Nebular-phase Spectra of Superluminous Supernovae: Physical Insights from Observational and Statistical Properties

Matt Nicholl1,2, Edo Berger1, Peter K. Blanchard1, Sebastian Gomez1, and Ryan Chornock3

1 Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA; matt.nicholl@cfa.harvard.edu
2 Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
3 Astrophysical Institute, Department of Physics and Astronomy, 251B Clippinger Lab, Ohio University, Athens, OH 45701, USA

Received 2018 August 1; revised 2018 November 6; accepted 2018 November 26; published 2019 January 24

Abstract

We study the spectroscopic evolution of superluminous supernovae (SLSNe) later than 100 days after maximum light. We present new data for Gaia16apd and SN 2017egm and analyze these with a larger sample comprising 41 spectra of 12 events. The spectra become nebular within 2–4 e-folding times after light-curve peak, with the rate of spectroscopic evolution correlated to the light-curve timescale. Emission lines are identified with well-known transitions of oxygen, calcium, magnesium, sodium, and iron. SLSNe are differentiated from other SNe Ic by a prominent O I λ7774 line and higher ionization states of oxygen. The iron-dominated region around 5000 Å is more similar to broad-lined SNe Ic than to normal SNe Ic. Principal component analysis shows that five “eigenspectra” capture ≥70% of the variance, while a clustering analysis shows no clear evidence for multiple SLSN subclasses. Line velocities are 5000–8000 km s⁻¹ and show stratification of the ejecta. O I λ7774 likely arises in a dense inner region that also produces calcium emission, while [O I] λ6300 comes from farther out until 300–400 days. The luminosities of O I λ7774 and Ca II suggest significant clumping, in agreement with previous studies. Ratios of [Ca II] λ7300/[O I] λ6300 favor progenitors with relatively massive helium cores, likely ≥6 M⊙, though more modeling is required here. SLSNe with broad light curves show the strongest [O I] λ6300, suggesting larger ejecta masses. We show how the inferred velocity, density, and ionization structure point to a central power source.

Key words: supernovae: general – supernovae: individual (SN2017egm, Gaia16apd, PS17aea)

1. Introduction

In recent years, much progress has been made in characterizing the new population of hydrogen-poor superluminous supernovae (SLSNe). These events first started to appear in wide-field, untargeted transient surveys (Chomiuk et al. 2011; Quimby et al. 2011; Gal-Yam 2012) and now constitute a few percent of the SNe classified each year. SLSNe have a median luminosity M ≈ −21 mag (Nicholl et al. 2015; De Cia et al. 2017; Lunnan et al. 2018), making them up to two orders of magnitude brighter than typical SNe, and sparking intense interest in the unexpectedly diverse outcomes of massive-star deaths.

SLSNe are now generally classified spectroscopically rather than photometrically. Their unique early-time spectra show a series of broad O II absorption lines superposed on a blue continuum, indicating hot, ionized ejecta. As they expand and cool, the spectra evolve to resemble those of lower-luminosity SNe Ic (Pastorello et al. 2010; Inserra et al. 2013), though there may be subtle differences (Liu & Modjaz 2017; Quimby et al. 2018). Thus, SLSNe are best described as SNe Ic (explosions of stripped stars lacking hydrogen and helium) that manage to stay hot (T ≥ 10,000 K) over several weeks or months, allowing them to attain higher luminosities.

The favored source of additional heating is a central engine, such as the spin-down of a rapidly rotating magnetar (Kasen & Bildsten 2010; Woosley 2010). It has been shown that this model can reproduce both the light curves (Chatzopoulos et al. 2013; Inserra et al. 2013; Nicholl et al. 2017c) and early spectra (Dessart et al. 2012; Howell et al. 2013; Mazzali et al. 2016) of SLSNe. Maximum-light observations, however, only probe the outer layers of the ejecta, because of a large optical depth to the center. Over time, recombination reduces the optical depth, and by t_neb ~ 360 days (M_ ej/10 M_⊙)¹/² (v/10⁴ km s⁻¹)⁻¹ after the explosion, where M_ ej is the ejecta mass and v is the expansion velocity, the ejecta become largely transparent (see review by Jerkstrand 2017). Once this so-called nebular phase is reached, it is possible to directly probe with spectroscopy the conditions at the center of the explosion, to constrain the composition and distribution of material and to search for any hydrodynamic signatures of the explosion mechanism. However, such observations are challenging because the SN will have faded substantially over the time t_neb and hence nebular spectroscopy is only currently possible for SLSNe at z ≤ 0.2.

Nicholl et al. (2016a) and Jerkstrand et al. (2017), following earlier work by Milisavljevic et al. (2013), showed that the few existing nebular spectra of SLSNe resemble those of broad-lined SNe Ic—thought to be engine-driven explosions and sometimes accompanied by long gamma-ray bursts (LGRBs). This was interpreted as evidence for a similar internal structure for SLSNe and LGRB SNe, suggesting that they may arise from similar progenitors. They also inferred large ejected masses of ≥10 M⊙ for some SLSNe. Since those initial nebular observations, the available sample of SLSNe with nebular-phase observations has increased as a number of recent nearby SLSNe have evolved to sufficiently late times. For example, Quimby et al. (2018) recently published a large spectroscopic sample of SLSNe from the Palomar Transient Factory (PTF), including many late-phase spectra.

In this paper, we undertake a systematic observational study of SLSN nebular spectra. We present new data for Gaia16apd/ SN 2016eay (z = 0.1013) and SN 2017egm (z = 0.0307), two nearby SLSNe that have been well studied at earlier phases (Kangas et al. 2017; Nicholl et al. 2017a, 2017b; Yan et al. 2017b; Bose et al. 2018), and combine these with all available published spectra of SLSNe obtained more than 100 days after maximum light.
We describe our observations and the processing of new and archival data in Section 2. In Section 3, we present the spectral sequence and mean population properties, including line identifications and comparisons to SNe Ic. We then apply machine learning techniques to characterize the diversity of SLSNe in Section 4. The line profiles are used to investigate the distribution of material in Section 5, and their luminosities and ratios are used to infer ejecta conditions in Section 6. We discuss the implications of our findings in the context of SLSN models in Section 7 and summarize our conclusions in Section 8.

2. Data

2.1. Gaia16apd

We observed Gaia16apd on 2017 May 17 (399 rest-frame days after peak luminosity) with GMOS on Gemini North (Hook et al. 2004). We used the R150 grating with a central wavelength of 7000 Å and the GG455 blocking filter to prevent second-order contamination. The data were reduced using the Gemini pipeline in PYRAF, to apply bias and flat-field corrections, determine a wavelength solution, and calibrate the relative flux with a standard star observed in the same setup. Narrow emission lines from the host galaxy were subtracted after fitting Gaussian profiles.

The spectrum is shown in Figure 1 and clearly contains contributions from both Gaia16apd (broad lines) and its host (continuum). After Gaia16apd had faded below detectability, as indicated by late-time imaging (P. Blanchard et al. 2018, in preparation), we obtained a host galaxy spectrum on 2017 November 26 (574 rest-frame days) using the Blue Channel spectrograph at MMT (Schmidt et al. 1989), with the 300 line mm\(^{-1}\) grating. We reduced the spectrum using standard PYRAF packages. This spectrum is also plotted in Figure 1 and shows no signs of residual SN contamination.

Finally, we subtracted the host spectrum from the earlier SN spectrum. The Blue Channel spectrum does not extend to the redder wavelengths covered by GMOS, so we used a simple linear extrapolation of the host galaxy light between 8000 and 10000 Å. The result of the subtraction is shown in Figure 1.

2.2. SN 2017egm

SN 2017egm is the second-closest known SLSN, at z = 0.0307, and is therefore ideal for a late-time study. We obtained multiple spectra spanning 126–353 days after maximum in the rest frame. Observations were carried out using the Ohio State Multi-Object Spectrograph (OSMOS; Martini et al. 2011) on the 2.4 m Hiltner telescope at MDM observatory, Blue Channel and Binospec on MMT, and GMOS on Gemini North. Binospec\(^5\) is a newly commissioned imaging spectrograph with dual fields of view. We used the 270 line mm\(^{-1}\) grating, providing spectral coverage from 3900 to 9240 Å at ~6 Å resolution.

OSMOS and Blue Channel data were de-biased, flat-fielded, and wavelength- and flux-calibrated in IRAF/PYRAF, while for the Gemini data we used the GMOS pipeline. Binospec data were reduced with a dedicated pipeline, based on the pipeline for the MMT Magellan Infrared Spectrograph (Chilingarian et al. 2013).

Almost uniquely among SLSNe, SN 2017egm occurred in a massive spiral galaxy, rather than the metal-poor dwarfs that typically host these explosions (Chen et al. 2017; Nicholl et al. 2017b; Bose et al. 2018; Izzo et al. 2018). Host galaxy light was removed when extracting the 1D spectra using low-order polynomials along the direction of the slit, fitted to regions of the 2D host spectrum on either side of the SN spectral trace. While the host light profile appears smooth on the spectrograph, we cannot exclude the possibility that a bright region underlying the SLSN location could be contributing some excess continuum. However, this can only be tested once SN 2017egm has completely faded.

The resultant spectra are shown in Figure 2. The earliest spectrum at 126 rest-frame days still shows a relatively blue continuum. The spectrum evolves quickly between the first two epochs, but slowly thereafter. Initially, the spectrum is very similar to SN 2015bn (Nicholl et al. 2016b), but as it evolves, it is better matched by LSQ14an (Jerkstrand et al. 2017).

2.3. SLSNe from PTF

Quimby et al. (2018) recently published a spectroscopic sample of all SLSNe classified in PTF. While their rigorous analysis of the spectral properties and line evolution focuses only on the photospheric phase, their sample also includes many late-time spectra. The authors have made these spectra available via WISEREP (Yaron & Gal-Yam 2012), with various levels of processing. We downloaded the spectra that have had host emission lines removed, but with no further smoothing. We carry out our own smoothing simply to ensure that all spectra in our analysis are subject to consistent processing (see Section 2.5).

Many of the spectra of PTF events show clear evidence of host galaxy continuum, similar to that shown for Gaia16apd in Figure 1. Fortunately, most of these galaxies were observed by

---

\(^5\) The closest is SN 2018bsz at z = 0.0267 (Anderson et al. 2018), but that event is too young for a nebular study at the time of writing.

---

Figure 1. Gemini/GMOS spectrum of Gaia16apd, an SLSN at z = 0.1013, at 399 rest-frame days after maximum light. We subtract a host spectrum obtained with MMT/Blue Channel, with a linear extrapolation redward of 8000 Å.
Perley et al. (2016), after the SLSNe had faded. These host spectra are also available from WISeREP. The galaxies are typically faint, and the host spectra are often noisy. To avoid adding additional noise to the SLSN spectra, we median-filtered the host spectra over 100 Å windows (after manually removing any emission lines), leaving only smooth continua. While this heavy smoothing will result in a failure to remove any narrow stellar absorption features present in the SLSN spectra, no such features are visible above the noise, nor would they have a significant impact on our analysis. Moreover, we do not expect strong absorption features in the host spectra, since these galaxies are dominated by young stellar populations (Lunnan et al. 2014; Leloudas et al. 2015; Angus et al. 2016; Chen et al. 2017; Perley et al. 2016; Schulze et al. 2018). Both the SLSN and host spectra were scaled to match photometry from De Cia et al. (2017) and Perley et al. (2016) before finally subtracting the host continua from the SLSN data.

