

New Insights into Classical Novae

Laura Chomiuk,¹ Brian D. Metzger,² and Ken J. Shen³

¹Center for Data Intensive and Time Domain Astronomy, Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA; email: chomiukl@msu.edu

²Department of Physics, Columbia University, New York, NY 10027, USA; email: bdm2129@columbia.edu

³Department of Astronomy and Theoretical Astrophysics Center, University of California, Berkeley, California 94720, USA; email: kenshen@astro.berkeley.edu

Annu. Rev. Astron. Astrophys. 2021. 59:391–444

First published as a Review in Advance on
June 18, 2021

The *Annual Review of Astronomy and Astrophysics* is
online at astro.annualreviews.org

<https://doi.org/10.1146/annurev-astro-112420-114502>

Copyright © 2021 by Annual Reviews.
All rights reserved

Keywords

thermonuclear explosions, white dwarf stars, cataclysmic variable stars, (radiative) shocks, relativistic particle acceleration, gamma rays

Abstract

We survey our understanding of classical novae—nonterminal, thermonuclear eruptions on the surfaces of white dwarfs in binary systems. The recent and unexpected discovery of GeV gamma rays from Galactic novae has highlighted the complexity of novae and their value as laboratories for studying shocks and particle acceleration. We review half a century of nova literature through this new lens, and conclude the following:

- The basics of the thermonuclear runaway theory of novae are confirmed by observations. The white dwarf sustains surface nuclear burning for some time after runaway, and until recently, it was commonly believed that radiation from this nuclear burning solely determines the nova’s bolometric luminosity.
- The processes by which novae eject material from the binary system remain poorly understood. Mass loss from novae is complex (sometimes fluctuating in rate, velocity, and morphology) and often prolonged in time over weeks, months, or years.
- The complexity of the mass ejection leads to gamma-ray-producing shocks internal to the nova ejecta. When gamma rays are detected (around optical maximum), the shocks are deeply embedded and the surrounding gas is very dense.

**ANNUAL
REVIEWS CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

- Observations of correlated optical and gamma-ray light curves confirm that the shocks are radiative and contribute significantly to the bolometric luminosity of novae. Novae are therefore the closest and most common interaction-powered transients.

Contents

1. INTRODUCTION	392
2. MASS LOSS IN NOVAE	395
2.1. Nuclear Burning in Novae	396
2.2. Mechanisms of Mass Ejection	400
2.3. Multiwavelength Observations of Mass Loss	403
3. SHOCK SIGNATURES ACROSS THE ELECTROMAGNETIC SPECTRUM	412
3.1. GeV Gamma Rays	413
3.2. X-Rays from Hot Plasma	419
3.3. Nonthermal Radio Emission	421
4. THEORY OF INTERNAL RADIATIVE SHOCKS IN CLASSICAL NOVAE	422
4.1. Dynamics of Dual Radiative Shocks	423
4.2. Relativistic Particle Acceleration	425
4.3. Radiative Signatures of Radiative Shocks	426
4.4. Multidimensional Effects and Dust Formation	427
5. BRINGING IT TOGETHER: IMPLICATIONS AND OPEN QUESTIONS	429
5.1. Long-Term Evolution of White Dwarf Binaries	429
5.2. Implications for Extragalactic Shock-Powered Transients	431
6. OPEN QUESTIONS, FUTURE WORK	432

1. INTRODUCTION

Classical and recurrent novae (derived from the Latin *stella nova* or *new star*) are luminous eruptions that take place in binary star systems in which a white dwarf (WD) accretes matter from a nondegenerate stellar companion (Gallagher & Starrfield 1978). As an accreted layer accumulates on the WD surface, the density and temperature at its base rise, leading to an increase in the nuclear burning rate. Under circumstances that depend sensitively on the WD mass and accretion rate, the layer undergoes unstable (runaway) nuclear burning once it reaches a critical mass (e.g., Starrfield et al. 1972, Prialnik & Kovetz 1995, Townsley & Bildsten 2004). The resulting energy release causes the accreted envelope to expand enormously, ultimately leading to its ejection, often along with heavier elements dredged up from deeper layers of the WD. Novae, with an estimated frequency of $\sim 20\text{--}70$ eruptions per year in our Galaxy (e.g., Darnley et al. 2006, Shafter 2017), are the second most common type of thermonuclear eruptions after Type I X-ray bursts from neutron stars (Galloway & Keek 2017). However, even with the advent of synoptic time-domain surveys, the discovery rate of Galactic novae¹ remains modest ($\sim 5\text{--}15 \text{ year}^{-1}$), probably due to gaps in optical monitoring and dust obscuration in the Galactic plane.

¹ See <https://asd.gsfc.nasa.gov/Koji.Mukai/novae/novae.html> for a compilation of recent Galactic novae.

Among the brightest transients in the night sky, novae were sometimes called guest stars (see Duerbeck 2008 for a historical perspective and Hoffmann et al. 2020 for recent work). Given their storied role in the history of astronomy, it is striking—despite substantial observational and theoretical progress—that our understanding of these common transients remains incomplete. Nothing highlights this better than the nearly universal² surprise that accompanied the discovery of GeV gamma-ray emission from novae by the Large Area Telescope (LAT) on NASA's *Fermi Gamma-Ray Space Telescope* (*Fermi*; Abdo et al. 2010, Ackermann et al. 2014). Evidence has long existed for complex mass-loss patterns and internal shocks within nova outflows, but the energetic importance of these shocks was not fully appreciated until recently.

In the standard paradigm, novae are almost exclusively driven by thermal emission from the hot WD. After the thermonuclear runaway (TNR), a shell of gas is expelled from the WD, expanding into the surrounding environment at hundreds to thousands of km s⁻¹. The thermal and kinetic energies of the ejecta, whether released in a short-lived episode or as a longer, continuous wind, are powered by radiation from nuclear burning on the surface of the WD. Novae are best known as optical transients with light curves that rise rapidly to maximum and decay over timescales of days to months (Payne-Gaposchkin 1957). However, as mass loss subsides and the ejecta dilute, they become increasingly transparent to radiation at shorter wavelengths and the spectral energy distribution of the nova shifts to the ultraviolet (UV; Gallagher & Code 1974). Eventually, the photosphere recedes far enough inward that the WD can be observed as a luminous supersoft X-ray source, sustained by residual nuclear burning for weeks to years (Kahabka & van den Heuvel 1997). The expanding ejecta are photoionized by the hot central WD, and produce thermal radio emission on timescales of years (Seaquist & Bode 2008). Many novae form dust in their ejecta, revealed by sudden rises in their infrared (IR) emission and sometimes by drops in their optical emission due to extinction along the line of sight (Gehrz 1988).

Although this basic picture remains largely intact, multiwavelength observations over the past decade have increasingly revealed a nonthermal, shock-powered side to novae, which is providing new insights into old mysteries about these events. The shocks occur either as multiple phases of ejecta collide during the eruption or as the ejecta crash into a preexisting medium surrounding the binary. The continuum gamma-ray ($\gtrsim 100$ MeV) emission observed by *Fermi*-LAT is clear evidence of the acceleration of relativistic particles by shocks (Martin & Dubus 2013). Shocks have also long been implied by X-ray observations of hot ($\sim 10^7$ – 10^8 K) presumably shock-heated gas in novae (O'Brien et al. 1994), observed weeks to months after eruption (Mukai et al. 2008). These X-rays might be absorbed earlier in the eruption and reprocessed into the ultraviolet–optical–infrared (UVOIR) bands, thus contributing to the optical light curve and its variability (Metzger et al. 2014, Li et al. 2017, Aydi et al. 2020b). Relativistic electrons accelerated at shocks also generate synchrotron radiation, sometimes seen as a distinct early component of radio emission (Taylor et al. 1987, Weston et al. 2016b, Finzell et al. 2018). Shock interaction may also play a crucial role in shaping the complex large-scale morphology of nova ejecta and may—via compression and hydrodynamical instabilities—generate the large densities and inhomogeneities needed for molecule and dust formation (Evans & Rawlings 2008, Derdzinski et al. 2017).

As rich and complex phenomena, novae are worthy of study in their own right. However, they also serve as bright, nearby probes of several physical processes, such as binary mass transfer, explosive nuclear burning, and radiative shocks, that are relevant to many other astrophysical

Large Area Telescope (LAT): instrument on the *Fermi Gamma-Ray Space Telescope* that monitors $\sim 60\%$ of the sky simultaneously at photon energies ~ 0.1 – 300 GeV

Supersoft X-ray sources: thermal X-ray sources powered by surface nuclear burning on white dwarfs with luminosities $L_X \approx 10^{36}$ – 10^{38} erg s⁻¹ and effective temperatures $T_{\text{eff}} \approx 10^5$ – 10^6 K

²However, evidence for relativistic particle acceleration was already present from radio synchrotron emission (e.g., Hjellming et al. 1986, Taylor et al. 1987, Rupen et al. 2001; Section 3.3) and through other, indirect inference (Tatischeff & Hernanz 2007).

Classical nova:

thermonuclear eruption from a white dwarf accreting hydrogen-rich material from a Roche lobe–overflowing main sequence or moderately evolved companion

systems. The conditions under which an accreting WD can gain mass, despite mass loss in novae, have implications for the viability of the single degenerate channel for thermonuclear supernovae (SNe; e.g., Type Ia or Type Ia_x; Nomoto 1982, Foley et al. 2013). The morphologies of nova ejecta—often described as bipolar outflow with an equatorial ring—resemble the outflows from other dynamical binary mass-transfer events, such as planetary nebulae (De Marco 2009) or those expected from the common-envelope phase of stellar mergers (Ivanova et al. 2013). A better understanding of novae could provide new insights into a wide variety of other—rarer and more distant—transient events, especially those that may be shock-powered, such as Type II_n (interacting) SNe, tidal disruption events, and stellar mergers.

Although we touch upon most aspects of the nova phenomenon in this work, a comprehensive review is not possible. Instead, motivated by areas of rapid observational development, we shape the discussion around two broad and connected themes: mass loss during novae (Section 2) and the newly appreciated role of shocks in shaping their electromagnetic emission (Sections 3 and 4). In Section 5, we discuss the implications for other WD phenomena and astrophysical transients and suggest strategies for answering critical open questions surrounding nova mass loss and shocks. Throughout the article, we point the reader to reviews that provide greater depth on specific topics. **Figure 1** shows a schematic timeline of physical processes and electromagnetic emission in novae that will be useful to refer to as we proceed.

Before diving in, we introduce a few common definitions. With one exception (the helium nova V445 Pup; see the sidebar titled Helium Novae), the donor stars in nova-hosting binaries are hydrogen rich. A classical nova is one in which the secondary is typically a main sequence star

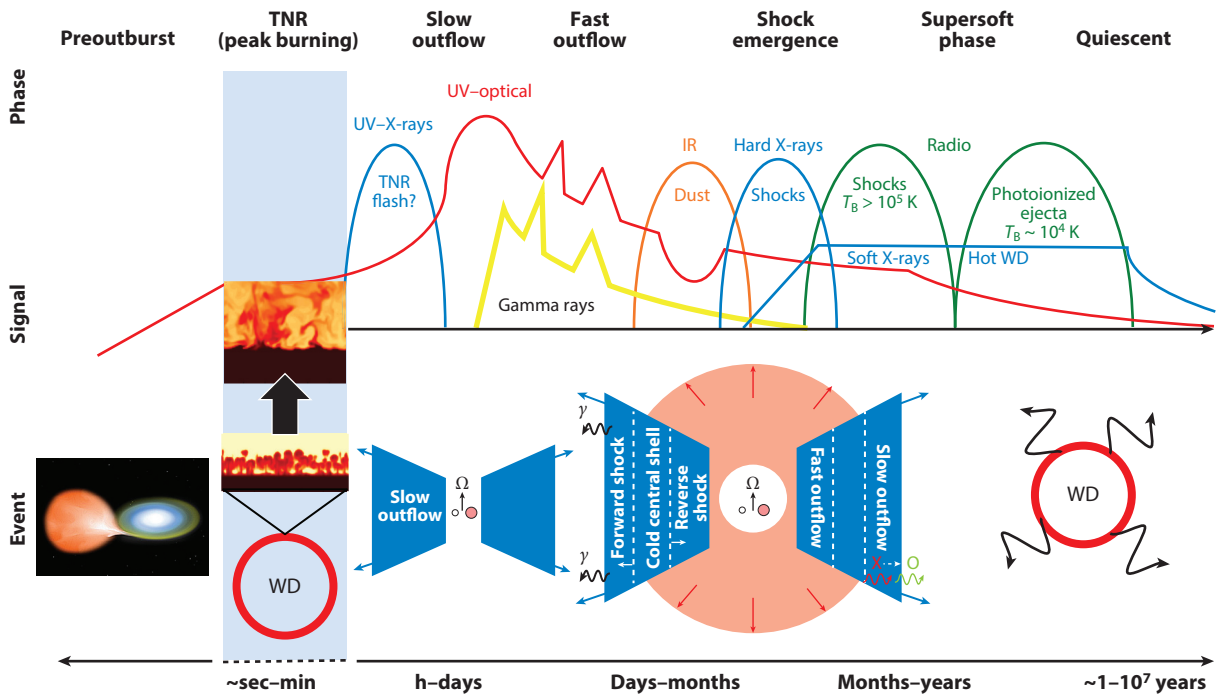


Figure 1

Schematic timeline of the physical processes and electromagnetic signals from novae. The figure includes modified images of convection and mixing during the TNR (figure adapted from Casanova et al. 2016; copyright ESO, with permission) and internal shocks (figure adapted from Metzger et al. 2015). Abbreviations: TNR, thermonuclear runaway; WD, white dwarf.

HELIUM NOVAE

The vast majority of classical and recurrent novae are the result of runaway hydrogen burning. However, qualitatively similar phenomena can occur on WDs because of unstable helium burning (Taam 1980, Shen & Bildsten 2009b). The best helium nova candidate is V445 Puppis (Ashok & Banerjee 2003). Early phases of its 2000 eruption showed notably hydrogen-deficient spectra containing prominent He, C, and Fe features (Iijima & Nakanishi 2008). Near-IR imaging of the ejecta showed a tightly collimated bipolar outflow and an equatorial dust disk (see Section 2.3.5; Woudt et al. 2009), similar to the ejecta geometry frequently observed in ordinary hydrogen novae (Section 2.3.5), but with higher ejecta speeds ($\sim 8,500 \text{ km s}^{-1}$). Controversy initially surrounded the interpretation of this event, because the nondetection of a postnova central object was taken as evidence for a destructive explosion (Goranskij et al. 2010). However, the dust around the eruption site has finally begun to clear, and a photometric period has recently been detected at $P \approx 1.8\text{--}3.7$ days, which is consistent with the orbital period of a He star transferring mass to a WD companion (D. Steeghs, private communication).

overflowing its Roche lobe; the host binaries are cataclysmic variables (CVs) with short-period orbits, $P \approx 1.4\text{--}10$ h (Warner 1995, Diaz & Bruch 1997; although moderately evolved donors are not uncommon, e.g., Darnley et al. 2012). CV mass transfer is thought to be relatively conservative, with at most a few percent of the transferred mass lost in outflows; searches for circumstellar material around CVs point to low-density surroundings (e.g., Froning 2005, Hoard et al. 2014). By contrast, an embedded nova is one in which the secondary is an evolved giant star, typically on a long orbit with period $P \gtrsim 100$ days (Mikolajewska 2010). They are described as embedded because interaction with the companion wind often shapes their observational signatures. Novae with red giant companions likely make up 20–40% of observed events (Williams et al. 2016). Symbiotic novae are the subset of embedded novae with eruptions that evolve very slowly, over decades or even centuries (Kenyon & Truran 1983).

The binary remains intact after a nova eruption, so all novae are expected to recur, with periods ranging from ~ 1 year to $\gtrsim 10^7$ years as required to accrete and accumulate a critical-mass envelope (e.g., Yaron et al. 2005). Recurrent novae are the subset of systems observed to undergo more than one eruption in recorded history, but are otherwise driven by the same physical processes as other novae. Hereafter, we use the word novae as a general term for all thermonuclear novae (classical and recurrent), unless additional specificity is warranted.

2. MASS LOSS IN NOVAE

Mass loss is key to driving the observational appearance of nova eruptions. The mass-loss rate and outflow velocity as functions of time—whether impulsive or in the form of a sustained outflow—control how the light curve and spectral energy distribution evolve. Stochastic or secular variability in these properties can generate internal shocks within the ejecta (O’Brien & Lloyd 1994, Mukai & Ishida 2001, Chomiuk et al. 2014a). Mass loss can carry angular momentum away from the binary, with implications for the long-term evolution of CVs and thermonuclear SN progenitors (Section 5.1). Despite decades of theoretical and observational work, the mechanisms giving rise to nova outflows, including the role of the binary companion, remain a topic of debate.

In addressing the issue of mass loss, we first describe the root cause of the eruption—the TNR (Section 2.1.1) and the subsequent steady-burning phase (Section 2.1.2), and observational tests of the basic theory. We then discuss several proposed mass-loss mechanisms in Section 2.2. In Section 2.3, we describe observations that probe mass loss in novae across the electromagnetic spectrum.

Cataclysmic variable (CV):

a mass-transferring binary system containing a white dwarf and main sequence secondary that is overflowing its Roche lobe

Embedded nova:

a nova where the white dwarf is fed material from a giant companion star, usually via wind accretion

Symbiotic novae: the subset of embedded novae that evolve slowly, over decades or even centuries

Recurrent nova:

a nova observed to undergo more than one thermonuclear eruption in recorded history

2.1. Nuclear Burning in Novae

Accretion rate (\dot{M}):
the rate of accretion
onto the white dwarf,
driven by mass transfer
from a binary
companion

The outcome of hydrogen-rich mass transfer depends on several properties of the accreting WD including its mass (M_{WD}), accretion rate (\dot{M}), and core temperature and the composition of the accreted gas (e.g., Fujimoto 1982a, Starrfield et al. 2000, Townsley & Bildsten 2004, Yaron et al. 2005, Nomoto et al. 2007, Shen & Bildsten 2009a, Chen et al. 2019). When significant mixing occurs between the WD and the accreted layer, the outcome also depends on the composition of the WD being carbon–oxygen (CO) or oxygen–neon (ONe). CO WDs are the expected remnants of stars with zero-age main sequence masses $M_{\text{ZAMS}} \lesssim (7\text{--}8) M_{\odot}$, whereas ONe WDs come from stars with $(7\text{--}8) M_{\odot} \lesssim M_{\text{ZAMS}} \lesssim (9\text{--}10) M_{\odot}$; these ranges are metallicity-dependent and theoretically uncertain (Doherty et al. 2015).

A nova eruption is due to a TNR, which is the unstable burning of hydrogen on the WD surface (Gallagher & Starrfield 1978). A TNR is the outcome of mass accretion at low rates, $\dot{M} \lesssim \dot{M}_{\text{stable}}$, where the threshold value for steady, thermally stable burning, $\dot{M}_{\text{stable}} \approx 4 \times (10^{-8} \text{--} 10^{-7}) M_{\odot} \text{ year}^{-1}$, is an increasing function of M_{WD} (**Figure 2**). For accretion rates just above \dot{M}_{stable} , hydrogen burns stably at the rate it is accreted, powering persistent soft X-ray emission (Fujimoto 1982b, Nomoto 1982, Nomoto et al. 2007, Shen & Bildsten 2007, Wolf et al. 2013). For still higher accretion rates ($\dot{M} \gtrsim 3 \dot{M}_{\text{stable}}$), the fusion rate cannot match the accretion rate,

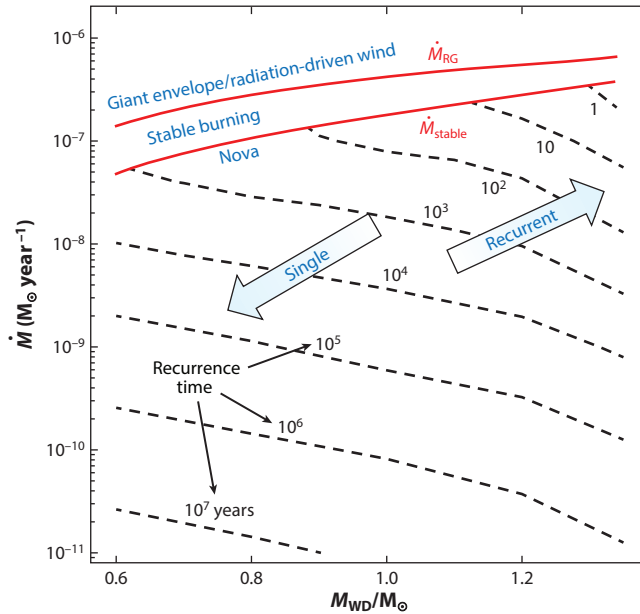


Figure 2

Depending on the WD mass and accretion rate (\dot{M}), a wide variety of phenomena are theoretically expected. The lower red line is \dot{M}_{stable} , the lowest accretion rate leading to stable burning for a given WD mass. At higher accretion rates, $\dot{M} > \dot{M}_{\text{RG}}$, hydrogen will still burn stably, but more slowly than the matter is accreted. The matter will pile up to form a red giant-like structure or it must be carried away in a wind. Below \dot{M}_{stable} , burning is unstable, resulting in nova eruptions. Dashed black lines are lines of constant nova recurrence time (labeled in units of years). Nova ejecta masses can be estimated as the product of recurrence time and \dot{M} . The stability lines are taken from Wolf et al. (2013). Recurrence timescales are calculated using the stellar evolution code MESA (v12115; Paxton et al. 2011) in a similar procedure as that described by Wolf et al. (2013), but with a central core temperature of 10^7 K and covering a larger range in accretion rates. Figure adapted with permission from Wolf et al. (2013); copyright AAS. Abbreviation: WD, white dwarf.

and the accreted matter piles up into an extended red giant-like structure or it is lost in a radiation-driven wind (Hachisu, Kato & Nomoto 1996). The CV hosts of classical novae typically have mass-transfer rates $\dot{M} \approx 10^{-10}$ – $10^{-8} M_{\odot} \text{ year}^{-1}$ (Patterson 1984), which place them in the unstable regime.

2.1.1. The thermonuclear runaway. Novae are challenging to model because myriad hydrodynamical effects such as convection, mixing, and instabilities combine with a variety of nuclear processes. However, the basic mechanism of the TNR can be understood through a comparison of several physical timescales:

- The accretion/recurrence timescale, $\tau_{\text{rec}} = M_{\text{acc}}/\dot{M}$, which depends on the accretion rate ($\dot{M} \approx 10^{-11}$ – $10^{-7} M_{\odot} \text{ year}^{-1}$) and the critical accreted mass for TNR ($M_{\text{acc}} \approx 10^{-7}$ – $10^{-3} M_{\odot}$). It is typically $\sim 10^4$ – 10^7 years, but it can be as short as 1 year.
- The nuclear timescale, $\tau_{\text{nuc}} = c_p T / \epsilon_{\text{nuc}}$, where T is temperature, c_p is the specific heat, and ϵ_{nuc} is the nuclear energy generation rate from hydrogen burning in the burning layer at the base of the accreted envelope. During peak burning in the TNR, τ_{nuc} can become as small as seconds if there is significant dredge-up from the underlying WD, which provides catalytic seed nuclei for CNO-cycle reactions.
- The convective overturn timescale, $\tau_{\text{eddy}} = H_p / v_{\text{conv}}$, which is the timescale over which eddies can rise over a pressure scale height, H_p , and redistribute the heat generated in the thin burning layer at the base of the accreted material. Convective velocities, v_{conv} , can reach hundreds of km s^{-1} at the peak of burning, implying a convective timescale of seconds, which is somewhat faster than the shortest nuclear timescales.
- The hydrodynamic timescale, $\tau_{\text{hd}} = H_p / c_s$, required for hydrostatic readjustment of the pressure scale height given the sound speed, c_s . At the peak of burning, τ_{hd} is a few tenths of a second and is the shortest relevant timescale.

During the accretion phase, $\tau_{\text{rec}} \ll \tau_{\text{nuc}}$ (for most novae, models predict proton–proton and ^3He reactions occurring over much of τ_{rec} , but at a low rate; Kovetz & Prialnik 1985, Townsley & Bildsten 2004, Shen & Bildsten 2009a). The envelope mass increases, along with the density and temperature at its base. After enough mass has been accreted, energy release from nuclear burning becomes too rapid to be transported by radiation and electron conduction, and the TNR phase begins with the onset of convection (Fujimoto 1982b). The TNR is a product of the thin-shell instability, brought on by the relatively small value of the specific thermal energy compared to the specific gravitational binding energy (i.e., the envelope cannot initially expand enough to quench the TNR). Although electron degenerate conditions can aid the onset of the TNR owing to the density’s lack of response to an increase in temperature, novae are ignited under ideal gas conditions as well. This occurs at higher accretion rates $\gtrsim 3 \times 10^{-9} M_{\odot} \text{ year}^{-1}$. An increase in the temperature under such conditions does lead to a stabilizing decrease in the density, but because the relevant nuclear burning rates scale much more strongly with temperature than density, the TNR can still occur.

The nuclear burning becomes very vigorous during the convective burning phase associated with the TNR, but the timescale ordering $\tau_{\text{hd}} < \tau_{\text{eddy}} < \tau_{\text{nuc}}$ is always preserved. No deflagration or detonation occurs, because convection is always able to efficiently redistribute the heat generated in the burning layer throughout the envelope. Eventually, the specific thermal energy in the envelope approaches the specific gravitational binding energy of the WD and the envelope expands, quenching runaway nuclear burning and causing a transition to a prolonged phase of steady burning (it is this transition that may result in the Early UV–X-ray Flash; see the sidebar titled Early UV–X-ray Flash from the Thermonuclear Runaway).

Recurrence timescale (τ_{rec}): recurrence time between novae, i.e., time required to build up, through accretion, the critical M_{acc} to initiate the thermonuclear runaway

Accreted mass (M_{acc}): the mass transferred to the white dwarf between novae

EARLY UV-X-RAY FLASH FROM THE THERMONUCLEAR RUNAWAY

In addition to the X-ray emission from the later stages of novae (Sections 2.1.2 and 3.2), a short-lived phase of UV/X-ray emission is predicted soon after the TNR, as the effective temperature T_{eff} rises prior to the envelope expansion (Hillman et al. 2014). The luminosity of this early UV/X-ray flash is close to the Eddington limit and can last from hours to days for typical parameters. Premaximum UV emission has been seen in some novae (e.g., Cao et al. 2012), but the early X-ray flash has yet to be unambiguously detected (Kato et al. 2016, Morii et al. 2016). This is unsurprising given its relatively short duration and occurrence prior to the visual outburst that generally triggers nova discoveries. Extant and future wide-field UV or X-ray monitors, such as MAXI (Negoro et al. 2016) and the *Einstein Probe* (Yuan et al. 2018), have a chance to detect this emission, which would confirm an important prediction of nova theory, and provide better constraints on the exact time of the TNR, the WD mass, and the amount of preexisting circumbinary material (e.g., via X-ray absorption).

Two types of nuclear timescales come into play during the TNR: those related to β decays (yielding timescale τ_β), and those related to proton capture reactions [$\tau_{(p,\gamma)}$]. As the density and temperature in the accreted envelope approach critical levels, nuclear reactions transition to the CNO cycle, which first operates in equilibrium [$\tau_\beta < \tau_{(p,\gamma)}$]. However, as the temperature at the base of the accreted envelope reaches $\gtrsim 10^8$ K, the timescales reverse [$\tau_\beta > \tau_{(p,\gamma)}$], and the CNO cycle is limited by the β -decay timescales (the hot CNO cycle). Convection transports the β -unstable nuclei to the outer cooler regions of the nova envelope, where they are preserved from destruction and available to decay at a later time (e.g., Starrfield et al. 2016). The resulting energy release drives expansion of the envelope and may contribute to mass ejection (Section 2.2.1).

