Bipolar planetary nebulae from outflow collimation by common envelope evolution

Yangyuxin Zou,1,⋆ Adam Frank,1 Zhuo Chen,2 Thomas Reichardt,3,4 Orsola De Marco,3,4 Eric G. Blackman,1 Jason Nordhaus,5,6 Bruce Balick,7 Jonathan Carroll-Nellenback,1 Luke Chamandy1 and Baowei Liu1

1Department of Physics and Astronomy, University of Rochester, Rochester, NY 14627, USA
2Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada
3Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia
4Astronomy, Astrophysics and Astrophotonics Research Centre, Macquarie University, Sydney, NSW 2109, Australia
5National Technical Institute for the Deaf, Rochester Institute of Technology, Rochester, NY 14623, USA
6Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, NY 14623, USA
7Department of Astronomy, University of Washington, Seattle, WA 98195, USA

Published by Oxford University Press on behalf of the Royal Astronomical Society

ABSTRACT

The morphology of bipolar planetary nebulae (PNe) can be attributed to interactions between a fast wind from the central engine and the dense toroidal-shaped ejecta left over from common envelope (CE) evolution. Here we use the 3D hydrodynamic adaptive mesh refinement (AMR) code ASTROBEAR to study the possibility that bipolar PN outflows can emerge collimated even from an uncollimated spherical wind in the aftermath of a CE event. The output of a single CE simulation via the smoothed particle hydrodynamics (SPH) code PHANTOM serves as the initial conditions. Four cases of winds, all with high enough momenta to account for observed high momenta pre-PN outflows, are injected spherically from the region of the CE binary remnant into the ejecta. We compare cases with two different momenta and cases with no radiative cooling versus application of optically thin emission via a cooling curve to the outflow. Our simulations show that in all cases highly collimated bipolar outflows result from deflection of the spherical wind via the interaction with the CE ejecta. Significant asymmetries between the top and bottom lobes are seen in all cases. The asymmetry is strongest for the lower momentum case with radiative cooling. While real post-CE winds may be aspherical, our models show that collimation via ‘inertial confinement’ will be strong enough to create jet-like outflows even beginning with maximally uncollimated drivers. Our simulations reveal detailed shock structures in the shock-focused inertial confinement (SFIC) model and develop a lens-shaped inner shock that is a new feature of SFIC-driven bipolar lobes.

Key words: hydrodynamics – binaries: close – planetary nebulae: general – planetary nebulae: individual: OH 231.8+04.2.

1 INTRODUCTION

Planetary nebulae (PNe) are remnants of the late stages of stellar evolution of low- and intermediate-mass stars. They are created by the interaction between a fast wind ejected during the post-asymptotic giant branch (AGB) evolutionary state of a star and the slow moving, previously ejected AGB material.

We begin with a short review of the history of the problem as some of the explicit issues this paper addresses have been outstanding in the community for some time. The earliest version of this paradigm called the Interacting Stellar Wind (ISW) model was first described by Kwok, Purton & Fitzgerald (1978). According to the ISW model the slow wind was considered to be a spherical outflow such as the superwind phase of an AGB star (v ≲ 10 km s⁻¹). As the star evolves from the AGB phase to the PN nucleus, the speed of the line-driven winds would then rise up to v ≳ 1000 km s⁻¹, sweeping the slow wind material into a dense shell while also forming a hot bubble of shocked fast wind material. The bright rims of PNe are then the ionized shells of swept up AGB material. The ISW model has proven itself highly successful at explaining the properties of spherical or mildly elliptical PNe (Schönberner et al. 2014).

Most PNe, especially young PNe, are not spherical, and a generalization of this model relied on ‘inertial confinement’ of the spherical fast wind by an aspherical AGB slow wind (Balick 1987). Analytic solutions (Icke 1988) and later simulations (Icke, Balick & Frank 1992a; Icke et al. 1992b; Frank et al. 1993; Frank & Mellema 1994a,b) verified the ability of this Generalized Interacting Stellar Wind (GISW) model to capture observed PN morphologies. In the GISW model, the variation from mildly elliptical to strongly bipolar butterfly-shaped PNe is depended on the pole-to-equator density contrast in the AGB slow wind (ε = ρ_p/ρ_e), as well as the specifics of its aspherical morphology.

After the introduction of the GISW model, other shaping processes were considered, including the role of magnetic fields (García-Segura et al. 1999; Blackman et al. 2001a; García-Segura, López & Franco 2005; Nordhaus, Blackman & Frank 2007; Balick, Frank & Liu 2019), clumps (Steffen & López 2004), and the possibility that the
fast winds are collimated close to the star (Lee & Sahai 2003). New observations also showed that the PN shaping process may begin when the central star is still in the post-AGB or pre-PN (PPN) phases, before nebular ionization occurs (Bujarrabal et al. 2001).

The importance of collimated outflows as the driving source of PPN and PN was emphasized in a series of papers that attempted to strongly link observational characteristics of bipolar lobes to numerical models of winds and/or jets (Balick et al. 2017, 2018; Akashi & Soker 2018; Estrella-Trijillo et al. 2019; Rechy-Garcia, Peña & Velázquez 2019; Schreier, Hill & Soker 2019). From these results it appears that applications of the GISW model must be capable of taking a wind from the central engine of any geometry and converting it into highly collimated flow that then drives highly collimated PPN or PN bipolar lobes.

In spite of these advances, the GISW model often remains a background aspect of newer scenarios for PN formation. Many magnetohydrodynamic (MHD) models, for example, still require a strong pole-to-equator density contrast to produce very tight-waisted nebulae (i.e. the bipolar lobes converge to narrow width region near the source). Thus a pre-existing toroidal density distribution still figures centrally in our understanding of the PN shaping process.

One question that has not yet been fully addressed in GISW models of PNe is the origin of asphericity in the AGB slow wind. Single star models for the production of a pole-to-equator density contrast have been proposed (Bjorkman & Cassinelli 1993) but most of these rely on rotation of the central star. The required rotation rates for these models have proven to be difficult to square with observations (García-Segura et al. 2014). Binary stars, on the other hand, offer a natural source of angular momentum that can be tapped to create toroidal-shaped flows, i.e. a slow wind that has most of the material confined around the equatorial plane, or a small pole-to-equator density ρp/ρ. Binary star models for AGB wind shaping (and non-spherical PN production) have long been part of the PN literature (Fabian & Hansen 1979; Morris 1981, 1990; Soker & Livio 1989; Livio 1993; Nordhaus & Blackman 2006). More recent 3D simulation work has shown how detached binary interactions can produce density distributions that can shape PPN and PN flows (Chen et al. 2017).

