an estimator of $B-V$ colour at t_0, the time of maximum brightness. This colour term is a catch-all parameter which includes both intrinsic colour of the supernova and extinction from host galaxy dust. An additional parameter, x_0, controls the absolute flux scaling of the SED, which for my purpose, is used to transform into an absolute magnitude. All epochs in the SALT2 model are defined with reference to t_0, such that t=-10 is 10 days before maximum light and t=10 is 10 days post-peak. Finally, since the SALT2 model only describes SNe Ia evolution from t=-20 days to t=+50 days, all simulations and fits to real data only include epochs within this time frame.
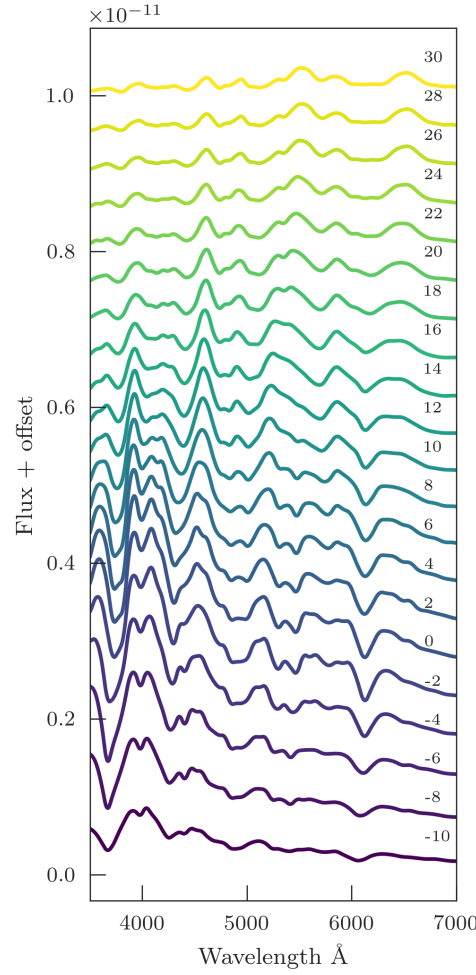
FIGURE 4.2: The SALT2 spectral time-series for a typical Ia from -10 to 30 days from peak.

Each normal SN Ia then has an absolute corrected magnitude $\mathcal{M}_{\mathrm{corr}}$ calculated through a stretch and colour correction from

$$\mathcal{M}_{\mathrm{corr}} = \mathcal{M} - \alpha x_1 + \beta C \qquad (4.1)$$

where $\mathcal{M}$ is the measured absolute magnitude, α and β are 'nuisance parameters' defining the x_1–luminosity and colour–luminosity relations which I take from Betoule et al. (2014) as $\alpha = 0.141$ and $\beta = 3.101$.

However, this functional form for correcting measured light curves can also be used to create synthetic light curves with only minor modifications. We can set an absolute B-band magnitude for an object in a population to be

$$M_B = -19.05 - \alpha x_1 + \beta C + \sigma_{\text{int}}. \tag{4.2}$$

where -19.05 is a typical uncorrected M_B for a SNe Ia. Including the term σ_{int}, which is a value randomly drawn from the normal distribution, adds variation to each object. This ensures that a population of artificial SNe show the same scatter in M_B as a real population after light curve corrections. The parameters for a set of cosmologically useful[1] Type Ia are shown in Table 4.1.

Parameter	Range
x_1	-3.0 to 3.0
Colour (C)	-0.3 to 0.3
Intrinsic dispersion (σ_{int})	Normal($\mu = 0, \sigma = 0.15$)

TABLE 4.1: SALT2 SN Ia model parameters used in the light curve simulations.

4.1.1 Bringing colour to a single band survey

PTF operated predominantly in R, with 83% of the observations in this filter. In Chapter 3.6 only R-specific efficiency grids were produced. This means that for the vast majority of the transients in PTF it is impossible to measure the colour evolution from the real time light curves. Whilst I must include colour for my light curve simulations from Equation 4.2, I cannot compare real PTF transients to this metric as it is not possible to measure it.

Parameters that go into the construction of light curves and which are measurable for real transients, such as x_1 or redshift, occupy a unique location within that parameter's axis on a multidimensional efficiency grid. Therefore, the underlying distribution of that metric is irrelevant; unphysical regions of the parameter space will never be sampled. Other than computational efficiency, it doesn't matter if those objects are simulated. However, if a simulation parameter cannot be measured, then the underlying distribution *must* be representative of a true population.

To illustrate this point we consider the example of simulating a population of 'red' supernovae, which are at the extreme, or beyond, the limits of those listed in Table 4.1 (i.e. $C \geq 0.2$). Redder supernovae are the intrinsically dimmer and faster evolving (Riess et al., 1996) objects in the SN Ia family. Even without an in-depth analysis of the efficiencies, we can intuitively conclude from my earlier work (Figure 3.8, top left panel) that these objects will be harder to recover. Therefore the mean efficiency for

---

[1] Objects that would be put on a Hubble diagram and used to constrain cosmological parameters

this population, marginalised over the all the measurable parameters, would be lower than for a more representative population.

To ensure my simulations are representative of a cosmologically useful population of SNe Ia, I used the SALT2 model fit parameters from the Betoule et al. (2014) sample. To their sample I fit a skewed-Gaussian of the form

$$f(C; A, \mu, \sigma, \gamma) = \frac{A}{\sigma\sqrt{2\pi}} e^{\frac{-(C-\mu)^2}{2\sigma^2}} \left\{ 1 + \text{erf}\left[\frac{\gamma(C-\mu)}{\sigma\sqrt{2}}\right] \right\} \qquad (4.3)$$

where C is the SALT2 colour, erf is the error function, A is the amplitude and the nuisance parameters μ, σ and γ are the location, scale and shape descriptors respectively. The resultant fit is shown in Figure 4.3 and the best-fit model parameters are displayed in Table 4.2.
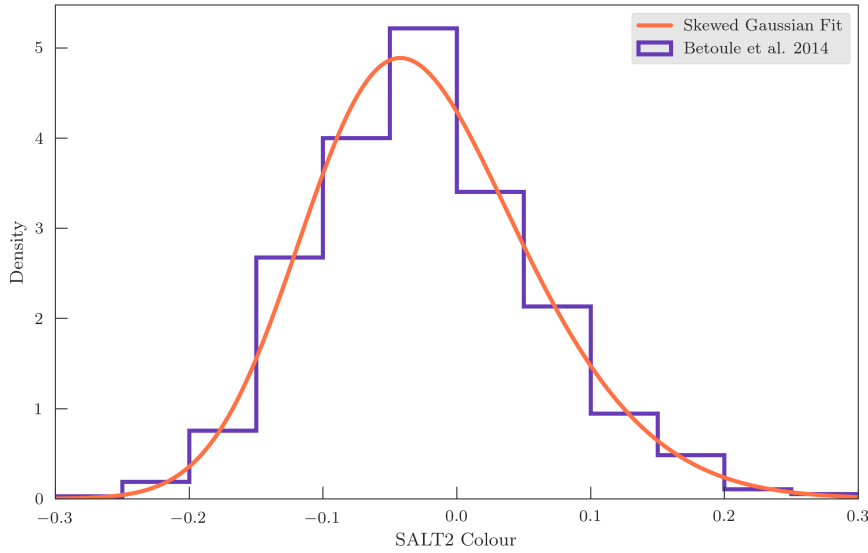
FIGURE 4.3: The colour distribution of the SNe Ia from Betoule et al. (2014) and the fitted skewed Gaussian. When I simulate light curves in later parts of this Chapter I draw the colour properties from this distribution. The underlying distribution of colour remains representative of a true population of SNe Ia.

Parameter	Value
A	0.994
μ	-0.103
γ	1.68
σ	0.115

TABLE 4.2: The best-fit parameters for the skewed Gaussian function from Equation 4.3 to the cosmologically useful sample of SNe Ia of Betoule et al. (2014).

From these results I obtained a function that describes the colour properties of a population of SNe Ia. Therefore, my later simulations of light curves always draw their colour

properties randomly from this distribution. If the later efficiency grids are marginalised over the other simulation parameters then the unseen colour distribution replicates the one I set here.

4.1.2 `sncosmo`

The python package `sncosmo` (Barbary, 2014) is a library for supernova cosmological analysis. It provides a wrapper around the SALT2 templates to not only generate synthetic spectra and light curves, but to also fit observed light curves. I added into `sncosmo` the R_{P48} filter transmission function (Figure 2.1) so that each spectrum created by the SALT2 templates could be converted into an observed magnitude by the P48 telescope. Additionally, cosmological corrections to the SED were applied through a user defined redshift. This scaled the flux of the supernovae and also shifted the SED by the appropriate 1+z amount, which handles the necessary k-corrections. As I used cosmological distances within these parameters I also defined a back-end cosmology. Throughout this work I make calculations under an assumption of a flat ΛCDM universe with H_0=70, Ω_M=0.3, and Ω_Λ=0.7.

`sncosmo` can also include observational effects such as Milky Way (MW) dust extinction or host galaxy extinction. We include MW extinction as measured by Schlegel et al. (1998).

`sncosmo` takes an array of epochs with a defined t_0 and creates the template SED for each of them. The SED is integrated through the desired bandpass, returning synthetic fluxes and magnitudes for each epoch and in each filter. This is considered a synthetic light curve. I will next utilise the efficiency grids from Chapter 3.6 to probabilistically cull epochs from the light curve which PTF may not have detected. The result of simulating hundreds of millions of light curves was a statistically representative population of artificial supernovae. I will discuss my methodology over the remaining parts of this Chapter and present my findings at the end. In Figure 4.4 I demonstrate the variation in B-band light curves for a population of random SNe Ia realisations from the parametrisations I discussed earlier.

4.2 Reconstructing PTF to simulate the Survey

PTF stores all the image metadata in a database which I refer to as the 'observing log'; accessible through basic SQL queries and spatially indexed so that efficient coordinate searches can be performed. Each stage of my pipeline will be described here and the final
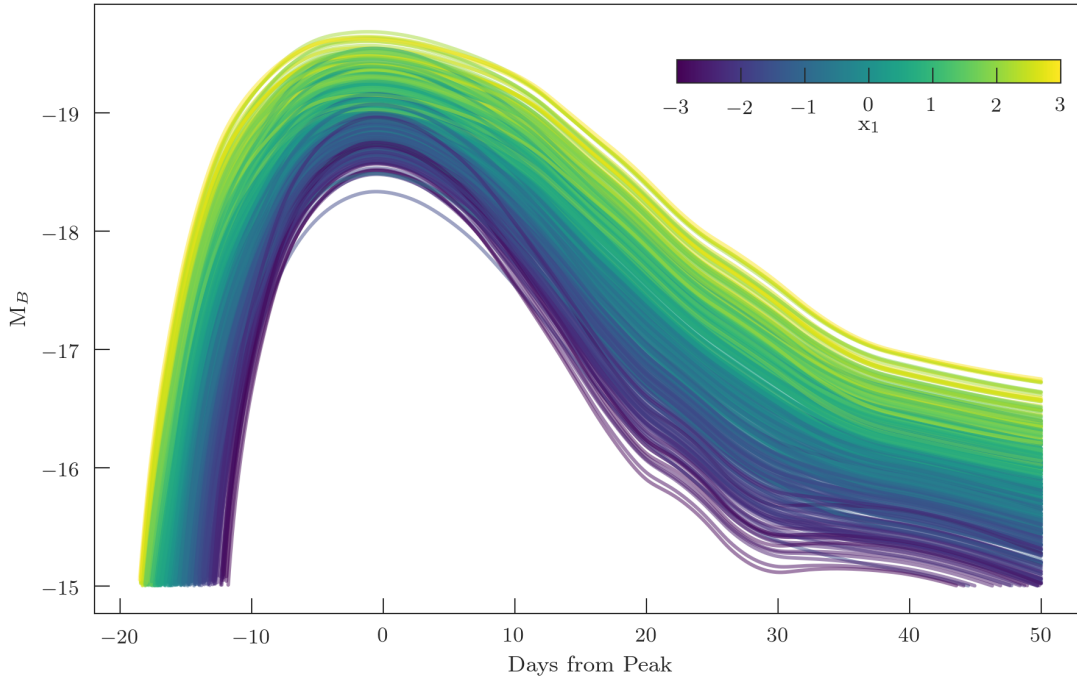
FIGURE 4.4: This Figure shows a randomly generated population of synthetic light curves generated from the SALT2 model parameters in Betoule et al. (2014). When the empirically derived light curve corrections of Equation 4.1 are applied to this sample their absolute magnitudes are centred on $M_B = -19.05$ with a scatter of 0.15 mag, just like a true population of normal SNe Ia.

stage of transient specific efficiencies will be presented in Section 4.4. The explanation of the pipeline follows the flow chart in Figure 4.5. I will sequentially explain the each of the steps from the initial randomised location and SALT2 parameters, following onto the PTF Observing Log stage and onwards to the entering of the data into the simulation output database. Throughout the pipeline explanation in this section, a light curve is used to explain or demonstrate concepts. This light curve features in Figures 4.5,4.8 & 4.9. It is a $x_1 = 0$, C=0, z=0.075 and M_R=-19.1 SN Ia.

4.2.1 The SN parameters

A random right ascension and declination for each simulated object is drawn from within an area defined by a box enclosing a well-observed region of the sky. Each simulation box is described in Section 4.3 and shown in Figure 4.10. The R.A. and Dec. were sampled from a uniform spherical distribution. The random redshifts were drawn uniformly from 0 to 0.1. Finally, the x_1 parameter was drawn from a uniform distribution between -3 to 3. σ_{int} from a normal distribution, $\mu = 0, \sigma = 0.15$. Colour, C, came from the distribution described in Chapter 4.1.
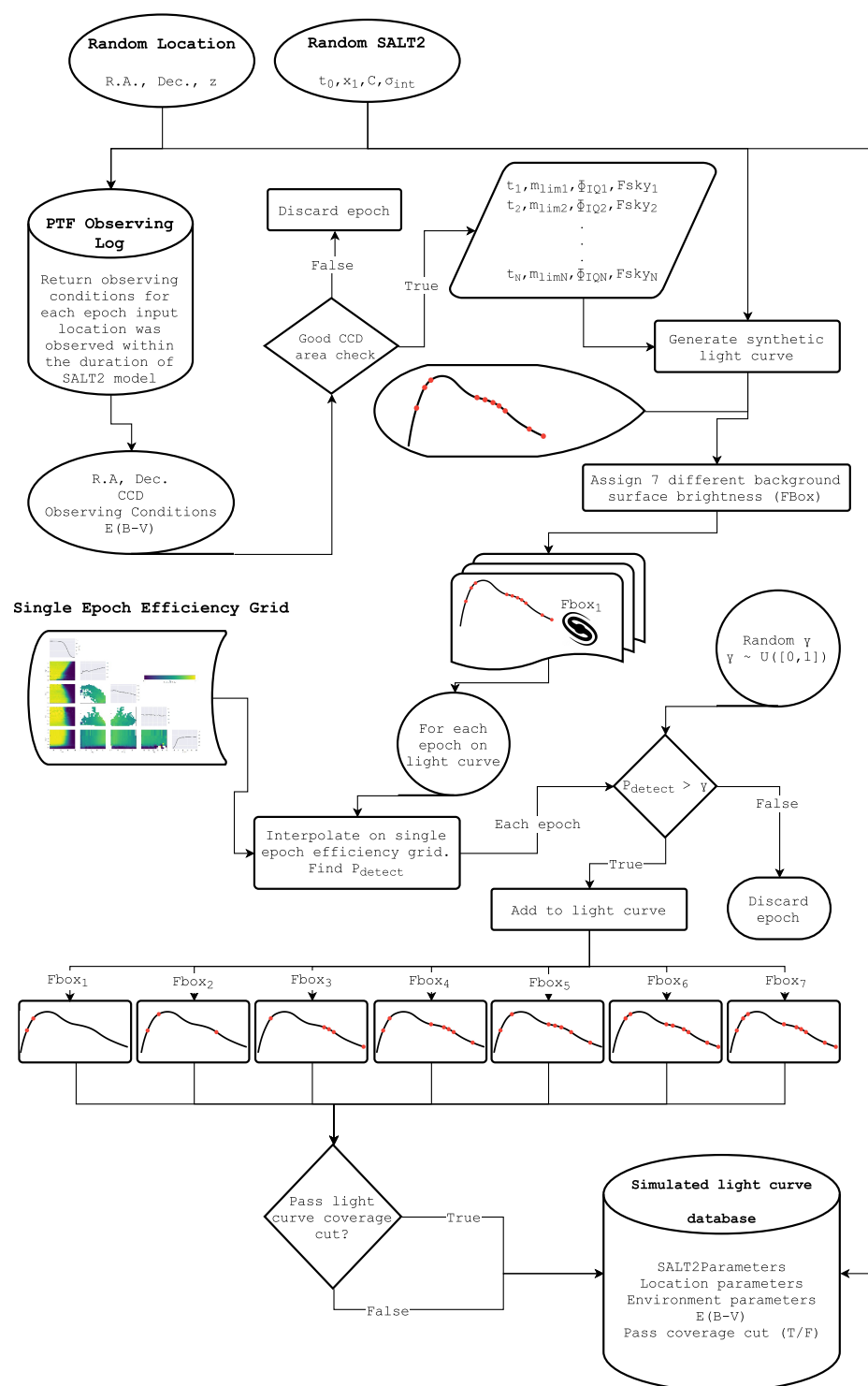
FIGURE 4.5: A flowchart of the pipeline for simulating SNe Ia light curves. The stages of the pipeline are described in Section 4.2. The only user inputs are the spatial and temporal boundaries for the simulation footprint. The pipeline runs completely in parallel on modern HPC architecture and was able to maximise the allocated resources on the Iridis4 super computer at the University of Southampton. It ran on up to 512 processing cores and made more than 1,000 simultaneous connections to the observing log database. It was able to generate 4.2 million light curves per hour.

4.2.2 Observing Log Database Query

Using the randomly generated R.A. and Dec., a query was made on the spatially indexed database containing the observing logs. The query searches all CCDs that observed the specified location within the valid range of the SALT2 model, just like the example query in Chapter 2.4.1.2. PTF returned to each field after 45 minutes for a second exposure, and both are returned here, hence later plots show two points, almost coincident, each night. The query also returned metadata used in later checks, including the observing conditions. Each epoch had its own value for $m_R^{\lim}$, Φ_{IQ} and F_{sky} and these were carried along the pipeline to be used in the single epoch recovery efficiencies.

4.2.3 CCD Defects

The second stage in the pipeline analysed each epoch and checked to see if the SN was observed in a 'good' region of a CCD. In addition to the defunct CCD3 there are other areas on the detector which I did not simulate. The CFH12K CCDs are well know to suffer from edge effects (Bloom et al., 2012) which affected the candidate recovery, caused an increase in spurious detections, and resulted in unreliable machine-learning scores. These edge effects are inherent to CCDs and are caused by guard rails along the edge of the wafers (Walter, 2015). Electrons held in pixels close to the edge feel a small electrostatic force, strong enough to affect the transfer of charges when the CCD is read-out. To address this I conservatively masked out the edges up to 75 pixels, shown in Figure 4.6. This decision was additionally motivated by the work in Chapter 5.3 to limit the association of spurious candidates to real light curves. I also masked out a strip in CCD0 caused by a chip defect known to also affect the machine learning scores of candidates.

These chip masks and the CCD0 defect are all defined in pixel coordinates whereas the simulated objects are in celestial coordinates. Furthermore, the observing logs do not store the full astrometric solution for each CCD, however the R.A. and Dec. for each corner pixel is recorded. I therefore applied a simple rotation matrix to map the pixel coordinate space onto the celestial coordinate space. Higher order corrections from a standard WCS solution are unnecessary on the scale of a single CCD and a further cosine correction to the Dec. component is, at worst, <1 pixel; therefore I do not make any further adjustments. After the rotation had been applied, a check was made to see if the object sat in a masked region for the CCD, and any epoch which did was discarded. The surviving epochs were sent to the next stage where a light curve and several 'host' backgrounds were constructed.
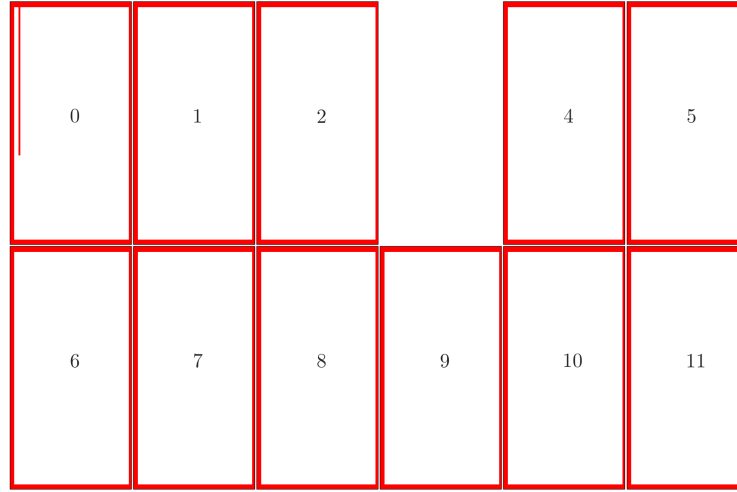
FIGURE 4.6: The CCDs on the CFH12K are not perfect. Spurious subtraction artefacts and machine learning confusion are all particularly susceptible to occurring towards the edges of a CCD. Additionally, a significant defect in CCD0 causes a strip of pixels which also induce similar problems. The regions on each CCD that are masked from both light curve simulation and light curve fitting are shaded in red.

4.2.4 Creating a light curve and assigning a Surface Brightness

At this point in the pipeline a light curve was constructed from the input parameters. This stage paired up the dates of observation, returned by the observing log, with the associated synthetic apparent R magnitude from the SALT2 model, making it possible to state that: For t_N (time) the simulated object had magnitude $(m_R)_N$ and was observed under conditions $(m_R^{\mathrm{lim}})_N$, $(\Phi_{\mathrm{IQ}})_N$ and $(F_{\mathrm{sky}})_N$, where N is a generic point on the light curve. This information is almost complete in describing the artificial PTF transient. However, as I discuss in Chapter 3.5.2.1 and show in Figure 3.9, it is important to consider the effects of the immediate surface brightness at the location of the object, a parameter I called F_{box} in flux-space or, in magnitude space, $m_R^{F_{\mathrm{box}}}$.

This introduces additional complexities. So far, all the parameters in the pipeline have either been taken directly from the observing logs, or constructed from an empirically derived and well-constrained model. The $m_R^{F_{\mathrm{box}}}$ metric requires a new measurement of the flux in a 3x3 pixel grid centred on the location of the transient, but without transient flux. The initial method considered was to use the reference image for each simulated epoch, and record the $m_R^{F_{\mathrm{box}}}$ parameter at the location of the artificial transient. However, this was very computationally demanding, being an I/O task. It was impossible to use this method and have any hope of simulating enough objects - recall that in Chapter 3.3 a similar task took ∼7s to process on each image.
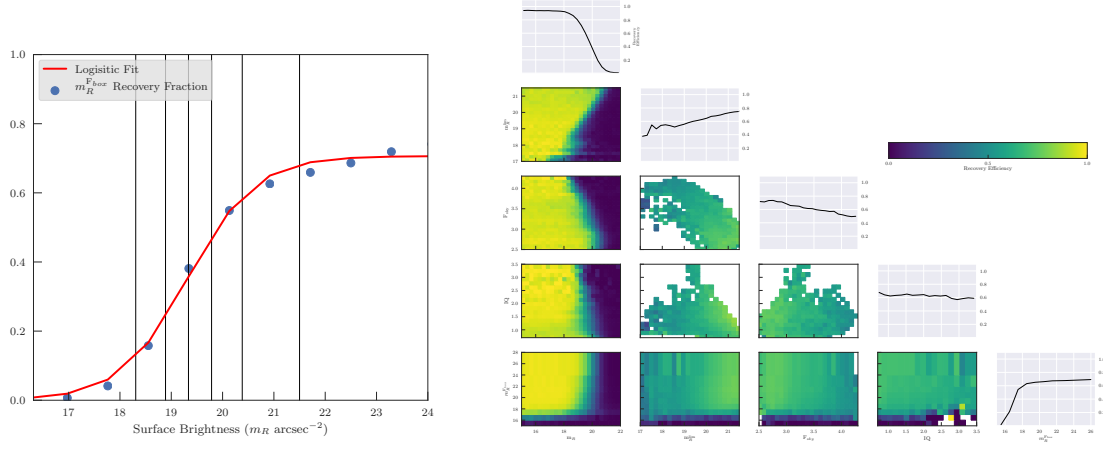
FIGURE 4.7: Left: The $m_R^{F_{\text{box}}}$ metric described in Chapter 3.5.2.1 was fit by a logistic function and split into 7 coarse bins whaich capture similar variation in the efficiency. These bins were then used to construct a new multi-dimensional efficiency grid, Right, which builds upon the grid created in Chapter 3.6 to include the $m_R^{F_{\text{box}}}$ metric. This new grid is used for the later interpolation stages

To solve this problem I created seven copies of the original light curve, and to each I assigned a randomly generated $m_R^{F_{\text{box}}}$ background. The rest of the pipeline was performed on each copy in parallel. Ultimately, the final product of this simulation was seven different efficiency grids, created from the same light curves, but each grid represents the scenarios where the event happened under different $m_R^{F_{\text{box}}}$ conditions. This solution maintained the speed of the pipeline at the expense of increasing the number of final data products. The processing time was the critical factor compared to the additional data storage requirements, which were negligible.

The $m_R^{F_{\text{box}}}$ parameter required an additional axis on the single-epoch recovery efficiency grid. Therefore, I took the 1D $m_R^{F_{\text{box}}}$ efficiency from Chapter 3.5.2.1 and fit a logistic function to the data. This function was then split into 7 different bins with each bin width capturing an equal range in efficiency. In other words, the $m_R^{F_{\text{box}}}$ bins are approximately equally spaced in efficiency and the vertical lines in Figure 4.7(left) show the bin boundaries. This coarse binning was then applied to the $m_R^{F_{\text{box}}}$ axis on a new multi-dimensional single-epoch efficiency grid. An identical method to the one I described in Chapter 3 was adopted. The new single epoch efficiency grid is shown in Figure 4.7 (right), and was adopted for the relevant part of the pipeline described in the next Section.

At this stage of the pipeline the simulated object has the most complete description I can provide within the framework of SALT2 and the efficiencies requirements. Therefore the next stage of the pipeline determined whether each epoch on the light curve would have been found.

4.2.5 Recovering light curve epochs

I now describe how my pipeline makes a probabilistic statement on the likelihood that each epoch on the light curve would have been detected by PTF. Each point on the light curve is associated with several key parameters defined in Chapter 3.2. These are: m_R m_R^{lim} Φ_{IQ} F_{sky} and, $m_R^{F_{\text{box}}}$.

These parameters define a location within the single-epoch, multi-dimensional recovery efficiency grid shown in Figure 4.7 (right). The value of them at that location is the fraction of the simulated fakes, with the same properties as the input epoch, recovered. That is to say, it describes the probability of the event being both detected and passing the machine learning cuts in Chapter 2.2.3. As the efficiency grid has a discrete binning, I use multidimensional liner interpolation to find the 'True' probability.

Sequentially running all light curve points for a transient through this process gives each point on the light curve its own probability of detection (P_{detect}). For each epoch, I randomly selected a number, λ, from a uniform distribution between 0 and 1 for comparison with that epoch's detection probability: if $\lambda \leq P_{\text{detect}}$, the SN is considered detected on that epoch, and if $\lambda > P_{\text{detect}}$ the SN is considered not detected. Figure 4.8 demonstrates this concept, showing typical observational metric locations on the efficiency grid for a demonstration SN Ia. Performing this on all light curve points replicates how PTF would have observed the object.

For clarity, in the illustration of Figure 4.8 the $m_R^{F_{\text{box}}}$ axis is not shown as it does not vary over the light curve, rather. The effect on a light curve of the culling process is shown in Figure 4.9 for the demonstration object we have followed throughout this chapter. The green points show epochs which had P_{detect} determined from Figure 4.8 to be greater then λ, whilst the opposite is true for the red points. The final light curve, which is passed onto the final stage, is shown by the blue points.

4.2.6 Coverage Cuts

The penultimate stage of the pipeline was designed to make an overall assessment as to whether the simulated object was 'discovered' or not. It is difficult to replicate this in PTF due to the human intervention in the pipeline. Instead, I make a conservatively-motivated coverage requirement, with the constraint that any cuts I make on simulated objects must be replicated on the real objects. Thus, the cut had to be stringent enough so that real SNe Ia would pass with little to no contamination from objects of unknown type, but also loose enough that a large number of SNe Ia would enter a sample.

It was decided that the final cut on the light curve coverage would be:
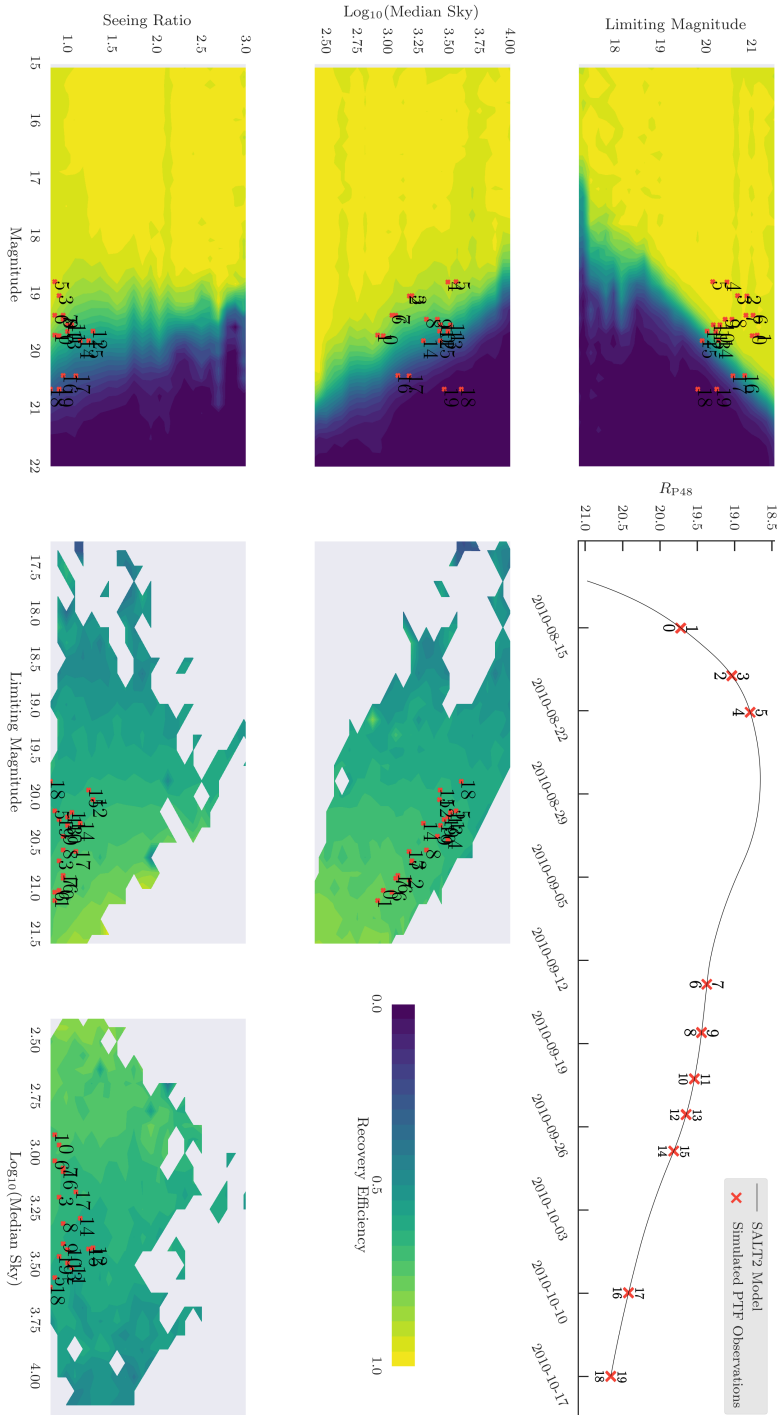
FIGURE 4.8: A demonstration of the implementation of the detection efficiency grid for an example transient. The SALT2 model is used to predict the apparent magnitude on every epoch on which PTF made an observation (top right panel) and the PTF database returns the observing conditions on that epoch. Each combination of apparent magnitude and observing conditions then has an associated efficiency P_{detect}, interpolated from the multi–dimensional efficiency grid; the figure shows the position of the point in various combinations of the grid dimensions, and the points are numerically labelled in the figure (PTF typically observes each position twice during a given night). If the value of P_{detect} is greater than or equal to a random number between 0 and 1, then that point is considered detected by PTF; otherwise the point is not considered detected. This process is repeated for each observation. The entire light curve can then be considered against appropriate selection criteria that determine the probability of whether the transient would be detected over the course of its evolution.
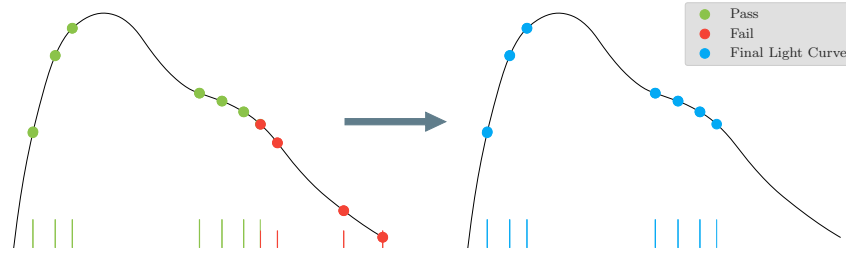
FIGURE 4.9: This figure shows a cartoon representation of how a light curve is culled. The points on the light curve are passed into the efficiency grid to find P_{detect} and compared to a random number. The points in green pass the cut and the points in red fail. The final light curve consists of the blue points and is passed to the final coverage cuts test in Section 4.2.6.

- Epochs with $m_R \leq 20$ mag are considered as points on a light curve

- At least 4 points on the light curve over the duration of the SALT2 model ($-20 \leq t_0 \leq +50$)

- At least 2 points before maximum and 2 points after maximum

- At least 12 hours between light curve points

These cuts are necessary to increase the confidence that the real sample of SNe Ia used in this study have a high purity. A less rigorous criteria would leave the SNe Ia sample open to larger amounts of contamination as it would not be possible to confidently exclude the contaminants. Furthermore, requiring two points both before and after maximum improves the ability to constrain the x_1 parameter which is used to define the SNe Ia sample.

The simulated light curve was checked to see if it met the criteria list above. If it did then the it was assigned a 'True' flag and oppositely, if it didn't, a 'False' flag. Again, the check was done in parallel for each light curve under the assumed $m_R^{F_{box}}$. All the data describing the simulated the light curves, and whether they were recovered, was stored in the PostgreSQL database. The pipeline was deployed on the Iridis4 supercomputer where I was able to simulate almost 4.2 million light curves per hour. In Section 4.4 of this Chapter I discuss how I convert the properties of my simulations into model specific efficiency grids for SNe Ia model comparisons.

4.3 Sky location and time periods

The next task is to define an area of sky and survey time period when PTF was able to maintain a regular cadence. The challenge of modelling PTF is unlike any similar study. Works such as Dilday et al. (2008, 2010); Perrett et al. (2010); Leaman et al. (2011); Graur et al. (2014), had well a defined cadence and/or a fixed area of observation.

I visually inspected each month in the 5DC experiment to find regularly observed areas of the sky with coverage of several weeks. I tried, in this regard, to be as unbiased as possible by not just drawing the smallest box around the most number of known SNe Ia. This would obviously place a bias in later calculations which perform volumetric calculations.

In Figure 4.10 I have shown a month-by-month breakdown of the sky in a tangent plane projection for the 5DC experiment. The colour scale for each month represents the number of times each field was observed compared to the most observed field. A broadly consistent colour across multiple months would correspond to a regular cadence and hence a more consistent probability of discovering a SN. Each of the boxes outlined by a different colour show the areas I chose to perform my simulations in. Other areas on the sky were explored but in many cases I chose not to pursue. The properties of the chosen areas are listed in Table 4.3.

	Box	Area (deg^2)	Nights
▪	2010 A	2521.81	170
▪	2010 B	1248.79	140
▪	2010 C	1680.52	109
▪	2010 D	611.37	149
▪	2011 A	1495.10	150
▪	2012 A	1664.89	109
▪	2012 B	918.45	78
▪	2012 C	299.97	139

TABLE 4.3: The the properties of the simulation areas shown in Figure 4.10. The stated areas are the total surveyed areas inside each box. The 'Nights' column states for what period the cadence was regular and, hence, the duration simulated.