The critical mass of the hydrogen layer needed to trigger the runaway spans a wide range, $M_{\text{acc}} \approx 10^{-7} - 10^{-3} M_\odot$, and is a decreasing function of both M_{WD} and \dot{M} (e.g., Yaron et al. 2005, Wolf et al. 2013). For a fixed M_{WD} , a higher value of \dot{M} implies higher temperatures at a given M_{acc} , so the ignition mass is smaller. For a fixed \dot{M} , a higher M_{WD} means higher densities and temperatures for a given value of M_{acc} , so the ignition mass is smaller.

The nova recurrence time, $\tau_{\text{rec}} = M_{\text{acc}}/\dot{M}$, is the time to build up the critical layer through accretion, and it also decreases with both M_{WD} and \dot{M} , as shown in **Figure 2**. Recurrence times can vary from ~ 1 year for rapidly accreting massive WDs approaching the Chandrasekhar limit ($M_{\text{Ch}} \approx 1.4 M_\odot$) to $\tau_{\text{rec}} \gtrsim 10^7$ years for slowly accreting, low-mass WDs ($M_{\text{WD}} \lesssim 0.8 M_\odot$). Because shorter recurrence times result in more frequent novae—and, thus, a higher detection rate in surveys—the observed nova sample is biased to higher WD masses and higher accretion rates (Truran & Livio 1986, Ritter et al. 1991, Iben et al. 1992a).

Observations of novae generally affirm these theoretical predictions. Recurrent novae ($\tau_{\text{rec}} \lesssim 100$ years; Schaefer 2010) tend to have small ejecta masses and occur on high-mass WDs (e.g., Diaz et al. 2010, Osborne et al. 2011, Orio et al. 2013, Page et al. 2015)—although there are confounding exceptions (e.g., T Pyx; Uthas et al. 2010, Nelson et al. 2014, Patterson et al. 2017). Observations of the steady-burning phase following the TNR (Section 2.1.2) imply that ejecta mass scales inversely with WD mass, as theoretically predicted (Wolf et al. 2013, Henze et al. 2014). However, other observational estimates of ejecta mass consistently yield values more than an order of magnitude greater than theoretically predicted, and show only weak correlations with other nova properties (Roy et al. 2012, Tarasova 2019). Much of this discrepancy may be addressed by more accurately and consistently correcting for the effects of clumping and aspherical geometries in the ejecta, which would drive observational estimates downward (Ribeiro et al. 2014, Wendeln et al. 2017).

Spectroscopic studies of nova ejecta often show overabundances in elements such as carbon, oxygen, and neon (e.g., Ferland & Shields 1978, Williams et al. 1985, Gehrz et al. 1998,

Schwarz et al. 2001, Downen et al. 2013). These metallicity enhancements cannot be the result of nuclear burning because the temperatures achieved during the TNR ($\sim \text{few} \times 10^8$ K) are not high enough to synthesize such heavy elements. Instead, they indicate that mixing must take place between the accreted hydrogen envelope and the underlying CO or ONe WD (Starrfield et al. 1978a, Prialnik et al. 1978). Enrichment of heavy elements into the burning region is also needed to generate outflows of sufficient mass and kinetic energy to be consistent with observations (Starrfield et al. 1998, José & Hernanz 1998). The luminosity from nuclear burning achieved during the peak of the TNR is ultimately limited by β -decay timescales (and hence is temperature insensitive), but it does scale with the abundance of CNO nuclei in the burning region.

Several mechanisms have been proposed to generate mixing at the interface between the accreted material and the underlying WD, which can operate gradually prior to the TNR or rapidly during the TNR. Pre-eruption diffusion can generate moderate enrichment (Prialnik & Kovetz 1984), but may not have sufficient time to operate in high- \dot{M} systems with short recurrence times ($\dot{M} \gtrsim 10^{-9} M_{\odot} \text{ year}^{-1}$; Livio & Truran 1987). Instabilities that feed off differential rotation in the accreted layer, such as the baroclinic instability, could also play a role in long-term gradual mixing (e.g., Kippenhahn & Thomas 1978, Fujimoto 1993). By contrast, Kelvin–Helmholtz instabilities driven by turbulent convection could cause rapid mixing at the onset of the TNR, akin to the process of convective overshoot in stellar evolution (Glasner et al. 1997). Recent multidimensional hydrodynamical simulations have demonstrated that this process can reproduce observed levels of metal enrichment in nova envelopes (e.g., Casanova et al. 2010, José et al. 2020). However, debate continues regarding whether such mixing will always be capable of mixing material through the helium-rich buffer layer generated after each nova by stable hydrogen burning (Section 2.1.2; e.g., Iben et al. 1992b, Starrfield et al. 1998, Denissenkov et al. 2013). The amount of mixing, and how it depends on parameters like \dot{M} and M_{WD} , remains one of the largest uncertainties in modeling nova TNRs (e.g., Starrfield et al. 2020).

Novae are not generally believed to be major contributors to Galactic-scale nucleosynthesis. Although classical novae occur $\gtrsim 10^3$ times more frequently than SNe, they typically eject $\lesssim 10^{-5}$ less mass per event ($M_{\text{ej}} \lesssim 10^{-4} M_{\odot}$ in novae versus $M_{\text{ej}} \gtrsim 1\text{--}10 M_{\odot}$ in SNe). Nevertheless, novae likely are major contributors of isotopes produced by β -decay bottlenecks in the CNO process, such as ^{13}C , ^{15}N , and ^{17}O (e.g., Starrfield et al. 1972, José & Hernanz 1998). They may also be significant contributors of radioactive nuclei such as ^{22}Na and ^{26}Al (José et al. 1997, Hernanz 2012). Recent spectroscopic observations have shown that novae can be prolific producers of ^7Be and its decay product ^7Li (e.g., Izzo et al. 2015, Tajitsu et al. 2015; see Della Valle & Izzo 2020 for a review), confirming early theoretical predictions (Starrfield et al. 1978b). José & Hernanz (2007) and Starrfield et al. (2016) provide focused reviews of nucleosynthesis in novae.

2.1.2. Supersoft X-rays from sustained nuclear burning. Following the TNR, much of the accreted envelope expands and is eventually expelled from the binary (Section 2.2). Once enough mass has been removed, the remaining envelope finds a much more compact hydrostatic solution and burns steadily until it reaches the minimum envelope mass for steady burning (Schwarz et al. 2011, Wolf et al. 2013). During this steady-burning period, which lasts for days to years, the WD maintains a luminosity determined by the core mass–luminosity relation of Paczyński (1970; $L \approx 10^{37}\text{--}10^{38} \text{ erg s}^{-1}$), which approaches the Eddington luminosity as M_{WD} approaches the Chandrasekhar mass $M_{\text{Ch}} \approx 1.4 M_{\odot}$ (see Gehrz et al. 1998 for a detailed discussion). The WD’s effective temperature,

$$T_{\text{eff}} \approx \left(\frac{L}{4\pi\sigma R_{\text{WD}}^2} \right)^{1/4} \approx 6 \times 10^5 \text{ K} \left(\frac{L}{10^{38} \text{ erg s}^{-1}} \right)^{1/4} \left(\frac{R_{\text{WD}}}{10^9 \text{ cm}} \right)^{-1/2}, \quad 1.$$

Eddington

luminosity: the critical luminosity, $L_{\text{edd}} = 4\pi GMc/\kappa$, above which outward radiation pressure exceeds gravity for material of mass M and opacity κ

Ejecta velocity (v_{ej}):
the ejecta velocity of
material expelled in a
nova eruption

ranges from $(0.3\text{--}1.5) \times 10^6$ K. T_{eff} is primarily dependent on M_{WD} (Wolf et al. 2013), because of the WD mass–radius relationship (Nauenberg 1972):

$$R_{\text{WD}} \approx 8.9 \times 10^8 \text{ cm} \left(\frac{M_{\text{WD}}}{M_{\odot}} \right)^{-1/3} \left[1 - \left(\frac{M_{\text{WD}}}{M_{\text{Ch}}} \right)^{4/3} \right]^{1/2}. \quad 2.$$

The nuclear-burning WD’s spectrum peaks in the extreme UV and very soft X-ray bands and is therefore called a supersoft source.

This supersoft X-ray emission is not immediately visible after the TNR, because the dense ejecta absorb the WD’s emission. As the expanding ejecta drop in column density, the photosphere recedes through the ejecta until the supersoft emission is finally revealed (typically on timescales of weeks to months, primarily depending on the ejecta mass and expansion velocity; Schwarz et al. 2011, Henze et al. 2014). This supersoft phase lasts for a time t_{off} that primarily depends on M_{WD} (Starrfield et al. 1974, Sala & Hernanz 2005, Wolf et al. 2013), after which the hydrogen layer is too low in mass to support a steadily burning solution, and the WD cools.

This basic picture has been confirmed by X-ray and UV observations. Early space-based UV observations showed that novae remain UV bright months after eruption, even after their optical emission had substantially faded (Gallagher & Code 1974). A decade later, similar results were found for soft X-rays (Ögelman et al. 1987; see also Shore et al. 1994). High-resolution X-ray grating spectra obtained with the *Chandra X-ray Observatory* and XMM-Newton (*X-ray Multi-Mirror Mission-Newton*) confirmed that the emission arises near the WD surface (e.g., Nelson et al. 2008, Rauch et al. 2010, Ness et al. 2011, Orio 2012). In some cases, the soft X-ray flux is continuum emission directly from the WD atmosphere, whereas in other cases it is in the form of strong emission lines, likely powered by photospheric emission that is obscured from the line of sight (e.g., Ness et al. 2013 and references therein). High-cadence and long-term monitoring of novae by the *Neil Gehrels Swift Observatory* (Gehrels et al. 2004) has substantially improved our understanding of X-ray emission from novae in recent years (see Osborne 2015 and Page et al. 2020 for reviews).

A measurement of the temperature of the supersoft source is the most direct strategy for estimating the WD mass during a nova eruption (Wolf et al. 2013), although caution should be exercised in fitting models to the X-ray spectrum (Krautter et al. 1996). The supersoft turn-off time, t_{off} , is sometimes taken as a proxy for WD mass, and broad agreement exists between theoretical predictions and the observed $T_{\text{eff}}\text{--}t_{\text{off}}$ relation (Henze et al. 2011, Schwarz et al. 2011, Wolf et al. 2013). However, there are significant discrepancies for lower-mass WDs (Henze et al. 2014), perhaps because t_{off} depends on the uncertain mechanism of envelope removal (Section 2.2; Wolf et al. 2013). Constraints on the WD mass, spin period, and properties of the burning layer could, in principle, also be derived from the short-period (30–70 s) supersoft oscillations that have been observed for several novae (e.g., Ness et al. 2015); however, attempts to match these oscillations to theoretical models have so far been unsuccessful (Wolf et al. 2018).

2.2. Mechanisms of Mass Ejection

The observable signatures of nova eruptions—and possibly even their overall radiated energy (Section 4)—are intimately tied to how mass is ejected from the binary. An important clue to the mass-loss mechanism(s) comes from the velocities of nova ejecta, which are observed to span a wide range $v_{\text{ej}} \approx 200\text{--}7,000$ km s^{−1} (e.g., Munari et al. 2011, Naito et al. 2012, Darnley et al. 2016), across different events and even within an individual nova eruption (Section 2.3.2). Most stellar outflows reach velocities comparable with the local escape speed, $v_{\text{ej}} \approx v_{\text{esc}} = (2GM/R_w)^{1/2}$ (e.g., Castor et al. 1975), where R_w is the outflow-launching radius and M is the central mass (in

novae, typically dominated by the WD). Turning this around, the ejecta velocities in novae imply outflow-launching radii of

$$R_w \approx 2.7 \times 10^{10} \text{ cm} \left(\frac{M}{M_\odot} \right) \left(\frac{v_{\text{ej}}}{1,000 \text{ km s}^{-1}} \right)^{-2}. \quad 3.$$

For comparison, the radius of a cold WD is $R_{\text{WD}} \lesssim 10^9 \text{ cm}$ (Equation 2), whereas the semimajor axis of a binary of period P is

$$a_{\text{bin}} \approx 3.5 \times 10^{10} \text{ cm} \left(\frac{M}{M_\odot} \right)^{1/3} \left(\frac{P}{\text{h}} \right)^{2/3}. \quad 4.$$

This simple analysis implies that different physical mechanisms may drive slow (about a few hundred km s^{-1}) and fast (about a few thousand km s^{-1}) outflows in novae. Indeed, several distinct mechanisms driving nova mass loss have long been proposed, described in Sections 2.2.1–2.2.3.

2.2.1. Impulsive ejection at the thermonuclear runaway. Observations of nova ejecta are often interpreted as a single, impulsive mass ejection coincident with the TNR (e.g., Seaquist & Bode 2008, Mason et al. 2018). Indeed, some hydrodynamical models of nova eruptions find that a portion of the envelope is ejected in the minutes to hours following the TNR, powered by radioactive heating from β -unstable nuclei (Starrfield et al. 2008). Early models found that this material is expelled as a shock-driven shell (Sparks 1969, Prialnik 1986), but later this picture was revised to a wind whose mass-loss rate declines by an order of magnitude over just a few hours (Prialnik & Kovetz 1992).

The mass and velocity of this initial ejection are primarily determined by the amount of mixing between the accreted envelope and the underlying WD (Starrfield et al. 1978a, 1998). Even when significant prompt ejection occurs in nova models, only a fraction of the envelope is expelled. If this were the only form of mass loss from novae, the sustained-burning supersoft X-ray phase would last for longer than observed (\sim centuries; Starrfield et al. 1978a). Combined with the \sim Eddington luminosities of novae after eruption, this implies a significant role for prolonged winds driven by radiation pressure (Section 2.2.2).

Rotation of the WD could in principle play a role in shaping the geometry of the prompt ejecta (e.g., Porter et al. 1998, Scott 2000). If the WD is rotating rapidly, the pressure of the burning layer—and hence the peak temperature achieved during the TNR—will depend on latitude, being significantly higher at the rotational poles than at the equator. As the rate of energy production is extremely sensitive to temperature, material ejected from the earliest phases of the eruption could show departures from spherical symmetry. Rotation could also play a role in shaping any centrifugally driven mass loss, if the expanding nova envelope—instead of simply conserving angular momentum—maintains corotation with the rapidly spinning WD surface via efficient magnetic coupling (Zhao & Fuller 2020).

2.2.2. Prolonged optically thick winds. The highest velocities seen in novae, $v_{\text{ej}} > 1,000 \text{ km s}^{-1}$, imply small launching radii $R_w \lesssim a_{\text{bin}}$. A compelling physical model for such fast outflows is an optically thick wind (i.e., where acceleration occurs deep below the photosphere) driven by radiation pressure from the near-constant luminosity of the nuclear-burning WD (Friedjung 1966, Bath & Shaviv 1976). The Eddington luminosity at a particular location in the nova envelope depends on temperature- and density-dependent mean opacities (as estimated, e.g., in the OPAL tables; Iglesias & Rogers 1996), and the most important wind driver is the iron opacity bump, which occurs at a temperature of $T \approx 1.6 \times 10^5 \text{ K}$ in the nova envelope (Kato & Hachisu 1994).

As the mass of the hydrogen-rich envelope gradually decreases owing to both nuclear burning and wind mass loss, the wind mass-loss rate also decreases (Kato & Hachisu 1994). This makes

the photosphere recede to smaller radii and higher temperatures, and the wind velocity increases as the wind-launching radius moves inward. Once the photospheric temperature exceeds 1.6×10^5 K, there are no more significant sources of opacity, and the wind ceases (Kato & Hachisu 1994). The duration of the optically thick wind phase can be days to months, with lower M_{WD} and larger ignition masses leading to longer wind phases.

Essentially all hydrodynamical models of nova eruptions to date are one-dimensional. Even putting aside the potential for large-scale asphericity in nova ejecta, this simplification may gloss over important physics in nova winds. Multidimensional (magneto)hydrodynamical models of massive star envelopes with radiative transport reveal that, near the Fe opacity bump, the interplay between convection and radiative transport is complex. Multidimensional effects and magnetic fields lead to clumpy envelope structures that qualitatively differ from those predicted by one-dimensional models (e.g., Jiang et al. 2017). If similar physical processes occur in nova outflows, they could drive the inferred clumpiness of nova ejecta (Sections 2.3.2 and 2.3.5).

Some novae reach peak luminosities that exceed the electron-scattering Eddington luminosity by up to an order of magnitude ($\sim 10^{39}$ erg s $^{-1}$) (e.g., Duerbeck 1981; Schwarz et al. 1998, 2001; Aydi et al. 2018; Skopal 2019).³ To achieve photon luminosities in excess of the Eddington limit requires an additional energy source in the hydrostatic WD atmosphere that is itself super-Eddington (Quataert et al. 2016). A substantial contribution to the luminosity of novae may come from internal shocks within the ejecta (Section 4.3), but even in this case super-Eddington kinetic luminosities would be required of the outflow. Alternatively, apparently super-Eddington luminosities can be achieved with nuclear burning if the Eddington luminosity is increased because of a decrease in the effective opacity (κ ; $L_{\text{edd}} \propto \kappa^{-1}$), which might be achieved if the WD atmosphere is clumpy or porous (e.g., Shaviv 2001).

2.2.3. Common-envelope-like mass loss. The orbital motion of the binary is another potential source of energy (e.g., MacDonald 1980). Around light curve maximum, the nova photosphere reaches radii of $\approx 10^{12}$ – 10^{13} cm, engulfing the binary system for novae with main sequence companions ($P \lesssim 12$ h). The proposed picture is broadly similar to a common-envelope phase, in which a compact secondary star enters the dilute envelope of a giant companion following a phase of unstable binary mass transfer (Paczynski 1976, Ivanova et al. 2013). The influence of the binary should be particularly apparent for the lowest ejecta velocities, $v_{\text{ej}} \lesssim 1,000$ km s $^{-1}$, for which the inferred launching radii reside near or exterior to the companion orbit ($R_{\text{w}} \gtrsim a_{\text{bin}}$; Equations 3 and 4).

Early generations of opacity tables implied that radiation-driven winds were insufficient to eject the envelope, and works by MacDonald et al. (1985) and Shankar et al. (1991) found that frictional drag in this common-envelope phase could power mass ejection in slow novae. However, other models that also included radiation-driven winds found that the envelope density near the binary orbit was too low for frictional heating to play a significant role (Kato & Hachisu 1991a,b). With the inclusion of the Fe opacity bump, radiation-driven winds seemed sufficient to remove the envelope (Kato & Hachisu 1994). Studies of the common-envelope phase in novae fell out of fashion, although Kato & Hachisu (2011) find that it may play a role in the slowest novae. However, we note that even in the common-envelope community, the efficiency and timescale of envelope removal remain uncertain (Ivanova et al. 2013), and modern common-envelope simulations fail to unbind stellar envelopes without contentious additional sources of energy like recombination (Ivanova 2018), jets (Soker 2017), or pulsation-driven shocks (Clayton et al. 2017).

³Kato & Hachisu (2005, 2007) propose that super-Eddington luminosities can be attained by including free-free emission in the wind. However, this violates energy conservation unless the wind is heated from below by some mechanism at a rate that itself exceeds the Eddington luminosity.

In principle, nature has provided a relatively clean test of the role of frictional drag in nova ejections, as embedded novae driven by wind accretion and with large orbital separations should not undergo a common-envelope phase. Indeed, embedded novae seem to be polarized to two extremes: very fast novae that evolve over just a week or so, and symbiotic novae that are among the slowest evolving transients known (evolving over decades to centuries; Mikolajewska 2010). The lack of embedded novae that evolve on intermediate timescales (approximately months) may be an indication that nova envelope removal is less efficient in long-period systems owing to the lack of frictional drag (Kenyon & Truran 1983).

Even if frictional drag from the binary is not a significant contributor to the expulsion of the envelope, binary orbital motions may still shape the ejecta, concentrating them in the equatorial plane (Hutchings 1972, Livio et al. 1990). As the WD envelope expands to encompass the binary orbit following the TNR, it will be spun-up and preferentially focused toward the binary orbital plane by centrifugal forces and the companion star's gravity. Mass loss may preferentially occur through the minimum in the gravitational potential near the outer L_2 Lagrange point, located in the orbital plane on the far side of the secondary companion. Mass that leaks through this nozzle emerges from the binary in a sprinkler-like spiral pattern and is accelerated outward by a combination of pressure gradients and nonaxisymmetric torques from the binary motion (Shu et al. 1979, Pejcha et al. 2016). As discussed in Section 2.3.5, novae are commonly observed to have bipolar outflows and equatorial rings or disks, which suggest that the binary orbit plays an important role in shaping, if not driving, outflows during at least some phases of the eruption. Spatially resolved images suggest that the remnants of slower novae are more elongated than those of higher-velocity novae (e.g., Slavin et al. 1995, O'Brien & Bode 2008), as expected if orbital motion shapes the ejecta.

2.3. Multiwavelength Observations of Mass Loss

Our understanding of nova mass loss is constrained by observations that span the electromagnetic spectrum. As described in Sections 2.3.1 and 2.3.2, novae are primarily known as optical transients, but infrared (Section 2.3.3) and radio observations (Section 2.3.4) provide unique insights on different components of the ejecta (neutral/solid and ionized phases, respectively).

2.3.1. Optical light curves. Novae are generally discovered as optical transients, and a large fraction of their total radiative output emerges in the optical. The luminosity increases by 8–15 mag from pre-outburst to peak in the V band, typically over just a few days (Payne-Gaposchkin 1957, Warner 1995). Peak absolute magnitudes are generally in the range $M_V = -5$ to -10 mag (Shafter 2017, Özdönmez et al. 2018), implying luminosities close to—or in excess of—the WD Eddington luminosity $L_{\text{edd}} \approx 10^{38}$ erg s $^{-1}$.

To first order, a simple nova model of expanding ejecta and constant bolometric luminosity powered by sustained nuclear burning on the WD (Section 2.1.2) predicts a smooth decline like that seen over most of V392 Per's evolution in **Figure 3**. The optical emission declines because the peak of the spectral energy distribution shifts blueward, into the UV and X-ray, as the ejecta thin and the photosphere recedes. At late times, the ejecta become optically thin and the optical luminosity is dominated by emission lines. Although this has been the standard picture for decades (e.g., Gallagher & Starrfield 1976, Shore et al. 1994), recent work has demonstrated that reprocessed emission from internal shocks may also contribute or even dominate the optical luminosity in some cases (see Section 4.3 for more discussion).

Nova light curves are often labeled according to their speed class, quantified as the time to decay by 2 or 3 visual magnitudes from peak brightness (denoted as t_2 or t_3 , respectively; Payne-Gaposchkin 1957). Speed classes range from very fast ($t_2 < 10$ days) to very slow ($t_2 \gtrsim 150$ days).

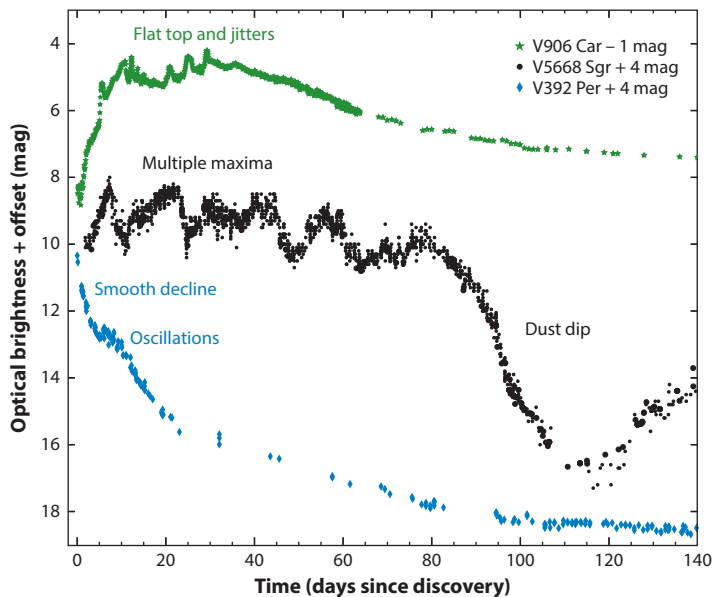


Figure 3

A selection of optical light curves of three novae (all detected by *Fermi*-LAT, the *Fermi Large Area Telescope*), illustrating the broad range of nova light curve morphologies. Common features are marked, several of which remain unexplained. V906 Car data are in the *R* band, from the BRITE (*BR*ight-*star* *T*arget *E*xplorer) satellite and the Stonybrook/SMARTS atlas of novae (Walter et al. 2012, Aydi et al. 2020b). The V5668 Sgr light curve is composed of *V* (larger circle symbols) and visual band (smaller circles) data from the AAVSO (American Association of Variable Star Observers; Kafka 2020) database. The V392 Per data are in the *V* band, as reported by the AAVSO. Figure created by E. Aydi.

Roughly speaking, the speed class is related to the timescale for envelope expansion and removal; a nova fades faster for lower ejecta masses or higher expansion velocities (for a controversial application, see the sidebar titled Maximum Magnitude–Rate of Decline Relationship?). Theory predicts that M_{ej} is a decreasing function of WD mass (Section 2.1.1), so faster novae are typically thought to arise from more massive WDs. However, the relationship between speed class and WD mass has relatively little direct empirical proof and should be applied with caution. Also note that t_2 is far from a perfect metric, given the complex evolution of many nova light curves. For example, the optical brightness of V5668 Sgr drops 2 mag below its maximum several times during its evolution (Figure 3).

Strope et al. (2010) present a catalog of 93 well-observed nova light curves, mainly from the American Association of Variable Star Observers database (AAVSO; Kafka 2020), that they use to classify light curves based on both their speed class and light curve shape. Although many novae (38% of their sample) show smoothly declining light curves as might be expected from a simple nova model, they note that the rest of the sample show curious features in their light curves like flat tops, oscillations, and jitters. Figure 3 shows three optical light curves illustrating these morphologies. With the exception of dust dips (Section 2.3.3), most of these features currently lack compelling theoretical explanations. Increasingly, the diversity and complexity of nova light curve structure—and the present lack of explanation for them—are being highlighted by the higher cadence and precision light curves from space-based facilities like the Solar Mass Ejection Imager on *Coriolis* (Hounsell et al. 2010, 2016), HI-1B on STEREO (*Solar Terrestrial Relations*

MAXIMUM MAGNITUDE–RATE OF DECLINE RELATIONSHIP?

An empirical relationship has long been claimed between the absolute optical magnitude of novae at maximum, M_V , and light curve decline time, such that more luminous novae have shorter t_2 or t_3 (e.g., McLaughlin 1945, de Vaucouleurs 1978, Capaccioli et al. 1989, Della Valle & Livio 1995). Although physical interpretations of this “maximum magnitude–rate of decline” (MMRD) relation have been proposed (Livio 1992), they rely on model-dependent predictions for how the peak nova luminosity scales with the WD mass. The MMRD relationship is frequently used to determine distances to Galactic novae, if the visual extinction is known. However, recent observations show a significant population of outliers from this relationship that have called the existence of the MMRD relation into question (particularly fast and faint novae; Kasliwal et al. 2011, Shara et al. 2017). There are also theoretical reasons to believe nova properties should not conform to a one-parameter family (e.g., Prialnik & Kovetz 1995). Given this uncertainty in the MMRD, substantial and hard-to-quantify uncertainties exist in the distances to most Galactic novae, although there have been recent promising advances using three-dimensional dust maps (Özdönmez et al. 2016) and *Gaia* astrometry (Schaefer 2018).