Throughout the extensive discussion of binary stars and PNe, common envelope (CE) evolution has been seen as a principle means of generating high pole-to-equator density contrast toroidal flows (Livio & Soker 1988; Nordhaus & Blackman 2006). CE interaction occurs when a more compact companion plunges into the envelope of a red giant branch (RGB) or AGB star (Paczynski 1976; Ivanova et al. 2013). The release of gravitational energy during the rapid orbital decay unbinds most of the AGB envelope creating an expanding toroidal flow that later becomes the aspherical slow wind needed for GISW models (Nordhaus & Blackman 2006). The ability to numerically simulate CE interaction to the point where the expanding envelope could serve as input for GISW models has, in the past, been hampered by the complexity of CE 3D flows. Over the last decade, however, a number of new simulation platforms have been developed for CE evolution including smoothed particle hydrodynamics (SPH; Passy et al. 2012; Nandez, Ivanova & Lombardi 2014; Ivanova & Nandez 2016; Nandez & Ivanova 2016; Reichardt et al. 2019), adaptive mesh refinement (AMR) fixed mesh methods (Ricker & Taam 2012; Kuruwita, Staff & De Marco 2016; Staff et al. 2016; Iaconi et al. 2017, 2018; Chamandy et al. 2018), and moving mesh methods (Oehlmann et al. 2016; Prust & Chang 2019).

In this paper, we use an SPH CE simulation as the input for GISW models. Our goal is to explore the capacity of CE ejecta to serve as the source for strongly collimated bipolar flows in PPN. This question was also explored in García-Segura, Ricker & Taam (2018) who used AMR CE simulations as input. That work showed that PN collimation would occur but their results were restricted to 2.5D axisymmetric models. Our studies, which build on those of Frank et al. (2018), are fully 3D and, as such, we are particularly interested in the development of asymmetric features in the bipolar outflows.

The organization of the paper is as follows. In Section 2, we describe the methods and model. In Section 3, we present our results and explore the role of shock-focused inertial confinement (SFIC) in generating bipolar outflows. In Section 4, we characterize and discuss asymmetries that develop, as well as momentum budget questions. We conclude in Section 5.

### 2 SIMULATION SET-UP

The initial condition in our simulation comes from a single CE simulation (Iaconi et al. 2017, 2018) produced by the smoothed particle hydrodynamic (SPH) code PHANTOM (Price et al. 2017, 2018). The model used here is the 1.1 million particle simulation except for its gravity. The model used here is the 1.1 million particle simulation in Reichardt et al. (2019, rendered in figs 1 and 2, and the top two rows in fig. 14 of their paper). The simulation involves a 0.88 M⊙ RGB star with a 0.392 M⊙ core, and a 0.6 M⊙ compact companion. Thus the final binary system of the two stellar cores has a total mass of about 1.8 M⊙. The SPH simulation lasted for ~14 yr of physical time, and the binary orbital separation had been approximately constant for the last ~2000 d of that period. The orbit-averaged final separation is about 20 R⊙.

Our simulation uses the AMR code ASTROBEAR (Carroll-Nellenback et al. 2013). We compare simulations of jets with two different values of momenta and cases with no radiative cooling versus application of optically thin emission via a cooling curve to the outflow. Table 1 summarizes the four models presented in this work. Each model runs on 120 cores. Wall clock time for models A, B, and C evolving to the scale presented here is about 5 d, and 10 d for model D.

<table>
<thead>
<tr>
<th>Model</th>
<th>Cooling</th>
<th>Momentum</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No</td>
<td>High</td>
</tr>
<tr>
<td>B</td>
<td>No</td>
<td>Low</td>
</tr>
<tr>
<td>C</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>D</td>
<td>Yes</td>
<td>Low</td>
</tr>
</tbody>
</table>

The AMR simulation domain is a cubic box with side length of 128 000 R⊙. The coarsest grid cells have a length of 1000 R⊙ on each side. There are seven levels of AMR, with each finer grid having a refinement level of two over the previous. Hence, the finest grid has cells of ~7.8 R⊙ per side. The data for these large and small grids, along with the fourth level grid, were obtained by interpolating the density, internal energy, and velocity from the SPH CE simulation to 3D grids. This interpolation was done with the SPH visualization tool, SPLASH (Price 2007). The mapping is not strictly conservative, the gas quantities are interpolated and sampled at each point on a specified grid.

After mapping the initial condition on to the grids, a point particle of 1 M⊙ is placed at the centre of the domain to represent the binary system. The point particle interacts with the gas only by its gravity and has a softening radius of 39 R⊙ (five times the size of a seventh level grid cell). Self-gravity of the gas is not included. Note that the final binary separation of the SPH simulation is within this gravitational softening radius. Thus we do not resolve the evolution of the binary any further or consider its effect on the fast wind in our simulation except for its gravity.

Table 1. Summary of the four models.
A spherical fast wind is initialized within a radius of 46.9 $R_\odot$ (six times the size of a seventh level grid cell, hereafter $r_{FW}$) around the central particle. Within that radius, density and temperature are set to fixed initial values, and are updated at each time step. This fast wind condition is initialized at the beginning of our simulation, but with zero outflow speed until the time we ‘turn on’ the outflow by giving it a uniform radial speed in all directions.

We note that the choice of a spherical fast wind is likely an oversimplification. As has been discussed in the literature, the wind may form from the primary core or via a disc around the primary, the secondary, or the binary (Blackman, Frank & Welch 2001b; Matt, Frank & Blackman 2006; Guidarelli et al. 2019).

During initialization, a floor temperature of 20 K and floor density of $4 \times 10^{-15}$ g cm$^{-3}$ have been applied to remove noise from mapping of the SPH data to the AMR grids. After the first time step, the allowed minimum values in the simulation are lowered for temperature to 10 K, and for density to $1 \times 10^{-22}$ g cm$^{-3}$.