2.4. Other SLSNe from the Literature

In addition to our new data and the host-subtracted PTF SLSN spectra, we include in our sample all published late-time spectra of SLSNe available from WISeREP (Yaron & Gal-Yam 2012) and the Open Supernova Catalog (Guillochon et al. 2017). These SLSNe (and the publications from which their spectra are taken) are SN 2007bi (Gal-Yam et al. 2009; Young et al. 2010), PTF12dam (Nicholl et al. 2013; Chen et al. 2015), SN 2015bn (Nicholl et al. 2016a, 2016b; Jerkstrand et al. 2017), LSQ14an (Inserra et al. 2017; Jerkstrand et al. 2017), and Gaia16apd (Kangas et al. 2017; Nicholl et al. 2017a). In total, our sample comprises 41 spectra of 12 SLSNe. We did not include SLSNe for which no SN features were visible above host galaxy light: SN 2011ke (Quimby et al. 2018) and PS16aqy (Blanchard et al. 2018). Two additional events with late-time spectra were not included because the wavelength range of their spectra did not cover the main lines of interest: PS1-14bj (Lunnan et al. 2016) and PTF09atu (Quimby et al. 2018). We also exclude the few SLSNe with a dominant Hα line in their spectra. These have been extensively analyzed by Yan et al. (2015, 2017a). We exclude these events because the hydrogen emission most likely probes circumstellar material, rather than the interior of the explosion, while obscuring important SN features. A list of all spectra is given in Table 1.

| SLSN       | \(z^a\) | \(\tau_b^b\) | \(\tau^c\) | Reference |
|------------|---------|-------------|------------|-----------|
| SN 2017egm | 0.0307  | 60          | 126        | This work |
|            |         |             | 141        | This work |
|            |         |             | 159        | This work |
|            |         |             | 191        | This work |
|            |         |             | 212        | This work |
|            |         |             | 353        | This work |
| Gaia16apd  | 0.1013  | 43          | 169        | Nicholl et al. (2017a) |
|            |         |             | 198        | Kangas et al. (2017)  |
|            |         |             | 399        | This work |
| SN 2015bn  | 0.1136  | 80          | 106        | Nicholl et al. (2016b) |
|            |         |             | 243        | Nicholl et al. (2016b) |
|            |         |             | 256        | Nicholl et al. (2016a) |
|            |         |             | 295        | Nicholl et al. (2016a) |
|            |         |             | 315        | Jerkstrand et al. (2017) |
|            |         |             | 343        | Nicholl et al. (2016a) |
|            |         |             | 392        | Nicholl et al. (2016a) |
| PTF12dam   | 0.1075  | 73          | 171        | Nicholl et al. (2013)  |
|            |         |             | 221        | Nicholl et al. (2013)  |
|            |         |             | 269        | Quimby et al. (2018)   |
|            |         |             | 324        | Quimby et al. (2018)   |
|            |         |             | 509        | Chen et al. (2015)     |
| LSQ14an    | 0.1637  | 85          | 111        | Insera et al. (2017)   |
|            |         |             | 149        | Insera et al. (2017)   |
|            |         |             | 365        | Jerkstrand et al. (2017) |
|            |         |             | 410        | Jerkstrand et al. (2017) |
| SN 2007bi  | 0.1279  | 85          | 367        | Gal-Yam et al. (2009)  |
|            |         |             | 471        | Young et al. (2010)    |
| PTF10hgi   | 0.0982  | 36          | 241        | Quimby et al. (2018)   |
|            |         |             | 315        | Quimby et al. (2018)   |
| PTF10mmn   | 0.1236  | 82          | 182        | Quimby et al. (2018)   |
|            |         |             | 213        | Quimby et al. (2018)   |
|            |         |             | 321        | Quimby et al. (2018)   |
|            |         |             | 527        | Quimby et al. (2018)   |
| PTF09cnd   | 0.2585  | 75          | 121        | Quimby et al. (2018)   |
| PTF10weg   | 0.190   | 40          | 147        | Quimby et al. (2018)   |
| PTF1hqr    | 0.0571  | 67          | 159        | Quimby et al. (2018)   |
|            |         |             | 350        | Quimby et al. (2018)   |
| PTF12hni   | 0.1056  | 27          | 298        | Quimby et al. (2018)   |

Notes.

\(a\) Redshift.

\(b\) Light-curve e-folding timescale, in rest-frame days.

\(c\) Time at which spectrum was obtained, in rest-frame days since light-curve maximum.

2.5. Processing

Prior to our analysis, we corrected all spectra for Galactic extinction and de-redshifted them to the rest frame. We used the Schlafly & Finkbeiner (2011) calibration for dust extinction along the line of sight. Values for \(E(B - V)\) and \(z\) were obtained from the Open Supernova Catalog (Guillochon et al. 2017) or directly
from the file headers of the Quimby et al. (2018) spectra. We assume that extinction in the SLSN host galaxies is negligible (Lunnan et al. 2014; Leloudas et al. 2015; Nicholl et al. 2017c).

The signal-to-noise ratio (S/N) varies significantly between the spectra in our sample. In some cases, noisy or oversampled spectra make it difficult even to identify real spectral features. Additionally, it greatly facilitates a statistical analysis to have all spectra on a common wavelength grid. We therefore apply a two-step interpolation and smoothing procedure to our spectra.

First, we smoothed each spectrum using a Savitsky–Golay filter implemented in SCIPPY. This algorithm replaces the flux in each pixel with an interpolated value based on a polynomial fit of order $n$ to the neighboring pixels within a window $\pm w$. After experimenting with a range of values, we found that using $w = 19$ pixels (i.e., fitting to a region spanning 45 Å on either side of the target wavelength) and $n = 1–2$ gave satisfactory results. We examined each spectrum by eye to determine which value of $n$ achieved the best balance between accurately preserving the shape of real structure and suppressing the noise. In practice, we generally used $n = 2$ for spectra with a clean signal and $n = 1$ for noisy data.

We then linearly interpolated the spectra to a common dispersion scale of $\Delta \lambda = 5$ Å, spanning a wavelength range 3000–9000 Å in the rest frame of each SLSN. This range was chosen to cover all spectral lines of interest, while avoiding any flux calibration uncertainties at the extrema of either the SLSN or host spectra. Finally, we produced normalized spectra by dividing each spectrum by its mean flux. An example of this smoothing (and prior host subtraction) is shown in Figure 3.

## 3. Observed Properties

### 3.1. Spectral Sequence

All spectra in our sample are shown in Figure 4, following the processing described in Section 2.5 (a version with unsmoothed spectra is provided in Appendix C). Each spectrum is labeled with the date on which it was obtained, measured in rest-frame days relative to when that SLSN reached its peak brightness. We use the dates of bolometric maximum as given in the literature.

SLSNe show a broad range of timescales in their photometric and spectroscopic evolution (Nicholl et al. 2015), so displaying spectra in order of time after peak is not necessarily a fair comparison. Instead, we rank our spectra using a normalized timescale $t/t_d$, where $t$ is the rest-frame phase from maximum light and $t_d$ is the exponential decay time of the light curve for each event, which ranges over $\approx 25–90$ days. Nicholl et al. (2015) and De Cia et al. (2017) provide tabulated values for most of this sample, while individual values for Gaia16apd and SN 2017egm come from Kangas et al. (2017) and Bose et al. (2018), respectively.

Virtually all of these spectra are at normalized phases $t/t_d > 2$. We show in the left panel of Figure 4 those spectra with $t/t_d < 4$ (23 spectra) and in the right panel those with $t/t_d > 4$ (18 spectra). This immediately leads to an important insight. Spectra on the left often show residual continua and have a weak or poorly developed feature at 6300 Å. The strongest line in these spectra is generally at 7300 Å, usually attributed to [Ca II] $\lambda 7300$ (really a doublet with components at 7291 and 7324 Å). The later spectra in the right panel, on the other hand, show little continuum, and the strongest line is nearly always at 6300 Å, consistent with [O I] $\lambda 6300$ (also a doublet, at 6300 and 6364 Å).

Therefore, we conclude that SLSNe can reach a pseudonebular phase, with several prominent emission lines but some residual continuum, after $\approx 2$ light-curve timescales from maximum light. By $\approx 4$ decline times, SLSNe tend to be fully nebular. The nature of the continuum is not certain—it could arise from an internal photosphere, circumstellar interaction, or possibly a forest of overlapping broad lines. The fact that [O I] $\lambda 6300$ is not fully developed during the continuum phase suggests that the ejecta are still relatively dense ($n_e \sim 10^{3} cm^{-3}$), so a residual photosphere may be most likely (however, see Section 5).

In 7 of the 10 SLSNe with spectra at $t/t_d > 4$, the strongest line is the $\lambda 6300$ line, with ratios $\lambda 6300/\lambda 7300 \gtrsim 1$–2. However, there are three that show weak $\lambda 6300$ features and prominent $\lambda 7300$ emission, with $\lambda 6300/\lambda 7300 < 0.3$. These events are PTF10hgi, LSQ14an, and SN 2017egm, including their spectra at quite late phases of 241–410 days, or $4.3t_d<8.8t_d$. Jerkstrand et al. (2017) and Insetra et al. (2017) showed that the strong $\lambda 7300$ feature in LSQ14an is likely due to contamination by [O II] $\lambda \lambda 7320$, 7330. The case of SN 2017egm is even more surprising, as its spectrum is initially almost identical to the prototypical SN 2015bn, but as it evolves, the line at 6300 Å never becomes dominant. We explore these outliers further in Sections 5 and 6.

### 3.2. Mean Spectrum

We begin by constructing the average nebular spectrum of SLSNe. We sum the normalized flux of all processed spectra in each 5 Å bin and then divide by the number of spectra contributing in that bin. The sums of fluxes and spectra per bin are shown in the top panel of Figure 5. In order that the mean spectrum not be biased by the few events with the most data, we also constructed a sum taking only one spectrum per SLSN.

---

De Cia et al. (2017) give the time to decline by 1 mag in $g$ band; we convert to $t_d$ using an empirical relation $t_d \approx 1.25 t_d, 1_{mag}$, based on three events in common with Nicholl et al. (2015).
choosing the spectrum with the best wavelength coverage. As shown in Figure 5, this made no appreciable difference to the result, so we take the mean of the full sample as our average spectrum.

In the bottom panel of Figure 5, we show the mean and standard deviation, as well as individual means for the spectra at $t/t_d < 4$ and $t/t_d > 4$. These two subsets are mostly similar and lie within the $1\sigma$ contours of the overall mean. The most significant difference is at 6300 Å, which is comparable in strength to the $\lambda 7300$ line when averaged over the whole sequence, but the ratio between these lines varies from $\sim 0.5$ in the earlier spectra to $\sim 2$ in the latter. The relative lack of evolution outside of the $\lambda 6300$ line, over significant periods of time, was noted for specific SLSNe by Jerkstrand et al. (2017) and Inserra et al. (2017); here we show that this holds true for a much larger sample of the population.

### 3.3. Line Identifications

The main advantage of constructing the average spectrum is to identify significant features that are common in SLSNe but may be difficult to detect in individual events at modest S/N. In Figure 6, we overplot spectral lines of common species observed in SNe. For each species, we searched the NIST Atomic Spectra Database (Kramida et al. 2015) for transitions between low-lying energy levels, with a particular focus on forbidden lines close to the ground state.
Most lines are matched using only five elements: oxygen, calcium, iron, magnesium, and sodium. We also detect possible contributions from carbon and sulfur. While the strongest oxygen lines are from neutral O I, there are weaker features consistent with O II and O III. The higher-ionization features overlap with iron lines, so their identification is not always clear—we investigate this further in the next section.

Calcium and iron appear singly ionized, while oxygen and magnesium appear neutral. This is mostly consistent with their relative ionization potentials: oxygen has a much greater ionization potential (13.6 eV) than these other elements (6–8 eV). We do not see Fe III (the second ionization potential of iron is 16.2 eV), which is responsible for the strongest line in the nebular spectra of SNe Ia (Axelrod 1980). Although magnesium has a lower ionization potential close to that of neutral iron, Mg II has few distinct lines in the optical. Asymmetry in the O I \(\lambda 7774\) line profile may be an indication of contamination from Mg II \(\lambda\lambda 7877, 7896\), as we will discuss in Section 5. Mg I \(\lambda 4571\) is both a recombination line and an efficient coolant, so it is often strong even when most magnesium is ionized (Jerkstrand 2017). None of the species listed above predict additional strong lines not seen in our spectra. We apply a similar consistency check to test for weak lines of hydrogen or helium. While some features match the locations of these lines, they predict many unseen features, so it is not possible to interpret any of the lines as hydrogen or helium in a self-consistent manner. Thus, we rule out any substantial contribution from these two species.

We also investigate any subtle differences in line identifications as a function of time, by comparing the mean spectra at \(t/t_d < 4\) and \(t/t_d > 4\). Figure 7 shows close-ups around several features identified above. The line at \(\approx 4000\) Å can be attributed to a blend of Ca II H and K and the S II \(\lambda 4069\) doublet. The line center shifts slightly toward S II at later phases. The Fe II blend at \(\approx 5250\) Å likely contains a contribution from Mg I \(\lambda 5183\) (Nicholl et al. 2013; Jerkstrand et al. 2017)—this line also changes as the spectra evolve, indicating that the relative contribution from magnesium likely increases over time.