For each of the areas traced out in Figure 4.10 I simulated at least 3×10^7 light curves following the procedure described from Chapters 4.1 onwards. Each of the 8 different areas has a set of 7 efficiency grids associated with it - one for each assumed $m_R^{F_{box}}$. The efficiency grids were stored with a unique identifier associated with each area and $m_R^{F_{box}}$. In the next Section I describe how the results of the simulations are used to construct these 56 SNe Ia efficiency grids. This is the final stage of replicating PTF before we can being an assessment of the real objects PTF discovered.
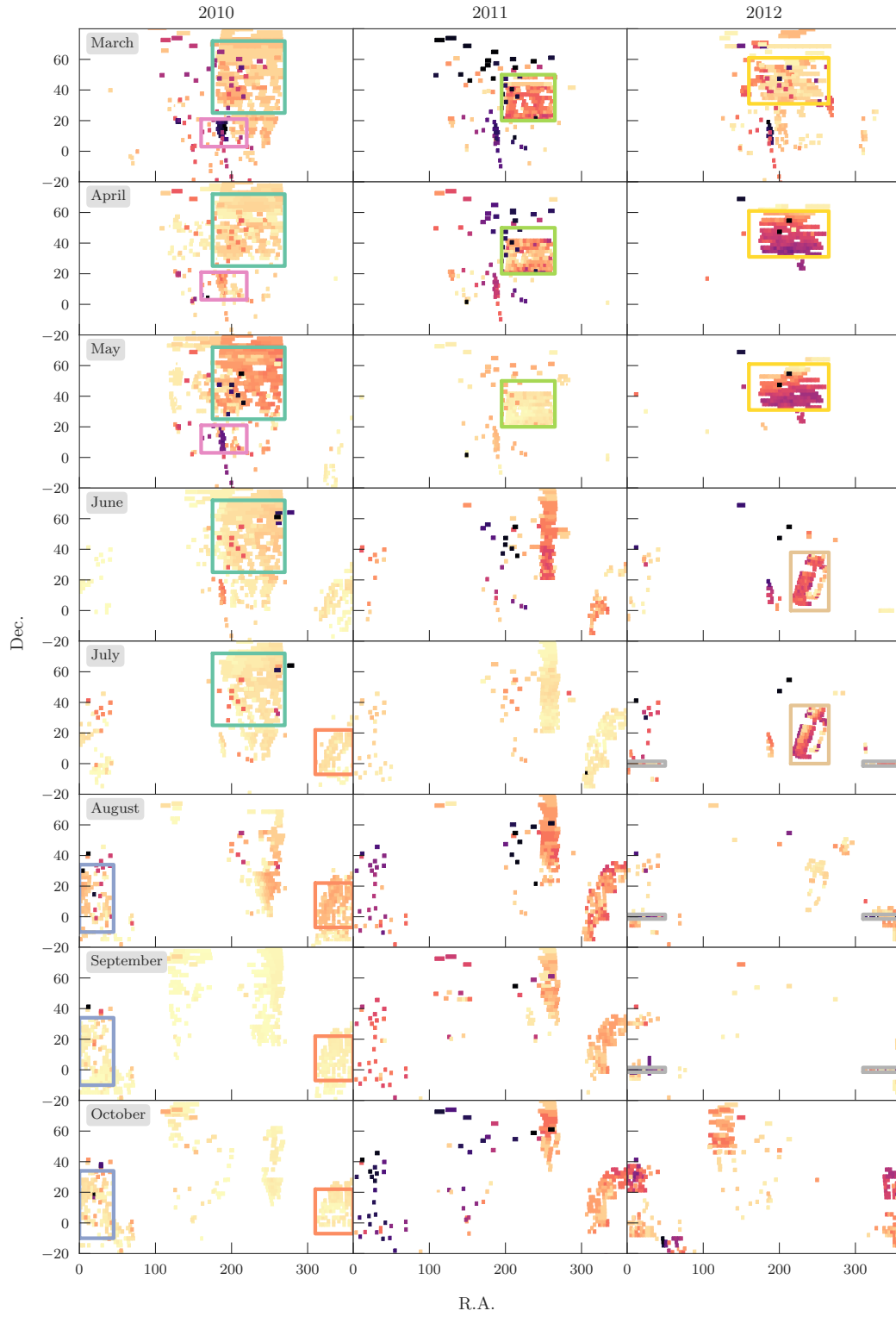
FIGURE 4.10: Each column in this plot shows a different year during the survey and the rows are the months the 5DC cadence experiment was operational. Each cell is a tangent plane project of the observing footprint of PTF. The colour scale tracks the number of observations compared to the most frequently observed that month. Each of the coloured boxes are areas on the sky I have chosen to simulate light curves in. They are areas which were consistently observed and had a good follow-up program operating at that time.

4.4 SNe Type Ia Efficiency

The efficiency grids were constructed from the SALT2, redshift and location parameters of the simulated objects. The database contains the properties for each of the simulated light curves along with a True/False flag for each $m_R^{F\text{box}}$ which states whether the objects passed the coverage cuts or not. The efficiency grid is created by binning the parameters that form each axis and calculating the fraction of objects found in each bin to the total number simulated. This process is identical in nature to the work I detailed in Chapter 3.6 for the multidimensional efficiency grids of the fake point sources. Except, here I will analyse the recovery efficiency as a function of the light curve parameters and not for the fake point sources. I will detail the specifics of the process and any uncertainties in the efficiencies will be treated under the prescription of Chapter 3.5.1.1.

The parameters that form the axes on the efficiency grid are: redshift, x_1, the absolute magnitude in R (M_R, from the synthetic spectrum), right ascension and declination. A final distinction is made for each of the surface brightness backgrounds $m_R^{F\text{box}}$. I marginalised over t_0 as the simulation areas were chosen to have a roughly constant detection efficiency in time. The R.A. and Dec. parameters were binned to approximately the resolution of a single field. The redshift, x_1, and M_R metrics did not require a fine binning as their behaviour was well behaved and showed the expected trends, discussed in the next section, therefore 20 equally spaced bins was adequate.

4.4.1 Spatial Recovery Efficiency

In this section I explore the spatial dependency of the recovery efficiency. This was done to decide whether R.A. and Dec. were important parameter and, if so, at what resolution should they be sampled.

In Figure 4.11 I show the spatial dependency for each simulation box to a resolution finer than a single CCD. It is worth noting that the malfunctioned chip, CCD3, is captured at this resolution. The absolute minimum requirement to pass the coverage cuts was 4 observations in a 70 day period. The figure shows what percentage of the time this was achieved in the light curve simulations. Even with this loose definition it is clear that there is a field-by-field variation in the efficiency. Hence supporting the requirement for R.A. and Dec. to be parameters on the efficiency grid. If the fields were found to be uniformly efficient then it wouldn't be necessary to include them.

Figure 4.12 shows the tangent plane projection of the spatial efficiencies at a resolution of approximately a field per pixel. Effects from the other grid parameters, discussed in Section 4.4.2 were also included. The black areas in each simulation are either due
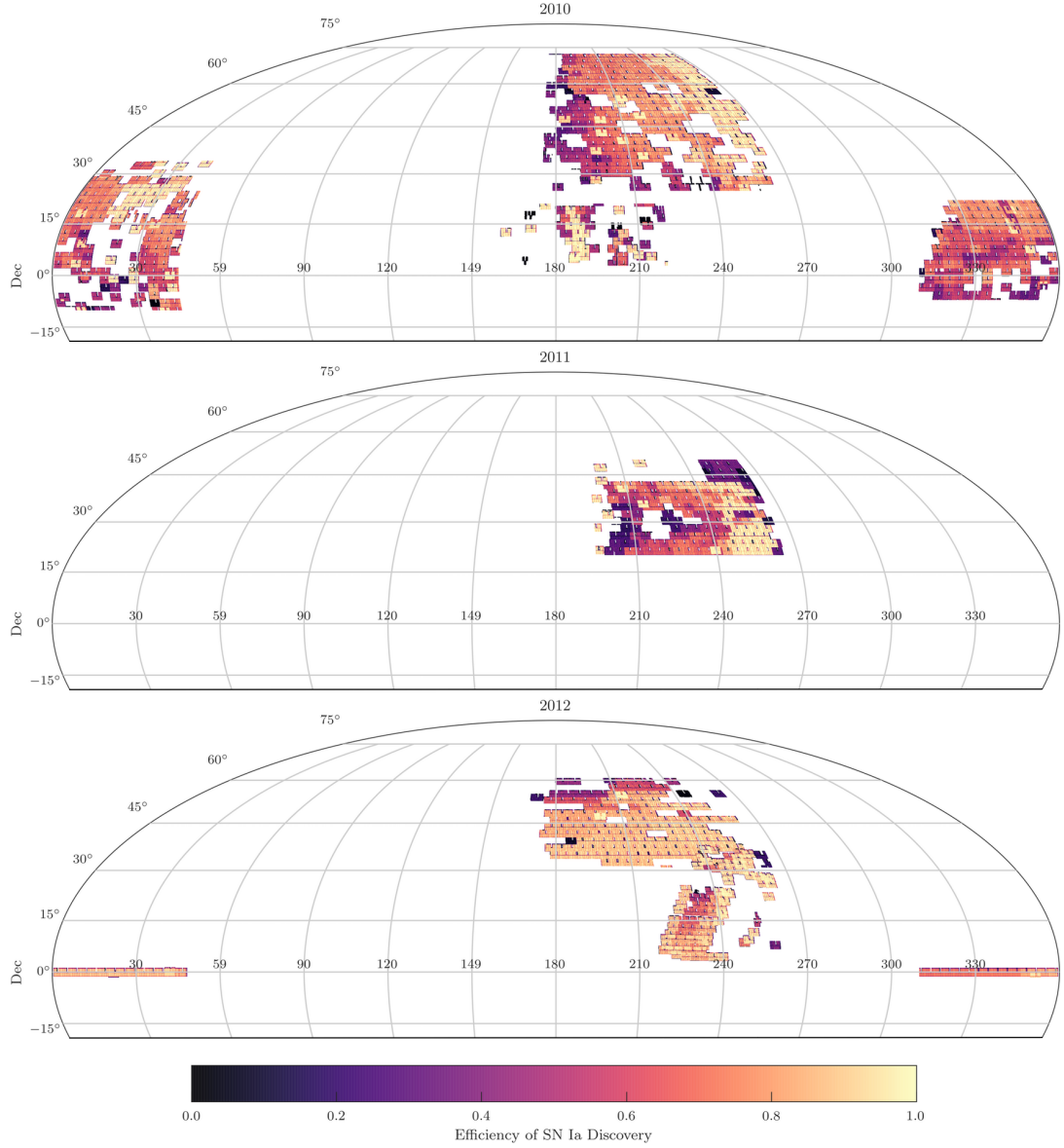
FIGURE 4.11: Many millions of points were simulated and checked to see if they pass the most basic cut of being observed at least 4 times separated by 12 hours in the time span of a SALT2 model. In this plot there is no dependence on actual SALT2 model parameters, only on cadence. This plot is an analysis of the cadence during each footprint's time span and shows not all fields are treated equal. Due to computational limitations when extra dimensions are included, the final efficiency grids are at a coarser resolution.

to gaps in the observing footprint, hence zero probability of discovering a transient, or because PTF did observe the area but could never recover a light curve which met the criteria cuts. These zero efficiency areas do not contribute to areas stated in Table 4.3.
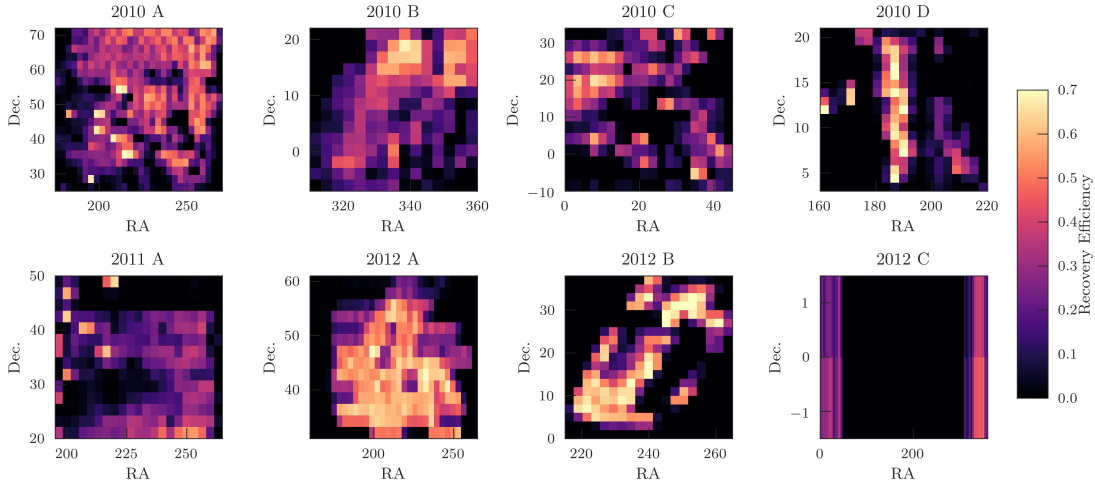
FIGURE 4.12: These plots show the final resolution of the different simulation areas. The resolution is of approximately the size of a single field and cannot be any finer. This is because a finely spaced grid in multiple dimensions looses statistical power in each bin when spread too thinly. Increasing the resolution to a CCD scale would only served to increase the number of light curves which require simulation without adding much extra useful information.

4.4.2 1D Efficiencies

In this Section I explore the 1D efficiencies for each of the different simulation boxes depicted in Figure 4.10. Unlike for the single epoch recovery efficiencies in Chapter 3.5.2, there is no single metric that can describe how well PTF found, say, a redshift 0.05 SN Ia across the whole of its operation. Fundamentally this is because the light curve simulations captured both the intrinsic supernova properties and the actual operation of the survey itself. As I have shown there are wildly different strategies for the various parts of the sky and for different years.

To find the 1D efficiencies I took the results of the simulations from Section 4.2 and studied the fraction of light curves that passed the coverage cuts (N_{rec}) in each bin to the number of objects simulated (N_{sim}) in that bin. This ratio is the definition of the recovery efficiency, ϵ_{SN},

$$\epsilon_{SN}\left(\varphi\right) = \frac{N_{rec}(\varphi)}{N_{sim}(\varphi)} \tag{4.4}$$

where φ denote the light curve and observing metrics: $R.A., Dec., z, x_1, M_R, m_R^{F_{box}}$

I analysed the 1D efficiencies for the measurable light curve parameters; Redshift, x_1 and M_R. It is clear from a visual inspection of Figure 4.13 that each of the parameter efficiencies show the expected trends. The approximately linear shift seen between the simulations was due to unseen-metrics, such as cadence or a uniform sky sampling, that were still captured in the simulation. By marginalising over the spatial parameters to produce these 1D efficiencies, I carry the mean spatial efficiency with them.

Figure 4.13 shows the expected trends: objects that are further away (higher redshift) are more difficult to recover; brighter objects are easier to discover; higher x_1 are easier to detect.
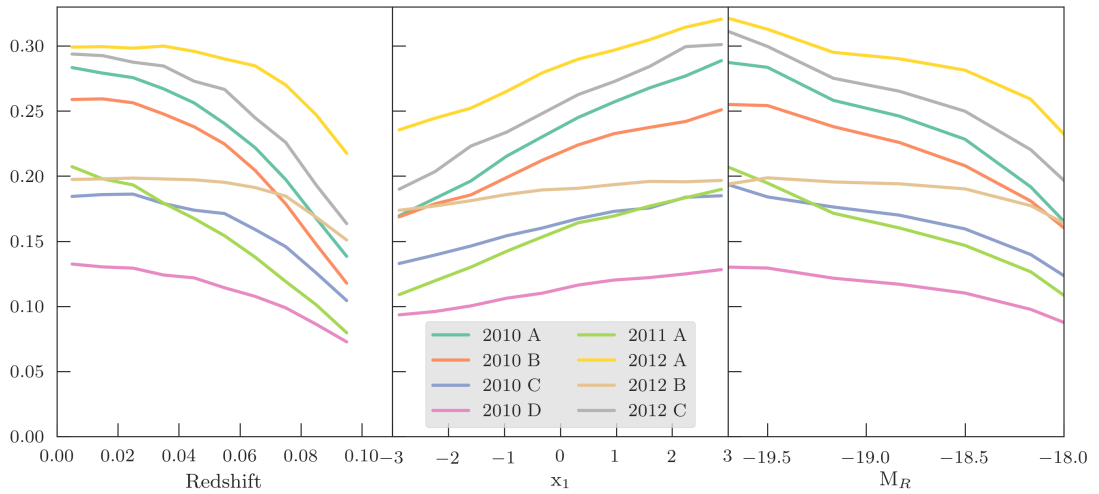
FIGURE 4.13: The 1D Efficiencies for the 3 metrics measurable from the model light curves are shown. The expected trends can be seen where higher z, faster evolving and fainter objects are found less frequently. The shift between each of the simulation boxes is due to a marginalisation over the spatial dependence of the efficiency. This shows that R.A. and Dec. need to be considered in any efficiency study.

4.4.3 Recovering Colour

In Section 4.1.1 I discussed the need to simulate a known colour distribution so that the underlying population of supernovae was representative of a true cosmological sample. Similarly, it is important that the input colour distribution was well recovered after the light curve simulations had passed through the pipeline as there were no cuts on colour.

In Figure 4.14 I show the colour distribution for the recovered simulated objects. The distribution was compared to the input distribution in Section 4.1.1. With just a visual inspection it is clear that the colour distribution was preserved throughout the pipeline and did not significantly affect the recovery efficiencies. This preservation is encouraging,

it means that final efficiency grid will maintain a representative colour distribution when real objects are compared to it.
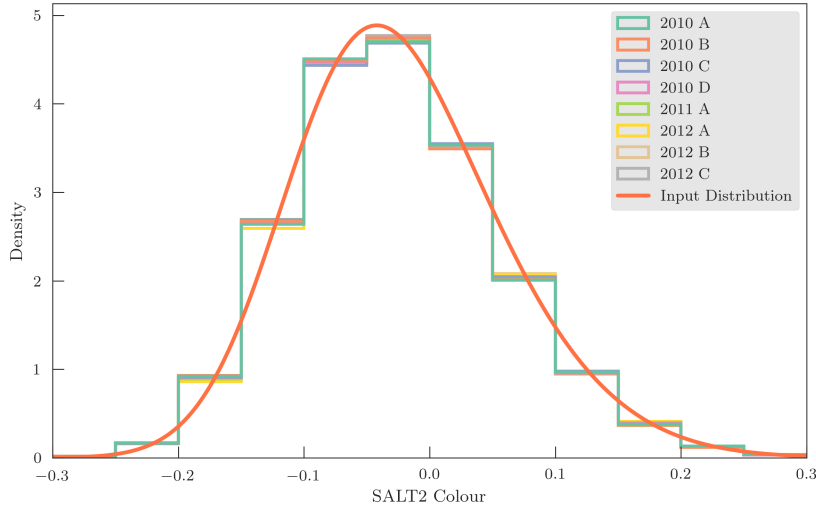
FIGURE 4.14: The histograms show the recovered colour distribution from the light curve simulation of only the objects that passed all the quality and coverage cuts. The over plotted line is the original distribution all the object's colour was drawn from. To see an identical distribution returned is encouraging as it shows that the efficiency grid maintains the underlying colour properties of a true population of SNe Type Ia. These distributions were expected to match as there were no light curve cuts made on colour as it is impossible to measure it for real PTF transients.

Now that I have established the dominant factors, in 1D, affecting the recovery of an SNe Ia light curve it is time to explore the multidimensional parameter space in much the same way I did in Chapter 3.6. These final, multidimensional efficiency grids must contain enough information to be able to fully describe the probability that PTF recovered a SNe Ia light curve as a function of its light curve properties and environment.

4.5 Multidimensional SNe Type Ia Efficiency Grids

I extend the analysis for the single parameter recovery efficiency grids and present the multidimensional efficiency grids. These grids were constructed from the parameters analysed in the previous sections and used to describe the recovery fraction of the simulated SNe Ia light curves. In principle the methodology to create these grids is identical to Chapter 3.6. Utilising a multidimensional approach to the analysis of the simulation results allows for situations to be studied which cannot be described by any single parameter - say a faint, low-stretch SNe at high redshift, observed in a field with high cadence, and hosted in a bright galaxy. Each real SNe Ia discovered by PTF occupies a unique location inside one of the relevant multidimensional efficiency grids. The value at

that location is the probability that PTF would have recovered an object similar to it. I denote this light curve recovery efficiency as, ϵ_{SN}. For example, a real transient found to have $\epsilon_{SN} = 0.2$ would mean that this 1 object represents a population of 5 objects, where PTF missed the other 4.

For each $m_R^{F_{box}}$ a separate grid was created as a function of redshift, x_1, M_R, R.A. and Dec. The parameters were binned in an identical way to the 1D analysis I described in Section 4.4.2 and used to create a multidimensional parameter space. The results of the simulation were then used to fill the space according to Equation 4.4.
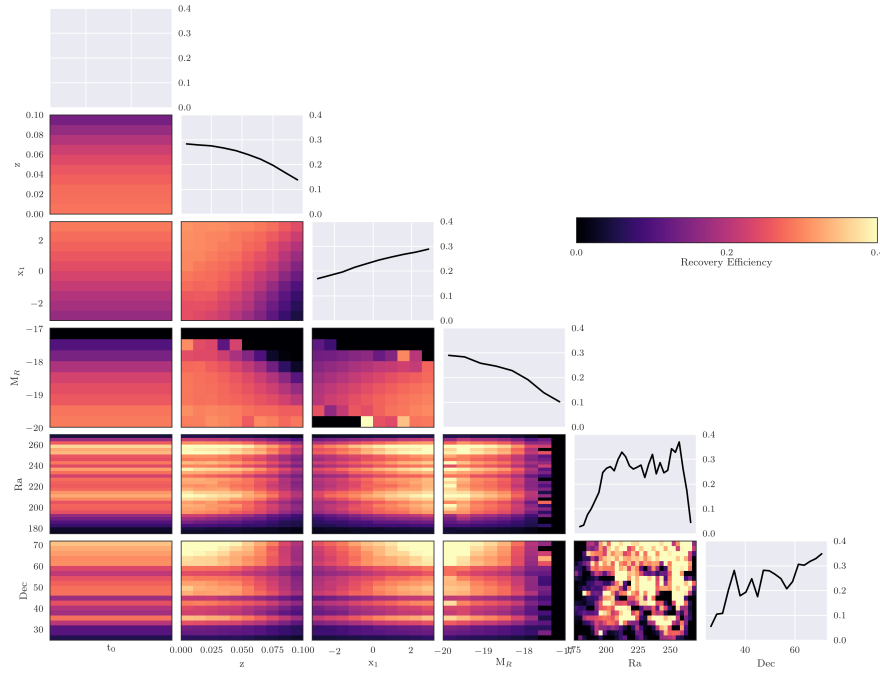
FIGURE 4.15: The final efficiency grid for the 2010 A field for the faintest $m_R^{F_{box}}$. The diagonals are the 1D efficiencies for each of the grid parameters and the off diagonals are efficiencies for combinations of those parameters. The t_0 parameter was marginalised over, so the left-most column has no variation in the efficiency.

All 56 efficiency grids are presented in Appendix A, and the 2010 A field is discussed here. Figure 4.15 shows the grid with the faintest $m_R^{F_{box}}$ i.e. the high probability of detecting the transient. The first axis is the time of peak brightness, which was marginalised over. The diagonal elements in the Figure are the 1D efficiencies, identical to those presented in Section 4.4.2. The off diagonals are constructed from combinations of the remaining grid axes. It should not be forgotten that throughout this final stage of the Chapter the different $m_R^{F_{box}}$ grids are produced in parallel and treated identically. Each $m_R^{F_{box}}$ for the simulation boxes is produced from the same population of simulated light curves but with a different underlying assumption on the immediate background surface brightness. The treatment of these additional grids is schematically shown in Figure 4.16

where each simulation box can be thought of as a stack of grids where each layer in the stack represents the $m_R^{F_{\mathrm{box}}}$ bins.
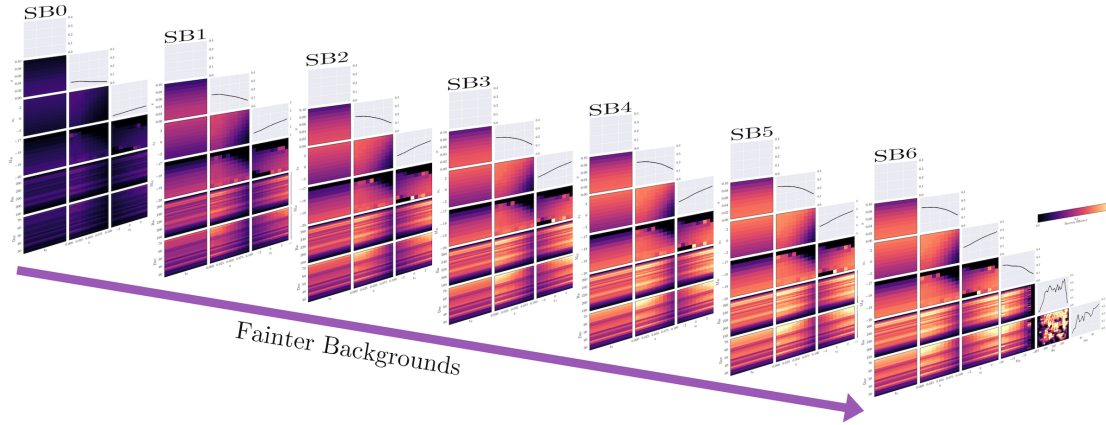
FIGURE 4.16: Each of the light curves was simulated with 7 different assumed backgrounds. I show the seven different grids as a stack here for the 2010 A simulation area. A real object will exist in just one of these layers depending on the $m_R^{F_{\mathrm{box}}}$ measured for the object. Then a normal interpolation can be performed on that grid to find the efficiency. The SB0 layer shows the recovery efficiency for the brightest background bin and the SB6 grid is the faintest. The expected trend is observed where, overall, the fainter background result in simulated objects being recovered more frequently.

The grids are now established and ready for the the real SNe Ia objects discovered by PTF. The real objects were selected from the relevant sky box areas and had to have exploded during the simulation time. They then had their own $m_R^{F_{\mathrm{box}}}$ calculated from real observational data so that the relevant slice of the efficiency grid stack could be selected. Then for that efficiency grid the measured light curve parameter and redshift occupied a unique location in the grid. The value of ϵ_{SN}, found by interpolating on the grid, is the efficiency of discovery for that object which is used later as a weight factor in the eventual rate calculations.

4.6   Summary

In this Chapter I have quantified the performance of PTF for the Type Ia Supernova transient class. This work provides the means to calculate individual weights for SNe Ia, essential for the calculation of rates and luminosity functions. Specific efficiency grids were created for different observing footprints during the survey. Quantifying how well PTF could recover SNe Ia in each footprint involved the simulation of hundreds of millions of individual light curves. During the pipeline development 2×10^9 light curves were simulated to highlight the effect of including different parametrisations as part of the simulations. The complexity of the pipeline grew from initially just quantifying efficiency as a function of the SALT2 parameters to ultimately including spatial

information, local environment surface brightness and CCD defects. This level of detail in an efficiencies simulation is, to-date, unparalleled for a sky survey. The final efficiency grids, used in the publication of this work's findings, consisted of 4.6×10^8 generated light curves, each analysed against 7 different background metrics. This totalled 3.2×10^9 objects which were processed in the pipeline. 56 different efficiency grids were produced so that any real transient occurring in one of the simulation footprints could have its efficiency quantified.

The light curves were generated from the SALT2 model and constructed under the prescription of being cosmologically useful. In this regard I used the work of Betoule et al. (2014) to constrain the boundaries of the model parameters. A serious consideration of the treatment of colour in the simulation was undertaken. PTF cannot measure colour and so no quality cuts or efficiencies could be applied as a function of this parameter. The underlying population of simulated objects therefore needed to be representative of a sample of SNe Ia. To achieve this I used the (Betoule et al., 2014) sample and fit a skewed-Gaussian distribution. For all simulations of SNe Ia the colour properties of the population were derived from this distribution.

The observing logs were combined with the model light curves to check the conditions the artificial object was observed under. Combining the observing log, the synthetic magnitude and $m_R^{F\text{box}}$ together allowed for each epoch on the light curve to be checked against the single epoch recovery efficiencies I derived in the previous Chapter. A Monte Carlo simulation of these light curves and culling the single epoch points produced a population of statistically representative SNe Ia. Efficiency grids as a function of the object's spatial parameters (R.A. and Dec.), SALT2/light curve parameters (x_1, z, M_R) and the environment parameter ($m_R^{F\text{box}}$) were constructed.

These efficiency grids are now ripe for exploitation. In the next Chapter I will discuss the real Type Ia Supernovae found and spectroscopically confirmed by PTF. To these objects I fit the SALT2 model and find their $m_R^{F\text{box}}$ from real observational data. Additionally, PTF was not spectroscopically complete and so I performed a photometric search of all transients recorded by PTF to find the objects well constrained as SNe Ia. Finally, I performed a massive search of every single candidate in the PTF database to find objects that were completely missed during the survey and didn't even receive a PTFName. Once these objects were found and measured they occupy a region in one of the 56 efficiency grids and, thanks to this Chapter, I am able to assign true weights to them.

Chapter 5

The Type Ia Supernova Sample

After the efforts of creating fake point sources, artificial night skies and simulated supernovae, in this chapter we will explore the real SNe Ia that form the analysis sample of this work. The exploration of supernova samples from PTF can broadly be categorised into 3 groups: the 'spectroscopically confirmed' sample, the 'photometric' sample and the 'missed' sample.

Spectroscopic objects are those which were followed up and had a spectrum taken. Subsequent classifications were obtained from line identification explained in Chapter 1.1. For the purposes of SNe Ia, this classification method can be considered unambiguous for an object to enter the sample. The spectra also provide a redshift for each object, which allows this parameter to be fixed in model fits.

Photometric objects are transients that received a PTF name but were not spectroscopically followed up. Therefore, the only assessments that can be made about the objects are derived from their light curve properties and possible host information. Confusion can arise regarding the true nature of the object, especially if the light curve is sparsely sampled. Furthermore, without redshift information it becomes difficult to resolve the degeneracy between faint nearby objects, and those which are brighter but further away. Such effects contribute to the systematic uncertainties.

Finally, there are the most elusive and difficult to classify objects - those completely missed by PTF. These are objects which were never detected in real time and were never followed up. They are devoid of any automated contextual assessment, nor did they ever receive a review from the collaboration members. While the transients in this category are few and far between, they are by far the most challenging of all to classify.

In this Chapter I detail the process of searching for and selecting supernovae from each of these categories. The identification of the SNe Ia is necessarily thorough, as missing

just a few objects will bias later calculations of rates to lower values. Furthermore, the need to find the missed objects is compounded by the fact that they are typically the hardest to find. Consequently, the efficiencies derived for those objects would carry the most weight into the rate calculations. The effects of missing photometric objects can be mitigated by having a large sample with high spectroscopic completeness, and in this regard the PTF sample is exceptional.

The layout of the Chapter is as follows. I first discuss the catalogues containing the spectroscopically confirmed supernovae in PTF. Subsequently I explain how I retrieved the real-time photometry to produce light curves. Following this, I discuss how I fit the SALT2 model to the light curves and apply coverage cuts. To further constrain the model fit parameters I also performed a more sophisticated photometric reduction and refit the same model. The second sample category, the photometric objects, were analysed through a similar multi-stage procedure except the input sample was much larger as it contains all unclassified PTF transients. These were filtered down to the objects best matching the SN Ia model. Finally, the 'missed' objects were searched for by re-analysing more than 1×10^9 candidate detections from the image subtractions. This final stage recreated light curves for all known supernovae, AGN, and variable stars - which were easily filtered out. Crucially, however, light curves were constructed for objects missed by the automated pipeline.

5.1   The Spectroscopic Sample

PTF spectroscopically classified almost 1,900 supernovae. Each spectrum taken by a follow up program was uploaded to the PTF marshal. Members of the collaboration were able to visually inspect the data and assign a classification to the object within the marshal. It was possible to update a classification as new data were obtained, all whilst retaining a record of previous assessments. To make the most complete sample, I utilised the results of these classifications and included any object that received a SNe Ia tag at any point during the survey. This produced 1,249 unique objects across all cadence experiments, redshifts and Ia sub-types. Of these, 606 were at a redshift of z≤0.1 (Maguire et al., 2014), beyond this the number of spectroscopic objects decreases (Figure 2.5) as the objects become too faint to prioritise follow-up telescope time (or may not be detected by the automated pipeline). In Figure 2.5 the distribution of spectroscopically confirmed objects is shown to track the volume of the universe out to a redshift of ∼0.08–0.09. All objects with z≤0.1 had their light curves assessed in the next stage of the analysis.
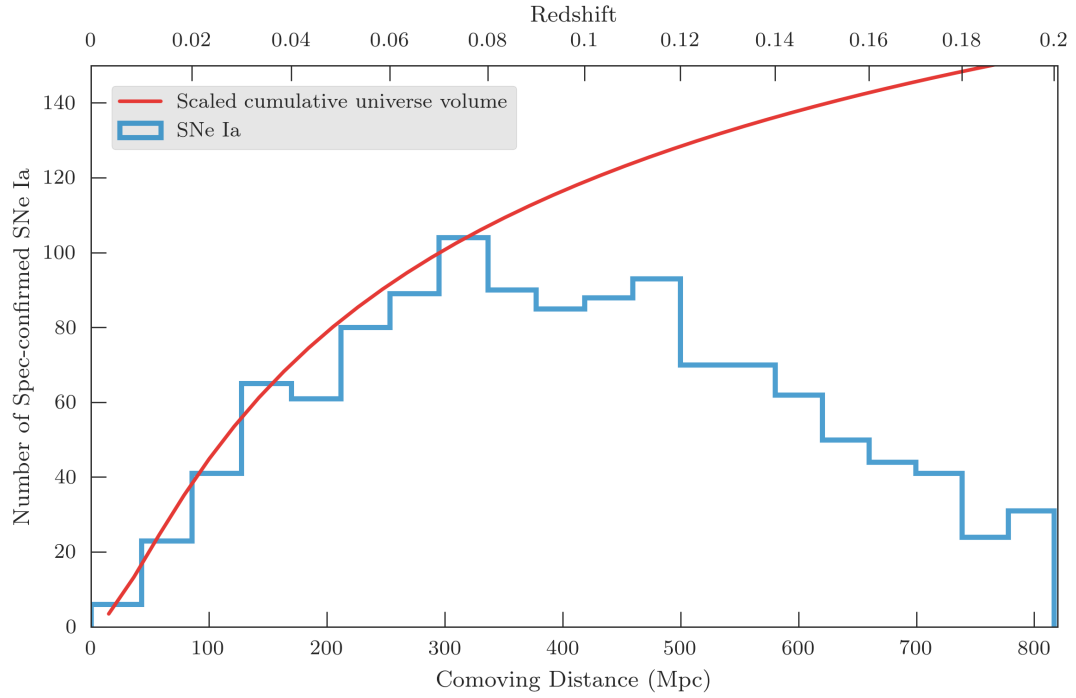
FIGURE 5.1: The volume element of the universe, in red, is compared to the redshift distribution, in blue, of the spectroscopically confirmed SNe Ia. The distribution of objects tracks the general trend of volume until z~0.04, at which point a noticeable deviation is seen. Confidently, we can claim that PTF is spectroscopically complete to this distance. Whilst the general increasing trend is seen out to z~0.04, it no longer track the volume growth. Beyond z~0.11 the overall number spectroscopy classified falls away. This is due to the faintness of objects at this distance becoming too resource costly to justify a high quality follow up observation.

5.1.1 Generating the real time light curve

As part of the PTF data structure, each object with a PTF name retained a record of the candidates which were found to be spatially coincident with the initial detection. Concurrent with the PTF name to candidate association, metadata on the reference image was also joined to the information available for each object and its light curve data. I then require that the real-time photometric measurements were only used if:

- R_{P48} points had a Real-Bogus score ≥ 0.07 (see Chapter 2.2.3)

- Brighter than 20 mag

- Not in the masked CCD regions (see Chapter 4.2.3)

- Median date of the reference image stack was at least 25 days prior to the observation

This ensured that a real light curve would look identical to any simulated object with the same intrinsic properties.

In Figure 5.2 I have shown several PTF real-time light curves, typical of the sample and cadence, out of the 606 spectroscopically confirmed objects. Once the suitable light curves had been constructed from the real time photometry, a check was performed on the coverage of the light curve. If the object had fewer than 4 points on the light curve separated by less than 12 hours then the object was rejected from further analysis. This coverage requirement was the most basic cut and did not depend on the SALT2 model parameters, hence it was conducted before light curve fitting.

5.1.2 Light Curve Fitting

In this section I discuss the procedure of fitting the SALT2 model to the light curves. Through performing a model fit to the observations I obtained values for the parameters that define the efficiency grids in Chapter 4.4. These fit parameters occupy a location within a multidimensional efficiency grid, where the value at that location is the probability that PTF recovered a simulated object similar to the real SN.

To execute the model fits I once again used the `python` package `sncosmo`. I use the Markov Chain Monte Carlo (MCMC) fitter based on the Python package `emcee` (Foreman-Mackey et al., 2013). To remain consistent with the process of simulating light curves I adopt the SALT2 model. For the purposes of fitting light curves, the randomly generated trial values are the SALT2 model parameters; z, t_0, x_0, x_1 and C. Each of these parameters are used to create a synthetic spectral time series. The synthetic spectrum at each observational epoch is integrated through the R_{P48} filter to generate a synthetic photometric point. Comparing the observed to the synthetic photometry, over the whole light curve, provides a χ^2 goodness-of-fit metric. A likelihood function is applied to the χ^2 metric so that when the χ^2 is minimised the likelihood is maximised. For fitting the light curve I use a likelihood function, $\mathcal{L} = \exp(-\chi^2/2)$. The next iteration chooses different SALT2 parameters and finds the likelihood for that combination, $\mathcal{L}_{\mathrm{proposed}}$. The ratio of the two likelihoods are compared, $\mathcal{L}_{\mathrm{proposed}}/\mathcal{L}_{\mathrm{current}}$ and a random number, λ, is drawn uniformly from 0 to 1. If the ratio is $>\lambda$ then the the current likelihood is replaced with the proposed one, if not, then the current likelihood is retained. The redshift was fixed at the spectroscopic redshift, determined through either the supernova features or through contaminant host galaxy flux. An initial guess of the time of peak, t_0, was determined from the MJD of the point with the largest flux. The value of t_0 was bound between limits such that the earliest observation aligns with the latest SALT2 phase and the latest observation aligns with the earliest phase. The colour term was left

free as a single band light curve cannot be constrained in colour. The remaining terms x_0 and x_1 were also left as free parameters. The choice on the number of samples to take during the MCMC routine is somewhat arbitrary and, in principle, could run forever. In practise, a trial-and-error process was performed on a random selection of objects to assess how many samples were required to converge on a solution. In most cases, convergence to a solution was seen for far fewer samples than $\sim$10,000, whereas the most sparsely observed supernovae would almost never converge. Computing resources and time were of no concern at this stage of the process and so I set a limit of 50,000 samples for each light curve fit.

