



New Mass Estimates for Massive Binary Systems: A Probabilistic Approach Using Polarimetric Radiative Transfer

Andrew G. Fullard¹ , John T. O'Brien¹ , Wolfgang E. Kerzendorf¹ , Manisha Shrestha² , Jennifer L. Hoffman³ ,
Richard Ignace⁴ , and Patrick van der Smagt^{5,6}

¹ Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA; fullarda@msu.edu

² Astrophysics Research Institute, Liverpool John Moores University, Liverpool Science Park IC2, 146 Brownlow Hill, Liverpool, L3 5RF, UK

³ Department of Physics and Astronomy, University of Denver, Denver, CO 80210 USA

⁴ Department of Physics and Astronomy, East Tennessee State University, Johnson City, TN 37614 USA

⁵ Machine Learning Research Lab, Volkswagen Group, Munich, Germany

⁶ Faculty of Informatics, Eötvös Loránd University, Budapest, Hungary

Received 2022 February 1; revised 2022 February 18; accepted 2022 February 19; published 2022 May 5

Abstract

Understanding the evolution of massive binary stars requires accurate estimates of their masses. This understanding is critically important because massive star evolution can potentially lead to gravitational-wave sources such as binary black holes or neutron stars. For Wolf–Rayet (WR) stars with optically thick stellar winds, their masses can only be determined with accurate inclination angle estimates from binary systems which have spectroscopic $M \sin i$ measurements. Orbitally phased polarization signals can encode the inclination angle of binary systems, where the WR winds act as scattering regions. We investigated four Wolf–Rayet + O star binary systems, WR 42, WR 79, WR 127, and WR 153, with publicly available phased polarization data to estimate their masses. To avoid the biases present in analytic models of polarization while retaining computational expediency, we used a Monte Carlo radiative-transfer model accurately emulated by a neural network. We used the emulated model to investigate the posterior distribution of the parameters of our four systems. Our mass estimates calculated from the estimated inclination angles put strong constraints on existing mass estimates for three of the systems, and disagree with the existing mass estimates for WR 153. We recommend a concerted effort to obtain polarization observations that can be used to estimate the masses of WR binary systems and increase our understanding of their evolutionary paths.

Unified Astronomy Thesaurus concepts: Massive stars (732); Binary stars (154); Polarimetry (1278); Bayesian statistics (1900); Radiative transfer (1335)

1. Introduction

Massive stars enrich the interstellar medium with heavy elements via stellar winds and supernova explosions. Their high-mass remnants, such as neutron stars and black holes, have been detected with gravitational-wave observations (e.g., Abbott et al. 2016, 2017a). The evolutionary paths that massive stars take to reach different types of supernovae, or even core collapse without a luminous event, are strongly dependent on their masses and mass-loss rates (Puls et al. 2008; Langer 2012; Abbott et al. 2016; Woosley 2019).

A large fraction of sufficiently massive (18–80 M_{\odot}) individual stars are expected to evolve to the Wolf–Rayet (WR) stage (Sander et al. 2019), which is characterized by strong emission line spectra and high mass-loss rates ($\sim 10^{-5} M_{\odot} \text{ yr}^{-1}$; Crowther 2007). However, the optically thick winds of WR stars make it impossible to measure their masses using $\log g$. Fortunately, massive stars typically occur in binary systems (Mason et al. 2009; Sana et al. 2014) so it is possible to obtain mass estimates using orbital models. Massive binaries can evolve into compact binaries that emit gravitational waves (Abbott et al. 2017b). Aside from using binaries for mass estimation, binary interactions can have a strong effect on the pre-supernova masses and structure of massive stars (e.g., Laplace et al. 2021) and more than 50% of massive binaries interact during their lifetime

(Sana et al. 2012, 2013). Mass transfer as part of these interactions reduces the mass at which stars can reach the WR stage (McClelland & Eldridge 2016). It is therefore important to constrain the masses of massive binary systems so that we can understand whether or not the WR components were formed as a result of binary interactions.

Only three massive Milky Way binaries with WR stars have masses derived using both spectroscopic and visual binary orbits: γ^2 Velorum (Lamberts et al. 2017), WR 133 (Richardson et al. 2021), and WR 140 (Thomas et al. 2021). Fifteen Milky Way WR binary stars have mass estimates derived from spectroscopy, photometry, and polarimetry, although their uncertainties are not well constrained (Crowther 2007). Outside the Milky Way, some progress has been made to determine WR + O masses in the Magellanic Clouds using radiative-transfer models of WR and O star spectra (Shenar et al. 2016, 2019).

Measuring the masses of noneclipsing binary stars is a difficult task. The most robust method is to combine spectroscopic radial velocity orbital solutions that provide $M \sin i$ values with astrometric orbits to obtain the inclination angle of the system, and thus derive masses. Obtaining the inclination angle requires alternative solutions in the absence of a visual orbit. Methods that have been used to determine inclination angles without a visual orbit include photometric models that are reliant on wind eclipses (e.g., Lamontagne et al. 1996), stellar spectral synthesis models (e.g., Martins et al. 2005), models of colliding winds (e.g., Hill et al. 2000), and polarimetric models (e.g., Brown et al. 1978). These methods span a wide range of uncertainties, with the most precise being



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

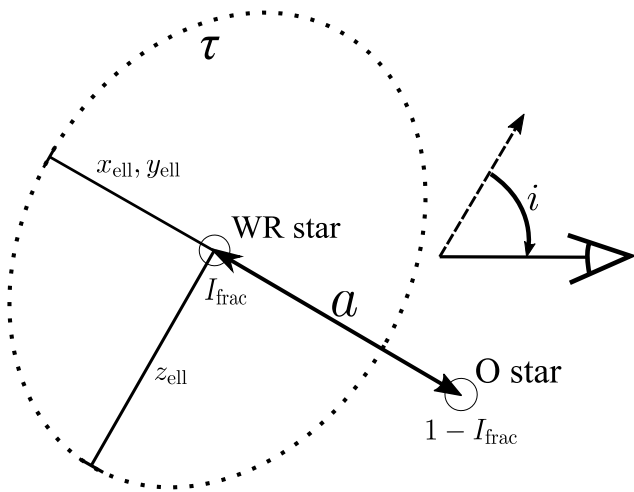


Figure 1. Schematic diagram of the binary model. The diagram is presented such that the system is viewed at an inclination angle of 60° measured from the normal to the orbital plane. A definition of the model inclination angle i is shown in the bottom right. The stars are the labeled circles of radius r_* ; the dotted ellipse is the scattering region of electron scattering optical depth τ . The diagram is not to-scale ($r_* \ll a$). Other parameters are described in Section 3.1 and Table 2.

the photometric models (on the order of 50% formal uncertainty in WR mass; Lamontagne et al. 1996) and the least precise the models based on stellar spectral synthesis. The true uncertainties of the photometric models may be larger (see footnote 11 of Lamontagne et al. 1996). The stellar spectral synthesis method does not report uncertainties, but rather a likely mass range for the O star based on its spectral class, and that informs the inclination angle of the system to determine the WR mass in concert with previous inclination angle estimates (Vanbeveren et al. 1998, 2020). Furthermore, the methods do not consistently agree on the masses for a given WR + O system, exacerbated by the range of reported uncertainties (or lack thereof). For example, previous estimates for the mass of WR 79 from the photometric and polarimetric methods disagree by more than 3σ (Lamontagne et al. 1996). This motivates the more robust method that we describe in this paper.

The polarimetric method mentioned previously works by modeling the time-varying linear polarization signal that occurs when the hot winds of the WR star act as electron-scattering regions for photons produced both by the WR star and its companion (see Figure 1, showing an electron-scattering region and WR + O illumination sources). This produces a constant linear polarization from the WR star when the wind is asymmetric (Brown & McLean 1977). Illumination by the secondary star produces the time-varying linear polarization, which is partly dependent on the orbital inclination angle i (Brown et al. 1978), and this provides a way to measure inclination angles and thus masses. To recover accurate uncertainties for the inclination angle, we must determine the posterior distribution of the inclination angle through repeated model evaluations.

There are two primary methods for modeling polarimetric signals: analytic methods and numerical radiative-transfer methods. Analytic models suffer from a variety of challenges, including a bias toward high inclination angles and limitations to optically thin scattering regions ($\tau < 0.01$; Aspin et al. 1981; Brown et al. 1982; Simmons et al. 1982; Wolinski & Dolan 1994; Buchner 2019). In contrast, numerical radiative-transfer methods are not biased by inclination angle and can

include the effects of high optical depths such as multiple scattering (e.g., Hoffman et al. 2003) though these simulations are computationally expensive. We have used an enhanced version of the Monte Carlo radiative-transfer (MCRT) code SLIP (Hoffman 2007; Huk 2017; Shrestha et al. 2018) to perform numerical simulations of binary stars in circular orbits, for cases where one star has a scattering region surrounding it. To obtain sufficient signal-to-noise for numerical polarimetric estimates, which are often much less than 1% of the total emitted intensity, SLIP requires large numbers of photons to be propagated through the simulation, slowing its evaluation times. To reconstruct the posterior distribution of the model parameters, we must evaluate the model many times. This motivates the adoption of acceleration methods.

One such method to accelerate models is that of emulators based on neural networks. Emulators relate the approximated model to its input parameters. Neural networks are capable of rapidly approximating any model (see, e.g., Cybenko 1989; Hornik et al. 1989). In our case, the trained neural network rapidly generates an approximated polarimetric SLIP model based on the input parameters. This means the emulator can be rapidly sampled to obtain many models and infer parameters using Bayesian statistics with maximum-likelihood testing for problems that lie within the range of the training data. This technique has been successfully applied to spectral models (Czekala et al. 2015; O’Brien et al. 2021; Kerzendorf et al. 2021).

In this paper, we use an enhanced version of the SLIP radiative-transfer code accelerated by an emulator to produce a robust probabilistic estimate of the inclination angles and thus masses of four WR + O binary systems in the Galaxy with existing polarimetric observations. In Section 2, we describe the polarimetric data and the WR + O systems. In Section 3, we describe our radiative-transfer model and emulator procedure. In Section 4, we present our results and discuss their implications. Appendix A describes the SLIP code in detail, and we validate our model in Appendix B. In Appendix C, we provide additional details about our emulator method. The full parameter spaces for our investigations of each system are presented in Appendix D.

2. Data

We investigated four WR + O binary systems with publicly available, phased polarization data. We sourced the data from St-Louis et al. (1987; WR 42 and WR 79) and St-Louis et al. (1988; WR 127 and WR 153). All of the objects were observed with the Minipol polarimeter at Las Campanas, Chile from 1985 to 1986 (Frecker & Serkowski 1976), using a blue filter with central wavelength 4700 Å and FWHM 1800 Å. Table 1 lists the objects and their estimated orbital characteristics. We refer to the systems via their WR catalog number as defined in Crowther (2015).⁷

These systems have similar orbital periods (~ 1 week) and circularized orbits. They are split into two each of WC- and WN-type WR stars with O-type companions. We used the orbital phases calculated by St-Louis et al. (1987, 1988) despite the determination of newer ephemerides (e.g., Nazé et al. 2021), because any uncertainties in the ephemerides are greatly amplified by the large number of orbits of each system that have occurred since the polarimetric data were taken. The ephemerides in St-Louis et al. (1987, 1988) are closer in date to

⁷ <http://pacrowther.staff.shef.ac.uk/WRcat/index.php>

Table 1
WR + O System Spectral Types and Orbital Data

WR	Spectral Type	Period (d)	Source	$i(^{\circ})$	Source	$\Omega(^{\circ})$	Source
42	WC7 + O7V	7.8912	1	36–44	2, 3	63.1–72.8	6
79	WC7 + O5-8	8.8911	1	29–45	2, 3	103.2–111.2	6
127	WN5o + O8.5V	9.5550	2	55–90	4	83.6–103	7
153	WN6o/CE + O3-6	6.6887	3	65–78	5	93.5–95.4	7

Note. Spectral types are taken from Crowther (2015).

References. 1: Hill et al. (2000); 2: Lamontagne et al. (1996); 3: Hill et al. (2002); 4: de La Chevrotière et al. (2011); 5: Demers et al. (2002); 6: St-Louis et al. (1987); 7: St-Louis et al. (1988).

the observations, and thus the effects of uncertainties on the orbital phase are lowest.

3. Methods

For this investigation, we used an enhanced version of the MCRT code SLIP (Hoffman 2007; Huk 2017; Shrestha et al. 2018, 2021) that includes binary star capabilities (Fullard 2020). Appendix A gives more information about the design of SLIP. Using this code, we created a model system and trained a neural-network emulator in order to analyze the observations.