Observatory; Eyres et al. 2017, Thompson 2017), *BRITe* (*BR*ight-*star* *T*arget *E*xplorer Aydi et al. 2020b), and, in the future, *TESS* (*Transiting Exoplanet Survey Satellite*).

For example, many novae exhibit multiple distinct optical maxima, also called jitters or flares (e.g., Bianchini et al. 1992; Strobe et al. 2010; Walter et al. 2012; Aydi et al. 2019, 2020b). During optical flares, the radius of the photosphere often appears to temporarily expand (e.g., Tanaka et al. 2011, Munari et al. 2015, Aydi et al. 2019), behavior that is accompanied by the appearance of new absorption line systems and suggestive of distinct mass ejection events (see Section 2.3.2; although Williams 2016 tries to explain them with ballistic, expanding clumps). There are several mechanisms that could in principle give rise to such variability in the outflow properties. One is a change in the nature of the mass-loss mechanism, like a global transition from a slower, quasi-hydrostatic common-envelope phase to a faster wind (Kato & Hachisu 2011). Alternatively, some studies have found that contraction of the envelope may accelerate nuclear burning at the base of the envelope, triggering envelope re-expansion and a second mass-loss episode (e.g., Prialnik & Livio 1995, Hillman et al. 2014), although this has not been found in other work (e.g., Townsley & Bildsten 2004, Denissenkov et al. 2013). As we shall discuss in Section 3.1, there is evidence from gamma-ray observations that optical flares in at least some cases are powered by internal shocks (see Sanyal 1974 for an early discussion of this scenario).

2.3.2. Optical spectra. Novae have been studied using optical spectroscopy for more than a century (McLaughlin 1943, Payne-Gaposchkin 1957). Over the course of an eruption, nova spectra transition from being photosphere dominated to emission line dominated as the ejecta thin, and the ionization states of the lines generally increase. The widths, shapes, and evolution of spectral features hold information about ejecta dynamics and morphology. Here, we focus on spectroscopic constraints on mass ejection in novae, but note that the relative strengths of emission lines also constrain the abundances of the ejecta and can reveal whether the nova occurred on a CO or ONe WD (see Gehrz et al. 1998 for a review of abundance measurements). The optical spectra of novae are strongly affected by physical processes that occur in the UV, where the bulk of the WD luminosity is emitted and the ejecta opacity from atomic transitions is highest. Shore (2008) provides a description of the UV spectral evolution that accompanies what is more commonly observed in the optical.

Novae discovered early in their evolution are generally first seen during a photospheric or fireball phase, in which the nova envelope expands and the optical light curve rises, as the photosphere radius grows and the effective temperature decreases (Gehrz 1988, Hauschildt et al. 1994). This expansion is accompanied by the onset of mass loss, which is variously modeled as an impulsive shell ejection (e.g., Shore 2014, Mason et al. 2018), a continuous wind (Kato & Hachisu 1994), or a combination of the two (e.g., Friedjung 1987). As the nova transitions through optical maximum, most spectral lines show P Cygni profiles [e.g., **Figure 4b**, subpanel *i*; see also the sidebar titled Transient Heavy-Element Absorption (THEA) Features]. With time, the emission component strengthens and the absorption weakens, which is consistent with a receding photosphere. Around maximum, the ejecta are dense enough to rapidly recombine, and the Strömgren sphere powered by the nuclear-burning WD is ionization bounded, such that the outer ejecta are neutral (Beck et al. 1990, Williams 1990). The dominant form of opacity is line blanketing by nearly neutral metals (sometimes called the iron curtain, because most of the opacity is from Fe lines; Shore & Aufdenberg 1993, Shore et al. 1994). Modeling of spectra before and around maximum finds that the density profiles of nova ejecta require a sharp outer profile ($\rho \propto r^{-n}$ with $n \approx 15$ for the outermost ejecta; Hauschildt et al. 1994) and a much shallower profile for the inner ejecta ($n \approx 3$; Hauschildt et al. 1992). This density profile implies that the optical photospheric radius can vary by over a factor of ~ 100 with wavelength and thereby sample a range of temperatures and ionization states (Hauschildt et al. 1995); it is therefore often called a pseudophotosphere.

As the ejecta expand and densities drop, the ionization front moves outward (in both mass and physical coordinates), until eventually the ejecta become fully ionized (Williams 1990). Spectral features transition from low ionization states around maximum (e.g., H I, O I, Na I, Fe I; **Figure 4b**, subpanel *i,ii*) to highly ionized species. As the density drops, the features also migrate from permitted transitions to forbidden transitions (**Figure 4b**, subpanel *iii*). Eventually, the nova spectrum relaxes to a nebular phase characterized by forbidden lines from ionization states expected in a $\sim 10^4$ -K gas (e.g., [N II], [O III]). The ejecta can remain ionized long after the supersoft source turns off, frozen in by the now low densities and long recombination times.

The evolution of the spectral line profiles is less well explained. Prior to optical peak, the emission lines show P Cygni profiles characterized by low velocities, 200–1,000 km s^{−1} (**Figure 4c**, subpanel *i*). Shortly after the nova reaches maximum light, a broad emission component emerges with roughly double the width of the initial P Cygni profile, whereas the absorption component from the preexisting P Cygni profile remains superimposed on top of the broader emission (**Figure 4c**, subpanel *ii*). This general spectral evolution has long been recognized (e.g., McLaughlin 1943, Gallagher & Starrfield 1978), and has recently been revisited by Aydi et al. (2020a), who show that it is nearly universal across diverse novae. The field has long grappled with interpreting the coexistence of broad emission and lower-velocity absorption, along with the abrupt transition of the spectral profiles over a matter of days. The longest-standing and most cohesive explanation is to have two physically distinct—slow and fast—ejecta components (Section 2.3.5) (McLaughlin 1947; Friedjung 1987, 2011; Friedjung & Duerbeck 1993). Additional components can subsequently appear in the line profiles, usually at higher velocities (e.g., Hutchings 1970b). Their appearance is frequently associated with new flares in the optical light curve (Csák et al. 2005, Tanaka et al. 2011).

Another curiosity is that spectral components are often seen to accelerate to higher velocities (e.g., Hutchings 1970a, Shore et al. 2011). This phenomenon may reflect driving of clumps or shells by radiation pressure (e.g., Williams et al. 2008), or it can potentially be explained by inner ejecta expanding faster than the outer ejecta (so that as the photosphere recedes inward in mass, it samples faster material; Friedjung 2011). If the line-forming regions are thin shells compressed by

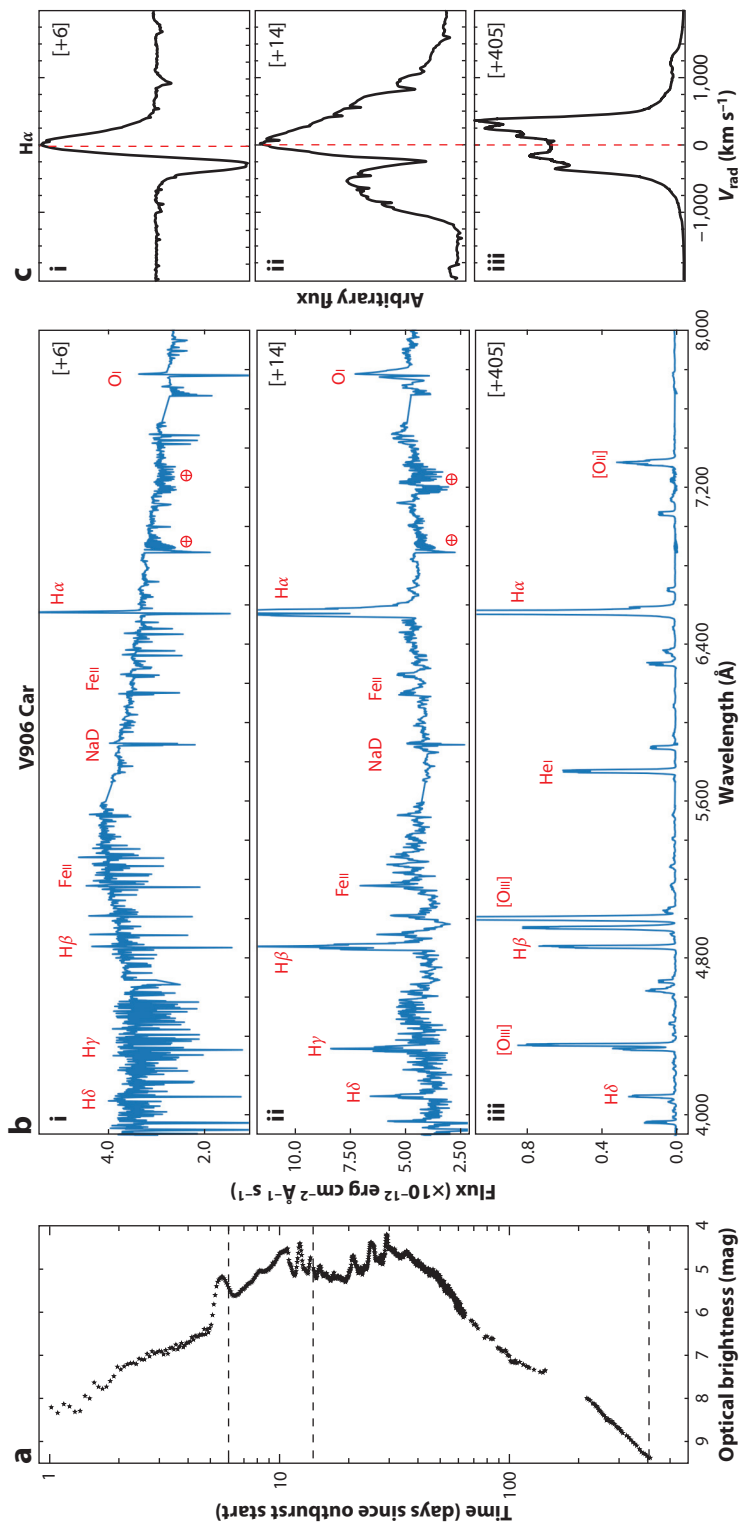


Figure 4

The optical evolution of nova V906 Car (2018), showing (a) the optical (*R*-band) light curve, (b) optical spectra, and (c) close-ups detailing spectra around the H α line. Panel b, subpanel i is from 5 days before light curve maximum (6 days after the start of eruption) and shows a photospheric spectrum with relatively narrow P Cygni profiles. Panel b, subpanel ii is from 3 days after light curve maximum and shows the strengthening and broadening of emission lines. The absorption component from the previous spectrum is still superimposed on the emission line. Panel b, subpanel iii is from more than a year after maximum, and shows a nebular spectrum dominated by high-excitation and forbidden emission lines. Data from Aydi et al. (2020b); figure created by E. Aydi.

TRANSIENT HEAVY-ELEMENT ABSORPTION (THEA) FEATURES

Narrow absorption features from heavy elements (e.g., FeII, TiII, CrII; FWHM $\approx 30\text{--}300\text{ km s}^{-1}$) are seen in high-resolution spectra of some novae around and after maximum light (Williams et al. 2008, Williams & Mason 2010). The physical origin of these THEA features is uncertain. Williams & Mason (2010) proposed that the absorbing material originates from a massive preexisting circumbinary disk, inspired by the model of Taam & Spruit (2001). However, IR searches rule out such disks around most CVs, finding $\lesssim 10^{-10} M_{\odot}$ in circumbinary material (e.g., Hoard et al. 2014). Williams (2012) proposed that the THEA lines originate in material irradiated or ablated from the secondary star during the nova; however, simulations find that mass loss from the secondary is likely not sufficient to produce the THEA lines (Figueira et al. 2018).

The possibility remains that the THEA lines are associated with the early slow ejecta (e.g., the narrow P Cygni component in **Figure 4b**, subpanel *i*; Aydi et al. 2020a). Some novae show hints that the THEA features are not present at the start of the eruption (Williams 2012), and the velocity and timing of their appearance matches the premaximum and/or principal components long recognized in nova spectra (McLaughlin 1943, Williams 2013). Future high-resolution spectra from early in nova eruptions will test if the THEA lines originate in nova ejecta or have a more exotic source.

internal shocks, momentum added to the shell by the WD outflow leads to acceleration (Steinberg & Metzger 2020; Section 4).

Spectroscopy also reveals that the morphologies of nova ejecta are complex. The ejecta are clumpy, as optical line ratios yield estimates of volume filling factors $\lesssim 0.1$ in the line-emitting gas (e.g., Ederoclite et al. 2006, Shore et al. 2013). Structures in the spectral lines often persist at a particular velocity for months during eruption and have been interpreted as originating in discrete clumps within the ejecta (Williams 2013, Mason et al. 2018). The coexistence of a wide range of ionization species indicates a striking range of densities and temperatures within the ejecta and is again often explained as self-shielding clumps (Saizar & Ferland 1994, Williams 1994). A variety of evidence at other wavelengths also supports nova ejecta being highly clumpy. This includes clearly resolved inhomogeneities in images of nova ejecta (Section 2.3.5) and large-amplitude variability in the supersoft X-rays due to time-variable absorption (likely by clumps in the ejecta; Osborne et al. 2011, Page et al. 2020). Finally, emission line profiles, especially in the nebular stage, often imply bulk asphericity in the ejecta. Double- or triple-peaked line profiles are common and usually interpreted as bipolar ejecta, sometimes surrounded by an equatorial torus or disk (Hutchings 1972, Ribeiro et al. 2013, Shore et al. 2013).

2.3.3. Dust and molecule formation. Many novae show dramatic drops in their optical light curves on timescales of 20–100 days after the eruption (e.g., V5668 Sgr in **Figure 3**). These features arise from obscuration of the optical photosphere by dust forming within the expanding ejecta.

These so-called dust dips are accompanied by a simultaneous increase in the mid-IR emission, originating from the outer ejecta coincident with the newly formed dust layer (e.g., Hyland & Neugebauer 1970, Gehrz et al. 1980). Dust formation is rapid, with grains growing to large sizes ($\sim 1\text{ }\mu\text{m}$) compared to the dust in the interstellar medium over month-long timescales (e.g., Helton et al. 2010, Gehrz et al. 2018). It is sometimes preceded by the formation of molecules like carbon monoxide (CO), which can be detected in near-IR spectra of novae within weeks of eruption (Banerjee & Ashok 2012). The molecular and dust phases of nova ejecta, and the complex chemistry therein, are reviewed by Gehrz (1988) and Evans & Rawlings (2008).

The timescale for dust formation can be understood from basic considerations. The equilibrium temperature in an ejected shell is set by the radiation of the WD of luminosity L and given by $T_{\text{eq}} = (L/4\pi\sigma R^2)^{1/4}$, where $R = v_{\text{ej}}t$ is the ejecta radius at time t after the start of expansion and v_{ej} is the velocity. Dust formation should take place when T_{eq} reaches the condensation temperature for solids, $T_s \approx 1,200$ K, at time

$$t_{\text{dust}} \equiv \left(\frac{L}{4\pi\sigma v_{\text{ej}}^2 T_s^4} \right)^{1/2} \approx 76 \text{ days} \left(\frac{L}{10^{38} \text{ erg s}^{-1}} \right)^{1/2} \left(\frac{v_{\text{ej}}}{10^3 \text{ km s}^{-1}} \right)^{-1} \left(\frac{T_s}{1,200 \text{ K}} \right)^{-1/2}. \quad 5.$$

This estimate is in reasonable accord with the observed timescale for dust formation. A more detailed calculation accounts for the fact that the dust grains do not reradiate the absorbed light of the nova with perfect efficiency, and thus do not share the effective temperature of the radiation (e.g., Evans et al. 2017).

However, the nova ejecta are irradiated by the UV and X-ray luminous WD, leading to hostile conditions for the formation of molecules and dust. The chemistry leading to dust formation requires carbon atoms to be neutral (e.g., Rawlings & Williams 1989) and shielded from the WD's harsh radiation (e.g., Bath & Harkness 1989). This is generally only possible if the gas density is $\gtrsim 10^9\text{--}10^{10} \text{ cm}^{-3}$ (e.g., Gehrz & Ney 1987), which is higher than one would predict if the thickness of the ejecta shell were comparable with its radius ($R_{\text{dust}} \approx v_{\text{ej}}t_{\text{dust}}$) at the dust formation epoch. Clumps in the ejecta (as observed with optical spectroscopy and imaging; Sections 2.3.2 and 2.3.5) may be dense enough to serve as sites of dust formation (Woodward et al. 1992). As we discuss in Section 4.4, shocks within the nova ejecta also offer a mechanism to produce regions of dense, self-shielded gas (Derdzinski et al. 2017).

Although optical dust dips are seen in only $\sim 20\%$ of novae (Strope et al. 2010), many more show signs of dust formation in the form of IR excess (Gehrz et al. 1998). A population study of how many novae show dust signatures is currently lacking but is increasingly feasible in an era in which multiband optical–IR light curves are common. Though some of the variation in dust signatures may be due to differences in the amount of dust formed (Gehrz 1988), inclination effects from aspherically distributed dust formation probably play an important role. In the well-resolved helium nova V445 Pup, some dust emission is seen in a bipolar outflow, but the primary evidence for dust is seen in absorption, as an equatorial disk that blocks the light of the binary (**Figure 5d**; Woudt et al. 2009). In the slow nova V1280 Sco, emission from warm dust is again structured in two distinct lobes, but it is unclear if an absorbing equatorial dust disk is located between them or if all the dust is located in the bipolar outflow (Chesneau et al. 2012). The sites of dust formation in novae remain poorly understood, and additional high-resolution IR and millimeter observations are needed to understand the link between dust formation and bulk ejecta properties.

2.3.4. Thermal radio emission. Nova ejecta emit free–free thermal radiation at radio wavelengths, which is often modeled as an isothermal expanding HII region (Seaquist & Bode 2008). Temperatures of $\sim 10^4$ K are maintained throughout the ejecta by the ionizing radiation from the central WD (Section 2.1.2), despite the loss of thermal energy due to adiabatic expansion (Cunningham et al. 2015).

Nova radio light curves evolve over months to years, brightening as the radio photosphere expands with the ejecta and then fading as the ejecta become optically thin at radio wavelengths and continue decreasing in density (Seaquist & Palimaka 1977, Hjellming et al. 1979). Free–free opacity decreases with increasing frequency, so emission at higher frequencies emerges from deeper in the ejecta and evolves more rapidly. The spectral energy distribution of the free–free emission has been observed over >2 decades in radio frequency, from 1 GHz to 600 GHz (Iverson

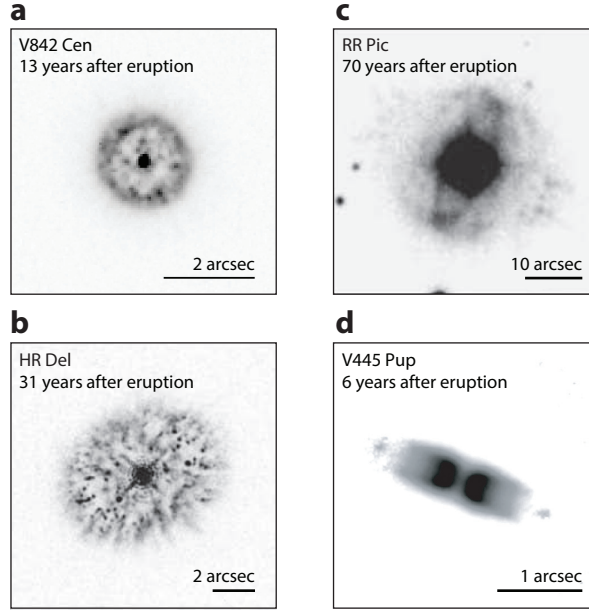


Figure 5

Optical and IR images of novae, highlighting the diversity of structures observed in their ejecta. The time elapsed between eruption and imaging is noted in the top-left corner of each panel. (a) V842 Cen (1986), observed with the HST (*Hubble Space Telescope*) and consistent with a clumpy near-spherical ejection (image obtained from the Hubble Legacy Archive). (b) HR Del (1967) imaged with HST, revealing “cometary tails” extending from clumps, prolate ellipticity, and on closer analysis, bipolar morphology. (c) RR Pic (1925) observed with the Anglo-Australian Telescope and exhibiting a clear equatorial ring and bipolar outflows. (d) Helium nova V445 Pup (2000) imaged using adaptive optics on the Very Large Telescope, showing a collimated bipolar outflow with distinct high-velocity knots at the outer extent of each lobe. The V842 Cen, HR Del, and RR Pic images trace the $H\alpha + [\text{NII}]$ emission line flux, while the V445 Pup image was obtained in the near-IR K band and traces both warm dust and plasma. Panel a obtained from the Hubble Legacy Archive, panel b adapted from Harman & O’Brien (2003), panel c adapted from Gill & O’Brien (1998), and panel d adapted with permission from Woudt et al. (2009); copyright AAS.

et al. 1993, Hjellming 1996), although there is a poorly explored possibility that dust emission can contribute at >100 GHz (e.g., Nielbock & Schmidtobreick 2003).

The radial profile of the ejecta density $\rho(r, t)$ and the total mass of ionized gas can be measured by monitoring of the radio spectral evolution. The radio data are generally well fit with a density profile, $\rho \propto r^{-2}$ or r^{-3} (e.g., Seaquist et al. 1980, Weston et al. 2016b), which is consistent with models of UV and optical spectra (e.g., Hauschildt et al. 1992; Section 2.3.2). The derived ejecta profiles show a cutoff at an inner radius and can be explained as either a single homologous ejection or a prolonged wind (Hjellming et al. 1979, Kwok 1983). In addition, by tracking the expanding photosphere, radio light curves can trace when the ejecta began expanding, sometimes revealing surprising weeks-long delays between the rise of the optical light curve and the ejection of the radio-emitting material (Nelson et al. 2014, Linford et al. 2017).

Ejecta masses obtained by integrating the inferred density profile are typically in the range of $M_{\text{ej}} \approx 10^{-5} - 10^{-3} M_{\odot}$ (Roy et al. 2012, Wendeln et al. 2017). Because the free-free luminosity traces the emission measure of the ionized gas ($n_e^2 dl$, where n_e is the number density of electrons and dl is the path length through the ejecta), ejecta mass estimates require assumptions about the

filling factor, which can be <1 due to ejecta clumping or aspherical geometry. Analyses of optical spectra imply filling factors less than unity (Section 2.3.2), implying that many ejecta mass estimates in the literature are in fact upper limits, because they assume it is unity. Spherical symmetry is generally assumed when modeling the nova's radio light curve, but morphokinematic codes like *SHAPE* now allow this assumption to be relaxed (Steffen et al. 2011, Ribeiro et al. 2014). In the future, radio images from high-resolution interferometers (Section 2.3.5) can inform the choice of ejecta morphology and be analyzed along with radio light curves and optical spectroscopy to provide better measurements of the distribution and total mass of ejecta.

2.3.5. High-resolution imaging. As they expand, the nova ejecta grow to an angular diameter,

$$\theta = 0.115 \text{ arcsec} \left(\frac{d}{1 \text{ kpc}} \right)^{-1} \left(\frac{v_{\text{ej}}}{1,000 \text{ k s}^{-1}} \right) \left(\frac{t}{100 \text{ days}} \right), \quad 6.$$

where t is the time since the start of expansion and d is the source distance. This means that high-resolution facilities like the *Hubble Space Telescope* (HST), radio/millimeter interferometers, and near-IR adaptive-optics imagers can resolve nova ejecta starting several months to years after eruption, whereas ground-based optical imaging can resolve novae years to decades after eruption.

The largest samples of nova images have been collected at optical wavelengths, usually using an $\text{H}\alpha + [\text{NII}]$ narrow-band filter, at relatively late times ($\sim 5\text{--}50$ years; e.g., Cohen 1985, Slavin et al. 1995, Gill & O'Brien 1998, Downes & Duerbeck 2000). Most resolved nova ejecta are consistent with clumpy, elliptical shells (e.g., V842 Cen in **Figure 5a**; see O'Brien & Bode 2008 for more examples). A minority show more complex structures, like bipolar overdensities and equatorial rings (e.g., RR Pic in **Figure 5c**). When imaged at enough resolution, the emission usually resolves into clumps, and sometimes cometary tail features extend radially outward from the clumps (e.g., HR Del in **Figure 5b**), which could be the result of a fast wind moving past a slower inhomogeneous shell (Lloyd et al. 1995).

High-resolution imaging at earlier times is more informative, because it can trace how the ejecta evolve over time and resolve structures that may be too low-density to detect at late times, thereby constraining the mass ejection processes (Section 2.2) and the origin of shocks (Section 4). Early images of nova ejecta are consistent with an elliptical expanding shell, and ellipticity that changes with time, from either rounder to more prolate (Paresce et al. 1995, Schaefer et al. 2014) or more prolate to rounder (Taylor et al. 1988, Pavelin et al. 1993, Hjellming 1996). This evolution is suggestive of at least two distinct flows with different morphologies (Taylor et al. 1988, Chochol et al. 1997), but the clearest evidence comes from V959 Mon, where the binary has an edge-on inclination convenient for imaging (Page et al. 2013). Chomiuk et al. (2014a) present multiepoch radio images of V959 Mon, and capture an orthogonal flip in the major axis of the ejecta, as shown in **Figure 6**. They interpret the bi-lobed ejecta imaged on day 126 as a fast bipolar flow (elongated along the horizontal axis in **Figure 6a**, subpanel *ii*), which diffuses and fades by day 615 when the emission is dominated by a slower, equatorially concentrated flow (elongated along the vertical axis in **Figure 6a**, subpanel *iii*). HST imaging confirms an edge-on equatorial torus and faster, bipolar flows in V959 Mon (Sokoloski et al. 2016). The helium nova V445 Pup showed a similar morphology (**Figure 5d**), and, in a clear indication of nonimpulsive mass ejection, Woudt et al. (2009) showed that the high-velocity knots visible at the extremities of the bipolar outflow were not ejected until 345 days after the start of eruption. Novae that are near enough for well-resolved, early imaging campaigns are rare, and every future example should be exploited as a test of whether the bipolar outflow and equatorial disk configuration can explain the morphologies of all novae.

