Galactic seismology: the evolving "phase spiral" after the Sagittarius dwarf impact

Joss Bland-Hawthorn[1,2]⋆ and Thor Tepper-García[1,2,3]

[1]*Sydney Institute for Astronomy, School of Physics, A28, The University of Sydney, NSW 2006, Australia*
[2]*Center of Excellence for All Sky Astrophysics in Three Dimensions (ASTRO-3D), Australia*
[3]*Centre for Integrated Sustainability Analysis, School of Physics, The University of Sydney, NSW 2006, Australia*

4 September 2020

ABSTRACT
One of the most important discoveries to emerge from the ESA *Gaia* astrometric survey is the "phase spiral" pattern detected in the $z - V_z$ plane throughout the solar neighbourhood. In Galactic cylindrical coordinates (R, ϕ, z), individual stars have velocities (V_R, V_ϕ, V_z) and oscillation frequencies $(\Omega_R, \Omega_\phi, \Omega_z) = (\kappa, \Omega, \nu)$. The phase spiral is most evident when each point of the $z - V_z$ phase plane is represented by either $\langle V_R \rangle$ or $\langle V_\phi \rangle$, averaged over the volume, indicating that the disturbance has locked the vertical oscillations with the in-plane epicyclic motions. We revisit Binney & Schönrich (2018) who use an analytic model of a disc-crossing satellite (2×10^{10} M$_\odot$) to explain the phase spiral and its coupled behaviour in $\langle V_R \rangle$ and $\langle V_\phi \rangle$. With the aid of the AGAMA (initialization) and RAMSES (simulation) codes, we explore a high-resolution N-body model that is true to the analytic calculation. We find that the impulse triggers a superposition of two distinct bisymmetric ($m = 2$) modes – a density wave and a bending wave – that wrap up at different rates. Stars in the *faster* density wave wrap up with time T according to $\phi_D(R, T) = (\Omega_D(R) + \Omega_o) T$ where ϕ_D describes the spiral pattern and $\Omega_D = \Omega(R) - \kappa(R)/2$. While the pattern speed Ω_o is small, it is non-zero indicating that this is a dynamic density wave. The *slower* bending wave wraps up according to $\Omega_B \approx \Omega_D/2$ producing a corrugated wave, a phenomenon now well established in the Galaxy. The bunching effect of the density wave triggers the "phase spiral" as it rolls up and down on the bending wave ("rollercoaster" model). The phase spiral emerges slowly about 380 Myr after impact, and symbiotically in $\langle V_R \rangle$ and $\langle V_\phi \rangle$. It appears to be a long-lived, disc-wide phenomenon that continues to evolve over the entire 1 Gyr simulation, becoming increasingly wound up with the passage of time.

Key words: Surveys – the Galaxy – stars: dynamics – stars: kinematics – methods: N-body simulations – methods: analytic

1 INTRODUCTION

In April 2018, the ESA *Gaia* astrometric mission (Perryman et al. 2001; Prusti et al. 2016; Soubiran et al. 2018) presented its second data release (DR2). These observations have revolutionized both stellar and Galactic studies (DR2: Gaia Collaboration et al. 2018). After only two years of observations, the remarkable precision of *Gaia*'s measured stellar parameters has led to new discoveries and new fields of study. These data are being complemented *inter alia* by large spectroscopic surveys – RAVE (Steinmetz et al. 2006), APOGEE (Majewski et al. 2017), *Gaia*-ESO (Gilmore et al. 2012), LAMOST (Deng et al. 2012) and GALAH (De Silva

et al. 2015). In all respects, this is the golden age of Galactic archaeology (Freeman & Bland-Hawthorn 2002).

We highlight the remarkable discovery of an unexpected phase-space signal in the local stellar disc by the *Gaia* team (Antoja et al. 2018). In a volume element defined by $(\Delta R, R \Delta\phi, \Delta z) = (\pm 0.1, \pm 0.1, \pm 1)$ kpc^3 centred on the Sun, Antoja et al. (2018) detect a coherent spiral pattern in the phase plane defined by z and V_z. This phenomenon[1] is in-

---

[1] This feature has been variously referred to as the "snail" (Antoja et al. 2018), the "snail shell" (Li & Shen 2020), the "Gaia spiral" (Darling & Widrow 2019a), the phase-plane spiral (Binney & Schönrich 2018) and the phase-space spiral (Michtchenko et al. 2019). Consistent with the traditional use of "phase mixing" rather than "phase-space mixing," we adopt the term "phase spiral" (e.g. Khanna et al. 2019; Hunt et al. 2019; Mackereth et al.