The line center of O I \(\lambda 6300\) also shifts over time. This line is a doublet, with wavelengths of 6300 and 6364 Å. When the line is optically thick, these components have approximately equal strengths, but in the optically thin regime their ratio is 3:1. Thus, the shift of the central wavelength to the blue shows that as the line becomes stronger after \(t/t_d \sim 4\), it also becomes

![Figure 6. Line identifications in SLSNe. All atomic data were retrieved from NIST. All of the strong lines can be associated with oxygen, calcium, magnesium, sodium, and iron, with possible contributions from carbon and sulfur. No hydrogen or helium is observed.](image-url)
Figure 7. Zoom-in around blended lines from Figure 6. The red dashed line corresponds to $t/t_d < 4$, while the blue–green dotted line corresponds to $t/t_d > 4$. The peak wavelengths shift over time as the ratios between components evolve.

optically thin. The final panel of Figure 7 shows the Ca II near-IR (NIR) triplet, which may be blended with [C I]. However, we see no significant evidence for a change in the relative contribution from [C I] between the early and late nebular phase.

3.4. Comparison to SNe Ic

We conclude this section with a comparison to normal and broad-lined SNe Ic, to which SLSNe are known to share similar spectroscopic properties in both the photospheric (Pastorello et al. 2010; Inserra et al. 2013; Liu et al. 2017; Quimby et al. 2018) and nebular (Gal-Yam et al. 2009; Nicholl et al. 2016b; Jerkstrand et al. 2017) phases. We retrieve from the Open Supernova Catalog (Guillochon et al. 2017) all spectra of SNe Ic obtained more than 100 days after maximum light. We exclude any spectra with clear evidence for host galaxy contamination in the continuum. Because SNe Ic exhibit a faster evolution than SLSNe, as we will show, we further include any additional spectra of these same SNe Ic obtained $\gtrsim 50$ days after maximum. This results in 80 spectra of 15 SNe Ic, which are shown in Appendix C. We construct a mean spectrum for SNe Ic following the method in Section 3.2. To facilitate a fair comparison, we also tried omitting SN Ic spectra with $t/t_d > 11$, the latest observed normalized phase in our SLSN sample (we assume $t_d \approx 20$ days for typical SNe Ic and 57 days for the slowly evolving SN 2011bm; Valenti et al. 2012; Nicholl et al. 2015). However, excluding these data made no significant difference to the mean spectrum.

We show a comparison of the mean Sn Ic spectrum to that of SLSNe in Figure 8. SNe Ic in the nebular phase generally exhibit a ratio of $[\text{Ca II}] \lambda 7300/[\text{O I}] \lambda 6300 \sim 0.5$, so we use the mean of SLSN spectra at $t/t_d > 4$ when the ratio is similar. That the median age of the Sn Ic spectra is only 150 days, compared to 359 days for SLSNe, immediately highlights the slower spectroscopic evolution of SLSNe compared to other stripped-envelope explosions.

While there is virtually a one-to-one correspondence in identified spectral lines, we note three key differences between the two classes. First, SLSNe exhibit a much stronger $\text{O I}\lambda 7774$ line, often with a shoulder on the red side that is not seen in lower-luminosity SNe Ic (see also Milisavljevic et al. 2013; Nicholl et al. 2016b). Second, SLSNe have a strong line at 5000 Å, the only feature that appears in one class but not the other. In Section 3.3, we interpreted this line as a possible blend of $\text{Fe II}$ and $\text{O III}$. The fact that the line does not appear in SNe Ic, which do display the other strong iron lines, indicates a substantial contribution from $\text{O III}$ in the SLSNe. This could be explained if SLSNe have higher temperature and ionization in the oxygen-rich ejecta than do SNe Ic, despite the fact that the SLSNe here have been expanding for twice as long as the SNe Ic at the epochs of observation. This line was also identified as $\text{O III}$ by Lunnan et al. (2016), Nicholl et al. (2016b), Jerkstrand et al. (2017), and Inserra et al. (2017).

Finally, SLSNe are clearly elevated in flux, on average, over the iron-dominated region between $\approx 4000$ and $5500$ Å. This was first pointed out by Gal-Yam et al. (2009) and Milisavljevic et al. (2013) in the context of SN 2007bi, the first SLSN with a nebular spectrum (though this was partly attributable to host galaxy contamination; Jerkstrand et al. 2017). Recently, Moriya et al. (2018) showed that if some SLSNe were powered by fallback accretion, they would need to accrete $>1 M_\odot$ onto the central compact remnant to sustain their high luminosities. An observable consequence should be a deficiency of iron-group elements in the late-time spectra. The fact that SLSNe exhibit an excess, rather than a deficiency, in the iron lines therefore disfavors a fallback model.
between ≈4000 and 5500 Å that is similar to that of SN 1998bw. Thus, while SLSNe do exhibit an excess in this region compared to average SNe Ic, there is no evidence for an extreme fraction of iron-group elements above those seen in broad-lined SNe Ic (see also recent work by Blanchard et al. 2018, on the late-time light curves of SLSNe).

We conclude this section by noting that these differences between SLSNe and other SNe Ic are in some ways quite subtle, despite SLSNe being more luminous than SNe Ic by up to two orders of magnitude even at nebular times. If a large contribution to the luminosity of SLSNe was from circumstellar interaction, it is difficult to imagine how one would hide all obvious interaction features in the nebular spectrum (this is in contrast to the subset that show late-time hydrogen emission; Yan et al. 2017a).

4. Machine Learning Analysis

Machine learning algorithms are becoming increasingly valuable in astrophysical research. Recently, Inserra et al. (2018) applied some of these techniques to investigate a statistical basis for classifying SLSNe during the photospheric phase. In this section, we will analyze the nebular spectra of SLSNe using tools from the PYTHON package SCIKIT-LEARN (Pedregosa et al. 2012).

4.1. Principal Component Analysis

We begin by applying principal component analysis (PCA) to our spectra. Fundamentally, PCA transforms a set of observations (in our case, spectra) consisting of a number of variables (the flux in each wavelength bin), which are typically correlated (the fluxes in neighboring bins are of course highly correlated), into a set of uncorrelated vectors, i.e., the “principal components” (Pearson 1901; Hotelling 1933). These form an orthogonal basis for the space: any observation \( \mathbf{O}_i \) can be represented by

\[
\mathbf{O}_i = \sum_j a_{ij} \mathbf{C}_j,
\]

where \( \mathbf{C}_j \) is the \( j \)th principal component and \( a_{ij} \) the coordinates of the observation in the PCA basis.

For \( n \) observations (41 spectra) with \( p \) variables (1200 wavelength bins), the number of components is the lesser of \( n - 1 \) and \( p \) (i.e., 40). However, all components are not equally informative. The first component is chosen to explain the maximum amount of variance in the observations, and each successive component explains as much of the remaining variance as possible while remaining uncorrelated with the previous components. Thus, later components capture progressively weaker features, and eventually only noise, in the data.

We use the PCA method as implemented in the package SCIKIT-LEARN.DECOMPOSITION to transform our set of observed spectra into a basis of principal components. We find that over 70% of the sample variance can be explained using only the first five “eigenspectra.” These are shown in Figure 9. The components largely correspond to distinct spectral features. Almost 25% of the sample variance is explained by a component that primarily consists of [Fe II] lines around 5000 Å. Perhaps this is to be expected, since this is the most complicated region of the spectrum consisting of many blended lines, but it may also suggest a diversity in iron-group element production between different SLSNe.

The next three components are dominated successively by [O I] \( \lambda \)6300, the Ca II NIR triplet, and [Ca II] \( \lambda \)7300, together accounting for over 40% of the variance. The fifth component does not have a single line that is quite so dominant, but it does exhibit the strong \( \text{O I} \lambda \)7774 line. All strong lines identified in our spectra are included among these five eigenspectra, and we exclude any further components from our analysis, leaving a five-parameter model following Equation (1), where the free parameters are the coefficients, \( a_j \), for \( j = 1–5 \).

We verify this model by randomly excluding three spectra as a test sample, reconstructing the PCA basis from the remaining training sample (the results are indistinguishable compared to using the full sample), and fitting the test sample with the eigenspectra. The results are shown in Figure 10. We find that the PCA basis effectively captures the key features of the spectra.

4.2. Clustering

A key question in the study of SLSNe is whether all of these events form a continuous class, or if there are separate
subpopulations with different progenitors and/or power sources (Gal-Yam 2012; Nicholl et al. 2013, 2015, 2017c; De Cia et al. 2017; Inserra et al. 2018; Lunnan et al. 2018; Quimby et al. 2018). An interesting recent result from Inserra et al. (2018) is the possibility that SLSNe with slower light-curve evolution may show a flatter velocity gradient, though the sample of events with sufficient measurements remains small.

The velocity gradient probes the density structure in the outer layers of the ejecta. In the nebular phase, we have access to the whole ejecta, including the innermost regions that are most sensitive to differences in nucleosynthesis or hydrodynamic effects of different explosion mechanisms. Thus, it provides a valuable test of diversity in these events. Here we employ the same machine learning algorithm used by Inserra et al. (2018) to look for separate populations in our data.

\( k \)-means clustering partitions data into \( k \) clusters of equal variance, where each element belongs to the group whose mean is nearest to that element (Steinhaus 1956; Forgy 1965; MacQueen 1967; Lloyd 1982). The algorithm requires that \( k \) be specified in advance—it is therefore important to include some additional metric to determine the optimal number of clusters.

We reduce the dimensionality of our spectra from 1200 to 5 using the PCA decomposition described above, and then we apply the KMEANS method in SCIKIT-LEARN.CLUSTER. The decomposition is plotted in the PCA parameter space in Figure 11. We also tested the algorithm on the unreduced data set, finding that the assignment of spectra to clusters was essentially unchanged.

We search for the optimal number of clusters using three methods: the Bayesian information criterion (BIC; Schwarz et al. 1978), the elbow criterion, and silhouette analysis. The details can be found in Appendix A, but overall there is no strong evidence for clustering (\( k > 1 \)) in the data.

In the interest of completeness, we plot the clustering solution for \( k = 2 \) in Figure 12. The color coding indicates cluster membership, but the plot is otherwise identical to Figure 11. Clearly, the “clusters” are not well separated. Most significantly, individual SLSNe evolve between the two clusters; thus, cluster membership is primarily just a function of temporal evolution rather than physical differences between SLSNe. We therefore find no evidence for multiple subpopulations.

### 4.3. Prospects for Classification from Late-time Spectra

Spectroscopic classification of transients tends to be geared toward young events whenever possible, since this generally provides a path to greater scientific returns. However, due to factors such as unfavorable sky location, uneven survey cadence, and weather-related gaps, it is not uncommon for SNe to be typed quite late after explosion. If the light-curve peak is missed by photometric surveys, it is not possible to classify a target as an SLSN using the original \( M < -21 \) mag definition, the early O II absorption lines, or even the statistical definition proposed by Inserra et al. (2018). To avoid missing events entirely, classification using late-time spectra will be important, particularly in the era of the Large Synoptic Survey Telescope, for example, to improve our knowledge of the SLSN volumetric rate.

Quimby et al. (2018) showed that despite the strong similarity between SLSNe and normal SNe Ic during some phases of their evolution, there were sufficient subtle differences that they could be distinguished statistically. They did this through a rigorous process of template comparison, quantifying the best matches and determining threshold matching scores to separate SLSNe from the general SN population. Here we suggest an alternative (or complementary) approach, using PCA. This method is particularly suitable for...
late times, when emission lines are present in the spectra, but it may be appropriate for earlier phases too given an appropriate training set.

The method is simple, as we demonstrate using a test case. PS17aea was discovered by the Pan-STARRS Survey for Transients (Huber et al. 2015) on 2017 January 3, and we flagged it as a bright source in a faint galaxy. However, there were no observations of the field for over 200 days preceding this initial detection. We obtained a spectrum on 2017 January 28 using the Inamori-Magellan Areal Camera and Spectrograph (Dressler et al. 2011) on the 6.5 m Magellan Baade telescope. The spectrum showed narrow emission in Hα, Hβ, and [O III] from the host galaxy, from which we measured a redshift \( z = 0.104 \). The observed absolute magnitude at this distance is \(-18.8\) mag, which is not exceedingly bright for many well-known SN classes.