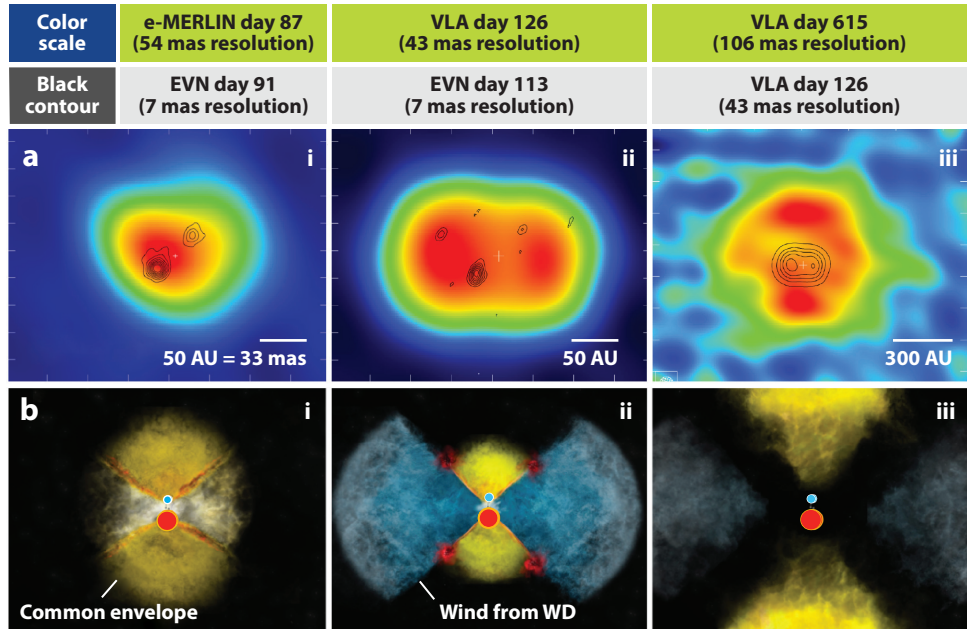


Figure 6

Images of nova V959 Mon show an ejecta geometry consistent with a slowly expanding equatorial torus, a faster bipolar outflow, and shocks where the two ejections collide. Panels in row *a* show radio observations (labeled at *top*), and panels in row *b* are a series of artist's representations. In column *i*, thermal nova ejecta are barely resolved at early times (color scale in panel *a*, column *i*), and synchrotron knots are superimposed on the ejecta (*black contours* in panel *a*, column *i*, likely tracing shocks and particle acceleration; Section 3.3). In column *ii*, the thermal ejecta are concentrated in two lobes (color scale), and the synchrotron knots (*black contours*) surround the lobes. This is consistent with a fast wind funneled along the binary's poles, crashing into dense equatorial material and producing shocks at the interface (*red regions* in panel *b*, column *ii*). In column *iii*, the polar outflow has faded, and the brightest, densest material is now oriented perpendicularly, corresponding to the slower equatorial material (color scale; the *black contours* here represent thermal emission from day 126, for comparison). This can be explained if the WD stops powering a wind at late times, and the ejecta drift off into space. Radio images adapted from Chomiuk et al. (2014a); artist's impressions created by B. Saxton/NRAO/AUI/NSF and reproduced with permission. Abbreviation: WD, white dwarf.

3. SHOCK SIGNATURES ACROSS THE ELECTROMAGNETIC SPECTRUM

Many of the observational probes of nova ejecta discussed in Section 2.3, especially optical spectroscopy and high-resolution imaging, support the presence of distinct outflow components that span a range of velocities, which may be launched with substantial delays following TNR and which tend to accelerate as the eruption proceeds. Interactions of these components with each other—or with preexisting material surrounding the binary—will inevitably give rise to shocks. Because the typical sound speeds in the $\sim 10^4$ K nova ejecta are $c_s \approx 10$ km s $^{-1}$, whereas the ejecta velocities are $\gtrsim 100$ –1,000 km s $^{-1}$, the resulting shocks are highly supersonic. These shocks heat the plasma to millions of degrees and accelerate a small fraction of particles to relativistic speeds, generating thermal and nonthermal emission, respectively.

This section describes the observational evidence for shocks in novae. To aid our discussion, **Figure 7** shows the optical depth τ_{abs} through the ejecta as a function of photon frequency, along with the shock-powered spectral energy distribution, for conditions characteristic of a week after

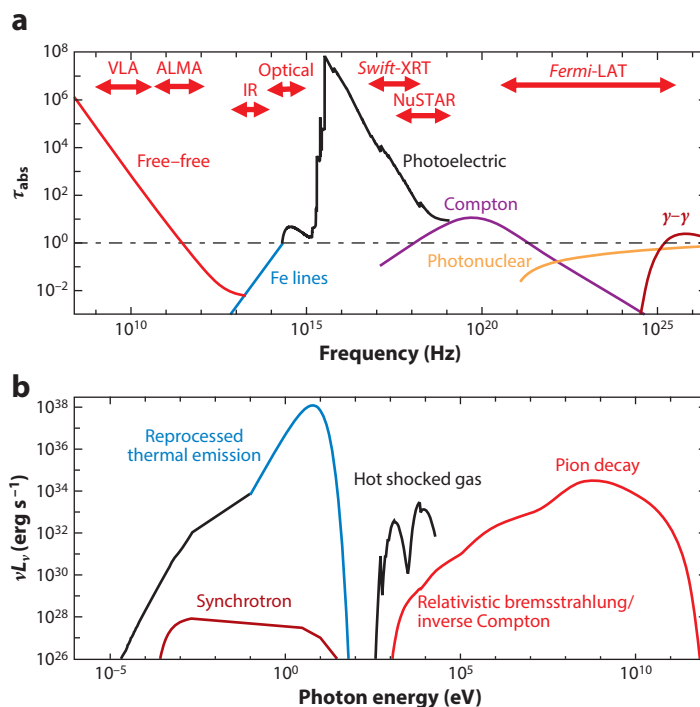


Figure 7

The processes governing opacity and emission from internal shocks in nova ejecta, from radio to gamma-ray frequencies. (a) Absorptive optical depth to the shocked region, measured outward through the nova ejecta, as a function of photon energy (frequency); $\tau_{\text{abs}} = 1$ is marked as a horizontal dot-dashed line. This is for a relatively early epoch in a nova eruption, when the gas column ahead of the shocks is $\sim 10^{25}$ cm $^{-2}$. We have assumed solar composition gas with a temperature of 10^4 K upstream of the shock, and a nova thermal luminosity of 1.5×10^{38} erg s $^{-1}$ that primarily emerges at optical wavelengths. The free-free and photoelectric opacities were calculated from CLOUDY (Ferland et al. 2013); the effective optical depth due to inelastic Compton scattering was calculated using a Monte Carlo method (I. Vurm, private communication); the optical depth due to photonuclear pair creation with the ejecta and γ - γ pair creation on the nova optical light were calculated using standard expressions (e.g., Zdziarski & Svensson 1989). (b) The spectral energy distribution of shock-powered emission in novae, accounting for absorption by the ejecta. This is based on one-dimensional internal shock simulations from Steinberg & Metzger (2020), combined with nonthermal emission from Vurm & Metzger (2018). A hadronic model is assumed for the gamma-ray emission, whereas the radio emission results from electron acceleration at the shock with an assumed efficiency of $\epsilon_e \approx 0.01$. A possible extension of the nonthermal gamma rays to \gtrsim TeV energies accessible to Cherenkov telescopes is not shown (Section 4.2). Panel adapted from Steinberg & Metzger (2020).

eruption when the shocks are still relatively deeply embedded. Places in the spectrum where τ_{abs} reaches a minimum, such as the optical-IR and GeV gamma-ray bands (Section 3.1), will reveal shock-powered emission first. However, as the shocks propagate to larger radii where τ_{abs} is smaller, shock-powered X-ray (Section 3.2) and radio (Section 3.3) emission become observable.

3.1. GeV Gamma Rays

Fermi-LAT has detected continuum gamma-ray emission in the photon energy range of ~ 0.1 –10 GeV from over a dozen Galactic novae, as summarized in **Supplemental Table 1**

Supplemental Material >

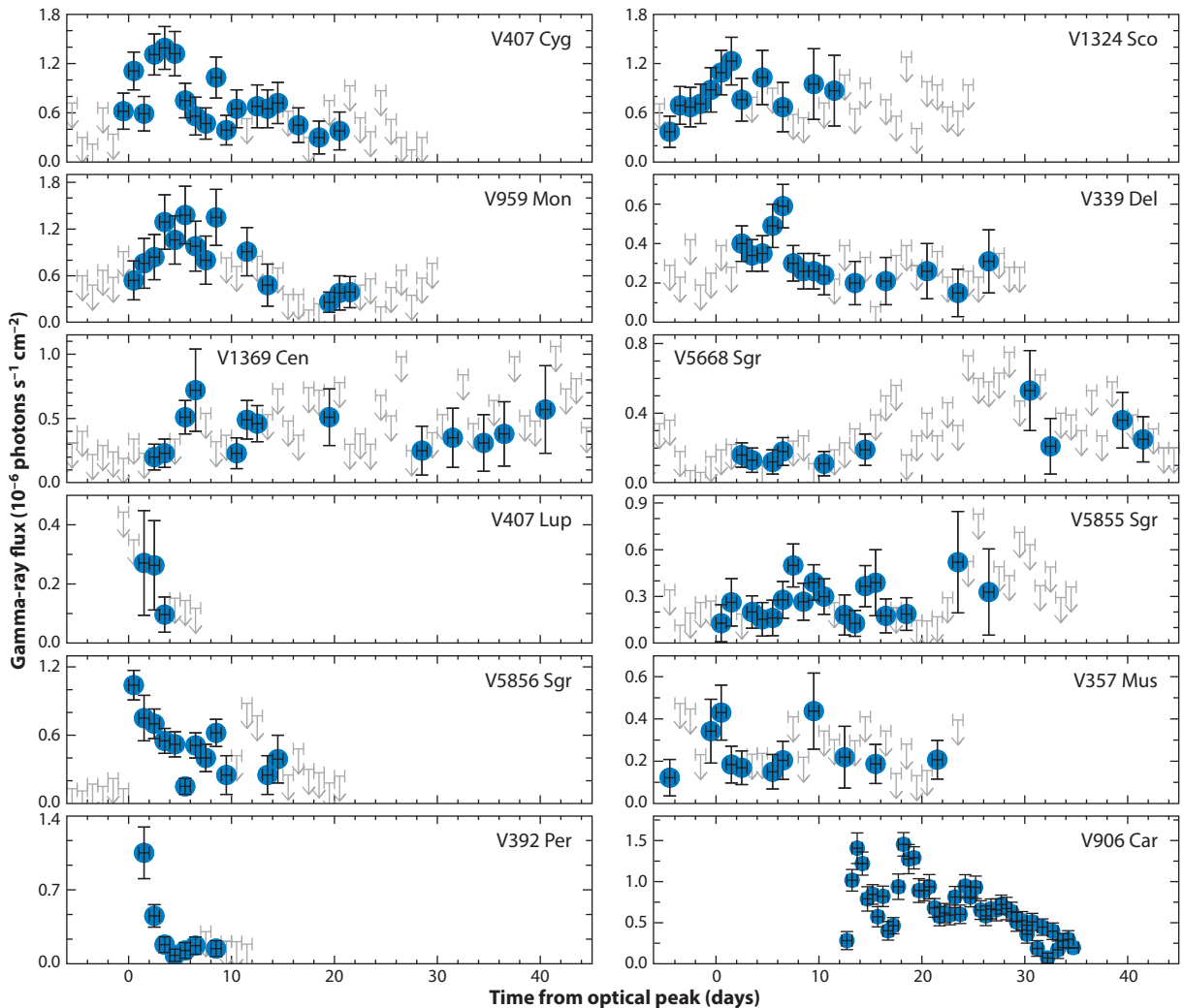


Figure 8

GeV gamma-ray light curves for 12 novae, as measured by *Fermi*-LAT (*Fermi* Large Area Telescope) for photon energies >100 MeV. Dates with $>2\sigma$ significance detections are marked as black dots, whereas 95% upper limits are plotted as gray arrows for nondetections. Plotted times are relative to optical light curve maximum, except in the case of V959 Mon, which was discovered with *Fermi* during solar conjunction; in this case $t = 0$ marks the first gamma-ray detection. Due to a solar panel issue, there are no *Fermi*-LAT data available before the gamma-ray detections of V392 Per and before or after the gamma-ray detections of V906 Car. References for the gamma-ray analysis are listed in **Supplemental Table 1**.

(Abdo et al. 2010, Ackermann et al. 2014, Cheung et al. 2016, Li et al. 2017, Aydi et al. 2020b), not to be confused with MeV gamma rays (see the sidebar titled MeV Gamma-Ray Emission from Radioactive Nuclei). The gamma rays are usually detected for a few weeks, starting around the time of optical light curve maximum (**Figure 8**).⁴

⁴Unless otherwise noted, all quoted uncertainties and plotted error bars are 1σ significance (i.e., 68%).

MeV GAMMA-RAY EMISSION FROM RADIOACTIVE NUCLEI

Gamma-ray emission at MeV photon energies has long been predicted from novae due to nuclear decay processes (Clayton & Hoyle 1974, Hernanz et al. 2002). This includes continuum and 511-keV line emission from e^-/e^+ annihilation, which are predicted to last a day or so following the TNR (Gomez-Gomar et al. 1998). Also predicted are emission lines from decays of individual isotopes such as ^7Be (478 keV) and ^{22}Na (1,275 keV) in CO and ONe novae, respectively. Unfortunately, no MeV gamma-ray emission has yet been detected from novae, despite decades of upper limits measured using the *Compton Gamma Ray Observatory* (e.g., Hernanz et al. 2000), the *Wind* satellite (Harris et al. 1991), *Swift*-BAT (*Swift* Burst Alert Telescope; Senziani et al. 2008), and INTEGRAL (*International Gamma-Ray Astrophysics Laboratory*; Siebert et al. 2018). Detection of MeV gamma-ray emission, which would provide crucial diagnostics of the TNR, requires a new generation of gamma-ray satellites with greater sensitivity [e.g., COSI (Compton Spectrometer and Imager), AMEGO (All-sky Medium Energy Gamma-ray Observatory); McEnery et al. 2019, Tomsick et al. 2019] or an extremely nearby event ($\lesssim 1$ kpc).

The gamma rays are most naturally understood as nonthermal emission from relativistic particles accelerated at shocks, likely through the process of diffusive shock acceleration (e.g., Blandford & Ostriker 1978). Typical detected gamma-ray luminosities are in the range of $L_\gamma \approx 10^{34} - 10^{36} \text{ erg s}^{-1}$, and the total energies emitted in gamma rays are $E_\gamma \approx 10^{41} - 10^{42} \text{ erg}$ (Ackermann et al. 2014, Cheung et al. 2016), which are much smaller than the bolometric outputs of novae. However, only a small fraction of the power of the shock goes into relativistic particles, and only a fraction of the energy in relativistic particles is radiated as detectable gamma rays, which means that the shock luminosities needed to produce the gamma rays can be comparable with the bolometric output, $\sim 10^{38} \text{ erg s}^{-1}$.

The high signal-to-noise gamma-ray light curves of V5856 Sgr and V906 Car reveal striking correlations between the optical and gamma-ray light curves (Li et al. 2017, Aydi et al. 2020b). For example, V906 Car showed maxima in gamma rays alongside optical flares (**Figure 9a,b**). A correlation analysis finds that the brightenings occur nearly simultaneously, with the optical lagging behind the gamma rays by $5.3 \pm 2.7 \text{ h}$ (Aydi et al. 2020b). Moderate (2σ) evidence also exists for correlated optical and gamma-ray light curves in V339 Del and V5855 Sgr, despite the poorer gamma-ray statistics (Li et al. 2017). **Figure 9c** shows the time evolution of the gamma-ray-to-blackbody luminosity ratio for a sample of gamma-ray-detected novae, revealing a range of values $L_\gamma/L_{\text{BB}} \approx 3 \times 10^{-4} - 10^{-2}$ (here L_{BB} is derived from a blackbody fit to the optical-IR spectral energy distribution, as a rough estimate of the bolometric luminosity). This ratio is relatively constant for a given nova, although in V906 Car, L_γ/L_{BB} appears to peak during optical flares (**Figure 9**). This correlation naturally results if a portion of the optical luminosity—particularly during strong flares—is reprocessed shock power (Section 4.3).

3.1.1. Do all novae emit gamma rays? The first gamma-ray detection was from the 2010 nova in the symbiotic system V407 Cyg (Abdo et al. 2010), where the shocks were interpreted as the collision between the nova ejecta and the dense wind of the Mira giant companion (e.g., Nelson et al. 2012, Martin & Dubus 2013; see the sidebar titled Shocks in Embedded Novae). However, essentially all the other high-confidence gamma-ray detections have been classical novae with main sequence companions. The winds from main sequence stars are weak, and mass transfer onto the WD is quite conservative, so there is too little circumbinary material to generate powerful shocks. Therefore, the shocks must be internal to the nova ejecta. In fact, Martin et al. (2018) find that internal shocks could even give rise to much of the gamma-ray emission from V407 Cyg without needing to invoke any interaction with the giant wind.

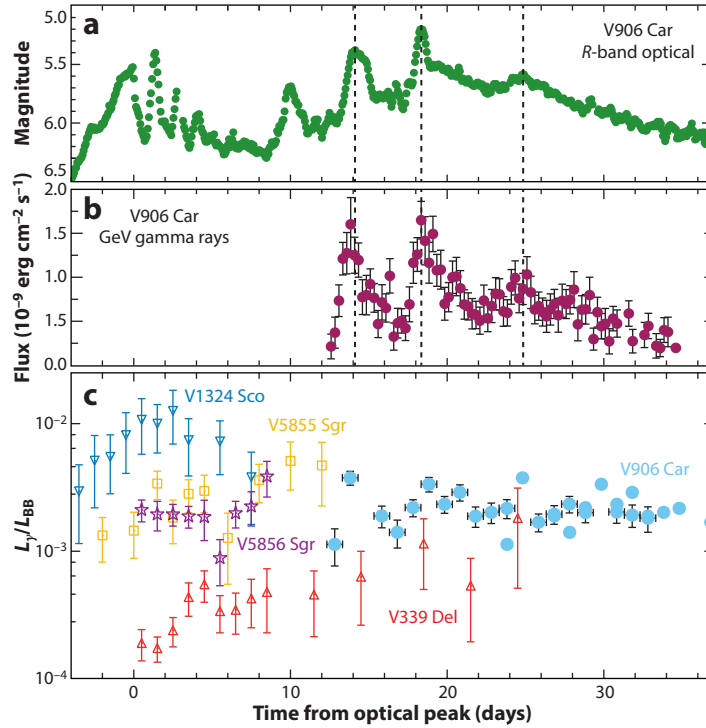


Figure 9

The correlated flares in the optical and gamma-ray light curves of nova V906 Car. (a) Optical *R*-band light curve of V906 Car, observed at high cadence and precision using the BRITE (*BR*ight-*star* *T*arget *E*xplorer) constellation of nanosatellites. The three optical flares that occurred during the *Fermi* observations are marked with black dotted lines. (b) *Fermi*-LAT (*Fermi* Large Area Telescope) gamma-ray light curve (>100 MeV) of V906 Car. Unfortunately, *Fermi*-LAT was not operational outside the times with plotted data. (c) The ratio of gamma-ray-to-blackbody luminosity as a function of time. V906 Car (blue circles) is compared with other *Fermi*-detected novae: V5856 Sgr (purple stars), V5855 Sgr (gold squares), V339 Del (red triangles), and V1324 Sco (blue triangles). Figure adapted from Aydi et al. (2020b), with L_γ/L_{BB} values from Metzger et al. (2015) and Li et al. (2017).

Although it is plausible that all novae host internal shocks and emit gamma rays, the rate of gamma-ray-detected novae ($\sim 1 \text{ year}^{-1}$) is only a small fraction of the optical discovery rate ($\sim 5\text{--}10 \text{ year}^{-1}$). The *Fermi* detections tend to be of relatively nearby novae ($\sim 2\text{--}5 \text{ kpc}$) and of marginal significance (**Supplemental Table 1** and **Supplemental Figure 1**), suggesting that many novae go undetected due to the limited flux sensitivity of *Fermi*-LAT.

Nevertheless, the hypothesis that all novae possess the same gamma-ray luminosity can be firmly excluded, as Franckowiak et al. (2018) showed that the gamma-ray luminosities span at least two orders of magnitude. However, to date no clear correlations have been found between gamma-ray luminosity and nova properties like t_2 (Franckowiak et al. 2018; **Supplemental Figure 1**). There is a suggestion of an anticorrelation between the total gamma-ray energy emitted and the duration of the gamma-ray emission (Cheung et al. 2016). In an internal shock model (Section 4), we expect the gamma-ray luminosity to increase with ejecta mass and velocity, and future studies should test for such relationships.

SHOCKS IN EMBEDDED NOVAE

Shocks have long been expected and observed in embedded novae, arising from the interaction of the nova ejecta with the giant companion's wind. The most famous example of these shocks is in RS Oph, which showed bright radio synchrotron emission and hot X-ray emitting plasma during its 1985 and 2006 eruptions (e.g., Hjellming et al. 1986, O'Brien et al. 2006, Sokoloski et al. 2006). In the past decade, the sample of well-observed embedded novae has grown dramatically, establishing a class of novae with similar radio and X-ray signatures to RS Oph (e.g., Nelson et al. 2012, Linford et al. 2015, Delgado & Hernanz 2019, Orio et al. 2020).

These external shocks with circumstellar material in symbiotic systems suffer significantly less absorption than shocks in classical novae that are presumably internal to the ejecta. For example, V407 Cyg is the only gamma-ray-detected nova that showed 1–10-keV X-rays simultaneous with gamma rays (see Section 3.2). The gamma-ray luminosities of these embedded novae may be consistent with those of classical novae, but embedded novae are rarer and so tend to be discovered at larger distances (**Supplemental Figure 1**). V407 Cyg is the only one with a high-significance *Fermi*-LAT detection, whereas three other embedded novae—V745 Sco, V1535 Sco, and V3890 Sgr—show hints of gamma-ray emission (Cheung et al. 2014, Buson et al. 2019), but at $<3\sigma$ significance (Franckowiak et al. 2018; K. Li, private communication).

Supplemental Material >

3.1.2. Leptonic or hadronic gamma rays? The time-integrated *Fermi*-LAT gamma-ray spectra of novae are typically modeled as a power law with an exponential cutoff, $dN/dE \propto E^{-\Gamma} \exp(-E/E_c)$, where E_c is the cutoff energy at high photon energies and Γ is the photon index. The best-fit values for individual novae span the range $\Gamma \approx 1.3$ –2.3 (**Supplemental Table 1**). Although some novae are best fit with $E_c \approx 2$ –8 GeV, others are well described by a simple power law with no cutoff energy. Franckowiak et al. (2018) show that all pre-2016 novae are consistent with a single universal spectrum with $\Gamma = 1.71 \pm 0.08$ and $E_c = 3.2 \pm 0.6$ GeV (excluding the embedded nova V407 Cyg from the sample, which exhibited a significantly harder photon index than the classical novae). Two more recent novae, V5856 Sgr and V906 Car, were particularly bright in gamma rays, enabling high-fidelity spectral fits and yielding values of $\Gamma \approx 1.7$ –1.9 and $E_c \approx 3$ –8 GeV, which is broadly consistent with the universal spectrum (**Figure 10**; Li et al. 2017, Aydi et al. 2020b). Both spectra also show a deficit below a few hundred MeV compared to a simple power law, the origin of which we return to below. Although the evidence for a high-energy spectral cutoff is marginal based on the LAT data alone, a deep upper limit on the 0.1–10-TeV emission from V339 Del, obtained with MAGIC (Major Atmospheric Gamma-ray Imaging Cherenkov) simultaneously with a LAT detection, necessitates a spectral cutoff or steepening in the 10–100-GeV range (Ahnen et al. 2015).

Physical models for the gamma-ray emission divide into hadronic and leptonic scenarios, depending on whether the emitting particles are primarily relativistic ions or electrons. In leptonic scenarios, the accelerated electrons emit gamma rays via bremsstrahlung emission due to interaction with ambient protons and electrons:

$$e^{\pm} + p \rightarrow e^{\pm} + p + \gamma; \quad e^{\pm} + e^{-} \rightarrow e^{\pm} + e^{-} + \gamma. \quad 7.$$

Leptonic gamma rays are also produced via inverse Compton (IC) scattering of optical photons:

$$e^{\pm} + \gamma_{\text{opt}} \rightarrow e^{\pm} + \gamma. \quad 8.$$

In the hadronic scenario, ions collide with ambient ions such as protons, producing pions that decay into gamma rays as

$$\begin{aligned} p + p &\rightarrow \pi^0 \rightarrow \gamma + \gamma \\ &\rightarrow \pi^{\pm} \rightarrow \mu^{\pm} + \nu_{\mu} \rightarrow e^{\pm} + \nu_e + \nu_{\mu}, \end{aligned} \quad 9.$$

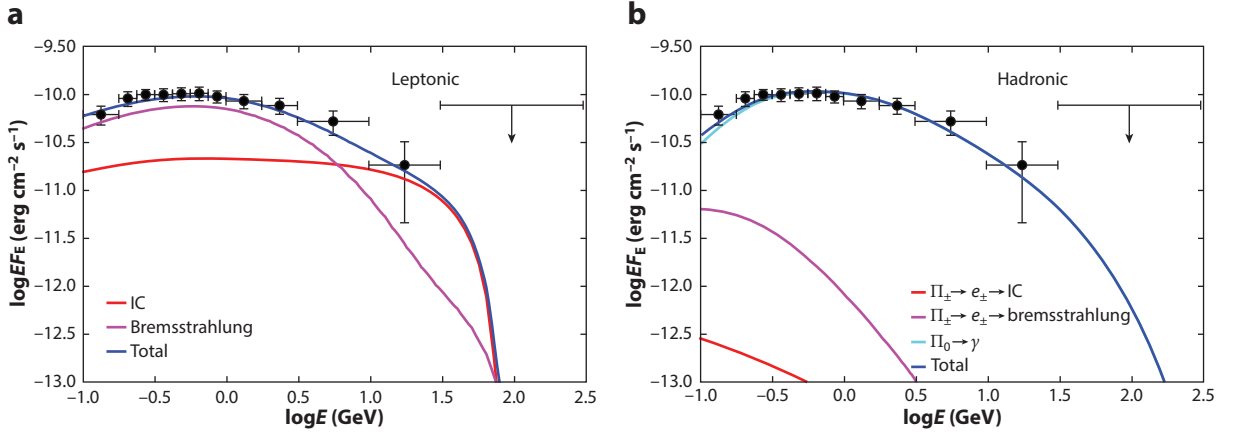


Figure 10

The *Fermi*-LAT (*Fermi* Large Area Telescope) GeV spectral energy distribution of V5856 Sgr (*black points*), compared with gamma-ray emission models (*solid lines*). (a) Leptonic model fit, which assumes electrons accelerated at the shock with a $q = 1.8$ momentum spectrum. (b) Hadronic model fit, which assumes protons accelerated at the shock with $q = 2.4$. Both models assume a maximum particle energy of 100 GeV. Although both models describe the data well, the hadronic model is favored because it more robustly accounts for the break in the spectrum at low energies set by the pion creation threshold ($m_\pi c^2 = 135$ MeV). The leptonic model also has several efficiency issues. Figure adapted from Li et al. (2017). Abbreviation: IC, inverse Compton.

where $\approx 1/3$ and $\approx 2/3$ of the inelastic p - p collisions go through the π^0 and π^\pm channels, respectively (Kelner & Aharonian 2008). Leptons produced in the π^\pm channel can then produce gamma rays through the bremsstrahlung and IC processes, although the π^0 channel is expected to dominate the gamma-ray luminosity (Figure 10). As a useful approximation, generating a photon of energy E requires a proton of energy $10E$, so to produce gamma rays up to ~ 10 GeV requires particle acceleration up to ~ 100 GeV (Kelner & Aharonian 2008). In leptonic scenarios, the photon energy is comparable with that of the primary electron, so particle acceleration up to ~ 10 GeV is sufficient.

Figure 10 shows the *Fermi*-LAT gamma-ray spectrum of V5856 Sgr fit using the leptonic or hadronic models. Relativistic particles (electrons or protons, respectively) are assumed to be accelerated at the shock with an energy spectrum $dN/dp \propto p^{-q}$, where $p = \beta\gamma_p$ is the particle momentum, whereas γ_p and $\beta = v/c$ are the Lorentz factor and particle velocity, respectively. As described by Vurm & Metzger (2018), the thermodynamic evolution of the particles and their secondaries, including various emission and energy-loss processes (e.g., IC, bremsstrahlung, synchrotron, Coulomb scattering), can be followed downstream behind the shock to calculate the emerging radiation spectrum. The best fit models require $q \lesssim 2$ in the leptonic scenario and $q \approx 2.4$ in the hadronic scenario.