⋆ Contact e-mail: jbh@physics.usyd.edu.au

dicative of a system that is settling from a mildly disturbed state to a stationary configuration through the process of phase mixing (Lynden-Bell 1967).

Since its discovery two years ago, up to a dozen independent research papers seek to explain or shed light on the phase-spiral phenomenon. Explanations to date invoke three different mechanisms: (i) partial coherence imposed by the recent dissolution of several star clusters (Candlish et al. 2014; Michtchenko et al. 2019); (ii) vertical oscillations excited by the strong buckling of a stellar bar (Khoperskov et al. 2019); and (iii) vertical and in-plane oscillations induced by a massive disc-crossing perturber (Binney & Schönrich 2018; Darling & Widrow 2019a). The first two mechanisms are of internal origin; the third external mechanism has received the majority of attention to date, not least because the Sgr dwarf galaxy is observed to be undergoing disruption as it crosses the disc. This is the focus of our attention because we consider it to be the most likely interpretation for the phase-spiral phenomenon (Laporte et al. 2019; Bland-Hawthorn et al. 2019). The effect is now observed over too large a surface area (see below) to be a coherent phenomenon arising from dissolving star clusters. Moreover, the extremely tight alignment of the plane of the stellar bar with the inner disc plane (Bland-Hawthorn & Gerhard 2016) imposes strong constraints on the idea of a buckling bar.

Three independent observing campaigns confirm the phase-space pattern and study it in more detail. With the aid of the GALAH survey, Bland-Hawthorn et al. (2019, hereafter B19) dissected the phase spiral as a function of orbit actions, stellar ages and stellar abundances. They were able to detect the pattern over a larger radial range ($7 < R < 9$ kpc) and azimuthal angle ($\delta\phi \lesssim \pm 15°$). The LAMOST survey provides some evidence for a changing phase-spiral pattern over the radial range $6 < R < 12$ kpc, as predicted in action-based, analytic models (e.g. B19, Fig. 20) and observed in the highest contrast numerical simulation to date (Laporte et al. 2019, hereafter L19). Both Li & Shen (2020) and B19 present evidence that the phase spiral is mostly confined to cold orbits and not associated with the hotter, α-enhanced disc, further underscored by the 2-component disc model in B19.

In these studies, the spiral pattern in phase space is found to be somewhat patchy and uneven. This is due in part to the well-documented kinematic (Widrow et al. 2012; Williams et al. 2013; Carlin et al. 2013) and density asymmetries (Yanny & Gardner 2013; Slater et al. 2014; Xu et al. 2015) when comparing the upper and lower sides of the Galactic plane. However, when the $z - V_z$ plane is encoded as the angular momentum about the z-axis, $\langle L_z \rangle$ (action $\langle J_\phi \rangle$), Khanna et al. (2019) find that the phase-spiral pattern is smooth and varies systematically as a function of disc location. This suggests that the emergence of the phase spiral is somehow related to the angular momentum properties of the disc, in accord with the BS18 model.

Galactic seismology,[2] a subset of Galactic archaeology,

2019). The advantage of the more compact language becomes apparent when used as an adjective, e.g. phase-spiral evolution, phase-spiral dynamics, etc.
[2] Widrow et al. (2012) coined the term 'galactoseismology' to describe the study of dynamical modes arising from disc perturba-

is useful for exploring past interactions and structural properties of the Galaxy. Specifically, we seek a better understanding of how a differentially-rotating stellar disc responds to a powerful impulse and how this sets up the phase-spiral phenomenon. If a satisfactory model can be achieved, it must surely provide evidence that dark matter, which is dominant in the regions of interest, absorbs energy and momentum from moving massive bodies like ordinary matter. Then we would have conclusive evidence that dark matter comprises non-degenerate fermions rather than something altogether more exotic (q.v. Hui et al. 2017; Milgrom & Sanders 2008).