The spectrum is shown in Figure 13, after applying the same processing steps described in Section 2.5. The top panel shows a comparison to SN 2015bn at 106 days after maximum (Nicholl et al. 2016b), to which PS17aea displays an intriguing degree of similarity. However, when trying to classify this spectrum using GELATO,\(^7\) the top 10 matches include not only SN 2015bn but also a number of normal and broad-lined SNe Ic, SNe IIn, and even a peculiar SN Ia.

The presence of several emission features suggests a nebular character for PS17aea. Comparing to the mean spectrum of SLSNe at \( \approx 100–200 \) days, we find that PS17aea lies within the \( 1\sigma \) contours across essentially the full wavelength range and is therefore a good candidate for an SLSN evolving into the nebular phase. This would imply that the explosion occurred well before the first PanSTARRS Survey for Transients detection.

Comparison to the mean spectrum is a good start, but to proceed further, we can appeal to PCA. We fit the spectrum of PS17aea with the eigenspectra in Figure 9. We find that the model fits PS17aea as well as it fits the SLSN spectra in Figure 10. With only five components, the model does not capture the complexity seen around the \([O I] \lambda 6300\) line (see also SN 2015bn) but otherwise does an excellent job. For comparison, we construct an identical model for SNe Ic, using the comparison sample described in Section 3.4 and limiting the model to the first five eigenspectra as for SLSNe (this captures \( 91\% \) of the variance in the sample). Unlike the SLSN model, the SN Ic model does not satisfactorily capture some key features of the data, particularly the [Ca II] \( \lambda 7300 \) line and the strength of the [Fe II] bump. The reduced \( \chi^2 \) scores are \( \sim 2 \) for the SLSN model and \( \sim 5 \) for the SN Ic model. To ensure that this is not simply a consequence of different normalized phases between the SLSN and SN Ic samples, we also tried fitting the data with an SN Ic PCA model restricted to spectra at \( t/t_d < 11 \) (matching SLSN spectra) and found that using younger SNe Ic did not improve the fit quality. We therefore conclude that PS17aea was indeed a likely SLSN at relatively low redshift.

This test case shows the power of using PCA to resolve ambiguous SN classifications, even with a relatively modest training sample. In the future, further progress can be made by compiling large samples of all transient types, for example, from the Open Supernova Catalog (Guillochon et al. 2017), and using these to implement more advanced machine learning algorithms.

### 5. Line Profiles and Ejecta Distribution

Having investigated the statistical properties of nebular-phase SLSN spectra, we now move on to analyze the properties of individual events and spectral lines. The shapes of emission lines are an important diagnostic of the spatial distribution of material. Jerkstrand (2017) summarized the predicted line profiles for simple geometries such as uniform spheres and disks. Most notable is a flat-topped line if the emission comes from a hollow shell.

A double-peaked line profile would suggest a bipolar explosion viewed at a large angle from the dominant axis; this has been seen in several stripped-envelope SNe (Mazzali et al. 2005; Maeda et al. 2008). Taubenberger et al. (2009) conducted a detailed study of \([O I] \lambda 6300\) emission profiles in SNe Ic and found considerable diversity between events, including symmetric Gaussian profiles with and without an additional narrow core, symmetric double peaks from a bipolar geometry, and more complex asymmetric multi-peaked profiles.

#### 5.1. Continuum Subtraction

Before making velocity and flux measurements for spectral lines, any continuum must be removed. For each line in every spectrum, we manually identified \( 50–100\) Å wide regions to the blue and red of the line wings and approximated the continuum by fitting a linear function to these regions. A typical example is shown in Appendix B. For isolated lines (essentially anything redward of \( 6000\) Å), this process is fairly straightforward, but in the case of blended lines such as \([Mg II] \lambda 4571\), the “pseudo-continuum” consists of many blended lines, making it difficult to estimate its true level. In this case, flux measurements are likely a lower limit. To test the sensitivity of our velocity and flux measurements to the choice of continuum estimate, we additionally employed an automated procedure where we defined

---

\(^7\) https://gelato.tng.iac.es
two 100 Å regions centered at the local minimum flux on each side of the line to serve as the continuum level. We found that differences in derived properties were \( \lesssim 10\% \) compared to when we examined each SLSN manually.

5.2. Individual SLSNe

Figure 14 shows cutouts (in velocity space) around four of the strongest lines, \([\text{O I}] \lambda 6300\), \([\text{O I}] \lambda 7774\), \([\text{Ca II}] \lambda 7300\), and \([\text{Mg I}] \lambda 4571\), for all 12 SLSNe in the sample. We discuss these on a case-by-case basis below.

5.2.1. SN 2015bn

SN 2015bn serves as an excellent baseline for comparison, with high-S/N spectra well sampled from \( \approx 100 \)–400 days. The line profiles are approximately Gaussian at all times. The most interesting characteristic is the narrowness of \([\text{O I}] \lambda 7774\) relative to \([\text{O I}] \lambda 6300\). \([\text{O I}] \lambda 7774\) is primarily a recombination line and is stronger at higher densities. Thus, Nicholl et al. (2016a) suggested that the narrow \([\text{O I}] \lambda 7774\) emission originated in a high-density region at a low velocity coordinate, i.e., close to the center of the ejecta.
5.2.2. Gaia16apd

The line profiles in Gaia16apd are more complicated. Especially at later times, [O I] λ6300 and [Ca II] λ7300 show some indication of a broad base and slightly narrower core. Taubenberger et al. (2009) found that this morphology is common in SNe Ic and can be interpreted as either an axisymmetric explosion viewed edge-on (with the emitting material in a torus around the equator) or an ejecta distribution with a dense core.

O I λ7774 shows an even more interesting profile, with a shoulder or possible double peak on the red side, offset by ~4000 km s⁻¹ from the center. Emission lines with secondary peaks have been seen previously in SNe Ic, and Taubenberger et al. (2009) propose that it could be attributable to large-scale clumping, the ejection of massive blobs of material, or even a unipolar jet—though the last interpretation seems unlikely when the other lines do not share a similar profile. In particular contrast to the other lines, Mg I λ4571 shows a simple Gaussian profile.

5.2.3. LSQ14an

LSQ14an shows a large contrast between the flux at 7300 Å and that at 6300 Å. While Mg I λ4571 exhibits a simple Gaussian shape, the other lines show a high degree of asymmetry. The peaks of the oxygen lines are consistent with zero velocity, but the profiles cut off sharply on the blue side and exhibit a broad, red shoulder out to several thousand kilometers per second.

The very strong line at 7300 Å has sufficient S/N that it is clearly resolved into a double peak, with the components separated by ~2300 km s⁻¹. This feature is almost always attributed to [Ca II] λ7300, but given the strength of the line, and other lines of ionized oxygen in LSQ14an, Inserria et al. (2017) and Jerkstrand et al. (2017) suggested that this is a blend with [O II] λλ7330, 7300. The similarity in shape between this line and the [O I] λ6300 and O I λ7774 lines confirms that the blended feature is most likely dominated by [O II].

The oxygen-emitting region of LSQ14an thus appears to be highly asymmetric and with higher ionization than most SLSNe.

5.2.4. SN 2017egm

SN 2017egm initially exhibits very similar nebular-phase properties to SN 2015bn (this despite the much faster evolution in the light curve and the unusual metal-rich host galaxy; Nicholl et al. 2017b; Bose et al. 2018). However, the relative strength of the [O I] λ6300 line in SN 2017egm does not increase with time, and as with LSQ14an, we suggest that the strongest feature at 7300 Å is heavily contaminated by [O II].

There is also a hint of a narrow core to [O I] λ6300, and in the final spectrum it appears somewhat asymmetric, being broader on the red side, though less so than LSQ14an. The other lines are well developed and appear largely Gaussian, with the O I λ7774 line narrower than [O I] λ6300.

5.2.5. PTF12dam

The [O I] λ6300 line in PTF12dam strengthens slowly and starts out very broad, but it reaches a Gaussian profile by 324 days. The Mg I λ4571 line is broader than the others and develops a hint of a flat top as the spectrum evolves.

5.2.6. SN 2007bi

SN 2007bi shows fairly straightforward line evolution (note that the [Ca II] λ7300 in the later epoch is at low S/N and the implied line shape is likely contaminated by a spurious noise feature). The O I λ7774 line is again conspicuously narrow.

5.2.7. PTF09cnd

The [O I] λ6300 line in PTF09cnd has a possible flat-topped profile, which would indicate that the emission comes from a shell. However, given that there is only one epoch of spectroscopy at 121 days, it is also possible that we are seeing a transitional phase before the line fully develops. The latter interpretation is consistent with the high [Ca II] λ7300/ [O I] λ6300 ratio and overall very broad [O I] λ6300 profile. Support for a genuine flat top comes from Mg I λ4571, which also shows a hint of such a profile, while even O I λ7774 is broader and less Gaussian than in the rest of the sample. The peak of [Ca II] λ7300 appears to be blueshifted, which could arise if there is a large optical depth blocking emission from the receding part of the ejecta.

5.2.8. PTF10hqi

Like LSQ14an and SN 2017egm, the PTF10hqi spectra are dominated by the feature at 7300 Å, which at the observed strength is most likely a blend of [Ca II] λ7300 and [O I]. If most oxygen is ionized, this would help explain why [O I] λ6300 is so weak (barely above the noise at the earlier epoch, 241 days). At later times, [O I] λ6300 and O I λ7774 appear to have similar strengths and profiles, with a possible shoulder on the red side; however, in this case the λ7300 feature is a smooth Gaussian.

5.2.9. PTF10nnn

PTF10nnn shows features largely typical of the class, with Gaussian profiles and a narrow O I λ7774, once again with an asymmetric red shoulder. Mg I λ4571 in this case is weak and appears contaminated by spurious noise features.

5.2.10. PTF10vww

The lines in PTF10vww are on the broad side and appear to depart from simple Gaussian profiles, though we caution that, as with PTF09cnd, the only spectrum is at a fairly early phase (147 days) and has been heavily smoothed. Nevertheless, [O I] λ6300 shows a possible flat top, and it is the only event that shows this in the [Ca II] λ7300 profile too. There is also a hint of the same in Mg I λ4571. The O I λ7774 line is more typical, with a weak but detectable excess on the red side.

5.2.11. PTF11hrq

PTF11hrq has a prototypical line evolution with smooth Gaussian profiles, including a narrow O I λ7774 line.

5.2.12. PTF12hnj

The only clearly detected emission line in the spectrum of PTF12hnj is [O I] λ6300. There is a plausible narrow O I λ7774 centered at the correct wavelength, but this is equally consistent with noise. There is no obvious [Ca II] λ7300 emission at 298 days, while Mg I λ4571 falls outside the observed wavelength range.
5.3. Summary of Line Shapes

Overall, most SLSNe show smooth Gaussian line profiles without strong evidence for double peaks or flat tops, particularly in the [O I] λ6300 line that is most commonly used to diagnose geometry (Mazzali et al. 2005; Maeda et al. 2008; Taubenberger et al. 2009). This suggests that ejecta asymmetry is most likely modest. Polarization has been detected in the emission from SN 2015bn (Inserra et al. 2016; Leloudas et al. 2017) and SN 2017egm (Bose et al. 2018), implying that despite the smooth line profiles, there is certainly some asymmetry in these events, comparable to broad-lined SNe Ic. LSO14am stands out in our analysis as a candidate for a particularly asymmetric explosion.

The line profiles may also offer a clue to the nature of the spectral continuum at $t/t_d < 4$. From an observer’s point of view, an internal photosphere covers more of the receding side of the ejecta, predicting early blueshifted line profiles that move toward the rest wavelength with time as more of the ejecta becomes optically thin. The fact that we do not see significant blue-to-red movement in the emission profiles for any SLSNe suggests either that the source of continuum is external (i.e., interaction) or that the photosphere is already below a velocity coordinate $\lesssim 500$ km s$^{-1}$ (the typical size of our uncertainties) by 100 days after maximum light.