The hadronic model naturally predicts a low-energy spectral turnover as observed, near the pion rest mass energy of $m_\pi c^2 \approx 135$ MeV that sets the threshold energy for p - p interactions (Equation 9). A similar break can appear in leptonic models due to Coulomb interactions competing for the energy of low- γ_p electrons, but the location of this feature must be fine-tuned because it depends on the optical radiation field.

Other considerations also favor the hadronic scenario. A strong magnetic field near the shock is required to confine and accelerate the particles (Section 4.2; Equation 18). However, electrons embedded in such a magnetic field cool quickly via synchrotron emission, suppressing their contribution to the gamma-ray luminosity and, hence, making leptonic emission channels inefficient

(Li et al. 2017). The shallow spectral indices $q < 2$ required in leptonic models are also in tension with the values of $q > 2$ predicted by diffusive shock acceleration (Blandford & Ostriker 1978) or inferred from observations of other astrophysical shocks such as SN remnants (Green 2019). Nevertheless, such low values of q are required for the leptonic scenario to avoid an energy crisis: Electron Lorentz factors $\gamma_p \gtrsim 10^3\text{--}10^4$ are needed to produce the gamma rays, but for $q > 2$ the total energy is dominated by low- γ_p electrons. Leptonic models also predict an extension of the LAT gamma-ray emission down to the hard X-ray band (Vurm & Metzger 2018), which is in tension with NuSTAR (*Nuclear Spectroscopic Telescope*) observations of V906 Car (Aydi et al. 2020b). Naively, the correlation between gamma-ray and optical light curves (**Figure 9**) could be interpreted within leptonic scenarios as time-variable IC emission driven by the fluctuating optical target photon background (Equation 8). However, electrons with sufficient energy to generate the observed gamma rays are already in a deeply fast-cooling regime (Metzger et al. 2015), and hence the nonthermal energy is converted to gamma-ray luminosity with maximal efficiency, regardless of optical luminosity. Models that systematically fit nova gamma-ray spectra by following particle acceleration and emission within fully time-dependent, one-dimensional hydrodynamical models also conclude that the hadronic channel dominates (Martin et al. 2018).

3.2. X-Rays from Hot Plasma

In addition to the supersoft X-ray emission that arises from residual nuclear burning on the WD surface (Section 2.1.2), novae often emit harder X-rays ($\gtrsim 1$ keV; e.g., Lloyd et al. 1992, Krautter et al. 1996, Mukai & Ishida 2001, Orio et al. 2001, Mukai et al. 2008). The harder component of the X-ray spectrum is distinct from the supersoft X-ray component and is well modeled by free-free emission and atomic lines from an optically thin $10^7\text{--}10^8$ K thermally emitting plasma. This X-ray emission also evolves differently from the supersoft X-ray component and is visible both before and after the supersoft phase. These observations imply a distinct physical origin for the harder X-ray component and lead to an association with shocks.

A strong shock generated by the collision of two flows at velocity v heats gas to a characteristic temperature (Landau & Lifshitz 1959),

$$kT_{\text{sh}} \simeq \frac{3}{16} \mu m_p v^2 \approx 1.2 \text{ keV} \left(\frac{v}{10^3 \text{ km s}^{-1}} \right)^2, \quad 10.$$

where we have taken $\mu = 0.62$ for the mean molecular weight of solar composition material. Thermal plasma emission models of nova X-ray spectra imply temperatures⁵ $kT \approx 1\text{--}10$ keV (Mukai et al. 2008), pointing to shock velocities $\approx 1,000\text{--}3,000$ km s^{−1}. These velocities are broadly similar to the difference in velocity between the fast and slow components inferred from optical spectroscopy (e.g., **Figure 4**), suggesting that collisions between these two components may be the origin of the shocks.

Figure 11 shows 1–10-keV X-ray light curves for eight gamma-ray-detected novae, primarily obtained with the X-ray Telescope (XRT) on *Swift*. The shock-generated X-ray emission typically becomes detectable about a month after eruption and reaches a peak luminosity $L_X \approx 10^{33}\text{--}10^{34}$ erg s^{−1}, which is $\lesssim 10^{-4}$ of the bolometric luminosity and $\lesssim 1\%$ of the gamma-ray luminosity. With the exception of embedded nova V407 Cyg (see the sidebar titled Shocks in Embedded Novae), no X-ray emission has been observed in the 1–10-keV band simultaneous with a *Fermi*-LAT detection (Gordon et al. 2021).

⁵In X-ray astronomy, temperatures are often given in units of keV, calculated as kT , where k is the Boltzmann constant (1.38×10^{-16} erg/K). 1 keV = 1.16×10^7 K.

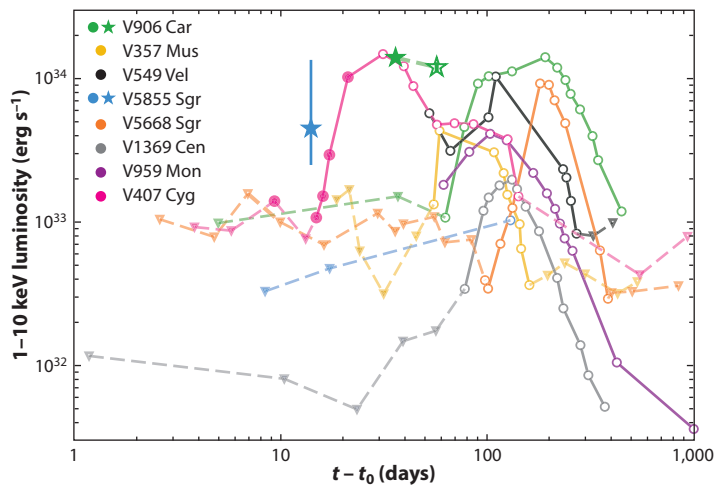


Figure 11

Hard X-ray (1–10 keV) light curves for eight gamma-ray-detected novae. If an observation is concurrent with gamma-ray detection, the symbol is filled. Detections with *Swift*-XRT (*Swift* X-ray Telescope) are marked as circles and are not corrected for internal absorption by the nova ejecta, whereas *Swift* upper limits are denoted with triangles. NuSTAR (*Nuclear Spectroscopic Telescope*) detections are marked as stars and are corrected for internal absorption (Nelson et al. 2019, Sokolovsky et al. 2020b). Assumed distances are listed in **Supplemental Table 1**. Figure created by A. Gordon.

Supplemental Material >

What is responsible for the delayed rise of the harder X-rays? In some cases, shock formation itself may be delayed due to a lag between the TNR and expansion of the bulk ejecta (e.g., Chomiuk et al. 2014b, Nelson et al. 2020). However, the early onset of the gamma-ray emission in many novae indicate that some shocks occur well before the first X-ray detection. In such cases, the lack of X-rays is attributed to absorption that delays detectable X-ray emission until the shocks emerge into an environment with $\tau_{\text{abs}} \lesssim 1$ (**Figure 7**). This is supported by observations of the hard X-ray spectrum, which shows that the absorbing column density decreases with time after eruption (e.g., Mukai & Ishida 2001).

The shocks can be detected earlier in harder X-rays ($\gtrsim 10$ keV), because of the decreasing ejecta opacity at higher photon energies (**Figure 7**). NuSTAR has detected hard (3–79 keV) X-rays from three novae simultaneously with *Fermi*-LAT (see **Figure 11**) (Nelson et al. 2019; Sokolovsky et al. 2020a,b). The hard X-ray emission is best described with optically thin thermal plasma models with inferred temperatures, $kT \approx 6\text{--}11$ keV, and absorbing column densities of gas $\Sigma \approx 10^{23}\text{--}10^{24}$ cm $^{-2}$, that will easily prevent detection by *Swift*-XRT. This X-ray emission is too soft to be a low-energy extension of the nonthermal *Fermi*-LAT emission (Vurm & Metzger 2018).

Even after correcting for absorption, the observed X-ray luminosities, $L_X \approx 10^{33}\text{--}10^{34}$ erg s $^{-1}$, are ~ 100 times lower than the gamma-ray luminosities detected concurrently and orders of magnitude smaller than theoretically expected (Nelson et al. 2019, Sokolovsky et al. 2020b). One solution to this mystery is that the NuSTAR-detected shocks are distinct from those powering the gamma rays (or arise from only a portion of the shock front), in which case the latter must be even more deeply embedded. Alternatively, the X-ray emission from radiative shocks in novae are significantly suppressed relative to the naive expectation of one-dimensional models (Section 4.4). Regardless, this implies the bulk of the shock luminosity is reprocessed into UVOIR emission (Section 4.3).

Column density of gas (Σ): integral of gas number density $n(l)$ along the path length between the source and observer l , such that $\Sigma = \int n(l)dl$

3.3. Nonthermal Radio Emission

In addition to the thermal radio emission discussed in Section 2.3.4, many novae also display non-thermal synchrotron radio emission that must also be associated with shocks (Taylor et al. 1987; Krauss et al. 2011; Chomiuk et al. 2014a; Weston et al. 2016a,b; Finzell et al. 2018). The primary evidence for synchrotron emission is the high radio surface brightness sometimes observed in the first months of eruption. Radio surface brightness is characterized by the brightness temperature,

$$\frac{T_B}{K} = 1,200 \left(\frac{S_\nu}{\text{mJy}} \right) \left(\frac{\nu}{\text{GHz}} \right)^{-2} \left(\frac{\theta}{\text{arcsec}} \right)^{-2}, \quad 11.$$

where S_ν is the observed flux density at frequency ν . The angular diameter of the source, θ , can be estimated if the distance, expansion velocity, and expansion time are known, or by directly resolving the emission with, e.g., very long baseline interferometry. On timescales of $\lesssim 100$ days, brightness temperatures inferred for novae are often $\gtrsim 10^5$ – 10^6 K, whereas the maximum T_B of a photoionized thermal emitter is $\sim 10^4$ K. This implies an additional emission source, from either shock-heated thermal gas or shock-induced synchrotron emission. Pioneering work by Taylor et al. (1987) modeled the early radio emission from QU Vul as shocks between a high-velocity outflow and slower earlier ejecta, which is similar to the scenario emphasized by this review.

X-ray observations of novae reveal that shocks can heat gas to temperatures $T_{\text{sh}} \gtrsim 10^7$ K, exceeding the observed radio brightness temperatures (Section 3.2). However, the column of hot gas available at any time is generally insufficient to explain the observed radio fluxes, as constrained by both X-ray observations (e.g., Weston et al. 2016b) and modeling of nova shocks (Metzger et al. 2014, Vlasov et al. 2016). The high radio luminosities and brightness temperatures instead favor synchrotron radiation from relativistic electrons gyrating in magnetic fields behind shocks.

Figure 12 shows an example of this high- T_B early emission during the first ~ 100 days of the gamma-ray-detected nova V1324 Sco (Finzell et al. 2018). Several novae have shown similar distinct, early-time radio maxima that rival the later, $\sim 10^4$ K thermal maxima in flux (Taylor et al. 1987, Weston et al. 2016b). Other novae show more subtle excesses at early times that nevertheless cannot be explained by $\sim 10^4$ K expanding ejecta (Chomiuk et al. 2014a, Aydi et al. 2020b).

Optically thin synchrotron radiation is characterized by a broadband power-law spectrum, $F_\nu \propto \nu^\alpha$, where the spectral index α is determined by the energy spectrum of the emitting relativistic electrons ($dN/d\gamma_p \propto \gamma_p^{-p}$) as $\alpha = -(p - 1)/2$. Typically, $\alpha \approx -0.5$ to -1 is observed for synchrotron emission in relativistic shocks, implying $p = 2$ – 3 (e.g., SNe, SN remnants; Weiler et al. 2002, Green 2019). However, such a negative spectral index rarely describes the high brightness temperature emission in novae (**Figure 12**). Because of free-free absorption by the nova ejecta, as with the X-rays, early radio emission from the shock is absorbed (**Figure 7**). The flux at lower frequencies will be preferentially suppressed, leading to a rising spectrum ($\alpha > 0$). For a single temperature-absorbing medium with a fixed radial column, the low-frequency spectral cutoff is exponential. However, if absorption is important in the multitemperature postshock cooling layer where the radio emission is being produced, this causes the spectrum of radiation emerging from the shock (i.e., before absorption by the rest of the ejecta) to be flatter than that of optically thin synchrotron emission (Metzger et al. 2014).

Radio emission (unlike gamma rays or X-rays) can be imaged at very high angular resolutions ($\sim 10^{-3}$ arcsec), allowing long baseline interferometers to reveal where the shocks are taking place in the nova ejecta. In V959 Mon, synchrotron knots were observed to hug the relatively slow thermal ejecta in the orbital plane and hypothesized to originate in the collision of fast polar ejecta with this slower equatorial material (**Figure 6**; Chomiuk et al. 2014a). Although the synchrotron emission over most of the shock surface is invisible due to free-free absorption (**Figure 7**), portions of the shock front may peek out at low column densities and appear as synchrotron knots.

Spectral index (α): the spectral energy distribution of radio continuum emission, described as $S_\nu \propto \nu^\alpha$, where S_ν is the flux density at frequency ν

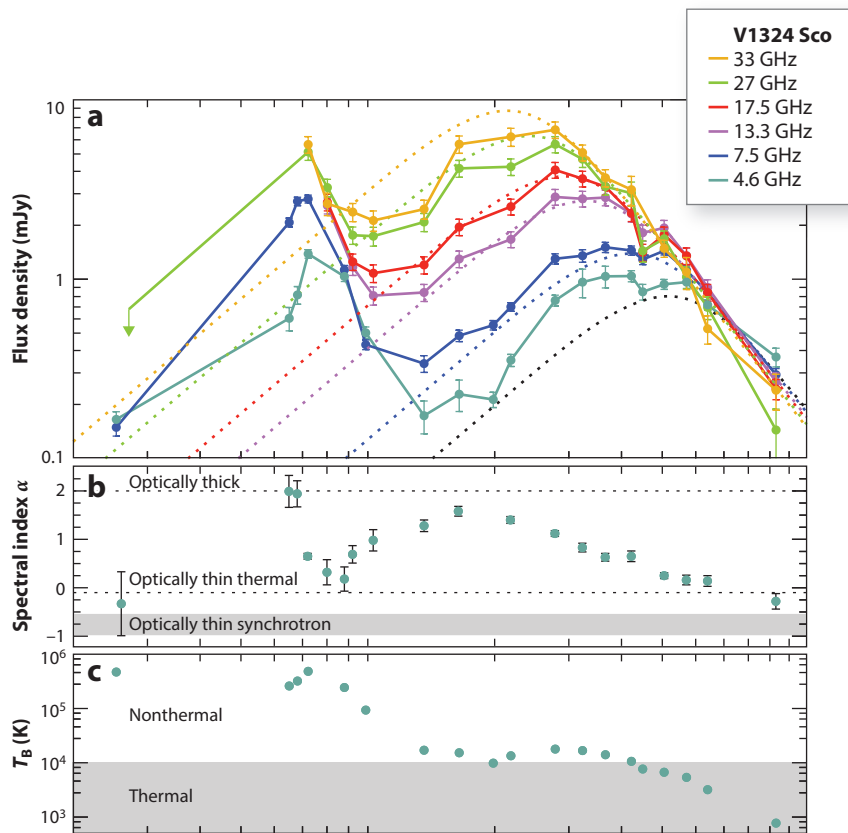


Figure 12

The radio light curve of gamma-ray-detected nova V1324 Sco shows two distinct maxima. The first is likely associated with synchrotron emission, showing $T_B \approx 5 \times 10^5$ K and peaking on day ~ 70 . The second is powered by thermal emission with $T_B \approx 10^4$ K, and peaks ~ 1 year after eruption. (a) Multifrequency light curves observed at 4.6–33 GHz with the Karl G. Jansky Very Large Array (VLA). A simple “Hubble flow” fit (*dashed lines*) demonstrates that the second maximum is reasonably well fit as expanding thermal ejecta (e.g., Seaquist & Bode 2008). (b) The spectral index of the radio emission as a function of time. (c) The brightness temperature of the 4.6 GHz emission, assuming an expansion velocity of $1,000 \text{ km s}^{-1}$ starting on day 0. Figure adapted with permission from Finzell et al. (2018); copyright AAS.

4. THEORY OF INTERNAL RADIATIVE SHOCKS IN CLASSICAL NOVAE

The previous sections have described observational evidence in support of internal shocks as common—if not ubiquitous—features of nova outflows. The most powerful shocks likely take place in the orbital plane of the binary, when an initial slow, equatorially focused outflow of velocity v_s of about a few hundred km s^{-1} is hit from behind by a fast, more spherically symmetric outflow of velocity v_f of about a few thousand km s^{-1} (Section 2.3.5). The slow flow may arise from the outer Lagrange point as the nova envelope first encases the binary orbit (Section 2.2.3), whereas the fast flow may be a radiation-driven wind (Section 2.2.2). Whatever their origin, multiple episodes of such slow–fast transitions and collisions can take place in a single nova event, giving rise to distinct optical and gamma-ray flares (Figure 9).

As we demonstrate in Section 4.1, in the first weeks of many nova eruptions, the internal shocks are likely to be radiative—that is, they will radiate away their thermal energy faster than it can be reconverted back into kinetic energy via PdV expansion. This section presents the basic theory of internal radiative shocks and how they can explain observations of novae. We start by considering a one-dimensional toy model of the shock interaction (Section 4.1) and then outline particle acceleration (Section 4.2) and radiative signatures (Section 4.3). In Section 4.4, we discuss multidimensional effects not captured by a one-dimensional picture. **Figure 13** shows a schematic diagram of the shock structure and its electromagnetic emissions.

4.1. Dynamics of Dual Radiative Shocks

We model both the fast and slow flows as freely expanding winds with constant mass-loss rates of \dot{M}_f and \dot{M}_s . The fast outflow is assumed to be spherically symmetric, whereas the slow outflow is confined to a fractional solid angle $f_\Omega < 1$ in the orbital plane. The number densities in the outflows at radius r are

$$n_s = \frac{\dot{M}_s}{4\pi f_\Omega r^2 \mu m_p v_s} \quad \text{and} \quad n_f = \frac{\dot{M}_f}{4\pi r^2 \mu m_p v_f}. \quad 12.$$

Given that novae typically eject $M_{\text{ej}} \approx 10^{-5}$ – $10^{-4} M_\odot$ over timescales of weeks (Section 2), characteristic values are $\dot{M}_f, \dot{M}_s \approx 10^{-6}$ – $10^{-5} M_\odot \text{ week}^{-1}$. This model is sufficient to explain the basic physics, even if the actual flows are more complex.

We define $t_{\text{sh}} = 0$ as the time when the fast outflow turns on abruptly and collides with the earlier slow flow, driving a dual forward–reverse shock structure outward in radius. As verified below, over at least the first few weeks of interaction, both shocks are likely to be radiative and, hence, radiate all the energy dissipated by the collision. The cooled gas behind the shocks will collect in a thin, cold (and clumpy; Section 4.4) shell of mass M_{cs} and velocity v_{cs} . Because the postshock gas cools so quickly as to exert effectively no pressure, the velocities of the forward and reverse shocks in the WD frame also equal v_{cs} . The mass and momentum of the cold shell thus grow, as (e.g., Metzger et al. 2014)

$$\frac{dM_{\text{cs}}}{dt_{\text{sh}}} = f_\Omega \dot{M}_f \left(\frac{v_f - v_{\text{cs}}}{v_f} \right) + \dot{M}_s \left(\frac{v_{\text{cs}} - v_s}{v_s} \right), \quad 13.$$

$$\frac{d}{dt_{\text{sh}}} (M_{\text{cs}} v_{\text{cs}}) = f_\Omega \dot{M}_f (v_f - v_{\text{cs}}) + \dot{M}_s (v_{\text{cs}} - v_s). \quad 14.$$

Numerical integration of these equations reveals an approximate steady state ($dv_{\text{cs}}/dt_{\text{sh}} \approx 0$) in cases when the shell gains most of its momentum from the fast wind, but most of its mass from the slow ejecta shell. In this case,

$$v_{\text{cs}} \approx \left(\frac{\dot{M}_f}{\dot{M}_s} \frac{v_f f_\Omega}{v_s} \right)^{1/2} v_s \equiv \xi v_s, \quad 15.$$

yielding $\xi \sim \text{few}$ for typical parameters (e.g., $\dot{M}_f \approx \dot{M}_s$; $f_\Omega \approx 0.3$; $v_f \approx 10v_s$). The radius of the cold shell thus grows as

$$R_{\text{cs}} \approx v_{\text{cs}} t_{\text{sh}} \approx 5.4 \times 10^{13} \text{ cm} \left(\frac{\xi}{3} \right) \left(\frac{v_s}{300 \text{ km s}^{-1}} \right) \left(\frac{t_{\text{sh}}}{1 \text{ week}} \right), \quad 16.$$

and must lie outside the binary orbit in classical novae on the timescales of the observed gamma rays (Equation 4).

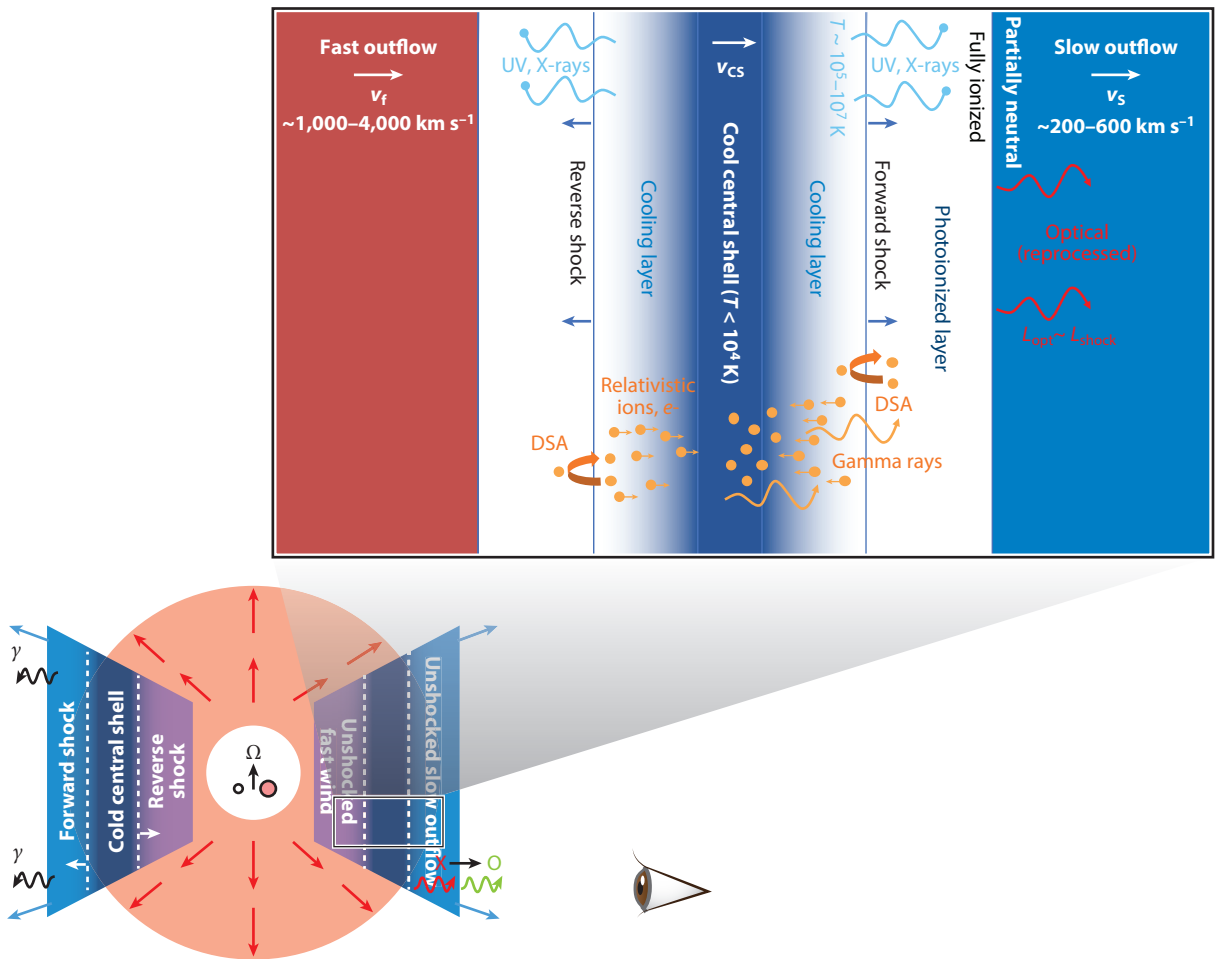


Figure 13

Anatomy of a radiative internal shock, zoomed in on a small patch where a fast outflow of velocity v_f collides with a slower outflow of velocity v_s . A forward shock is driven forward into the slow outflow, whereas a reverse shock propagates backward into the fast outflow. Each shock heats the gas immediately behind it to temperatures $\gtrsim 10^6$ – 10^7 K, but the gas quickly cools via UV and X-ray emission in a narrow cooling layer. Due to thin-shell instabilities, the cool gas collects in a thin central shell that is corrugated on a scale of approximately its thickness (not drawn here; see **Supplemental Video 1**). UV and X-rays from the shocks are absorbed by the partially neutral upstream flow or by the cool central shell, and their emission is reprocessed into the UVOIR and contributes to the optical light curve (Section 4.3). Most of the relativistic particles (electrons and ions; *orange dots*) accelerated at the shocks are advected into the cold central shell, where they emit gamma rays. Both optical and gamma-ray production occur with little delay, resulting in correlated optical and gamma-ray light curves. Figure adapted from Metzger et al. (2015). Abbreviations: DSA, diffusive shock acceleration; UVOIR, ultraviolet–optical–infrared.

Supplemental Material >

Expansion time

(t_{exp}): the timescale on which the radius of the central shell and shocks will double,
 $t_{\text{exp}} = R_{\text{cs}}/v_{\text{cs}}$

We can now confirm that the shocks are radiative, if we compare the expansion time, $t_{\text{exp}} = R_{\text{cs}}/v_{\text{cs}} = t_{\text{sh}}$, to the radiative cooling time of the shocked gas, $t_{\text{cool}} = 3kT_{\text{sh}}/[2n_{\text{sh}}\Lambda(T = T_{\text{sh}})]$. Here, T_{sh} (Equation 10 with $v = v_{\text{cs}}$) and n_{sh} are the temperature and density of the shocked gas, respectively, and $\Lambda \approx 3 \times 10^{-27} (T_{\text{sh}}/\text{K})^{1/2} \text{ erg cm}^3 \text{ s}^{-1}$ is the free–free cooling rate (although line cooling can dominate for low shock velocities).

Taking $n_{\text{sh}} = 4n_{\text{f}}(r = R_{\text{cs}})$ for the reverse shock ($n_{\text{sh}} \approx 10^{10} \text{ cm}^{-3}$ for fiducial parameters at $t_{\text{sh}} \sim 1$ week, where the factor of four comes from the shock jump conditions), we have

$$\frac{t_{\text{cool}}}{t_{\text{exp}}} \approx 0.1 \left(\frac{\xi}{3} \right)^2 \left(\frac{\dot{M}_{\text{f}}}{10^{-5} M_{\odot} \text{ week}^{-1}} \right)^{-1} \left(\frac{v_{\text{f}}}{2,000 \text{ km s}^{-1}} \right)^2 \left(\frac{v_{\text{s}}}{300 \text{ s}^{-1}} \right)^2 \left(\frac{t_{\text{sh}}}{1 \text{ week}} \right). \quad 17.$$

The reverse shock is thus radiative ($t_{\text{cool}} \ll t_{\text{exp}}$) over the first few weeks for fiducial parameters. The forward shock can remain radiative for even longer, owing to its lower velocity and higher density of the slow flow.