In what follows, we start by reviewing the Binney & Schönrich (2018, hereafter BS18) model (Sec. 2), which provides the impetus for our new work, before discussing the overall strategy (Sec. 3). The main goal is to design and carry out a numerical simulation that bridges between the simplicity of an analytic model (BS18) and the complexity of a realistic numerical simulation (e.g. L19). In Sec. 4, we discuss the adopted parameters of the Galactic model and the intruder. The Galactic model is set up carefully with specific initial conditions that guarantee long-term stability in an N-body simulation. In Sec. 5, with reference to web links to our simulations, we describe our analysis and the new results. In the final discussion (Sec. 6), we show that the new simulations are a major improvement over earlier work (L19, B19), although many questions remain.

2   BINNEY & SCHÖNRICH (2018) MODEL

BS18 develop a purely analytic (toy) model to investigate the response of the Galactic stellar disc to the perturbation induced by an external massive object. They employ action-based modelling defined by distribution functions $f(\mathbf{J})$ that are analytic functions of the action integrals $\mathbf{J} = (J_R, J_\phi, J_z)$. With recourse to the Action-based GAlaxy Model Architecture (AGAMA) package (Vasiliev 2019), they produce a highly self-consistent, multi-component model of the Galaxy. Their vertical density profile through the solar neighbourhood is an excellent match to the observed two-component disc profile (Gilmore & Reid 1983). This is important if the model is to match the vertical actions (J_z) and frequencies ($\Omega_z \equiv \nu$) given the anharmonic nature of the vertical gravitational potential. They choose phase-space coordinates for a million stars in the solar neighbourhood and backwards integrate them to a time when the perturber first appears. The intruder is moving with a downward velocity of $\vec{v}_P \approx 300$ km s^{-1}, has a mass of $M_P = 2 \times 10^{10}$ M$_\odot$ and impacts the disc at $R_P = 18$ kpc along the Sun's radius vector. They are then able to compute the acceleration imposed by the intruder on the particles relative to the unperturbed acceleration.

BS18 observe that the largest off-planar perturbation is expected at $\phi - \phi_o \sim \pm\pi/2$ because these stars are either moving towards ($\phi - \phi_o < 0$) or away ($\phi - \phi_o > 0$) from the perturber as it transits the disc. (Here, $\phi_o = 0$ because the intruder transits along the Sun's radius vector.) Thus, stars approaching $\phi = \phi_o$ feel a net upward pull, and stars receding from $\phi = \phi_o$ feel a net downward pull. This action sets up a strong low-order bending mode across the disc. As

tions, both internal and external. This term is used interchangeably with Galactic seismology when it relates to the Milky Way.
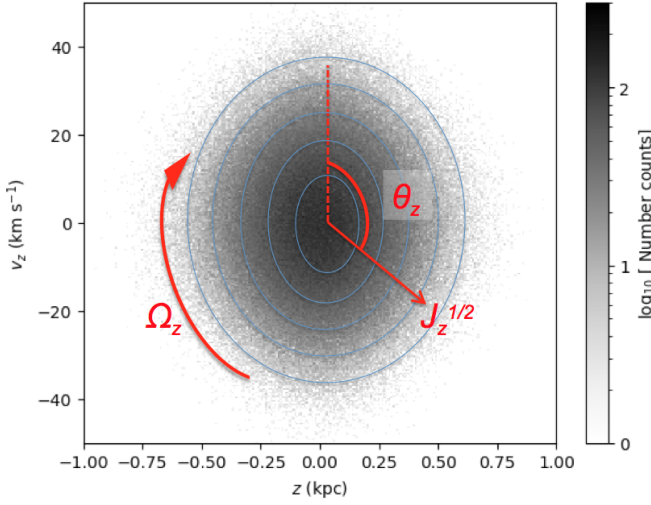
Figure 1. The natural coordinate system in the $z - V_z$ plane: the conjugate angle θ_z is the angle from the vertical; the radial distance of a star from the centre of this plane is proportional to the action $\sqrt{J_z}$. These are overlaid onto the projected stellar density in the solar neighbourhood where the simulated distribution arises from the vertical dispersion σ_z profile and the scale height of the disc (Sec. 4). The blue ellipses are surfaces of section for orbits with a fixed vertical action J_z. The vertical disc frequency Ω_z of the orbit determines the rate at which a star moves around the centre $(z, V_z) = (0, 0)$.
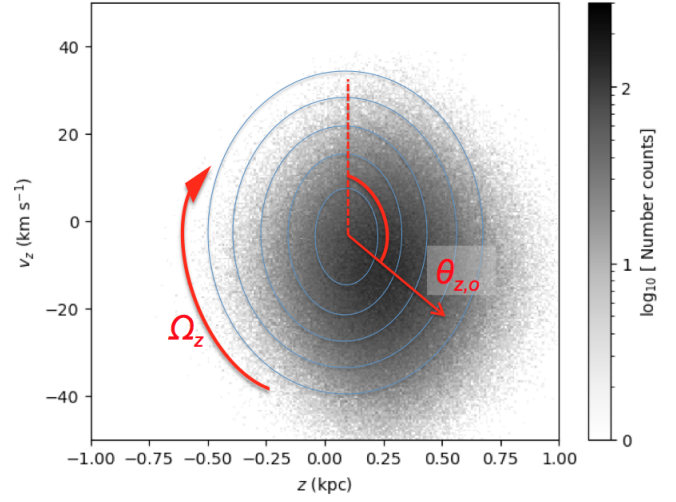