The first part of the simulation is a quiescent period of evolution to let the CE ejecta relax on the AMR grids. During this 3000-d quiescent phase, the ejecta expand by about four times in length scale without significant change of the morphology. This also ensures that, later in the simulation, bipolar lobes of the nebulae grow into regions already swept by the expanding ejecta, reducing possible artefacts from the SPH to AMR mapping. Fig. 1 shows the density distribution before turning on the fast wind. The fast wind boundary allows material to pass both ways, but whatever comes within the regions already swept by the expanding ejecta, reducing possible artefacts from the SPH to AMR mapping. Fig. 1 shows the density distribution before turning on the fast wind. The fast wind boundary condition show fallback towards the point source within the boundary during the entire quiescent period. This infalling material reflects the fact that our choice of fast wind boundary is arbitrary and some gas in the densest inner regions would still remain bound to the binary perhaps in an accretion disc that may be the source of the wind (Kuruwita et al. 2016). We note that the choice of fast wind boundary modifies the amount of CE material that falls back. In another test run where the fast wind boundary is doubled, $r_{FW} \sim 0.4$ au, the total fallback is about 0.23 $M_\odot$. These changes do not have a significant effect on our results, however, as a test run that has no quiescent period at all shows a morphology of the bipolar lobes that is qualitatively the same (see Fig. B1).

Parameters for the fast wind injection are chosen with the problem of momentum excess for PPN in consideration. We chose a radial speed reasonable for a post-AGB wind, and set the total mechanical luminosity ($L_{mech} = M_{FW} v_{FW}^2$) be comparable with the radiative luminosity of such a star. Therefore, all models have high enough momenta in the fast wind to account for observed high momenta PPN outflows. We discuss the momentum budget questions later in Section 4.2. Table 2 summarizes parameters related to the two momentum fluxes used in the simulation. The fast wind is turned on instantaneously after the quiescent period, with a radial speed of 300 km s$^{-1}$. For each momentum flux, we run two models comparing the case of no radiative cooling (models A and B) versus applying the Dalgarno–McCray cooling curve (Dalgarno & McCray 1972) to material above 30 000 K (models C and D). This critical temperature is chosen to let cooling happen mostly in material interior to the bipolar outflow but not so much in the ejecta. We note our choice is arbitrary for the purpose of comparative study, while Dalgarno–McCray cooling becomes significant above 12 000 K.

Table 2. Simulation’s input parameters.

<table>
<thead>
<tr>
<th>Models</th>
<th>High momentum</th>
<th>Low momentum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{FW}$ (g cm$^{-3}$)</td>
<td>$1 \times 10^{-11}$</td>
<td>$5 \times 10^{-13}$</td>
</tr>
<tr>
<td>$v_{FW}$ (km s$^{-1}$)</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Quiescent period (d)</td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>$M_{FW}$ (M$_\odot$ yr$^{-1}$)</td>
<td>$6.42 \times 10^{-4}$</td>
<td>$3.21 \times 10^{-5}$</td>
</tr>
<tr>
<td>$T_{FW}$ (K)</td>
<td>30 000</td>
<td>30 000</td>
</tr>
<tr>
<td>$L_{mech} (L_{\odot}) \geq L_{AGB}$</td>
<td>$4.75 \times 10^{3}$</td>
<td>$2.38 \times 10^{2}$</td>
</tr>
<tr>
<td>($M_{FW} v_{FW}$/$L_{mech} c$)</td>
<td>2000</td>
<td>100</td>
</tr>
</tbody>
</table>
a prism) whose length in $z$ increases. When mapping the SPH data, the base level grid covers the entire domain, while the fourth level grid occupies a cubic box of $\sim 37$ au per side, and the seventh level grid occupies a cubic box of $\sim 7$ au per side, both centred around the point particle. Intermediate levels are placed around (by the AMR routines) so that finer grids are nested within coarser grids, and the maximum refinement changes by one level between regions of different resolution. This cubic region of higher refinement ($\geq 4$), though a small fraction of the entire simulation domain, contains the entire CE ejecta in its initial condition and later on the central PN nucleus. The other part of the control is the prism with a square base equal to the seventh level box, but its height along the $z$-axis grows, leading ahead the extent of the outflow material. A critical outflow density is set to capture the extent of the outflow material. Refinement is forced to the seventh level within this prism, and gradually reduced outside by the AMR. A snapshot with grid patches can be found in Appendix C.

3 RESULTS

Here we present a description of the results from our four numerical models. We first describe the two adiabatic models in Sections 3.1 and 3.2. In Section 3.3, we elaborate on the detailed structure of the shocks in the outflow. Then, in Sections 3.4 and 3.5, we point out features that change when the system is allowed to cool via optically thin radiative emission. For each model, we choose four snapshots to represent the time evolution. The times are chosen such that the lobes are of comparable size for each model even if this means the snapshots are taken at different times across models. The snapshots are selected to match the physical height of the top lobe in model D, because it is the slowest evolving model and we stop our simulation at 7000 d. Note that we have run models A and B to a larger physical scale (but less time) and we do not see significant change in the morphology. Thus the expansion of the bipolar outflows eventually becomes homologous. In Section 4.1, we address the asymmetries in the models.

Before we begin discussion of the results we remind the reader of the basic shock structure in G1SW models. As in any colliding flow, two shocks form as the fast wind interacts with the previously deposited slow moving material. An outer shock, moving radially outward from the central wind source, is created as slow wind material is accelerated and swept up via the ram pressure of the interior fast wind. In addition, an inner shock forms moving back to the fast wind source via the deceleration of the fast wind. The inner shock converts fast wind kinetic energy into thermal energy of post-shock fast wind material (this energy will be radiated away if cooling is present). If both the fast and slow wind are spherical, the two shocks will be spherical. If, however, the slow wind is aspherical, then the inner shock can take on a convex semi-elliptical shape. This distortion of the inner shock is important because the radial fast wind streamlines no longer encounter the shock normally but instead strike it at an angle. Streamlines passing through a shock at an acute angle are refracted away from the shock normal. In this way an aspherical slow wind in G1SW models leads to a ‘SFIC’ of the fast wind. This leads to well-collimated jets (Icke et al. 1992a,b). The stronger the pole-to-equator contrast ($e$) in the slow wind, the more effective is the shock focusing of fast wind material into bipolar jets (Frank 1999). The formation of jets and bipolar outflows from SFIC has been studied in a variety of contexts from PN to star formation (Frank 1994; Mellema 1994, 1995; Mellema & Frank 1995). In Section 3.3, we will use this basic understanding of shock structure to interpret our models.