3.1. Model Binary System

Our model binary system is described using four parameters: I_{frac} , τ , z_{ell} , and i . Figure 1 shows a schematic diagram of the model. Here, I_{frac} is the fractional intensity of the central photon source (the WR star) relative to the total number of photons. Thus, the O star intensity is $1 - I_{\text{frac}}$. τ is the electron scattering optical depth of the scattering region at $i = 90^{\circ}$. z_{ell} is the half-height of the ellipsoidal scattering region as measured normal to the orbital plane, in arbitrary units. x_{ell} and y_{ell} are the x and y extents of the scattering region and are held constant and equal to each other. If $z_{\text{ell}} = x_{\text{ell}} = 1.5$, the scattering region is spherical. If $z_{\text{ell}} \neq x_{\text{ell}}$, the scattering region can be considered prolate or disk-like. i is the inclination (viewing) angle, measured from the normal to the orbital plane. The remaining parameters shown in Figure 1 are held constant and are described in Table 2. SLIP produces simulations of the fractional q and u polarization signal produced by the electron scattering in this model, normalized from the Stokes Q and U vectors as $q = Q/I$ and $u = U/I$.

The z_{ell} parameter should not be thought of as the true geometry of the scattering region. Rather, it is a measure of the asymmetry of the scattering region or the density distribution within it (e.g., a disk-like scattering region produces polarization signals similar to a spherical scattering region with equatorially enhanced density). We assume only electron scattering is an important contributor to the polarization signal, which is a reasonable assumption for the case of a highly ionized WR star wind. To reduce the number of parameters required for the model, we assume a constant electron density in the scattering region and a constant albedo of unity, as well as total ionization of the wind. The uniform density means that for scattering regions that are asymmetric, there is a nonlinear relationship between z_{ell} and τ as a function of i . The constant albedo of one means that the polarization signal is independent of wavelength.

We assume that the stars are sufficiently separated (parameter a) so that eclipse effects are negligible, which is appropriate for the range of expected inclination angles and

Table 2
Constant Model Parameters for the Emulator

Parameter	Value	Description
r^*	4.65×10^{-4}	Stellar radii
a	3.1	Binary separation
$x_{\text{ell}}, y_{\text{ell}}$	1.5	x, y extent of the scattering region

Note. Lengths are displayed as relative values (see Figure 1).

Table 3
Sample Space of the Model

Parameter	Min	Max	Description
I_{frac}	0.0	1.0	Fractional intensity of the WR star
τ	0.0	1.0	Electron scattering optical depth
z_{ell}	0.2	3.1	z extent of the scattering region
i	0°	180°	Orbital inclination angle

periods of our chosen objects. Fox (1994) showed that occultation is only important in extremely close binary systems, where the separation of stars is less than 10 times the radius of the primary. In the case of WR 127, eclipsing inclination angles are unlikely as they would result in a mass for the system that is significantly lower than any other estimate (see comparisons and discussion in Section 4). An exploration of eclipse effects would require varying star-separation and star-size parameters, with a commensurate increase in the number of model runs required to train the emulator.

The sample space of our model, tabulated in Table 3, was chosen to represent the four WR + O binaries described in Section 2. The range of scattering region geometries encapsulated by z_{ell} transitions from disk-like ($0.2 < z_{\text{ell}} < 1.5$) through to prolate ($1.5 < z_{\text{ell}} < 3.1$) distributions. For the WR + O star binaries we have investigated in this paper, the scattering region approximates various geometries (or, equivalently, density distributions) of the WR star wind. The optical depth range covers scattering from optically thin material (the outer regions of WR star winds) to more optically thick material (closer to the location of emission lines that drive the wind; Sander et al. 2020). The inclination angle range is valid for Stokes Q and U polarization measurements, which can distinguish between above and below the orbital plane, or equivalently the orbital direction about the orbital plane normal vector.

3.2. Transforming to the Observational Frame

When a binary system is observed in polarized light, two additional factors arise that we do not model using SLIP. These

are interstellar polarization (ISP) and the position angle of the binary system’s orbit as projected onto the sky.

The ISP in the Milky Way at optical wavelengths is described by the empirical Serkowski law (Serkowski et al. 1975). We treat it as constant with time for our purposes, because the observations were taken over a timescale that is much shorter than the expected timescale of ISP change ($\sim 30\text{--}50$ Myr; Voshchinnikov & Hirashita 2014).

In our case of broadband filter polarimetry, we add ISP to the q and u models separately by calculating the q_{ISP} and u_{ISP} at the peak wavelength of each filter. This ISP estimate is therefore a flux-weighted ISP across the observed filter.

The observed position angle of the binary orbit manifests itself as a rotation of the observed polarization in the q – u plane, and thus applying this rotation to the model moves it to the observational frame. The rotation is applied to the model using

$$q = p \cos 2(\theta - \theta_R), \quad (1)$$

$$u = p \sin 2(\theta - \theta_R), \quad (2)$$

where $p = \sqrt{q_{\text{em}}^2 + u_{\text{em}}^2}$ (total polarization), θ is the position angle of the observation, and θ_R is the rotation angle of the model. Following Villar-Sbaiffi et al. (2005), we recover the angle of the line of ascending nodes, Ω , using $\Omega = (\theta_R + \pi)/2$.

A final correction is needed to match the model’s orbitally phased polarization signal to the observed phases. We project the model onto an orbital phase grid of 40 points, shifted to the center of each orbital phase bin that is output from SLIP. We linearly interpolate the model onto the observed phases so that the likelihood function (see Section 3.4.2) is valid for each observed phase ϕ .

3.3. Emulator Pipeline

Typical run times for our model in the sample space are around 2–4 hr (see Appendix C.1). This makes model evaluations intractable for producing probabilistic estimates of parameters. We accelerate the process with an *emulator*, which is a machine-learning method, in this case a neural network, that is trained to relate model parameters to the output of the model. Such emulators have been successfully applied to astrophysics problems (e.g., Type Ia supernovae; O’Brien et al. 2021). The emulator learns the relationship between model parameters and output by fitting to a parameter set and its associated model outputs (the *training data*). The user chooses the training data, and these constrain the region over which the emulator is valid.

The construction of an emulator thus requires training data but also a specified neural-network architecture. We trained the emulator on the model described in Section 3.1. The emulator training set is described in Appendix C.1 and the emulator architecture is described in Appendix C.2. The emulator increases the model evaluation speed by a factor of $\sim 10^8$ over running SLIP for a typical model in the sample space in Table 3. The emulator representation of SLIP for the training set is accurate to within 5% of the SLIP model.

3.4. Parameter Inference

Two components are required for a rigorous parameter inference from our model: a prior distribution and a likelihood function. Obtaining the posterior distribution requires the selection of a prior probability distribution across the parameter space and a likelihood function to compare the data with the

Table 4
WR + O ISP Estimates

WR Number	q_{ISP} (%)	u_{ISP} (%)
42	-0.085 ± 0.164	-1.13 ± 0.179
79	-0.284 ± 0.020	-0.214 ± 0.025
127	0.618 ± 0.149	0.661 ± 0.140
153	-0.182 ± 0.170	4.16 ± 0.018

Note. Calculated from Fullard et al. (2020), except for WR 42 (see text).

model. We describe the prior distributions in Section 3.4.1. Polarimetric models are produced by the emulator from samples drawn from the prior distribution. We calculate the likelihood of an individual prior input using a χ^2 likelihood function, described in Section 3.4.2. We describe our exploration of the posterior via Monte Carlo sampling of the priors in Section 3.4.3.

3.4.1. Priors

We constructed the Bayesian prior of the parameters listed in Table 2 using information about the target object. We sampled the parameters I_{frac} , τ , and z_{ell} uniformly over the entire sample space of the model, because we do not have existing constraints on these parameters that are narrower than the sample space. We chose a normal distribution for the prior of the inclination angle, i , with the standard deviation set as the half-range of previous inclination angle estimates for each system in Table 1, and the mean taken to be the center of the range. In this way, we incorporate existing information about the system, including photometric and spectroscopic measurements. I_{frac} is allowed to vary over the entire emulator sample space.

For the transformation parameters described in Section 3.2, we sample q_{ISP} and u_{ISP} from Gaussian distributions with the peaks located at the estimated ISP from Fullard et al. (2020) for each target. We take σ values from the propagated uncertainty of q_{ISP} and u_{ISP} , using the uncertainties reported by Fullard et al. (2020). For WR 42, we refit the ISP following the methods in Fullard et al. (2020) without the intrinsic constant polarization component, because our model should produce intrinsic constant polarization by varying I_{frac} , τ , and z_{ell} . The ISP values are presented in Table 4. The transformation parameter, θ_R , is taken as a uniform prior from $-\pi/2 \leq \theta_R \leq \pi/2$ because existing estimates for the directly related orbital parameter, Ω , were produced with the biased analytic polarization model. Note that this specific range occurs because the polarization position angle is degenerate (repeating in π instead of 2π rotations in the q – u plane).

3.4.2. Likelihood

We formulated the likelihood function with a log-likelihood function

$$\log \mathcal{L}(\mathbf{t}) = -\frac{1}{2} \sum_{\phi} \left[\left(\frac{q_{\text{em}}(\mathbf{t}) - q_{\text{obs}}}{\sigma_q} \right)^2 + \left(\frac{u_{\text{em}}(\mathbf{t}) - u_{\text{obs}}}{\sigma_u} \right)^2 \right], \quad (3)$$

where q_{em} and u_{em} are the predicted q and u Stokes vectors from the emulator using parameters \mathbf{t} . q_{obs} and u_{obs} are the observed Stokes vectors, and σ_q and σ_u are the associated uncertainties of the observed Stokes vectors. The likelihood

function is summed over ϕ , the orbital phase of the individual q_{obs} and u_{obs} data points. Stokes q and u are independent, hence we simply sum them as part of the logarithmic transformation of the likelihood function.

The transformation to the observational frame described in Section 3.2 replaces q_{em} and u_{em} in Equation (3) with transformed versions. The transformations are applied in the following order: interpolation to the observed phases, then rotation to the observed position angle, then addition of q_{ISP} and u_{ISP} to q_{em} and u_{em} .

3.4.3. Sampling

Following the methodology of O’Brien et al. (2021), we used the UltraNest⁸ (Buchner 2021) nested sampling package to achieve robust sampling of the posterior distribution with the MLFriends algorithm (Buchner 2016, 2019). The posterior distribution was explored with 800 live points. It converged (based on the UltraNest measures) after $\sim 25,000$ iterations and required $\sim 6 \times 10^6$ model evaluations for each target. The convergence of the sampling process typically took 2 hr.

4. Results and Discussion

We sampled the posterior distribution following the prescription described in Section 3.4.3, which produced a six-dimensional posterior distribution across the parameter space (one dimension for each parameter). Figure 2 shows the posterior distributions for the masses of the WR star components, derived from the inclination angles using $M_{\text{WR}} \sin^3 i$ measurements taken from Davis et al. (1981), Seggewiss (1974), de La Chevrotière et al. (2011), and Lamontagne et al. (1996). Existing mass estimates for the systems we investigated are based on spectral class (Shao & Li 2016; Vanbeveren et al. 2020), analysis of their polarimetric behavior (St-Louis et al. 1987, 1988), colliding-wind models (Hill et al. 2000), and photometric eclipse models (Lamontagne et al. 1996). Of these existing mass estimates, the spectral-class models do not provide uncertainties.

Figure 3 shows the posterior distributions of masses for the O star components. Table 5 shows the 90% credible interval(s) for the component masses of each object, and we discuss these results below. The credible intervals were calculated using the highest-density interval algorithm of the ARVIZ Python package (Kumar et al. 2019). We display the full set of posterior distributions as corner plots in Appendix D.

4.1. WR 42

The WR mass 90% credible interval calculated from our method for WR 42 ($16.2\text{--}44.0 M_{\odot}$) agrees with existing mass estimates from all of the methods displayed in Figure 2, taking into account their uncertainties. For the previous polarimetric analysis (St-Louis et al. 1987), the estimate was $8\text{--}14 M_{\odot}$; however, this did not fully take into account the bias in the analytical polarimetric model (Wolinski & Dolan 1994) and thus the true uncertainty is larger. The photometric method mass estimate (Lamontagne et al. 1996) did not provide WR mass uncertainties, though the reported inclination angle of $40.3^{\circ} \pm 2.9$ produces a WR mass range of $11\text{--}16 M_{\odot}$. It is important to note that the inclination angle uncertainty here is the “formal” uncertainty and the real values can be larger (see footnote 11 of Lamontagne et al. 1996). The wind model of

Hill et al. (2000) produces a very uncertain mass estimate of $54 \pm 43 M_{\odot}$ so we provide it for the sake of completeness. Our derived 90% credible interval for the O star mass agrees with the existing mass estimates from spectral-class models, encapsulating the Vanbeveren et al. (2020) estimated mass range of $27\text{--}37 M_{\odot}$. The WR mass reported by Vanbeveren et al. (2020) is “adopted” as $16 M_{\odot}$, based on the previous reported inclination angles for the system (and thus similarly uncertain as the polarimetric and photometric mass estimates).