Figure 2. In the BS18 model, the impact causes the local stellar distribution to "slosh" in the $z - V_z$ plane with respect to the underlying large-scale potential. Each star now has a different vertical frequency Ω_z compared to its position in the unperturbed distribution and starts to circulate out of equilibrium (phase mix) in the $z - V_z$ plane.

the perturber transits the disc, stars near the crossing point experience a strong in-plane tug towards the perturber's trajectory, which binds the in-plane and vertical actions.

With the aid of Figs. 1-3, we illustrate a key insight from the BS18 model. The coordinate system of the $z - V_z$ plane is shown in Fig. 1 superposed on the surfaces of section for different stellar orbits confined to tori about the Galactic Centre. In Fig. 2, sloshing in the $z - V_z$ plane induced by the impact leads to a stellar overdensity that is offset with respect to the original stratification at a mean angle of $\theta_z \approx \theta_{z,o}$ (Fig. 2). Whereas before stars had constant values of J_z that were independent of the conjugate angle θ_z, the entire distribution now depends on θ_z and the assigned J_z associated with the new ellipse. Since for a disc potential, Ω_z declines with $\sqrt{J_z}$ (e.g. Candlish et al. 2014, BS18), a phase-spiral structure starts to develop. Crucially, BS18 note that the phase spiral is not apparent in stellar overdensity, although L19 detect a weak signal, in agreement with the original study (Antoja et al. 2018). Another mechanism is needed for the pattern to emerge so clearly in $\langle V_\phi \rangle$ and $\langle V_R \rangle$.

In Fig. 3, we illustrate how elliptic orbits from the inner and outer disc are able to enter the local volume. The vertical oscillation frequency Ω_z declines smoothly with increasing radius because of the disc's exponential decline in the local surface density. Stars from the inner disc reach us at their aphelia, and these are characterised by lower $\langle V_\phi \rangle$ and higher $\langle \Omega_z \rangle$; the converse is true for stars coming in from the outer disc. The dislodged stars trace ellipses in the $z - V_z$ plane with a conjugate angle that grows with time t as $\theta_z = \Omega_z t + \theta_{z,o}$ (Fig. 2). But Ω_z is tied to $\langle V_\phi \rangle$ (and $\langle V_R \rangle$) such that stars with lower $\langle V_\phi \rangle$ move ahead of stars with higher $\langle V_\phi \rangle$. Thus, the phase spiral is defined by a change in $\langle V_\phi \rangle$ (and $\langle V_R \rangle$) as we move clockwise around each ellipse in the $z - V_z$ plane.
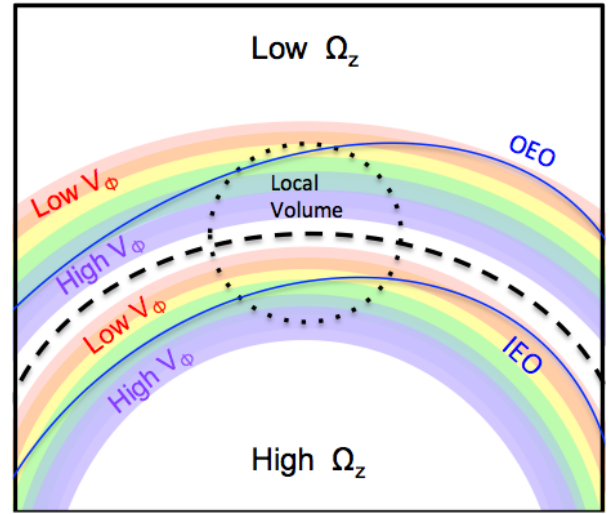