In Chapter 4.1, I was only able to simulate objects within the valid time range of the SALT2 model, -20 to +50 days. To maintain consistency in the treatment of simulated and real objects, the fits were performed as part of a two stage process. Firstly, the SALT2 model was fit to the entire data, and then the best fit value of t_0 was used to trim the light curve down to those in the aforementioned range. The second stage re-performed the fit on the trimmed light curve using the previous best-fit parameter values as initial guesses. From this secondary fit the new best-fit model parameters, along with M_R and the fit covariance matrix, were stored for later processing. The best fit parameters were taken as the median value of each parameter from the sample distributions, and the 1σ uncertainties determined from the boundaries containing 68.3% of the probability along a projection of the parameter axis.

5.1.3   Coverage Cuts

The next stage of the process was to determine whether or not the objects passed the coverage cuts defined in Chapter 4.2.6. The best-fit time of peak brightness, t_0 was used to check whether there were at least two epochs before and after peak brightness. The number of SNe Ia passing the coverage cuts was 227.

The light curves, after coverage cuts, in Figure 5.2 are shown with their best fit SALT2 model light curves over-plotted. The inset table displays their fit parameters I later use to calculate the efficiency of their detection. A table of all the model fit results to the objects with at least 4 data points with $\leq$20mag can be found in Appendix B.

From the real time photometry it is clear that often the SALT2 model was not always well constrained. To place better constraints on the underlying SN parameters I next performed a more rigorous photometric analysis described in the next section.
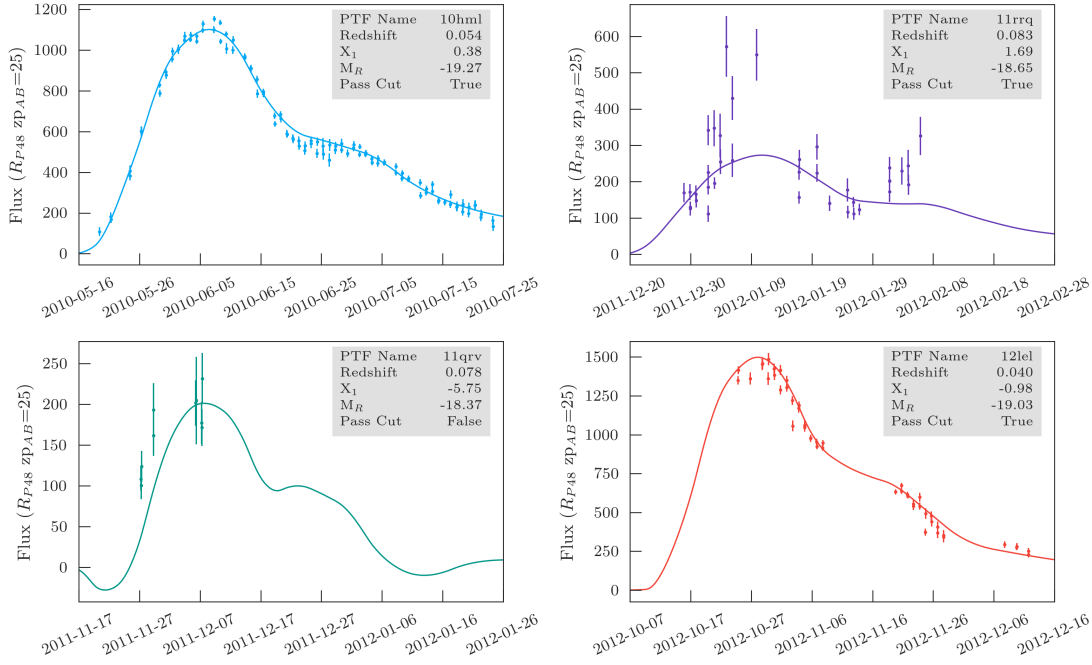
FIGURE 5.2: Four example real-time light curves selected from the spectroscopic sample of SNe Ia. The top left object, 10hmv, is exquisitely sampled for many weeks from explosion and occurred when PTF concentrated on a single footprint. Contrast this with the bottom right object, 12lel, which was also well observed on the decline, but has no rise-time information as it exploded before PTF observed its field. The final two objects demonstrate light curves which are not immediately obvious SNe Ia. Their light curves are ambiguous and it was only through their spectra that they were easily identified as SNe Ia. Each object has been fit with the SALT2 model and the best-fit parameters are shown in the inset boxes to each light curve.

5.1.4 Improving the Photometry

To improve the photometry, on each of the qualifying supernovae I utilise a pipeline widely used within the PTF community, `ptfphot`. The pipeline has produced high quality photometry for SN studies such as Maguire et al. (2011, 2012); Pan et al. (2014); Firth et al. (2015); Dimitriadis et al. (2017) and was therefore a suitable choice for this work.

The pipeline is based on the same image subtraction principles as the real time pipeline described in Chapter 2.2.2. A high-quality deep reference image was created from a combination of observations taken prior to the supernova's explosion. The PSF of each science and reference image was measured from isolated stars. The PSFs were matched and the reference was subtracted from the science image. The position of the SN was determined from the highest signal-to-noise epochs, typically to within 0.05–0.1 pixels.
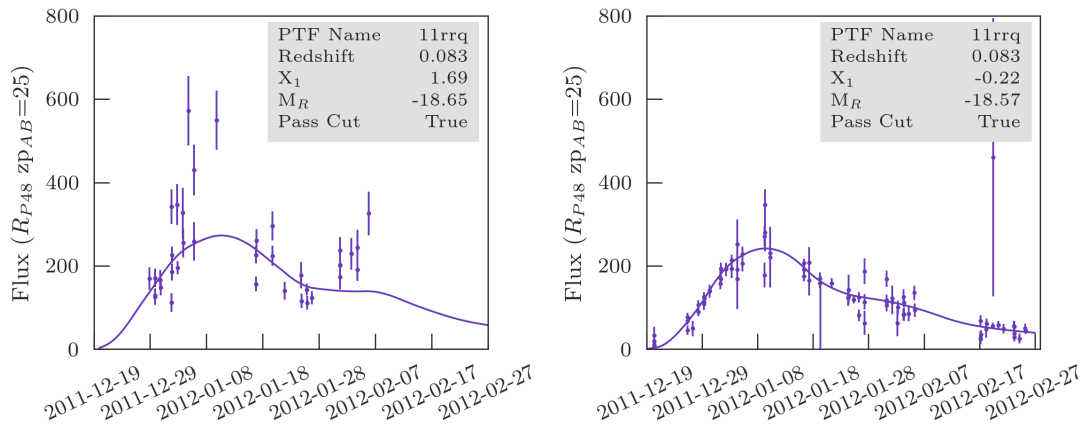
FIGURE 5.3: The real time photometry (left) and the `ptfphot` photometry (right) are compared. Clearly the right hand panel better constrains the light curve model and hence I perform this improved photometry on all objects.

Once established, PSF photometry is forced at that location in all images. The photometry was calibrated to SDSS if they share the field, or to the catalogues of Ofek et al. (2012a) if they do not, just like for the real time pipeline.

After the photometry had been performed, the resulting light curves were put through an identical fitting routine to the real time data described in Section 5.1.2 and the data products were similarly stored. In Figure 5.3, I compare the products of the real time photometric pipeline to the improved photometry. Immediately it is clear that `ptfphot` improves the photometry.

The real time light curves were constructed following the same selection criteria placed upon the simulated light curves. This stems from the fact that the entire efficiencies analysis was performed on the real time data, demanding that selection criteria derived from photometric coverage on real objects must also only come from real time data. Conversely, the improved photometry is not restricted in this a way as it is only performed on the objects passing the real time cuts and is only used to greater constrain the underlying SALT2 parameters which describe the 'True' SN Ia.

Furthermore, and as I demonstrate in sections 5.2 and 5.3, the initial selection must be made on the real time photometry as running `ptfphot` on the hundreds of thousands of photometric and missed candidates is impossible. It is essential that all objects, regardless of being spectroscopic, photometric or artificial, are treated in a consistent way.

5.2 The PTF Name Photometric Sample

The photometric sample of supernovae consists of objects identified by the automated transient search and assigned a PTF name. However, these objects were not spectroscopically followed up and hence the nature of the transient is unknown. Some of the objects in this photometric catalogue were most likely added to a follow up queue, but were never observed as other transients took precedence. There are almost 47,000 distinct objects in this catalogue and searching all of them by eye would be an overwhelming task. In this section, I describe the process I developed to find and extract the SNe Ia from the other transient types and spurious detections.

5.2.1 The PTF Name database

To begin the task of discovering photometric SNe Ia, I first spatially associated every PTF named object to all candidate detections made at that location and joined on the observing log information. There were several initial cuts I was able to make on the dataset: the requirement for a light curve to have at least 4 points brighter than 20 mag, with RB$\geq$0.07, and separated by 12 or more hours. This first cut is independent of transient classes and model fitting. Additionally, this also sets a lower limit on the duration of the light curve of: 4 points separated by 12 hours must leave at least 36 hours between the first and most recent observation. Therefore, all PTF named objects were cut from the sample if they were shorter than 36 hours and failed to meet the magnitude and coverage cuts. Furthermore, the reference images must have had a median stack date of at least 25 days prior to each observation.

These initial cuts culled the original $\sim$47,000 objects down to just 5,051. In Figure 5.4, I have randomly chosen 9 example light curves from the sample of 5,051 that met my minimum criteria. Clearly, there are objects here which are simply junk and the result of poor subtractions or non-SN variation. However, a couple of the objects, 10sur and 12jso, do show general supernova-like features - i.e. a steady rise or decline, or rise followed by a decline. The objects which do display trends like these, and perhaps are SNe Ia, would provide a better fit to the SALT2 model compared to spurious detections. This motivated the next stage of the search.

5.2.1.1 Photometric sample light curve fitting

5,051 supernovae are a reasonable number of objects to quickly fit with the SALT2 model. Unlike the spectroscopic objects in Section 5.1.2, I was unable to fix a redshift
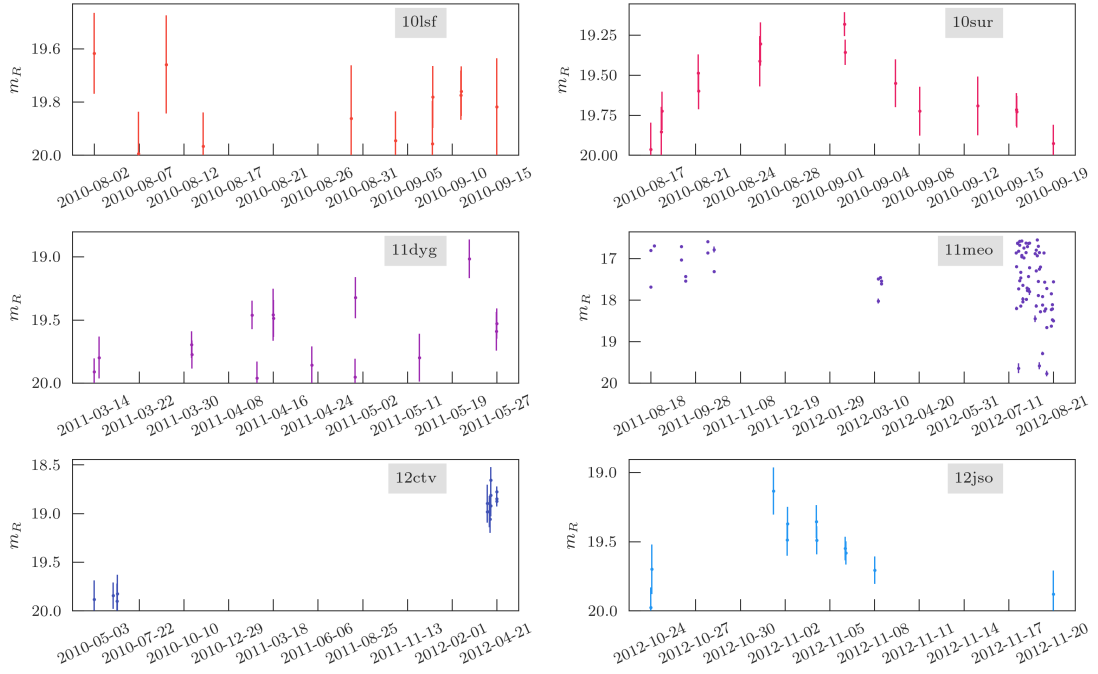
FIGURE 5.4: Example objects which were missing a spectral classification but were assigned a PTF name. Some of the objects are junk, whereas others clearly have transient features.

for the objects; however, I did set an upper limit of z~0.3. The only other change to the previous fitting method was to reduce the number of MCMC samples to 3,000. This was done to reduce to overall processing time and, at this stage, the fit was only a rough guide to check for supernova-like evolution.

The fit results to a selection of typical objects are shown in Figure 5.5. In Figure 5.6, I show the distribution of the reduced χ^2 of the 5,051 objects. The long tail to the distribution shows many objects are an extremely poor fit, and I cut objects from further analysis if their χ^2_{red} was $\geq$150. As a comparison distribution, the spectroscopic sample objects are also shown. To cull the sample even further I cut objects with $x_1 > 8$, this removes 75% of the objects from the sample. This was a conservative cut on the x_1 parameter and removed all objects which have a similar light curve shape to the bottom row of objects in Figure 5.5, which are clearly not SNe Ia.

After these cuts, 1,117 light curves remained, which I then visually inspected for SN-like properties. This retained 216 objects. I show thumbnail light curves in Figure 5.7 for these. 13 of the 216 objects appeared to have missed a cluster of points and instead a light curve was fit to a possible spurious bright detection. These few objects were assigned a pass flag where I later inspected them in more detail.
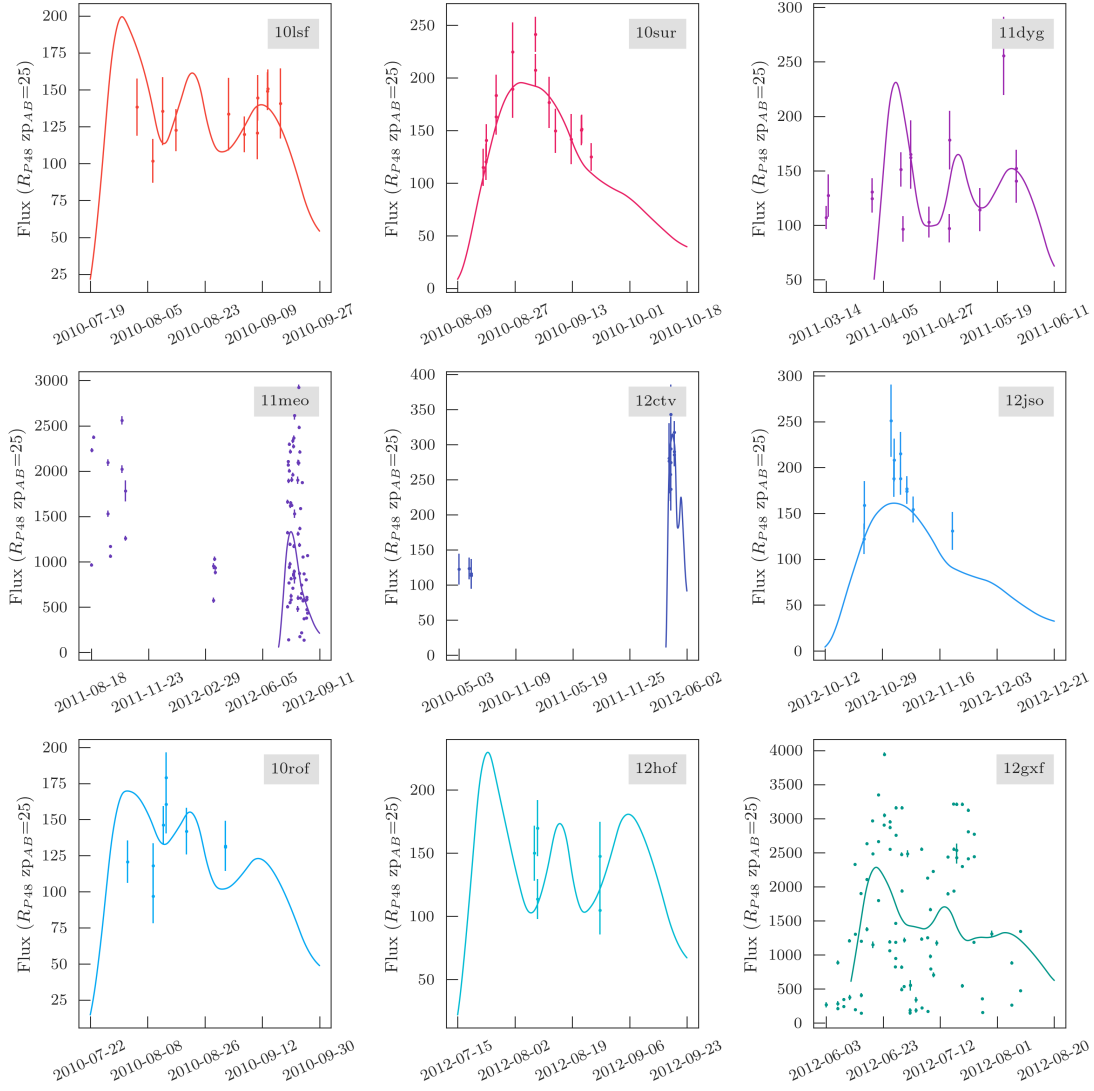
FIGURE 5.5: Example of light curves fits to the many thousands of objects that were fit with the SALT2 model. The light curve fits are extremely instructive at filtering the non-supernova light curves form the objects I was interested in.

There remains a final, and easily applicable, cut to perform. Since the aim of this chapter was to find objects to enter the final SNe Ia sample, only objects that occurred in the simulation boxes of Chapter 4.3 were useful. I performed this cut and the sample of 216 reduced to 31. Whilst this cut may have been the most useful to perform first, it was a curiosity into the objects that PTF didn't take a spectrum of that led to this being one of the final cuts. In the next section I try to further constrain the nature of these objects by studying any host galaxy information.
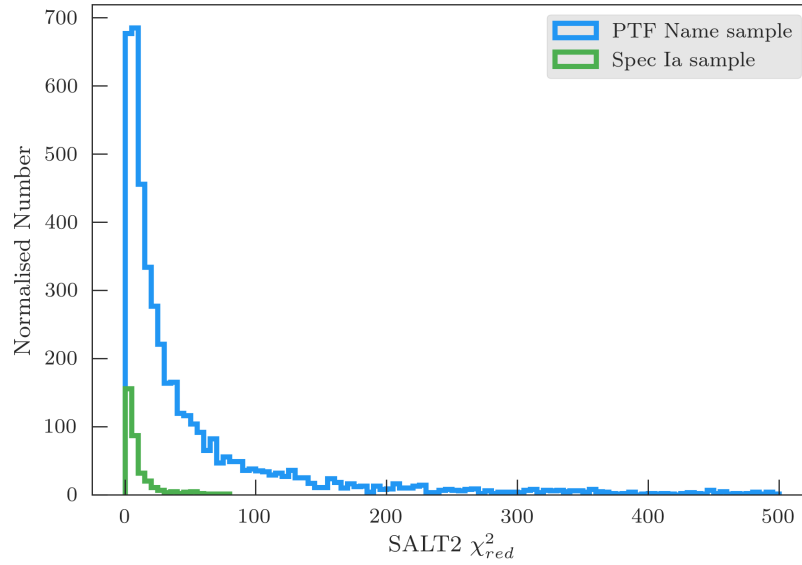
FIGURE 5.6: The distribution of the χ^2_{red} of the SALT2 fits for the photometric sample are compared to the spectroscopic sample. The spectroscopic objects shown here do not necessarily meet the cuts to make the final sample. This was done to see where a goodness-of-fit cut could be made to confidently cull the photometric sample without loosing potential objects.

5.2.2 Host Galaxy information

The objects passed onto this stage were queried against the 'Galaxy' table in the SDSS (York et al., 2000; Eisenstein et al., 2011) database to return the properties of the nearest galaxy to each supernova. Additionally, image stamps at the location of each object were produced to ensure the right host was paired to the object. The stamps, and the object light curve, can be seen in Figure 5.8. The cross-hairs are centred on the location of the object and any red squares show objects with a spectrum. Such was the extensive coverage of SDSS, some of the photometric supernovae I investigated were associated to galaxies with a spectrum. The objects with a host redshift are listed in Table 5.1.

PTF Name	R.A.	Dec.	Host z
10eei	195.0611	38.92302	0.072
10fmi	184.1685	66.78963	0.086
10hmq	201.5008	7.250227	0.086
10lrp	222.8477	35.76095	0.079
12cci	213.8721	39.90248	0.067
12gcn	230.9323	8.437248	0.075
12gfj	242.4574	29.72575	0.049

TABLE 5.1: The redshifts of the galaxies hosting the photometrically identified SNe Ia

FIGURE 5.7: Thumbnail images of the light curves I visually inspected and chose for further analysis. These objects existed across multiple cadence experiments and time during the survey.
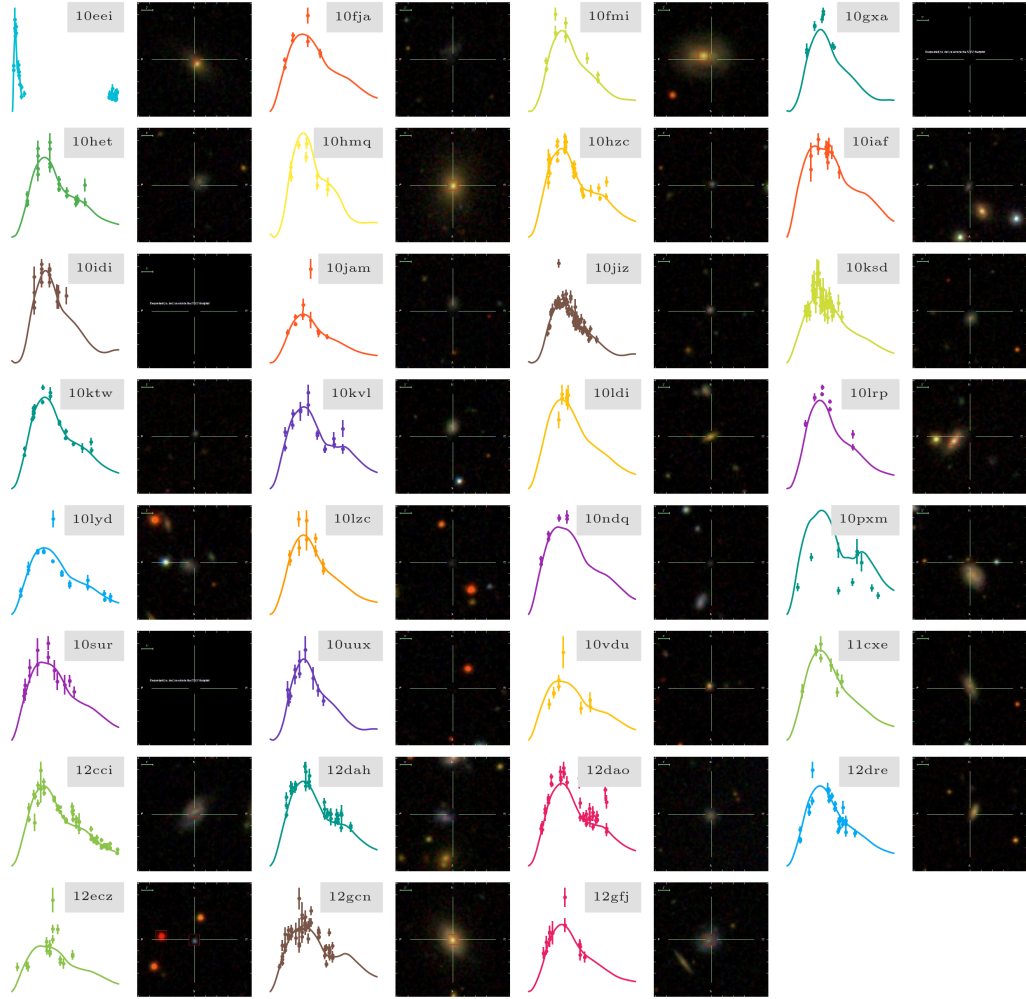
FIGURE 5.8: Each of the photometrically identified objects, which passed the cuts and occurred in a simulation box are shown. The galaxy stamps show the environments of the object where the cross-hairs show the location of the transient. If there are any spectral data available for the galaxy then it is identified by a red box on the target.

5.2.3 Improved photometry and light curve fitting

The objects shown in Figure 5.8 all had their photometry improved by the `ptfphot` pipeline, as detailed in Section 5.1.4. Following this, the light curves were refit according to the methods laid out in Section 5.1.2. For the objects in Table 5.1 the model redshift was fixed and the other parameters were treated identically to the spectroscopic sample. The remaining objects were fit with redshift as a free parameter but the number of MCMC samples was increased to compensate with for extra dimension.

Objects that did not meet the same light curve coverage and fit parameter requirements as the spectroscopic sample were cut. In total 6 objects, listed in Table 5.2, remained
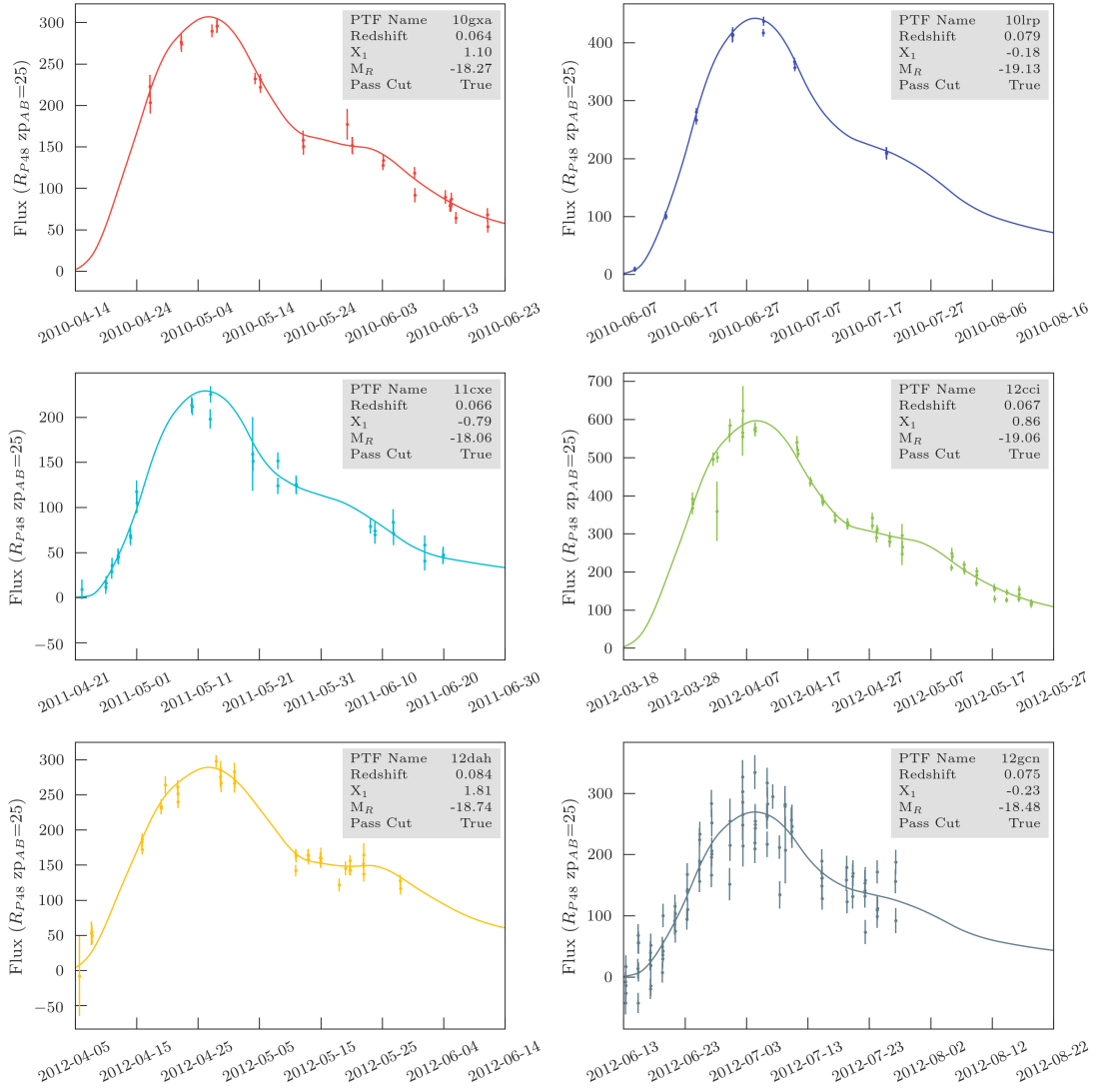
FIGURE 5.9: The 6 object light curves that made the final SNe Ia sample are derived from the improved photometry and fit with the SALT2 model. The inset box displays the best fit model parameters.

to form the sample of photometric SNe Ia. Their higher quality light curves and best fitting light curve can be seen in Figure 5.9.

It is interesting to note that I checked the PTF Marshal for each of these objects. PTF12cci had two comments: the first on the light curve rise stating that it is probably a SN Ia and a spectrum should be obtained. The second comment came after the secondary maximum on the light curve to claim it is clearly a SN Ia. It is encouraging that my method of recovery has picked up an object for which independent evidence also suggests as SN Ia.

Finally, the objects listed in Table 5.2 had their reference images analysed as described in Section 5.4, just like for the spectroscopic sample.

PTF Name	R.A.	Dec.	z	t_0	X_1	M_R	Spec z
10lrp	222.8476	35.76095	0.079	2455375.47	-0.18	-19.13	True
12gcn	230.9322	8.437248	0.075	2456111.91	-0.23	-18.48	True
10gxa	248.6179	57.60668	0.064	2455321.18	1.10	-18.27	False
11cxe	250.8159	40.53592	0.066	2455693.39	-0.79	-18.06	False
12dah	230.4538	50.07262	0.084	2456043.48	1.81	-18.74	False
12cci	213.8721	39.90248	0.067	2456024.95	0.86	-19.06	True

TABLE 5.2: The final set of photometrically identified objects which entered the SNe Ia and their fit parameters are listed here.

5.3   Searching for Missed Supernovae

I will now discuss the efforts made to discover transients that were completely missed by the automated pipeline. These objects didn't receive a PTF name, therefore they were never assessed by the human scanners or had any type of follow up observation. The objects presented at the end of this chapter are new discoveries. This search for missing objects is a necessary part of forming a complete catalogue of SNe Ia, and, just because PTF missed the object, doesn't mean I can discard them for my sample. The single epoch efficiencies, that form the foundation of this entire work, quantify only how well the transient detection pipeline performed on a candidate with various measurable properties. Stringing candidates together to form light curves, human interpretation of the light curve, and external follow up are impossible to simulate and therefore cannot be quantified. Furthermore, performing this search was motivated by the limitations of the automatic detection pipeline. As discussed in the Chapter 2.2.3, the transient discovery pipeline discarded a candidate if nothing was detected between 45 minutes and 6 days of that epoch. This leaves open the potential for transients to be missed if the weather was poor for an extended period, or even if a prior spurious detection is spatially coincident to an actual event.

This search constructed light curves for all spatially similar events in the candidates database which contains 1,008,603,464 separate detections. These light curves included all PTF detections including contaminant sources such as AGN and variable stars. After a process of filtering, the light curves were assessed by fitting the SALT2 model to further throw away junk, until a manageable number was left for visual inspection.

The layout of this section is as follows; first I will detail the process of extracting the candidate detections from the table of photometry. Through various selection criteria and joins to complementary data sources I will demonstrate how only useful data will be utilised. Following this, the data intensive process of clustering candidates into potential sources will be described. Then I will explain how, through the analysis of external catalogues, I removed known sources. The remaining sources and their light curves are

either the real missed transients or pseudo-transients, i.e. junk. Finally, I will describe the method of extracting the objects PTF missed, improving their photometry and fitting the SALT2 model. A presentation of the supernovae I discovered will conclude the chapter.

5.3.1 Filtering the candidates

Throughout this process I will refer to the tables and connections shown in Figure 5.10. The figure is a visualisation of the database schema I modelled on the remarkably easy to follow London Underground[1] Tube map. Each table in the schema represents a different source of data and the 'tube lines' represent the connections made between common identities in each table. The colour coding on the lines describes the data type travelling along each line. The shaded backgrounds show tables where the sources of the data were internal to PTF, or drawn externally. Finally, the connecting 'stations' show points in the schema that required external processing to filter or add extra information to each table.

The process started in the candidate table containing the $>1\times10^9$ sources found by PTF. Objects were filtered out to only include the ones observed in R, later than 2010 January 01, with Milky Way extinction $E(B-V)\leq0.1$, Real-Bogus≥0.07 and brighter than 20 mag. This reduced the sample down to 4.8×10^7 candidate points and, in addition, 1.48×10^8 candidates which had negative flux in the subtraction images. There are many more negative subtractions than positives because the machine learning algorithms were not trained on negatives.

5.3.2 Clustering the candidates

Many of the candidate detections were spatially coincident with one-another when marginalised over time. So far the data are just listed in the tables without knowledge of the location of the other points. To cluster candidates together is to create an object. To prepare for the clustering of points I re-engineered the data, as described in Chapter 2.4, into a spatial format within PostGIS. To perform the clustering I used the Density-Based Spatial Clusters of Applications with Noise (`DBSCAN`) algorithm (Ester et al., 1996).

---

[1]https://tfl.gov.uk/

FIGURE 5.10: The schema diagram detailing the algorithm to search for missing supernovae. The objects are identified starting the from the candidates table in the top left and each connecting 'line' signifies a connection between the tables. The 'stations' through which multiple lines are joined involve an external processing step require to perform a filtering process on the data. The process of identifying and clustering the candidate objects generally flows from left to right. Once the final clusters and the candidates within those clusters are identified in the 'bcand_matching' table the light curves are created in a right to left direction

5.3.2.1 DBSCAN

DBSCAN is a neighbourhood scanning algorithm that discovers clusters of arbitrary shape with high computational efficiency. DBSCAN is an attractive algorithm for the purpose of spatially clustering astrophysical sources because it doesn't require prior knowledge (or an estimate) of the number of unique clusters. The only two parameters to set are the maximum distance between points, ϵ, which forms the cluster, and the minimum number of associated points to define a cluster, minPts.

The spatial data were processed using the DBSCAN algorithm with $\epsilon=3''$ and minPts$=4$. The ϵ parameter incorporates the requirement PTF used for spatial pairing and the minPts value was to meet the absolute minimum coverage requirement of 4 light curve points. DBSCAN is extremely scalable and I was able to process both the positive and negative flux candidate tables across multiple cores in <20 minutes. A limitation of the PostGIS implementation of DBSCAN was that each processor returned a cluster ID in an integer sequence starting from 0, therefore multiple clusters had the same ID. To create a globally unique ID I used the Cantor pairing function to generate a number from system processor ID (PID) and the cluster ID (CID) as

$$\text{ID}(\text{PID}, \text{CID}) = \frac{1}{2}(\text{PID} + \text{CID})(\text{PID} + \text{CID} + 1) + \text{CID} \qquad (5.1)$$

Each cluster is made up of points with both flux and time information. This defines a light curve, allowed me to perform additional light curve coverage checks.

I then enforced the requirement of 4 points separated by $\geq$12 hours, leaving 1.22×10^6 unique clusters. The centroid of each cluster and the 3σ scatter were recorded so that other objects could be matched to them.

5.3.3 Matching the clusters with known objects

Within the 1.22×10^6 clusters, lie all PTF transients already discovered, and therefore it was necessary to filter them out. To achieve this, I took the catalogues of known objects used in the photometric object search, and spatially matched it to the cluster centroids. Any cluster spatially coincident with an object in the catalogue was discarded.

Additionally, PTF maintained a catalogue of point sources used in image calibration (Ofek et al., 2012a). These stars were also matched with a $3''$ radius to the clusters and, again, rejected if paired. Additionally, a catalogue of stars from SDSS that overlapped with the PTF fields was obtained. Anything in the SDSS star catalogue that overlapped within $3''$ also caused the cluster to be excluded. The final catalogue of known

objects contained AGN and was compiled by the iPTF collaboration. The catalogue contains sources predominantly from the Infrared Astronomy Satellite (IRAS) mission (Neugebauer et al., 1984), Parkes radio telescope (Wright & Otrupcek, 1990), X-ray Multi-Mirror Mission Newton (XMM-Newton) (Bianchi et al., 2009) and the Fermi-LAT AGN catalogues (Ackermann et al., 2015). Whilst I was not involved in the creation of any of these catalogues, I still used them as an excellent source of objects to filter out clusters with known sources within $3''$. Any unknown AGN in my cluster sample were also removed by spatially matching the positive candidates to the negative ones. The stochastic variability of AGN means that at any epoch the current brightness may be brighter or fainter than in the reference image, thus producing a light curve of multiple positive and negative subtractions. Any cluster that had 2 negative flux points between the first and last positive flux points were also discarded. After these steps the number of clusters was reduced down to 29,367, a suitably small number for light curve fitting.

5.3.3.1   Light curve fitting

The real time photometry for each of the remaining clusters was fit with the SALT2 model with a method identical to that described in Section 5.2.1.1. The derived fit parameters were assessed identically to Section 5.2.1, and only those with a portion of their light curve occurring in one of the simulation boxes of Chapter 4.3 qualified for the next assessment stage. This reduced the number of object light curves down to 486, and I visually inspected each one. Only 23 of these light curves displayed any transient light curve features with both a rise, decline and the coverage cuts detailed in Section 5.1.3.

Each of the 23 objects had SDSS stamps generated to see if there was any host galaxy spectral information to determine redshifts. Figure 5.11 shows the host stamps and the clustered object light curves. Following this, the `ptfphot` photometry pipeline was run on the objects. This had quite a significant effect on the sample with almost all objects being revealed as spurious detections.