It is interesting to note that the polarimetric measurement from St-Louis et al. (1987), produced using the same data, overlaps (within the St-Louis et al. 1987 polarimetric uncertainties) a small secondary peak in our posterior distribution. This may indicate a local minimum was found with the analytic model. If the uncertainty in $M \sin^3 i$ reported by Davis et al. (1981) for WR 42 is used to calculate the mass range, this extends both the WR and O star mass 90% credible intervals to $25\text{--}48 M_{\odot}$ and $19\text{--}81 M_{\odot}$, respectively.

Fullard et al. (2020) reported that the mean polarization signal of WR 42 had a large constant intrinsic component in both q and u . As noted in Section 3.4.1, we fit the ISP without the intrinsic polarization terms. As a result, the model predicts a strongly asymmetric WR wind that is more disk-like, or equivalently a symmetric WR wind with a strong density increase parallel to the orbital plane. This may explain the intrinsic polarization, and potentially represents the asymmetry produced by the colliding wind structures proposed by Hill et al. (2000).

4.2. WR 79

We find multiple peaks in the WR 79 mass posterior distribution. This creates a series of three modes with associated credible intervals, which we report from low to high mass in Table 5. The first credible interval ($8.3\text{--}12.8 M_{\odot}$) agrees with both the spectral-class-model mass lower limit of $10 M_{\odot}$ (Vanbeveren et al. 2020) and the photometric-model mass estimate of $9\text{--}13 M_{\odot}$ (Lamontagne et al. 1996). The second, more mathematically probable posterior peak and associated credible interval ($15.3\text{--}62.8 M_{\odot}$) matches the mass estimate from the colliding-wind model of Hill et al. (2000). However, we note that the wind-model mass estimate is very uncertain at $27 \pm 17 M_{\odot}$ (Hill et al. 2000) and therefore it also overlaps with the first credible interval. The third credible interval ($89.5\text{--}152.2 M_{\odot}$) is very unlikely, and suggests masses well outside any previous model. None of the intervals agree with the previous polarimetric mass estimate that used the same data (St-Louis et al. 1987), confirming that the basic analytic model is unlikely to provide accurate results for binary systems with low inclination angles. Note that our model recovers an inclination angle greater than 90° , which is possible for polarimetry because the sign of Stokes U breaks the degeneracy present in spectroscopic orbital models. This inclination angle is equivalent to a reversed orbital motion compared to the model’s motion in the inclination range $0^{\circ} \leq i \leq 90^{\circ}$.

The physical implications of the multiple modes are of some interest. Scale renderings of the models are shown in Figure 4. The first inclination angle (mass) mode (top in Figure 4) corresponds to models with an extremely asymmetric WR wind, low electron scattering optical depth of $\tau < 0.1$, and the majority of emission ($\sim 95\%$) originating from the O star. The asymmetric wind can also be considered to be a strong polar density enhancement. However, such an enhancement is unlikely. Distortions to the WR

⁸ <https://johannesbuchner.github.io/UltraNest/>

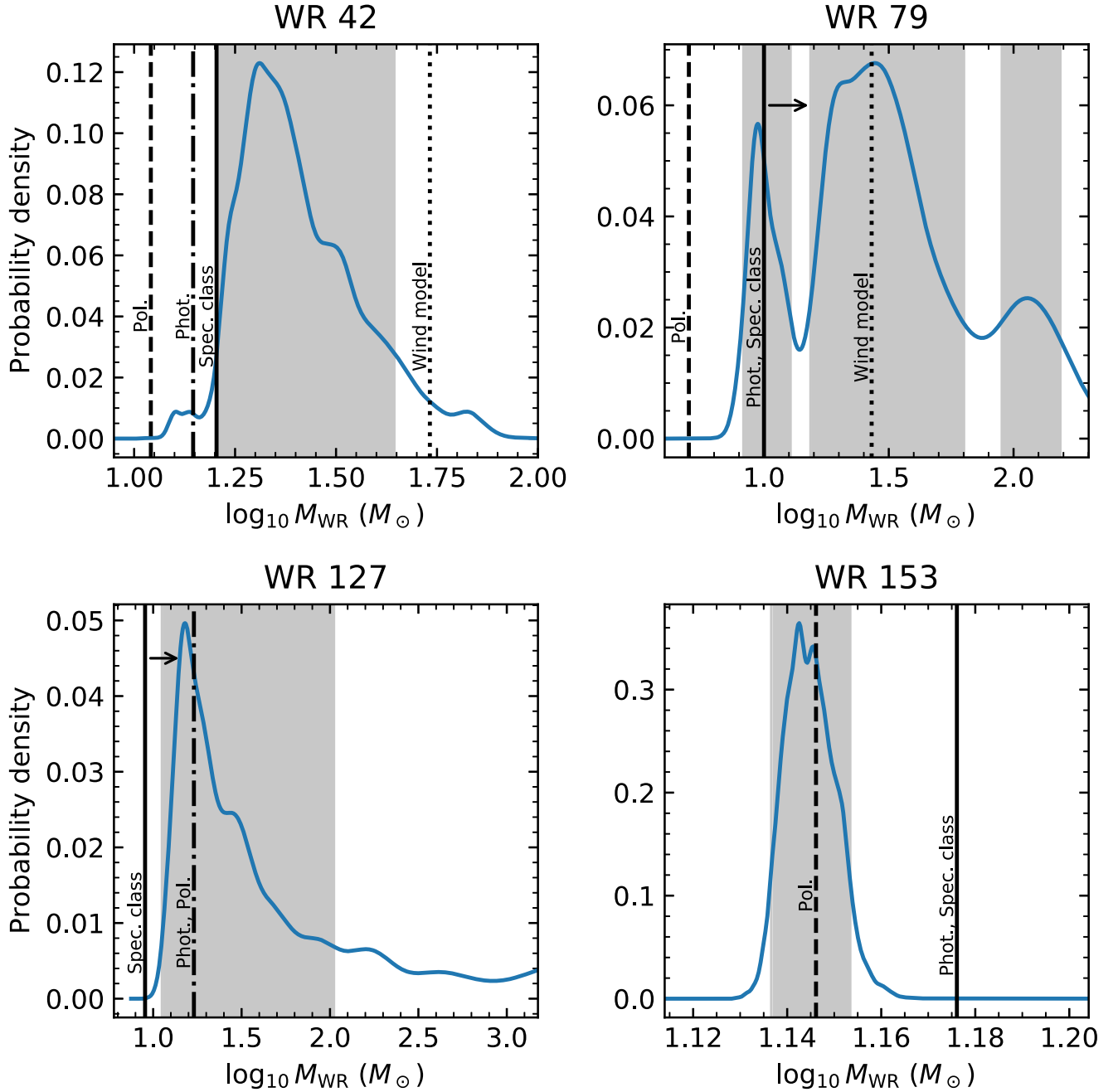


Figure 2. WR 42, WR 79, WR 127, and WR 153 WR star mass posterior distributions, shown as solid curves. The gray shaded regions show the 90% credible interval of the distribution calculated using the highest-density interval algorithm. Vertical lines show previous mass estimates for the systems and are labeled with the method that was used to obtain the estimate. Spec. class: spectral-class mass estimates from Vanbeveren et al. (2020; WR 42, WR 79, WR 127) and Shao & Li (2016; WR 153). Arrows show lower limits. Phot: photometric mass estimates from Lamontagne et al. (1996). Pol: polarimetric estimates from St-Louis et al. (1987; WR 42, WR 79) and St-Louis et al. (1988; WR 127, WR 153). Wind model: mass estimates using a spectroscopic colliding-wind model from Hill et al. (2000).

wind caused by wind–wind collisions are observed to occur primarily in the plane of the orbit (e.g., Callingham et al. 2020) and models agree (e.g., Hill et al. 2000; Lamberts et al. 2017). Corotating interaction regions (see, e.g., Mullan 1984) are another potential source of wind asymmetries, but their effect on polarization is small unless they have an optical depth of order unity (Carlos-Leblanc et al. 2019), which would disagree with this first mass mode.

The second mass mode (bottom in Figure 4) corresponds to models with a nearly spherical WR wind (with $z_{\text{ell}} > x_{\text{ell}}$, or equivalently a slight polar density enhancement), an electron

scattering optical depth of $\tau \sim 0.3$, and the majority of emission ($\sim 70\%$) originating from the WR star. The second WR mass mode at $\sim 27 M_{\odot}$ would further increase the similarities between WR 42 and WR 79, which are identical in WR type, similar in O star type, and with orbital periods only 1 day apart. The corresponding second mode O star mass of $\sim 60 M_{\odot}$ is compatible with the upper estimate for the O star from Vanbeveren et al. (2020). Another implication of the higher WR star mass is that the WR star would be more likely to have reached the WR stage without interacting with its companion, based on single-star evolutionary tracks (Sander et al. 2019).

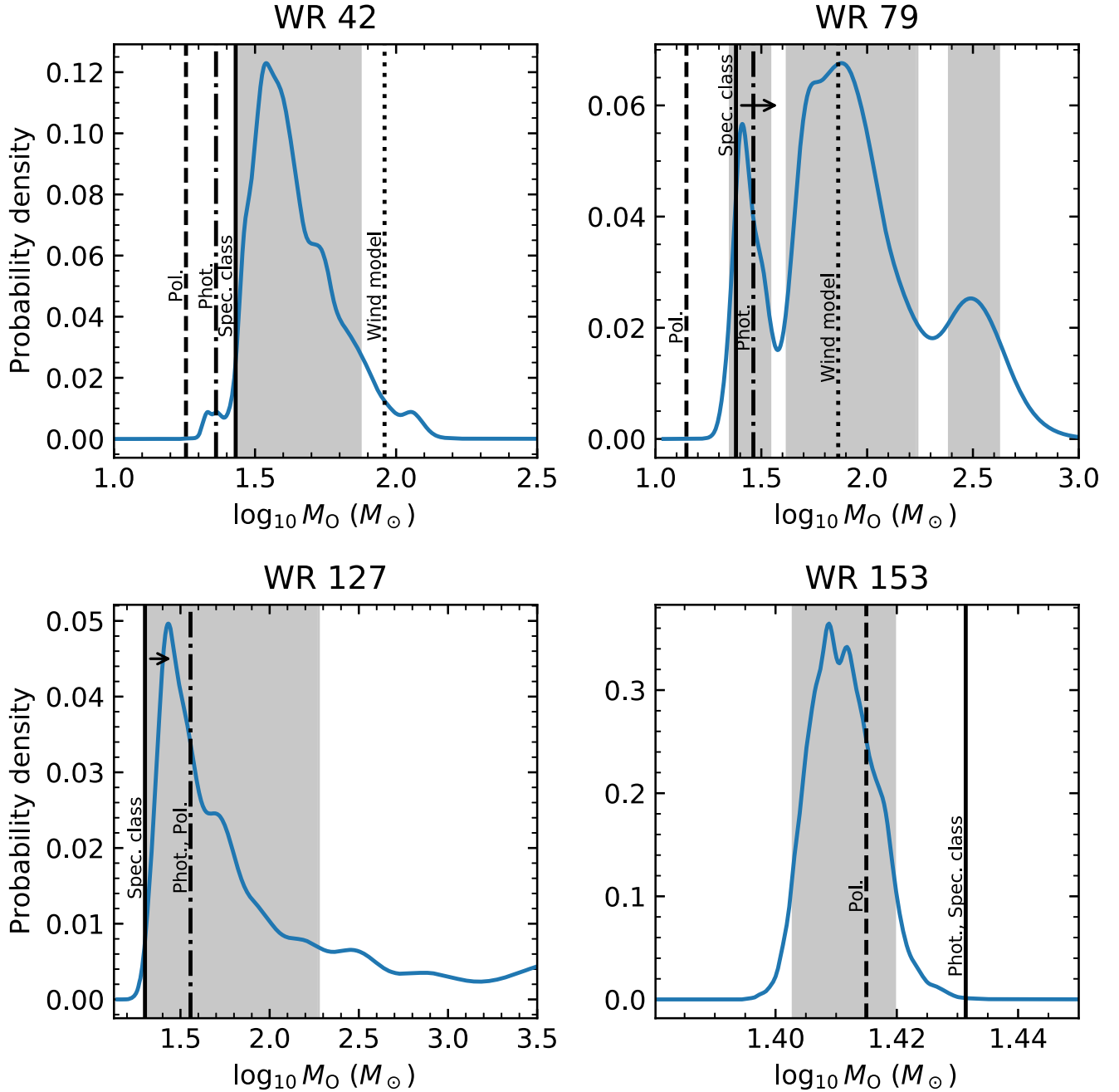


Figure 3. WR 42, WR 79, WR 127, and WR 153 O star mass posterior distributions, shown as solid curves. The gray shaded regions show the 90% credible interval of the distribution. Vertical lines show previous mass estimates for the systems and are labeled with the method that was used to obtain the estimate (see the caption to Figure 2).