4.2. Relativistic Particle Acceleration

A small fraction of the charged particles that enter the shocks can be accelerated to relativistic velocities by the diffusive shock acceleration process, if a strong and turbulent magnetic field deflects particles back and forth across the shock front in the locally converging flow (e.g., Blandford & Ostriker 1978; see Blasi 2019 for a review). The magnetic field strength near the reverse shock can be estimated from equipartition arguments, $B_{\text{sh}} = v_{\text{cs}} \sqrt{6\pi \epsilon_B \mu_p m_p n_{\text{f}}}$, yielding $B_{\text{sh}} \approx 3 \text{ G}$ for fiducial parameters ~ 1 week after eruption (Section 4.1) and an efficiency of magnetic field amplification $\epsilon_B = 0.01$. A promising mechanism for amplifying the magnetic field in nonrelativistic shocks is the cosmic-ray current-driven instability (Bell 2004), where streaming relativistic particles amplify fluctuations in the magnetic field. Numerical plasma simulations of this instability imply $\epsilon_B \gtrsim 10^{-4}$ – 10^{-2} , depending on the ion acceleration efficiency and the Alfvén Mach number (e.g., Caprioli & Spitkovsky 2014b).

Up to what energy can particles be accelerated? In diffusive shock acceleration, as cosmic rays gain greater and greater energy E , they can diffuse back to the shock from a greater downstream distance, z , because of their larger gyroradii $r_g = E/eB_{\text{sh}}$. The maximum energy to which particles are accelerated before escaping, E_{max} , is found by equating the upstream diffusion time $t_{\text{diff}} \sim D/v_{\text{cs}}^2$ to the minimum of various particle loss timescales. These include the downstream advection timescale $t_{\text{adv}} \sim z_{\text{acc}}/v_{\text{cs}}$, where z_{acc} is the width of the acceleration zone, and (in hadronic scenarios) the pion creation timescale $t_{\pi} = (n_{\text{sh}} \sigma_{\pi} c)^{-1}$, where $\sigma_{\pi} \sim 2 \times 10^{-26} \text{ cm}^2$ is the inelastic cross section for p – p interactions (Kamae et al. 2006). For example, equating $t_{\text{diff}} = t_{\text{adv}}$ and taking $D \approx r_g c/3$ as the diffusion coefficient (Caprioli & Spitkovsky 2014b), one obtains

$$E_{\text{max}} \sim \frac{3eB_{\text{sh}}v_{\text{cs}}z_{\text{acc}}}{c} \approx 4 \times 10^5 \text{ GeV} \left(\frac{z_{\text{acc}}}{R_{\text{cs}}} \right) \left(\frac{\xi}{3} \right) \times \left(\frac{\epsilon_B}{10^{-2}} \right)^{1/2} \left(\frac{\dot{M}_{\text{f}}}{10^{-5} M_{\odot} \text{ week}^{-1}} \right)^{1/2} \left(\frac{v_{\text{f}}}{2,000 \text{ km s}^{-1}} \right)^{-1/2} \left(\frac{v_{\text{s}}}{300 \text{ km s}^{-1}} \right) \quad 18.$$

if we normalize the radial width of the acceleration zone, z_{acc} , by the shock radius, R_{cs} (Equation 16).

In the case of fully ionized, nonradiative (adiabatic) shocks, it may be justified to take $z_{\text{acc}} \sim R_{\text{cs}}$, in which case particle acceleration up to $E_{\text{max}} \gtrsim 10^{14} \text{ eV}$ would appear possible. However, in novae, z_{acc} may be limited to much smaller values by several effects (e.g., Metzger et al. 2016). For example, thermal cooling compresses the postshock region to a characteristic width $z_{\text{cool}} \sim v_{\text{cs}} t_{\text{cool}} \sim (t_{\text{cool}}/t_{\text{exp}}) R_{\text{cs}} \lesssim (10^{-2}$ – $10^{-1}) R_{\text{cs}}$ (Equation 17), and would imply $E_{\text{max}} \lesssim 10^4 \text{ GeV}$ if strong magnetic fields and particle acceleration are limited to this region (e.g., due to damping of cosmic-ray-driven instabilities by neutral-ion collisions; Metzger et al. 2016). Energy loss

Radiative cooling timescale (t_{cool}): timescale on which the shocked gas cools, $3kT_{\text{sh}}/(2n_{\text{sh}}\Lambda [T = T_{\text{sh}}])$

Maximum energy (E_{max}): the maximum energy to which relativistic particles are accelerated

Efficiency of particle acceleration (ϵ_{rel}):
the fraction of the shock kinetic energy that is transferred to accelerating particles to relativistic speeds

Supplemental Material >

due to pion creation may also limit the value of E_{max} , particularly in the calorimetric limit that $t_{\pi} \ll t_{\text{adv}}$ (see Section 4.3). Consistent with this, the gamma-ray spectra of some novae show evidence for high-energy cutoffs in their spectra above 10 GeV (**Supplemental Table 1**; Ahnen et al. 2015), which in hadronic models requires an intrinsic cutoff of the accelerated particle spectrum above $E_{\text{max}} \sim 10E_{\gamma} \approx 100$ GeV, which is much less than suggested by Equation 18 for $z_{\text{acc}} \sim R_{\text{cs}}$.

The nova ejecta can be sufficiently dense during the period of gamma-ray emission such that the timescale for relativistic ions or electrons to radiate their energy in \sim GeV gamma rays is short compared to the ejecta expansion time (e.g., Metzger et al. 2015). In other words, the shocks are radiative in terms of both their nonthermal and thermal particles. This calorimetric feature of nova shocks enables multiwavelength observations to probe the fraction of the shock's kinetic luminosity placed into relativistic particles, ϵ_{rel} , or the efficiency of particle acceleration. Because both forward and reverse shocks are radiative, the bolometric luminosity (L_{bol} ; dominated by the UVOIR bands around optical maximum) provides an upper limit on the total shock power (an upper limit because the remaining luminosity originates from the WD). Likewise, a fraction ϵ_{γ} of the relativistic particle energy is emitted as gamma rays in the LAT bandpass ($\epsilon_{\gamma} \approx 0.3$ – 0.4 in the hadronic scenario; Vurm & Metzger 2018). Thus, one can place a lower limit of $\epsilon_{\text{rel}} \gtrsim \epsilon_{\gamma}^{-1}(L_{\gamma}/L_{\text{bol}})$, given the distant-independent luminosity ratio $L_{\gamma}/L_{\text{bol}}$ (Metzger et al. 2015).

Figure 9c shows the time evolution of the ratio L_{γ}/L_{BB} for several gamma-ray novae, where L_{BB} is a best-fit blackbody luminosity that we take as an estimate of L_{bol} . The ratio varies from 3×10^{-4} in V339 Del to $\sim 10^{-2}$ in V1324 Sco, implying $\epsilon_{\text{rel}} \gtrsim 10^{-3}$ – 3×10^{-2} . In cases like V5856 Sgr and V906 Car, in which the optical–gamma-ray correlation demonstrates that a large fraction of the bolometric luminosity is powered by the shocks (**Figure 9**), the lower limit on ϵ_{rel} instead becomes an equality, enabling a measurement of $\epsilon_{\text{rel}} \approx (2$ – $5) \times 10^{-3}$ (Li et al. 2017, Aydi et al. 2020b).

If the gamma-ray emission from novae is generated via the hadronic channel (Section 3.1.2), then an inescapable prediction is a luminosity of ~ 0.1 – 10 GeV neutrinos from π^{\pm} decay that nearly mirrors the gamma-ray luminosity (Razzaque et al. 2010, Metzger et al. 2016). For an integration time of a few weeks (comparable with the duration of the observed gamma-ray emission), IceCube DeepCore (Abbasi et al. 2012) can reach a sensitivity to neutrinos of $\approx 10^{-5}$ – 10^{-4} GeV cm $^{-2}$ s $^{-1}$. Unfortunately, this is $\gtrsim 4$ orders of magnitude greater than the predicted neutrino fluxes, $\approx 10^{-8}$ GeV cm $^{-2}$ s $^{-1}$, of the present sample of gamma-ray novae (Alex Pizzuto, private communication), making a detection unlikely with current-generation experiments. If the gamma-ray spectra in some novae extend to much greater than TeV energies (Section 5.3), then the prospects of a neutrino detection would be improved.

4.3. Radiative Signatures of Radiative Shocks

When $v_{\text{f}} \gg v_{\text{s}}$, the reverse shock is more powerful than the forward shock, and its luminosity,

$$L_{\text{sh}} \approx \frac{9}{32} f_{\Omega} \dot{M}_{\text{f}} v_{\text{f}}^2 \approx 1.0 \times 10^{38} \text{ erg s}^{-1} \left(\frac{f_{\Omega}}{0.3} \right) \left(\frac{\dot{M}_{\text{f}}}{10^{-5} M_{\odot} \text{ week}^{-1}} \right) \left(\frac{v_{\text{f}}}{2,000 \text{ km s}^{-1}} \right)^2, \quad 19.$$

is comparable with the bolometric output of the nova near maximum. Because the shocks are radiative, L_{sh} must escape as radiation. However, the appearance of this radiation is strongly influenced by the medium ahead of the shocks.

The column density of gas separating the forward shock ($r = R_{\text{cs}}$) from an observer is

$$\Sigma = \int_{R_{\text{cs}}}^{\infty} \mu n_{\text{s}} dr \approx 6 \times 10^{23} \text{ cm}^{-2} \left(\frac{\xi}{3} \right)^{-2} \left(\frac{\dot{M}_{\text{s}}}{10^{-5} M_{\odot} \text{ week}^{-1}} \right) \left(\frac{v_{\text{s}}}{300 \text{ km s}^{-1}} \right)^{-2} \left(\frac{t_{\text{sh}}}{1 \text{ week}} \right)^{-1}. \quad 20.$$

The column density to the reverse shock will be even higher, $\Sigma \approx 10^{24}\text{--}10^{26} \text{ cm}^{-2}$, because of the additional contribution from the central swept-up shell.

Figure 7 shows the optical depth of the ejecta to absorption when $\Sigma \approx 10^{25} \text{ cm}^{-2}$, as expected roughly a week after the start of eruption for fiducial parameters. At radio frequencies, the optical depth due to free–free absorption is extremely high at these early times, but it will decrease rapidly as $\propto t^{-3}$ or faster as the shocks propagate outward, eventually enabling the escape of nonthermal synchrotron emission on timescales of months (Section 3.3). Likewise, the opacity in the far-UV and X-ray bands is enormous when $\Sigma \approx 10^{25} \text{ cm}^{-2}$, but it decreases as $\Sigma \propto t^{-1}$ or faster as the shocks propagate outward, eventually allowing the escape of X-rays (Section 3.2). In the first days to weeks of eruption, the substantial UV and X-ray luminosity of the shocks ($L_{\text{sh}} \approx 10^{37}\text{--}10^{38} \text{ erg s}^{-1}$ near peak; Equation 19) will be absorbed and reprocessed to wavelengths that can escape, contributing to the UVOIR (**Figure 9**). There is absorption in the OIR bands from a combination of free–free absorption and atomic absorption lines (Section 2.3.2), but it is significantly lower than the opacity at radio or X-ray wavelengths.

Shock emission combined with $\tau_{\text{abs}} < 1$ in both the optical and *Fermi* bands is the key to producing correlated optical and gamma-ray light curves in novae (**Figure 9**). The correlation implies that a significant fraction of the nova optical luminosity around maximum is powered by shocks (Metzger et al. 2014, Li et al. 2017, Aydi et al. 2020b), rather than by radiation from the nuclear-burning WD (Section 2.1.2). Some of the puzzling variability observed in nova optical light curves (Section 2.3.1; **Figure 3**) may be attributable to variability in the luminosity of the shocks (e.g., Equation 19), powered by fluctuations in the mass-loss rate and/or velocity of the fast or slow outflows.

A delay of days is observed between the start of the optical rise in novae and the onset of gamma-ray emission (**Figure 8**). This delay could be the timescale over which the fast wind is established, which then drives the shock interaction with the prior slow flow. Alternatively, shock interaction could start even earlier than gamma rays are observed, but at such large column densities ($\Sigma \gg 10^{26} \text{ cm}^{-2}$) that attenuation by photonuclear pair creation initially absorbs the gamma rays (Li et al. 2017, Franckowiak et al. 2018; see **Figure 7**). When the column densities are so high, the shocks are also mediated by radiation instead of collisionless plasma processes, leading to suppression of nonthermal particle acceleration (e.g., Blandford & Payne 1981) and, hence, gamma-ray emission.

By contrast, the faithful tracking of optical and gamma-ray light curves, with optical lagging behind the gamma rays by $\lesssim 8 \text{ h}$ in V906 Car (**Figure 9**), implies that particles are accelerated promptly, and conversion of cosmic-ray to gamma-ray energy is effectively instantaneous. The timescale required for the gamma-ray-emitting particles to be accelerated to high energies at the shock is at most the advection time through the cooling layer (see derivation of Equation 18), which is a small fraction of the expansion time. Likewise, once accelerated, the timescale for relativistic ions to interact with the dense gas behind the shocks and produce pions, and hence gamma rays, is extremely short ($\lesssim 1 \text{ day}$; Metzger et al. 2015, Martin et al. 2018, Vurm & Metzger 2018). A small lag in the arrival time of the optical radiation relative to the gamma rays, such as the $5.3 \pm 2.7\text{-h}$ lag observed in V906 Car (Aydi et al. 2020b), can arise owing to the finite diffusion time of optical photons through the ejecta, placing a constraint on the density profile ahead of the shocks.

4.4. Multidimensional Effects and Dust Formation

Although much insight can be gleaned from one-dimensional models, the structure of radiative internal shocks is intrinsically multidimensional. On the largest scales, the interaction between a fast outflow and an equatorially concentrated slower flow will not only generate a dual shock structure in the equatorial plane (**Figure 13**) but also may redirect the fast flow into a bipolar or hourglass structure, which is consistent with the large-scale morphology of nova ejecta inferred

from high-resolution imaging (**Figures 5 and 6**). This is qualitatively similar to the interacting winds scenario for the morphologies of planetary nebulae (Soker & Livio 1989).

On smaller scales, rapid cooling behind the shocks results in enormous increases in mean gas density by a factor up to $\mathcal{M}^2 \sim 10^4 (v/1,000 \text{ km s}^{-1})^2$, where $\mathcal{M} \equiv v/c_s \approx 100$ is the Mach number of the shock of velocity v . The compressed material collects into a thin shell with a thickness that is typically less than a few percent of its radius (e.g., Steinberg & Metzger 2018).

Radiative shocks are prone to various instabilities that lead to complex dynamical and geometrical structures and play an important role in shaping their emission relative to adiabatic shocks. These include thermal instabilities due to radiative cooling that can transform the temperature profile of the postshock cooling region from a smooth gradient into a turbulent multiphase plasma with distinct pockets of hot and cool gas (and little warm gas in between; e.g., Chevalier & Imamura 1982). The dense thin layer of gas behind dual radiative shocks is also subject to a non-linear thin-shell instability (Vishniac 1994). This occurs because lateral perturbations of the interaction front cause material to be diverted in such a way that the nominally planar shock front is given a corrugated shape (Strickland & Blondin 1995). **Supplemental Videos 1 and 2** show the evolution of the temperature and density structure behind dual radiative shocks created by the head-on collision of two flows, illustrating the effects of these instabilities.

The high compression ratios and clumpy structure of the shocked ejecta may produce conditions conducive to grain nucleation and dust formation (Derdzinski et al. 2017; Section 2.3.3). The cold central shell achieves densities $\gtrsim 10^{14} \text{ cm}^{-3}$, which can be maintained as the shell expands to large enough radii for the ambient temperature to drop sufficiently to enable dust nucleation. Derdzinski et al. (2017) find that dust grains readily grow to diameters $\sim 1 \mu\text{m}$ under these conditions and yield dust masses sufficient to explain the observational signatures of dust in novae. A prediction of this scenario is that dust extinction events will be preferentially observed for novae viewed within the binary plane; future imaging of nova dust emission with ALMA or IR interferometry can test this idea (Section 2.3.5).

Another consequence of instabilities at radiative shocks is a reduction of the thermal X-ray luminosity as compared to the expectation from a one-dimensional analysis that most of the shock power will be radiated by gas of temperature $\sim T_{\text{sh}}$ (Equation 10). First, the corrugated shape of the shock front imprinted by the thin-shell instability reduces T_{sh} from the case of a shock front that is strictly perpendicular to the upstream velocity. Second, the hot postshock gas can drive weak shocks into underpressurized cold regions in the thin shell, transferring energy from the hot to cool phase before it can be radiated and thus further reducing the effective temperature of the shock emission (Steinberg & Metzger 2018). Both effects may contribute to the unusually weak thermal X-ray emission seen from novae, even at times when luminous gamma-ray emission points to powerful shocks (Section 3.2; **Figure 11**).

Finally, the ion acceleration efficiencies $\epsilon_{\text{rel}} \sim 0.3\text{--}1\%$ inferred for nova shocks (Section 4.2; **Figure 9**) are notably lower than the $\sim 10\%$ efficiency typically inferred for cosmic-ray acceleration in SN remnants (Ellison et al. 2007). However, the nature of the upstream magnetic field, and the radiative nature of the shocks in novae, are different from the SN case. We expect that the magnetic field advected from the WD will be oriented perpendicular to the direction of the outflow velocity and thus parallel to the radially propagating shock front (similar to the Parker spiral magnetic field topology of the solar wind). Particle-in-cell plasma numerical simulations of nonrelativistic magnetized shocks indicate that ion acceleration is substantially suppressed for this upstream magnetic field geometry (Caprioli & Spitkovsky 2014a). Nevertheless, due to the corrugated shape of radiative shock fronts (**Supplemental Videos 1 and 2**), local patches of the shock front may satisfy the conditions for efficient ion acceleration and a shock-front-averaged efficiency of $\sim 1\%$ may be obtained (Steinberg & Metzger 2018).

5. BRINGING IT TOGETHER: IMPLICATIONS AND OPEN QUESTIONS

In this final section, we discuss the implications of nova mass loss and shocks for the long-term evolution of WD binaries (Section 5.1.1) and more energetic astrophysical transients (Sections 5.1.2 and 5.2). We close by reviewing the most important open questions in nova studies, as well as promising avenues for addressing them (Section 5.3).

5.1. Long-Term Evolution of White Dwarf Binaries

Novae induce mass loss (and, potentially, angular momentum loss) from accreting WD binaries. This has implications for how these binaries evolve and the end product of the WD itself.

5.1.1. Implications for cataclysmic variable populations. In standard models (e.g., Rappaport et al. 1983), mass transfer in CVs with orbital periods $P \approx 3\text{--}12$ h is driven by the loss of angular momentum due to braking by the secondary's magnetized wind. At shorter orbital periods ($P \lesssim 2$ h), the secondary no longer maintains a strong or ordered magnetic field, and angular momentum loss is dominated by gravitational wave radiation. As the period of a CV shrinks over time, the mean mass-transfer rate generally decreases [which is consistent with observations (Patterson 1984, Pala et al. 2017); although see the sidebar titled Nova “Feedback” on the Companion Star and Mass Transfer], and thus the interval between nova eruptions should increase. By contrast, if the donor is a subgiant or giant, mass transfer is instead governed by the companion's nuclear evolution, driven by the expansion of the secondary's envelope or its wind.

Does a nova eruption drive the WD and its companion closer together or further apart? If mass is lost from the binary (in the form of nova ejecta) but angular momentum is conserved, the binary separation will expand (Frank et al. 1992); this may lead to a decline in accretion rate and is the origin of the hypothesis that CVs hibernate after a nova eruption (Prialnik & Shara 1986, Shara et al. 1986, Kovetz et al. 1988, Hillman et al. 2020). However, if angular momentum is removed from the binary (i.e., through common-envelope-like interaction with the ejecta; Section 2.2.3), the binary separation may decrease (Livio et al. 1991). In addition to frictional angular momentum

NOVA “FEEDBACK” ON THE COMPANION STAR AND MASS TRANSFER

Although the standard picture of CV evolution predicts that the mass-transfer rate \dot{M} should mainly depend on the binary orbital period, CVs actually show a broad range of accretion rates at a given period, spanning more than two orders of magnitude (Patterson 1984, 2011; Warner 1987). Novae may be the cause of this diversity, because, in addition to producing discrete changes in orbital separation (Section 5.1.1), the nova eruption ablates and irradiates the companion star (Kovetz et al. 1988, Figueira et al. 2018). Irradiation of the secondary may lead to expansion of its atmosphere and a temporarily enhanced \dot{M} for decades to millennia after the nova, until the WD has cooled down (e.g., Mróz et al. 2016, Hillman et al. 2020).

If the impact of the nova on the secondary star is sufficiently strong, a positive feedback process can be established, in which the enhanced \dot{M} results in more frequent novae and even greater heating of the companion (Knigge et al. 2000). This effect would be stronger for smaller binary separations and may explain the recurrent nova T Pyx, which has a very short orbital period, $P = 1.83$ h, and a quiescent bolometric luminosity an order of magnitude greater than expected from CV theory (Patterson et al. 2017, Godon et al. 2018). Such positive feedback from nova cycles could lead to faster erosion of the companion star and the disappearance of short-period binaries from the CV population (Patterson et al. 2017).

loss, other processes like the companion star’s magnetic field braking on the nova ejecta (Martin et al. 2011), asymmetric expulsion of nova ejecta (Nelemans et al. 2016, Schaefer et al. 2019), or angular momentum extracted by a circumbinary disk (Taam & Spruit 2001, Liu & Li 2016) may decrease the orbital separation during or following a nova eruption.

Although the nova’s impact on the orbital separation should be measurable by comparing the orbital period before and after nova eruption, such measurements have proved challenging, usually because we lack excellent time-series photometry for the prenova binary. The literature now includes ~ 10 measurements of period changes across nova eruptions (**Supplemental Table 2**). They yield a mix of positive and negative values and show no clear correlation with nova or binary properties (Schaefer 2020). As a sample, however, they imply that—at least sometimes—angular momentum loss wins out over mass loss, and nova eruptions can drive their binaries to smaller separations. In the future, deep all-sky high-cadence surveys (e.g., the Legacy Survey of Space and Time with the *Vera C. Rubin Observatory*) will improve prospects for measuring changes in orbital period across nova eruptions.

Understanding how much angular momentum is lost in nova eruptions—and how it scales with system parameters like WD mass—is a critical ingredient in modeling CV populations. Although a wide range of observational evidence supports the standard model of CV evolution (e.g., Townsley & Bildsten 2003, Knigge 2006, Schreiber et al. 2010), several important discrepancies remain between observations and models. The CV space density is observed to be ~ 10 – 100 times lower than theoretically predicted (e.g., Patterson 1998, Pala et al. 2020), the minimum CV orbital period is observed to be longer than theoretically predicted (e.g., Gänsicke et al. 2009, Knigge et al. 2011), and the masses of WDs in CVs are observed to be larger than those in their progenitor population (Zorotovic et al. 2011). Schreiber et al. (2016) argue that all these discrepant properties can be understood if CV binaries experience additional angular momentum loss beyond magnetic braking and gravitational waves during their evolution. As discussed above, nova eruptions provide several potential sinks of angular momentum. These are starting to be included in models of CV populations, but they must act preferentially on systems with low-mass WDs to explain the discrepancies between CV observations and theory (Nelemans et al. 2016, Liu & Li 2016). Conveniently, nova angular momentum losses may be more severe for binaries with lower-mass WDs, because lower-mass WDs typically generate more slowly expanding, higher-mass ejecta (Section 2.1.1) and may, therefore, have stronger common-envelope interactions (Section 2.2.3). One speculative—and intriguing—mechanism for explaining the dearth of low-mass WDs in the CV population is merger of the WD with its companion brought on by slow nova eruptions.

5.1.2. Progenitors of thermonuclear supernovae. A Type Ia SN results from the thermonuclear explosion of a CO WD. One way to ignite runaway carbon burning in the center of a WD is by gradually increasing its mass up to a critical value near the Chandrasekhar limit through accretion from a nondegenerate binary companion (Whelan & Iben 1973, Nomoto 1982, Nomoto et al. 1984). Because the maximum initial mass of a CO WD is $\sim 1.1 M_{\odot}$ (Doherty et al. 2015), the viability of this single degenerate channel depends on whether the accreting WD can accumulate and retain $\gtrsim 0.3 M_{\odot}$ of material. Even if a CO WD can grow to $\sim M_{\text{Ch}}$, it may not produce a detonation, but instead lead to a pure deflagration event (recently hypothesized to explain Type Iax SNe, a subtype of thermonuclear explosions that are dimmer and less energetic than ordinary Type Ia SNe; Foley et al. 2013).

The overabundance of C, O, and Ne observed in the ejecta of many novae is strong evidence for mixing between the underlying WD and the accreted envelope (Section 2.1.1), so that the net effect of a nova is almost certainly to decrease rather than increase the WD mass due to the excavation and ejection of WD material. Only the most massive WDs and those with

M31N 2008-12a: TYPE Ia SUPERNOVA PROGENITOR?

A recurrent nova in the Andromeda Galaxy, M31N 2008-12a, exhibits a recurrence period of ~ 1 year, the shortest known to date (Darnley et al. 2015, Henze et al. 2015). This is one of the best known candidates for a single-degenerate Type Ia SN progenitor, as the WD's very high effective temperature during the supersoft X-ray phase implies a mass close to the Chandrasekhar limit ($\sim 1.5 \times 10^6$ K; Darnley et al. 2016), and the very fast recurrence time is only achieved by nova models for the most massive WDs accreting at very high rates (**Figure 2**). M31N 2008-12a is also surrounded by a nova superremnant that is 130 pc in diameter and is probably inflated by $\sim 10^5$ recurrent nova eruptions (Darnley et al. 2019). Estimates suggest that the M31N 2008-12a WD could reach the Chandrasekhar mass and thus undergo core burning in less than 20,000 years (Darnley et al. 2017). However, whether the outcome will be a Type Ia SN, as opposed to a pure deflagration (e.g., Type Iax SN) or accretion-induced collapse event, depends on the core composition of the WD and the outcome of core ignition.

accretion rates just below \dot{M}_{stable} (**Figure 2**) should host gentle novae that do not lead to net mass loss from the WD (e.g., Yaron et al. 2005). These will also be the novae with the shortest recurrence times, and indeed, several recurrent nova systems are observed to host massive WDs and may have $M_{\text{acc}} > M_{\text{ej}}$ (Thoroughgood et al. 2001, Osborne et al. 2011, Page et al. 2015; see also the sidebar titled M31N 2008-12a: Type Ia Supernova Progenitor?). However, it is unclear if the compositions of the WDs in recurrent nova systems are primarily ONe or CO. An ONe WD cannot be a Type Ia SN progenitor, because the end result of its growth toward the Chandrasekhar limit is generally thought to be an accretion-induced collapse to a neutron star instead of a thermonuclear explosion (Nomoto & Kondo 1991, Schwab et al. 2015; however, see Jones et al. 2019). It also appears that there are not enough recurrent novae (or steady supersoft sources accreting at $\gtrsim \dot{M}_{\text{stable}}$) in nearby galaxies to explain the Type Ia SN rate (e.g., Gilfanov & Bogdán 2010). Furthermore, even if WDs manage to retain a significant fraction of the accreted mass, the accreted mass is typically too small to achieve M_{Ch} , especially in old stellar populations (e.g., Ruiter et al. 2009, Maoz & Mannucci 2012).