The most complicated line is O I $\lambda 7774$, which is often significantly narrower than the other strong oxygen line, [O I] $\lambda 6300$. It also shows a shoulder or secondary maximum in at least half of the SLSNe, but always on the red side. This cannot be explained in terms of an oxygen-rich ejecta blob or unipolar jet, since in either case we expect that the excess would be equally likely to be blueshifted.

Thus, the most likely explanation is contamination by another line. As shown in Figure 6, there is a doublet of Mg II at 7877 and 7896 Å, which could contribute to the red shoulder. Since this asymmetry is not typically seen in SNe Ic, stronger Mg II emission may constitute another difference between SLSNe and SNe Ic. As this line arises from an excited state of Mg II, a possible explanation is higher line excitation in SLSNe, due to either a higher temperature or some nonthermal effect (e.g., Mazzali et al. 2016). A future test will be to compare Mg II lines in the UV and NIR between SLSNe and SNe Ic. UV lines at 2795 and 2802 Å may contribute a large fraction of the cooling (Jerkstrand et al. 2017), while NIR lines at 9218 and 10914 Å result from the lower level of the $\lambda \lambda 7877, 7986$ transition.

5.4. Velocities

We now use these same lines to infer the velocity structure of the SLSN ejecta. Velocity measurements during the nebular phase have the advantage of probing deeper into the ejecta, compared to the photospheric phase, and hence measure the bulk velocities of different emitting layers. To proceed, we assume that all lines can be represented by Gaussian profiles, which works well in the majority of cases. In a few cases (typically at $t \lesssim 150$ days), the [O I] $\lambda 6300$ line does not exhibit a profile that can be reasonably represented with a Gaussian (see Appendix B). We exclude such cases from our analysis here.

We fit to [O I] $\lambda 6300$, [Ca II] $\lambda 7300$, O I $\lambda 7774$, and Mg I] $\lambda 4571$, accounting for the doublet nature of lines when appropriate. We assume a ratio of 3:1 between the 6300 and 6364 Å components of [O I], i.e., that the line is optically thin (assuming a 1:1 ratio made no significant changes to the inferred velocities). We fit to the unsmoothed spectra, in order to better quantify the error in the fit. We report the FWHM of each Gaussian, taking as our uncertainty the variance returned by SCIPY.CURVE.FIT.

An alternative measure of line velocity is the full-width at zero intensity (FWZI). This has a number of advantages over the FWHM: it is less sensitive to the shape of the line profile and better captures the full velocity extent of each ion. However, the FWHM is preferred here since it is less sensitive to the exact placement of the continuum. Taking the zero-flux level as the point where the line intensity falls below 1% of maximum, FWZI $\approx 2$ FWHM for a Gaussian profile. The FWHM contains 90% of the emitted flux, but we caution that 10% of each ionic species may be located at velocities extending to twice the values reported here.

The velocity evolution of the [O I] $\lambda 6300$ and [Ca II] $\lambda 7300$ lines is shown in Figure 15. We plot against the rest-frame phase, with and without normalizing to $t_p$, and in both cases compare to SNe Ic. Using both observed and normalized timescales allows for a fair comparison, and we see that while SLSNe initially exhibit much greater [O I] $\lambda 6300$ velocities than SNe Ic, this discrepancy largely disappears once we account for the slower light-curve evolution of the SLSNe.

While the overall trend of initially decreasing and then relatively flat velocity is exhibited by each line, the absolute scales are quite different. The [O I] $\lambda 6300$ velocities start out much higher than [Ca II] $\lambda 7300$ (and the other lines; see discussion below), up to $\sim 16,000$ km s$^{-1}$. The width drops quickly (and smoothly) as the line becomes more prominent and peaked; a typical plateau velocity for [O I] $\lambda 6300$ is $\sim 7000$ km s$^{-1}$. [Ca II] $\lambda 7300$ is the easiest line to measure in general, since it is strong and typically Gaussian-shaped starting from our earliest spectra. This line shows a very smooth evolution from around 10,000 km s$^{-1}$ at 100 days to 5000–7000 km s$^{-1}$ when its velocity plateaus.

We calculate the mean velocities of [O I] $\lambda 6300$, [Ca II] $\lambda 7300$, O I $\lambda 7774$, and Mg I] $\lambda 4571$ in 50-day bins (rest-frame phases) and bins of $\Delta(t/t_d) = 2$ (normalized phases); these are plotted in Figure 16 (all individual measurements for these four lines are shown in Appendix C). This clearly shows that [O I] $\lambda 6300$ and O I $\lambda 7774$ emission arises in regions with significantly different velocity, in many cases until 400 days (or $\sim 5 t_d$) after maximum. O I $\lambda 7774$ may approach 10,000 km s$^{-1}$ at early times but quickly settles at $\sim 5000$ km s$^{-1}$. [Ca II] $\lambda 7300$ emission appears to come from the same velocity zone. Mg I] $\lambda 4571$ shows the flattest velocity evolution among the strong lines, remaining between $\approx 7000$ and 12,000 km s$^{-1}$ at all times.

The observed hierarchy in velocities is consistent with explosion models of massive stars. Examining the models of stripped SNe used by Jerkstrand et al. (2014), oxygen is distributed across a wide range of zones of different elemental composition: it is abundant in the O/Si/S zone (Ne-burning ashes) and O/Ne/Mg zone (C ashes), and above this we expect unburned oxygen from the progenitor and/or an O/C zone (explosive He burning). Calcium is present in the Si/S (O burning) and O/Si/S zones, and [Ca II] $\lambda 7300$ is expected to be
the primary coolant of the Si/S zone (Jerkstrand et al. 2017). Magnesium is mostly in the O/Ne/Mg zone.

Figure 16 therefore represents a consistent picture in which [O I] λ6300 initially arises in the outer layers, but as time progresses we see a larger contribution from the deeper O/Ne/Mg zone, apparent when the [O I] λ6300 velocity crosses the Mg I λ4571 velocity at ~300 days. Later, we see emission from the O/Si/S zone after ~400 days, when the average velocity of [O I] λ6300 becomes comparable to O I λ7774 and [Ca II] λ7300. The O I λ7774 emission, in contrast, arises primarily from the O/Si/S zone at all times and is therefore a useful diagnostic of the density in this layer. These results will be important for our analysis in Section 7.

6. Line Luminosities and Physical Constraints

The luminosities of—and ratios between—spectral lines can be used to derive important constraints on the physical conditions in the ejecta, including the temperature, density, and mass of the emitting material. Clearly, such insights offer clues about the progenitors and power sources of SLSNe. Progress has been made through constructing detailed models for the line emission in SLSNe (Gal-Yam et al. 2009; Jerkstrand et al. 2017) and by employing analytic relations to convert integrated luminosities into physical parameters (Nicholl et al. 2016a; Jerkstrand et al. 2017). In both cases, the analysis points toward rather massive explosions, with inferred oxygen masses of $\gtrsim 10 M_\odot$ for some slowly evolving...
SLSNe. These events are thought to most likely originate from the high-mass end of the SLSN population (Nicholl et al. 2015, 2017c), whereas here we include SLSNe with a range of light-curve properties. Full radiative transfer modeling is beyond the scope of this work, but we can make progress by analyzing line ratios in a moderate-sized sample of SLSNe for the first time. We will make heavy use of relations provided by Jerkstrand et al. (2014, 2015, 2017) and Jerkstrand (2017).

6.1. Integrated Fluxes

We measure the integrated luminosities of various diagnostic lines: \([\text{O I}] \lambda 6300, \text{O I} \lambda 7774, [\text{O I}] \lambda 5577, \text{Mg I} \lambda 4571, [\text{Ca II}] \lambda 3730, \text{and the Ca II NIR triplet (8498, 8542, and 8662 Å). Most of these are reasonably well isolated and can be measured by direct numerical integration, after subtracting a linear fit to the continuum as in Section 5. The exception is [O I] \lambda 5577, which is weak and blended with [Fe II] \lambda 5528. In this case we estimate the luminosity using a double-Gaussian fit. The centers of the Gaussians are fixed at the rest wavelengths of these two lines, and their widths are assumed to be equal.}

Again we analyze the original unsmoothed spectra and use the observed S/N to calculate the uncertainties. While we have endeavored to scale all spectra to contemporaneous photometry before making measurements, in some cases large interpolations/extrapolations in light curves were required. We therefore impose an error floor of 20%, added in quadrature to the statistical error, to account for systematic errors in the absolute flux calibration or placement of the continuum. Note that uncertainties in continuum level affect our flux measurements at the \(\lesssim 10\%\) level; see Section 5.1.

The results are plotted in Figures 17 and 18. In the first case we show luminosities plotted against rest-frame phase from maximum, whereas in the latter we normalize the times by \(t_d\). We only show measurements for which the luminosity is measured at \(S/N > 3\) (i.e., the line is detected in the spectrum at \(> 3\sigma\) significance, including the additional systematic error in flux calibration/continuum placement).

The luminosities of the strongest lines can be as high as \(\sim 10^{42} \text{erg s}^{-1}\) at \(\sim 100\) days, dropping to \(\sim 10^{40} \text{erg s}^{-1}\) by 400–500 days. In general, this decrease is quite smooth in time, though this is less clear in the case of the noisy [O I] \(\lambda 5577\) line. The smooth evolution is particularly striking when accounting for the differences in light-curve decline rates: when plotted against \(t/t_d\), the measurements for all SLSNe cluster quite tightly around the same exponential slope (Figure 18). This indicates that the rate of fading in the spectrum is very closely tied to that of the light curve. Comparing to late-time SLSN bolometric light curves (Chen et al. 2015; Nicholl et al. 2016b), a strong line such as [O I] \(\lambda 6300\) typically accounts for around 5% of the total SLSN luminosity.

More instructive is to examine the ratios between line luminosities. This also has the advantage of being independent of the absolute flux normalization of the spectra, facilitating comparison with SNe Ic. In Figure 19, we show the luminosities of key diagnostic lines [Ca II] \(\lambda 7300, \text{O I} \lambda 7774, \text{and [O I]} \lambda 5577, \) normalized to the luminosity of [O I] \(\lambda 6300\). With the exception of [O I] \(\lambda 5577\), the luminosity in each line drops from \(\sim 1–2\) to \(\lesssim 0.5\) times the [O I] \(\lambda 6300\) luminosity (the same is true for Mg I \(\lambda 4571\); see Appendix C).

Compared to SNe Ic, the observed ratios of [O I] \(\lambda 6300/\lambda 7300\) and \(\lambda 7774/\lambda 5577\) appear significantly enhanced in SLSNe. Indeed, the early prominence of [Ca II] \(\lambda 7300\) relative to [O I] \(\lambda 6300\) and unusually strong [O I] \(\lambda 7774\) are some of the defining features of SLSN nebular spectra (Section 3). Figure 19 shows that this is partially attributable to the overall slow spectroscopic evolution of SLSNe compared to SNe Ic. When the phase is expressed relative to the light-curve timescale, we see that most of the SN Ic observations are at \(t/t_d > 5\), by which time the line ratios in SLSNe have decreased to comparable values to the SNe Ic. Accounting for the relative light-curve timescales mitigates for the fact that SLSNe stay hotter for longer than SNe Ic, and the eventual similarity in the line ratios may indicate that the structure and composition of SLSNe and SNe Ic are not dissimilar.

6.2. \([\text{O I]} \lambda 5577/\text{[O I]} \lambda 6300\) Ratio and Temperature

The [O I] \(\lambda 5577/\text{[O I]} \lambda 6300\) ratio appears fairly constant in time, at a typical value of 0.1–0.2. Assuming that these lines arise in the same zones, the ratio is sensitive only to the temperature and the optical depths in the lines. After \(\approx 200\)
days, the centroid of the [O I] λ6300 line is always close to 6300 Å, implying that the line is optically thin (Figure 7). The optical depth in [O I] λ5577 is lower than [O I] λ6300; therefore, both lines should be optically thin at late times (Jerkstrand et al. 2014), and we can use their ratio as a thermometer. There are significant caveats to this approach (Jerkstrand 2017)—for example, [O I] λ5577 may not form in local thermodynamic equilibrium (LTE) at late times, and the ratio can also depend on clumping.

Assuming that the lines are optically thin and form in LTE, the observed ratios ≲0.2 imply temperatures $T \lesssim 5000$ K, using Equation (2) from Jerkstrand et al. (2014), with an overall plausible range of 4000–6000 K for ratios of 0.05–0.3.