4.3. WR 127

The posterior for the mass of the WR component of WR 127 (90% credible interval $11.1\text{--}102.5 M_\odot$) is in agreement with Vanbeveren et al. (2020), since they provide only a lower limit. We also agree with both photometric and polarimetric estimates. However, our method provides an upper limit on the mass of $49.5 M_\odot$. It also increases the mass of the lower limit to $11.1 M_\odot$. The most probable mass is $\sim 14 M_\odot$. The same is true for the O star, with its most probable mass near $\sim 25 M_\odot$, at the upper end of the $17\text{--}24 M_\odot$ estimate given by Vanbeveren et al. (2020).

Our mass estimate also agrees with both the previous polarimetric (St-Louis et al. 1988) and photometric

(Lamontagne et al. 1996) methods, which reported WR masses of $11\text{--}21 M_\odot$ and $12\text{--}16 M_\odot$, respectively. Although these are lower uncertainty mass estimates than our presented values, the previous polarimetric estimate was published before the complete bias analysis of Wolinski & Dolan (1994) and thus the uncertainty is likely to be larger in reality. Furthermore, the photometric model assumes a spherical WR wind and formal uncertainties. This is unlikely for WR 127 given the evidence for a possible wind collision from X-rays (see Section 4.5 and Nazé et al. 2021).

4.4. WR 153

We find that WR 153 has a significantly lower mass (90% credible interval $13.7\text{--}14.2 M_\odot$) than measured by photometric

Table 5
Our New WR + O System Mass Estimates

WR Number	$M_{\text{WR}} (M_{\odot})$	$M_{\text{O}} (M_{\odot})$	$M \sin^3 i$ Source
42	16.2–44.0	27.4–74.6	1
79 (mode 1)	8.3–12.8	22.5–34.8	2
79 (mode 2)	15.3–62.8	41.8–170.9	2
79 (mode 3)	89.5–152.2	243.7–414.2	2
127	11.1–102.5	19.8–183.1	3
153	13.7–14.2	25.3–26.3	4

Note. Derived from our radiative-transfer models, including all modes for WR 79.

References. 1: Davis et al. (1981); 2: Seggewiss (1974); 3: de La Chevrotière et al. (2011); 4: Lamontagne et al. (1996).

and spectral-class models (where we define “significant” as outside the 90% credible interval), with a low uncertainty compared to the other stars in the sample. The photometric-model inclination angle estimate is similarly certain, with a mass range of $14.5\text{--}14.7 M_{\odot}$ (Lamontagne et al. 1996), though the true uncertainty is likely to be larger. Similarly to Vanbeveren et al. (2020), Shao & Li (2016) simply present an adopted mass for the WR star of $15 M_{\odot}$ with no associated uncertainty. Our posterior distribution agrees with the analytic polarimetric estimate using the same data (St-Louis et al. 1988), which is unsurprising, given that the high inclination angle of the system reduces the bias of the analytic model. The O star mass is low for its spectral type, which is expected to be greater than $30 M_{\odot}$ (Vanbeveren et al. 1998). However, there is not currently a measure of its luminosity class, which makes this expectation very uncertain. Furthermore, measurement uncertainties for $M \sin^3 i$ would broaden the range of masses, potentially including the mass estimate from Shao & Li (2016), Demers et al. (2002) produced an additional measurement of $M \sin^3 i$, but it was only a lower limit.

4.5. Wind Asymmetry

Our findings indicate that all four systems we investigated have some form of asymmetry in their winds. It is unsurprising to find asymmetry in the winds of close massive stellar binaries because of the possibility of wind–wind interaction between the components. Hill et al. (2000) modeled the colliding winds in both WR 42 and WR 79, and found that optical line variations could be explained by the presence of conic wind collision regions, making it plausible that our detection of wind asymmetries in these systems is caused by wind–wind collision. Recent work by Nazé et al. (2021) has shown that WR 127 exhibits X-ray emission that indicates the presence of wind collision regions, though WR 42 was a nondetection and WR 153 showed only faint X-rays compared to the expectation for a wind collision. Thus, the possible wind asymmetry we find for WR 153 may not be caused by wind–wind collision. These results also cast some doubt on the presence of wind–wind collision regions in WR 42. Instead, the wind asymmetries could be caused by rotation of the WR star expanding the wind equatorially (Vink & Harries 2017) or optically thick corotating interaction regions producing areas of enhanced density in the wind (Carlos-Leblanc et al. 2019).

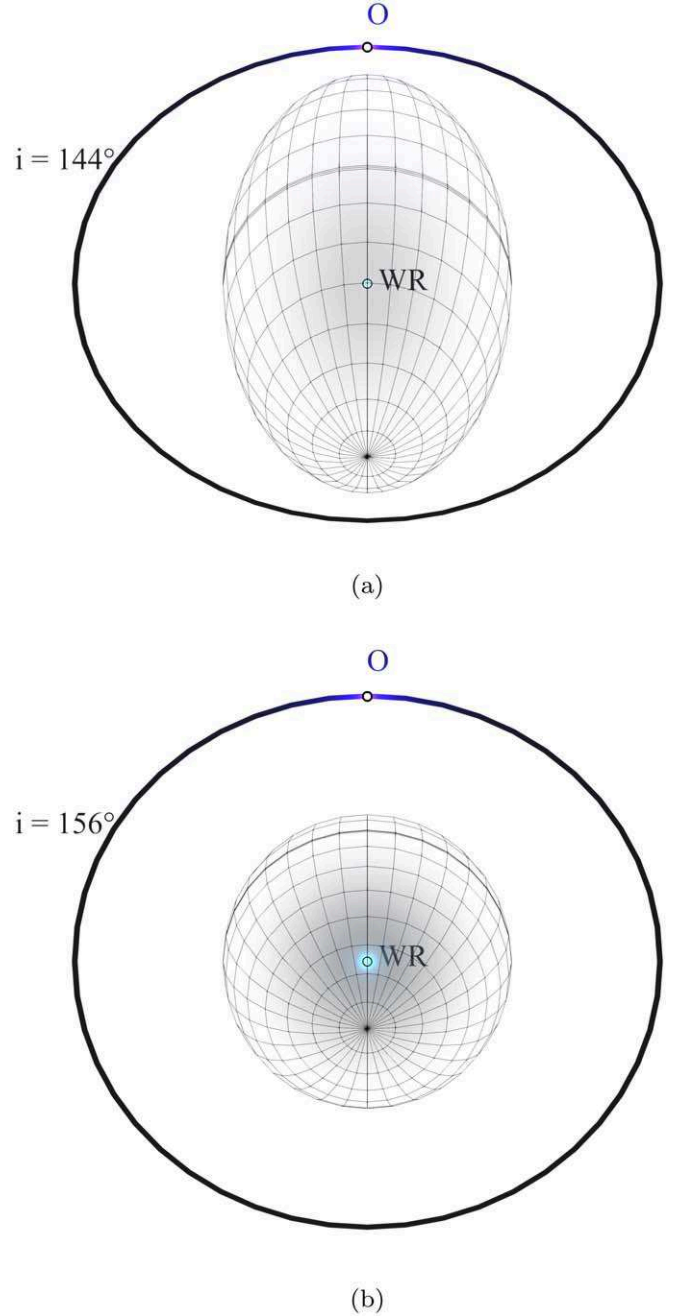


Figure 4. (a) WR 79 representative model of the first posterior mode of mass $8.3\text{--}12.8 M_{\odot}$ at $i = 144^\circ$. (b) WR 79 representative model of the second posterior mode of mass $15.3\text{--}62.8 M_{\odot}$ at $i = 156^\circ$. Both figures are at orbital phase 0.5 and to scale. The triple lines on the scattering regions show the orbital planes, and the semitransparent fills represent their optical depths. The thick black lines are the projected O star orbits. The stars have relative emission equivalent to each model.

4.6. Other Model Parameters

Our posterior distribution reveals relationships between parameters. Across all models, we note relationships between $\tau\text{--}z_{\text{ell}}$, $I_{\text{frac}}\text{--}\tau$, $I_{\text{frac}}\text{--}z_{\text{ell}}$, and $q_{\text{ISP}}\text{--}u_{\text{ISP}}$. In the case of $\tau\text{--}z_{\text{ell}}$ and $I_{\text{frac}}\text{--}\tau$, the two-dimensional posterior distributions show strong exponential-like relationships. The $I_{\text{frac}}\text{--}z_{\text{ell}}$ correlation is similar, but generally less strong. This is unsurprising, given that the probability of a photon scattering in a path length l is $1 - \exp(-\tau)$, and $\tau \propto l$. z_{ell} changes the path length and thus

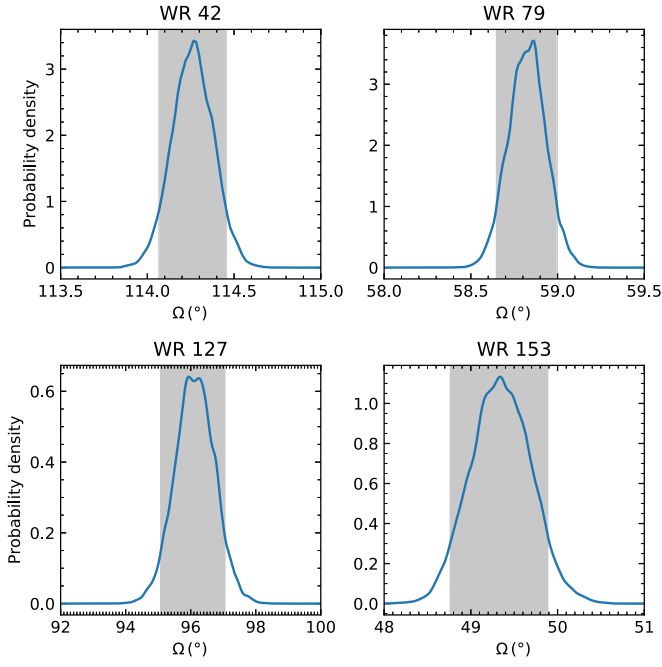


Figure 5. Posterior distributions of the longitude of the ascending node, Ω , for WR 42, WR 79, WR 127, and WR 153, shown as solid lines. The shaded area in each panel shows the 90% credible interval.

the required optical depth to reach a given polarization. Therefore, z_{ell} and τ are directly related. Then, a more asymmetric scattering region produces more polarization for a given I_{frac} , which produces the $I_{\text{frac}}-z_{\text{ell}}$ relationship; the relationship between $\tau-z_{\text{ell}}$ follows from $I_{\text{frac}}-z_{\text{ell}}$ and $\tau-z_{\text{ell}}$. Finally, the relationship $q_{\text{ISP}}-u_{\text{ISP}}$ occurs because the rotation of the model by θ_R transforms the polarization directly between q and u .

As part of the modeling process, we also recover ISP and Ω posterior distributions. Our ISP posteriors are strongly influenced by the priors, though all the systems show some deviation from the existing estimates in Fullard et al. (2020). We present the maximum-likelihood estimates in Table 6. The uncertainties for the distributions are displayed in Appendix D as part of the corner plots, split into upper and lower 34% quantiles.

Figure 5 shows the posterior distribution of Ω for each system. The gray shaded regions show the 90% credible interval of the distributions. Our posterior distributions for Ω are in complete disagreement with the previous estimates presented in Table 1 for all but WR 127 (i.e., the previous estimates lie outside the plotted range of Ω in Figure 5). However, this is not surprising given that the sources of these estimates did not provide uncertainty estimates for Ω that take into account the bias on the analytic model that was used to derive Ω (Wolinski & Dolan 1994).

The final parameters we recover are I_{frac} and τ , the fraction of photons arising from the WR star and the optical depth of the scattering region, respectively. Our method confirms that these parameters are related to each other and to z_{ell} . This means that the posterior distributions are difficult to interpret as standalone results. For example, in the case of WR 79, the I_{frac} distribution peaks at $z_{\text{ell}} \sim 1.6$ and $\tau \sim 0.2$. There are also peaks in the posteriors that correspond to values of $\tau \sim 0$, $z_{\text{ell}} \sim 3.0$, and $I_{\text{frac}} \sim 0$. Although these two possible models are peaks in the posterior distributions, there is a range of possible models

Table 6

WR + O System ISP Maximum-likelihood Estimates from Our Analysis

WR Number	q_{ISP} (%)	u_{ISP} (%)
42	−0.52	−0.49
79	−0.31	−0.24
127	0.69	0.68
153	−0.04	4.15

Note. Uncertainties are displayed in Appendix D as part of the corner plots.

between these peaks because of the relationship between the parameters. If we disregard these relationships and focus on the median τ for each object, we typically find low optical depths, indicating that the broadband polarization arises in the outer regions of the WR star wind.