Figure 3. Stars from the inner and outer disc are able to reach the solar neighbourhood along elliptic orbits. We show an inner elliptic orbit (IEO) and an outer elliptic orbit (OEO); these precess about the Galactic Centre but remain confined to an annulus about the guiding radius (R_g) at their midpoint. In the BS18 model, the tangential velocity V_ϕ of a star in the Local Volume is intimately associated with the vertical frequency Ω_z of the star's orbit. Stars arriving from the inner disc typically have lower V_ϕ but higher Ω_z due to the higher disk surface density; the converse is true for stars arriving from the outer disc. Stars from the inner disc circulate faster in the $z - V_z$ plane (Fig. 1) compared to stars from the outer disc. (The chosen colour scheme is consistent with BS18.)

3   MOTIVATION & STRATEGY

The motivation for our new work can be stated as follows. We seek to understand the phase-spiral phenomenon with the simplest possible model that is true to the BS18 toy model while incorporating the advantages of the N-body approach. First, our work demands a more rigorous treatment for setting up the initial conditions. We require the coldest possible synthetic disc that comes into equilibrium at the start of the simulation and does not develop significant substructure or change its properties over the course of many disc rotations. Darling & Widrow (2019b) demonstrate that the emergence of the phase spiral is a competition between phase mixing and the disc's self gravity. Consistent with our approach, a lower self gravity leads to a stronger signal (higher contrast with respect to the background population) although a higher self gravity appears to sustain the signature for longer.

Secondly, we require a better treatment of the impulse approximation by an N-body solver. In particular, the heavy point mass must be 'live' on its first and only approach, and during the transit, and thereafter it must not continue to disturb the disc. In our earlier work (B19), the disc response can be more erratic in the post-transit phase due to the sustained force applied by the intruder. While there is an in-built asymmetry to our approach, this is true of all Sgr-based orbit families because of dynamical friction and the perturber's mass loss along its orbit (Law et al. 2005; Nichols et al. 2014). The underlying mechanisms involved in the disc response to an extended, disrupting perturber are obfuscated by many overlapping processes (e.g. L19, B19).

In contrast, the BS18 toy model uses a theoretical description of the Galaxy that is analysed in terms of distribution functions defined in action space, as discussed in Sec. 2. The latter model, while computationally efficient, lacks the realism of the actual system, most notably the complex recoil of both Sgr and the Milky Way to their mutual attraction and, importantly, the self-gravity and 'elasticity' of the dominant stellar disc. How the vertical and in-plane motions emerge and work together during the interaction are unclear (D'Onghia et al. 2016; Darling & Widrow 2019b). But by using a model that bridges the divide between analytic and N-body models, we seek to shed more light on this process.

If we are to compare the BS18 model meaningfully with an N-body simulation, the basic parameters of their set-up must be matched closely (Sec. 4.1). In our model, a point mass is used in place of an extended perturber to maintain the comparison and is placed on a hyperbolic orbit to minimize ongoing interactions with the disc. Aguilar & White (1985) note that in order to recreate the impulse approximation within an N-body simulation, the interval during which the mutual gravitational forces are significant is short compared to the crossing time within each system (see also Binney & Tremaine 2008, §8.2). The "impulsiveness" is further ensured (see Sec. A) by a perturber mass that is constant until the point of disc transit, and then declines exponentially in time such that the mass is negligible at the second disc crossing (Sec. 4.1). In the next section, we explore a new suite of models to simulate a clean impulse of a disc-crossing perturber.