Figure 5.12 shows 4 examples of light curves from the improved photometry. Clearly 3 of the objects are noise similar to most of the other objects in the sample. However, one shows a transient light curve evolution and was fit with the SALT2 model as previously described. Whilst the fit parameters did return values consistent with SNe Ia properties, the fitted redshift and its uncertainty was well beyond the limit of my sample at z>0.2. This object, and the others I performed the fit on, did not make my final sample. Whilst the presence of a galaxy in most of the stamps promoted the hypothesis that these objects were transients, their locations are also not unexpected. Galaxies are notoriously difficult to perform perfect image subtraction on. Their diffuse structure, bright cores, and, in
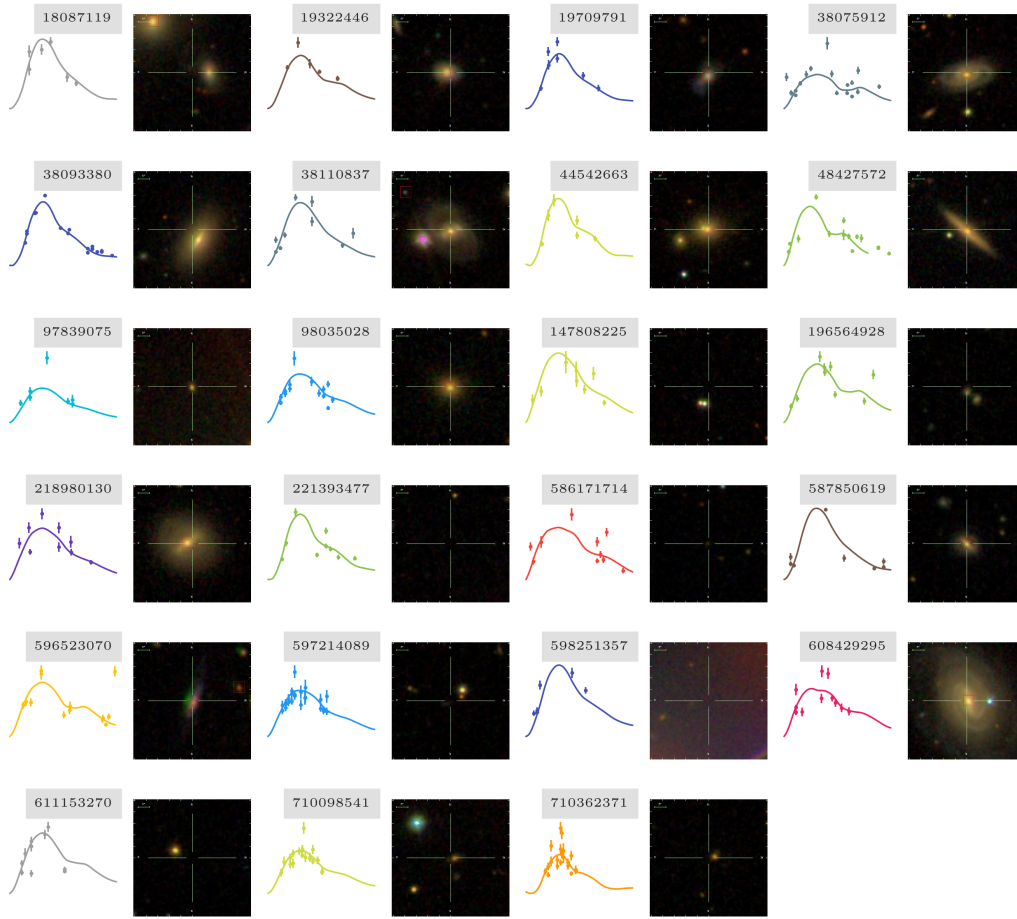
FIGURE 5.11: Light curves and hosts for the completely missed objects I identified in the pipeline as potential new transients. The galaxy stamps show the host locations of the transients and the cross-hairs identify the exact location of the cluster. The red boxes identify galaxies with a spectrum.

some cases, their variability all contribute to an increased probability of subtraction artefacts.

I conclude that during its operational period PTF clearly found the vast majority of the nearby (z<0.1) SNe Ia as two independent searches - the real time survey and this work - did not find any obvious objects. There exists one final object to analyse which I have designated the name PTF38093380.

5.3.3.2   PTF 38093380

The only missed transient of note in my fields, was designated the ID: 38093380. Its light curve, shown in Figure 5.13, is beautifully sampled especially in the exponential decay of the tail. The SALT2 model was fit to the object's photometry. The resulting parameters
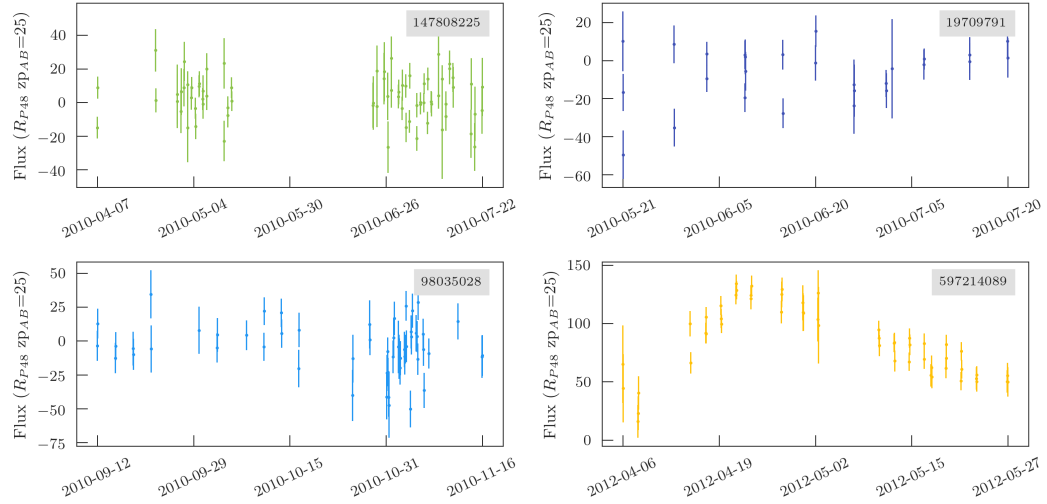
FIGURE 5.12: The potential objects were passed through the `ptfphot` pipeline. Most of the objects were rejected after the photometry was improved. However a couple of objects stilled showed light curve features, such as the object in the bottom right panel. These objects were fit with a SALT2 model to see if they met the criteria to enter the final sample.
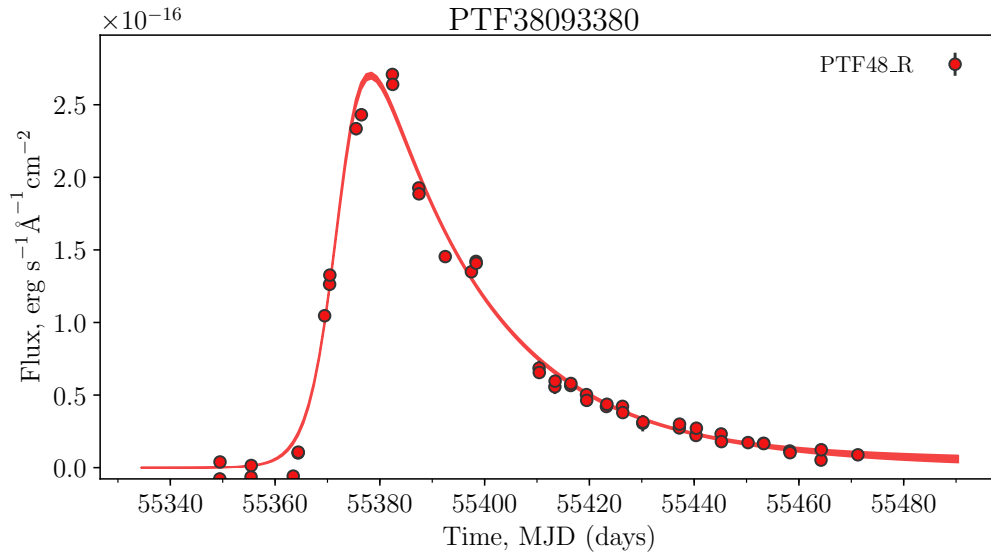
FIGURE 5.13: A light curve generated by the COCO routine (Firth et al., in prep). The data points are from the `ptfphot` photometry routine. This object did not provide a good fit to a SALT2 model. However, the object provides an extremely good fit to a SESN model parametrisation.

did not pass the requirements to enter the final sample as an SNe Ia. An exploration into the likely nature of the transient was under taken. The model parametrization of Karpenka et al. (2013) was fit using `MultiNest` (Feroz et al., 2009, 2013) as part of the COCO routine (Firth et al., in prep). Several stripped-envelope supernova (SESN) spectrophotometric templates were used to construct a model light curve which was fit to the observed photometry. The results of this routine provided an excellent fit to a stripped-envelope supernova model. Hence, the exclusion of this object from my final sample was justified. Although, without a spectroscopic redshift, the absolute magnitude of the object is unknown.

5.4 Measuring local surface brightness from the observations

The model parameters extracted from the SALT2 fits, along with the spectroscopic redshift, are able to define a location in each slice of the multidimensional stack I visualised in Figure 4.16. The slices represent the different $m_R^{F_{\mathrm{box}}}$ bins discussed in Chapter 4.2.4. Therefore, the only question left is: 'Which real object is related to which slice?'. After this stage I could fully quantify an object's detection efficiency.

The parameter $m_R^{F_{\mathrm{box}}}$ is the surface brightness, in a 3x3 pixel box, at the location of the supernovae, before the supernova exploded. This could not be measured in real time and hence necessitates the use of the high quality reference images generated as part of the `ptfphot` pipeline. The method behind this task was simple; a 3x3 box was centred on the SN position and the flux summed in this area. The resulting $m_R^{F_{\mathrm{box}}}$ is found through

$$m_R^{F_{\mathrm{box}}} = -2.5 \log\left(\frac{F_{\mathrm{box}}}{9.18}\right) + \mathrm{zp} \qquad (5.2)$$

where F_{box} is the sum of the counts in a 3x3 pixel box of area $9.18''^2$ and zp is the zeropoint of the image. These $m_R^{F_{\mathrm{box}}}$ measurements were stored for each object along with associated SALT2 model fits parameters.

In Figure 5.14 I show the normalised distribution of $m_R^{F_{\mathrm{box}}}$ for the spectroscopic sample of SNe Ia. As expected, the vast majority (93.5%) of the spectroscopically confirmed objects were found in the two most efficient detection bins where $m_R^{F_{\mathrm{box}}} > 20.38 \ m_R$ arcsec^{-2}. Of course, these objects are the easiest to find and so the benefit of including $m_R^{F_{\mathrm{box}}}$ on the efficiency grids is highlighted by the objects in brighter host bins which will receive a higher weight in the final calculations.
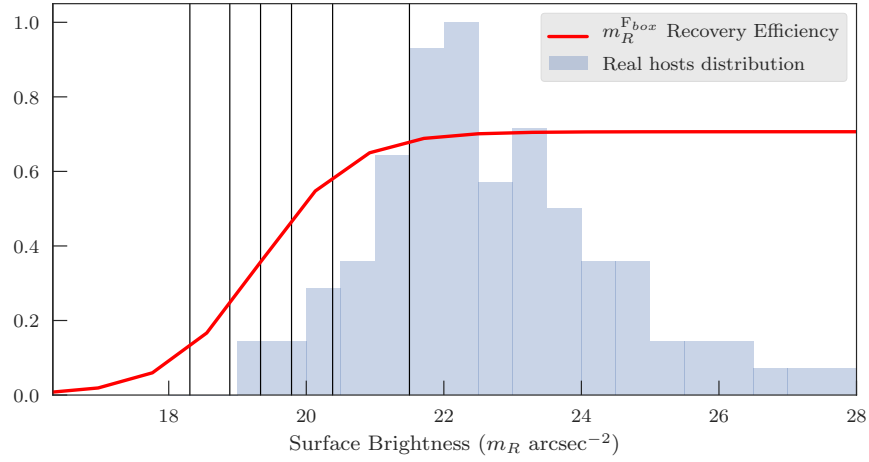
FIGURE 5.14: The distribution of the $m_R^{F_{box}}$ parameter for the spectroscopic sample of objects. The red line shows the 1D recovery efficiency as a function of the $m_R^{F_{box}}$ parameter and the black vertical lines are the bin edges on the final efficiency grids. Most of the spectroscopic objects were found in the most efficient bin.
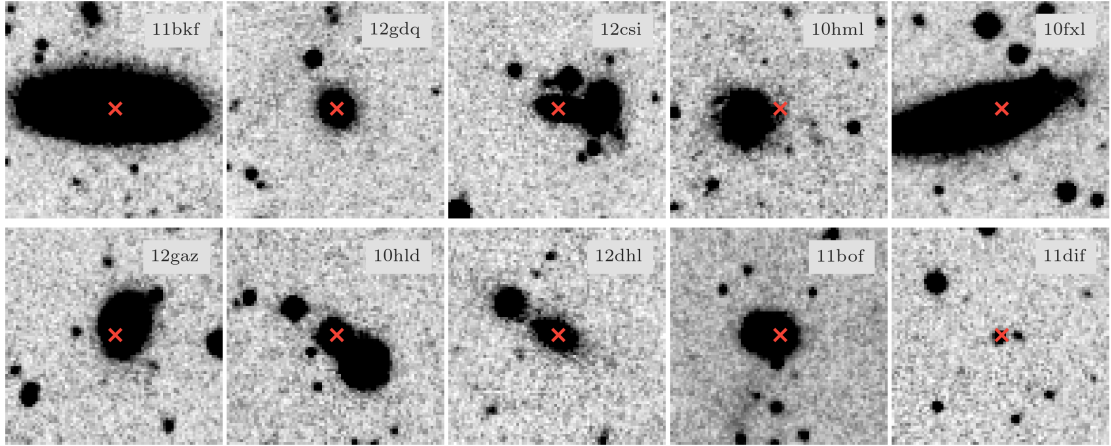
FIGURE 5.15: Example galaxy stamps from which the $m_R^{F_{box}}$ parameter was measured.

Several example cut-out stamps (80×80 pixels) of the host regions for spectroscopic SNe are shown in Figure 5.15. PTF11bkf is located amongst the highest local surface brightness region of any object in my sample with $m_R^{F_{box}} = 19.08 m_R$ arcsec^{-2}. Other objects, such as 12csi and 10hld appear to be in merging systems and have relatively bright $m_R^{F_{box}}$. Objects such as 10hml or 11dif, which appear in the outskirts or more remote locations, fall into the most efficient $m_R^{F_{box}}$ bin.

5.5 Summary

In this chapter I have studied the sample of supernovae PTF discovered. The objects are split into 3 source categories; a spectroscopic sample, a photometric sample and a missed sample. The spectroscopic objects make up the vast majority of the SNe Ia that enter the final sample of objects. These objects had their class confirmed using a spectrum and their classification for this purpose is considered unambiguous. The second category, the photometric sample, contains objects that are well fit by a SNe Ia model but have no spectrum and therefore a potential for a misidentification. Furthermore, not all the host galaxies of the sample have a spectroscopic redshift, which left this parameter unconstrained in model fits. Finally, the totally missed sample contains objects that the PTF automatic detection algorithm did not identify as a transient. These objects are the hardest to find and the most ambiguous to classify reliably.

In the search for missed objects every single detection made by SExtractor during the real time pipeline was re-analysed, where new and existing light curves were generated from the clustering of data. Ultimately zero new sources, that could enter the final sample, were discovered.

All transients, regardless of their origin, had both their real time and improved photometry analysed and fit with a SALT2 model. Objects were required to pass the real time coverage cuts before they could enter the final sample. Constraints on the light curve parameters were derived from the improved photometry fits. The best fit model parameters and their uncertainties were retained so that realisations of each light curve could be incorporated into a Monte Carlo simulation in the next chapter.

The immediate local surface brightness to the transient was measured from the reference images generated by the photometric pipeline. From this the values of $m_R^{F_{\text{box}}}$ were found. A combination of the SALT2 model fits and the $m_R^{F_{\text{box}}}$ value describe locations in the light curve efficiency grid parameter space created in Chapter 4.

I have now determined all the information needed to weight all supernovae in my final sample. In the next Chapter I pull together all of the work performed in the previous chapters to provide efficiency weighting of the real objects. From this I will derive the volumetric Type Ia supernova rate out to a redshift of z=0.09.

Chapter 6

The Volumetric Rate of Type Ia Supernovae

In this chapter I pull together the work of the previous chapters to calculate the SNe Ia volumetric rate - the heart of this thesis. The principles of rate calculations are simple, but as we have explored thus far in the thesis, the practicalities are complex. From a naïve perspective, a volumetric rate is the number of supernovae observed in a given volume during a fixed time period. Of course, PTF wasn't a 100% efficient sky survey and so the work of Chapters 3 and 4 were performed to quantify the performance of the transient discovery system. This performance was captured in the efficiency metric which provides weights to the real supernovae that PTF discovered.

To calculate a rate, a real supernova needs to be assigned a weight based on how likely its discovery was. These weights act to account for the objects PTF missed either as a result of the cadence patterns, weather conditions, coverage cuts or even all three. The weights are calculated by comparing the real SNe to simulated objects with similar properties. For instance, if a real object is most similar to a simulated object that has an efficiency of $\epsilon=0.2$, then that one real object represents a true population of five objects, where PTF missed the other four. The value for epsilon was found through a comparison of the real SNe to the simulated objects in Chapter 4.

Rates at low redshift, z<0.1, are dominated by low number statistics. Surveying a large volume at low redshift requires a wide field of view and typically a fast cadence. Rates from intermediate/high redshift surveys (e.g. Perrett et al., 2012) can survey a substantially larger volume from a smaller field of view and recover many more supernovae. PTF produced a catalogue of low redshift SNe unrivalled in a combination of quality and size, thus improving the statistical uncertainties on rate measurements.

The layout of this chapter is as follows; Firstly, I will explain the rate equation used to calculate a volumetric rate. Then, I will detail the process of deriving the efficiency weights for the final sample of SNe Ia. With this information available it is possible to solve the rate equation and produce the volumetric rate of SNe Ia. Following this, I will review both the statistical and systematic uncertainties associated with this calculation. After a comparison with SNe Ia rates from other studies, I will unify the results and fit a model delay-time distribution to the combined data.

6.1   The rate calculation

The volumetric SN Ia rate is calculated as the summation of N weighted supernovae in a given timespan, ΔT, and a fixed comoving volume, V. The weights are the inverse of the detection efficiency for the ith object in the sample, ϵ_i, and the $(1 + z_i)$ factor accounts for time dilation so that each object is considered in its rest frame. The rate is therefore expressed as

$$r_{\text{V}}(z) = \frac{1}{\text{V}\Delta\text{T}} \sum_{i=1}^{N} \frac{1 + z_i}{\epsilon_i} \tag{6.1}$$

and the volume V is

$$\text{V} = \frac{\Theta}{41253} \frac{4\pi}{3} \left[\frac{c}{H_0} \int_{z_1}^{z_2} \frac{dz'}{\sqrt{\Omega_M(1 + z')^3 + \Omega_\Lambda}} \right]^3 \text{Mpc}^3 \tag{6.2}$$

where Θ is the area searched in deg^2, z_1 and z_2 are the lower and upper redshift limits placed upon sample cuts, c is the speed of light, and H_0, Ω_M, and Ω_Λ are the present Hubble, matter density, and vacuum density parameters respectively.

Throughout this chapter the universe is assumed flat and the cosmological parameters are: $H_0 = 70\text{km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$

6.2   The Final Supernova Sample

In this section I will describe the process of extracting suitable SNe from the objects investigated in Chapter 5. The catalogue shown in Appendix B contains all SNe Ia I investigated. Some of these objects will not meet the requirements to be 'cosmologically useful' and therefore not used as part of the rate calculation. I had already set a requirement on the x_1 parameter to meet the Betoule et al. (2014) criteria. However, I have not yet set a definition of the absolute magnitude of a SNe Ia to enter the sample.

6.2.1 M_R selection

The SNe Ia I simulated were created in B-band. PTF operated in R-band and without colour information it is impossible to construct a selection criteria for the M_B of a PTF object. However, it is possible to work the other way and use the Betoule et al. (2014) sample to translate their sample in R-band and thus set an M_R selection criterion.

I took the Betoule et al. (2014) catalogue of SALT2 model parameters. From this I performed 15,200 realisations of the data, drawing the SALT2 parameters from the covariance matrix of the model fits. A synthetic spectrum at peak was created and integrated through the R_{P48} transmission function to find M_R. The resulting distribution is shown in Figure 6.1.
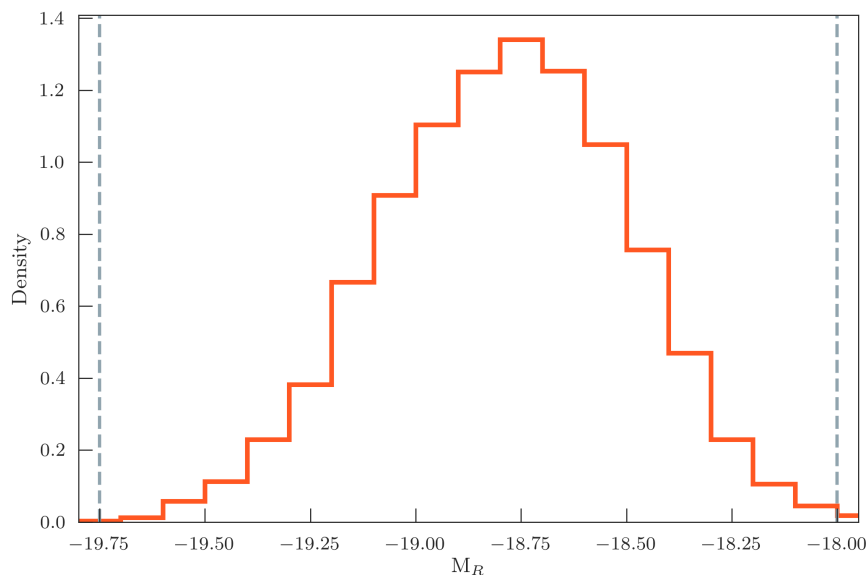
FIGURE 6.1: A Monte Carlo simulation of 15,200 realisations of the Betoule et al. (2014) cosmological SNe Ia data were analysed. Each object had its spectrum at peak integrated through the R_{P48} transmission function to produce the peak M_R magnitude. The distribution is shown by the orange line and the vertical dashed lines show the chosen boundaries for the PTF SNe Ia sample. These boundaries contain more than 99.5% of the data.

The peak of this distribution was found to be M_R=-18.75 with more than 99.5% of the data contained between the limits of -19.75<M_R<-18.0. This naturally sets a convenient boundary for the objects I consider in the rate calculation.

6.2.2 Final definition of a cosmologically useful SN Ia in R-band

Drawing together all the information gathered over the course of this work it is now appropriate to set the formal definition of a SN Ia to enter the sample for the rates

analysis. These boundaries are presented in Table 6.1 and is used both for the rate calculation in Section 6.4 and the rate Monte Carlo simulation in Section 6.5.

Parameter	Boundary
M_R	-19.75 to -18.0
x_1	-3 to 3
z	0.015 to 0.09

TABLE 6.1: The parameters each SNe Ia must meet to make it in to the rate calculations.

Setting these boundaries allows for the objects in Appendix B to be extracted only if they satisfy the requirements of Table 6.1. From this sample 90 objects passed the requirements.

For these 90 objects to be used as part of Equation 6.1 they need to be assigned a weight via the efficiency parameter ϵ. This process is explored in the next section.

6.3 Efficiency Weighting

In this section I describe the process of assigning weights to the 90 SNe from the sample described above. This method takes advantage of the efficiency grids created in Chapter 4 to compare each object to the simulated objects which closely resemble it. The efficiency grids were built from the simulation of many hundreds of millions of artificial light curves in different simulation areas on the sky. The SALT2 parameters, redshift and coordinate metrics parametrised the grid in a multidimensional space. Each real supernova occupies a distinct location inside the grid defined by the measurements made upon it in Chapter 5. The weighting was calculated by comparing the number of simulated objects recovered at each location to the number that were simulated. This defines the ϵ parameter.

This task was performed sequentially on each object in the catalogue and Figure 6.2 schematically shows the process for a demonstration SN. Firstly appropriate stack efficiency grids were chosen from Chapter 4.4. Each stack describes how efficiently the supernovae passed the selection criteria as a function of the $m_R^{F_{\mathrm{box}}}$ parameter. The real SN being analysed had a corresponding $m_R^{F_{\mathrm{box}}}$ measured in Chapter 5.4 which was used to select the appropriate slice from the stack. This slice was transferred to the next stage where the object's measured light curve parameters define a location inside the grid. A multidimensional linear interpolation was performed at that location to find the fraction of identically parametrised simulated light curves that were recovered. This value, ϵ, was retained and the next object in the list was processed in an identical manner.
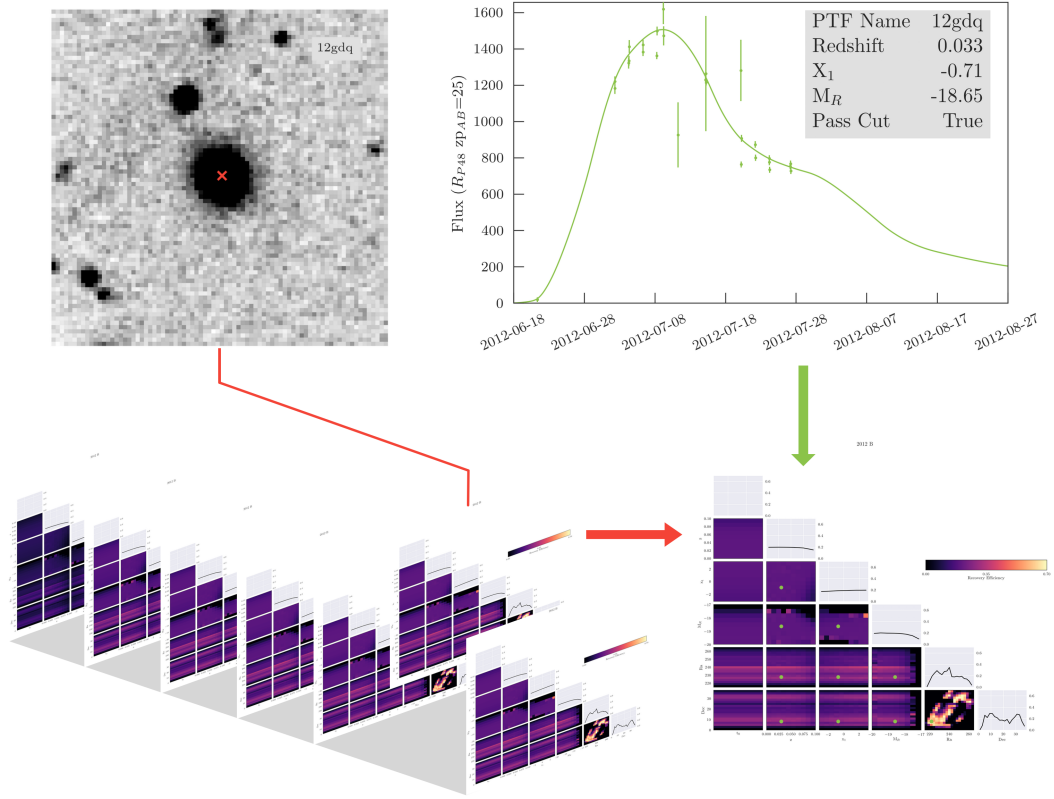
FIGURE 6.2: A schematic of the technique used to determine the object efficiency weighting. The appropriate stack of efficiency grids are chosen from Section 4.4. The host environment parameter (top left), $m_R^{F_{\mathrm{box}}}$, selects the slice from the stack that best describes it. The light curve fit parameters (top right) are used to perform a multi-dimensional linear interpolation on the grid. This step finds the fraction of objects that were most like it in the simulation that were recovered. The green dots show the location of the light curve parameters in the efficiency grid. The value at that location provides the efficiency parameter ϵ.

Once all the objects were processed a catalogue of 90 ϵ values were associated to their respective SNe. At this stage, enough information was readily available to calculate the volumetric rate. Before the summation in Equation 6.1 was performed the statistical uncertainties for the population were analysed.

6.3.1 Statistical Uncertainties

A sample size of 90 objects is sufficiently large that standard Poisson uncertainties, $\sqrt{N}$, could be applied. However, this does not take into account the weighting on each object in the sample. A weight, ω, is defined as $1/\epsilon$. For example, an object with ϵ=0.1 carries 5 times the weight of an object with ϵ=0.5. Therefore, both the final weighted sample and the associated uncertainties are dominated by the low efficiency objects.

To provide a meaningful statistical uncertainty, the weights were normalised such that the new weights, ω_n, summed to 1. The statistical uncertainty, σ_{stat}, then follows as

$$\sigma_{\text{stat}} = \sqrt{\sum \omega_n^2} \times N \tag{6.3}$$

where N is the total number of SNe in the sample. Note that when all weights are equal the $\sqrt{N}$ behaviour is returned. Throughout this chapter, any statistical uncertainties quoted on the rates are calculated through this method.

6.4 The Volumetric Supernova Type Ia Rate

In this section I present the culmination of all the work thus far in this thesis: the volumetric SNe Ia rate. The simulation boxes in Chapter 4.1 were treated in unison and each area converted to a volume through Equation 6.2, with the redshift limits from Table 6.1. From the catalogue of 90 objects passing the cuts to enter the sample the redshift and efficiencies were used as part of the summation. From this I can present the volumetric rate of SNe Ia out to a redshift of z=0.09 as

$$2.42(\pm 0.29) \times 10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} \, h_{70}^3.$$

The uncertainties presented are the statistical uncertainties on the sample of 90 weighted objects. The volume-weighted mean redshift of the sample was z=0.073. In the next section I will investigate the systematic uncertainties associated with this calculation.

6.5 Rate Monte Carlo Simulation and Uncertainties analysis

The rate presented above was calculated from the best-fit SALT2 values and the redshift for each object in the SNe Ia sample. Of course, there is also an uncertainty in each of the SALT2 fit parameters. The model covariance matrix from the fitting process in Chapter 5.1.2 was used in conjunction with the best-fit values to draw realisations of the SALT2 model for each object. These new values occupy slightly different locations inside the multidimensional efficiency grid, thus the objects are assigned different weights. Propagating these effects through the entire calculation and repeating the process many times builds up a distribution of rate results. Additionally, edge cases are able to enter or leave the sample as a new realisations of their light curve parameters result in new

objects passing or failing the cuts in Table 6.1. This effect will be particularly sensitive to objects that sit just on the boundary of passing or failing. A consequence of this is that not only will the sample size change but also the statistical uncertainties on each measurement. Such an analysis into the systematics of a rate calculation would be the most detailed investigation performed to date.

To conduct a study into the rate systematics I took the catalogue of SNe Ia and the corresponding covariance matrix for the model fit. The best-fit value for each parameter were the centroids for the multivariate normal distribution from which a randomised realisation was drawn. The covariance of the fit was preserved in the realisations. The new sample was checked to see which objects passed the cuts in Table 6.1 and those that did were passed onto the rate calculation stage. The rate was calculated in an identical way to Section 6.4 and, along with it, the statistical uncertainty on the sample, the volume weighted mean redshift and the sample size were logged.

In total, 5.6×10^6 realisations of the rate were calculated. The distribution of results is shown in Figure 6.3. The quoted value of the rate is taken as the median of the distribution and the uncertainties capture 68% of the probability around the mean. The volumetric rate and the systematic uncertainties were found to be

$$2.43^{+0.33}_{-0.19} \times 10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} h_{70}^3.$$

The change in the number of objects included in the rate realisation is rather small. Whilst the best-fit parameters produced a sample size of 90 objects, the median number from the Monte Carlo simulation was 89. Of course, as the sample size and the efficiency distribution changed, so did the statistical uncertainty. The distribution of the sample size and statistical uncertainties from the Monte Carlo simulation are presented in Figure 6.4. The median uncertainty was found to be $2.9^{+0.39}_{-0.24} \times 10^{-6}$. This value has uncertainties which are not included in later analysis .

This section has calculated the final volumetric SNe Ia rate out to a redshift of z=0.09 from the Palomar Transient Factory. The uncertainties on this measurement were provided by a Monte Carlo simulation containing millions of rate calculations drawn from realisations of the supernova sample available in Appendix B. I can present the final rate to be

$$2.43 \left(^{+0.29}_{-0.29} \text{ stat} \right) \left(^{+0.33}_{-0.19} \text{ sys} \right) \times 10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} h_{70}^3.$$
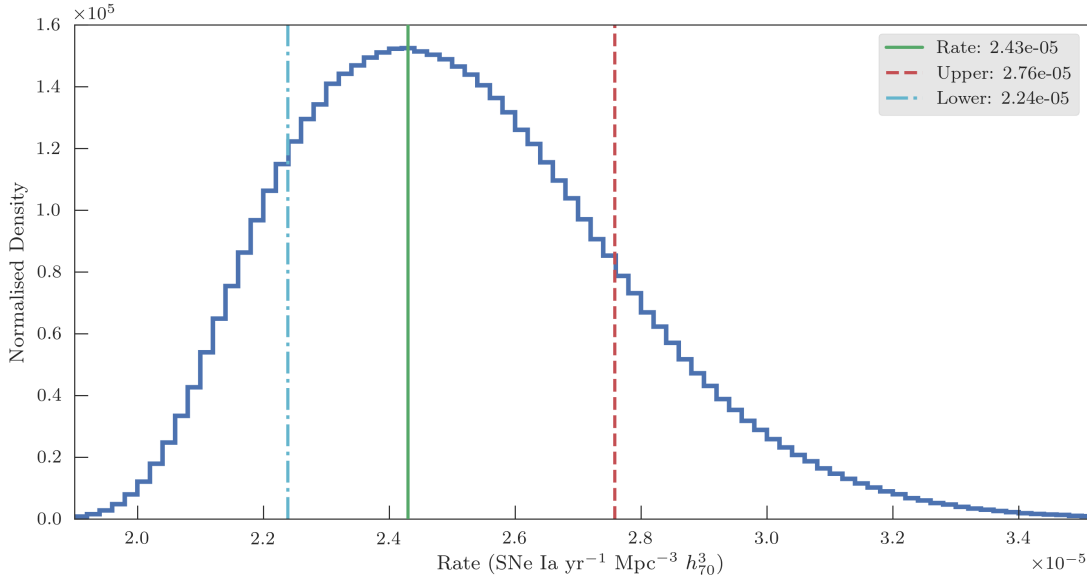
FIGURE 6.3: The distribution of rate values after performing $> 5.6 \times 10^6$ realisations of the rate. Each rate calculated from a new sample of objects drawn from the uncertainties on the model fits. The solid green line represents the median value of the distribution. The red dashed and blue dot-dash represent the upper and lower 1σ values respectively.

6.6 Comparison to other rates

To place this value of the PTF rate into context I will, in this section, compare similar surveys and their volumetric rate results. An overview of the different supernova search techniques was presented in Chapter 1.6. For the purposes of comparing my rate I will compare only results obtained from similar searches. This necessarily excludes the low-redshift rates of Cappellaro et al. (1999) and Li et al. (2011a) for being targeted surveys. Additionally, I exclude the SDSS DR7 rate of (Graur & Maoz, 2013) as this was conducted as a search for contaminant SNe Ia flux in galaxy spectra. I do note however, that their findings are in broad agreement with mine. The PTF rate and those from the comparison surveys are shown in Figure 6.5.

The Supernova Legacy Survey (SNLS) was a multi-colour, rolling, high-redshift search with a primary goal of improving supernova sample statistics for cosmology (Astier et al., 2006; Sullivan et al., 2011a). As an offshoot to this, their 5 year search produced 691 SNe Ia for their a rates analysis (Perrett et al., 2012). The sample was comprehensive enough to cover a large redshift range of $0.1 \leq z \leq 1.1$ and is, to-date, the best constraint of the rate evolution for intermediate redshifts.

The SNe Ia rates of Graur et al. (2011) are a reanalysis of Poznanski et al. (2007) from the Subaru Deep Field (Kashikawa et al., 2004). This high-redshift search was conducted
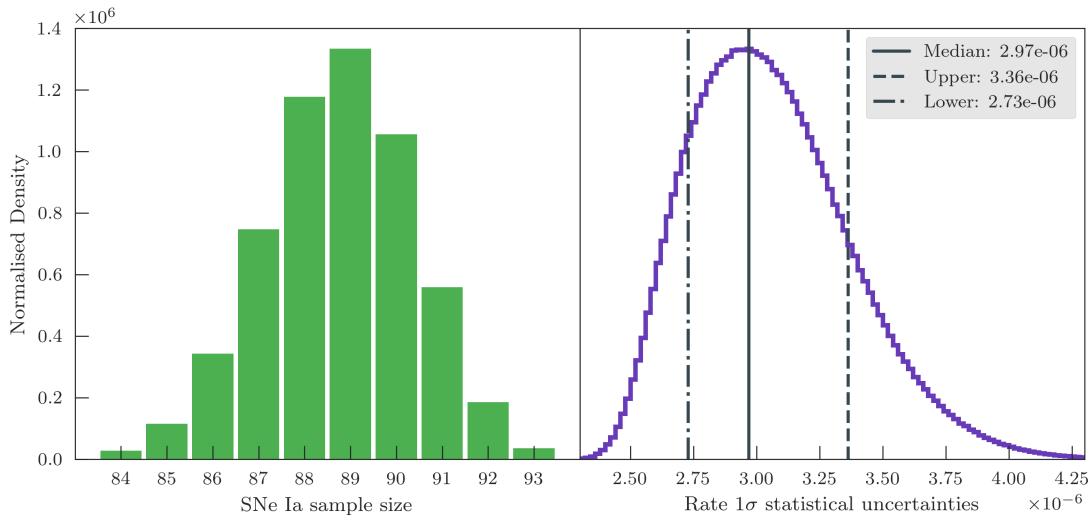
FIGURE 6.4: The left-hand histogram shows the distribution of the sample size, peaking at 89 objects. The right hand panel shows the distribution of statistical uncertainties associated with each rate realisation. The median value of this distribution was 2.97×10^6.

in the R, i$'$, and z$'$ bands to observe the rest frame blue emission from the supernovae. The search for supernovae was conducted on 4 epochs over a period from 2005 March to 2008 June. In total they found 28 SNe Ia in the redshift range z=1–1.5 and 10 in z=1.5–2.0.