The collimator for our flows is the highly aspherical CE ejecta (Fig. 1). In particular note the presence of what we will call ‘funnels’ above and below the central core deep in the CE ejecta. These are relatively low-density polar regions bounded by high-density equatorial walls and form due to the angular momentum transfer during the CE orbital insiprual of the companion. The pole to equator contrast at the end of the quiescent phase is $e = \rho_f / \rho_e \approx 0.126$ at radius $r \sim 0.2$ au (against the fast wind radius), and the minimum is $e = 1.67 \times 10^{-4}$ at radius $r \sim 0.4$ au.

3.1 Model A: adiabatic, high-momentum outflow

Model A has a high-density outflow of $6.42 \times 10^{-4}$ $M_\odot$ yr$^{-1}$ and evolves adiabatically. Fig. 2 shows the density evolution and Fig. 3 shows the temperature evolution over 920 d after the onset of the fast wind. All frames are sliced at the middle of the $y$-axis. The PN core (i.e. the binary) is at the centre of the frames (coordinate $x$, $y$, $z = 0$) and the binary orbit is in the $x-y$ plane. White contours, representing temperature of 10 000 K, are plotted over in the density plots as this is a good indication of the outline of the bipolar outflow.

The first thing to notice is that although the fast wind is injected spherically into the simulation grids, two highly collimated lobes emerge shortly after we turn on the fast wind. On the leading edge of the lobes is a shell of hot CE ejecta (AGB) material swept up by the outward going bow shock, represented by the light blue regions just behind the outer temperature contour in Fig. 2. Interior to this shell is a tenuous hot bubble of gas at temperature around 10$^5$ K and filled with shocked ejecta material. At early times, this hot bubble has almost uniform density around $10^{-13}$ g cm$^{-3}$, and later grows into a hollow cavity with density below $10^{-15}$ g cm$^{-3}$.

We see the SFIC process produces collimated jets that run from the central PN nucleus up and down along the polar axis. Fig. 3 shows abrupt temperature changes across a pair of lens-shaped inner shocks going backward into the fast wind. Kinetic energy is converted into thermal energy in the shocked fast wind material, which produces turbulence at the far end of the jets. Interior to the lens-shaped region, the jet cools with distance from the PN nucleus due to geometrical dilution. We note that in later stages of the simulation, the bipolar lobes grow laterally extending beyond the refinement region defined by the prism described in Section 2. For example, in Fig. 2, the sides of the bipolar lobes in light blue are in a region with four levels of AMR refinement. This lack of resolution might have suppressed turbulence (compare the density maps in Fig. C1).

When we use a tracer (not shown in figures) to track the fast wind, we find that the material originally injected near the equatorial plane has been redirected by the funnel in the dense CE ejecta into the collimated jets. To explore the ability for the CE ejecta to redirect the wind, we performed an additional test run. In this case the wind was confined to the equatorial plane rather than being spherically symmetric. Even in this case, using a maximally uncollimated wind we find the interaction with CE ejecta again drives bipolar lobes (see Appendix A).

The presence of asymmetries is an important result of our models, as PPN/PN observations often show significant differences between lobes in bipolar flows on small scales. On larger scales we see the upper lobe has a more pointed tip than the lower lobe, while the lower lobe is slightly wider at its mid-point. On a smaller scale, the detailed structures of the filaments are visibly different between the lobes. Similar asymmetries also present in an orthogonal
Planetary nebulae from common envelope evolution

Figure 2. Model A: density \((\text{g cm}^{-3}, \text{in log scale})\) plots of the high-momentum, no radiative cooling outflow taken at 3260, 3480, 3700, and 3920 d in the simulation (260, 480, 700, and 920 d after turning on the fast wind). White contours mark constant temperature at 10000 K. Note the outer contour coincides with the outward-moving bow shock of the bipolar outflow, while the inner contour indicates a lens-shaped shock moving backward into the fast wind material. More details on the shock structures in Section 3.3.

Figure 3. Model A: temperature (K, in log scale) plots of the high-momentum, no radiative cooling outflow taken at the same times as Fig. 2. The hot (\(\sim 10^5\) K) region defining the outer rim of the bipolar lobes is where the outer shock drives into the CE ejecta. A pair of low-temperature, lens-shaped regions behind the elliptical shock are fast wind material that cools via geometric dilution. Note that near the base of the outflow, colder material from the CE ejecta is dragged into the lobes. This region helps define the funnel in the CE ejecta that is responsible for the overall collimation of the bipolar lobes.

cut through the \(y-z\) plane. We quantify the asymmetries later in Section 4.1.

Note that near the base of the bipolar lobes we see cold ejecta material being dragged into the lobes. This region helps define the funnel in the CE ejecta that is responsible for the overall collimation of the bipolar lobes. The entrainment of collimating material into the lobes is an effect that was seen even in early SFIC studies (Icke et al. 1992a).
Figure 4. Model B: density (g cm$^{-3}$, in log scale) plots of the low-momentum, no radiative cooling outflow taken at 3640, 4060, 4380, and 4780 d (640, 1060, 1380, and 1780 d after turning on the fast wind). White contours mark constant temperature at 10,000 K. It takes much longer for the low-momentum outflow to evolve to similar physical scale compared to the high-momentum case.

Figure 5. Model B: temperature (K, in log scale) plots of the low-momentum, no radiative cooling outflow taken at the same times as Fig. 4. Note that the inner shock in the top lobe has been pushed back closer to the centre and has an irregular shape.

3.2 Model B: adiabatic, low-momentum outflow

Model B has an adiabatically evolving, low-density outflow with mass-loss rate of $3.21 \times 10^{-5} M_\odot \text{yr}^{-1}$. Other parameters are set the same as in model A. Fig. 4 shows the density evolution and Fig. 5 shows the temperature evolution on the same scale as for the previous section.