In order to achieve the fidelity of the phase spiral in the BS18 models, these authors recommend that any N-body

simulation use $\sim 10^8$ disc particles, an order of magnitude larger than any matched, impulsive disc simulation to date. BS18 observe a synthetic phase spiral over an area in the $z - V_z$ plane that is comparable to Antoja et al. (2018), i.e. ±60 km s^{-1} and ±1 kpc, a fourfold improvement over L19 and B19. After following BS18's recommendation, we also achieve phase spiral patterns over the smaller region.

4   N-BODY SIMULATIONS

4.1   The model

We conduct a series of N-body simulations using the tried and tested RAMSES simulation package (version 3.0 last described by Teyssier 2002). The Galaxy is approximated by a three-component system consisting of a dark-matter (DM) host halo, a stellar bulge, and a cold stellar disc; a complete list of parameters is given in Tab. 1. The stellar disc is sampled with 5×10^7 particles, the largest impulsive model of this kind to date. The model Galaxy has properties that are consistent with their observed counterparts (q.v. Bland-Hawthorn & Gerhard 2016). Its total mass is $M_{\rm tot} \approx 1.45 \times 10^{12}$ M$_\odot$ with a distribution that yields a nearly flat rotation curve out to $R = 20$ kpc; the tangential velocity remains bounded within 240 to 250 km s^{-1} (Fig. 4, top right). The density and integrated mass profiles of each component are also shown (top left), along with the surface density profile of the disc alone (bottom right).

The LAMOST and GALAH studies reveal that the phase spiral is mostly confined to stellar populations with the coldest orbits (B19, Li & Shen 2020). In order to enhance the contrast of the phase-spiral signature in the stellar component, we achieve one of the coldest, long-lived disc simulations to date (Sec. 4.1). The disc is stabilised by our choice of Toomre (1964)'s stability parameter

$$Q_\star \equiv \frac{\sigma_R \, \kappa}{3.36 \, G \, \Sigma_\star} = 1.5, \tag{1}$$

where κ and $\Sigma_\star$ are the epicyclic frequency and stellar surface density, that provides local axisymmetric stability if we ensure $Q \gtrsim 1.3$ throughout the disc. Our choice of Q and disc mass fraction (Sec. 4.3) ensures that no bar or spiral arms form in isolation over a time frame of 4 Gyr (Fujii et al. 2018, see Sec. 4.3 below). These complicating factors, that were not considered by BS18, are deferred to a follow-up paper. The stellar dispersion profiles $\sigma_R(R)$ and $\sigma_z(R)$ are presented in Fig. 4 (bottom left); these apply to the unperturbed Galaxy model in isolation, but they are identical to the corresponding properties of the interacting Galaxy at the start of the simulation.

In line with BS18, the Sagittarius dwarf galaxy (Sgr) is represented by a point-like object. The point mass is placed on a hyperbolic orbit that intersects the Galactic plane at ~ 18 kpc, at which point Sgr is travelling at ~ 330 km s^{-1}. These constraints are consistent with Sgr's trefoil orbit parameters at the last crossing (e.g. Tepper-García & Bland-Hawthorn 2018). The motivation for using a hyperbolic orbit is to limit the interaction between the Galactic disc and the point mass to a single impact (impulse approximation) and then trace the disc's evolution over a timespan of up to 1 Gyr. This extended time frame is in line with earlier work

Table 1. Galaxy model parameters: In columns 1 and 2, we show the Galactic component and the intended functional forms; these are only approximate because they share the same gravitational potential. In column 3, the function variable is shown where (R, z) are cylindrical coordinates, and r is the spherical radius. The total mass, scale length and cut-off radius are indicated in columns 4, 5, 6 respectively. Column 7 is the number of collisionless particles used in the simulation.