Finally, and most comparable to the PTF rate, is the work of Dilday et al. (2010) from the SDSS-II Supernova Survey (Frieman et al., 2008). This branch of the overall survey was carried out on stripe-82 with a 4 day cadence. Their sample totalled 516 SNe Ia, of which 270 were spectroscopically confirmed out to a redshift of z∼0.3. The lowest redshift bin in their study contained only 4 objects, whilst the bin closest in redshift to the PTF result contained 31 SNe. Their statistical uncertainties dominated the low-redshift error budget (z≤ 0.2) and were calculated as the 1σ Poisson uncertainty for N objects in each redshift bin. Clearly, the larger sample of PTF SNe (89 objects) provided improvements on the Poisson uncertainty. Furthermore, the PTF rate measurement explored a more detailed parameter space in the error analysis to include uncertainty in the SALT2 model fit for each SN, and its relation to the associated efficiencies. Such an analysis was not performed for the SDSS calculation, instead model light curves were simulated in the observational data to provide object efficiencies as a function of a single parameter - redshift. Capturing the efficiency in a single parameter excludes information from the rate calculation that considers object-to-object variation. For example, one would assume intrinsically brighter SNe in SDSS would have a higher efficiency compared to fainter events, at the same redshift. These subtleties were lost in the SDSS calculation, but captured in the multidimensional analysis performed for
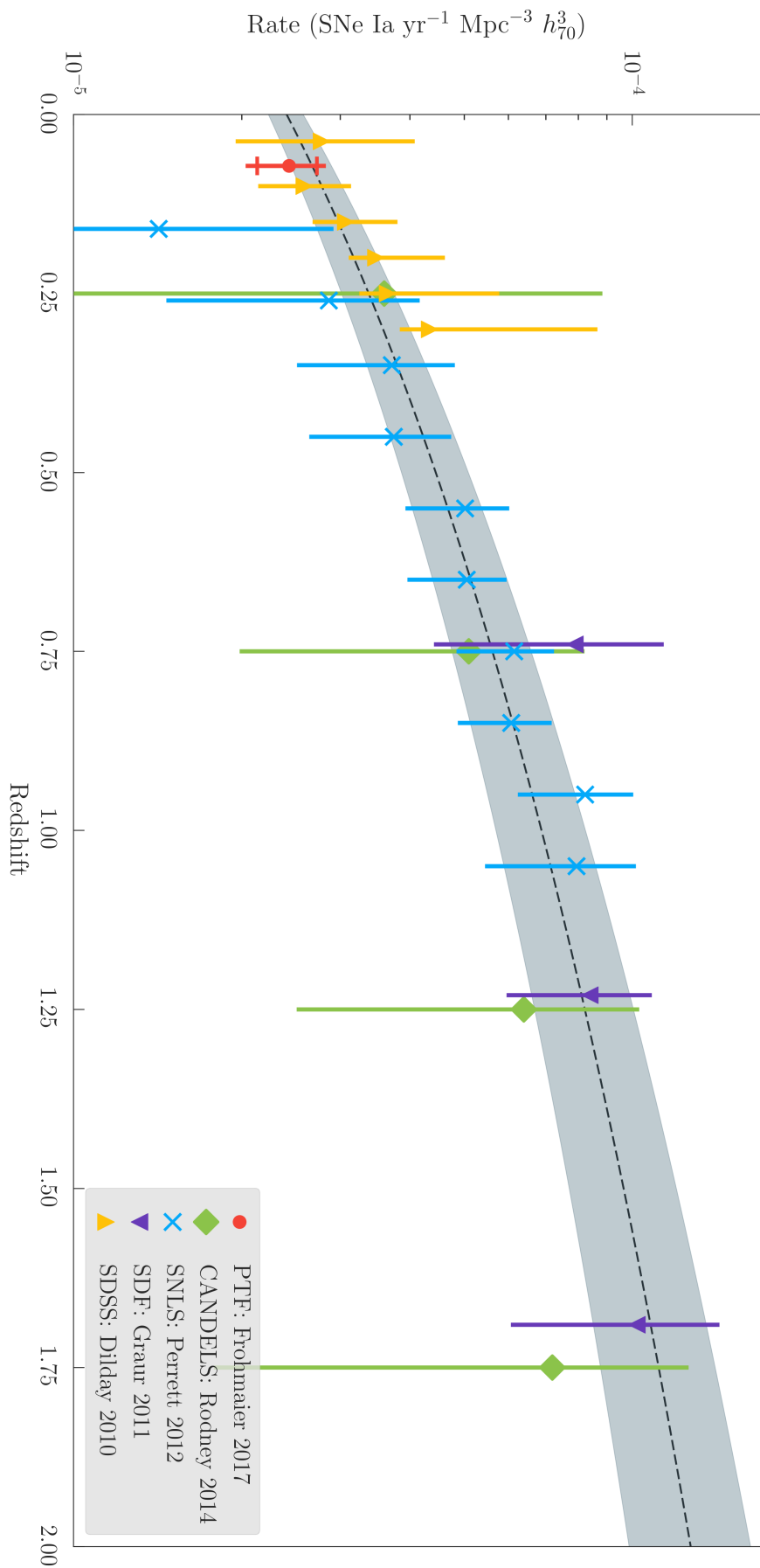
FIGURE 6.5: The PTF volumetric rate of SNe Ia at a volume weighted mean redshift of $z=0.073$ is shown at the red dot. The error bar shows the combined systematic and statistical error bar. The 'cap' along the error bar shows the contribution from the statistical uncertainties. To place the rate in context, several other rates calculated from similar surveys or techniques are shown. The dashed line shows the best-fit rate evolution parametrised in Equation 6.4, the shaded area is the 1σ uncertainty on the fit.

the PTF rate. Marginalising over these other effects in SDSS potentially causes the systematic uncertainties to be underestimated in their work. Overall, the PTF rate not only benefited from a larger, more statistically significant sample, but also from a more detailed error analysis as a multi-dimensional function of several key parameters which determine the detection efficiency, ϵ, in Equation 6.1. Hence, the PTF rate provides the tightest ever constraints on the low-redshift SNe Ia rate.

The rates, and uncertainties, from these various studies were combined. Where necessary the rates were adjusted by a small scaling to ensure that the adopted cosmologies were identical. This process is described in Section 6.6.1. A simple power-law parametrisation of the rate evolution was modelled as

$$\text{SNR}_{Ia}(z) = r_0(1+z)^\alpha \tag{6.4}$$

and fit to the combined dataset. The data were only considered in a fit if the redshift was <1. This necessarily excludes the final bin of the SNLS rate, which is in-keeping with their own analysis, due to incompleteness in their sample. The resulting fit found $r_0 = 2.27(\pm 0.19) \times 10^{-5}$ and $\alpha = 1.70(\pm 0.21)$. This result is consistent with the combined sample from Perrett et al. (2012).

6.6.1   Cosmology Corrections

Different surveys choose to adopt different cosmologies for their own analyses. The cosmological parameters adopted in this work are commonly used, thus requiring no corrections. However, where they do differ, a scaling is applied to match the square of the Hubble parameter[1] for each cosmology at the desired redshift. The Hubble parameter is calculated via the Friedmann equation for a multicomponent universe

$$\frac{H^2}{H_0^2} = \frac{\Omega_{0,m}}{a^3} + \frac{1 - \Omega_{0,m} - \Omega_{0,\Lambda}}{a^2} + \Omega_{0,\Lambda} \tag{6.5}$$

where H is the Hubble parameter at the scale factor, $a = \frac{1}{1+z}$. H_0 is the present value of the Hubble parameter, along with $\Omega_{0,m}$ and $\Omega_{0,\Lambda}$ the present day values of the matter density and vacuum density parameter respectively.

---

[1] Volumetric rates depend on H_0^3

6.7 The Delay-Time Distribution

Further to the simple rate evolution described in Section 6.6, it is possible to combine the the rate measurements from the previous section and the cosmic star formation history[2] (SFH) to constrain the delay-time distribution (DTD) for a SNe Ia population. The DTD describes the supernova rate as a function of time from a δ-function-like burst of star formation. The SNe Ia are the impulse response to the star formation and the time between the two is governed by the underlying progenitor physics. Without a direct observation of an event pre-explosion, supernova rates provide the best insight into possible progenitor systems. The supernova rate as a function of cosmic time can be modelled as the convolution of the DTD (Ψ) and the cosmic SFH

$$\mathrm{SNR}_{Ia}(t) = \int_0^t \mathrm{SFH}(t - \tau)\Psi(\tau)d\tau \qquad (6.6)$$

Different theoretical progenitors produce different DTDs. For example, two distinct populations of progenitors have been speculated in a two-component model (Mannucci et al., 2006) with so-called 'prompt' and 'tardy' coefficients, following star-formation and stellar mass respectively. This is an intriguing model that offers the temptation to attribute such a system to the single- and double-degenerate progenitor scenarios. However, it would be wrong to currently do so, as the biggest limitations lie with poorly constrained high redshift rates (e.g. Graur et al., 2014) - where the prompt channel would dominate - and significant model variation in the SFH, also towards high redshift.

Alternatively, a system composed of two WDs that merge during an in-spiral dictated purely by the radiation of gravitational waves would lead to a delay-time following a power-law t^{-1} distribution (see Greggio, 2005, and the Introduction for a derivation). Whilst a power-law of this form favours the DD scenerio, it does not exclude a contribution from a SD channel (Hachisu et al., 2008).

There are many other delay-time models and proposed distributions in the literature, often covered in reviews (Scannapieco & Bildsten, 2005; Maoz et al., 2012, 2014), however a theoretical delay-time is only as good as the observations and rate calculations that support it. Motivated by this, I explored a simple $\Psi_1 t^{-\beta}$ delay-time distribution as part of Equation 6.6. The combined rates sample used in this part of the analysis is described in Section 6.6.

The current limitation to this study is the variation in the different SFHs, shown in Figure 6.6. Such systematic uncertainties on the DTDs have been extensively explored

---

[2]The star formation history is the star formation rate (SFR) per unit volume as a function of cosmic time.
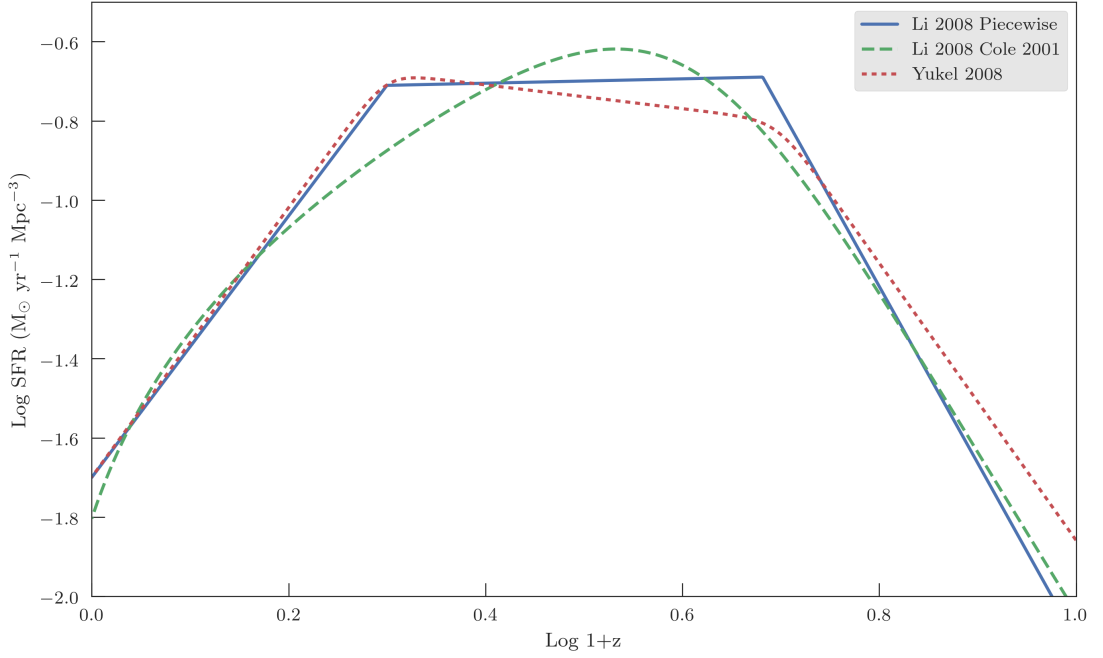
FIGURE 6.6: Star formation rate parametrisations explored as part of the DTD study. These are parametrised by Li (2008) as a piecewise fit and following the form of Cole et al. (2001). The Yüksel et al. (2008) parametrisation is derived from a slightly different dataset and was also included. Clearly the SFHs agree well at low-redshift but significantly differ beyond z>0.5. This motivated the DTD investigation to consider each of these SFHs to investigate systematics in the results.

(Förster et al., 2006; Graur et al., 2011, 2014) so to account for this I performed the model fit to the data using each SFH in turn. The three SFH parametrisations used were the Li (2008) as both a piecewise fit and one which follows the Cole et al. (2001) functional form, and the Yüksel et al. (2008) SFH.

The principles behind finding the DTD were simple and executed through a basic MCMC routine. An initial trial value for Ψ_1 and β were chosen. The first 40Myr of the DTD was set to zero for a hypothetical 8M⊙ star to evolve into the first WD. This DTD was then convolved with the SFH to produce a presumed SNe Ia rate. The combined PTF, SDSS, SDF, SNLS and CANDELS dataset, described in Section 6.6, were compared to this model rate to produce a χ^2 goodness-of-fit metric. This entered the routine's likelihood function and subsequent MCMC iterations were performed for the function to converge onto a solution.

SFH	β
Li 2008 piecewise	0.95($\pm$0.07)
Li 2008 Cole 2001	1.07($\pm$0.1)
Yüksel	0.94($\pm$0.08)

TABLE 6.2: Best-fit $t^{-\beta}$ values for the DTD.

The best fitting rate evolution for each SFH are shown in Figure 6.7 and are clearly well constrained at low redshift. The best fitting power-laws are all consistent with a $\beta = 1$, regardless of the SFH parametrisation, and are shown in Table 6.2. Such a power-law is consistent with a double degenerate (DD) progenitor scenario and found by a multitude of similar studies (Dilday et al., 2010; Perrett et al., 2012; Graur & Maoz, 2013; Maoz et al., 2012; Graur et al., 2014). The largest uncertainties in the DTD come from the SFH discrepancies. The total 1σ uncertainty is shown by the grey shaded region in Figure 6.7.

My PTF rate adds an anchor to the low redshift regime and undoubtedly the data are described here by a t^{-1} DTD. Further improvements in our understanding of the progenitor systems to SNe Ia would come from equally well constrained high redshift rates. This would motivate new DTD models that could handle different progenitor contributions that dominate at different redshifts.
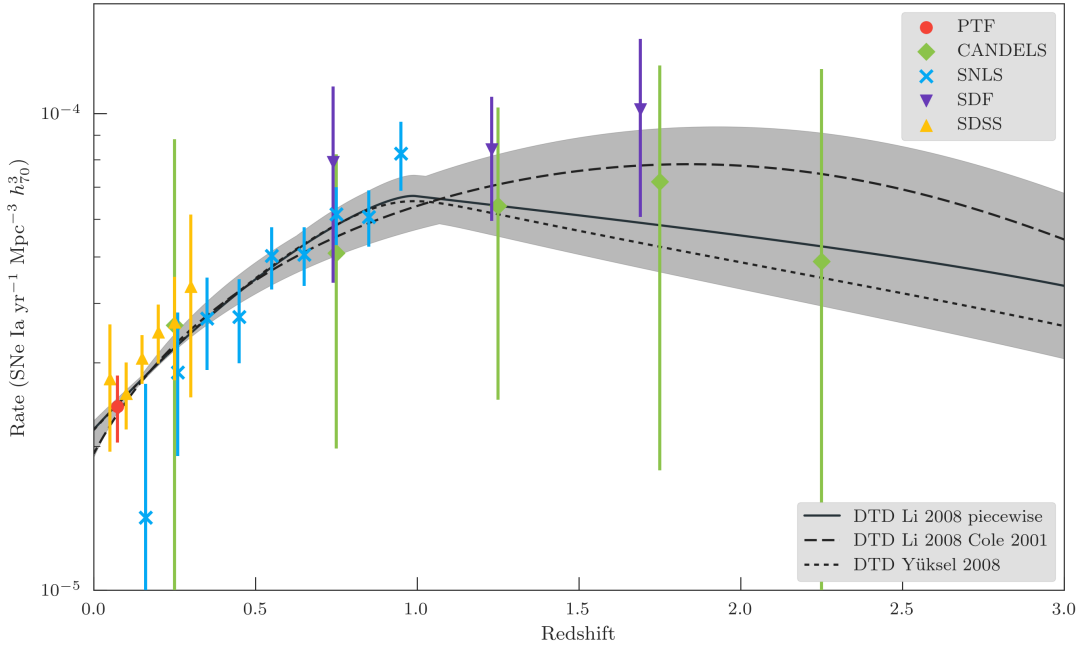
FIGURE 6.7: The best fit rate evolutions for a simple power-law $t^{-\beta}$ DTD. All values of β were consistent with 1. This generically explains the DD progenitor scenario. The shaded grey regions show the combined 1σ uncertainties on the fit resulting from the different SFH models. Clearly the data are well constrained at low redshift. Better rate studies and/or SFH constraints would help show where the SNe Ia rate turnover occurs as currently it is seen broadly between z=1–2.

6.8 Efficiency weighted functions

The final investigation in this chapter involves the weighting of both luminosity and x_1 distributions by the efficiencies found for each object. Un-weighted luminosity functions suffer from observational bias, as the brighter objects are typically easiest to find. The efficiencies calculated for each object in the rate sample reveals the 'True' number of events the single object represents.

I have already presented the justification for performing a Monte Carlo simulation of light curve realisations to account for systematic uncertainties in the light curve model fits in Section 6.5. Therefore 100,000 samples were created resulting in $> 8.8 \times 10^6$ objects drawn from the observed population of SNe Ia. Each object was weighted by the inverse of its efficiency, ϵ. The resulting histograms for x_1 and M_R are shown in Figure 6.8. The normalised distributions both before and after efficiency weighting were compared.

Immediately the luminosity function stands out for its seemingly bi-modal distribution (Figure 6.8, lower right panel). The main peak is seen at $M_R = -18.85$ but the second peak is at $M_R = -18.0$ - right at the sample cut off. The increase seen in the weighted distributions suggests that a few low-efficiency objects dominate the faint end of the sample. The same is not seen at the other end of the distribution. This could possibly be an effect of 91bg-like SNe, which are just on the cusp of the sample parameters, entering realisations of the distribution. Their under-luminous nature would naturally see them assigned a high weighting, and thus the emergence of a second peak in the distribution. There are no objects brighter than -19.4 that will ever make the sample, and even if they did their brightness naturally assigns them a high efficiency and thus a small contribution to the weighted sample. The x_1 distribution peaks at -0.25, although is rather more broad than the M_R distribution. The slight upturn in the distribution at the low x_1 end is due to the contribution of low efficiency objects for the same reasons as the M_R distribution.

6.9 Summary

In this chapter I have put all the pieces of the puzzle together to calculate the volumetric rate of SNe Ia out to a redshift of z=0.09. The sample of SNe Ia described in Chapter 5 were compared to the simulated objects of Chapter 4 to provide weights for each object. The SNe that passed the cuts described in Table 6.1 entered the rate calculation of Equation 6.1. Solving this for the best-fit sample of SNe Ia produced a rate of
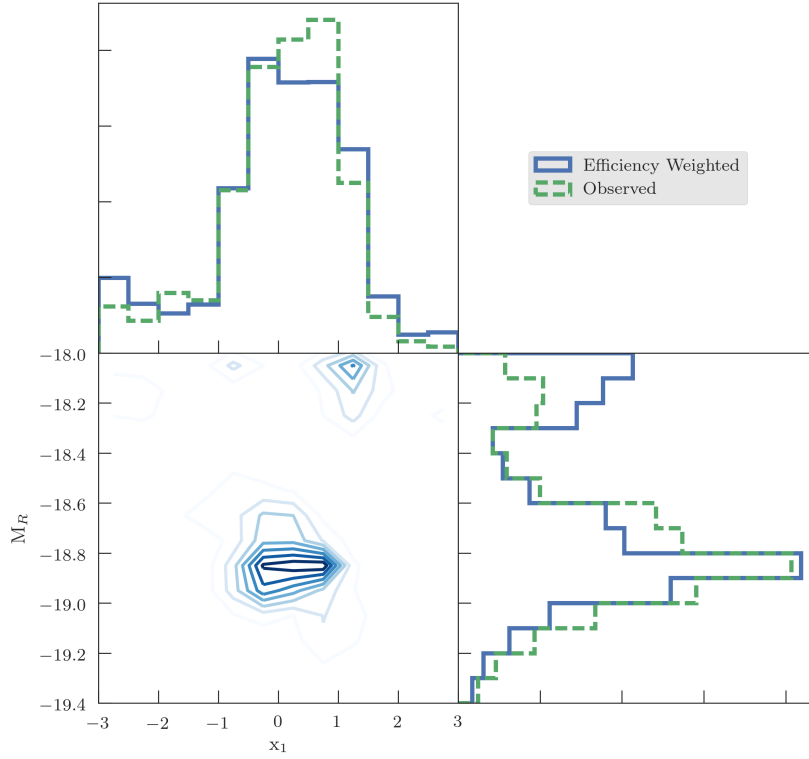
FIGURE 6.8: The efficiency weighted histograms for the x_1 (top panel) and M_R (bottom right) are shown as the solid lines. The dashed lines show the raw distribution before efficiency weighting. The efficiencies have a strong contribution to the lower end of the luminosity function as few very low efficiency objects boost the sample.

$$2.42(\pm0.29)\times10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} \ h_{70}^3.$$

where the quoted uncertainties are the 1σ statistical uncertainties calculated from the distribution of weights according to Equation 6.3.

The best-fit sample of objects does not necessarily capture all the objects which could have made the final sample. Each object has a uncertainty on the SALT2 model fit. Therefore, to investigate the systematics associated with this, millions of realisation of the rate were performed. For each realisation the potential sample of SNe were re-drawn from the multivariate normal distribution of their model fit covariance. The sample was then checked to see which of these objects passed the selection criteria and the rate was recalculated. This process built up a distribution of rates and found the rate to be

$$2.43 \left(^{+0.29}_{-0.29} \text{ stat}\right) \left(^{+0.33}_{-0.19} \text{ sys}\right) \times10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} \ h_{70}^3.$$

where the systematic uncertainties were calculated by finding the 68% confidence interval around the mean. The new quoted rate value is the median of this distribution. The number of objects that passed the cuts was liable to change each realisation. The mean number of objects that entered the rate was 89. The realisation built up a distribution of statistical uncertainties with a median 1σ uncertainty of 2.97×10^6 SNe Ia yr^{-1} Mpc^{-3}.

Furthermore, a simple rate evolution of the form $r_0(1_+z)^\alpha$ was fit to a combined sample from other surveys. I found $r_0 = 2.27(\pm0.19) \times 10^{-5}$ SNe Ia yr^{-1} Mpc^{-3} h_{70}^3 and $\alpha = 1.70(\pm0.21)$, these are consistent with other combined rate analyses.

Following on from this rigorous calculation of the rate and the uncertainties I explored fitting a model delay-time distribution. The DTD reveals underlying properties of the progenitor scenario that contributed to the measured SNe Ia rate population. I fit a hypothetical rate evolution by convolving the DTD, described by a simple power-law, of the form $\Psi_1t^{-\beta}$ to the cosmic SFH. I chose to use 3 different parametrisations of the SFH to investigate the systematics with the SFH that differ at high redshift (z>1).

The best fitting β values are shown in Table 6.2 and are all consistent with β=1. This is consistent with a progenitor population dominated by the double degenerate scenario.

In the next chapter I will present further applications of the efficiency study but on a different population of transient; calcium-rich supernovae. These SNe are the newest additions to the known supernova classes. Sample sizes, across all surveys, are small but the content of this thesis provides a unique opportunity to investigate these enigmatic objects.

Chapter 7

Calcium-rich supernovae in PTF

The efficiencies framework developed in this thesis was designed to accept any transient model light curve. The obvious place to start was from the most well studied-transients, SNe Ia. Results from that investigation demonstrated the effectiveness of the methodology.

Naturally, it is now compelling to investigate the 'weirdos' in the data. One of the key science achievements of PTF was to begin to probe the previously unsurveyed 'gap' between classical novae and supernovae (Kasliwal, 2011, and Figure 1.1). Among these so-called 'gap transients' were one of the most recent additions to the SN family: Calcium rich supernovae (Ca-rich SNe or Ca-rich gap transients). The first observations of these transients were made in in the early 2000s by Filippenko et al. (2003), who identified them as SN Ib- or Ic-peculiar with strong calcium emission. Such is their elusive nature, it wasn't until SN2005E (Perets et al., 2010) and objects discovered in PTF (Kasliwal, 2012) that they obtained their own class. In this chapter, I investigate the rates of Ca-rich supernovae and the implications for the calcium abundances in the intra-cluster medium (ICM).

Kasliwal (2012) defined the five characteristics of a Ca-rich gap transient to be

1. Peak luminosities between classical novae and supernovae

2. Faster photometric evolution (rise and decline) than supernovae

3. Photospheric velocities comparable to supernovae

4. Early evolution to the nebular phase

5. Nebular spectrum dominated by calcium emission

The literature sample of Ca-rich gap transients remains small, with fewer than ∼15 candidate SNe (Filippenko et al., 2003; Kawabata et al., 2010; Perets et al., 2010, 2011; Valenti et al., 2014). These include potential Ca-rich transients that lack the photometric and spectroscopic coverage to conclusively determine their type (e.g. SN2000ds, SN2001co, SN2003H, SN2003dg, SN2003dr, and SN2005cz). PTF added a further 5 to the known sample of Ca rich events: PTF10iuv, PTF11bij (Kasliwal et al., 2012), PTF11kmb and PTF12bho (Lunnan et al., 2017). The PTF sample of Ca-rich transients is the largest set of confirmed events from a single survey. This places me in a unique position to calculate the `first` rate of Ca-rich SNe. The method for calculating the Ca-rich rate is as follows.

- Ca-rich objects were identified from the real-time photometry.

- Improved photometry was obtained from external sources.

- A set of template light curves were created to describe the Ca-rich population.

- The template light curves were placed in simulations of PTF observations.

- The templates were observed and either 'discovered' or not.

- The fraction of templates, as a function of several parameters, that were discovered defined the recovery efficiency.

- The real Ca-rich transients were compared to the simulations and weights for each real object were obtained.

- These weights were used to calculate the Ca-rich rate.

This chapter also investigates the intra-cluster medium elemental abundances. Studies of the intra-cluster medium show a Ca/Fe over-abundance (de Plaa et al., 2007; Mernier et al., 2016). The origin of this is unknown, and traditional supernova classes (e.g. CCSNe, SNe Ia) cannot account for this discrepancy. In Figure 7.1, I have shown the observed abundance ratios (X/Fe) of silicon, sulphur, argon, calcium, iron and nickel. The combined nucleosynthetic products of CCSNe and SNe Ia are also shown in the figure to highlight the discrepancy between the models and the observations (de Plaa et al., 2007). Mulchaey et al. (2014) showed that a non-negligible contribution from Ca-rich transients may be able to provide the necessary Ca to solve this problem. Furthermore, the remote locations of these transients could efficiently pollute the ICM with Ca-rich ejecta, as it doesn't have to escape the gravitational potential of a host galaxy (Zaritsky et al., 2004). The rate of Ca-rich SNe I present here may provide more insight into the ICM abundance problem.
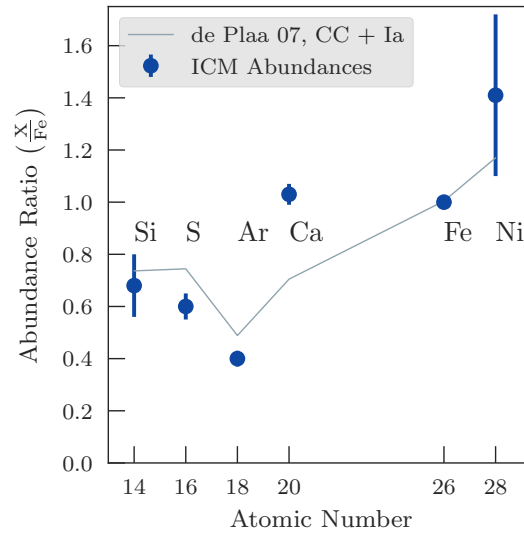
FIGURE 7.1: The observed abundance ratios of the ICM (de Plaa et al., 2007). The grey line shows the contribution from CCSNe and SNe Ia. These SNe types clearly do not produce the required abundance ratio of Ca/Fe.

I performed the model-fits, real-time photometry analysis, light curve efficiency simulations, final rate calculation and ICM abundance matching. The extended photometry, from which I created Ca-rich light curve models, was provided by K. Maguire. The work of this chapter will be published in Frohmaier et al. (in prep.).

7.1 The Ca-rich Sample

It should be noted that, the Ca-rich analysis only includes spectroscopically confirmed Ca-rich SNe in PTF. I do not search the PTF catalogues for additional objects, as this was done by Lunnan et al. (2017). They identified PTF11kmb and PTF12bho as new members of this Ca-rich class. They further speculated that PTF10hcw was a potential candidate from the emerging [Ca II] lines in the early spectrum, similar to 2005E. However, with only two P48 light curve points, and one from the P60, the data were not conclusive and the object would not pass my quality cuts. Hence, it is not considered as part of this analysis.

Out of necessity, PTF09dav was also excluded from the sample. The reasons were two-fold. Firstly, I demonstrated in Chapter 3.5.2.2 that the efficiencies in 2009 were unreliable. Secondly, PTF09dav shows detections of scandium and strontium in its spectrum (Sullivan et al., 2011b). Not only are these features not seen in other Ca-rich transients, but scandium has never even been seen in a WD explosion before.

This leaves the sample of objects for light curve analysis as: PTF10iuv, PTF11bij, PTF11kmb, PTF12bho.

7.1.1 Real-time observations

Each of the four Ca-rich supernovae had light curves produced from the real-time photometry, following the description in Chapter 5.1.1. The R-band light curves are shown in Figure 7.2. Immediately, it is clear that three of the objects show transient evolution, characterised by both a rise and decline. The fourth object, 11bij, only has a few detections spread across several hours, and they are close to the detection limit.

Quality cuts were imposed, such that a light curve must have at least four points separated by 12 hours. This excluded 11bij from the later rate calculations. Indeed, 11bij was only confirmed as a transient after routine follow up observations of the field, some 10 days later, by the P60 telescope. Whilst 11bij was excluded from the rates analysis, it was still used to construct template light curves when the photometry was improved in Section 7.1.1.1. The final sample of Ca-rich SNe that entered the rates calculation included: PTF10iuv, PTF11kmb, and PTF12bho.
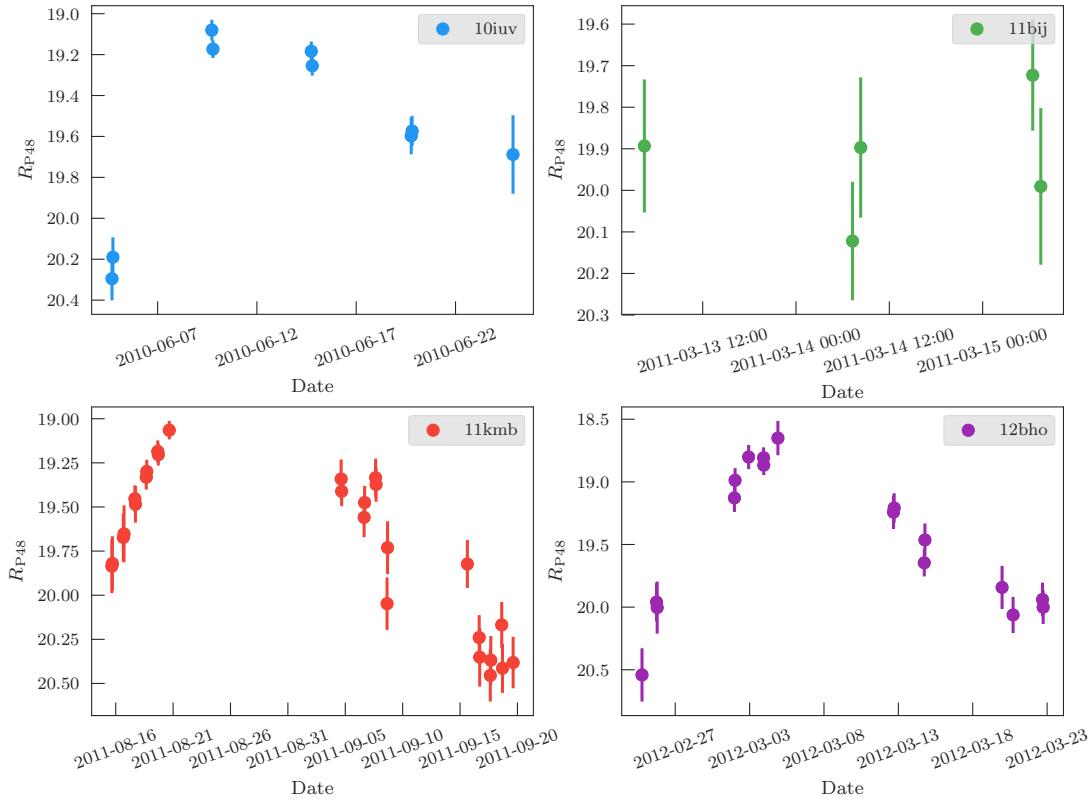
FIGURE 7.2: The real-time photometry for the Ca-rich objects discovered by PTF.

7.1.1.1 Improved Photometry

Follow-up observations of the Ca-rich SNe provided additional photometric points on the light curve. Observations from the P60 telescope (Cenko et al., 2006) were converted to approximate R_{P48} magnitudes. Furthermore, any calibrated spectra of the targets had spectrophotometry performed on them by integrating the SED through the R_{P48} system throughput. These photometric points were also placed on the light curve. The photometry for all the template light curves is shown in Figure 7.3.

7.2 Creating a Ca-rich light curve templates

Unlike for SNe Ia, there is no empirically or physically derived model to describe the evolution of Ca-rich transients. In order for a transient to be analysed within the efficiencies framework, a model light curve was needed. A phenomenologically-motivated model introduced by Bazin et al. (2009) was fit to the R-band light curves. The model is of the functional form

$$f(t) = A \frac{e^{-(t-t_0)/\tau_{\mathrm{fall}}}}{1 + e^{-(t-t_0)/\tau_{\mathrm{rise}}}} + B \tag{7.1}$$

where A is a normalisation factor and B is a constant offset. τ_{rise} and τ_{fall} are the rise and fall times of the light curve respectively. t_0 is a parameter related to the time of maximum brightness by

$$t_{\mathrm{max}} = t_0 + \tau_{\mathrm{rise}} \ln \left(\frac{\tau_{\mathrm{fall}}}{\tau_{\mathrm{rise}} - 1} \right). \tag{7.2}$$

This model was fit to the photometry with a constraint that $\tau_{\mathrm{rise}}/\tau_{\mathrm{fall}} \leq 1$. This was to ensure the light curve made physical sense, i.e. a rise followed by a decline. The best fit parameters and properties of each object are shown in Table 7.1.

Name	R.A.	Dec.	z	M_R	A	B	t_0	τ_{rise}	τ_{fall}
10iuv	259.2261	31.5643	0.020	-15.66	-5.25	22.64	-6.97	-3.53	-4.01
11bij	194.7432	37.3866	0.035	-15.69	-12.66	30.43	-14.98	-5.81	-6.15
11kmb	335.7234	36.2935	0.017	-15.61	-4.01	20.73	3.05	-5.22	-12.60
12bho	195.3193	28.0218	0.023	-15.94	-4.18	22.00	-6.30	-3.39	-3.78

TABLE 7.1: The best-fit results for each of the Ca-rich SNe. M_R is the peak absolute magnitude. The final five columns show the values of the Bazin et al. (2009) model for this fit.