6.3. [O I] λ6300 and Ejecta Mass

Since oxygen is expected to account for 60%–70% of the ejecta in stripped-envelope SNe (e.g., Maurer & Mazzali 2010), inferring the oxygen mass gives a good handle on the total mass. Jerkstrand et al. (2017) found that the luminosity in the [O I] λ6300 line in some SLSNe could only be reproduced by models with $\gtrsim 10 M_\odot$ of oxygen. In the case of SN 2015bn, Nicholl et al. (2016a) used a scaling relation to estimate an oxygen mass of $\sim 9 M_\odot$ for a temperature of 4900 K.

Such analytic relations are fraught with danger, however, as the mass is an exponential function of the temperature, which is itself uncertain owing to the weakness of [O I] λ5577 and the difficulty of deblending this line with [Fe II]. SN 2015bn was an unusual case in that the spectra were of very high S/N and the [O I] λ5577 line was well resolved. Applying Equation (3) from Jerkstrand et al. (2014) results in ejecta masses spanning many tens of solar masses, and with comparably large error bars. This ignores the effect of clumping, which can further skew mass estimates by factors of several.

Jerkstrand et al. (2017) show how the line luminosity varies as a function of mass, energy deposition, and clumping and argue that luminosity comparable to SN 2015bn, SN 2007bi, and LSQ14an requires at least $\sim 10 M_\odot$ of ejecta (assuming a
carbon-burning composition). One solid conclusion we can draw is that none of the SLSNe in our sample have \([O\,I]\) \(\lambda 6300\) luminosities that significantly exceed these three events (Figure 17). Thus, the ejecta masses estimated for these slowly evolving SLSNe (see also PTF12dam) likely represent the upper end of the SLSN mass distribution. This agrees with results from light-curve modeling (Nicholl et al. 2017c).

6.4. \([Ca\,II]\) \(\lambda 7300/\,[O\,I]\) \(\lambda 6300\) Ratio and Core Mass

The ratio \([Ca\,II]\) \(\lambda 7300/\,[O\,I]\) \(\lambda 6300\) is sometimes used to estimate the helium core mass of the progenitor, with a larger luminosity in \([O\,I]\) \(\lambda 6300\) indicating a more massive core. However, this method has many caveats: the ratio depends sensitively on the distribution of calcium in the ejecta, the phase of observation, and the velocity of expansion, with higher velocities increasing \([Ca\,II]\) \(\lambda 7300\) relative to \([O\,I]\) \(\lambda 6300\) (Fransson & Chevalier 1989; Houck & Fransson 1996; Jerkstrand 2017).

While keeping the above in mind, it is still illustrative to compare our observed ratios to spectral models. In Figure 20 we plot the ratio from the stripped-envelope SN models of Jerkstrand et al. (2015). Those models were calculated for a velocity of 3500 km s\(^{-1}\), lower than that observed in SLSNe. In the models of Fransson & Chevalier (1989), \([Ca\,II]\) \(\lambda 7300/\,[O\,I]\) \(\lambda 6300\propto v\), so we use that scaling here, adopting a velocity range of 5000–8000 km s\(^{-1}\) appropriate for SLSNe.

Because mixing between zones can strongly affect the line ratios, the comparison is more reliable at times \(\gtrsim 300\) days, when the velocity profiles indicate that oxygen emission and calcium emission are coming from a similar region in the ejecta. This is reflected in the spread of observed ratios at early times, before the eventual convergence to \(\sim 0.5\). While absolute

---

**Figure 18.** Same as Figure 17, but with phases normalized by the light-curve decline timescale \(t_d\). The line luminosities for all SLSNe now cluster tightly around relatively well defined decay slopes.
values should not be taken literally, we find that the highest-mass model, with a 5.9 \( M_\odot \) helium core, gives the best match to most events, but the range in the observations and the simplicity of the model mean that there is at least a factor \( \approx 2 \) uncertainty in this estimate.

Plotting the ratios as a function of \( t/t_d \), SN 2017egm, LSQ14an, and PTF10hgi appear even more discrepant in their [Ca II] \( \lambda7300 \)/[O I] \( \lambda6300 \) ratios, compared to the other SLSNe and to the model predictions. This strengthens the conclusion that the \( \lambda7300 \) feature is dominated by [O II] in these events. We note that if the [Ca II] \( \lambda7300 \) line is contaminated by [O II] (but at a lower level) in other SLSNe, the true ratio of [Ca II] \( \lambda7300 \) to [O I] \( \lambda6300 \) must be even lower, which would favor larger core masses.

### 6.5. [Ca II] \( \lambda7300 \)/Ca II NIR and Electron Density

In Figure 21, we show the ratio between the strongest calcium lines, [Ca II] \( \lambda7300 \) and the Ca II NIR triplet. Jerkstrand et al. (2017) pointed out that the triplet was unusually strong in SN 2015bn relative to [Ca II] \( \lambda7300 \), yet the lines do appear to originate in the same velocity zone (Nicholl et al. 2016a). If the \( \lambda7300 \) line has a significant component from [O II], the true ratio of the NIR triplet to...
[Ca II] λ7300 is even lower and thus more discrepant with normal SNe Ic.

For our sample, we measure typical values of [Ca II] λ7300/ Ca II NIR ∼ 0.5. Comparing to the detailed models from Jerkstrand et al. (2017), this indicates an electron density $n_e \gtrsim 10^6$ cm$^{-3}$ and a temperature >4000 K. Jerkstrand et al. (2017) also showed that to attain this electron density with a reasonable total mass for the emitting zone, the ejecta must be significantly clumped, i.e., the filling factor is $f \ll 1$.

Our velocity analysis suggested that O I λ7774 emission arises from deep in the ejecta, overlapping the calcium-rich zones. O I λ7774 is a recombination line and can therefore provide an independent check on the density and filling factor: $L_{\lambda 7774} \propto n_e^2 f$. However, this line can remain optically thick for hundreds of days, and thus at earlier epochs there may be a significant contribution from scattering (Jerkstrand et al. 2015).

Using recombination coefficients from Jerkstrand et al. (2015) and Maeder & Mazzali (2010) and assuming a velocity of 6000 km s$^{-1}$, we show in Figure 22 that the O I λ7774 luminosity at 300–500 days is consistent with $10^{14} \lesssim n_e^2 f \lesssim 10^{16}$ cm$^{-6}$. For modest filling factors $f < 0.1$, this translates to a density $n_e > 10^7$ cm$^{-3}$. Or fixing $n_e > 10^6$ cm$^{-3}$ based on the calcium lines, we have $f < 1$, i.e., the observed lines cannot be reproduced without some degree of clumping.

We can verify this with a simple calculation: for a fiducial $5 M_\odot$ expanding at 6000 km s$^{-1}$ for 400 days, the number density of ions is $\sim 10^7$ cm$^{-3}$ (assuming an oxygen-dominated composition). This is the maximum possible electron density for unclumped ejecta, if the ejecta are singly ionized on average. Thus, to achieve $n_e > 10^6$ cm$^{-3}$ requires $f < 0.1$ (i.e., the volume occupied by the ejecta is only 10% of the available expansion volume). If the oxygen is instead mostly neutral, as may well be the case based on the dominant [O I] λ6300 line, the situation requires more extreme clumping.

Mg I λ4571 is also a recombination line and may be a useful indicator of the electron density at a higher velocity coordinate; however, this line also contributes significant cooling, and thus detailed spectral models are needed to disentangle these two components (Jerkstrand et al. 2017).

7. **Discussion**

Here we combine the results from the previous sections to piece together a plausible physical picture of SLSNe evolving through the nebular phase. A schematic diagram is shown in Figure 23 and elaborated on below.

### 7.1. Density and Ionization Structure as a Clue to the Power Source

The two strongest optical lines, [O I] λ6300 and [Ca II] λ7300, have similar critical densities, $n_e \sim 10^6$ cm$^{-3}$ (e.g., Li & McCray 1993; Fesen et al. 1999), so it is interesting that the [Ca II] λ7300 line often develops earlier. One way to suppress [O I] λ6300 emission would be if oxygen was mainly concentrated in a dense region where collisional de-excitation is dominant, but we know that this is not the case since we see [O I] λ6300 at higher velocities (where relative density should be
lowest) compared to other lines, including $[\text{Ca II}]\,\lambda 7300$ (Figure 16).

Another possibility is that most of the oxygen is initially ionized, and $[\text{O I}]\,\lambda 6300$ develops as the neutral fraction becomes substantial. This interpretation is also problematic, since the ionization potential of O I is greater than that of Ca II, so this scenario $[\text{Ca II}]\,\lambda 7300$ might also be suppressed. Moreover, for the long-lived power input needed to sustain SLSN luminosities, Margalit et al. (2018) recently found that the ionization structure remains frozen into the ejecta, such that a dramatic change in the neutral fraction of oxygen is unlikely on the timescales probed here.

The absolute flux in the $[\text{O I}]\,\lambda 6300$ line also argues against the above scenarios, since the line luminosity requires a large neutral oxygen mass and decreases smoothly in time even as the line becomes strongly peaked. This suggests that $[\text{Ca II}]\,\lambda 7300$ is enhanced, rather than $[\text{O I}]\,\lambda 6300$ being suppressed. Indeed, it has long been known that some SLSNe exhibit strong $[\text{Ca II}]\,\lambda 7300$ emission even before the onset of the nebular phase (Gal-Yam et al. 2009).

In Figure 18, we showed that the evolution in line luminosities is strongly coupled to the decline timescale of the light curve, suggesting that the lines are reprocessing the input from a continuous and decaying power source. Calcium resides at a relatively low velocity coordinate, likely in the Si/S and O/Si/S zones from comparison to explosion models, whereas oxygen is abundant in the O/Si/S zone and the outer O/Ne/Mg and O/C zones. Therefore, the luminous $[\text{Ca II}]\,\lambda 7300$ line appearing early in the evolution can be explained if a large amount of energy is being reprocessed by the inner regions of the ejecta. This is a natural consequence of a central engine power source, the most popular model for interpreting SLSN light curves.

Furthermore, 2D simulations of SLSNe show that a central engine can hollow out the interior of the ejecta and create a complex density structure due to instabilities in the compressed fluid (Chen et al. 2016). This could potentially provide the low filling factor required by certain line ratios, and rarified interclump regions can more quickly reach electron densities below the critical density for $[\text{Ca II}]\,\lambda 7300$—effectively, allowing parts of the innermost ejecta to become “nebular” before the outer layers.

A strong $[\text{Ca II}]\,\lambda 7300$ line during the early or pseudo-nebular phase may therefore be a signature of a central engine. However, this does raise a further question of how emission from a low velocity coordinate can be observed while the SN still has a photospheric component (i.e., continuum). Jerkstrand et al. (2015) show wavelength-dependent optical depths and photon escape probabilities for one of their stripped SN models, covering all the ejecta zones of interest here. While not tuned specifically for SLSNe, their results demonstrate that for reasonable parameters photons with $\lambda \gtrsim 6000$ Å can easily

Figure 23. Cartoon showing the structure of an SLSN in the nebular phase. Zone velocities and the regions responsible for key emission lines are labeled. $[\text{O I}]\,\lambda 6300$ emission comes from progressively deeper layers over time. The inner zones have a low filling factor due to an engine-blown bubble and resultant clumping by fluid instabilities—this region likely produces the $[\text{Ca II}]\,\lambda 7300$ and $[\text{O I}]\,\lambda 7774$ lines that are prominent in SLSNe, while an engine-driven oxygen ionization front, frozen in mass coordinate, allows production of $[\text{O II}]$ and $[\text{O III}]$. If the ionized region extends very far out, the spectrum can be dominated by $[\text{O II}]$, as seems to be the case in SN 2017gme, LSQ14an, and PTF10hgi.
traverse the ejecta by 200 days, even while the optical depth remains high at bluer wavelengths. Future radiative transfer calculations should help to test this hypothesis.

In cases where [Ca II] λ7300 turns on even earlier, such as SN 2007bi, SN 2015bn, and LSQ14an (Gal-Yam et al. 2009; Nicholl et al. 2016b; Inserra et al. 2017), additional effects may be required. For example, a nonspherical geometry could also serve to produce paths through the ejecta with a wide range of optical depth. The line profiles of LSQ14an support a significantly aspherical explosion, while (spectro)polarimetry of SN 2015bn has also revealed asymmetry (Inserra et al. 2016; Leloudas et al. 2017). However, the line profiles for most SLSNe suggest that asymmetry is probably modest.