It takes longer for the lobes in model B to evolve to a similar size compared to model A, as would be expected given the lower value of the injected momentum. The first snapshot (leftmost) of model B is taken 380 d later than the first snapshot shown for model A, and the last snapshot (rightmost) is taken 860 d later than the last snapshot for model A.
In terms of morphology, once again we see ambient material piles up just after the bow shock, but in thinner layers compared to model A. The bow shock is aligned quite well to the 10 000 K temperature contour in Fig. 4. Inside the bow shock, the lobes appear more homogeneous in this case, both in density and in temperature. Features that have been pointed out in Section 3.1, such as the central pillars of jet, and the lens-shaped inner shocks, are still present but exhibit a different morphology. The inner shock in the top lobe has been pushed back closer to the centre and has an irregular shape like a candle flame, which can still be identified as the abrupt temperature change from less than 104 K to about 105 K in Fig. 5. This is to be expected given the lower momentum injection rate. Note that the inner shock in the bottom lobe, however, can hardly be recognized in the very turbulent flow. The jets are also narrower and quickly become more turbulent as they propagate outward.

Large-scale asymmetry is more apparent in model B. The bottom lobe has a smaller volume compared to the top one. Starting from the second panel shown in Figs 4 and 5, the leading edge of the bottom lobe has turned slightly to the right of the polar axis (the side of positive $x$). Thus the funnel in the CE ejecta, which is the driver of the collimation, is not symmetric. The low-momentum runs, with their lower ram pressure, are more sensitive to the density distribution in the funnel as we will discuss in Section 4.1. We also note that the enhanced asymmetry between the high- and low-momentum runs will be more exaggerated when we consider the impact of cooling.

### 3.3 Shocks in the GISW model

As discussed at the beginning of Section 3, the classical GISW model involves the development of three discontinuities: an outer shock moving into, and accelerating, the ambient medium; an inner shock moving back towards the central engine and decelerating the fast wind; and a contact discontinuity separating the two shocks. When the AGB/slow wind material is aspherical, these discontinuities all become aspherical as well. Furthermore, SFIC occurs when the inner shock takes on an elliptical morphology that collimates the post-shock fast wind flow. While the basic physics of these shocks has been understood since the studies of Icke et al. (1992a,b), Mellema & Frank (1995), and Frank & Mellema (1996), the use of AMR methodologies in ASTROBEAR in the current work allows us to explore the evolution of the flows for much longer times and see the inner shock dynamics at much higher resolution. As we shall see the high density contrasts of CE ejecta produce new behaviours not seen in previous studies.

In Fig. 6, we present a pseudo-colour image Mach number with velocity vectors tracing the flow pattern. While the right-hand panel shows the entire bipolar flow, the left-hand panel zooms in around the central nucleus on a scale 10 times smaller. Both panels are taken from model A at 480 d after the start of the outflow (the second panel in Figs 2 and 3).

We begin with the left-hand panel that shows the flow close to the fast wind source. Note first the low Mach number CE ejecta (dark blue, $M \sim 2$). The initial funnel in the CE ejecta has now become the base of the bipolar flow. Within the funnel and surrounding the central wind source we see the elliptical inner shock that goes backwards to the central source and shocks the fast wind material that passed through. This shock is quite small with a semimajor axis length of just $\sim 0.2$ au and an eccentricity of about 0.95. The position of this elliptical shock front remains stationary during the entire simulation. Within the elliptical shock we see highly supersonic, freely expanding fast wind material whose Mach number is increasing with distance from the central source, while the physical speed of the gas ($v_{\text{gas}} > 250$ km s$^{-1}$) stays close to the initial speed of the injected spherical wind ($v_{\text{FW}} = 300$ km s$^{-1}$). Such an increase in Mach number is to be expected as the fast wind temperature decreases with radius as $T \propto r^{-(2/3 \gamma - 1)} (T \sim r^{-4/3}$ for monatomic ideal gas) due to geometrical dilution.

Immediately beyond the elliptical shock, the Mach number in the post-shock fast wind material drops by a factor of 10, as the...
light blue region around the yellow/red ellipse. Note that while the Mach number drops as would be expected at a shock, it remains above unity indicating the flow is not subsonic ($M \sim 2$). Thus, as expected from SFIC theory, the oblique inner shock leads to the production of a flow that is both supersonic and collimated.

The relatively lower Mach number of the immediate post-shock fast wind also means it has a relatively high thermal energy content. Thus as this collimated flow expands along the $z$-direction, its thermal energy will lead to a lateral expansion of the flow. This geometric expansion will, once again, lead to a gradual increase in the Mach number. After the initial collimation at the small-scale inner shock, we see the development of a large-scale diverging supersonic flow that fills the lobes. It is the interaction of flow with both the CE ejecta and downstream fast wind material that creates the lens-shaped shocks discussed in Sections 3.1 and 3.2. The development of the lens-shaped shocks is a new feature of SFIC-driven bipolar lobes and could only be seen by tracking the flows with high resolution (AMR). In addition it can only form in a high density contrast environment such as that created by CE ejecta.

We now focus on the right-hand panel of Fig. 6 that shows the entire bipolar flow. Once again the spherical dark blue region is the slowly expanding, low Mach number CE ejecta. The outer parts of the lobes are also in dark blue, and a steep jump in Mach number from around 2 to 10 (dark blue to dark green) demarcates the bow shock separating the undisturbed CE ejecta and material that has been accelerated by the outer shock. The yellow-green pillars running up the centre of the lobes represent fast wind collimated by the small-scale inner shock whose Mach number has increased with the flow’s expansion. This material has high velocity ($\sim 260 \text{ km s}^{-1}$) and relatively high Mach numbers ($M \geq 15$).

To summarize, we see three layers of shocks in the bipolar lobes. From the outer boundary and moving inward, these are: the bow shock or outer/forward shock; the lens-shaped inner/backward shock; and the small-scale elliptical shock around the central region, which is also a backward shock.

### 3.4 Model C: cooling, high-momentum outflow

Energy loss or ‘radiative cooling’ in GISW models leads to the expectation that thermal energy gained behind shocks will be lost, leading to decreased pressure support for these regions of the flow. Thus rather than extended post-shock shells we expect to see thin, cold dense regions immediately following the shocks. Such thin cold ‘slabs’ are prone to a variety of instabilities including thin-shell modes (Vishniac & Ryu 1989).