5. Conclusions

We have found probabilistic estimates for the mass of four WR + O binary systems, with a new method based on an emulated radiative-transfer model. Our results are different to previous estimates for WR 153, and we provide robust upper limits in all cases, as well as information about degeneracies in the model. Our disagreement with existing polarimetric estimates for these systems shows that care must be taken when drawing conclusions using masses derived from analytic polarimetric models.

We have shown that our method can break the orbital inclination degeneracy in $M \sin^3 i$, but this still requires spectroscopic determination. Thus, we recommend improvements to spectroscopic measurements of $M \sin^3 i$ to obtain more certain mass estimates for the four systems presented in this paper, and others like them. Phased polarimetric measurements are also a requirement for this method, and thus more accurate polarimetric observations are necessary to apply this method to more targets. We are obtaining accurate phased polarimetric observations of a larger sample of Milky Way WR + O systems (Fullard et al. 2018; Johnson et al. 2019), and will apply this method to them in the future. In principle, the model can also be applied to the 30% of Be stars that have companions (Oudmaijer & Parr 2010), because their disks produce polarization signals. Further improvements to our model could be made with the incorporation of elliptical orbits, to investigate systems such as WR 133 that already have polarimetric observations available.

With the advent of telescopes like the Polstar UV spectro-polarimetric satellite (Scowen et al. 2021), it will be critical to have polarization models that can be explored with statistical approaches to understand the observations. UV observations of WR + O binary systems can provide new constraints for the priors presented here.

We thank Selma de Mink for her insightful comments about the content of this manuscript.

We gratefully thank Johannes Buchner for UltraNest assistance.

TACC is operated by the Extreme Science and Discovery Environment (XSEDE), which we access under allocation AST-120067 to J.H.

J.H. is grateful for funding from the National Science Foundation under award AST-1816944, and acknowledges that

the University of Denver occupies land within the traditional territories of the Arapaho, Cheyenne, and Ute peoples.

R.I. acknowledges funding support for this research from a grant with the National Science Foundation, AST-2009412.

The data for the models in this article are available in a Zenodo repository at <https://doi.org/10.5281/zenodo.6226123>.

Software: ARVIZ, JUPYTER (Thomas *et al.*, 2016), LMFIT, TENSORFLOW, SCIKIT-LEARN, BLENDER (Blender Online Community 2021), CORNER (Foreman-Mackey 2016), ULTRANEST.

Contributor roles. Conceptualization: Andrew Fullard, Wolfgang Kerzendorf. Data curation: Andrew Fullard. Formal Analysis: Andrew Fullard, Patrick van der Smagt. Investigation: Andrew Fullard, Patrick van der Smagt. Methodology: Andrew Fullard. Project administration: Wolfgang Kerzendorf. Resources: Institute for Cyber-Enabled Research at Michigan State University, Texas Advanced Computing Center. Software: Andrew Fullard, Jennifer Hoffman, Manisha Shrestha, Patrick van der Smagt, John O’Brien. Supervision: Wolfgang Kerzendorf. Visualization: Andrew Fullard. Writing original draft: Andrew Fullard. Writing review & editing: Andrew Fullard, John O’Brien, Patrick van der Smagt, Jennifer Hoffman, Richard Ignace, Wolfgang Kerzendorf, Manisha Shrestha.

Appendix A

SLIP Monte Carlo Radiative-transfer Code

SLIP is a Fortran + MPI code based on the MCRT method outlined in Whitney (2011). The simulation grid is a uniformly spaced spherical polar coordinate system in r , θ , and ϕ . Photon packets are emitted from user-specified locations in the grid and propagate through the user-defined scattering region. At each grid cell the optical depth is integrated until the photon scatters or exits the grid cell. The photons are collected as they exit the simulation limits and are binned into different observational directions in θ and ϕ , with uncertainties calculated via Poisson statistics. In this way, a single model can be viewed from multiple angles, reducing computation time. For additional details of the SLIP code, see Huk (2017) and Shrestha et al. (2018).

The original SLIP code included a single photon source at the center of the simulation grid. We used an enhanced version of the code that includes an additional photon source located at an arbitrary distance along the x -axis ($\theta = 90^\circ$, $\phi = 0^\circ$) from the central source (each source can be a finite size or a point). The number of photons emitted from each source is controlled as a fraction of the total photon count. We simulate circular orbits by simply moving the “observer” around the grid in the ϕ direction. In Appendix B, we provide validation of these new capabilities against the existing analytic model of Brown et al. (1978), while being aware of the inherent limitations described in Appendix B.

Appendix B

Validating the SLIP Binary Model

To test the validity of the enhanced SLIP code, we computed models with three circumstellar geometries that can be fit with the analytic model of Brown et al. (1978); that is, a spherical distribution of electron scattering region, a circular disk-like distribution in the orbital plane with an opening angle of 1.8° (a single θ angle bin width in height), and a prolate distribution. The prolate distribution is described by an ellipsoid with the

major axis parallel to the z -axis, described by

$$\frac{x^2}{0.2^2} + \frac{y^2}{0.2^2} + \frac{z^2}{0.6^2} = 1, \quad (\text{B1})$$

where x , y , and z refer to the standard Cartesian coordinate system that SLIP transforms to its spherical coordinate system. The denominators are the values of x_{ell} , y_{ell} , and z_{ell} , the x , y , and z axes of the ellipsoid.

The finite nature of the SLIP grid means that the disk-like distribution is a single grid cell thick in the θ direction. All three geometries were centered on the source located at the origin. Both stellar sources emitted an equal proportion of the photons, and both were set as point sources so eclipse effects are not important and photons are not absorbed within either star. We set the optical depth of all these CSM distributions at $\theta = 90^\circ$ to $\tau = 0.1$ for $0 < \phi < 2\pi$ to match the optically thin assumption of the analytic model (Carlos-Leblanc et al. 2019 found that this was the limit between optically thin and optically thick for their similar MCRT models). Note that the CSM density is set to be constant and proportional to the optical depth at $\theta = 90^\circ$, so the prolate model shows an increase in optical depth toward $\theta = 0^\circ$ and 180° because of its increased radius toward those angles. Model calculations were made on the Stampede high-performance computing cluster at the Texas Advanced Computing Center,⁹ with 8×10^9 photons per simulation. The output was binned to 23 inclination bins and 40 orbital phase bins.

We fit the analytic model to the numerical results using LMFIT (Newville et al. 2014) to minimize χ^2 , and algebraically calculated the orbital parameters i and Ω from the Fourier coefficients following Drissen et al. (1986). The Fourier coefficients q_i and u_i are found by fitting the following equations:

$$q = q_0 + q_1 \cos \phi + q_2 \sin \phi + q_3 \cos 2\phi + q_4 \sin 2\phi, \quad (\text{B2})$$

$$u = u_0 + u_1 \cos \phi + u_2 \sin \phi + u_3 \cos 2\phi + u_4 \sin 2\phi, \quad (\text{B3})$$

where ϕ is the azimuthal coordinate and $\phi/2\pi$ is the equivalent orbital phase. The fits had median χ_ν^2 values of 3.30, 2.85, and 0.90 for the disk-like, spherical, and prolate density distributions, respectively (using $\chi_\nu^2 = \chi^2/(N - 10)$, where N is the number of data points and 10 is the number of fit variables). The χ_ν^2 value of 0.9 suggests some overfitting or overestimate of uncertainties in the case of the disk-like density distribution.

The comparisons between the numerical and analytic models are shown in Figures 6, 7, and 8 for the disk-like, spherical, and prolate density distributions, respectively. In all Figures 6–8, the color of the points corresponds to the inclination angle of the SLIP model. The uncertainty of the polarization increases as the inclination angle approaches zero (i.e., perpendicular to the orbital plane) because the amplitude of polarization is directly proportional to the orbital inclination angle (Brown et al. 1978), and a lower amplitude increases the Monte Carlo noise in the simulation because fewer photons have a polarization signal. In Figure 6, the significantly higher uncertainty of the SLIP model arises due to the extremely thin disk-like scattering region. This reduces the polarization from the model by an order of magnitude relative to the spherical and prolate density distributions, and thus the uncertainty is proportionally larger.

⁹ <https://www.tacc.utexas.edu/systems/stampede>

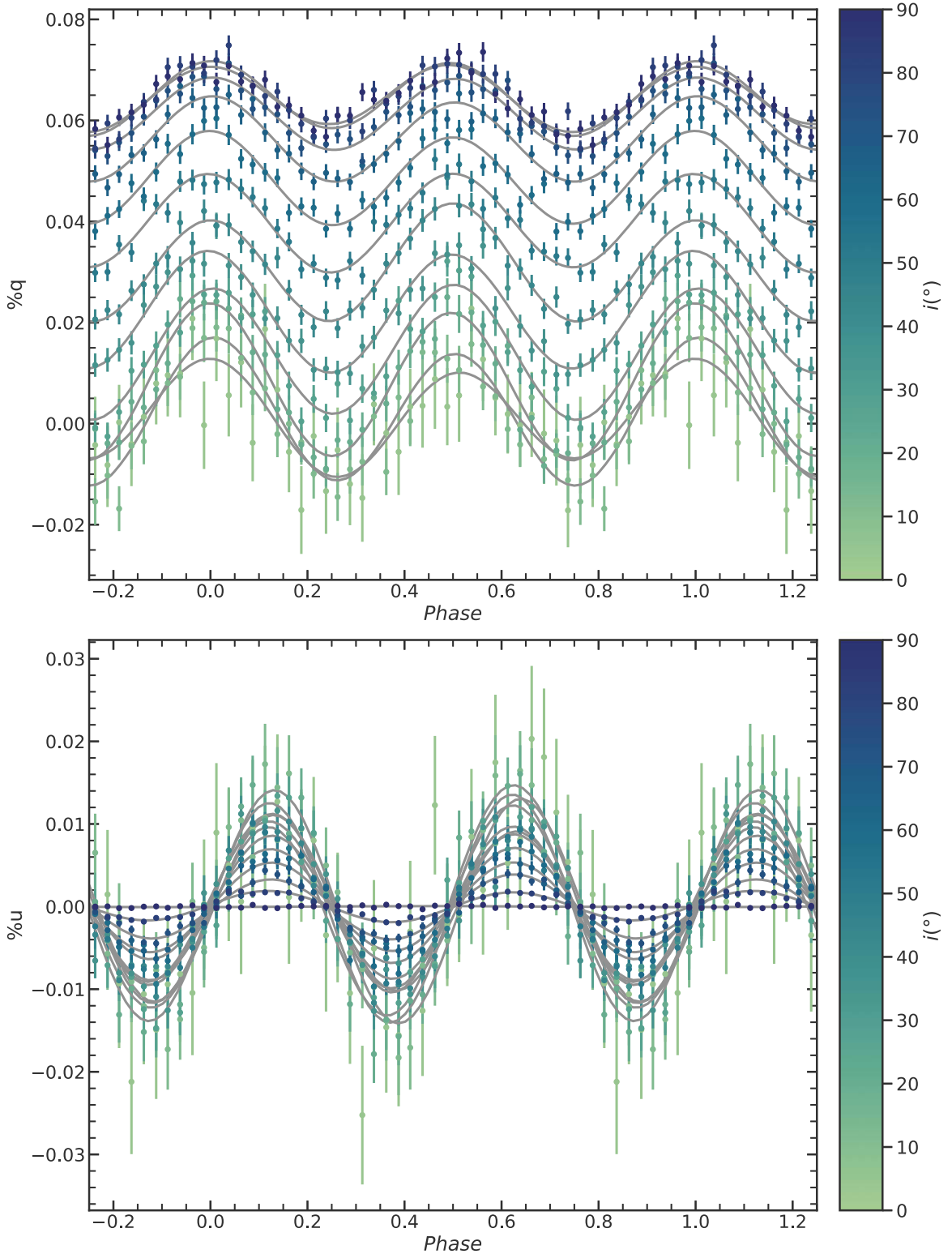


Figure 6. Comparison of SLIP results for a thin disk to the analytic model of Brown et al. (1978). The circles represent SLIP results for a range of inclination angles from 0° to 90° as displayed in the color bar. The data have been fit by the model of Brown et al. (1978; gray lines).

The fit results for a selection of inclination angles are presented in Figure 9 for the thin disk, sphere, and prolate CSM models. As described by Wolinski & Dolan (1994), the analytic model requires extremely high precision measurements relative to the polarization variation amplitude to recover accurate derived parameters as the inclination angle decreases from 90°

to 0° . Our results agree with those of Wolinski & Dolan (1994); despite the high accuracy of the numerical models, the truly low inclination angle numerical models are fitted with a biased higher inclination angle by the analytic model. The disk model is most strongly affected because it has the lowest amplitude of polarization variation of the models.