Component	Function	Variable	Total mass $M_{\rm tot}$ $(10^{10}\ {\rm M_\odot})$	Radial scale length r_s (kpc)	Cut-off radius r_c (kpc)	Particle count N (10^6)
DM halo	NFW	r	140	15	300	20
Stellar bulge	Hernquist	r	1.5	0.6	2.0	4.5
Stellar disc	Exp, sech^2	R, z	3.4	3.0	40	50

Notes: The NFW and Hernquist functions are defined elsewhere (Navarro et al. 1997; Hernquist 1990). The scale height of the stellar disc is $z_t \approx 250$ pc; the local instability parameter is everywhere $Q \gtrsim 1.3$ (Toomre 1964).
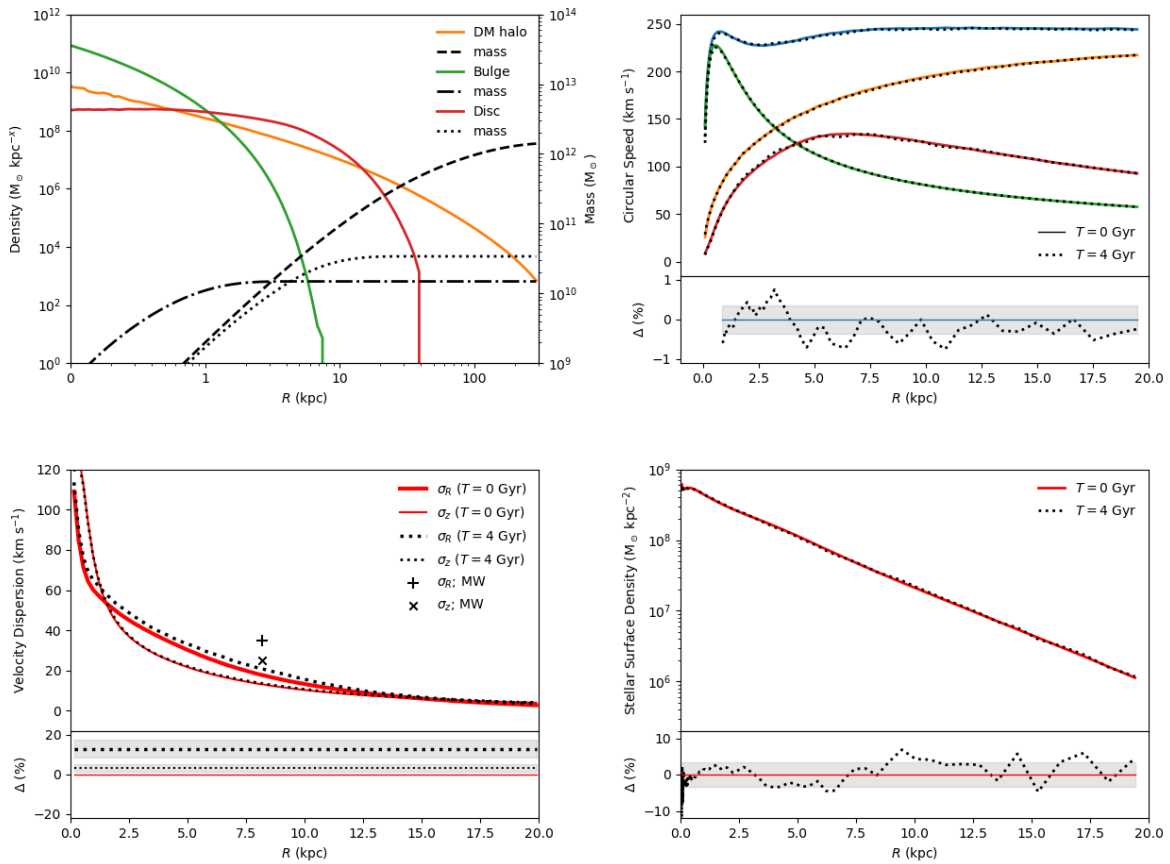