To investigate the role of radiative cooling in our simulations in model C we apply the Dalgarno–McCray cooling curve in all grids with $T > 3 \times 10^4 \text{ K}$ to the condition of model A. Figs 7 and 8 are the density and temperature evolving 1600 d after starting the outflow. Insulation of Figs 7 and 8 demonstrates that the cooling does create thinner, colder, and denser post-shock regions. Compared with Model A we see the bipolar lobes have a reduced width that occurs as lateral expansion is inhibited by the loss of thermal energy to optically thin radiation. The bipolar shell is also clearly fragmented as is the interior pillar of the jet formed from the elliptical inner shock. It is of particular interest to note that the degree of asymmetry between the two lobes is greater with cooling (model C) than without (model A). This is likely due to the fact that the cooling triggers instabilities that are sensitive to small-scale details of the interaction of fast wind with the CE ejecta, the latter of which is inherently 3D and asymmetric, a point we return to in the next section.

**Figure 7.** Model C: density ($\text{g cm}^{-3}$, in log scale) evolution of the high-momentum outflow with radiative cooling. Panels, from left to right, are taken at 3460, 3840, 4200, and 4600 d (460, 840, 1200, and 1600 d after turning on the fast wind). White contours mark constant temperature at 10 000 K. Note the highly turbulent flow in the central jets.
3.5 Model D: cooling, low-momentum outflow

Large-scale asymmetry becomes most apparent in model D, where we apply cooling to the low-momentum flow from model B. Figs 9 and 10 show evolution in density and temperature, respectively, for 4000 d after the outflow is turned on. The outflow to the negative $z$-direction has been choked for more than 1500 d, so the bottom lobe is only about one-third compared to the top lobe in its $z$-extent. Also, we see the bottom lobe has been redirected to the right-hand side of the polar axis. While this was apparent in model B, it is also observed here to a greater extent especially in the central two panels where almost the entire lobe has been shunted to the positive $x$ side of the top lobe’s nominal symmetry axis.

Spatial evolution of the top lobe in model D is also the slowest. It takes 4000 d for the top lobe to reach the same height in the $z$-direction, while the other three models take less than 2000 d. This is as expected since for model D we have both a low-momentum injection rate and the presence of radiative cooling.

While the scale of these simulations is still smaller than most observed PPN, it is worth noting the similarity between the morphology of this model and OH 231.8+04.2\(^2\) that shows bipolar lobes of distinctly unequal size (Sánchez Contreras et al. 2018). In addition, a central ‘disc’ (dark lane) is apparent in the nebula. Finally, structures within the lobe termed the ‘spine’ and ‘skirt’ by Balick et al. (2017) appear to find analogues with the jets and lens-shaped shock seen in our models.

We note that we also carried out runs with no quiescent period and once again the bottom lobe develops much later compared to the top lobe. Thus the strong asymmetry between top and bottom occurs regardless of the inclusion of a quiescent period between CE interaction and fast wind injection.

4 DISCUSSION

4.1 Quantification of the asymmetries

In this section, we try to quantify the degree of asymmetry using parameters of length, volume, and cross-section associated with the lobes.

Instead of the white temperature contour used in the density plots, here we define the lobes as regions filled with material from the fast wind outflow. We use tracers to keep track of material in the outflow. Tracers act as labels attached to the injected fast wind material. They are attached to the outflow gas density in a 1:1 ratio, where there is higher density, there are more tracers. A minimum tracer density at $10^{-15}$ gc m\(^{-3}\) distinguishes the regions between outflow material and ambient material. Visually, this threshold density makes an envelope around the central jet pillar inside the PN lobes, but within the white temperature contour in the density plots.

Spatial extents of the lobes are shown in Fig. 11 using (a) the extremum $z$-coordinates and (b) position of the centre-of-mass along the $z$-direction. For each model, the solid line represents the top lobe and the dashed line represents the bottom lobe. Comparing the momentum flux under the same cooling assumption (A to B and C to D), we see the low-momentum cases not only lag the high-momentum cases, but their bottom lobes also suffer a stronger effect compared to the top lobes. On the other hand, adding a cooling term does not affect the top-to-bottom asymmetry much besides that it slows down the spatial extension (compare model A to C and B to D). The left-hand panel in Fig. 12 shows the integrated volume in each lobe. Similarly, the higher the momentum flux, or the more energetic

\(^2\)The morphology of OH 231.8+04.2 is tantalizingly similar to the asymmetries we simulate. However it should be pointed out that this nebula contains an AGB Mira and a distant companion, making it a far cry from a classic post-CE system. For an alternative model of this system, which however cannot account for the asymmetry of the nebula, see Staff et al. (2016).
Figure 9. Model D: density (g cm$^{-3}$, in log scale) evolution of the low-momentum outflow with radiative cooling. Panels, from left to right, are taken at 4760, 5580, 6400, and 7000 d (1760, 2580, 3400, and 4000 d after turning on the fast wind). White contours mark constant temperature at 10 000 K. The bottom lobe has hardly made way out of the CE ejecta for over 2000 d after injecting the fast wind.

Figure 10. Model D: temperature (K, in log scale) evolution of the low-momentum outflow with radiative cooling, taken at the same times as Fig. 9.

(no radiative cooling) the outflow is, the larger the lobes will be. The right-hand panel in Fig. 12 picks up more asymmetries in small scales. It is the measurement of cross-section of the lobes taken at points of 1/3 (dashed lines) and 2/3 (solid lines) of their z-extents (Fig. 11, left). Positive side on the y-axis is for the top lobes, and negative side is for the bottom lobes.

A plausible hypothesis for developing asymmetries between the top and bottom lobes is that while the fast wind is purely spherical, the CE substrate calculated via the SPH simulation is not expected to be top–bottom symmetric. While some symmetry breaking will be due to the numerical algorithm, the underlying physics of the CE interaction is that of a highly turbulent 3D flow that should
Planetary nebulae from common envelope evolution

Figure 11. Extents of the lobes in $z$-direction. (a) Left: absolute length of the lobes. (b) Right: position of the centre of mass. For both panels, solid lines are the top lobes and dashed lines are the bottom lobes (in absolute value). Note that once the slopes flatten and the lobe expansion will appear to be homologous.