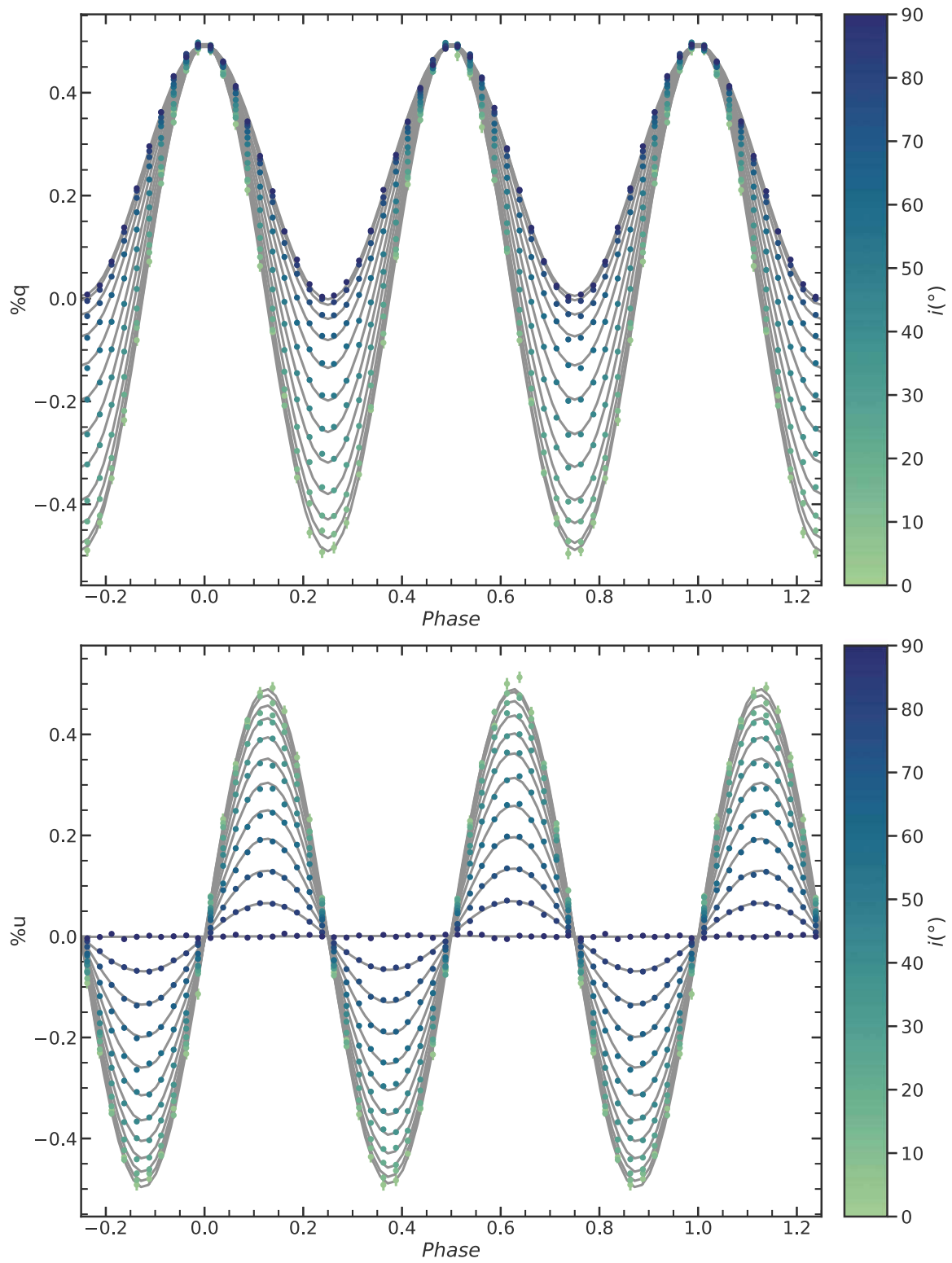


Figure 7. Comparison of SLIP results for a sphere to the analytic model of Brown et al. (1978). The circles represent SLIP results for a range of inclination angles from 0° to 90° as displayed in the color bar. The data have been fit by the model of Brown et al. (1978; gray line).

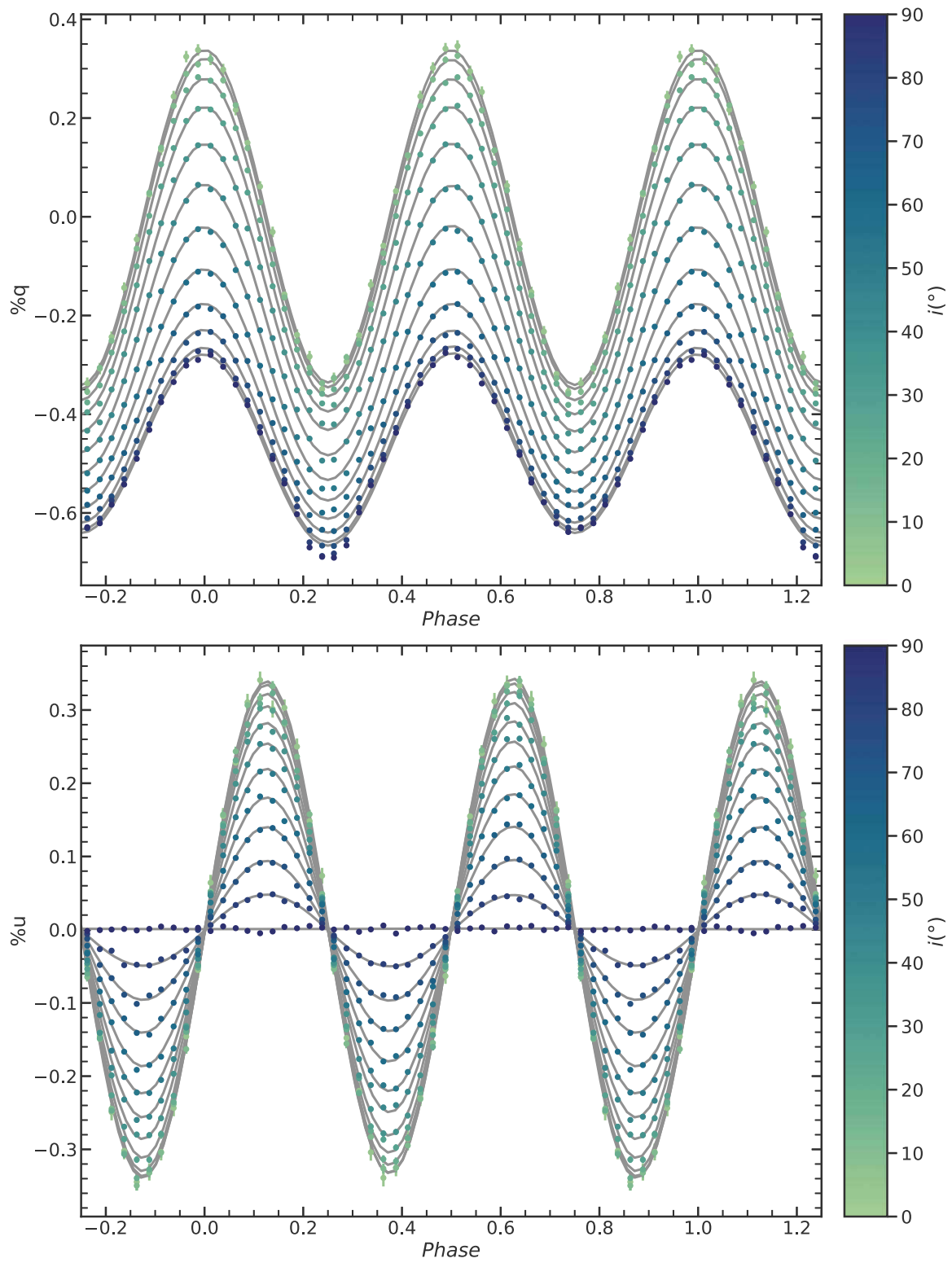


Figure 8. Comparison of SLIP results for a prolate ellipsoid (extended in the direction perpendicular to the orbital plane) to the analytic model of Brown et al. (1978). The circles represent SLIP results for a range of inclination angles from 0° to 90° as displayed in the color bar. The data have been fit by the model of Brown et al. (1978; gray line).

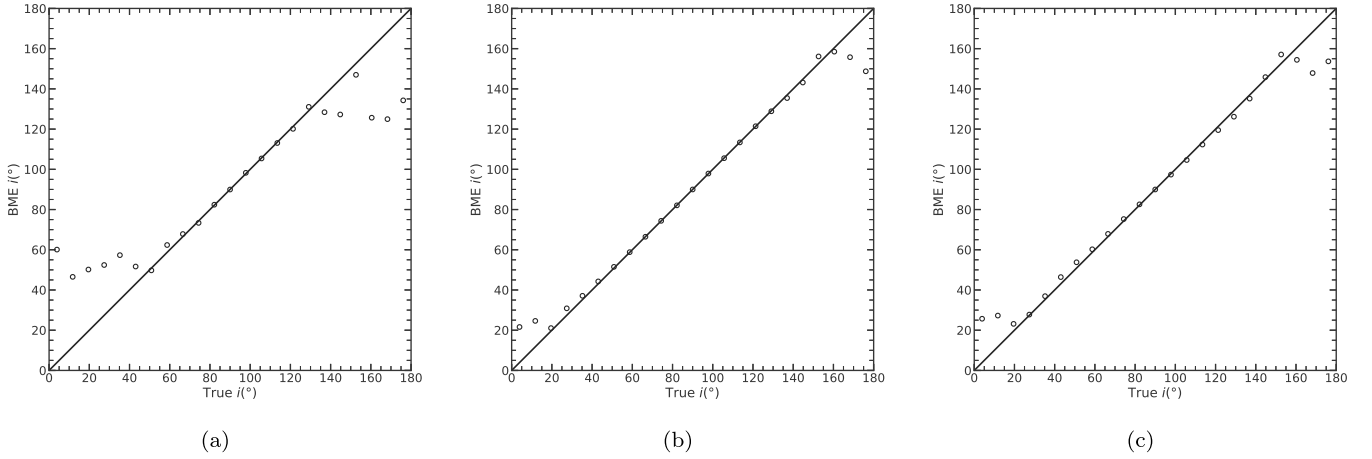


Figure 9. Comparison of SLIP inclination for a thin disk (a), a sphere (b), and a prolate ellipsoid (c) distribution of scattering material to the analytic model of Brown et al. (1978). The points show the true SLIP model inclination vs. the calculated Brown et al. model inclination. The black lines show unity.

Appendix C Emulator Details

Here we present the details of our emulator, including its training set and architecture.

C.1. Training Set

The data set used to train the neural network was produced using 4517 runs of the modified SLIP MCRT code using uniform random sampling of the parameter space described in Table 3.

To create the training set, we used the MSU high-performance computing cluster provided by the Institute of Cyber-Enabled Research. Each run took up to 4 hr across eight CPUs with 2×10^8 photons per CPU. This photon count was chosen to produce uncertainties in q and u below $\sim 0.05\%$ with a median uncertainty of $\sim 0.03\%$. The resulting training set including all inclination angle bins has size $406,530 \times 40$ because of the uniform sampling of i in each model and because each model has 40 total ϕ (phase) bins.

C.2. Emulator Architecture

The emulator is based on the proven DALEK emulator methods that have previously been applied to the TARDIS radiative-transfer code (Kerzendorf & Sim 2014; Kerzendorf et al. 2021). It was built using Tensorflow (Martín et al. 2015) and scikit-learn (Pedregosa et al. 2011) tools to provide the neural network and data preprocessing, respectively.

The neural-network architecture was investigated using a hyperparameter search of 2000 possible architectures on a cluster with 54 GPUs running Polyaxon, on-premise at the

Volkswagen Group Machine Learning Research Lab. The resulting best architecture was chosen based on the total loss (here, using the mean-squared error) and the number of steps required to train. The neural network consists of five hidden layers of width 30, with softplus activation (where $\text{softplus}(x) = \log[\exp x + 1]$) and Glorot normal initialization (Glorot & Bengio 2010). No dropouts or layer normalization was deemed necessary by the hyperparameter search.

The data and model input parameters were scaled using the StandardScaler function of scikit-learn. The data were randomly split into 90%/10% train/validation sets. Optimal neural-network architectures were those that had a minimal error on the validation set. The training was performed on an nVidia RTX 2080Ti graphics card with 12 GB of VRAM, and took approximately 19 hr to reach 20,000 epochs. Of these 20,000 steps, the training result with the lowest total validation set loss was chosen to represent the best result.

Two emulators were produced from the training set, one for Stokes q and one for Stokes u . Validation of the emulator against random samples of the training set showed that the two emulators returned consistent model parameters independently of each other.