The improved photometry and the best-fit model are shown in Figure 7.3. Clearly, the Bazin et al. (2009) model describes the light curve evolution well. The dataset was
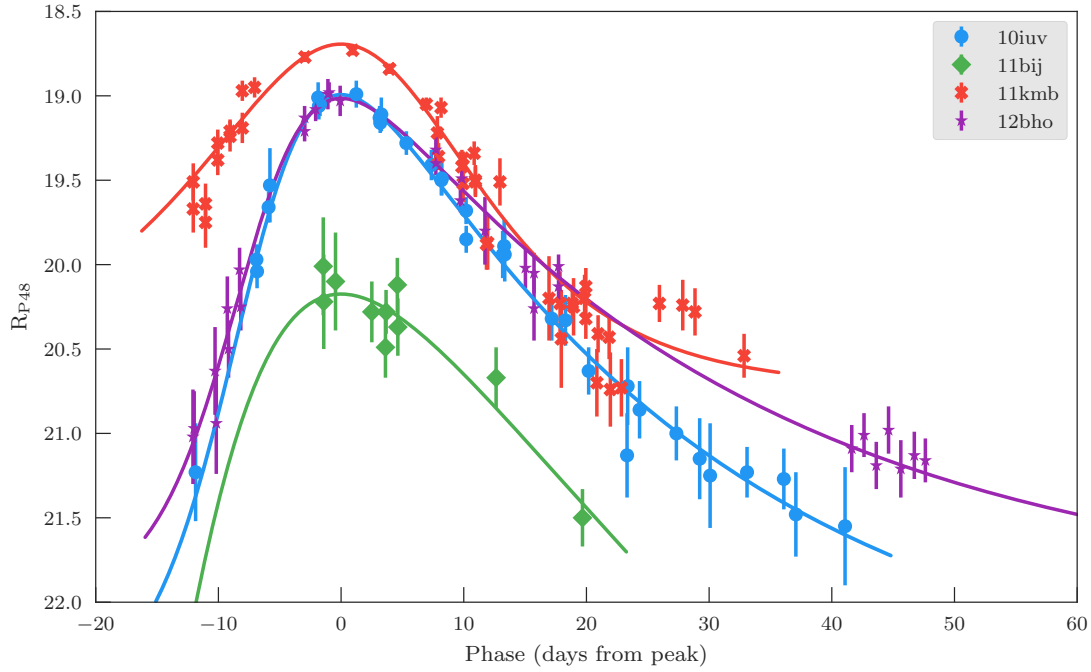
FIGURE 7.3: A phenomenological light curve model introduced by Bazin et al. (2009) was fit to a combined set of photometry.

too small to create a universal light curve template capable of describing any Ca-rich transient. Therefore, each template was used separately in later simulations.

For the simulation stage, the light curves were translated into their absolute magnitudes. This was done using their spectroscopic redshifts, and the resulting light curves are shown in Figure 7.4. There is remarkable homogeneity in the objects, which, just like for SNe Ia, might provide an insight into the explosion mechanism. However, with only four objects analysed, the samples are far too small to draw any significant conclusions. The light curves shown in Figure 7.4 are considered as templates for the simulations.

7.3 Simulations

In this section I will detail the process of calculating Ca-rich specific recovery efficiencies. This is extremely similar to the process presented in Chapter 4.2 for the SNe Ia.

Firstly, areas and time spans were chosen such that the template events could be simulated to determine the recovery efficiencies. This was achieved by visually inspecting the PTF observing footprints around the time each Ca-rich SN was first detected. Areas were selected to be large and consistent in cadence. This presented a challenge for PTF12bho, which was discovered during a DyC experiment. Typically in the DyC
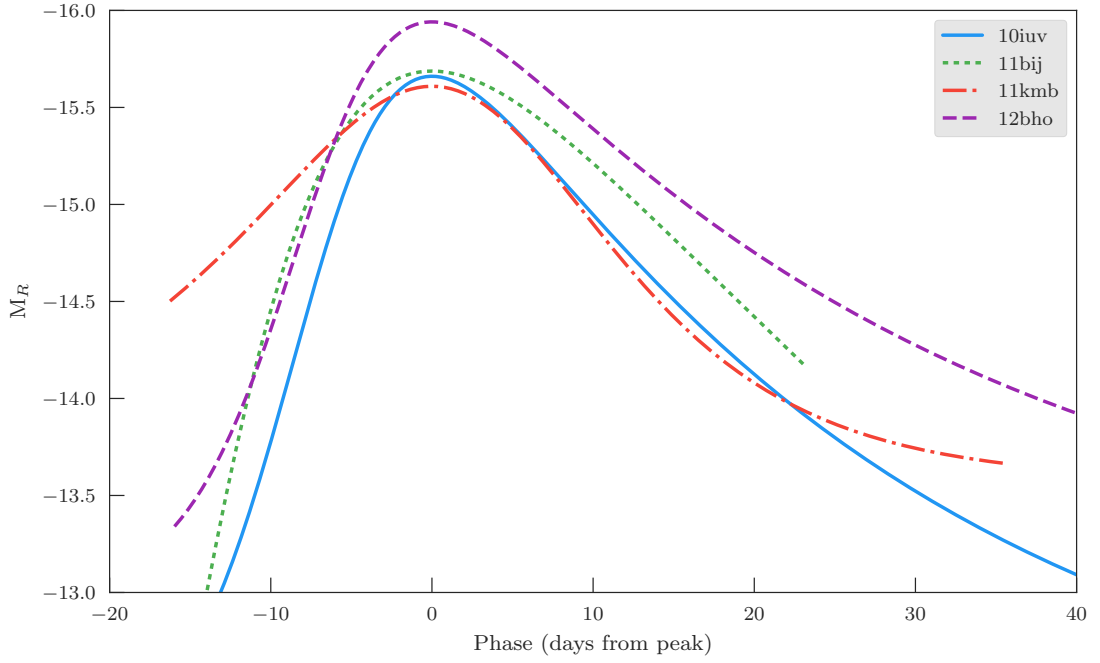
FIGURE 7.4: The absolute magnitudes of the Ca-rich SNe.
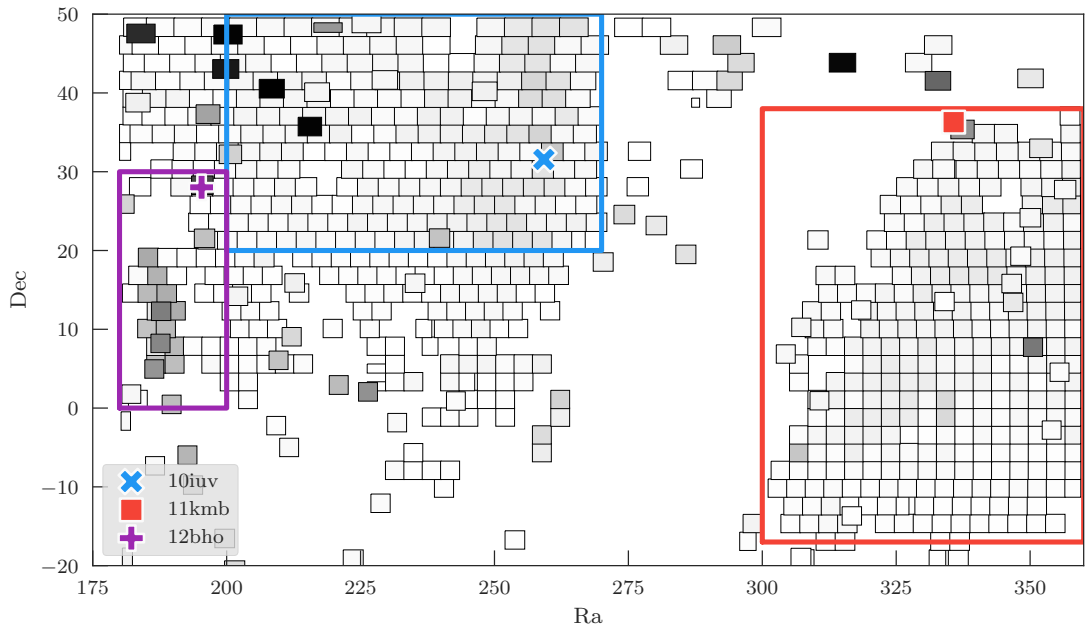
FIGURE 7.5: The area of the sky the simulations were performed in are bound by the boxes. The locations of each Ca-rich SN are marked. The areas were chosen to have a constant cadence for the longest duration achievable.

experiments, the target fields were not tiled together. This left large parts of the sky unobserved, and simulating SNe there was an inefficient use of computing resources. However, small but consistent footprints were found and are shown in Figure 7.5, along with a description for each area in Table 7.2.

	Object	Area ($\deg^2$)	Nights
■	PTF10iuv	1262.91	165
■	PTF11kmb	1480.81	135
■	PTF12bho	227.68	160

TABLE 7.2: The areas and number of simulated night for each object. The areas are a description of the footprints that were able to find Ca-Rich SNe.

The observing footprints defined the areas from which recovery efficiencies were determined through the simulation of light curves. The principles behind this process for the Ca-rich SNe was as follows. A randomly drawn R.A. and dec. set the location for the event, and both a random explosion date and redshift were assigned. A redshift was drawn uniformly such that $0.0035 \leq z \leq 0.04$. These limits are between the redshift a Ca-rich SNe would saturate the PTF detector, and when the SNe would be too faint to detect. The R.A., dec. and explosion date were specifically drawn to be within the constraints of the footprint being simulated in.

Each Ca-rich template was sequentially evaluated at these spatio-temporal metrics. The PTF observing logs were queried for all observations of the location, for a total period of 15 days pre- and 40 days post-peak. Each template was converted from M_R to an observed magnitude using the random redshift. Time dilation corrections were not applied to the light curve, because the effects of $(1+z)$ at these redshifts are small. The epochs PTF would have observed the simulated event were compared to the template light curve to find the apparent R-band magnitude, and the observing log returned the necessary observing conditions. The $m_R^{F_\text{box}}$ parameter (Chapter 3.5.2.1) was assumed to fall under the faintest regime, as these events are remote from their host galaxies.

The synthetic light curve was passed through the single-epoch recovery efficiencies, and the detectability on each object calculated as for the SNe Ia (Chapter 4.2.5).

Each light curve was then checked to see if it passed the coverage cuts. If it did, the object was considered 'discovered'. All the templates were simulated in equal proportion across all the areas under consideration. In total $>2.\times10^7$ Ca-Rich SNe templates were simulated.

7.4 The Rate of Ca-rich SNe

In this section, I calculate the rate of spectroscopically confirmed Ca-rich SNe in PTF. Firstly, efficiency grids were constructed so that the real transients could have weights assigned. The details of constructing the multi-dimensional efficiency grids are presented in Chapter 4.5. The efficiency grids describe the fraction of simulated Ca-rich transients in each parameter bin that passed the coverage cuts, compared to the number that were simulated. An efficiency grid was constructed for each area simulated in Figure 7.5.

The axes of the efficiency grid were redshift, R.A. and Dec. Once again, the time axis was marginalised over, as the footprints were designed to be consistently sampled in time. Unlike for the SNe Ia, the light curve model parameters were not used as axes. This is because the simulation captures the variation in the templates and the final efficiency is calculated as an average over all the templates. Furthermore, with only three Ca-rich objects in this calculation, the statistical uncertainties dominate over any systematics associated with template variation. The spatial efficiencies are shown in Figure 7.6 with the location of the Ca-rich transients marked.
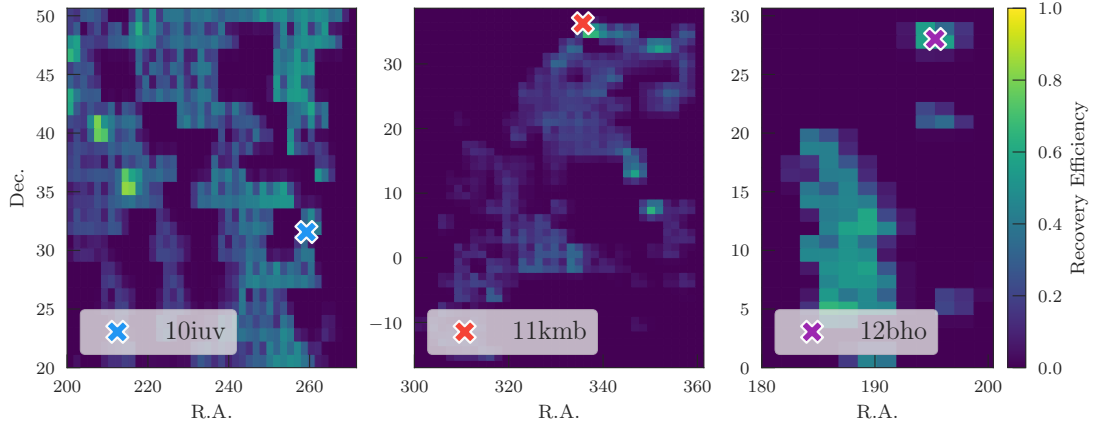
FIGURE 7.6: The spatial recovery efficiencies for each object I studied. The value at the location of each cross, and in an extra redshift dimension, is used to weight each object in the sample.

7.4.1 Ca-rich Rate Calculation

The rate equation for the Ca-rich rate is

$$r_{\mathrm{Ca}}(z) = \frac{1}{\mathrm{V}\Delta\mathrm{T}} \sum_{i=1}^{N} \frac{1 + z_i}{\epsilon_i} \tag{7.3}$$

This is identical to the SNe Ia rate equation and as such the volume is described by Equation 6.2. The redshift limits for the volume were set to be between 0.0035 and 0.036. The lower redshift limit was set for the distance the brightest template would saturate PTF. The upper redshift limit was set at the distance the brightest point on the faintest template falls below the 50% completeness limit of 20.3 mag, shown in Figure 3.11. The efficiencies, denoted by the parameter ϵ, were found by performing a multi-dimensional linear interpolation at the location of the real Ca-rich's R.A., Dec. and redshift. These values were used in Equation 7.3.

The efficiencies found for each Ca-rich SN were similar, hence I was able to use the 1σ poisson uncertainties for 3 objects. The final Ca-Rich rate was found to be

$$1.62(\pm 1.07) \times 10^{-5} \text{ SNe yr}^{-1} \text{ Mpc}^{-3} \ h_{70}^3.$$

The volume-weighted mean redshift of the sample was z=0.021.

This is the first Ca-rich rate to be calculated and, thus, there are no other measurements to compare to. Instead, I compare the Ca-rich rate to the SNe Ia rate evolution and find the rate to be ∼60% of the SNe Ia rate at the same redshift. Of course, the uncertainty on this result is large. In Figure 7.7 I compared the Ca-rich rate to the low-redshift SNe Ia rates from other sky surveys.

In the next section, I take advantage of recent modelling advances to explain faint-and-fast transient evolution. I compare model predictions of elemental yields to calcium in the intra-cluster medium. I then assess whether the rate of Ca-rich transients can solve the "cluster calcium conundrum".

7.5 Intra-cluster Medium Calcium Abundances

Recent efforts have been made to explain the observed light curves of under-luminous SNe through explosion modelling (Bildsten et al., 2007; Shen et al., 2010; Waldman et al., 2011; Woosley & Kasen, 2011; Sim et al., 2012). Invariably, these models require He on the surface of a CO or ONe WD. This explosion physics is similar to some SNe Ia models, but, for under-luminous events, the shock-wave doesn't propagate through the WD and trigger a core-explosion. These models predict ejecta masses in the region of 0.01 to 0.3 M$_\odot$.

Ca-rich SNe are an attractive source to solve the ICM Ca/Fe over-abundance because of their explosion products. The He detonation produces very small amounts of Fe, Co
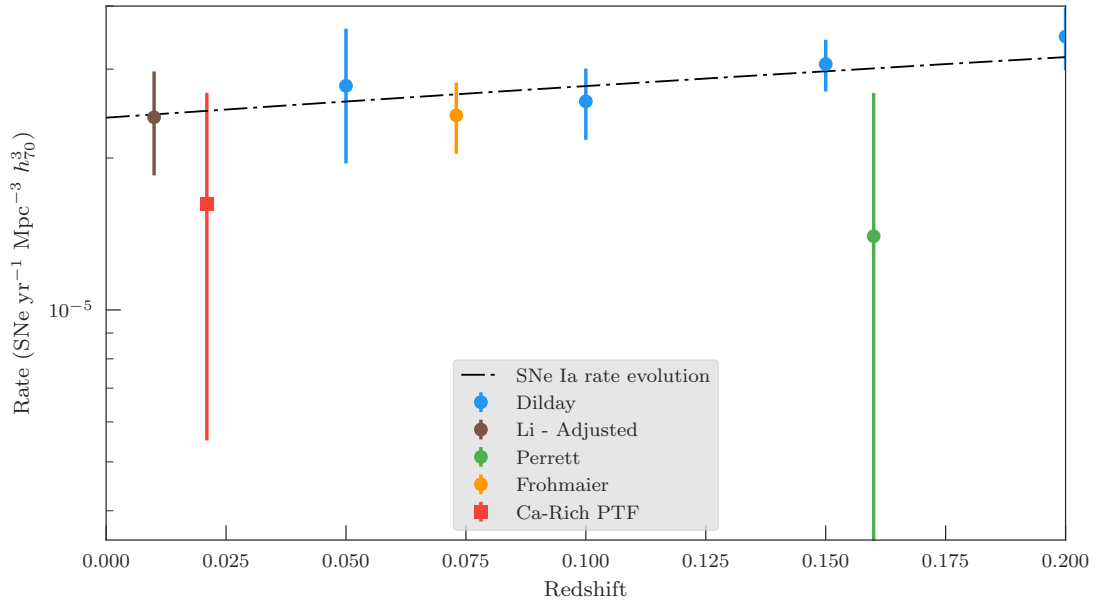
FIGURE 7.7: The rate of spectroscopically confirmed Ca-rich SNe in PTF. I find the rate to be ∼60% the SNe Ia rate at the same redshift. The Li et al. (2011a) was adjusted to only include 'normal' SNe Ia following the method of Perrett et al. (2012, also shown, in green). The SDSS rate is in blue (Dilday et al., 2010) and the PTF rate from Chapter 6 is in orange. The circle markers are SNe Ia rate results and the square marker is the Ca-rich SNe rate. The dot-dash line described the SNe Ia rate evolution calculated in Chapter 6.5

and Ni intermediate mass elements (IMEs), but relatively large amounts of Ca, Ti and Cr.

For this investigation, I adopt the model of Waldman et al. (2011). The bolometric luminosity of the prototypical Ca-rich transient, 2005E, was well reproduced by one of their proposed explosion scenarios. They found that a $0.45M_\odot$ CO WD with a $0.2M_\odot$ shell of He was the best match. However, they do note that the light curve decline is more rapid in their models than was observed. In total they produced 8 models, and I used each in the ICM abundance analysis.

7.6 ICM abundance matching

A sample of 22 clusters were observed by de Plaa et al. (2007) using the XMM-Newton X-Ray observatory. From these observations they derived abundance ratios (X/Fe) for silicon, sulphur, argon, calcium, iron and nickel. These abundances are shown in Figure 7.8.

To reproduce the observations, I considered the elements produced from three SN chan-
nels - Ca-Rich, CCSNe and SNe Ia. From theoretical models of the nucleosynthesis, I
considered the relative rates of each channel needed to best replicate the observations.
I introduce a Ca-rich explosion model into the analysis under the following method:

- Several Ca-rich explosion models were obtained from Waldman et al. (2011).

- The explosive products of Ca-Rich SNe, CCSNe and SNE Ia model channels were
  combined.

- The relative rates of each channel were calculated to best replicate the observed
  ICM abundances.

- The ratio of the Ca-rich SNe and SNe Ia rates (calculated in this thesis) were
  compared to the rates from the abundance predictions.

For each of the six elements, the abundances were fit using a least squares method to

$$\eta_{\text{tot}} \sum_i \alpha_i N_{i,j} = X_j \qquad (7.4)$$

where i represents the different SN types. η_{tot} is the total number of SNe needed to
replicate the observations, α_i is the fractional contribution, and X_j is the observed
abundance of the jth element. Finally, $N_{i,j}$ is calculated from the supernova models and
represents the number yield of each element produced

$$N_{i,j} = \frac{Y_{i,j}}{\mu_j \xi_j} \qquad (7.5)$$

where $Y_{i,j}$ is the predicted mass yield per supernova for each channel, μ_j is the mean
atomic mass of each element, and ξ_j is the proto-solar elemental abundance (Lodders,
2003). The mass yields for the SNe Ia were taken from the WDD2 model (Iwamoto
et al., 1999), and the yields for CCSNe from Nomoto et al. (2006). Each of the Ca-Rich
SNe yields from Waldman et al. (2011) were sequentially injected into the fitting routine,
along with the SNe Ia and CCSNe.

When considering the mass yields of calcium I included ^{44}Ti, as this decays into ^{44}Ca
on short cosmic timescales.

$$^{44}\text{Ti} \xrightarrow{\text{t}_{1/2} = 60 \text{ yr}} {}^{44}\text{Sc} \xrightarrow{\text{t}_{1/2} \sim 4 \text{ hr}} {}^{44}\text{Ca} \qquad (7.6)$$
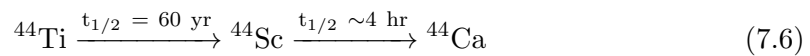
The χ^2_ν goodness-of-fit for a combination of the different SN models was analysed, along with the relative rate fractions, α_i. Assuming that *only* these three channels pollute the ICM then the results may provide insight into the potential progenitor system for Ca-Rich SNe.

The best-fitting model ($\chi^2_\nu = 0.97$) to the ICM abundances was found to involve a $0.6M_\odot$ CO WD with a $0.2M_\odot$ He shell (CO.6HE.2 model). The results can be seen in Figure 7.8 (left panel). The relative rate fractions were found to be $\alpha_{Ca} = 0.27$, $\alpha_{CC} = 0.26$ and $\alpha_{Ia} = 0.47$. Resultantly, this would mean the Ca-Rich rate would be 57% that of the SNe Ia rate. A trace of this rate is shown in Figure 7.9, compared to the SNe Ia rate evolution and the Ca-rich rate from Section 7.4.1.
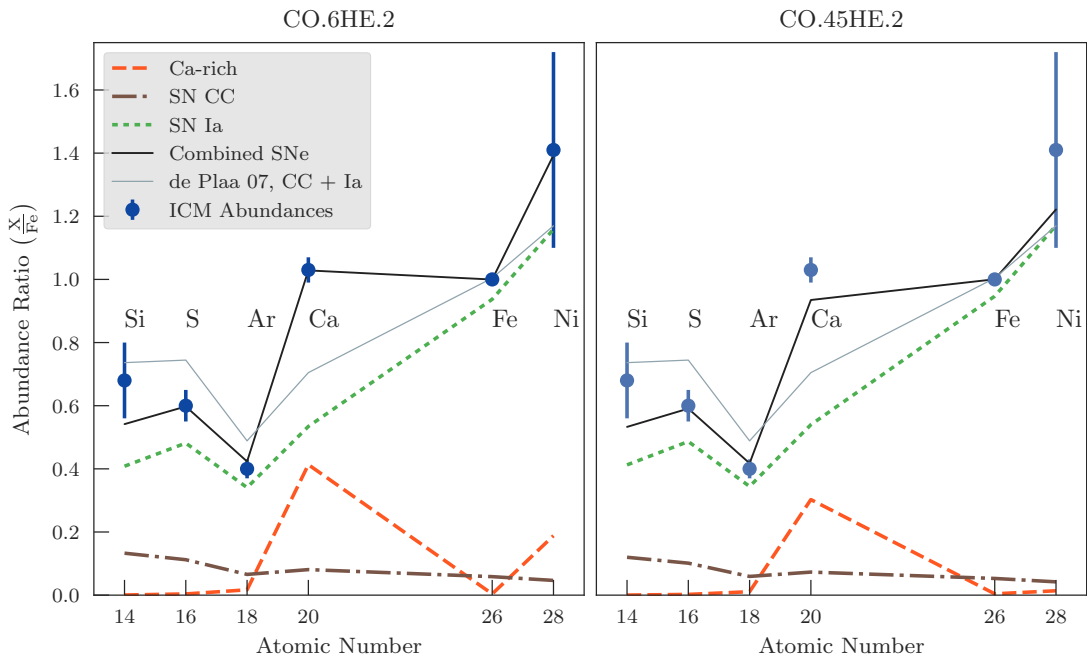
FIGURE 7.8: The intra-cluster abundances as calculated by de Plaa et al. (2007) are shown in the circle points. The thin grey line shows the abundances if only SN Ia and CCSNe are used to pollute the ICM. The various dashed lines show the contributions from different SN types. The black lines show the abundances if Ca-rich, CCSNe and SNe Ia are considered in the production of IMEs.

Another model of note is the one that best describes SN 2005E, CO.45HE.2. The results can be seen in Figure 7.8 (right). This model is perhaps the most instructive as it has been shown to reproduce the bolometric luminosity of a known Ca-rich object. However, the ICM Ca/Fe abundance is still under-estimated in this scenario by a small amount. This could possibly be explained by an additional significant SN source of Ca from other faint-and-fast transients, or uncertainties in explosion modelling. The recent discovery, lack of observations and early modelling attempts, make the latter more likely. For the CO.45HE.2 model, the relative rate fractions were, $\alpha_{Ca} = 0.16$, $\alpha_{CC} = 0.27$ and

$\alpha_{\mathrm{Ia}} = 0.57$. This means the Ca-rich rate would be 28% of the SNe Ia rate. This result is also shown in Figure 7.9. Consequently, the Ca-rich SNe rate I calculated in chapter is consistent with the rates required to solve the cluster calcium conundrum.
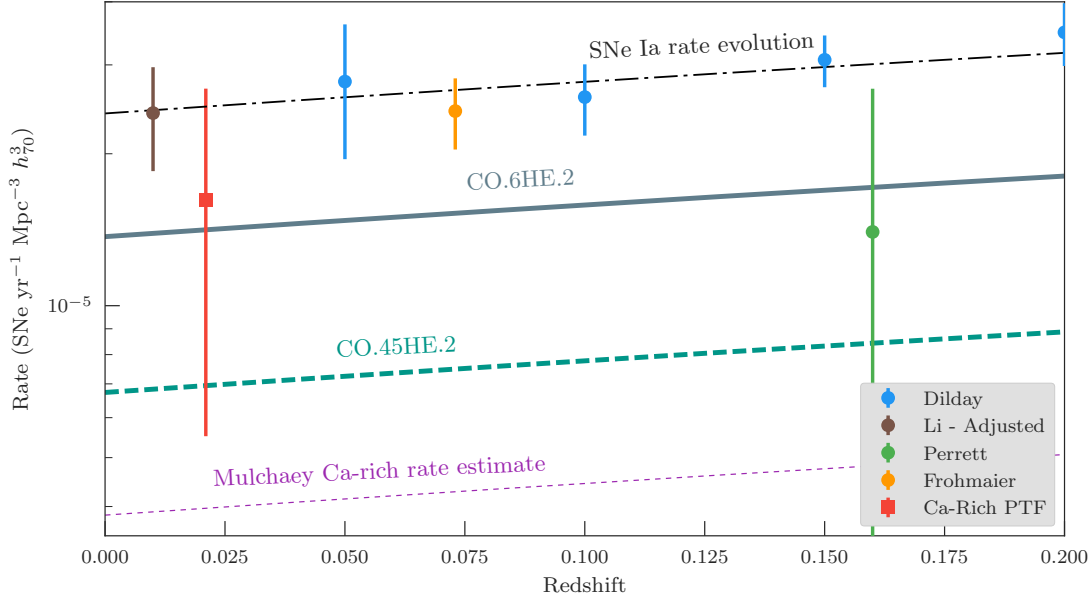
FIGURE 7.9: From the Ca abundances, the relative rates of the different models are shown in comparison to SNe Ia rates and the Ca-rich rate calculated in this chapter. The Mulchaey et al. (2014) rate estimation was calculated from ICM abundances in accordance with the method detailed in Section 7.6. However, their Ca mass yields from SN 2005E were higher than model predictions and hence their rate estimation is considerably lower.

7.7 Summary

In this chapter I have studied the most recently discovered class of thermonuclear explosion, Ca-Rich SNe. Using the PTF efficiencies framework I developed throughout this thesis, I calculated the rate of Ca-rich SNe. Furthermore, I used recent models that attempt to describe faint-and-fast SNe to reconcile the "cluster calcium conundrum".

To simulate the Ca-rich SNe, a set of template light curves were constructed. These light curves were modelled from photometry of four Ca-rich SNe discovered in PTF. The templates followed the parametrisation of Bazin et al. (2009). Only three of the objects discovered in real-time passed the light curve quality cuts. These three objects had their discovery footprints analysed to find areas and durations suitable for efficiency determination. Over 20 million template light curves were simulated in these areas and 'observed' under the frameworks explained in Chapter 4. These simulations determined the efficiency weights for the rate equation. The weights were summed and Equation 7.3

was solved. I found the rate of Ca-rich SNe, at a volume-weighted mean redshift z=0.021, to be

$$1.62(\pm 1.07) \times 10^{-5} \text{ SNe yr}^{-1} \text{ Mpc}^{-3} \ h_{70}^3.$$

Under the assumption that only Ca-rich SNe, CCSNe, and SNe Ia contribute to the intra-cluster medium elemental abundances, I determined the relative rates necessary to reproduce the observations. I used the CO WD, with an He shell, models of Waldman et al. (2011) to trial potential Ca-rich explosion scenarios. These models were mixed with SNe Ia (WDD2) and CCSNe nucleosynthesis models.

I found that a Ca-rich model capable of reproducing the ICM abundances involved a $0.6 M_\odot$ WD with a $0.2 M_\odot$ He shell. Under this prescription, the Ca-rich rate would be $\sim 57\%$ the SNe Ia rate. Furthermore, I studied the model that best described the prototypical Ca-rich SN, SN 2005E. If the entire Ca-rich population came from this model, a Ca-rich SN rate $\sim 28\%$ of the SNe Ia rate would be required. However, the observed Ca/Fe ICM abundances would still be higher than the known SNe channels can provide.

For further development to take place in this area, larger sample sizes of objects are needed. This would necessarily come from telescopes that have a high cadence and greater depth, so that these faint-and-fast transients can be discovered. Additionally, spectroscopic follow-up programs are also necessary to confirm the nature of the transient. If the rates of these objects are truly comparable to SNe Ia, then strategies need to be developed to assign follow-up priorities, especially as Ca-rich SNe evolve extremely quickly. It is only when better observations are obtained that explosion modelling of these objects can truly advance.

This concludes the science investigations of my thesis. In the next chapter I will provide perspectives on the achievements of this work and reflect on the implications for the future of SN research.

Chapter 8

Summary and Future Perspectives

8.1 Summary

In this thesis, I have performed an analysis of the entire real-time transient discovery infrastructure of the Palomar Transient Factory (PTF). This analysis quantified how efficiently PTF recovered transient events. By analysing simulations of billions of supernovae (SNe), in combination with the PTF detection efficiencies, I have demonstrated a framework for the analysis of the rates for any transient. The main focus on the thesis was on Type Ia supernovae (SNe Ia), for which I determined the local universe volumetric rate. I then extended this work to include an analysis of the rates of Ca-rich SNe. In this final chapter, I summarise my results.

8.1.1 Single Epoch Detection Efficiencies in PTF

The PTF real-time transient detection pipeline was evaluated through a method of 'fakes' injection. The fakes replicated a real transient event in the PTF imaging data. Approximately 7×10^6 fakes were placed into observational data chosen to be representative of the observing conditions experienced by PTF. The images containing the fake events were analysed through the PTF transient detection pipeline, identical to that used in the real operation of PTF.

The fraction of the fakes recovered by the detection pipeline, as a function of the fake properties and observing conditions, was defined as the recovery efficiency. I found:

- The PTF transient detection pipeline recovered $\sim 97\%$ of the bright fakes ($\mathrm{m}_R <$ 18.5 mag).

- The 50% completeness (where the pipeline recovers as many objects as it misses) was $\mathrm{m}_R = 20.3$ mag,

- The recovery efficiency as a function of observing conditions showed the expected trends that fakes were harder to recover when:

    - The seeing of the science image was poorer than the reference image.

    - The limiting magnitude was brighter.

    - The zeropoint was brighter.

    - The sky background counts were higher.

- Recovery of fakes showed a weak dependence on the moon illumination fraction and airmass.

I constructed an additional parameter that measured the immediate surface brightness in a $3'' \times 3''$ box at the location of the fake, $m_R^{F_{\mathrm{box}}}$. This measurement acted as a proxy for host galaxy brightness. Fakes in brighter environments were harder to recover.

The parameters which were most influential to the recovery fractions, m_R, m_R^{lim}, F_{sky}, Φ_{IQ}, and $m_R^{F_{\mathrm{box}}}$, were combined into a multidimensional efficiency grid allowing the recovery efficiency to be determined from a combination of these parameters.

8.1.2   Type Ia Supernovae Detection Efficiencies in PTF

The recovery efficiencies described above were linked to a transient-specific metric through simulations of model light curves. I first studied Type Ia supernovae (SNe Ia).

I created a framework whereby model light curves could be observed in a replica of PTF's operation. An artificial night sky was created, with the statistical properties of this sky derived from the recovery efficiencies. Light curves were simulated in this artificial sky, and recorded just as PTF would have observed them. The light curves were generated at random locations and with random host backgrounds. Each epoch on the simulated light curve had a model defined brightness, m_R, and a set of observing conditions taken from the observing logs. These epochs were sequentially compared to the recovery efficiency grid, returning the probability, P_{detect}, that PTF would have detected the SN on that epoch. The resultant light curve was checked to see if it passed my light curve coverage cuts. This required at least four points on the light curve separated by >12 hours, with

at least two points before maximum brightness and two points after. The details of the light curves that both passed and failed the coverage cuts were recorded.

The simulations were performed in eight different sky footprints throughout 3 years of the survey. SN Ia specific efficiency grids were created for each footprint constructed by calculating the fraction of simulated light curves that passed the coverage cuts, compared to the total number simulated. The efficiencies were parametrised by the redshift, light curve x_1 (width), M_R and the spatial location (R.A. and dec.)

The expected trends in the recovery efficiency as a function of the grid parameters were as expected: SNe Ia at a higher redshift were less likely to be recovered; Larger x_1 objects (higher stretch) were more likely to be recovered; Finally, the brightest objects were easiest to recover. The spatial dependence of the efficiency varied for each observing footprint according to the cadence strategy adopted by PTF. Resultantly, the locations observed more frequently recovered supernovae more efficiently.

The simulations are linked to real SNe by comparing the model fits to simulated properties. The result of the comparison provides the probability of the real object passing the light curve quality cuts. For example, if a real SNe Ia was found to have an efficiency of 0.2, then this single object represents a true population of five SNe where PTF missed the other four.

8.1.3   SNe Ia in PTF

I investigated populations of SNe Ia, discovered by PTF, to include in the final rates analysis. I categorised potential candidates into three groups; a spectroscopic sample, a photometric sample and a missed sample.

Objects in the spectroscopic sample were classified as SNe Ia from a PTF follow-up program. These were unambiguously identified and automatically entered into the analysis.

The photometric sample consisted of objects PTF discovered as potential transients, but didn't spectroscopically identify. In this circumstance, the SALT2 model was fit to the objects. A process of light curve quality cuts and visual inspections lead to six SNe that I considered to be convincing SNe Ia candidates.

The missed objects presented a different challenge. These objects were never identified in real-time and therefore lacked contextual assessment and spectroscopic classifications. To search for new objects, I took the entire PTF dataset of more than 1 billion candidates and filtered junk from the data. Clustering algorithms were used to search for new light curves. Known SNe, AGN and variable stars were discarded. The potential sample size

was small enough that SALT2 light curves were fit to the data, culled, and visually inspected. Ultimately, zero new SNe Ia were discovered.

8.1.4 The local universe SNe Ia rate

The rates of SNe Ia in the local universe were calculated in Chapter 6. The real SNe Ia entered the rates sample, if they occurred within a simulated sky footprint and their SALT2 model fit results produced:

- $-19.75 \leq M_R \leq -18.0$

- $-3 \leq x_1 \leq 3$

- $0.015 \leq z \leq 0.09$

The rates were analysed through a Monte-Carlo procedure by drawing realisations of all SNe Ia from within the model uncertainties. This built up a distribution of rates from which the systematic and statistical uncertainties were calculated.

The volume weighted mean redshift of the sample was z=0.073. The local universe volumetric SNe Ia rate was therefore found to be:

$$2.43 \left(^{+0.29}_{-0.29} \text{ stat} \right) \left(^{+0.33}_{-0.19} \text{ sys} \right) \times 10^{-5} \text{ SNe Ia yr}^{-1} \text{ Mpc}^{-3} \, h_{70}^3.$$

This result was combined with other SNe Ia rates from the literature. A simple delay-time distribution (DTD), of the form $\Psi_1 t^{-\beta}$, was convolved with the cosmic star formation history (SFH) and fit to the combined rates sample. Three different SFHs were used and all found results of $\beta \sim 1$. This is widely interpreted as being consistent with a double degenerate progenitor system dominating the low-redshift SNe Ia production channel.

8.1.5 Calcium-rich Supernovae

In Chapter 7, I focussed on Ca-rich SNe. These objects are faint, fast, and difficult for surveys to detect. PTF contains the largest confirmed sample from a recent sky survey: five.

A set of template light curves were created so that a population of Ca-rich SNe could be simulated in PTF. The method of deriving efficiencies was identical to the SNe

Ia analysis. More than 20 million Ca-rich light curves were simulated and population weights were calculated. The volume weighted mean redshift of the sample was z=0.021. The rate of Ca-rich SNe was calculated to be:

$$1.62(\pm1.07)\times10^{-5} \text{ SNe yr}^{-1} \text{ Mpc}^{-3} \ h_{70}^3.$$

The analysis was extended to investigate the intra-cluster medium Ca/Fe abundance ratios. This was done to reconcile the so-called "cluster calcium conundrum", where a higher Ca/Fe abundance ratio is observed than current SN classes (SNe Ia and CCSNe) can explain.

I used several Ca-rich explosion models and their nucleosynthetic yields, in combination with the SNe Ia and CCSNe models. Assuming these production channel are the only contributors to the ICM abundances, I was able to establish the relative rates of each channel.

I first investigated the the Ca-rich explosion model that could best replicate the observed ICM abundances. This was achieved by a $0.6M_\odot$ CO WD with a $0.2M_\odot$ He shell with a relative Ca-rich SN rate 57% of the SNe Ia rate.

I next investigated a model capable of reproducing the bolometric luminosity of SN 2005E. If Ca-rich SNe were typical of this model, then the relative rate of these events would need to be 28% of the SNe Ia rate. However, the Ca/Fe abundances would still be under predicted, by a small amount, in this scenario.

These predictions of the Ca-rich rates are consistent with the rate I calculated in PTF. This is an encouraging result and promotes Ca-rich SNe as a likely solution to the Ca/Fe abundance ratio.