Figure 12. (a) Volume of the lobes, solid lines for the top lobes and dashed lines for the bottom lobes. (b) Cross-section of the lobes taking at 1/3 (dashed lines) and 2/3 (solid lines) of the heights of the lobes, with positive $y$-axis for the top lobes and negative $y$-axis for the bottom lobes.

not be expected to maintain top–bottom symmetry. As we have emphasized above, the flows associated with CE interaction are inherently fully 3D. In particular, the binary stars’ movement through the CE gas sets up strong shearing shocks that lead to Kelvin–Helmholtz instabilities that drive local regions of turbulence. The complexity of the flows after repeatedly being shocked with the subsequent driving of unstable modes during each orbit creates local flow patterns that need not be axisymmetric or top–bottom symmetric. We note that such perturbations are seen not just in SPH simulations but also in AMR grid-based simulations (Chamandy et al. 2018, fig. 4; MacLeod, Ostriker & Stone 2018, fig. 1). In addition, even before the CE phase begins we can see the RGB and AGB stars, which will be the primary stars in CE interaction, will be convective. This means asymmetries should be expected to present at the start of CE evolution and therefore those asymmetries would also be expected in the ejecta into which our spherical wind propagates. Thus while the initial perturbations driving instabilities may have numerical seeds, the development of turbulent, asymmetric local flow conditions is to be expected. Therefore given the expected asymmetry in CE ejecta, it will manifest as a difference in wind obstruction between top and bottom and this will lead to asymmetries in the bipolar lobes’ evolution.

4.2 Angular distribution of linear momentum

In this paper, we have injected a wind with momentum outflow rates well in excess of that which can be supplied by radiation alone (see the last row of Table 2), even though the mechanical energy associated with this outflow is about the luminosity of a post-AGB wind (see the second to last row of Table 2). These values are chosen as input values because the bipolar outflows in PPN are observed to have very strong constraints of momentum (Bujarrabal et al. 2001; Sahai et al. 2008, 2017; Sahai & Patel 2015). Such outflows require another energy source besides radiation pressure that likely means extracting gravitational binding energy from within the binary, such as via accretion (Blackman & Lucchini 2014). This study addresses whether such outflows, even if highly uncollimated at the base, can emerge collimated.

In this context it is interesting to quantify the angular distribution of momentum that results from our collimation mechanism. By injecting the isotropic fast wind at fixed density and speed, we are inputting momentum to the simulation at a rate of $\dot{P} = 4\pi r_{FW}^2 \rho_{FW} v_{FW}^2$, uniformly distributed around the sphere with radius $r_{FW} \sim 0.2$ au. Most of this input momentum ($\sim 90$ per cent) gets absorbed by the CE ejecta through inelastic collision between the fast wind material...
and the ejecta, but some is redirected into the collimated jets and reinforces them. We compare the ratio between the momentum in the collimated jets and the inputting momentum from the same solid angle $\Omega$ subtended by the jet as

$$\frac{\text{Jet momentum flux}}{\text{Input momentum flux}} = \frac{\langle \rho_j v_j^2 \hat{z} \cdot \hat{r} \rangle |_\Omega}{\frac{1}{2} \rho_{FW} v_{FW}^2 (\Omega/4\pi)}.$$  \hspace{1cm} (1)

We measure the jet momentum at a sphere with radius 1 au from the central source, where the outflow has been strongly collimated, and the elliptical shock is interior to this region. Here, $\Omega$ is estimated from the traced outflow geometry, which is almost a cone shape. The numerator is computed by integrating ($\rho_j v_j^2 v_j z$) over the meshes along the sphere, then averaged over the area of the cone base ($\Omega^2/4\pi$). The denominator depends on the input density and speed, as well as the geometry. The factor of 1/2 in the denominator comes from isolating $\hat{z}$ component of the velocity.

For the two models without radiative cooling, we find the open angle of the collimated jet (i.e. $\Omega$) does not vary noticeably with time. We perform the above calculation for both top and bottom jets at four times throughout the evolution (same times as shown in Figs 2–5).

In the high-momentum model A, the collimated jet subtends a solid angle $\Omega \sim 0.37\pi$. For a single jet, this momentum ratio averages between 3.05 and 3.25. Given the spherical input momentum for the high-momentum case in Table 2, this gives momentum per unit time of the jet to be $\sim 500$ times $L_{\text{circ}}/c$.

The factor of focusing is stronger for less input momentum. In the low-momentum case, model B, we measure a smaller $\Omega \sim 0.2\pi$, and a higher momentum ratio between 6.1 and 7.8. Given the input spherical momentum for the low-momentum case in Table 2, this yields about 35–40 times $L_{\text{circ}}/c$ in each jet.

The calculations of this section have constrained the ratio of momentum in the collimated jet compared to just ‘cutting out’ that section of the solid angle of the initially spherically imposed momentum. The ratios that we find are greater than unity but less than 10. As such, matching the extremely high observed momenta in PPN outflows (Bujarrabal et al. 2001; Sahai et al. 2008, 2017; Sahai & Patel 2015) still requires an imposed isotropic momentum well in excess of what isotropic radiation from a star can provide. That is why we chose such high-momentum inputs for our isotropic source to supply such strong outflow such as accretion. We also note that by showing that a wind from the central PPN/PN engine of any geometry can be converted into a jet, our results can be used as inputs for models of larger scale bipolar outflows that begin with jets as initial conditions (Balick et al. 2017, 2018, 2019). This is an important point as it was not always clear that GISW models could produce the tight-waisted nebula seen in some PPN/PN. The models presented in this paper seem to indicate that GISW models can be the starting point for a wider variety of nebular models than previously thought.

As was discussed in Frank et al. (2018), we show symmetry breaking in the bipolar outflows driven by our GISW models. This occurs because the morphology of the bipolar lobes is shaped by the inherent asymmetries in the central funnel of the CE ejecta. These asymmetries affect the kinematic and energetic properties of the outflows. Thus our work may offer explanation for objects like OH 231.8+04.2 that shows significant asymmetries in their lobes (Sánchez Contreras et al. 2018).

It would also be important to carry out this simulation using CE simulations where most of the ejecta are formally unbound (for the problem of non-unbound CEs, see Reichardt et al. 2020). Finally, in order to properly model PN and PPN, it would be important to carry out a simulation that uses an AGB CE simulation, rather than an RGB one.

5 CONCLUSIONS

In this paper, we have shown that by beginning with full 3D CE simulation as an input, it is possible to drive strongly bipolar outflows from a maximally uncollimated (i.e. a spherical) wind from the post-AGB star or remnant binary system.