Appendix D Corner Plots

The corner plots, in Figures 10–13, show the six-dimensional posterior distributions of each parameter with two-dimensional posterior distributions for each pair of parameters. These distributions are available at Zenodo at <https://doi.org/10.5281/zenodo.6226123>.

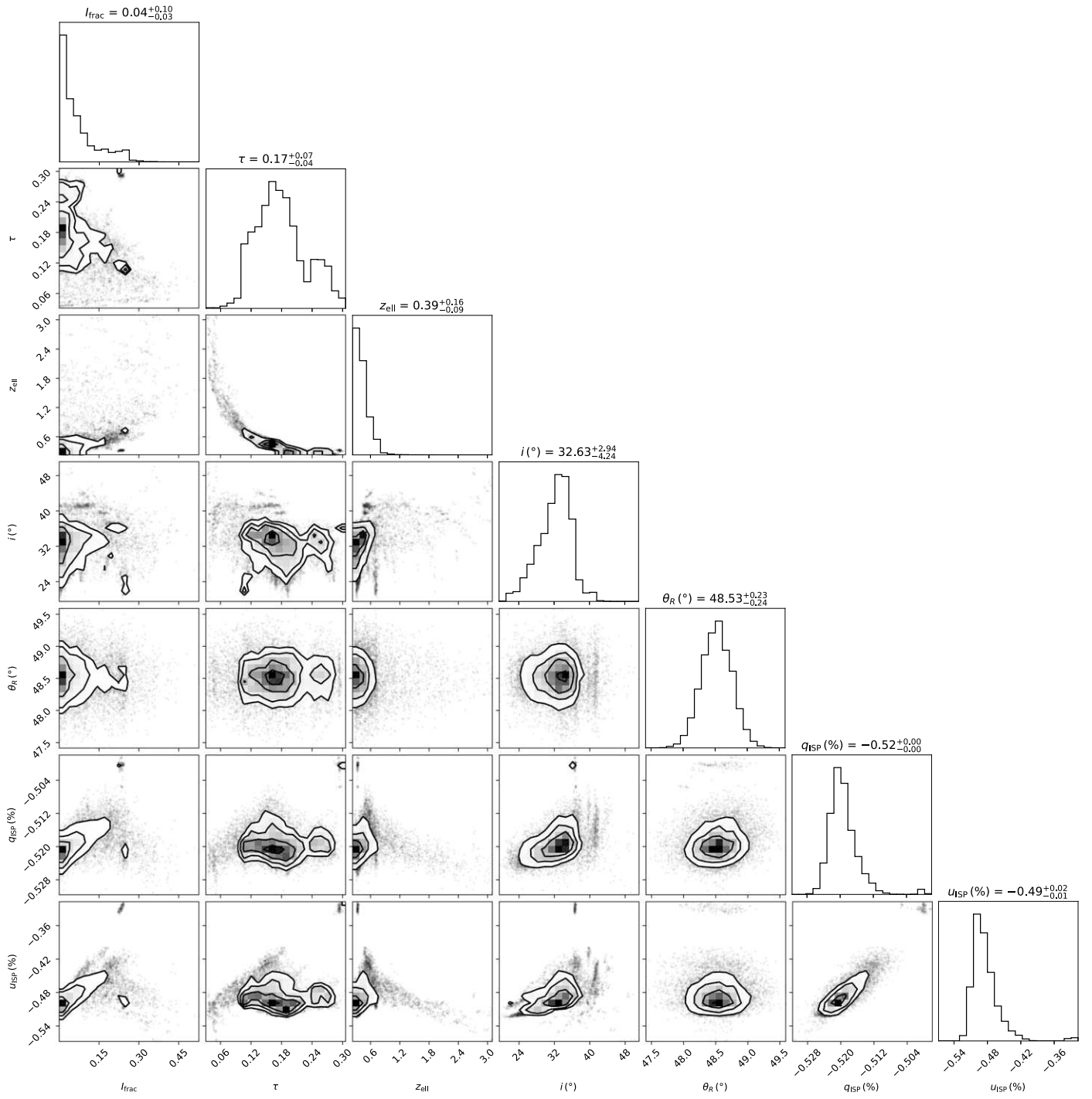


Figure 10. Corner plot of WR 42 showing posterior distributions for each parameter and their relationships.

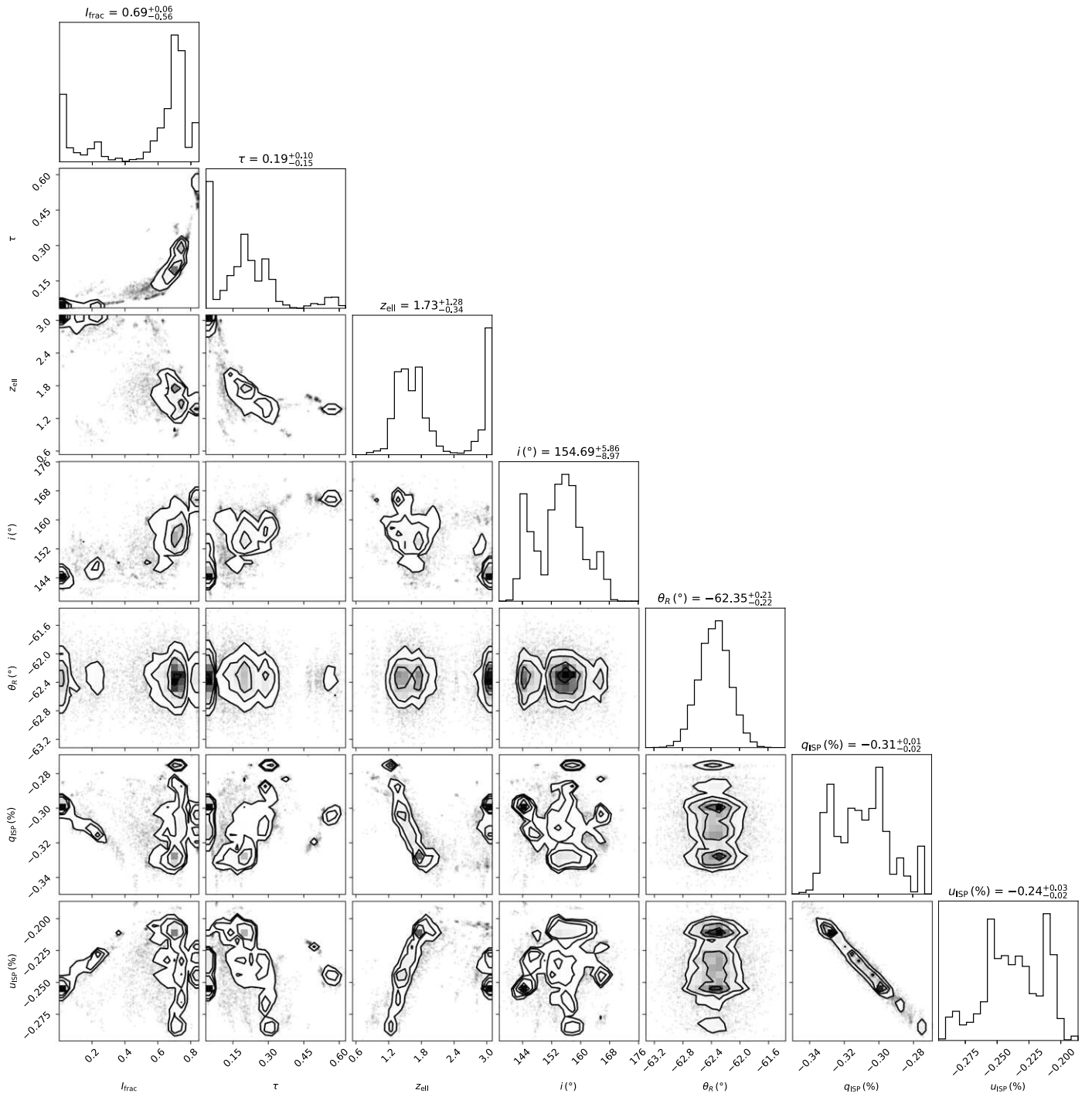


Figure 11. Corner plot of WR 79 showing posterior distributions for each parameter and their relationships.

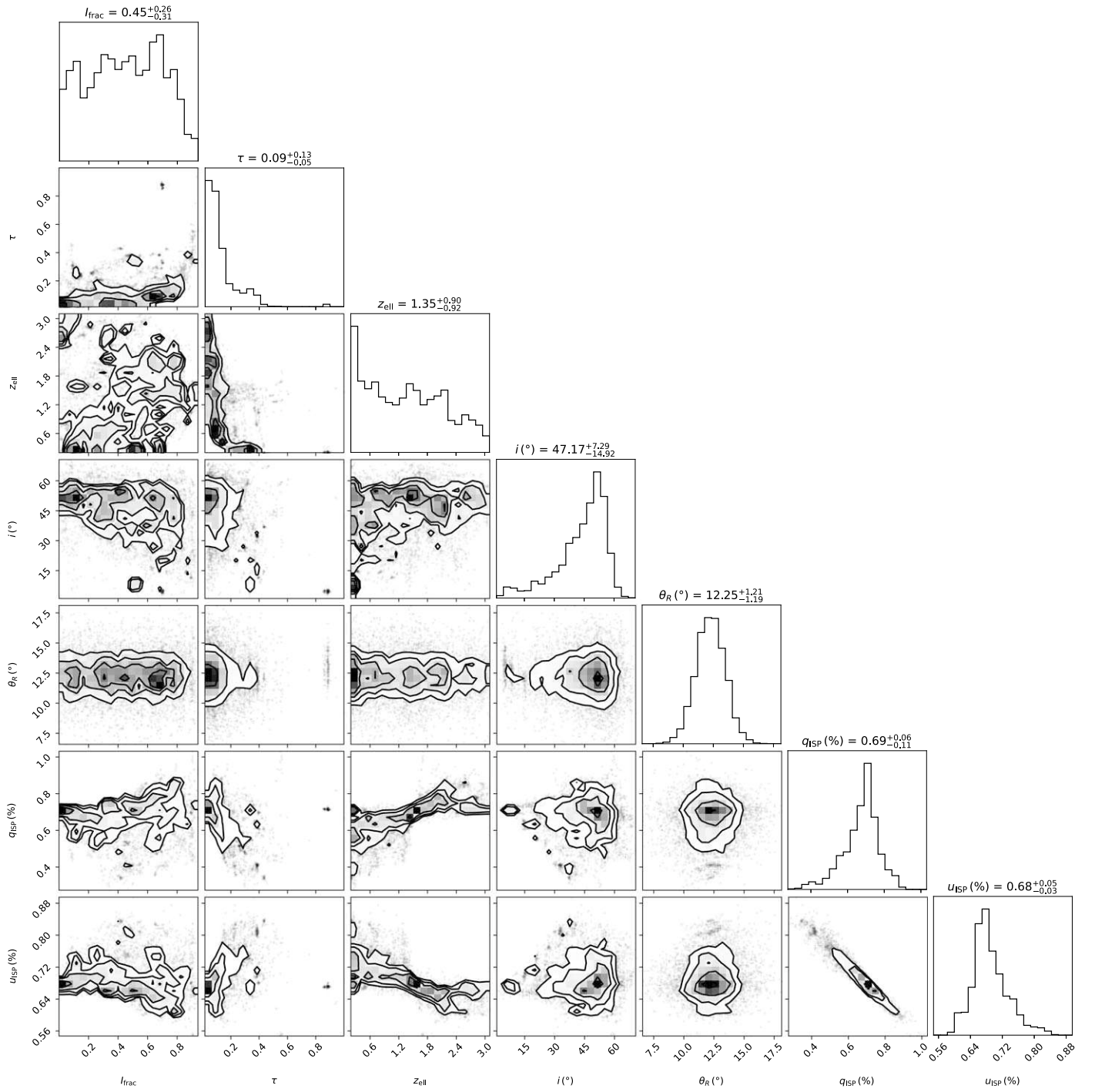


Figure 12. Corner plot of WR 127 showing posterior distributions for each parameter and their relationships.