8.2 Future Perspectives

The work presented in this thesis has been applied to two separate classes of transients: SNe Ia and Ca-rich SNe. In principle, there is no reason why other classes could not be studied in the same way. Population fractions are extremely sensitive to the underlying efficiencies of the survey. The efficiencies foundation for PTF was established in this work and, with suitable models for all transient classes, PTF could produce another significant study into population fractions.

Additionally, the methods presented in this thesis can easily be adopted by future sky surveys. The Zwicky Transient Factory is already planning to replicate my efficiencies

framework. The use of archival data, and analysing efficiencies in representative images, can be used on surveys that generate large volumes of data, e.g. Large Synoptic Survey Telescope (LSST Ivezic et al., 2009), VISTA Extragalactic Infrared Legacy Survey (VEILS Hönig et al., 2017), BlackGem(Bloemen et al., 2015) etc.

8.2.1 Superluminous Supernovae

Work has begun on superluminous supernova (SLSN) rates in PTF. The sample size of SLSNe from PTF is currently the largest of any public dataset (Perley et al., 2016). However, the numbers are still small, with only 32 discovered. The PTF SLSN-I sample contains more objects (8) than all other SLSN-I rate samples combined to-date. The lowest redshift SLSN-I rate, z=0.17, had a sample size of 1 object. The higher redshift rates of Prajs et al. (2017) (z=1.13) contained 3 objects, and Cooke et al. (2012) (z∼3) had 2. The uncertainties on all these rates are large and dominated by the statistical uncertainties. However, the rate evolution appears to tentatively follow the cosmic star formation history (SFH). Using the PTF sample, the rates can be significantly improved and used as a low redshift anchor. This work would have a significant impact in the understanding of SLSNe, as the only upcoming survey that will exceed the size of PTF's low-redshift sample is ZTF - these results are years away.

The methodology of calculating the rate is extremely similar to that presented in this thesis for the SNe Ia and Ca-rich rates. Preliminary results suggest the SLSN-I rate, at a volume weighted mean redshift of z=0.232 is $\rho_s lsn = 2.12(\pm 1.36)^{-8}$ SLSNe yr^{-1} Mpc^{-3}. Whilst the analysis currently lacks the rigour of the SNe Ia calculation, the SLSN result is unlikely to change significantly. The SLSNe rate work will be published in Frohmaier et al. (in prep).

8.2.2 Host Galaxies of Type Ia Supernova

The host galaxies are arguably more revealing to the nature of SNe Ia progenitors than volumetric rate studies. Using the efficiencies, it will be possible to calculate rates of SNe Ia as a function of the host galaxy properties, such as stellar mass and specific star formation rate. Also, investigating x_1 preference for active and passive galaxies is an illuminating insight for both the progenitors and cosmological corrections to SNe Ia. Such a study is beyond the scope of this thesis. Indeed, the content and results would be enough for an extremely interesting thesis of its own.

The large samples of low-redshift SNe Ia from PTF can be observed with high signal-to-noise spectroscopy of their host-environments. This data can help answer whether it is

the metallicity or age that drives the variation in the observed luminosities. Furthermore, delay-time distributions can be calculated using star formation histories specific to each galaxy. This probes a greater diversity in the ages of the stellar populations hosting the SNe Ia compared to the volumetric rates. Additionally, more sophisticated models of the delay-time can be adopted that are potentially capable of measuring the ratios of the 'prompt' and 'tardy' progenitor channels.

Beyond the progenitor problem, the dependence of SNe Ia luminosities on their hosts is the largest systematic uncertainty in SNe Ia cosmology. Using PTF to study the hosts of SNe Ia could likely play a critical role in understanding the environmental factors that correlate to cosmological SNe Ia properties. Ultimately, the most realistic route to further improve SNe Ia as dark energy probes is to understand how to make corrections from the host galaxy properties.

8.3 Thoughts on future transient survey strategies

The principles behind transient hunting from an optical sky survey will remain largely unchanged for the foreseeable future. High quality reference images will still be subtracted from a night's science image, but the products of the subtractions will be greatly improved with far fewer artefacts. Zackay et al. (2016) demonstrated a new image subtraction algorithm on PTF data. They claim that their algorithm produces such clean subtractions that the need for human scanners could become obsolete. If this method is shown to be as successful in large scale tests (which are currently being carried out), it may well reduce the computational overheads associated in transient searches.

The past few years have seen the rapid development of machine learning algorithms with applications to astronomical datasets. Already, surveys like Dark Energy Survey (DES) (Bernstein et al., 2012) are finding more SNe than can be spectroscopically observed. Therefore, sophisticated machine learning algorithms (e.g. Bernstein et al., 2012) have been adopted to perform photometric classifications. As we approach the LSST-era, photometric classification will become more important than ever.

Novel machine learning applications could also be applied to a study similar in ambition to this thesis. To analyse the PTF's recovery efficiency, I created grids containing the most influential metrics derived from the fakes injection process. Whilst these grids performed well for my needs, they aren't scalable to large dimensions. However, from a machine learning perspective, the sample of fakes can be considered as a training sample. All the magnitude data, image header metadata, and observing log information, could be fed into a neural-network classifier engine. When coupled with the binary information

of whether the fake was detected (with RB≥ 0.07) or not, the classifier would 'learn' and weight the inputs accordingly. From this, any combination of data (even with missing information) could be fed into the classifier and it would return the probability of detection, i.e. the recovery efficiency. The training stage would be computationally expensive, but the final result would be light-weight, robust, and easily implemented into the light curve analysis stage. Significant advances, to both the application and accessibility of these tools, have been made over the duration of this thesis (e.g. Abadi et al., 2015).

8.3.1   Versatility of the efficiencies framework

Any population study of objects collected from a sky survey must understand the biases in the dataset. This is largely achieved through an efficiency study to provide weights for observed objects. Previous studies (e.g. Dilday et al., 2008; Perrett et al., 2010) inserted realisations of light curves into observational data to determine the efficiency of recovery. Whilst this is a perfectly reasonable method of analysis, it is not scalable to larger datasets. The bottleneck to these methods comes from the I/O intensive nature of inserting fake point sources, which follow light curve evolution, across the time-series of images. Performing hundreds of thousands, possibly millions, of realisations for a single transient class in a dataset obtained from PTF would have been impossible in a reasonable time frame for this thesis.

The method adopted in this thesis of splitting efficiencies derived from artificial point sources in observational data, and efficiencies from transient light curve evolution, is a novel solution to these data intensive problems. This system is easily adopted by other transients surveys, and is currently in development for the Zwicky Transient Factory (Bellm, 2014). Of course, the methods discussed in this thesis can be applied to a sky survey designed to operate at any redshift. The fake point sources are easily tuned to represent objects within an instrument's sensitivity, and across any filter. Further studies can be performed by placing fakes below a single observation's detection limit in multiple images and stacking the result to create a deeper image. The fakes in these stacks can ultimately be used to determine the depth the survey can additionally go to probe the high redshift universe.

An intriguing extension to the F_{box} analysis, presented in Section 3.5.2.1, would be to conduct a similar experiment across multiple bands for different types of host galaxies. The single filter limitations of PTF didn't require such an investigation, but clearly for multi-band surveys the efficiencies as a function of colour for both the transient and host galaxy would be important considerations. Samples of fake point sources could be

added into galaxies with different photometric properties to build up a description of the recovery efficiency as a function of the SN host. This would be a key part of any analysis that would wish to calculate the rates of SNe Ia as a function of stellar mass, star formation rate, or other measurable galaxy metric.

The work of this thesis isn't just limited to the low-redshift, single filter, domain of PTF. The methods described throughout are applicable to any transient class, at any redshift, and hosted in any type of galaxy.

8.3.2 Gravitational Wave Electromagnetic Counterparts

It is not just population weighting that can be derived from the simulation of light curves. Optimised observing strategies can also be determined from an analysis of the number of model transients discovered as a function of the cadence. For example, simulations of the much theorised electromagnetic counterparts to gravitational waves could be simulated from model light curves (e.g. Kasen et al., 2015; Barnes & Kasen, 2013) and 'observed' by PTF. Such an experimental search has already been performed by DES to assist in looking for kilonovae in the first two seasons of their supernova search (Doctor et al., 2017). Ultimately, they found zero events, consistent with their predictions. However, PTF isn't necessarily bound by a fixed cadence experiment and can instead be used to simulate in both the 5DC and DyC fields. This opens up the possibility of answering questions on whether cadence or depth (using limiting magnitude as a proxy) is the more optimal metric to maximise, thus achieving the best chance of discovering kilonovae. Whilst I do not consider such an investigation in this work, the code and methodology described in this thesis certainly opens up this possibility for this future avenue of study.

Appendix A

Light Curve Efficiency Grids

In this appendix chapter, I show the efficiency grids for all 8 simulations boxes for the SNe Ia rate. Each simulation box is split into the the 7 different $m_R^{F_{\text{box}}}$ bins. Consequently, there are 56 efficiency grids shown.

FIGURE A.1: 2010 A efficiency grids

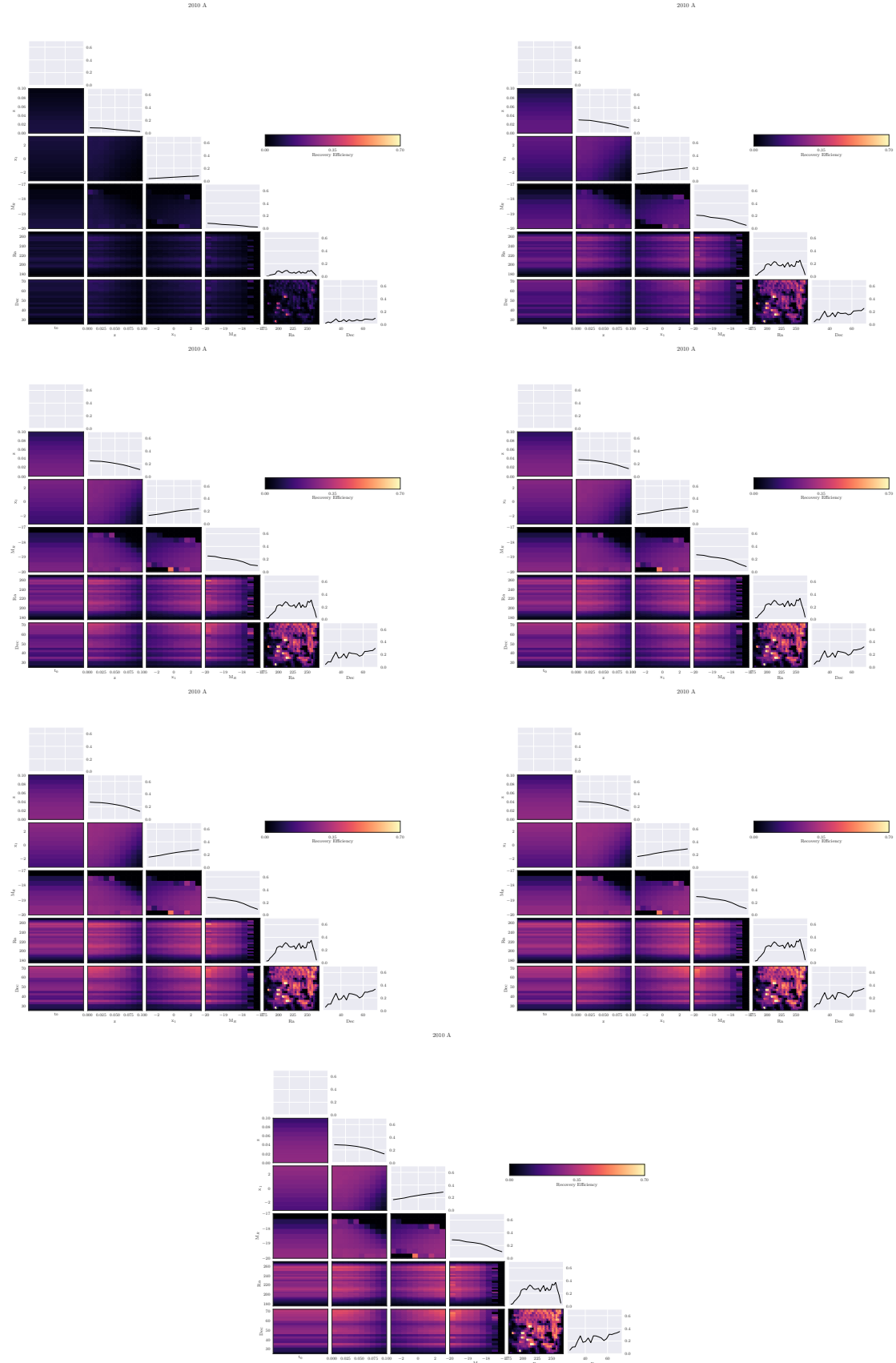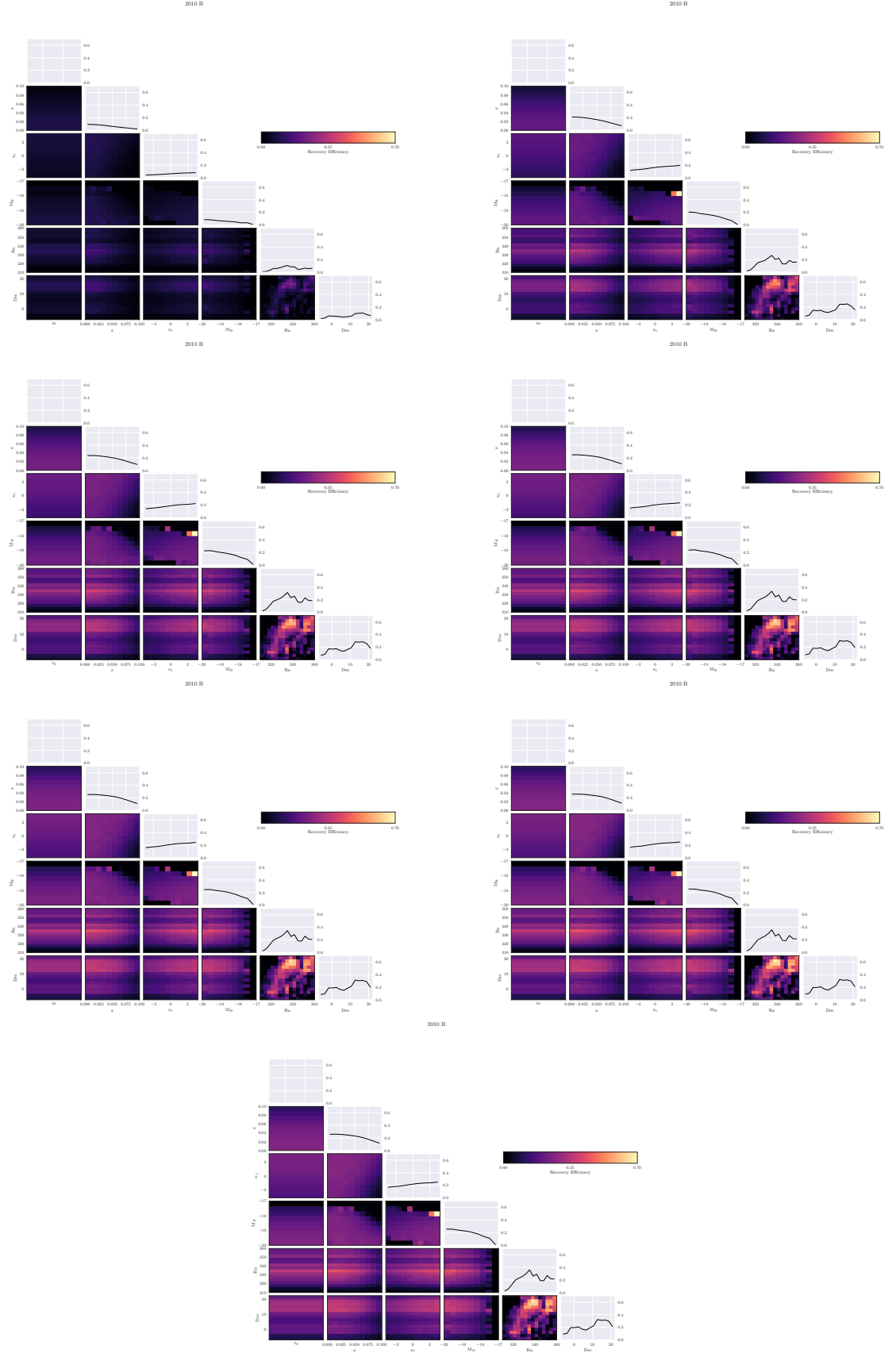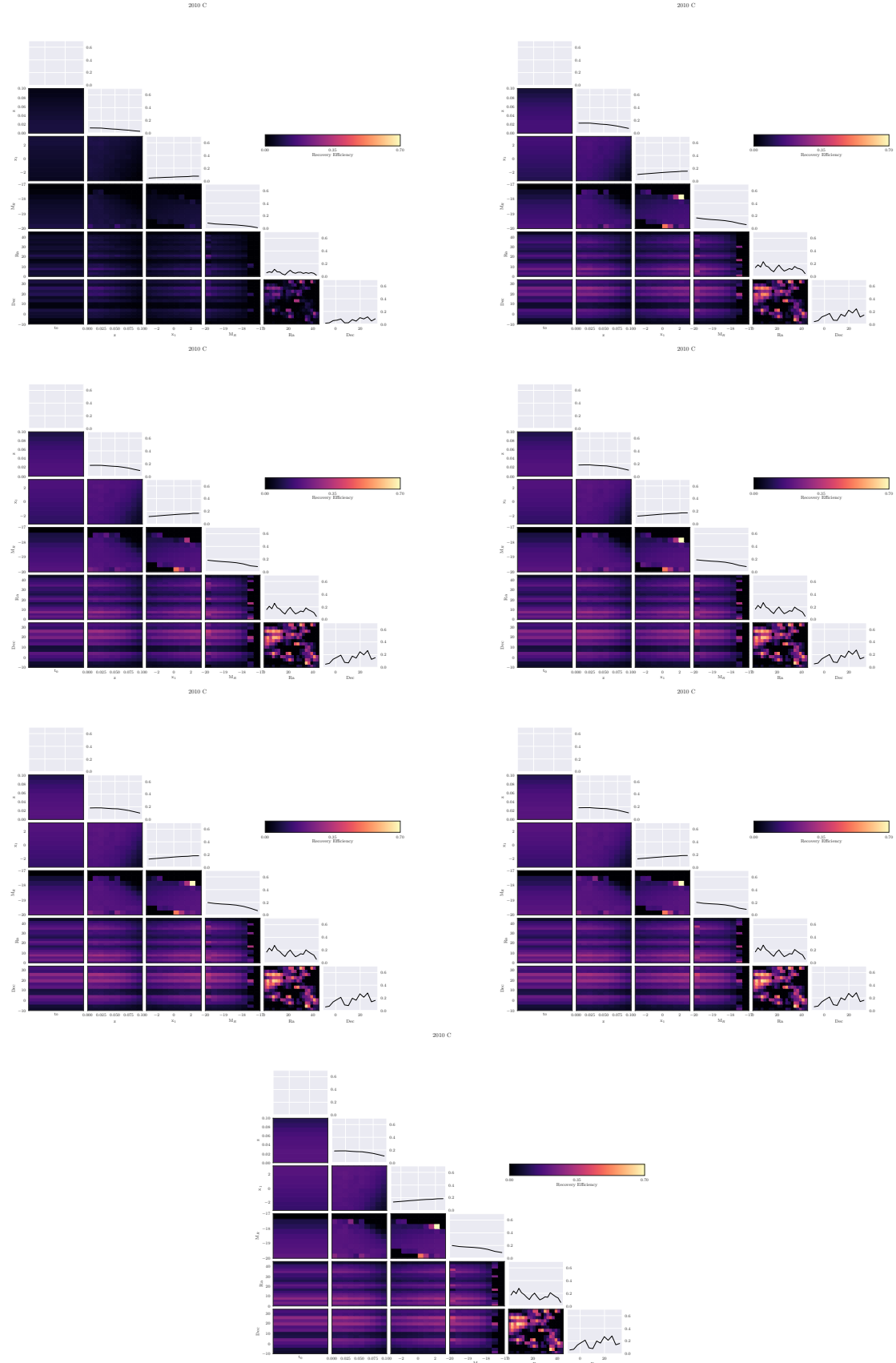FIGURE A.3: 2010 C efficiency grids
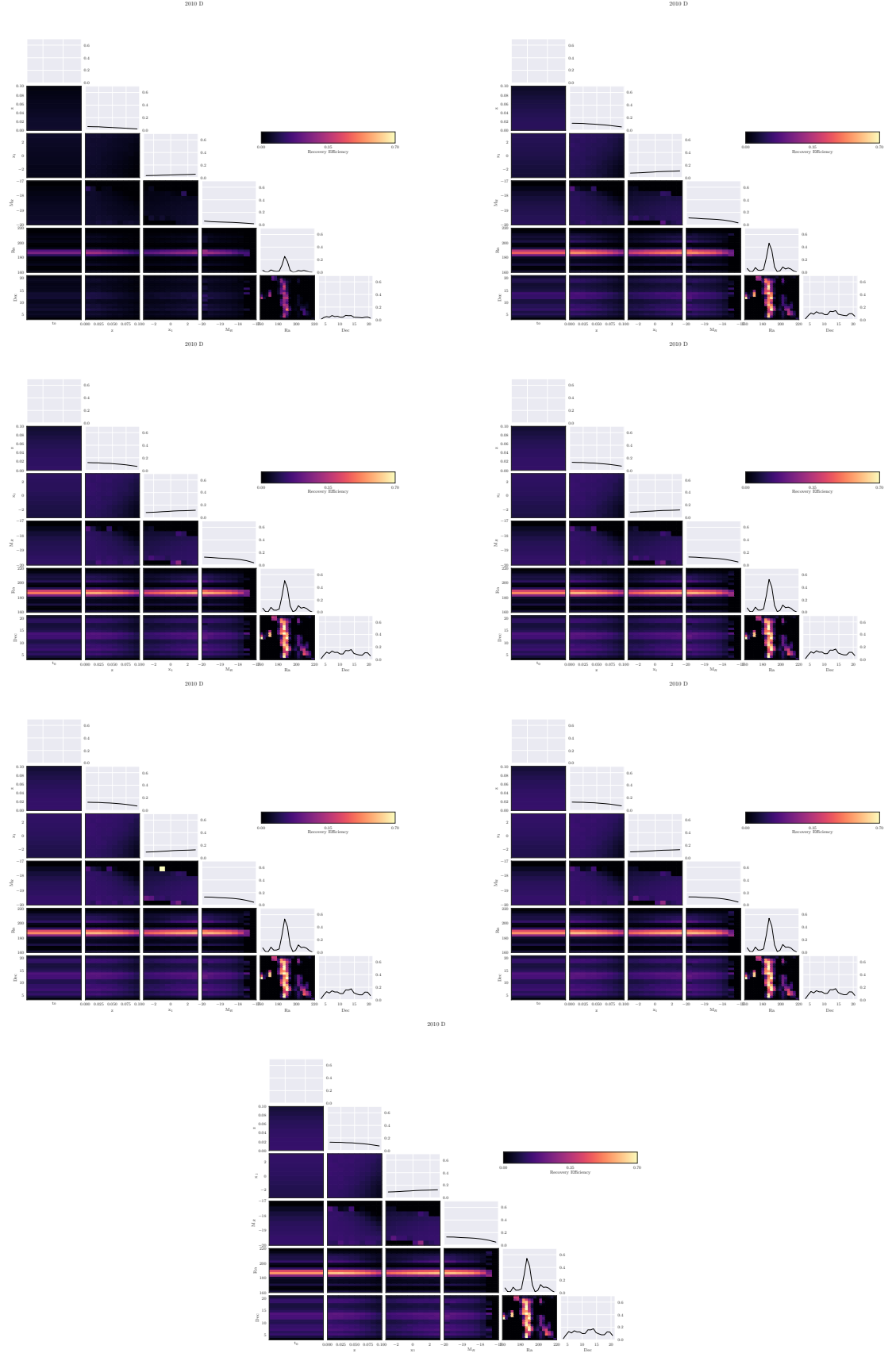
FIGURE A.4: 2010 D efficiency grids
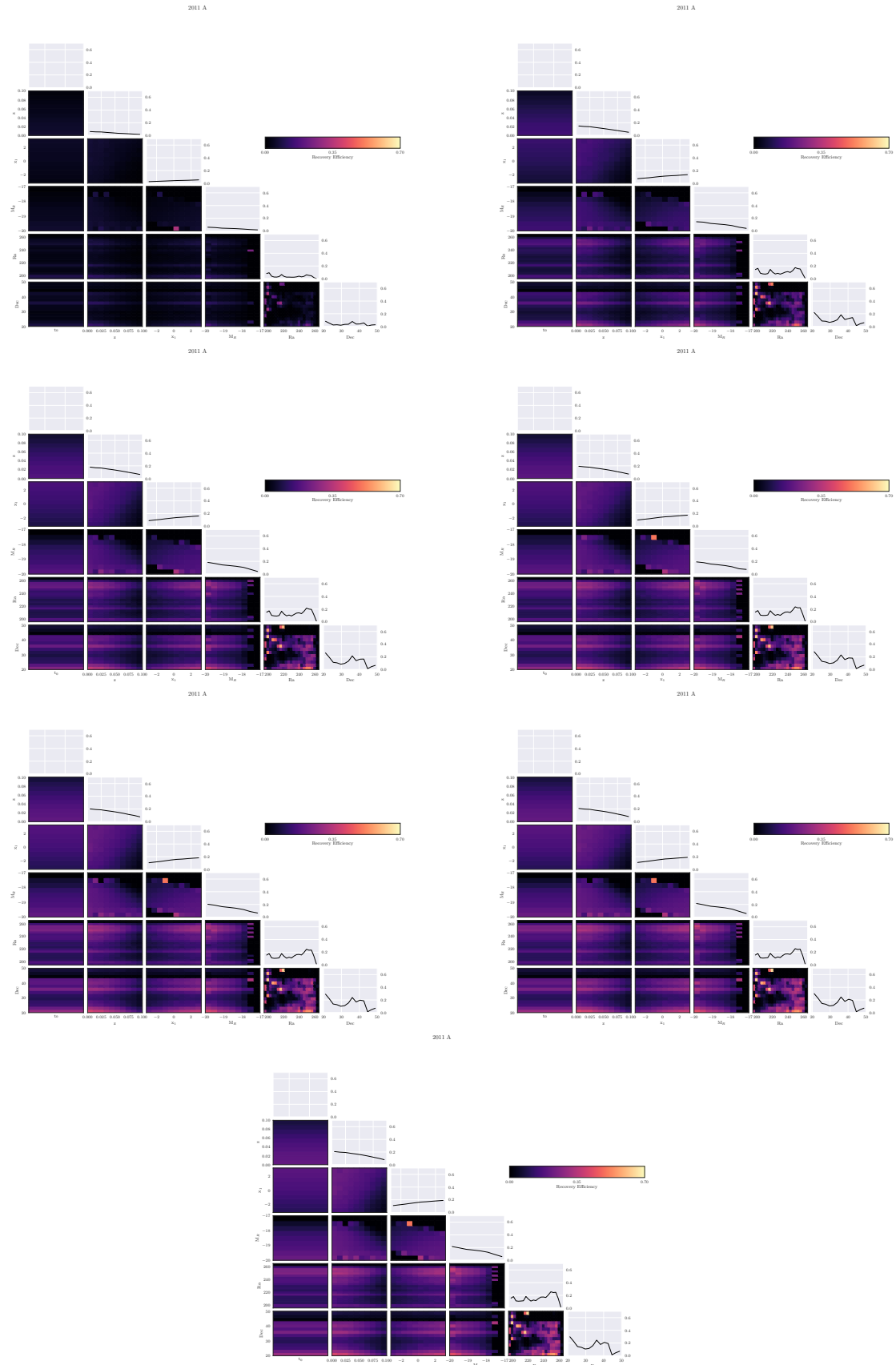
FIGURE A.5: 2011 A efficiency grids
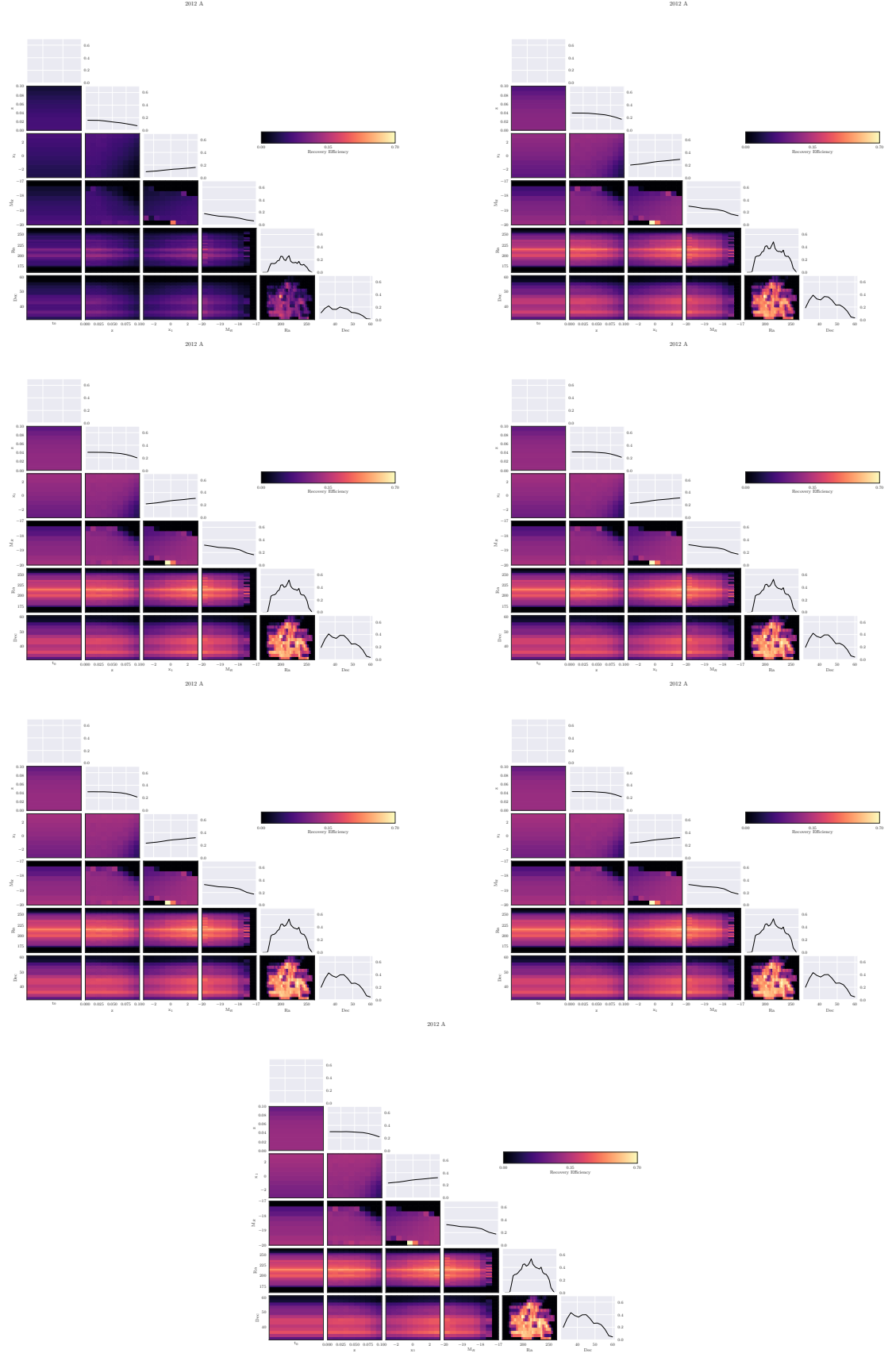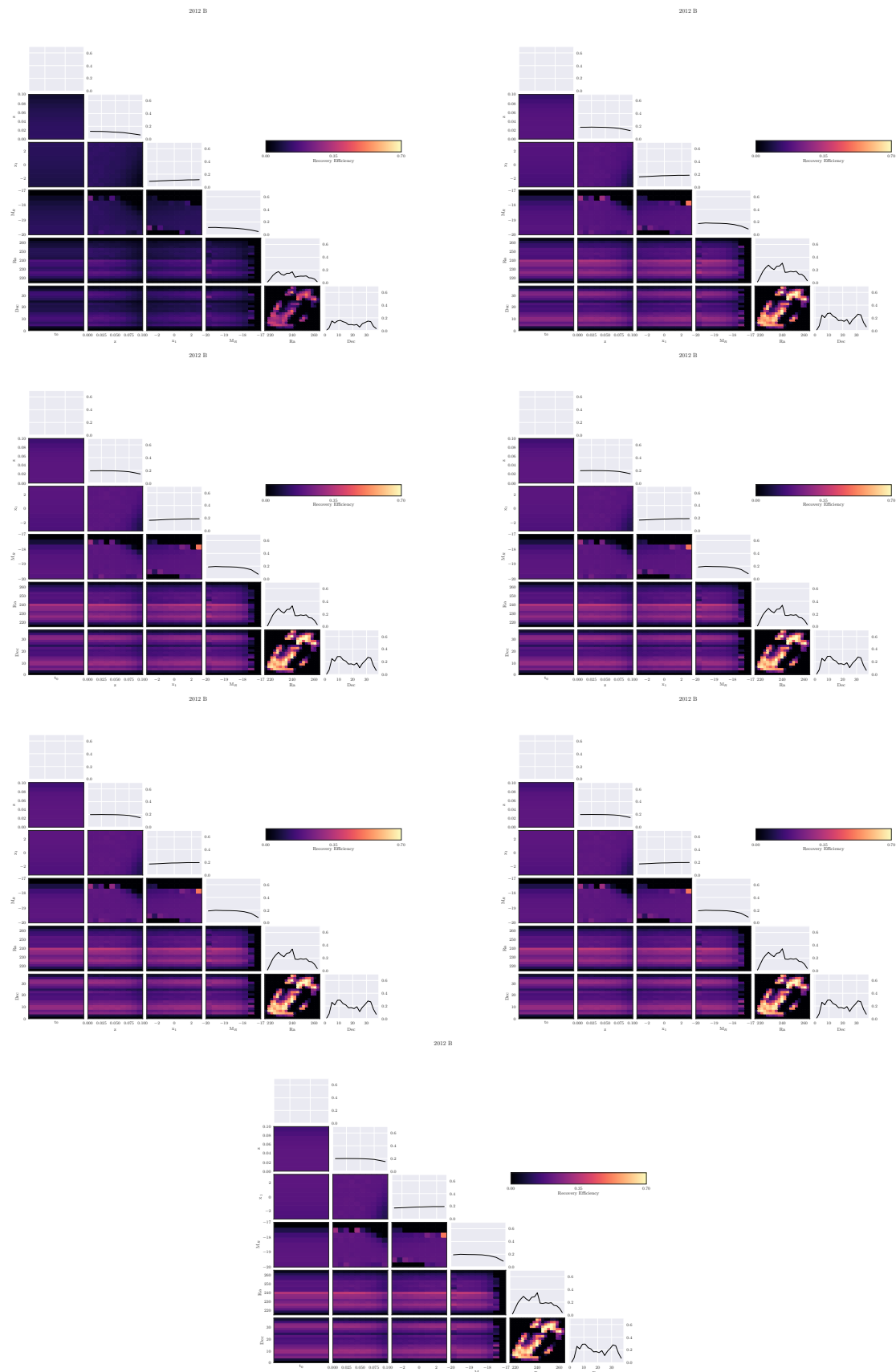
FIGURE A.7: 2012 B efficiency grids
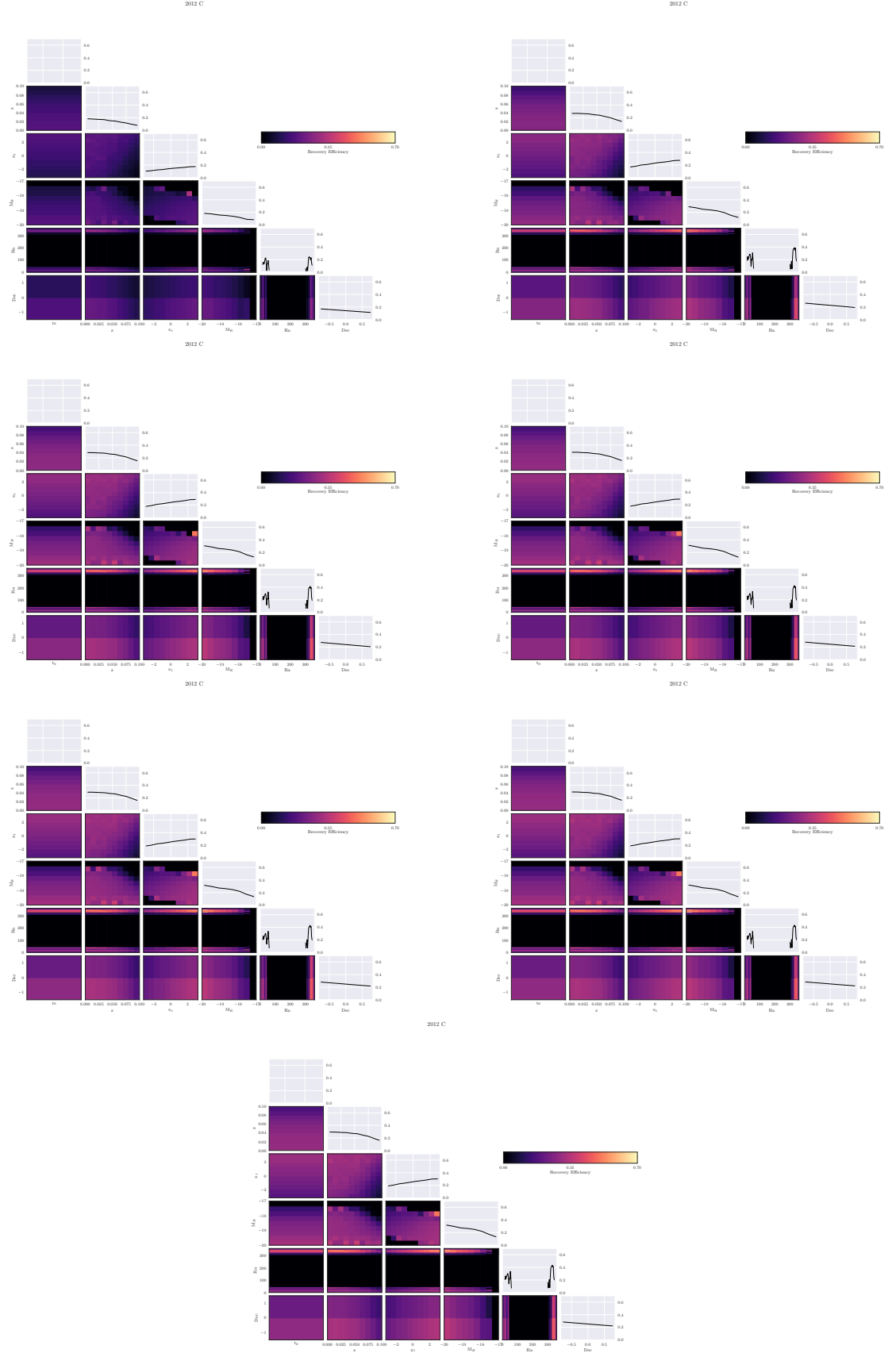
FIGURE A.8: 2012 B efficiency grids

Appendix B

Light Curve Fit Parameters

The table in this appendix chapter lists all the SNe Ia that entered into the rates analysis. The SALT2 model fits are listed, along with their 1σ uncertainties.