Even if accreted material can be retained on the WD through nova eruptions (e.g., Starrfield et al. 2020), the sustained burning phase following a nova eruption necessarily leaves a layer of helium ash that accumulates over time (e.g., Denissenkov et al. 2013). As in the hydrogen case, once this helium shell achieves a critical mass, it can undergo runaway helium burning, causing drastic expansion of the WD envelope and mass loss (Cassisi et al. 1998). Such helium novae can also occur in mass-transfer systems involving He-rich donors like V445 Puppis (see the sidebar titled Helium Novae).

5.2. Implications for Extragalactic Shock-Powered Transients

As time-domain surveys uncover new classes of astrophysical transients, it has become clear that the classical mechanisms of nuclear fusion and radioactive decay are insufficient to explain many of their luminosities and timescales. As an alternative, it has been proposed that many transients' light curves are instead powered by shock interaction for classes as diverse as Type IIIn SNe, Type Ia-CSM SNe, superluminous SNe, pulsational-pair instability SNe, stellar mergers, fast blue optical transients, and tidal disruption events (Woosley et al. 2007, Smith et al. 2008, Silverman et al. 2013, Metzger & Pejcha 2017, Gal-Yam 2019, Margutti et al. 2019).

For example, interaction-powered SNe (e.g., Types IIn and Ia-CSM) reach peak luminosities $\sim 5\text{--}100$ times brighter than SNe primarily powered by radioactive decay (i.e., Types Ia and Ib/c), and their optical light curves evolve more slowly (Kiewe et al. 2012, Silverman et al. 2013). The leading model is that these SNe are surrounded by large masses of circumstellar material, and the resulting shocks are deeply embedded and heavily absorbed such that the shock emission is reprocessed to optical wavelengths before emerging (e.g., Smith et al. 2008). However, it had not been clearly demonstrated that this mechanism can indeed produce the bulk of the optical luminosity—until the recent observations of correlated optical and gamma-ray emission in much closer classical novae (Figure 9).

With novae now demonstrating that reprocessed shock emission can dominate the luminosity of astrophysical transients, they can be used to study the physics of this reprocessing in detail. Novae demonstrate that optical and gamma-ray wavelengths provide the most direct glimpses of the shock (Figure 7). They reveal how shock power is converted to gamma-ray luminosity and by which mechanisms (Figure 10). Novae show how other wavelength regimes like X-ray and radio can provide useful insights into the shocks when studied with the understanding that most astrophysical transients are aspherical (e.g., Figure 6), that the emission from radiative shocks may not conform to expectations based on adiabatic shocks, and that optical depth effects play a key role in what we observe. Novae also provide potential real-time probes of dust formation in shock-compressed media (Sections 2.3.3 and 4.4), which is a phenomenon shared with other astrophysical systems such as colliding wind binaries (e.g., Tuthill et al. 1999) and stellar mergers (Munari et al. 2002, Metzger & Pejcha 2017).

The ejecta morphology proposed throughout this review—a slow, denser equatorial component and a faster, isotropic flow (which becomes funneled out the poles)—is common among astrophysical transients, especially those with binary progenitors. Similar ejecta morphologies occur in stellar mergers, SN 1987A (and its circumstellar environment), and kilonovae (Blondin & Lundqvist 1993, Pejcha et al. 2016, Metzger 2019). Novae allow us to observe the consequences of the two flows interacting, grounding predictions for these other more distant and exotic explosions.

Shocks in novae tend to possess a combination of higher densities and lower velocities than many other astrophysical shocks (e.g., SNe and SN remnants, gamma-ray bursts, etc.), but nevertheless, they sample a broad range of physical conditions. As such, nova shocks provide an opportunity to study how characteristics of the particle acceleration process (ϵ_{rel} , q , E_{max} ; Section 3.1.2) vary with shock properties in the radiative regime. In other shock environments, the cosmic-ray acceleration efficiency ϵ_{rel} and primary energy spectra are often difficult to measure, because of uncertainty in cosmic-ray escape fraction and propagation effects (Blasi 2013). However, this task is enormously simplified in novae because of their calorimetric nature (Section 4.2). As a recent example of lessons learned from novae and applied to other astrophysical transients, Fang et al. (2020) convert the optical luminosities of diverse transients into upper limits on their shock luminosities, energies in relativistic particles, and neutrino fluxes. Calibrating this conversion using particle acceleration properties measured from novae, Fang et al. (2020) found that known classes of shock-powered transients can only account for $<1\text{--}10\%$ of the high-energy neutrino background measured by IceCube (IceCube Collab. 2013).

5.3. OPEN QUESTIONS, FUTURE WORK

■ **How well does WD TNR theory match observed nova properties?** The basic predictions of nova theory—ejecta masses and kinetic energies, and how these scale with WD mass and accretion rate—still lack precise tests over a broad parameter space. Observers should work to grow

and diversify the sample of novae within which these properties are all accurately measured. Many data are already in hand, thanks to sustained efforts by, e.g., the Swift Nova-CV collaboration (e.g., Page et al. 2020), the Stony Brook/SMARTS Atlas (Walter et al. 2012), and the eNova collaboration (e.g., Roy et al. 2012). Future progress can be made by synthesizing these large samples and multiwavelength data sets. A first step would be a meta-analysis of ejecta mass measurements, in which discrepant estimates with different techniques are homogenized and the volume filling factor is properly accounted for.

On a related point, though several teams have developed codes for modeling nova eruptions over the past couple decades, no systematic comparison has ever been performed to assess points of agreement or conflict. Detailed code comparisons are a critical next step for establishing the biggest challenges and opportunities of the future.

■ **How does mixing between the accreted envelope and the underlying WD vary across novae, and how does it depend on WD mass and accretion rate?** Studies of diffusion-induced mixing and multidimensional simulations of shear-mixing driven by convection during the TNR have advanced our understanding of the mixing processes that are key for determining properties of the nova eruption. Future theoretical work should continue to examine these mechanisms throughout the entire parameter space of M_{WD} and \dot{M} , as well as investigate whether alternative processes, such as heating by convectively driven internal gravity waves, also contribute to mixing. Observers can probe this issue by systematically obtaining abundance measurements for large samples of novae using X-ray–UVOIR spectra and test theoretical predictions for how mixing scales with nova properties by comparing the measured level of WD pollution with properties of the outburst and of the binary system (e.g., M_{WD} and \dot{M}).

■ **How is mass ejected in nova eruptions: impulsive ejection, radiation-driven wind, common-envelope interaction, or a combination thereof?** Basic features in nova optical light curves like plateaus, jitters, and oscillations highlight that mass loss can persist over months and fluctuate in rate and velocity. Although the cadence and quality of observations are shedding new light on this complexity, theoretical work lags in comparison. Current models are largely limited to one dimension (necessarily neglecting the effects of the binary) and utilize simple parameterizations of mass loss. A new generation of three-dimensional simulations, that holistically model the physics of the TNR, convection, mass loss, and interaction with the binary, is critically needed. Such simulations are necessary for answering this—and many of the other open questions listed here—but will be computationally very challenging, as they will need to model enormous spatial and temporal dynamic ranges.

■ **What physical processes shape the 3D morphology of nova ejecta?** Although it is well established that nova ejecta are often aspherical, work is needed to connect this complex morphology with the physics of mass loss in nova eruptions. As theorists pursue multidimensional simulations, observers should pursue high-resolution imaging at a range of wavelengths, using the VLA (Very Large Array), ALMA (Atacama Large Millimeter/Submillimeter Array), IR adaptive optics/interferometry, and HST. Images should be obtained early and often during the eruption, to assess how the morphology of the ejecta changes as mass ejection persists and as the ejecta diffuse. Morphokinematic modeling software like SHAPE will remain critical to interpreting these observations. In the future, the Next Generation VLA will provide imaging of thermal emission at milliarcsecond scales (Murphy et al. 2018), enabling a large and exquisite survey of time-dependent nova morphologies.

■ **Where do shocks occur in the nova ejecta?** Throughout this review, we have suggested a picture of an equatorially focused slow flow impacted by a more isotropic faster flow. Although there is observational support for this picture, it is far from clear that the scenario is universal. Growing the sample of novae with high-resolution imaging will be helpful for testing the

universality of this scenario—especially if combined with imaging of synchrotron emission. In addition, studies of how nova properties depend on orbital inclination will be illuminating. For example, if shocks primarily occur in the orbital plane, we would expect radio, optical, and X-ray signatures of shocked gas to vary with orbital inclination.

A mystery persists as to why the X-rays observed concurrently with the GeV gamma rays are so underluminous compared to naive expectations. It is unclear if the corrugated and thermally unstable nature of radiative shocks can suppress the X-rays to the necessary degree or if different shocks dominate the observed X-ray and gamma-ray emission. Multidimensional simulations of shock interaction in novae, which capture the interaction of an isotropic fast flow with an equatorially concentrated slow flow at sufficient resolution to model relevant small-scale instabilities, would help address whether the full range of observational signatures is consistent with a single dominant shock structure.

■ **Where is dust produced in nova ejecta?** Future observations should test if dust is produced in the dense cool shell generated by radiative shocks in novae. One scenario put forth in this review predicts that deep dust dips in optical light curves should be correlated with shock signatures and occur in edge-on binaries. Kinematics of the dust—measured from molecular line profiles—would test if the dust is associated with the slow flow. Effort should also be invested in imaging the sites of dust formation with ALMA or near-IR interferometers.

■ **To what maximum energy do nova shocks accelerate particles?** Additional observational constraints on E_{max} —and if it varies across novae—would provide a valuable test of diffusive shock acceleration theories. Although there is evidence for a steepening in the gamma-ray spectra of novae above ~ 10 – 100 GeV from a single epoch in a single event (e.g., Ahnen et al. 2015), additional observations of novae at \sim TeV gamma-ray energies [with VERITAS (Very Energetic Radiation Imaging Telescope Array System), MAGIC, HAWC (High Altitude Water Cherenkov Gamma-Ray Observatory), and eventually the CTA (Cherenkov Telescope Array; CTA Consort. et al. 2019)] are needed to probe the highest particle energies accelerated at the shocks. If novae produce a significant luminosity of \sim TeV gamma rays (implying a high E_{max}), this would considerably improve their detection prospects as neutrino sources by IceCube and successor experiments.

■ **Why do the shock energies and gamma-ray luminosities span such a wide range?** Although the gamma-ray luminosities of novae span at least two orders of magnitude, we lack a clear explanation for this diversity. Furthermore, no clear correlations have been identified between the gamma-ray emission and other nova properties. Continued survey observations of GeV gamma rays are necessary to grow the sample of gamma-ray-detected novae to establish the full diversity of their properties. The kinetic power of the shocks in novae should be determined by the masses and relative velocities of the slow and fast flows, and radio imaging has the potential to separately measure the properties of these distinct flows (Figure 6).

■ **What powers super-Eddington luminosities in novae?** The apparently super-Eddington luminosities of some novae remain almost as much of a mystery as when they were first discovered, and the role of shocks (if any) in this behavior requires further elucidation. Multiwavelength observations, from the IR to the X-ray band, should be combined to measure bolometric light curves for a substantial sample of novae. With the availability of *Gaia* parallaxes and three-dimensional Galactic dust maps, these will enable the construction of bolometric luminosity time evolution with unprecedented accuracy. These can then be compared with shock signatures and the luminosity of the supersoft component to test the degree to which shocks contribute to the bolometric luminosities of novae. Comparisons between the bolometric luminosities of classical novae and embedded novae should also be carried out, as a test of the energetic contribution of a common-envelope phase in nova eruptions.

■ **Do CVs ever host dense circumbinary material?** Although the shocks in classical novae can be explained as internal shocks, and most observations of CVs are consistent with very low density circumbinary material, there are arguments to the contrary to explain THEA lines or to remove binary angular momentum. Novae themselves can serve as light bulbs to illuminate material preexisting the eruption. Observations of the early UV–X-ray flash and a better understanding of THEA lines have the potential to strongly constrain the presence of circumbinary material.

■ **Why are WDs in CVs so massive?** Perhaps the most important mystery plaguing CV populations is why the WDs in CVs are substantially more massive than their progenitor population. Future observations should test if nova eruptions serve as a significant sink of angular momentum and cause CV binaries with low-mass WDs to merge, removing them from the population. Systematically comparing eruption properties of classical novae and embedded novae can shed light on how strongly the binary orbit couples to the nova ejecta.

Measurements of how binary orbital period changes across a nova eruption should be pursued at scale with time-domain facilities like the *Rubin Observatory* to quantify how much angular momentum is lost in nova eruptions. Finally, we should test if all slow novae still contain a binary after eruption or if some slow novae might result in the merger of the WD and its companion. Tests of this hypothesis will be supported by a deepening understanding of the observational signatures of stellar mergers (e.g., luminous red novae) and improved coverage of the time-domain sky.

■ **What can we learn from helium novae?** The basic physics of novae powered by hydrogen burning should apply to helium novae as well, whether they arise from direct helium accretion or from the accumulation of helium ash due to hydrogen burning. Although much rarer than classical novae, helium novae can expand our understanding of all novae by extending the parameter space to higher ignition masses and ejecta velocities. The theoretical modeling discussed above should also be applied to helium novae, along with consideration of truly explosive phenomena (deflagrations or detonations) due to the removal of the β -decay constraint limiting the energetics of hydrogen novae. Such modeling will help observational efforts to recognize the subset of helium novae among the much larger class of hydrogen novae and select them for dedicated follow-up.

SUMMARY POINTS

1. The theory of nova eruptions as thermonuclear runaways (Section 2.1.1) has successfully explained many observations (Section 2.3). However, there remain observations—as fundamental as optical light curves—that lack agreed-upon explanations. The challenge is to understand how mass loss proceeds in nova eruptions (Section 2.2), including three-dimensional effects such as the interaction of the ejecta with the binary.
2. The ejecta in novae are usually far from spherical. The general morphology may be a slower denser flow in the orbital plane and a faster flow that escapes in the polar directions, but it remains unclear if this can explain the morphology of all novae.
3. Recent decades have shown dramatic improvements in our understanding of sustained nuclear burning (e.g., supersoft emission) on the WD following the nova (Section 2.1.2). These observations solidified a belief in the 2000s that radiation from the hot WD determines the bolometric luminosity and evolution of novae.
4. Recent (post-2010) detections of GeV gamma rays, and renewed attention to nova shocks, have shown that this paradigm is not complete. Models of nova shocks predicted—and observations have largely confirmed—that the internal shocks are

radiative, with the kinetic luminosity promptly transformed to radiative luminosity. Observations of correlated optical and gamma-ray light curves imply that shocks can significantly contribute to the bolometric luminosity and eruption evolution.

5. Novae are therefore the nearest and most common interaction-powered astrophysical transients, and the first to be detected and observed in detail from radio to gamma-ray wavelengths. They are valuable laboratories for understanding the physics—and observable signatures—of radiative shocks and relativistic particle acceleration.
6. To continue to make progress on nova mass loss and shocks, three-dimensional simulations that self-consistently include the full range of physics are urgently needed. On the observational side, high-resolution imaging will be critical for testing these simulations and further establishing the origin of nova shocks.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are deeply thankful to Elias Aydi for his work on figures and helpful comments. We offer special thanks to Alexa Gordon, Kwan-Lok Li, Elad Steinberg, and Indrek Vurm for supplying data to generate some of the figures, along with insights. We thank Pavel Denissenkov, Jay Gallagher, Boris Gänsicke, Yael Hillman, Justin Linford, Koji Mukai, Kim Page, Alex Pizzuto, Josiah Schwab, Jenő Sokolowski, Kirill Sokolovsky, Elad Steinberg, Dean Townsley, and Bill Wolf for useful information and helpful feedback. Special thanks to Chris Kochanek for thorough comments on an earlier draft of the text. We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research. **Figure 5** is based on observations made with the NASA/ESA *Hubble Space Telescope*, and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESA) and the Canadian Astronomy Data Centre (CADC/NRC/CSA).

L.C. acknowledges support from NSF awards AST-1751874 and AST-1907790, NASA awards Fermi-80NSSC18K1746 and NuSTAR-80NSSC19K0522, and a Cottrell fellowship of the Research Corporation. B.D.M. acknowledges support from NSF award AST-1615084 and the Simons Foundation through the Simons Fellowship (grant number 606260). K.J.S. acknowledges support from the NASA Astrophysics Theory Program (NNX17AG28G and 80NSSC20K0544).

LITERATURE CITED

- Abbasi R, Abdou Y, Abu-Zayyad T, et al. 2012. *Astropart. Phys.* 35:615–24
Abdo AA, Ackermann M, Ajello M, et al. 2010. *Science* 329:817–21
Ackermann M, Ajello M, Albert A, et al. 2014. *Science* 345:554–58
Ahnen ML, Ansoldi S, Antonelli LA, et al. 2015. *Astron. Astrophys.* 582:A67
Ashok NM, Banerjee DPK. 2003. *Astron. Astrophys.* 409:1007–15

- Atwood WB, Abdo AA, Ackermann M, et al. 2009. *Ap. J.* 697:1071–102
- Aydi E, Chomiuk L, Izzo L, et al. 2020a. *Ap. J.* 905:62
- Aydi E, Chomiuk L, Strader J, et al. 2019. arXiv:1903.09232
- Aydi E, Page KL, Kuin NPM, et al. 2018. *MNRAS* 474:2679–705
- Aydi E, Sokolovsky KV, Chomiuk L, et al. 2020b. *Nat. Astron.* 4:776–80
- Banerjee DPK, Ashok NM. 2012. *Bull. Astron. Soc. India* 40:243–65
- Bath GT, Harkness RP. 1989. In *Classical Novae*, ed. MF Bode, A Evans, pp. 61–72. Chichester, UK: Wiley
- Bath GT, Shaviv G. 1976. *MNRAS* 175:305–22
- Beck H, Gail HP, Gass H, Sedlmayr E. 1990. *Astron. Astrophys.* 238:283–95
- Bell AR. 2004. *MNRAS* 353:550–58
- Bianchini A, Friedjung M, Brinkmann W. 1992. *Astron. Astrophys.* 257:599–602
- Blandford RD, Ostriker JP. 1978. *Ap. J. Lett.* 221:L29–32
- Blandford RD, Payne DG. 1981. *MNRAS* 194:1041
- Blasi P. 2013. *Astron. Astrophys. Rev.* 21:70
- Blasi P. 2019. *Riv. Nuovo Cim.* 42:549–600
- Blondin JM, Lundqvist P. 1993. *Ap. J.* 405:337–52
- Bode MF, Evans A, eds. 2008. *Classical Novae*. Cambridge, UK: Cambridge Univ. Press. 2nd ed.
- Buson S, Jean P, Cheung CC. 2019. *Astron. Teleg.* 13114:1
- Cao Y, Kasliwal MM, Neill JD, et al. 2012. *Ap. J.* 752:133
- Capaccioli M, Della Valle M, D’Onofrio M, Rosino L. 1989. *Astron. J.* 97:1622–33
- Caprioli D, Spitkovsky A. 2014a. *Ap. J.* 783:91
- Caprioli D, Spitkovsky A. 2014b. *Ap. J.* 794:46
- Casanova J, José J, García-Berro E, Calder A, Shore SN. 2010. *Astron. Astrophys.* 513:L5
- Casanova J, José J, García-Berro E, Shore SN. 2016. *Astron. Astrophys.* 595:A28
- Cassisi S, Iben I Jr., Tornambè A. 1998. *Ap. J.* 496:376–85
- Castor JI, Abbott DC, Klein RI. 1975. *Ap. J.* 195:157–74
- Chen HL, Woods TE, Yungelson LR, et al. 2019. *MNRAS* 490:1678–92
- Chesneau O, Lagadec E, Otulakowska-Hypka M, et al. 2012. *Astron. Astrophys.* 545:A63
- Cheung CC, Jean P, Shore SN. 2014. *Astron. Teleg.* 5879:1
- Cheung CC, Jean P, Shore SN, et al. 2016. *Ap. J.* 826:142
- Chevalier RA, Imamura JN. 1982. *Ap. J.* 261:543–49
- Chochol D, Grygar J, Pribulla T, et al. 1997. *Astron. Astrophys.* 318:908–24
- Chomiuk L, Linford JD, Yang J, et al. 2014a. *Nature* 514:339–42
- Chomiuk L, Nelson T, Mukai K, et al. 2014b. *Ap. J.* 788:130
- Clayton DD, Hoyle F. 1974. *Ap. J. Lett.* 187:L101–3
- Clayton M, Podsiadlowski P, Ivanova N, Justham S. 2017. *MNRAS* 470:1788–808
- Cohen JG. 1985. *Ap. J.* 292:90–103
- Csák B, Kiss LL, Retter A, Jacob A, Kaspi S. 2005. *Astron. Astrophys.* 429:599–605
- CTA Consort., Acharya BS, Agudo I, et al. 2019. *Science with the Cherenkov Telescope Array*. Heidelberg, Ger.: CTA Consort.
- Cunningham T, Wolf WM, Bildsten L. 2015. *Ap. J.* 803:76
- Darnley MJ, Bode MF, Kerins E, et al. 2006. *MNRAS* 369:257–71
- Darnley MJ, Henze M, Bode MF, et al. 2016. *Ap. J.* 833:149
- Darnley MJ, Henze M, Steele IA, et al. 2015. *Astron. Astrophys.* 580:A45
- Darnley MJ, Hounsell R, Godon P, et al. 2017. *Ap. J.* 849:96
- Darnley MJ, Hounsell R, O’Brien TJ, et al. 2019. *Nature* 565:460–63
- Darnley MJ, Ribeiro VARM, Bode MF, Hounsell RA, Williams RP. 2012. *Ap. J.* 746:61
- De Marco O. 2009. *Publ. Astron. Soc. Pac.* 121:316–42
- de Vaucouleurs G. 1978. *Ap. J.* 223:351–63
- Delgado L, Hernanz M. 2019. *MNRAS* 490:3691–704
- Della Valle M, Izzo L. 2020. *Astron. Astrophys. Rev.* 28:3
- Della Valle M, Livio M. 1995. *Ap. J.* 452:704–9

- Denissenkov PA, Herwig F, Bildsten L, Paxton B. 2013. *Ap. J.* 762:8
- Derdzinski AM, Metzger BD, Lazzati D. 2017. *MNRAS* 469:1314–29
- Diaz MP, Bruch A. 1997. *Astron. Astrophys.* 322:807–16
- Diaz MP, Williams RE, Luna GJ, Moraes M, Takeda L. 2010. *Astron. J.* 140:1860–67
- Doherty CL, Gil-Pons P, Siess L, Lattanzio JC, Lau HHB. 2015. *MNRAS* 446:2599–612
- Downen LN, Iliadis C, José J, Starrfield S. 2013. *Ap. J.* 762:105
- Downes RA, Duerbeck HW. 2000. *Astron. J.* 120:2007–37
- Duerbeck HW. 1981. *Publ. Astron. Soc. Pac.* 93:165–75
- Duerbeck HW. 2008. See Bode & Evans 2008, pp. 1–15
- Ederoclite A, Mason E, Della Valle M, et al. 2006. *Astron. Astrophys.* 459:875–83
- Ellison DC, Patnaude DJ, Slane P, Blasi P, Gabici S. 2007. *Ap. J.* 661:879–91
- Evans A, Banerjee DPK, Gehrz RD, et al. 2017. *MNRAS* 466:4221–38
- Evans A, Rawlings JMC. 2008. See Bode & Evans 2008, pp. 308–34
- Eyres SPS, Bewsher D, Hillman Y, et al. 2017. *MNRAS* 467:2684–89
- Fang K, Metzger BD, Vurm I, Aydi E, Chomiuk L. 2020. *Ap. J.* 904:4
- Ferland GJ, Porter RL, van Hoof PAM, et al. 2013. *Rev. Mex. Astron. Astrophys.* 49:137–63
- Ferland GJ, Shields GA. 1978. *Ap. J.* 226:172–85
- Figueira J, José J, García-Berro E, et al. 2018. *Astron. Astrophys.* 613:A8
- Finzell T, Chomiuk L, Metzger BD, et al. 2018. *Ap. J.* 852:108
- Foley RJ, Challis PJ, Chornock R, et al. 2013. *Ap. J.* 767:57
- Franckowiak A, Jean P, Wood M, Cheung CC, Buson S. 2018. *Astron. Astrophys.* 609:A120
- Frank J, King A, Raine D. 1992. *Accretion Power in Astrophysics*. Cambridge, UK: Cambridge Univ. Press
- Friedjung M. 1966. *MNRAS* 132:317–36
- Friedjung M. 1987. *Astron. Astrophys.* 180:155–59
- Friedjung M. 2011. *Astron. Astrophys.* 536:A97
- Friedjung M, Duerbeck HW. 1993. In *Cataclysmic Variables and Related Objects*, NASA, Washington, pp. 371–412. SEE N95-27068 09-89
- Froning CS. 2005. In *The Astrophysics of Cataclysmic Variables and Related Objects*, ed. J-M Hameury, J-P Lasota, *ASP Conf. Ser.* 330:81–90. San Francisco: ASP
- Fujimoto MY. 1982a. *Ap. J.* 257:752–66
- Fujimoto MY. 1982b. *Ap. J.* 257:767–79
- Fujimoto MY. 1993. *Ap. J.* 419:768–75
- Gal-Yam A. 2019. *Annu. Rev. Astron. Astrophys.* 57:305–33
- Gallagher JS, Code AD. 1974. *Ap. J.* 189:303–14
- Gallagher JS, Starrfield S. 1976. *MNRAS* 176:53–61
- Gallagher JS, Starrfield S. 1978. *Annu. Rev. Astron. Astrophys.* 16:171–214
- Galloway DK, Keck L. 2021. In *Timing Neutron Stars: Pulsations, Oscillations and Explosions*, ed. TM Belloni, M Méndez, C Zhang. *Ap. Space Sci. Libr.* 461:209–62
- Gänsicke BT, Dillon M, Southworth J, et al. 2009. *MNRAS* 397:2170–88
- Gehrels N, Chincarini G, Giommi P, et al. 2004. *Ap. J.* 611:1005–20
- Gehrz RD. 1988. *Annu. Rev. Astron. Astrophys.* 26:377–412
- Gehrz RD, Evans A, Woodward CE, et al. 2018. *Ap. J.* 858:78
- Gehrz RD, Grasdalen GL, Hackwell JA, Ney EP. 1980. *Ap. J.* 237:855–65
- Gehrz RD, Ney EP. 1987. *PNAS* 84:6961–64
- Gehrz RD, Truran JW, Williams RE, Starrfield S. 1998. *Publ. Astron. Soc. Pac.* 110:3–26
- Gilfanov M, Bogdán Á. 2010. *Nature* 463:924–25
- Gill CD, O’Brien TJ. 1998. *MNRAS* 300:221–32
- Glasner SA, Livne E, Truran JW. 1997. *Ap. J.* 475:754–62
- Godon P, Sion EM, Williams RE, Starrfield S. 2018. *Ap. J.* 862:89
- Gomez-Gomar J, Hernanz M, Jose J, Isern J. 1998. *MNRAS* 296:913–20
- Goranskij V, Shugarov S, Zharova A, Kroll P, Barsukova EA. 2010. *Perem. Zvezdy* 30:4
- Gordon AC, Aydi E, Page KL, et al. 2021. *Ap. J.* 910:134