We find that CE ejecta provide the high pole-to-equator density contrast environment needed in the classical GISW model. Three layers of shocks are fully resolved in the simulations. These include the outer shock sweeping up the CE ejecta, the lens-shaped inner shock, and, most importantly, the small-scale, innermost, elliptical shock predicted by the GISW model. This SFIC shapes the spherical fast wind to become highly collimated bipolar outflow.

Our simulations start with high enough momenta in all four cases to account for observed high momenta PPN outflows. We assume other source to supply such strong outflow such as accretion. We see significant focusing of the input momentum into the jet, and the fraction being redirected depends on the input momentum.

We also note that by showing that a wind from the central PPN/PN engine of any geometry can be converted into a jet, our results can be used as inputs for models of larger scale bipolar outflows that begin with jets as initial conditions (Balick et al. 2017, 2018, 2019). This is an important point as it was not always clear that GISW models could produce the tight-waisted nebula seen in some PPN/PN. The models presented in this paper seem to indicate that GISW models can be the starting point for a wider variety of nebular models than previously thought.

As was discussed in Frank et al. (2018), we show symmetry breaking in the bipolar outflows driven by our GISW models. This occurs because the morphology of the bipolar lobes is shaped by the inherent asymmetries in the central funnel of the CE ejecta. These asymmetries affect the kinematic and energetic properties of the outflows. Thus our work may offer explanation for objects like OH 231.8+04.2 that shows significant asymmetries in their lobes (Sánchez Contreras et al. 2018).

ACKNOWLEDGEMENTS

This work used the computational and visualization resources in the Center for Integrated Research Computing (CIRC) at the University of Rochester and the computational resources of the Texas Advanced Computing Center (TACC) at The University of Texas at Austin, provided through allocation TG-AST120060 from the Extreme Science and Engineering Discovery Environment (XSEDE; Towns et al. 2014), which is supported by National Science Foundation grant number ACI-1548562. Financial support for this project was provided by the Department of Energy grants DE-SC0001063, DE-SC0020432, and DE-SC0020434, the National Science Foundation grants AST-1515648 and AST-1813298, and the Space Telescope Science Institute grant HST-AR-14563.001-A. YZ acknowledges financial support from University of Rochester Frank J. Horton Graduate Research Fellowship. ZC is grateful to the CITA National Postdoctoral Fellowship for financial support. TR acknowledges financial support from Macquarie University Research Excellence Scholarship. ODM acknowledges support from the Australian Research Council Future fellowship scheme FT120100452. EGB acknowledges additional support from KITP UC Santa Barbara funded by NSF grant PHY-1748958, and Aspen Center for Physics funded by NSF grant PHY-1607611. JN acknowledges support from NASA HST-AR-15044.

DATA AVAILABILITY

The data underlying this paper will be shared on reasonable request to the corresponding author.

REFERENCES


2866 Y. Zou et al.

MNRAS 497, 2855–2869 (2020)
Planetary nebulae from common envelope evolution

Frank A., Chen Z., Reichardt T., De Marco O., Blackman E., Nordhaus J., 2018, Galaxies, 6, 113

APPENDIX A: JET INJECTED IN EQUATORIAL PLANE

For comparison, we present a test run where the outflow is injected into the equatorial plane. The outflow is now confined in two opposite cones oriented along the z-axis. Outflow material is directed into ±z-directions, with a half-open angle of 15°, while the outflow density and speed are set the same as our high-momentum cases. Figs A1 and A2 show the evolution of this case in density and temperature, respectively.

As seen in other cases, bipolar lobes above and below the equatorial plane quickly developed along the z-axis. Even though this is another extreme case where outflow is injected in orthogonal directions to the lobes, it has still been redirected to the open funnel.
Figure A1. Density (g cm$^{-3}$, in log scale) maps for the test run where the outflow is confined in jets along the $x$-axis and injected into the equatorial plane. Panels are taken at 20 d apart after starting the fast wind. Note that scale is 10 times smaller compared to figures in the main text. This outflow is again collimated by the CE ejecta and bipolar lobes develop clearly after 40 d.

Figure A2. Temperature (K, in log scale) maps for the same test run shown in Fig. A1. Panels are taken at 20 d apart after starting the fast wind. Note the elliptical shock, which outlines the central cold region, is visible in all panels except in the first one.

APPENDIX B: TEST RUN WITHOUT THE QUIESCENT PERIOD

As part of our convergence and sensitivity study, we carried out a test run without the initial 3000-d quiescent phase. Here the spherical outflow starts at the beginning of the simulation with the same parameters as in model A. Time evolution of this run is shown in Fig. B1. It takes longer for the outflow to clear its way out of the central remnant and the bipolar lobes are about half the size compared to those in model A. The funnel is much denser since the CE remnant has no time to relax and expand. Top lobe in earlier time has a left tilt similar to those in models B and C.
APPENDIX C: AMR GRIDS AND REFINEMENT

To test the effect of the prism region of high refinement we performed a run that doubles the prism’s width to \( \sim 14 \text{ au} \). Other parameters were held the same as in model A. Fig. C1 shows the AMR grids plotted over density maps at 3700 d. The two left-hand panels are from the original model A, while the two right-hand panels are from the test run. While we see more small-scale features along the sides of the lobes (due to resolving Kelvin–Helmholtz instabilities) in the run with more refinement, the global flow pattern, including the shape and extent of the bipolar lobes, the central jet pillar, and the lens-shaped inner shock, remains the same for both runs.

Figure B1. Time evolution of a test run where outflow starts at the beginning of the simulation. Panels are snapshots taken at 260, 480, 700, 920, and 1080 d in the simulation. Pseudo-colour shows density (g cm\(^{-3}\), in log scale), and the white contours mark constant temperature of 10 000 K. As can be seen the evolution of the bipolar morphology is qualitatively similar.

Figure C1. Density maps and AMR grids taken at 3700 d in the simulation (same time as the third panel in Fig. 2). Two left-hand panels: model A. Two right-hand panels: test run that doubles the width of the refinement prism. Magenta squares mark the boundary of the rectangular prisms, where refinement is forced to be the highest level 7. The global flow pattern remains the same for both runs, while there are more small-scale features along the sides of the lobes in the right-hand panels.

This paper has been typeset from a \TeX/\LaTeX\ file prepared by the author.