PTF Name	R.A.	Dec.	z	σ_z	t_0	σt_0	x_0	σx_0	x_1	σx_1	M_R	Spec
10bjs	195.29679	53.81592	0.030	0	2455262.17	0.92	0.0075	0.0010	0.84	0.25	-19.08	T
10cko	182.81405	13.73375	0.070	0	2455266.93	0.16	0.0008	0.0002	0.61	0.17	-18.83	T
10cwm	188.23087	4.48647	0.080	0	2455271.97	0.49	0.0005	0.0001	0.61	0.33	-18.63	T
10cxk	194.85027	27.99625	0.016	0	2455277.17	0.07	0.0072	0.0013	-1.56	0.03	-17.88	T
10duy	209.29612	40.16330	0.079	0	2455283.93	0.12	0.0008	0.0002	1.13	0.12	-19.05	T
10duz	192.91459	14.43854	0.060	0	2455285.78	0.11	0.0011	0.0003	-0.26	0.17	-18.88	T
10fkk	249.90372	34.59231	0.080	0	2455308.21	0.13	0.0008	0.0002	1.22	0.16	-19.07	T
10fps	202.35442	11.79625	0.022	0	2455313.31	0.05	0.0051	0.0009	-2.11	0.04	-18.20	T
10fxl	253.19814	51.06246	0.030	0	2455318.44	0.05	0.0029	0.0006	-0.12	0.04	-18.63	T
10gjx	186.22570	20.08451	0.076	0	2455326.49	0.08	0.0009	0.0001	0.51	0.09	-18.89	T
10glo	188.27398	6.55629	0.075	0	2455322.89	0.14	0.0006	0.0001	1.14	0.09	-18.96	T
10gmg	246.24373	51.03931	0.063	0	2455324.98	0.09	0.0012	0.0003	0.95	0.09	-19.05	T
10gmj	250.78787	66.34763	0.082	0	2455315.58	0.33	0.0005	0.0001	-0.29	0.25	-18.74	T
10gnj	183.67408	10.95921	0.080	0	2455318.24	0.63	0.0003	0.0001	1.26	0.27	-18.05	T
10gqi	189.94853	59.94478	0.046	0	2455318.54	0.65	0.0004	0.0001	-1.51	0.42	-17.08	T
10gxa	248.61793	57.60668	0.064	0.027	2455321.18	0.38	0.0005	0.0001	1.10	0.20	-18.27	F
10hdn	223.10265	47.47645	0.070	0	2455343.39	0.12	0.0008	0.0002	0.45	0.14	-18.86	T
10hld	234.86376	50.09720	0.038	0	2455338.82	0.25	0.0008	0.0002	0.49	0.22	-17.53	T
10hmc	248.19088	35.06878	0.073	0	2455342.35	0.12	0.0010	0.0002	0.14	0.16	-19.28	T
10hml	199.95703	41.98378	0.054	0	2455353.00	0.03	0.0017	0.0003	0.54	0.04	-18.99	T
10hmq	201.50084	7.25023	0.086	0	2455325.25	0.64	0.0004	0.0001	-3.79	0.98	-18.81	T
10hmv	182.88745	47.27494	0.032	0	2455352.09	0.05	0.0035	0.0007	1.17	0.04	-18.41	T

PTF Name	R.A.	Dec.	z	σ_z	t_0	σt_0	x_0	σx_0	x_1	σx_1	M_R	Spec
10hoz	251.21953	52.61716	0.060	0	2455348.30	0.10	0.0010	0.0002	0.18	0.11	-18.69	T
10hrw	266.06446	52.14941	0.049	0	2455349.75	0.07	0.0014	0.0003	-0.42	0.08	-18.79	T
10idi	230.43049	69.44693	0.049	0.033	2455354.16	0.38	0.0003	0.0001	-1.62	0.36	-17.01	F
10ifj	214.84139	34.35397	0.076	0	2455354.66	0.18	0.0008	0.0002	1.21	0.15	-19.03	T
10inf	250.81277	32.67438	0.050	0	2455361.09	0.06	0.0016	0.0003	0.13	0.07	-18.78	T
10iyc	257.34094	44.39331	0.055	0	2455362.09	0.12	0.0016	0.0002	0.66	0.19	-18.88	T
10jdw	235.50131	47.59388	0.077	0	2455366.91	0.13	0.0006	0.0001	-0.56	0.18	-18.84	T
10jtp	257.74362	39.47452	0.067	0	2455365.03	0.22	0.0006	0.0001	-0.50	0.44	-18.58	T
10jwx	205.66374	56.46232	0.068	0	2455368.62	0.13	0.0010	0.0002	-0.85	0.19	-18.97	T
10kdg	199.30709	44.14429	0.062	0	2455370.68	0.06	0.0006	0.0001	-0.45	0.12	-18.42	T
10kiw	221.57535	47.20728	0.069	0	2455369.44	0.14	0.0004	0.0001	0.69	0.23	-18.17	T
10ksd	193.70280	29.22531	0.081	0.031	2455370.92	0.27	0.0002	0.0000	0.45	0.54	-17.78	F
10kvl	249.75849	32.49420	0.055	0.025	2455372.11	0.27	0.0004	0.0001	1.19	0.45	-17.68	F
10lot	234.66556	41.00552	0.025	0	2455381.93	0.11	0.0084	0.0018	0.44	0.08	-18.97	T
10lrp	222.84769	35.76095	0.079	0	2455375.47	0.17	0.0008	0.0002	-0.18	0.23	-19.13	T
10lxp	215.98668	55.72909	0.088	0	2455381.96	0.06	0.0007	0.0001	0.20	0.07	-19.24	T
10lya	355.38194	14.12365	0.064	0	2455380.76	0.19	0.0009	0.0002	0.09	0.25	-18.82	T
10mbk	214.26996	71.78987	0.065	0	2455381.70	0.30	0.0011	0.0003	0.59	0.39	-19.09	T
10mwb	259.45822	40.88113	0.030	0	2455391.34	0.08	0.0029	0.0007	-0.58	0.05	-18.61	T
10nct	220.29424	58.40726	0.052	0	2455391.22	0.43	0.0004	0.0001	0.36	0.63	-17.38	T
10ncu	267.25467	68.10514	0.070	0	2455381.00	0.33	0.0004	0.0001	-2.31	0.37	-18.14	T
10ndc	259.95907	28.69930	0.082	0	2455391.00	0.22	0.0007	0.0002	0.78	0.21	-19.08	T

PTF Name	R.A.	Dec.	z	σ_z	t_0	σt_0	x_0	σx_0	x_1	σx_1	M_R	Spec
10nlg	252.64366	60.27638	0.050	0	2455392.61	0.27	0.0008	0.0002	-0.38	0.44	-18.15	T
10nvh	323.00977	8.99325	0.068	0	2455394.24	8.19	0.0008	0.0003	0.39	0.33	-18.82	T
10otc	254.81823	68.18205	0.054	0	2455390.37	0.18	0.0009	0.0002	-2.95	0.30	-18.51	T
10pvi	330.50967	14.53622	0.080	0	2455410.83	0.15	0.0007	0.0001	0.23	0.13	-18.94	T
10qsc	323.58838	-5.06233	0.088	0	2455422.87	0.26	0.0006	0.0001	1.01	0.26	-19.03	T
10qwg	40.54116	2.44791	0.071	0	2455425.38	0.23	0.0008	0.0002	-0.41	0.23	-18.88	T
10qyx	36.80024	-4.53466	0.063	0	2455426.98	0.12	0.0008	0.0002	-1.39	0.15	-18.69	T
10rab	26.78089	-0.04985	0.085	0	2455426.41	0.35	0.0005	0.0001	0.14	0.36	-18.83	T
10rbp	19.15859	-1.82322	0.079	0	2455431.38	0.21	0.0006	0.0001	0.64	0.30	-18.85	T
10rhi	357.43239	13.04258	0.085	0	2455425.57	0.24	0.0006	0.0001	0.03	0.22	-18.93	T
10sur	38.64032	11.31923	0.055	0.041	2455435.67	0.51	0.0003	0.0001	0.55	0.67	-17.45	F
10tce	349.79316	9.19840	0.041	0	2455442.89	0.07	0.0025	0.0006	0.83	0.06	-18.85	T
10tqy	10.67300	24.75401	0.045	0	2455444.61	0.13	0.0019	0.0004	-1.76	0.11	-18.74	T
10trp	322.03327	9.85366	0.049	0	2455450.76	0.32	0.0009	0.0002	1.55	0.41	-18.16	T
10trs	3.83600	17.53316	0.073	0	2455442.31	0.17	0.0007	0.0002	-1.59	0.14	-18.87	T
10ubm	0.49701	21.82495	0.070	0	2455455.70	0.12	0.0011	0.0002	0.77	0.15	-19.23	T
10ucl	331.52023	15.50958	0.080	0	2455444.95	0.12	0.0006	0.0001	-0.72	0.15	-19.00	T
10ufj	36.41304	24.76477	0.073	0	2455457.36	0.10	0.0006	0.0001	-0.04	0.12	-18.65	T
10vdu	2.98668	25.65541	0.069	0.040	2455459.22	0.66	0.0002	0.0000	-0.87	1.19	-17.48	F
10vfo	18.77781	24.60494	0.089	0	2455459.24	0.38	0.0005	0.0001	-0.32	0.30	-18.86	T
10viq	335.08319	17.05617	0.031	0	2455452.94	0.03	0.0026	0.0001	0.69	0.04	-18.81	T
10wnm	5.51504	27.04063	0.066	0	2455477.67	0.15	0.0010	0.0002	0.43	0.23	-18.94	T

PTF Name	R.A.	Dec.	z	σ_z	t_0	σt_0	x_0	σx_0	x_1	σx_1	M_R	Spec
10wnq	12.29093	32.13874	0.070	0	2455473.45	0.23	0.0009	0.0002	-0.46	0.29	-18.97	T
10wof	353.17435	15.35881	0.053	0	2455475.08	0.13	0.0013	0.0003	0.05	0.21	-18.68	T
10wyq	19.60370	19.44617	0.080	0	2455480.18	0.16	0.0005	0.0001	0.47	0.25	-18.66	T
10xyl	5.10359	5.70160	0.056	0	2455477.95	0.19	0.0006	0.0001	-2.66	0.56	-18.21	T
11bas	199.19973	43.52035	0.085	0	2455641.06	0.29	0.0005	0.0001	0.24	0.23	-18.75	T
11biv	257.14127	22.07110	0.048	0	2455639.64	0.32	0.0009	0.0002	0.88	0.22	-18.07	T
11bkf	215.38064	23.94813	0.050	0	2455637.14	0.72	0.0012	0.0003	-3.22	0.37	-18.67	T
11bnx	247.58051	21.08535	0.060	0	2455653.87	0.19	0.0008	0.0002	-0.07	0.15	-18.65	T
11bof	238.26971	20.63598	0.026	0	2455655.70	0.10	0.0028	0.0002	0.57	0.08	-17.63	T
11bui	198.23498	47.45346	0.028	0	2455675.79	0.06	0.0038	0.0007	0.81	0.05	-18.83	T
11cao	244.69984	25.18791	0.040	0	2455671.11	0.08	0.0021	0.0005	0.28	0.09	-18.60	T
11cxe	250.81595	40.53593	0.066	0.036	2455693.40	0.24	0.0004	0.0001	-0.79	0.32	-18.06	F
11dif	213.96698	35.73999	0.059	0	2455703.67	0.37	0.0015	0.0002	0.62	0.14	-18.93	T
12bok	183.40480	46.48365	0.025	0	2456015.63	0.02	0.0041	0.0004	-0.48	0.02	-18.81	T
12cci	213.87211	39.90248	0.067	0	2456024.95	0.28	0.0012	0.0002	0.86	0.20	-19.06	T
12ccz	240.92515	36.99562	0.041	0	2456018.76	0.14	0.0015	0.0001	-2.94	0.10	-18.14	T
12cjg	205.32609	55.45218	0.067	0	2456025.53	0.16	0.0008	0.0002	-0.39	0.38	-18.80	T
12cks	215.68407	34.25452	0.063	0	2456028.94	0.12	0.0012	0.0003	0.99	0.10	-19.01	T
12cnl	197.78042	39.08212	0.047	0	2456034.64	0.07	0.0023	0.0003	1.25	0.05	-18.81	T
12csi	251.80760	33.30570	0.053	0	2456031.48	0.17	0.0007	0.0002	0.76	0.16	-18.03	T
12dah	230.45386	50.07263	0.084	0.018	2456043.49	0.21	0.0005	0.0001	1.81	0.25	-18.74	F
12dao	225.87210	44.67492	0.044	0.022	2456045.18	0.18	0.0005	0.0001	0.31	0.21	-17.38	F

PTF Name	R.A.	Dec.	z	σ_z	t_0	σt_0	x_0	σx_0	x_1	σx_1	M_R	Spec
12dco	230.13223	59.19573	0.075	0	2456040.50	0.64	0.0007	0.0002	0.42	0.76	-18.84	T
12dhb	244.23929	49.69736	0.057	0	2456042.06	0.10	0.0013	0.0003	0.36	0.12	-18.88	T
12dhk	198.79614	53.28167	0.070	0	2456047.69	0.24	0.0005	0.0001	3.09	0.47	-18.28	T
12dhl	200.73181	52.23347	0.057	0	2456041.42	0.14	0.0010	0.0002	-2.79	0.25	-18.78	T
12dre	215.55961	44.96836	0.051	0.047	2456054.20	0.40	0.0005	0.0001	1.76	1.00	-17.80	F
12dwm	199.57402	35.70713	0.053	0	2456053.52	0.23	0.0005	0.0001	-2.35	0.51	-17.85	T
12dxm	208.36032	43.91350	0.063	0	2456054.22	0.15	0.0007	0.0002	-1.67	0.16	-18.53	T
12eac	253.34410	36.27314	0.088	0	2456056.25	0.30	0.0003	0.0001	1.85	0.39	-18.39	T
12ecm	239.08879	36.53716	0.066	0	2456068.35	0.24	0.0009	0.0002	0.53	0.43	-18.83	T
12ecr	218.94382	45.18880	0.069	0	2456067.84	0.19	0.0007	0.0002	0.64	0.37	-18.72	T
12fuu	226.16831	6.07250	0.035	0	2456112.27	0.11	0.0017	0.0004	-0.01	0.13	-18.18	T
12gaz	234.41624	6.61612	0.071	0	2456113.69	0.16	0.0008	0.0002	-0.45	0.23	-18.84	T
12gcn	230.93229	8.43725	0.075	0	2456111.91	0.28	0.0005	0.0001	-0.23	0.42	-18.49	T
12gdq	227.89714	9.70946	0.033	0	2456116.99	0.11	0.0026	0.0007	-0.71	0.19	-18.65	T
12gfj	242.45735	29.72575	0.050	0	2456114.78	0.44	0.0004	0.0001	-0.27	0.46	-17.36	T
12ggb	234.60684	31.53579	0.060	0	2456117.02	0.09	0.0009	0.0002	0.58	0.14	-18.86	T
12gkn	231.41596	9.65157	0.077	0	2456120.42	0.14	0.0007	0.0001	1.03	0.23	-18.82	T
12gmv	41.25598	-0.73013	0.054	0	2456115.02	0.35	0.0010	0.0002	-0.04	0.14	-18.65	T
12hmx	341.78755	0.17444	0.085	0	2456150.80	0.21	0.0004	0.0001	-0.76	0.19	-18.53	T
12hwb	338.38972	-1.16259	0.056	0	2456171.35	4.07	0.0005	0.0005	42.94	90.50	-19.10	T
12iiq	42.53233	-0.26512	0.029	0	2456182.74	0.04	0.0031	0.0008	-1.29	0.04	-18.63	T
12ikt	18.67971	0.28531	0.044	0	2456187.79	0.05	0.0020	0.0005	-0.22	0.08	-18.85	T

Bibliography

Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, , , software available from tensorflow.org

Abt, H. A. 1983, ARA&A, 21, 343

Ackermann, M., Ajello, M., Atwood, W. B., et al. 2015, ApJ, 810, 14

Arcavi, I., Gal-Yam, A., Sullivan, M., et al. 2014, ApJ, 793, 38

Arnett, W. D. 1979, ApJ, 230, L37

—. 1982, ApJ, 253, 785

Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31

Baade, W., & Zwicky, F. 1934, Proceedings of the National Academy of Science, 20, 254

Baltay, C., Rabinowitz, D., Hadjiyska, E., et al. 2013, PASP, 125, 683

Barbary, K. 2014, sncosmo, v.0.4.2, Zenodo, doi:10.5281/zenodo.11938

Barbary, K., Aldering, G., Amanullah, R., et al. 2012, ApJ, 745, 31

Barbon, R., Cappellaro, E., & Turatto, M. 1989, A&AS, 81, 421

Barbon, R., Ciatti, F., & Rosino, L. 1979, A&A, 72, 287

Barnes, J., & Kasen, D. 2013, The Astrophysical Journal, 775, 18

Batygin, K., & Brown, M. E. 2016, AJ, 151, 22

Bazin, G., Palanque-Delabrouille, N., Rich, J., et al. 2009, A&A, 499, 653

Begelman, M. C., & Sarazin, C. L. 1986, ApJ, 302, L59

Bellm, E. 2014, in The Third Hot-wiring the Transient Universe Workshop, ed. P. R. Wozniak, M. J. Graham, A. A. Mahabal, & R. Seaman, 27–33

Benetti, S., Cappellaro, E., Mazzali, P. A., et al. 2005, ApJ, 623, 1011

Bernstein, J. P., Kessler, R., Kuhlmann, S., et al. 2012, ApJ, 753, 152

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22

Bianchi, S., Guainazzi, M., Matt, G., Fonseca Bonilla, N., & Ponti, G. 2009, A&A, 495, 421

Bildsten, L., Shen, K. J., Weinberg, N. N., & Nelemans, G. 2007, ApJ, 662, L95

Bloemen, S., Groot, P., Nelemans, G., & Klein-Wolt, M. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 496, Living Together: Planets, Host Stars and Binaries, ed. S. M. Rucinski, G. Torres, & M. Zejda, 254

Bloom, J. S., Richards, J. W., Nugent, P. E., et al. 2012, PASP, 124, 1175

Branch, D., & van den Bergh, S. 1993, AJ, 105, 2231

Brink, H., Richards, J. W., Poznanski, D., et al. 2013, MNRAS, 435, 1047

Cao, Y., Nugent, P. E., & Kasliwal, M. M. 2016, PASP, 128, 114502

Cappellaro, E., Evans, R., & Turatto, M. 1999, A&A, 351, 459

Cappellaro, E., Turatto, M., Benetti, S., et al. 1993, A&A, 268, 472

Cappellaro, E., Turatto, M., Tsvetkov, D. Y., et al. 1997, A&A, 322, 431

Cenko, S. B. 2017, Nature Astronomy, 1, 0008

Cenko, S. B., Fox, D. B., Moon, D.-S., et al. 2006, PASP, 118, 1396

Charisi, M., Bartos, I., Haiman, Z., et al. 2016, MNRAS, 463, 2145

Childress, M., Aldering, G., Antilogus, P., et al. 2013, ApJ, 770, 107

Cole, S., Norberg, P., Baugh, C. M., et al. 2001, MNRAS, 326, 255

Colgate, S. A., & McKee, C. 1969, ApJ, 157, 623

Conley, A., Sullivan, M., Hsiao, E. Y., et al. 2008, ApJ, 681, 482

Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1

Cooke, J., Sullivan, M., Gal-Yam, A., et al. 2012, Nature, 491, 228

Crocker, R. M., Ruiter, A. J., Seitenzahl, I. R., et al. 2017, Nature Astronomy, 1, 0135

Cuillandre, J.-C., Luppino, G. A., Starr, B. M., & Isani, S. 2000, in Proc. SPIE, Vol. 4008, Optical and IR Telescope Instrumentation and Detectors, ed. M. Iye & A. F. Moorwood, 1010–1021

Dahlen, T., Strolger, L., & Riess, A. G. 2007, in Bulletin of the American Astronomical Society, Vol. 39, American Astronomical Society Meeting Abstracts, 930

Dahlen, T., Strolger, L.-G., & Riess, A. G. 2008, ApJ, 681, 462

Dahlen, T., Strolger, L.-G., Riess, A. G., et al. 2004, ApJ, 613, 189

D'Andrea, C. B., Gupta, R. R., Sako, M., et al. 2011, ApJ, 743, 172

de Plaa, J., Werner, N., Bleeker, J. A. M., et al. 2007, A&A, 465, 345

Dilday, B., Kessler, R., Frieman, J. A., et al. 2008, ApJ, 682, 262

Dilday, B., Smith, M., Bassett, B., et al. 2010, ApJ, 713, 1026

Dimitriadis, G., Sullivan, M., Kerzendorf, W., et al. 2017, MNRAS, 468, 3798

Doctor, Z., Kessler, R., Chen, H. Y., et al. 2017, ApJ, 837, 57

Drake, A. J., Djorgovski, S. G., Mahabal, A., et al. 2009, ApJ, 696, 870

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, AJ, 142, 72

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in (AAAI Press), 226–231

Evans, R., van den Bergh, S., & McClure, R. D. 1989, ApJ, 345, 752

Faber, S. M., Phillips, A. C., Kibrick, R. I., et al. 2003, in Proc. SPIE, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 1657–1669

Faran, T., Poznanski, D., Filippenko, A. V., et al. 2014, MNRAS, 442, 844

Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601

Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2013, ArXiv e-prints, arXiv:1306.2144

Filippenko, A. V. 1989, PASP, 101, 588

—. 1997, ARA&A, 35, 309

Filippenko, A. V., Chornock, R., Swift, B., et al. 2003, IAU Circ., 8159

Firth, R. E., Sullivan, M., Gal-Yam, A., et al. 2015, MNRAS, 446, 3895

Foley, R. J. 2015, MNRAS, 452, 2463

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Förster, F., Wolf, C., Podsiadlowski, P., & Han, Z. 2006, MNRAS, 368, 1893

Frieman, J. A., Bassett, B., Becker, A., et al. 2008, AJ, 135, 338

Frohmaier, C., Sullivan, M., Nugent, P. E., Goldstein, D. A., & DeRose, J. 2017, ApJS, 230, 4

Gal-Yam, A. 2012, Science, 337, 927

Gal-Yam, A., Kasliwal, M. M., Arcavi, I., et al. 2011, ApJ, 736, 159

García-Berro, E., Badenes, C., Aznar-Siguán, G., & Lorén-Aguilar, P. 2017, MNRAS, 468, 4815

Graur, O., & Maoz, D. 2013, MNRAS, 430, 1746

Graur, O., Poznanski, D., Maoz, D., et al. 2011, MNRAS, 417, 916

Graur, O., Rodney, S. A., Maoz, D., et al. 2014, ApJ, 783, 28

Greggio, L. 2005, A&A, 441, 1055

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Gupta, R. R., Kuhlmann, S., Kovacs, E., et al. 2016, ArXiv e-prints, arXiv:1604.06138

Guy, J., Astier, P., Nobili, S., Regnault, N., & Pain, R. 2005, A&A, 443, 781

Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11

Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7

Hachisu, I., Kato, M., & Nomoto, K. 1996, ApJ, 470, L97

—. 2008, ApJ, 683, L127

Hamuy, M., Phillips, M. M., Maza, J., et al. 1995, AJ, 109, 1

Harkness, R. P., Wheeler, J. C., Margon, B., et al. 1987, ApJ, 317, 355

Hillebrandt, W., & Niemeyer, J. C. 2000, ARA&A, 38, 191

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19

Hönig, S. F., Watson, D., Kishimoto, M., et al. 2017, MNRAS, 464, 1693

Horesh, A., Kulkarni, S. R., Fox, D. B., et al. 2012, ApJ, 746, 21

Howell, D. A., Sullivan, M., Nugent, P. E., et al. 2006, Nature, 443, 308

Howell, D. A., Sullivan, M., Brown, E. F., et al. 2009, ApJ, 691, 661

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Iben, Jr., I., & Tutukov, A. V. 1984, ApJ, 284, 719

Inserra, C., Smartt, S. J., Jerkstrand, A., et al. 2013, ApJ, 770, 128

Ivezic, Z., Tyson, J. A., Axelrod, T., et al. 2009, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts #213, 366

Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, ApJS, 125, 439

Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Proc. SPIE, Vol. 7733, Ground-based and Airborne Telescopes III, 77330E

Karpenka, N. V., Feroz, F., & Hobson, M. P. 2013, MNRAS, 429, 1278

Kasen, D. 2010, ApJ, 708, 1025

Kasen, D., & Bildsten, L. 2010, ApJ, 717, 245

Kasen, D., Fernández, R., & Metzger, B. D. 2015, MNRAS, 450, 1777

Kashikawa, N., Shimasaku, K., Yasuda, N., et al. 2004, PASJ, 56, 1011

Kasliwal, M. M. 2011, PhD thesis, California Institute of Technology

Kasliwal, M. M. 2012, Publications of the Astronomical Society of Australia, 29, 482488

Kasliwal, M. M. 2012, PASA, 29, 482

Kasliwal, M. M., Kulkarni, S. R., Gal-Yam, A., et al. 2012, ApJ, 755, 161

Kawabata, K. S., Maeda, K., Nomoto, K., et al. 2010, Nature, 465, 326

Kessler, R., Marriner, J., Childress, M., et al. 2015, AJ, 150, 172

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Kokkotas, K. 2002

Kulkarni, S. R. 2013, The Astronomer's Telegram, 4807

Kuznetsova, N., Barbary, K., Connolly, B., et al. 2008, ApJ, 673, 981

Law, N. M., Kulkarni, S. R., Dekany, R. G., et al. 2009, PASP, 121, 1395

Leaman, J., Li, W., Chornock, R., & Filippenko, A. V. 2011, MNRAS, 412, 1419

Levitan, D., Kupfer, T., Groot, P. J., et al. 2013, MNRAS, 430, 996

Li, L.-X. 2008, MNRAS, 388, 1487

Li, W., Chornock, R., Leaman, J., et al. 2011a, MNRAS, 412, 1473

Li, W., Bloom, J. S., Podsiadlowski, P., et al. 2011b, Nature, 480, 348

Li, W., Leaman, J., Chornock, R., et al. 2011c, MNRAS, 412, 1441

Li, X.-D., & van den Heuvel, E. P. J. 1997, A&A, 322, L9

Liu, Y.-Q., Modjaz, M., Bianco, F. B., & Graur, O. 2016, ApJ, 827, 90

Lodders, K. 2003, ApJ, 591, 1220

Lunnan, R., Kasliwal, M. M., Cao, Y., et al. 2017, ApJ, 836, 60

Lyman, J. D., Levan, A. J., Church, R. P., Davies, M. B., & Tanvir, N. R. 2014, MNRAS, 444, 2157

Lyman, J. D., Levan, A. J., James, P. A., et al. 2016, MNRAS, 458, 1768

Maguire, K., Taubenberger, S., Sullivan, M., & Mazzali, P. A. 2016, MNRAS, 457, 3254

Maguire, K., Sullivan, M., Thomas, R. C., et al. 2011, MNRAS, 418, 747

Maguire, K., Sullivan, M., Ellis, R. S., et al. 2012, MNRAS, 426, 2359

Maguire, K., Sullivan, M., Pan, Y.-C., et al. 2014, MNRAS, 444, 3258

Mannucci, F., Della Valle, M., & Panagia, N. 2006, MNRAS, 370, 773

Mannucci, F., Della Valle, M., Panagia, N., et al. 2005, A&A, 433, 807

Maoz, D., & Mannucci, F. 2012, PASA, 29, 447

Maoz, D., Mannucci, F., & Brandt, T. D. 2012, MNRAS, 426, 3282

Maoz, D., Mannucci, F., & Nelemans, G. 2014, ARA&A, 52, 107

Mernier, F., de Plaa, J., Pinto, C., et al. 2016, A&A, 595, A126

Metzger, B. D. 2012, MNRAS, 419, 827

Modjaz, M., Blondin, S., Kirshner, R. P., et al. 2014, AJ, 147, 99

Monet, D. G., Levine, S. E., Canzian, B., et al. 2003, AJ, 125, 984

Moriya, T. J., & Maeda, K. 2016, ApJ, 824, 100

Mulchaey, J. S., Kasliwal, M. M., & Kollmeier, J. A. 2014, ApJ, 780, L34

Napiwotzki, R., Christlieb, N., Drechsel, H., et al. 2001, Astronomische Nachrichten, 322, 411

Neill, J. D., Sullivan, M., Balam, D., et al. 2006, AJ, 132, 1126

Neugebauer, G., Habing, H. J., van Duinen, R., et al. 1984, ApJ, 278, L1

Nomoto, K. 1982, ApJ, 253, 798

Nomoto, K., & Iben, Jr., I. 1985, ApJ, 297, 531

Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, Nuclear Physics A, 777, 424

Nugent, P., Cao, Y., & Kasliwal, M. 2015, in Proc. SPIE, Vol. 9397, Visualization and Data Analysis 2015, ed. D. L. Kao, M. C. Hao, M. A. Livingston, & T. Wischgoll, 939702

Nugent, P. E., Sullivan, M., Cenko, S. B., et al. 2011, Nature, 480, 344

Ofek, E. O., Laher, R., Law, N., et al. 2012a, PASP, 124, 62

Ofek, E. O., Laher, R., Surace, J., et al. 2012b, PASP, 124, 854

Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, PASP, 107, 375

Öpik, E. 1924, Publications of the Tartu Astrofizica Observatory, 25

Pain, R., Fabbro, S., Sullivan, M., et al. 2002, ApJ, 577, 120

Pakmor, R., Kromer, M., Taubenberger, S., et al. 2012, ApJ, 747, L10

Pakmor, R., Kromer, M., Taubenberger, S., & Springel, V. 2013, ApJ, 770, L8

Pan, Y.-C., Sullivan, M., Maguire, K., et al. 2014, MNRAS, 438, 1391

Paterno, M. 2004, Calculating Efficiencies and Their Uncertainties, Tech. Rep. FERMILAB-TM-2286-CD, Fermilab

Perets, H. B., Gal-yam, A., Crockett, R. M., et al. 2011, ApJ, 728, L36

Perets, H. B., Gal-Yam, A., Mazzali, P. A., et al. 2010, Nature, 465, 322

Perley, D. A., Quimby, R. M., Yan, L., et al. 2016, ApJ, 830, 13

Perlmutter, S., Gabi, S., Goldhaber, G., et al. 1997a, ApJ, 483, 565

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565

Perlmutter, S. A., Deustua, S., Gabi, S., et al. 1997b, in NATO Advanced Science Institutes (ASI) Series C, Vol. 486, NATO Advanced Science Institutes (ASI) Series C, ed. P. Ruiz-Lapuente, R. Canal, & J. Isern, 749

Perrett, K., Balam, D., Sullivan, M., et al. 2010, AJ, 140, 518

Perrett, K., Sullivan, M., Conley, A., et al. 2012, AJ, 144, 59

Phillips, M. M. 1993, ApJ, 413, L105

Phillips, M. M., Lira, P., Suntzeff, N. B., et al. 1999, AJ, 118, 1766

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13

Poveda, A., Allen, C., & Hernández-Alcántara, A. 2007, in IAU Symposium, Vol. 240, Binary Stars as Critical Tools and Tests in Contemporary Astrophysics, ed. W. I. Hartkopf, P. Harmanec, & E. F. Guinan, 417–425

Poznanski, D., Maoz, D., Yasuda, N., et al. 2007, MNRAS, 382, 1169

Prajs, S., Sullivan, M., Smith, M., et al. 2017, MNRAS, 464, 3568

Quimby, R. M., Kulkarni, S. R., Kasliwal, M. M., et al. 2011, Nature, 474, 487

Rau, A., Kulkarni, S. R., Law, N. M., et al. 2009, PASP, 121, 1334

Richardson, D., Jenkins, III, R. L., Wright, J., & Maddox, L. 2014, AJ, 147, 118

Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, ApJ, 473, 88

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Macri, L., Casertano, S., et al. 2011, ApJ, 730, 119

Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56

Rodney, S. A., Riess, A. G., Strolger, L.-G., et al. 2014, AJ, 148, 13

Röpke, F. K., Kromer, M., Seitenzahl, I. R., et al. 2012, ApJ, 750, L19

Rosswog, S., Kasen, D., Guillochon, J., & Ramirez-Ruiz, E. 2009, ApJ, 705, L128

Ruiter, A. J., Belczynski, K., & Fryer, C. 2009, ApJ, 699, 2026

Ruiter, A. J., Sim, S. A., Pakmor, R., et al. 2013, MNRAS, 429, 1425

Saio, H., & Nomoto, K. 1998, ApJ, 500, 388

Sako, M., Bassett, B., Becker, A., et al. 2008, AJ, 135, 348

Sako, M., Bassett, B., Becker, A. C., et al. 2014, ArXiv e-prints, arXiv:1401.3317

Santander-Garcia, M., Rodriguez-Gil, P., Corradi, R. L. M., et al. 2015, Nature, 519, 63

Scalzo, R., Yuan, F., Childress, M. J., et al. 2017, ArXiv e-prints, arXiv:1702.05585

Scannapieco, E., & Bildsten, L. 2005, ApJ, 629, L85

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Sell, P. H., Maccarone, T. J., Kotak, R., Knigge, C., & Sand, D. J. 2015, MNRAS, 450, 4198

Shappee, B. J., Prieto, J. L., Grupe, D., et al. 2014, ApJ, 788, 48

Shara, M. M., Doyle, T., Lauer, T. R., et al. 2017, ApJ, 839, 109

Shen, K. J., Bildsten, L., Kasen, D., & Quataert, E. 2012, ApJ, 748, 35

Shen, K. J., Kasen, D., Weinberg, N. N., Bildsten, L., & Scannapieco, E. 2010, ApJ, 715, 767

Shivvers, I., Modjaz, M., Zheng, W., et al. 2017, PASP, 129, 054201

Sim, S. A., Fink, M., Kromer, M., et al. 2012, MNRAS, 420, 3003

Smartt, S. J. 2009, ARA&A, 47, 63

Smartt, S. J., Valenti, S., Fraser, M., et al. 2015, A&A, 579, A40

Smith, R. M., Dekany, R. G., Bebek, C., et al. 2014, in Proc. SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 914779

Starrfield, S., Truran, J. W., Sparks, W. M., & Kutter, G. S. 1972, ApJ, 176, 169

Strolger, L.-G., Riess, A. G., Dahlen, T., et al. 2004, ApJ, 613, 200

Sullivan, M., Le Borgne, D., Pritchet, C. J., et al. 2006, ApJ, 648, 868

Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782

Sullivan, M., Guy, J., Conley, A., et al. 2011a, ApJ, 737, 102

Sullivan, M., Kasliwal, M. M., Nugent, P. E., et al. 2011b, ApJ, 732, 118

Tutukov, A. V., & Yungelson, L. R. 1981, Nauchnye Informatsii, 49, 3

Valenti, S., Yuan, F., Taubenberger, S., et al. 2014, MNRAS, 437, 1519

van den Bergh, S. 1991, Phys. Rep., 204, 385

van den Heuvel, E. P. J., Bhattacharya, D., Nomoto, K., & Rappaport, S. A. 1992, A&A, 262, 97

Waldman, R., Sauer, D., Livne, E., et al. 2011, ApJ, 738, 21

Walter, C. W. 2015, Journal of Instrumentation, 10, C05015

Webbink, R. F. 1984, ApJ, 277, 355

Whelan, J., & Iben, Jr., I. 1973, ApJ, 186, 1007

Woosley, S., & Janka, T. 2005, Nature Physics, 1, 147

Woosley, S. E. 2010, ApJ, 719, L204

Woosley, S. E., & Kasen, D. 2011, ApJ, 734, 38

Wright, A., & Otrupcek, R. 1990, in PKS Catalog (1990)

Yaron, O., & Gal-Yam, A. 2012, PASP, 124, 668

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Yuan, F., Kobayashi, C., Schmidt, B. P., et al. 2013, MNRAS, 432, 1680

Yüksel, H., Kistler, M. D., Beacom, J. F., & Hopkins, A. M. 2008, ApJ, 683, L5

Zackay, B., Ofek, E. O., & Gal-Yam, A. 2016, ApJ, 830, 27

Zaritsky, D., Gonzalez, A. H., & Zabludoff, A. I. 2004, ApJ, 613, L93

Zhang, B. R., Childress, M. J., Davis, T. M., et al. 2017, ArXiv e-prints, arXiv:1706.07573

Zwicky, F. 1942, ApJ, 96, 28