- Green DA. 2019. *J. Astrophys. Astron.* 40:36
- Hachisu I, Kato M, Nomoto K. 1996. *Ap. J. Lett.* 470:L97–100
- Harman DJ, O'Brien TJ. 2003. *MNRAS* 344:1219–26
- Harris MJ, Leising MD, Share GH. 1991. *Ap. J.* 375:216–20
- Hauschildt PH, Starrfield S, Austin S, et al. 1994. *Ap. J.* 422:831–44
- Hauschildt PH, Starrfield S, Shore SN, Allard F, Baron E. 1995. *Ap. J.* 447:829–47
- Hauschildt PH, Wehrse R, Starrfield S, Shaviv G. 1992. *Ap. J.* 393:307–28
- Helton LA, Woodward CE, Walter FM, et al. 2010. *Astron. J.* 140:1347–69
- Henze M, Darnley MJ, Kabashima F, et al. 2015. *Astron. Astrophys.* 582:L8
- Henze M, Pietsch W, Haberl F, et al. 2011. *Astron. Astrophys.* 533:A52
- Henze M, Pietsch W, Haberl F, et al. 2014. *Astron. Astrophys.* 563:A2
- Hernanz M. 2012. *Bull. Astron. Soc. India* 40:377–91
- Hernanz M, Gómez-Gomar J, José J. 2002. *Acta Astronaut.* 46:559–63
- Hernanz M, Smith DM, Fishman J, et al. 2000. In *The Fifth Compton Symposium*, ed. ML McConnell, JM Ryan. *AIP Conf. Ser.* 510:82–86. Melville, NY: AIP
- Hillman Y, Prialnik D, Kovetz A, Shara MM, Neill JD. 2014. *MNRAS* 437:1962–75
- Hillman Y, Shara MM, Prialnik D, Kovetz A. 2020. *Nat. Astron.* 4:886–92
- Hjellming RM. 1996. In *Radio Emission from the Stars and the Sun*, ed. AR Taylor, JM Paredes. *ASP Conf. Ser.* 93:174–81. San Francisco: ASP
- Hjellming RM, van Gorkom JH, Taylor AR, et al. 1986. *Ap. J. Lett.* 305:L71–75
- Hjellming RM, Wade CM, Vandenberg NR, Newell RT. 1979. *Astron. J.* 84:1619–31
- Hoard DW, Long KS, Howell SB, et al. 2014. *Ap. J.* 786:68
- Hoffmann SM, Vogt N, Protte P. 2020. *Astron. Nachr.* 341:79–98
- Hounsell R, Bode MF, Hick PP, et al. 2010. *Ap. J.* 724:480–86
- Hounsell R, Darnley MJ, Bode MF, et al. 2016. *Ap. J.* 820:104
- Hutchings JB. 1970a. *Publ. Astron. Soc. Pac.* 82:603
- Hutchings JB. 1970b. *Publ. Dom. Astrophys. Obs. Vic.* 13:397
- Hutchings JB. 1972. *MNRAS* 158:177–98
- Hyland AR, Neugebauer G. 1970. *Ap. J. Lett.* 160:L177–180
- Iben I Jr., Fujimoto MY, MacDonald J. 1992a. *Ap. J.* 384:580–86
- Iben I Jr., Fujimoto MY, MacDonald J. 1992b. *Ap. J.* 388:521–40
- IceCube Collab. 2013. *Science* 342:1242856
- Iglesias CA, Rogers FJ. 1996. *Ap. J.* 464:943–53
- Iijima T, Nakanishi H. 2008. *Astron. Astrophys.* 482:865–77
- Ivanova N. 2018. *Ap. J. Lett.* 858:L24
- Ivanova N, Justham S, Chen X, et al. 2013. *Astron. Astrophys. Rev.* 21:59
- Iverson RJ, Hughes DH, Lloyd HM, Bang MK, Bode MF. 1993. *MNRAS* 263:L43–46
- Izzo L, Della Valle M, Mason E, et al. 2015. *Ap. J. Lett.* 808:L14
- Jiang YF, Cantiello M, Bildsten L, Quataert E, Blaes O. 2017. *Ap. J.* 843:68
- Jones S, Röpké FK, Fryer C, et al. 2019. *Astron. Astrophys.* 622:A74
- José J, Hernanz M. 1998. *Ap. J.* 494:680–90
- José J, Hernanz M. 2007. *J. Phys. G Nucl. Phys.* 34:R431–58
- José J, Hernanz M, Coc A. 1997. *Ap. J. Lett.* 479:L55–58
- José J, Shore SN, Casanova J. 2020. *Astron. Astrophys.* 634:A5
- Kafka S. 2020. *Observations from the AAVSO International Database*, updated June 2020, retrieved June 5, 2020. <https://www.aavso.org>
- Kahabka P, van den Heuvel EPJ. 1997. *Annu. Rev. Astron. Astrophys.* 35:69–100
- Kamae T, Karlsson N, Mizuno T, Abe T, Koi T. 2006. *Ap. J.* 647:692–708
- Kasliwal MM, Cenko SB, Kulkarni SR, et al. 2011. *Ap. J.* 735:94
- Kato M, Hachisu I. 1991a. *Ap. J.* 373:620–23
- Kato M, Hachisu I. 1991b. *Ap. J.* 383:761–65
- Kato M, Hachisu I. 1994. *Ap. J.* 437:802–26

- Kato M, Hachisu I. 2005. *Ap. J. Lett.* 633:L117–20
- Kato M, Hachisu I. 2007. *Ap. J.* 657:1004–12
- Kato M, Hachisu I. 2011. *Ap. J.* 743:157
- Kato M, Saio H, Henze M, et al. 2016. *Ap. J.* 830:40
- Kelner SR, Aharonian FA. 2008. *Phys. Rev. D* 78:034013
- Kenyon SJ, Truran JW. 1983. *Ap. J.* 273:280–88
- Kiewe M, Gal-Yam A, Arcavi I, et al. 2012. *Ap. J.* 744:10
- Kippenhahn R, Thomas HC. 1978. *Astron. Astrophys.* 63:265–72
- Knigge C. 2006. *MNRAS* 373:484–502
- Knigge C, Baraffe I, Patterson J. 2011. *Ap. J. Suppl.* 194:28
- Knigge C, King AR, Patterson J. 2000. *Astron. Astrophys.* 364:L75–79
- Kovetz A, Prialnik D. 1985. *Ap. J.* 291:812–21
- Kovetz A, Prialnik D, Shara MM. 1988. *Ap. J.* 325:828–36
- Krauss MI, Chomiuk L, Rupen M, et al. 2011. *Ap. J. Lett.* 739:L6
- Krautter J, Ögelman H, Starrfield S, Wichmann R, Pfeffermann E. 1996. *Ap. J.* 456:788–97
- Kwok S. 1983. *MNRAS* 202:1149–57
- Landau LD, Lifshitz EM. 1959. *Fluid Mech.* Oxford, UK: Pergamon
- Li KL, Metzger BD, Chomiuk L, et al. 2017. *Nat. Astron.* 1:697–702
- Linford JD, Chomiuk L, Nelson T, et al. 2017. *Ap. J.* 842:73
- Linford JD, Ribeiro VARM, Chomiuk L, et al. 2015. *Ap. J.* 805:136
- Liu WM, Li XD. 2016. *Ap. J.* 832:80
- Livio M. 1992. *Ap. J.* 393:516–22
- Livio M, Govarie A, Ritter H. 1991. *Astron. Astrophys.* 246:84–90
- Livio M, Shankar A, Burkert A, Truran JW. 1990. *Ap. J.* 356:250–54
- Livio M, Truran JW. 1987. *Ap. J.* 318:316–25
- Lloyd HM, O’Brien TJ, Bode MF. 1995. *Ap. Space Sci.* 233:317–21
- Lloyd HM, O’Brien TJ, Bode MF, et al. 1992. *Nature* 356:222–24
- MacDonald J. 1980. *MNRAS* 191:933–49
- MacDonald J, Fujimoto MY, Truran JW. 1985. *Ap. J.* 294:263–70
- Maoz D, Mannucci F. 2012. *Publ. Astron. Soc. Aust.* 29:447–65
- Margutti R, Metzger BD, Chornock R, et al. 2019. *Ap. J.* 872:18
- Martin P, Dubus G. 2013. *Astron. Astrophys.* 551:A37
- Martin P, Dubus G, Jean P, Tatischeff V, Dosne C. 2018. *Astron. Astrophys.* 612:A38
- Martin RG, Livio M, Schaefer BE. 2011. *MNRAS* 415:1907–12
- Mason E, Shore SN, De Gennaro Aquino I, et al. 2018. *Ap. J.* 853:27
- McEnery J, van der Horst A, Dominguez A, et al. 2019. *Bull. Am. Astron. Soc.* 51:245
- McLaughlin DB. 1943. *Publ. Mich. Obs.* 8:149–94
- McLaughlin DB. 1945. *Publ. Astron. Soc. Pac.* 57:69–80
- McLaughlin DB. 1947. *Publ. Astron. Soc. Pac.* 59:244–50
- Metzger BD. 2019. *Liv. Rev. Relativity* 23:1
- Metzger BD, Caprioli D, Vurm I, et al. 2016. *MNRAS* 457:1786–95
- Metzger BD, Finzell T, Vurm I, et al. 2015. *MNRAS* 450:2739–48
- Metzger BD, Hascoët R, Vurm I, et al. 2014. *MNRAS* 442:713–31
- Metzger BD, Pejcha O. 2017. *MNRAS* 471:3200–11
- Mikolajewska J. 2010. Presented at the Physics of Accreting Compact Binaries Conference, Univ. Kyoto, Kyoto, Japan. arXiv:1011.5657
- Morii M, Yamaoka H, Mihara T, Matsuoka M, Kawai N. 2016. *Publ. Astron. Soc. Jpn.* 68:S11
- Mróz P, Udalski A, Pietrukowicz P, et al. 2016. *Nature* 537:649–51
- Mukai K, Ishida M. 2001. *Ap. J.* 551:1024–30
- Mukai K, Orio M, Della Valle M. 2008. *Ap. J.* 677:1248–52
- Munari U, Henden A, Banerjee DPK, et al. 2015. *MNRAS* 447:1661–72
- Munari U, Henden A, Kiyota S, et al. 2002. *Astron. Astrophys.* 389:L51–56

- Munari U, Ribeiro VARM, Bode MF, Saguner T. 2011. *MNRAS* 410:525–34
- Murphy EJ, Bolatto A, Chatterjee S, et al. 2018. In *Science with a Next Generation Very Large Array*, ed. E Murphy. *ASP Conf. Ser.* 517:3. San Francisco: ASP
- Naito H, Mizoguchi S, Arai A, et al. 2012. *Astron. Astrophys.* 543:A86
- Nauenberg M. 1972. *Ap. J.* 175:417–30
- Negoro H, Kohama M, Serino M, et al. 2016. *Publ. Astron. Soc. Jpn.* 68:S1
- Nelemans G, Siess L, Repetto S, Toonen S, Phinney ES. 2016. *Ap. J.* 817:69
- Nelson T, Chomiuk L, Roy N, et al. 2014. *Ap. J.* 785:78
- Nelson T, Donato D, Mukai K, Sokoloski J, Chomiuk L. 2012. *Ap. J.* 748:43
- Nelson T, Mukai K, Chomiuk L, et al. 2020. *MNRAS* 500:2798–812
- Nelson T, Mukai K, Li KL, et al. 2019. *Ap. J.* 872:86
- Nelson T, Orio M, Cassinelli JP, et al. 2008. *Ap. J.* 673:1067–79
- Ness JU, Beardmore AP, Osborne JP, et al. 2015. *Astron. Astrophys.* 578:A39
- Ness JU, Osborne JP, Dobrotka A, et al. 2011. *Ap. J.* 733:70
- Ness JU, Osborne JP, Henze M, et al. 2013. *Astron. Astrophys.* 559:A50
- Nielbock M, Schmidtobreick L. 2003. *Astron. Astrophys.* 400:L5–8
- Nomoto K. 1982. *Ap. J.* 253:798–810
- Nomoto K, Kondo Y. 1991. *Ap. J. Lett.* 367:L18–22
- Nomoto K, Saio H, Kato M, Hachisu I. 2007. *Ap. J.* 663:1269–76
- Nomoto K, Thielemann FK, Yokoi K. 1984. *Ap. J.* 286:644–58
- O’Brien TJ, Bode MF. 2008. See Bode & Evans 2008, pp. 285–307
- O’Brien TJ, Bode MF, Porcas RW, et al. 2006. *Nature* 442:279–81
- O’Brien TJ, Lloyd HM. 1994. *Ap. Space Sci.* 216:167–72
- O’Brien TJ, Lloyd HM, Bode MF. 1994. *MNRAS* 271:155–60
- Ögelman H, Krautter J, Beuermann K. 1987. *Ap. Space Sci.* 130:279–79
- Orio M. 2012. *Bull. Astron. Soc. India* 40:333–51
- Orio M, Behar E, Gallagher J, et al. 2013. *MNRAS* 429:1342–53
- Orio M, Drake JJ, Ness JU, et al. 2020. *Ap. J.* 895:80
- Orio M, Parmar A, Benjamin R, et al. 2001. *MNRAS* 326:L13–18
- Osborne JP. 2015. *J. High Energy Astrophys.* 7:117–25
- Osborne JP, Page KL, Beardmore AP, et al. 2011. *Ap. J.* 727:124
- Özdönmez A, Ege E, Güver T, Ak T. 2018. *MNRAS* 476:4162–86
- Özdönmez A, Güver T, Cabrera-Lavers A, Ak T. 2016. *MNRAS* 461:1177–201
- Paczynski B. 1970. *Acta Astron.* 20:47–58
- Paczynski B. 1976. In *Structure and Evolution of Close Binary Systems, Proc. IAU Symp.* 73, ed. P Eggleton, S Mitton, J Whelan, pp. 75–80. Cambridge, UK: Cambridge Univ. Press
- Page KL, Beardmore AP, Osborne JP. 2020. *Adv. Space Res.* 66:1169–92
- Page KL, Osborne JP, Kuin NPM, et al. 2015. *MNRAS* 454:3108–20
- Page KL, Osborne JP, Wagner RM, et al. 2013. *Ap. J. Lett.* 768:L26
- Pala AF, Gänsicke BT, Breedt E, et al. 2020. *MNRAS* 494:3799–827
- Pala AF, Gänsicke BT, Townsley D, et al. 2017. *MNRAS* 466:2855–78
- Paresce F, Livio M, Hack W, Korista K. 1995. *Astron. Astrophys.* 299:823–34
- Patterson J. 1984. *Ap. J. Suppl.* 54:443–93
- Patterson J. 1998. *Publ. Astron. Soc. Pac.* 110:1132–47
- Patterson J. 2011. *MNRAS* 411:2695–716
- Patterson J, Oksanen A, Kemp J, et al. 2017. *MNRAS* 466:581–92
- Pavelin PE, Davis RJ, Morrison LV, Bode MF, Ivison RJ. 1993. *Nature* 363:424–26
- Paxton B, Bildsten L, Dotter A, et al. 2011. *Ap. J. Suppl.* 192:3
- Payne-Gaposchkin CH. 1957. *The Galactic Novae*. New York: Interscience
- Pejcha O, Metzger BD, Tomida K. 2016. *MNRAS* 455:4351–72
- Porter JM, O’Brien TJ, Bode MF. 1998. *MNRAS* 296:943–48
- Prialnik D. 1986. *Ap. J.* 310:222–37

- Prialnik D, Kovetz A. 1984. *Ap. J.* 281:367–74
- Prialnik D, Kovetz A. 1992. *Ap. J.* 385:665–69
- Prialnik D, Kovetz A. 1995. *Ap. J.* 445:789–810
- Prialnik D, Livio M. 1995. *Publ. Astron. Soc. Pac.* 107:1201–6
- Prialnik D, Shara MM. 1986. *Ap. J.* 311:172–82
- Prialnik D, Shara MM, Shaviv G. 1978. *Astron. Astrophys.* 62:339–48
- Quataert E, Fernández R, Kasen D, Klion H, Paxton B. 2016. *MNRAS* 458:1214–33
- Rappaport S, Verbunt F, Joss PC. 1983. *Ap. J.* 275:713–31
- Rauch T, Orio M, Gonzales-Riestra R, et al. 2010. *Ap. J.* 717:363–71
- Rawlings JMC, Williams DA. 1989. *MNRAS* 240:729–40
- Razzaque S, Jean P, Mena O. 2010. *Phys. Rev. D* 82:123012
- Ribeiro VARM, Chomiuk L, Munari U, et al. 2014. *Ap. J.* 792:57
- Ribeiro VARM, Munari U, Valisa P. 2013. *Ap. J.* 768:49
- Ritter H, Politano M, Livio M, Webbink RF. 1991. *Ap. J.* 376:177–85
- Roy N, Chomiuk L, Sokoloski JL, et al. 2012. *Bull. Astron. Soc. India* 40:293–310
- Ruiter AJ, Belczynski K, Fryer C. 2009. *Ap. J.* 699:2026–36
- Rupen MP, Mioduszewski AJ, Dhawan V. 2001. *IAU Circ.* 7728:3
- Saizar P, Ferland GJ. 1994. *Ap. J.* 425:755–66
- Sala G, Hernanz M. 2005. *Astron. Astrophys.* 439:1061–73
- Sanyal A. 1974. *Ap. J. Suppl.* 28:115–45
- Schaefer BE. 2010. *Ap. J. Suppl.* 187:275–373
- Schaefer BE. 2018. *MNRAS* 481:3033–51
- Schaefer BE. 2020. *MNRAS* 492:3343–58
- Schaefer BE, Boyd D, Clayton GC, et al. 2019. *MNRAS* 487:1120–39
- Schaefer GH, Brummelaar TT, Gies DR, et al. 2014. *Nature* 515:234–36
- Schreiber MR, Gänsicke BT, Rebassa-Mansergas A, et al. 2010. *Astron. Astrophys.* 513:L7
- Schreiber MR, Zorotovic M, Wijnen TPG. 2016. *MNRAS* 455:L16–20
- Schwab J, Quataert E, Bildsten L. 2015. *MNRAS* 453:1910–27
- Schwarz GJ, Hauschildt PH, Starrfield S, et al. 1998. *MNRAS* 300:931–44
- Schwarz GJ, Ness JU, Osborne JP, et al. 2011. *Ap. J. Suppl.* 197:31
- Schwarz GJ, Shore SN, Starrfield S, et al. 2001. *MNRAS* 320:103–23
- Scott AD. 2000. *MNRAS* 313:775–82
- Sequist ER, Bode MF. 2008. See Bode & Evans 2008, pp. 141–66
- Sequist ER, Duric N, Israel FP, et al. 1980. *Astron. J.* 85:283–93
- Sequist ER, Palimaka J. 1977. *Ap. J.* 217:781–87
- Senziani F, Skinner GK, Jean P, Hernanz M. 2008. *Astron. Astrophys.* 485:223–31
- Shafter AW. 2017. *Ap. J.* 834:196
- Shankar A, Livio M, Truran JW. 1991. *Ap. J.* 374:623–30
- Shara MM, Doyle T, Lauer TR, et al. 2017. *Ap. J.* 839:109
- Shara MM, Livio M, Moffat AFJ, Orio M. 1986. *Ap. J.* 311:163–71
- Shaviv NJ. 2001. *MNRAS* 326:126–46
- Shen KJ, Bildsten L. 2007. *Ap. J.* 660:1444–50
- Shen KJ, Bildsten L. 2009a. *Ap. J.* 692:324–34
- Shen KJ, Bildsten L. 2009b. *Ap. J.* 699:1365–73
- Shore SN. 2008. See Bode & Evans 2008, pp. 194–231
- Shore SN. 2014. In *Stellar Novae: Past and Future Decades*, ed. PA Woudt, VARM Ribeiro, *ASP Conf. Ser.* 490:145–61. San Francisco: ASP
- Shore SN, Aufdenberg JP. 1993. *Ap. J.* 416:355–67
- Shore SN, Augusteijn T, Ederoclitte A, Uthas H. 2011. *Astron. Astrophys.* 533:L8
- Shore SN, Schwarz GJ, De Gennaro Aquino I, et al. 2013. *Astron. Astrophys.* 549:A140
- Shore SN, Sonneborn G, Starrfield S, Gonzalez-Riestra R, Polidan RS. 1994. *Ap. J.* 421:344–49
- Shu FH, Lubow SH, Anderson L. 1979. *Ap. J.* 229:223–41

- Siebert T, Coc A, Delgado L, et al. 2018. *Astron. Astrophys.* 615:A107
- Silverman JM, Nugent PE, Gal-Yam A, et al. 2013. *Ap. J. Suppl.* 207:3
- Skopal A. 2019. *Ap. J.* 878:28
- Slavin AJ, O'Brien TJ, Dunlop JS. 1995. *MNRAS* 276:353–71
- Smith N, Chornock R, Li W, et al. 2008. *Ap. J.* 686:467–84
- Soker N. 2017. *MNRAS* 471:4839–43
- Soker N, Livio M. 1989. *Ap. J.* 339:268–78
- Sokoloski J, Lawrence S, Crotts APS, Mukai K. 2016. In *Accretion Processes in Cosmic Sources, APCS2016*, September 5–10, PoS(APCS2016)021
- Sokoloski JL, Luna GJM, Mukai K, Kenyon SJ. 2006. *Nature* 442:276–78
- Sokolovsky KV, Aydi E, Chomiuk L, et al. 2020a. *Astron. Telegr.* 13900:1
- Sokolovsky KV, Mukai K, Chomiuk L, et al. 2020b. *MNRAS* 497:2569–85
- Sparks WM. 1969. *Ap. J.* 156:569–95
- Starrfield S, Bose M, Iliadis C, et al. 2020. *Ap. J.* 895:70
- Starrfield S, Iliadis C, Hix WR. 2008. See Bode & Evans 2008, pp. 77–101
- Starrfield S, Iliadis C, Hix WR. 2016. *Publ. Astron. Soc. Pac.* 128:051001
- Starrfield S, Sparks WM, Truran JW. 1974. *Ap. J. Suppl.* 28:247–70
- Starrfield S, Sparks WM, Truran JW, Wiescher MC. 2000. *Ap. J. Suppl.* 127:485–95
- Starrfield S, Truran JW, Sparks WM. 1978a. *Ap. J.* 226:186–202
- Starrfield S, Truran JW, Sparks WM, Arnould M. 1978b. *Ap. J.* 222:600–3
- Starrfield S, Truran JW, Sparks WM, Kutter GS. 1972. *Ap. J.* 176:169–76
- Starrfield S, Truran JW, Wiescher MC, Sparks WM. 1998. *MNRAS* 296:502–22
- Steffen W, Koning N, Wenger S, Morisset C, Magnor M. 2011. *IEEE Trans. Vis. Comput. Graph.* 17:454–65
- Steinberg E, Metzger BD. 2018. *MNRAS* 479:687–702
- Steinberg E, Metzger BD. 2020. *MNRAS* 491:4232–46
- Strickland R, Blondin JM. 1995. *Ap. J.* 449:727–38
- Strope RJ, Schaefer BE, Henden AA. 2010. *Astron. J.* 140:34–62
- Taam RE. 1980. *Ap. J.* 237:142–47
- Taam RE, Spruit HC. 2001. *Ap. J.* 561:329–36
- Tajitsu A, Sadakane K, Naito H, Arai A, Aoki W. 2015. *Nature* 518:381–84
- Tanaka J, Nogami D, Fujii M, Ayani K, Kato T. 2011. *Publ. Astron. Soc. Jpn.* 63:159–69
- Tarasova TN. 2019. *Astrophysics* 62:475–90
- Tatischeff V, Hernanz M. 2007. *Ap. J. Lett.* 663:L101–4
- Taylor AR, Hjellming RM, Seaquist ER, Gehrz RD. 1988. *Nature* 335:235–38
- Taylor AR, Pottasch SR, Seaquist ER, Hollis JM. 1987. *Astron. Astrophys.* 183:38–46
- Thompson WT. 2017. *MNRAS* 470:4061–65
- Thoroughgood TD, Dhillon VS, Littlefair SP, Marsh TR, Smith DA. 2001. *MNRAS* 327:1323–33
- Tomsick J, Zoglauer A, Sleator C, et al. 2019. *Bull. Am. Astron. Soc.*, 51:7
- Townsley DM, Bildsten L. 2003. *Ap. J. Lett.* 596:L227–30
- Townsley DM, Bildsten L. 2004. *Ap. J.* 600:390–403
- Truran JW, Livio M. 1986. *Ap. J.* 308:721–27
- Tuthill PG, Monnier JD, Danchi WC. 1999. *Nature* 398:487–89
- Uthas H, Knigge C, Steeghs D. 2010. *MNRAS* 409:237–46
- Vishniac ET. 1994. *Ap. J.* 428:186–208
- Vlasov A, Vurm I, Metzger BD. 2016. *MNRAS* 463:394–412
- Vurm I, Metzger BD. 2018. *Ap. J.* 852:62
- Walter FM, Battisti A, Towers SE, Bond HE, Stringfellow GS. 2012. *Publ. Astron. Soc. Pac.* 124:1057–72
- Warner B. 1987. *MNRAS* 227:23–73
- Warner B. 1995. *Cataclysmic Variable Stars*. Cambridge, UK: Cambridge Univ. Press
- Weiler KW, Panagia N, Montes MJ, Sramek RA. 2002. *Annu. Rev. Astron. Astrophys.* 40:387–438
- Wendeln C, Chomiuk L, Finzell T, Linford JD, Strader J. 2017. *Ap. J.* 840:110
- Weston JHS, Sokoloski JL, Chomiuk L, et al. 2016a. *MNRAS* 460:2687–97

- Weston JHS, Sokoloski JL, Metzger BD, et al. 2016b. *MNRAS* 457:887–901
- Whelan J, Iben I Jr. 1973. *Ap. J.* 186:1007–14
- Williams R. 2012. *Astron. J.* 144:98
- Williams R. 2013. *Astron. J.* 146:55
- Williams R. 2016. In *J. Phys. Conf. Ser.* 728:042001
- Williams R, Mason E. 2010. *Ap. Space Sci.* 327:207–17
- Williams R, Mason E, Della Valle M, Ederoclite A. 2008. *Ap. J.* 685:451–62
- Williams RE. 1990. In *Physics of Classical Novae, Lect. Notes Phys.*, Vol. 369, ed. A Cassatella, R Viotti, pp. 213–227. Berlin/Heidelberg: Springer
- Williams RE. 1994. *Ap. J.* 426:279–87
- Williams RE, Ney EP, Sparks WM, et al. 1985. *MNRAS* 212:753–66
- Williams SC, Darnley MJ, Bode MF, Shafter AW. 2016. *Ap. J.* 817:143
- Wolf WM, Bildsten L, Brooks J, Paxton B. 2013. *Ap. J.* 777:136
- Wolf WM, Townsend RHD, Bildsten L. 2018. *Ap. J.* 855:127
- Woodward CE, Gehrz RD, Jones TJ, Lawrence GF. 1992. *Ap. J. Lett.* 384:L41–45
- Woosley SE, Blinnikov S, Heger A. 2007. *Nature* 450:390–92
- Woudt PA, Steeghs D, Karovska M, et al. 2009. *Ap. J.* 706:738–46
- Yaron O, Prialnik D, Shara MM, Kovetz A. 2005. *Ap. J.* 623:398–410
- Yuan W, Zhang C, Ling Z, et al. 2018. *Proc. SPIE Conf. Ser.* 10699:1069925
- Zdziarski AA, Svensson R. 1989. *Ap. J.* 344:551–66
- Zhao X, Fuller J. 2020. *MNRAS* 495:249–65
- Zorotovic M, Schreiber MR, Gänsicke BT, et al. 2011. *Astron. Astrophys.* 536:L3