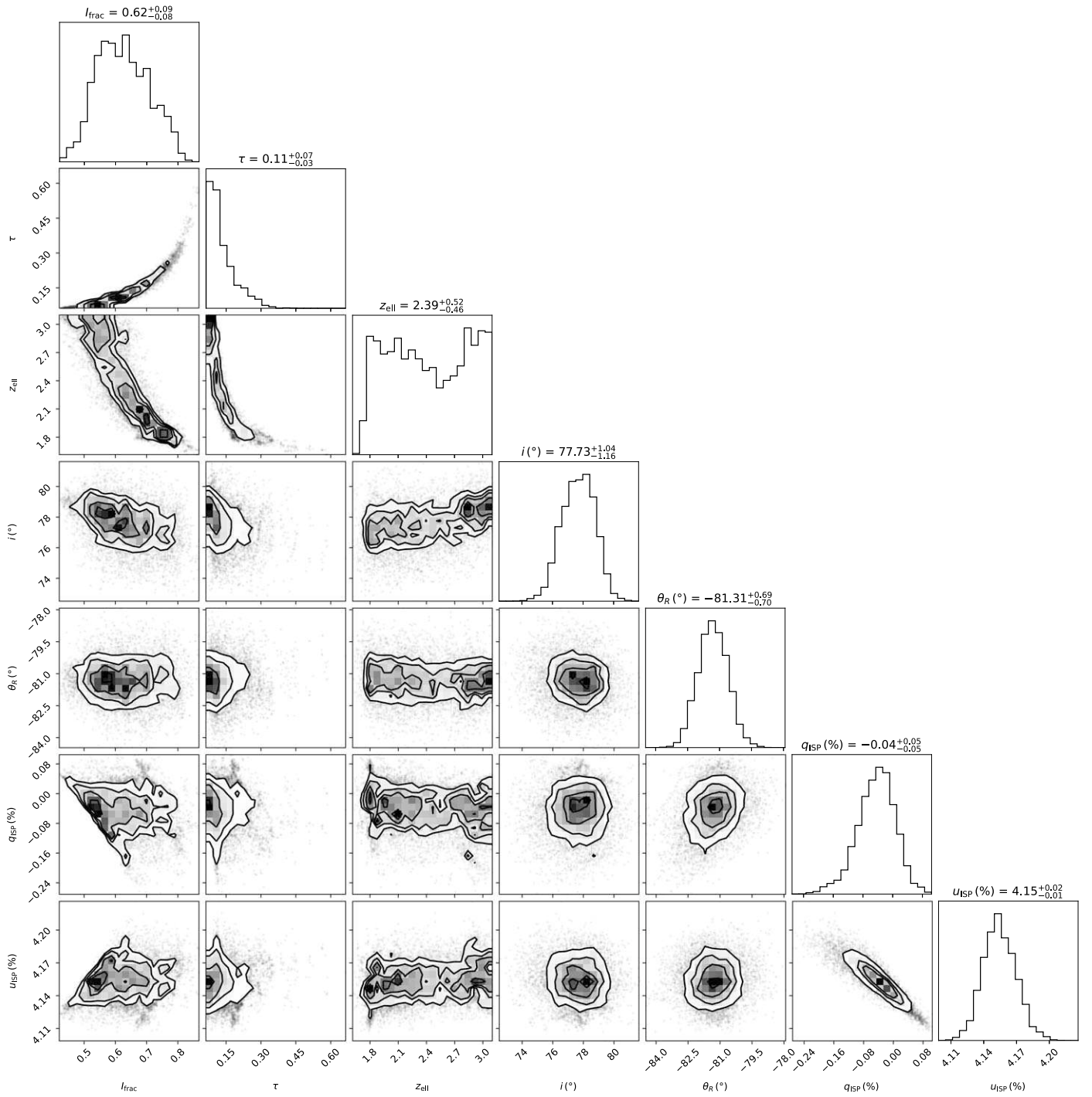









Figure 13. Corner plot of WR 153 showing posterior distributions for each parameter and their relationships.

ORCID iDs

Andrew G. Fullard  <https://orcid.org/0000-0001-7343-1678>
 John T. O'Brien  <https://orcid.org/0000-0003-3615-9593>
 Wolfgang E. Kerzendorf  <https://orcid.org/0000-0002-0479-7235>
 Manisha Shrestha  <https://orcid.org/0000-0002-4022-1874>
 Jennifer L. Hoffman  <https://orcid.org/0000-0003-1495-2275>
 Richard Ignace  <https://orcid.org/0000-0002-7204-5502>
 Patrick van der Smagt  <https://orcid.org/0000-0003-4418-4916>

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *ApJL*, **818**, L22
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *ApJL*, **848**, L12
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, *PhRvL*, **119**, 161101
 Aspin, C., Simmons, J. F. L., & Brown, J. C. 1981, *MNRAS*, **194**, 283
 Brown, J. C., Aspin, C., Simmons, J. F. L., & McLean, I. S. 1982, *MNRAS*, **198**, 787
 Blender Online Community 2021, Blender - a 3D modelling and rendering package, <http://www.blender.org>
 Brown, J. C., & McLean, I. S. 1977, *A&A*, **57**, 141
 Brown, J. C., McLean, I. S., & Emslie, A. G. 1978, *A&A*, **68**, 415
 Buchner, J. 2016, *Stat. Comput.*, **26**, 383
 Buchner, J. 2019, *PASP*, **131**, 108005
 Buchner, J. 2021, *JOSS*, **6**, 3001
 Callingham, J. R., Crowther, P. A., Williams, P. M., et al. 2020, *MNRAS*, **495**, 3323
 Carlos-Leblanc, D., St-Louis, N., Bjorkman, J. E., & Ignace, R. 2019, *MNRAS*, **489**, 2873
 Crowther, P. A. 2007, *ARA&A*, **45**, 177
 Crowther, P. A. 2015, in *Wolf-Rayet Stars: Proc. of an International Workshop held in Potsdam*, ed. W.-R. Hamann (Potsdam: Universitätsverlag Potsdam), **21**
 Cybenko, G. 1989, *Math. Control. Signals Syst.*, **3**, 303
 Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, *ApJ*, **812**, 128
 Davis, A. B., Moffat, A. F. J., & Niemela, V. S. 1981, *ApJ*, **244**, 528
 de La Chevrotière, A., Moffat, A. F. J., & Chené, A.-N. 2011, *MNRAS*, **411**, 635
 Demers, H., Moffat, A. F. J., Marchenko, S. V., Gayley, K. G., & Morel, T. 2002, *ApJ*, **577**, 409
 Drissen, L., Moffat, A. F. J., Bastien, P., Lamontagne, R., & Tapia, S. 1986, *ApJ*, **306**, 215
 Foreman-Mackey, D. 2016, *JOSS*, **1**, 24
 Fox, G. K. 1994, *ApJ*, **432**, 262
 Frecker, J. E., & Serkowski, K. 1976, *ApOpt*, **15**, 605
 Fullard, A. G. 2020, PhD Thesis, University of Denver, USA
 Fullard, A. G., Hoffman, J. L., DeKlotz, S., et al. 2018, *RNAAS*, **2**, 37
 Fullard, A. G., St-Louis, N., Moffat, A. F. J., et al. 2020, *AJ*, **159**, 214
 Glorot, X., & Bengio, Y. 2010, in *Proc. of the Thirteenth International Conf. on Artificial Intelligence and Statistics (JMLR Workshop and Conference Proceedings)*, Vol. 9, ed. Y. W. Tee & M. Titterton (Cambridge, MA: MIT Press), 249
 Hill, G. M., Moffat, A. F. J., & St-Louis, N. 2002, *MNRAS*, **335**, 1069
 Hill, G. M., Moffat, A. F. J., St-Louis, N., & Bartzakos, P. 2000, *MNRAS*, **318**, 402
 Hoffman, J. L. 2007, in *Revista Mexicana de Astronomía y Astrofísica Conf. Ser. 30*, ed. G. García-Segura & E. Ramírez-Ruiz (Córdoba: Instituto de Astronomía), **57**
 Hoffman, J. L., Whitney, B. A., & Nordsieck, K. H. 2003, *ApJ*, **598**, 572
 Hornik, K., Stinchcombe, M., & White, H. 1989, *NN*, **2**, 359
 Huk, L. N. 2017, PhD thesis, Univ. Denver
 Johnson, R. A., Fullard, A. G., Lomax, J. R., et al. 2019, *RNAAS*, **3**, 146
 Kerzendorf, W. E., & Sim, S. A. 2014, *MNRAS*, **440**, 387
 Kerzendorf, W. E., Vogl, C., Buchner, J., et al. 2021, *ApJL*, **910**, L23
 Kluyver, T., Ragan-Kelley, B., Perez, F., et al. 2016, in *Jupyter Notebooks - a Publishing Format for Reproducible Computational Workflows*, ed. L. Fernando & S. Birgit (Amsterdam: IOS Press), 87
 Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, *JOSS*, **4**, 1143
 Lamberts, A., Millour, F., Liermann, A., et al. 2017, *MNRAS*, **468**, 2655
 Lamontagne, R., Moffat, A. F. J., Drissen, L., Robert, C., & Matthews, J. M. 1996, *AJ*, **112**, 2227
 Langer, N. 2012, *ARA&A*, **50**, 107
 Laplace, E., Justham, S., Renzo, M., et al. 2021, *A&A*, **656**, A58
 Martins, F., Schaerer, D., & Hillier, D. J. 2005, *A&A*, **436**, 1049
 Martín, A., Ashish, A., Paul, B., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <https://www.tensorflow.org/>
 Mason, B. D., Hartkopf, W. L., Gies, D. R., Henry, T. J., & Helsel, J. W. 2009, *AJ*, **137**, 3358
 McClelland, L. A. S., & Eldridge, J. J. 2016, *MNRAS*, **459**, 1505
 Mullan, D. J. 1984, *ApJ*, **283**, 303
 Nazé, Y., Gosset, E., & Marechal, Q. 2021, *MNRAS*, **501**, 4214
 Newville, M., Stensitzki, T., Allen, D. B., & Ingargiola, A. 2014, LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python, v0.8.0, Zenodo, doi:10.5281/zenodo.11813
 O'Brien, J. T., Kerzendorf, W. E., Fullard, A., et al. 2021, *ApJL*, **916**, L14
 Oudmaier, R. D., & Parr, A. M. 2010, *MNRAS*, **405**, 2439
 Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, *JMLR*, **12**, 2825
 Puls, J., Vink, J. S., & Najarro, F. 2008, *A&ARv*, **16**, 209
 Richardson, N. D., Lee, L., Schaefer, G., et al. 2021, *ApJL*, **908**, L3
 Sana, H., de Mink, S. E., de Koter, A., et al. 2012, *Sci*, **337**, 444
 Sana, H., de Koter, A., de Mink, S. E., et al. 2013, *A&A*, **550**, A107
 Sana, H., Le Bouquin, J.-B., Lacour, S., et al. 2014, *ApJS*, **215**, 15
 Sander, A. A. C., Hamann, W.-R., Todt, H., et al. 2019, *A&A*, **621**, A92
 Sander, A. A. C., Vink, J. S., & Hamann, W.-R. 2020, *MNRAS*, **491**, 4406
 Scowen, P. A., Gayley, K., Neiner, C., et al. 2021, *Proc. SPIE*, **11819**, 1181908
 Seggewiss, W. 1974, *A&A*, **31**, 211
 Serkowski, K., Mathewson, D. S., & Ford, V. L. 1975, *ApJ*, **196**, 261
 Shao, Y., & Li, X.-D. 2016, *ApJ*, **833**, 108
 Shenar, T., Hainich, R., Todt, H., et al. 2016, *A&A*, **591**, A22
 Shenar, T., Sablowski, D. P., Hainich, R., et al. 2019, *A&A*, **627**, A151
 Shrestha, M., Neilson, H. R., Hoffman, J. L., & Ignace, R. 2018, *MNRAS*, **477**, 1365
 Shrestha, M., Neilson, H. R., Hoffman, J. L., Ignace, R., & Fullard, A. G. 2021, *MNRAS*, **500**, 4319
 Simmons, J. F. L., Aspin, C., & Brown, J. C. 1982, *MNRAS*, **198**, 45
 St-Louis, N., Drissen, L., Moffat, A. F. J., Bastien, P., & Tapia, S. 1987, *ApJ*, **322**, 870
 St-Louis, N., Moffat, A. F. J., Drissen, L., Bastien, P., & Robert, C. 1988, *ApJ*, **330**, 286
 Thomas, J. D., Richardson, N. D., Eldridge, J. J., et al. 2021, *MNRAS*, **504**, 5221
 Vanbeveren, D., De Donder, E., Van Bever, J., Van Rensbergen, W., & De Loore, C. 1998, *NewA*, **3**, 443
 Vanbeveren, D., Mennekens, N., Heuvel, E. P. J. v. d., & Bever, J. V. 2020, *A&A*, **636**, A99
 Villar-Sbaffi, A., St-Louis, N., Moffat, A. F. J., & Pirola, V. 2005, *ApJ*, **623**, 1092
 Vink, J. S., & Harries, T. J. 2017, *A&A*, **603**, A120
 Voshchinnikov, N. V., & Hirashita, H. 2014, *MNRAS*, **445**, 301
 Whitney, B. A. 2011, *BASI*, **39**, 101
 Wolinski, K. G., & Dolan, J. F. 1994, *MNRAS*, **267**, 5
 Woosley, S. E. 2019, *ApJ*, **878**, 49