O I λ7774 emission arises at a similar velocity coordinate to [Ca II] λ7300. This is also readily interpretable in the central engine scenario, as the engine can both compress and ionize the inner parts of the ejecta, creating high-density ionized regions that are favorable for oxygen recombination. Margalit et al. (2018) showed that for an $L \propto r^{-2}$ magnetar engine, the ionization front remains at roughly constant mass coordinate as the ejecta expand, consistent with our finding that the O I λ7774 recombination line comes predominantly from the same velocity zone at all times.

This is in contrast to the [O I] λ6300 emission, which clearly comes from progressively deeper-lying regions in the ejecta over time. We suggest that this emission initially comes from outer ejecta that is relatively undisturbed by the central engine overpressure and that the gradual movement to a lower velocity coordinate is simply a result of the density decreasing under free expansion. If the O I ionization front is located in the O/Ne/S layer, [O I] λ6300 eventually reaches a comparable velocity to O I λ7774, as we observe in some of our latest spectra.

The presence of this ionization front can have other effects on the spectra of SLSNe. In particular, the unusually strong [O III] emission in SLSNe compared to SNe Ic can be explained by persistent oxygen ionization by the central engine. We can also account for some of the diversity among SLSNe by appealing to the location of the ionization fronts. If some engines succeed in ionizing oxygen all the way through the ejecta, this would allow for suppression of [O I] λ6300 throughout the nebular phase and provide a source of [O II] to enhance the λ7300 line. Since Ca II should also be significantly ionized in this scenario, the 7300 Å line in LSQ14an, SN 2017egm, and PTF10hgi is likely dominated by [O II].

Why does this occur only in a few SLSNe? Jerkstrand et al. (2017) showed that in low-mass models of SLSNe a process of “runaway ionization” can occur when a large amount of energy is deposited in a small ejecta mass. Nicholl et al. (2017b) suggested that SN 2017egm may be among the lowest-mass SLSNe, making it a candidate for this process. LSQ14an, on the other hand, is more likely an SLSN with a large ejecta mass—in this case, the asymmetry suggested by the line profiles may result in lower-density regions that are easier to ionize. PTF10hgi is one of the least luminous SLSNe, so it may be an example of a low-mass event, though Quimby et al. (2018) detected helium in the photospheric spectrum, so envelope stripping is not as extreme as SN 2017egm. In both LSQ14an and PTF10hgi, strong [O III] emission may be an indication of low density.

While this narrative seems consistent with the evolution in velocity and luminosity for the key lines, as well as the diversity in ionization state between events, a central engine may have more difficulty explaining the shape of the emission-line profiles. As discussed in Section 5, a spatial distribution of ions with a hollow cavity is predicted analytically to emit flat-topped lines, whereas most of the lines we see are close to Gaussian in appearance. Velocity and density gradients, mixing, or ejecta asymmetries could help to mitigate this problem. However, solving it fully will require radiative transfer calculations for ejecta with a central cavity, to compare with the observations.

The simple schematic picture sketched out here should be used as a starting point for detailed modeling of SLSN nebular spectra. Jerkstrand et al. (2017) have produced model spectra for pure oxygen ejecta and carbon-burning compositions, and their results have informed much of the analysis here. To model all of the ejecta zones simultaneously is a complex computational feat, but we hope that our results can inform plausible choices for the velocities, densities, filling factors, and energy deposition in future calculations of synthetic spectra.

We conclude this subsection by noting that SLSNe powered by circumstellar interaction could also produce dense regions (a cold dense shell and resultant instabilities) and ionizing photons (from the shock fronts) to potentially reproduce the line ratios we see here. However, such a model may not so naturally account for the velocity structure we observe.

7.2. Comments on Progenitors

Since nebular spectroscopy probes the innermost layers of the explosion, this technique provides the most direct fingerprints of the pre-SN progenitor star. In Section 6, we showed that for many events the [Ca II] λ7300/[O I] λ6300 ratio was reasonably well matched by helium cores of $\sim 3.5 - 5.9 M_\odot$, though this ignored various complicating factors in deriving theoretical ratios from explosion models. On the observational side, we have found that [O II] emission can have a large effect on the 7300 Å luminosity, such that the true core mass may often be $> 6 M_\odot$.

The light-curve models of SLSNe from Nicholl et al. (2017c) suggested an average ejecta mass of $4.8 M_\odot$—if we assume that the helium core mass is equal to the ejecta mass plus a typical neutron star mass of $1.4 M_\odot$, this suggests a helium core of $\sim 6.2 M_\odot$, in broad agreement with our findings here. However, we note that the distribution of ejecta masses inferred from light-curve modeling is quite broad, including several events with masses of $2 - 3 M_\odot$. These estimates are complicated by the use of an uncertain gray opacity (Arnett 1982), which may lead to systematic offsets between masses inferred from light curves and spectroscopy.

There is also likely a selection effect favoring nebular spectroscopy of high-mass events: since slower-fading SLSNe tend to be more luminous at peak (Inserra & Smartt 2014), they are easier to observe at nebular times even when corrected for the longer $t_p$. Nebular follow-up of larger samples of SLSNe, spanning a range of decline rates, will be required to determine whether the overall mass range is as broad as that implied by light-curve models.

Many events have [O I] λ6300 luminosities that exceed $10^{41}$ erg s$^{-1}$ for long periods of time. Jerkstrand et al. (2017) showed that in their models this required an oxygen mass of $\sim 10 M_\odot$, or even greater depending on the energy deposition. In our measurements, SN 2017egm, PTF10hgi, and PTF12hni
have significantly lower [O I] $\lambda 6300$ luminosities and therefore may be examples of SLSNe from lower-mass progenitors, though ionization of oxygen also likely plays a role in setting the line luminosity, as we discussed above.

Another important clue for diagnosing progenitor mass is the abundance of iron-group elements. In Section 3.2, we found that SLSNe have an enhanced flux around $\sim 5000$ Å compared to typical SNe Ic, indicating a larger mass of iron-group elements. In fact, this region of the spectrum looks similar to broad-lined SNe Ic such as SN 1998bw, as previously noted for specific SLSNe by Nicholl et al. (2016a) and Jerkstrand et al. (2017).

A larger iron mass is generally linked to a more massive progenitor, as very massive stars are expected to synthesize more heavy elements in explosive burning (e.g., Umeda & Nomoto 2008). Thus, SLSNe and broad-lined SNe Ic appear to come from more massive stripped stars than do normal SNe Ic. This is also supported by analytic light-curve fits, which have found larger ejecta masses in broad-lined SNe Ic and SLSNe than in SNe Ic (e.g., Drout et al. 2011; Cano 2013; Nicholl et al. 2015, 2017c; Taddia et al. 2015). Differences in initial metallicity of the progenitor cannot account for the increased metal abundance in SLSNe and broad-lined SNe Ic, since these explosions prefer lower-metallicity environments than do normal SNe Ic.

However, our PCA showed that the iron plateau exhibits more variation between SLSNe than any other spectral feature, which may reflect a diversity in iron-group abundance between SLSNe (though temporal effects are also responsible for much of the variation). Therefore, even if SLSNe could result from stars that are on average of a similar mass to the progenitor of SN 1998bw, or at least produce a comparable iron mass, there is likely significant scatter.

Since SLSNe seem to come from relatively massive stripped-envelope stars, one important question is how they lose their hydrogen layers prior to explosion. Possibilities include line-driven winds, eruptive outbursts, interaction with a binary companion, or chemically homogeneous evolution. A useful observational handle is the distance to any hydrogen-rich material, presumably lost by the star prior to explosion.

Yan et al. (2017a) determined that the few SLSNe that do have late-time hydrogen signatures in their spectra typically reach this material at radii of $\sim 10^{16}$ cm. For our sample, we estimate the radii of emitting material for each SLSN at each epoch simply by multiplying observed line velocities by the phase from peak plus a rise time for each SLSN (Gal-Yam et al. 2009; Nicholl et al. 2013, 2016b, 2017a; De Cia et al. 2017; Inserra et al. 2017; Bose et al. 2018). We use velocities for both [O I] $\lambda 6300$ (the fastest material) and [Ca II] $\lambda 7300$ (for a more conservative estimate).

The results are shown in Figure 24. Most of the SLSNe in our sample reach radii of at least a few $\times 10^{16}$ cm over the timescales of the observations, comparable to the radii where hydrogen was detected by Yan et al. (2017a). Therefore, the fact that our spectra do not show hydrogen is not because the radii probed are smaller (they are not), but presumably because a fairly low fraction of SLSN progenitors, $\sim 3/15 = 20\%$, have significant hydrogen-rich material within $\sim 10^{16}$ cm. If hydrogen is ejected explosively, e.g., by pulsational pair-instability eruptions (Woosley et al. 2007; Woosley 2017), a typical velocity is expected to be $\sim 1000$ km s$^{-1}$. Thus, to reach at least a few $\times 10^{16}$ cm, the eruption must have occurred at least about a decade prior to explosion. In the case of a wind with velocity 10–100 km s$^{-1}$, the hydrogen envelope must be completely lost by 100–1000 yr before explosion (roughly the end of helium burning).

To summarize this section, SLSN progenitors are likely massive stars with helium cores $> 6 M_{\odot}$—and often $10 M_{\odot}$ or more—to produce enough oxygen for the observed [O I] $\lambda 6300$ luminosity (Jerkstrand et al. 2017). If the nebular-phase luminosity between $\sim 4000$ and $5000$ Å is a good indication of the total iron-group mass, itself an indication of progenitor mass, SLSN progenitors appear to be on average similar to that of SN 1998bw, though with considerable scatter. Envelope loss is complete at least a decade before explosion, and possibly much earlier.

8. Conclusions

We have conducted a systematic study of the observed properties of SLSN spectra as they evolve through the nebular phase. Our sample comprised 41 spectra of 12 SLSNe, including both fast- and slow-evolving events.

After applying a consistent interpolation and smoothing procedure to all spectra and normalizing the observed phase by the different decay rates of the SLSN light curves, we found that all events could be reasonably well described in terms of a single spectral sequence, ordered by this normalized phase. Our GMOS spectra of Gaia16apd (399 days after peak) and SN 2017egm (353 days after peak) are among the latest obtained for SLSNe, especially when considering the relatively fast light-curve timescales of these events.

We analyzed these spectra in terms of their statistical properties and the velocity and luminosity evolution of specific lines. For convenience, we summarize our main conclusions as follows:

1. SLSN spectra become dominated by nebular features within two to four $e$-folding times after their light-curve peaks. This provides a means early in the evolution of an
individual event to plan the optimal time for nebular-phase follow-up.
2. The main emission lines are easily identified with well-known transitions of oxygen, calcium, magnesium, sodium, and iron—the same species typically seen in normal and broad-lined SNe Ic.
3. [O I] λ6300 is initially weaker than [Ca II] λ7300 and takes longer to develop a Gaussian line profile, but it usually becomes the strongest optical line at later times. The ratio of the λ6300/λ6364 components indicates that this line is optically thin during most of the nebular phase.
4. Compared to SNe Ic, SLSNe are differentiated by a prominent O I λ7774 recombination line, often along with [O II] and [O III], indicating higher ionization, and the presence of oxygen in regions with significant variation in electron density.
5. SLSNe also show elevated flux compared to normal SNe Ic, on average, over the iron-dominated part of the spectrum between 4000 and 5500 Å. However, the SLSN iron region is similar to some broad-lined SNe Ic such as SN 1998bw.
6. PCA showed that ≥70% of the variance in SLSN spectra could be attributed to five eigenspectra, corresponding roughly to emission from Fe II, the Ca II NIR triplet, [O I] λ6300, [Ca II] λ7300, and O I λ7774, in order of decreasing variance.
7. We find no compelling evidence for clustering of SLSNe into subpopulations based on their nebular spectra. A K-means clustering analysis with two assumed clusters separates the spectra as much by phase relative to explosion as by the actual SLSNe to which these spectra belong.
8. Most SLSNe show no strong asymmetry in their [O I] λ6300, [Ca II] λ7300, or Mg I λ4571 line profiles; LSQ14an is an exception to this. However, at least half of the SLSNe in our sample exhibit an excess on the red side of O I λ7774, likely attributable to Mg II λλ7877, 7896.
9. The ejecta structure and composition inferred from the widths of the strongest lines are consistent with explosion models of massive stars, with calcium toward the center and magnesium farther out, and oxygen spanning a wide range of ejecta zones. [O I] λ6300 arises from faster (further out) regions of the ejecta than O I λ7774 for several hundreds of days.
10. The luminosity evolution in all lines decreases with time. After normalizing the spectral phase by the light-curve decline timescale, all SLSNe show virtually the same decline rate in their line luminosities, indicating that the nebular lines are reprocessing the same power source responsible for the earlier luminosity evolution of SLSNe. The ratios between many lines look similar to SNe Ic, provided that one takes into account the relatively slower light-curve evolution of SLSNe.
11. The [Ca II] λ7300/[O I] λ6300 ratio matches some models with masses of ~3.5–5.9 M☉, but the measured ratio may be contaminated by [O II] λλ7320, 7330 emission, such that the true core mass is likely even higher in some events. However, other events appear to be indicative of a population extending to lower masses, and selection effects may account for a large number of high-mass events in the nebular sample.
12. The O I λ7774 line and the ratio between calcium lines indicate a region of high electron density with low filling factor at a low velocity coordinate. We suggest that this could be due to the hydrodynamic impact of a central engine.
13. The ionization structure shows no evidence of changing over several hundred days. This is expected in the engine-powered models recently calculated by Margalit et al. (2018). In some SLSNe, oxygen may be ionized in a large fraction of the ejecta throughout the nebular phase—these events are SN 2017egm, LSQ14an, and PTF10hgi.
14. The radii reached by the ejecta on the timescales of these data are comparable to the radii where some SLSNe encounter hydrogen-rich material (Yan et al. 2015, 2017a). That our spectra show no sign of hydrogen indicates a diversity in when the hydrogen layer is lost by the progenitors.

There remain important outstanding questions, including an accurate calibration of the ejecta masses for better comparison with light-curve models and determining the processes in the pre-SN stellar evolution that expel the stellar envelope and likely allow the formation of a central engine. It also remains to be demonstrated whether our suggested ejecta distribution is consistent with the observed smooth line profiles, and whether [Ca II] λ7300 emission from deep-lying zones can escape the ejecta early in the transition to the nebular phase. Thus, a full understanding of SLSN progenitors will require more detailed modeling of their late-time spectra. Our analysis here provides an observationally motivated starting point for exploring the model parameter space.

We thank an anonymous referee for many insightful comments that improved the paper. M.N. is supported by a Royal Astronomical Society Research Fellowship. We thank Ashley Villar for comments on the manuscript, Daniel Fabricant and Igor Chilingarian for help with Binospec observations, and Ting-Wan Chen for providing a spectrum of PTF12dam. The Berger Time-Domain Group at Harvard is supported in part by the NSF under grant AST-1714498 and by NASA under grant NNX15AES0G. This paper is based on work supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE144152. R.C. acknowledges support from NASA Chandra grant award no. GO7-18046B. Based on observations (Proposal IDs GN-2017A-FT-15 and GN-2018A-FT-109) obtained at the Gemini Observatory acquired through the Gemini Observatory Archive and processed using the Gemini IRAF package, which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil). This paper uses data products produced by the OIR Telescope Data Center, supported by the Smithsonian Astrophysical Observatory. Some observations reported here were obtained at the MMT Observatory, a joint facility of the Smithsonian Institution and the University of Arizona. This paper includes data gathered...
with the 6.5 m Magellan Telescopes located at Las Campanas Observatory, Chile. This work is based in part on observations obtained at the MDM Observatory, operated by Dartmouth College, Columbia University, Ohio State University, Ohio University, and the University of Michigan. We thank J. Rupert for obtaining the MDM data. STSDAS and PyRAF are products of the Space Telescope Science Institute, which is operated by AURA for NASA.

Software: scikit-learn (Pedregosa et al. 2012), scipy (Jones et al. 2001), matplotlib (Hunter 2007), PyRAF, IRAF (Tody 1993).

Appendix A
Evaluating \( k \) in K-means Clustering

The \( K \)-means clustering algorithm requires that the number of clusters, \( k \), in which the data should be partitioned must be specified in advance. Therefore, it is essential to use some metric to evaluate whether these \( k \) clusters are separated in a meaningful way, and ultimately to determine the optimal number of clusters. Three such methods are described here and shown in Figure 25.

First, we evaluate the BIC (Schwarz et al. 1978) for different values of \( k \). This behaves similarly to the Bayes factor (the likelihood ratio between models, if neither model is favored a priori), but the BIC is not normalized. It strongly penalizes additional model complexity (i.e., more clusters), such that a lower BIC score is better (Kass & Raftery 1995). We find that the BIC decreases quite smoothly as \( k \) increases, with no local minima indicating a strongly preferred number of clusters, though there is a slight inflection at \( k = 3 \). We plot the BIC only up to \( k = 5 \), as a much larger number of clusters would not be meaningful in a data set of this size.

We next employ the elbow criterion. This method looks at the variance in the data explained as a function of \( k \), parameterized as the sum of squared residuals from the mean within each cluster. A flattening in this function indicates a preferred value of \( k \). The function plotted in Figure 25 does not flatten up to \( k = 5 \), leaving no strong incentive for choosing any particular \( k \).

Finally, we conduct a silhouette analysis using SKLEARN.METRICS. This analysis calculates a score for each data point based on how close it lies to the mean of the other clusters. The silhouette score ranges from \(-1\) to 1, with 1 indicating that the clusters are well separated. A score of 0 indicates that a point is on the boundary between clusters. For \( k > 3 \), we find that some clusters have low membership and score negatively; thus, a larger number of clusters is ruled out. We show silhouette plots for \( k = 2 \) and \( k = 3 \) in Figure 25. In both cases, the average scores are low, \( \approx 0.2 \). With \( k = 3 \), we find one large cluster and two rather smaller clusters, whereas for \( k = 2 \) the clusters are approximately equal in size.

Overall, there is no strong evidence for \( k > 1 \) clusters in the data.
Appendix B

Line Measurements

Velocities were measured by Gaussian fits. For some of spectra in the sample, particularly at earlier times, the [O I] λ6300 line is not well represented by a Gaussian. These are not included in velocity plots for [O I]λ6300. Specifically, we exclude SN 2015bn at 106 days, SN 2017egm at 126 days, PTF12dam at 171–221 days, PTF09cnd at 121 days, and LSQ14an and PTF10hgi (at all phases). An example is shown in Figure 26.

In Section 5.1 we described the process by which we approximated the continuum level for each spectral line before making measurements. Figure 27 shows an example of this.
Figure 26. Example of Gaussian fits to the [O i] λ6300 line. The left panel shows PTF12dam at 221 days. At this phase the line has not yet developed a Gaussian profile, and we exclude it from the velocity analysis. The right panel shows the same SLSN at 324 days. At later phases, [O i] λ6300 can be well fit with a Gaussian.

Figure 27. Example of continuum subtraction, for SN 2015bn at 392 days. Regions marked in red were used to fit a linear continuum to each line. Blue (orange) lines correspond to the spectrum before (after) continuum subtraction.
As the continuum placement can be uncertain for blended lines such as Mg I $\lambda 4571$ and O I $\lambda 5577$, we include an additional 20% systematic error in all line measurements to account for this.

Figure 28. Unsmoothed spectra of SLSNe. The only processing here is removal of host galaxy light and correction to rest frame.

Figure 29. Comparison sample of normal and broad-lined SNe Ic.

Appendix C
Additional Plots

Most of the spectra shown in this paper had been smoothed using the process described in Section 2.5. In Figure 28, we
plot the original unsmoothed spectra for reference. The analysis also made use of a comparison sample of SNe Ic, obtained from the Open Supernova Catalog. These spectra are shown in Figure 29.

In Sections 5 and 6, we plotted selected velocities and line ratios against the phase from maximum light, with and without phase normalization. For completeness, we here show how all velocities (Figures 30 and 31) and ratios (Figures 32 and 33) in our sample evolve as a function of both rest-frame phase and the normalized phase \( t/t_d \), where \( t_d \) is the exponential decline timescale of the light curve.

Figure 30. All velocity measurements plotted against phase from maximum light.
Figure 31. Same as Figure 30, but plotted in terms of normalized phase, $t/t_d$. 
Figure 32. Ratios of selected spectral lines to [O I] $\lambda 6300$. 
Figure 33. Same as Figure 32, but plotted in terms of normalized phase, $t/t_d$. 

ORCID iDs

Matt Nicholl @ https://orcid.org/0000-0002-2555-3192
Edo Berger @ https://orcid.org/0000-0002-9392-9681
Peter K. Blanchard @ https://orcid.org/0000-0003-0526-2248
Sebastian Gomez @ https://orcid.org/0000-0001-6395-6702

References

Anderson, J., Pessi, P., Dessart, L., et al. 2018, arXiv:1806.10609
Angus, C. R., Levan, A. J., Perley, D. A., et al. 2016, MNRAS, 458, 84
Arnett, W. D. 1982, ApJ, 253, 785
Axelrod, T. S. 1980, PhD thesis, California Univ.
Blanchard, P. K., Nicholl, M., Berger, E., et al. 2018, ApJ, 865, 9
Bose, S., Dong, S., Pastorello, A., et al. 2018, ApJ, 853, 57
Cano, Z. 2013, MNRAS, 434, 1098
Chatzopoulos, E., Wheeler, J. C., Vinko, J., Horvath, Z., & Nagy, A. 2013, ApJ, 773, 76
Chen, K.-J., Woosley, S. E., & Sukhbold, T. 2016, ApJ, 832, 73
Chen, T.-W., Nicholl, M., Smartt, S., et al. 2017, A&A, 602, A9
Chen, T.-W., Schady, P., Xiao, L., et al. 2017, ApJL, 849, L4
Chen, T.-W., Smartt, S. J., Jerkstrand, A., et al. 2015, MNRAS, 452, 1567
Chilingarian, I., Brown, W., Fabricant, D., et al. 2013, adass XXII, 475, 189
Chomiuk, L., Chornock, R., Soderberg, A., et al. 2011, ApJ, 743, 114
De Cia, A., Gal-Yam, A., Rubin, A., et al. 2018, ApJ, 860, 100
Dessart, L., Hillier, D. J., Waldman, R., Livne, E., & Blondin, S. 2012, MNRAS, 426, L76
Dressler, A., Bigelow, B., Hare, T., et al. 2011, PASP, 123, 288
Drout, M. R., Soderberg, A. M., Gal-Yam, A., et al. 2011, ApJ, 741, 97
Fesen, R. A., Gerardy, C. L., Filippenko, A. V., et al. 1999, AJ, 117, 725
Forgy, E. W. 1965, Biometrics, 21, 768
Fransson, C., & Chevalier, R. A. 1989, ApJ, 343, 323
Gal-Yam, A. 2012, Sci, 337, 927
Gal-Yam, A., Mazzali, P., Ofek, E., et al. 2009, Natur, 462, 624
Guillochon, J., Parrent, J., Kelly, L. Z., & Margutti, R. 2017, ApJ, 835, 64
Hook, I., Jørgensen, I., Allington-Smith, J., et al. 2004, PASP, 116, 425
Hotelling, H. 1933, Journal of Educational Psychology, 24, 417
Houck, J., & Fransson, C. 1996, ApJ, 456, 811
Howell, D., Kasen, D., Lidman, C., et al. 2013, ApJ, 779, 98
Huber, M., Chambers, K., Flewelling, H., et al. 2015, ATel, 7153, 1
Hunter, J. D. 2007, CSE, 9, 90
Inserra, C., Bulla, M., Sim, S., & Smartt, S. 2016, ApJ, 831, 79
Inserra, C., Nicholl, M., Chen, T.-W., et al. 2017, MNRAS, 468, 4642
Inserra, C., Prijs, S., Gutierrez, C., et al. 2018, ApJ, 854, 175
Inserra, C., Smartt, S., Jerkstrand, A., et al. 2013, ApJ, 770, 128
Inserra, C., & Smartt, S. J. 2014, ApJ, 796, 87
Izzo, L., Thöne, C. C., García-Benito, R., et al. 2018, A&A, 610, A11