Figure 4. (Top left) The density and integrated mass profiles of our three-component Galaxy model. The curves corresponding to the bulge's and disc's density distribution truncate because of the limited spatial extension of these components. The DM halo extends out to 300 kpc. Note that the density distribution of the halo and bulge are given in $\rm M_\odot\ kpc^{-3}$ (volume density), while the disc's density distribution is given in $\rm M_\odot\ kpc^{-2}$ (surface density). (Top right) The flat rotation curve (blue) of the synthetic Galaxy taken from the N-body simulation showing an acceptable resemblance to the Milky Way (Bland-Hawthorn & Gerhard 2016). The contributions to the rotation are also shown: bulge (green dot-dashed line), disc (red dotted line), dark halo (orange dashed line). The sub-panel shows the relative change (in percent) in the total rotation curve between these time frames. The gray-shaded area indicates the rms value around the mean. (Bottom left) The radial σ_R (solid) and vertical σ_z (dot-dashed) velocity dispersion profiles of the cold, synthetic disc extracted from the N-body simulation. In comparison to the Galaxy (Bland-Hawthorn & Gerhard 2016), $\sigma_R \approx 35$ km s^{-1} (+) and $\sigma_z \approx 25$ km s^{-1} (×) at the solar circle ($R_\odot = 8.2$ kpc), as indicated. The sub-panel shows the *mean* relative change (in percent) in each of the velocity dispersions between these time frames (horizontal lines: $\Delta(\sigma_R) \approx 13\%$; $\Delta(\sigma_z) \approx 3\%$) and the corresponding rms value around the mean (gray-shaded area). (Bottom right) The surface density of the disc: note that the red line is identical to the red curve in the top left panel. The sub-panel displays the relative change in the surface density between these two time frames. The gray-shaded area indicates the rms value around the mean.

that shows the phase spiral may be long-lived (Darling & Widrow 2019b, B19).

The adopted orbital parameters given above imply that the perturber is moving on a bound orbit, and must therefore cross the disc more than once in the allotted time frame. To circumvent this problem, we impose on Sgr the following mass evolution:

$$M_P(t) = M_P(0) \times \begin{cases} 1 & t < t_\circ \\ \exp\left[-(t - t_\circ)/\tau_s\right] & t \geqslant t_\circ \end{cases}$$

where $M_P(0) = 2 \times 10^{10}$ M$_\odot$ and $\tau_s = 30$ Myr, roughly a disc crossing time. The first disc crossing happens at ~ 100 Myr so we set $t_\circ = 150$ Myr. Thus the intruder mass remains constant until 40–50 Myr after the first disc transit, and declines by a factor of $\sim 10^6$ at the time of the second disc crossing ($t \approx 450$ Myr). The perturber's post-transit impact on the disc is entirely negligible. We experimented with these parameters and found that the impulse approximation is robust to our assumptions (Sec. A).

4.2 Initial conditions and evolution

In this work, we pay particular attention to the initial conditions (ICs) that specify our N-body models at the outset. A recurrent problem with N-body simulations to date, with rare exceptions (Widrow & Dubinski 2005), is that the initially specified, multi-component system evolves to a new configuration before long-term stability is achieved; the simulator must then accept a model that is less than ideal and not what was specified. This has been a longstanding problem with numerical simulations that the AGAMA code originally set out to resolve (Vasiliev 2019).

We exploit AGAMA to generate the positions and velocities for all components of our model (Table 1) in a self-consistent fashion. This guarantees that the ICs are virtually stationary from the outset[3].

The ICs of the perturber are calculated using a backwards integration scheme in the two-body approximation where the Galaxy is described by an extended rigid body (for details, see Tepper-García et al. 2020). This is subject to the condition that the point mass crosses the Galactic plane after 95 Myr at a Galactocentric distance of about 18 kpc along the negative x-axis, with its velocity vector perfectly perpendicular to the Galactic plane and a magnitude of 330 km s^{-1}. The initial position and velocity of the perturber with respect to the Galactic Centre thus obtained are $\vec{r}_P \approx (-11.1, 0, 28.6)$ and $\vec{v}_P \approx (-145.4, 0, -220.2)$ respectively.

The compound (Galaxy + point mass) N-body ICs are evolved with the RAMSES code (Teyssier 2002), which incorporates adaptive mesh refinement (AMR). The system is placed into a cubic box spanning 600 kpc on a side. The AMR grid is maximally refined up to level 14, implying a limiting spatial resolution of 600 kpc$/2^{14} \approx 36$ pc. The total simulation time is 1 Gyr for the phase-spiral enactments. We also explored a more expensive simulation with 2 pc resolution, but there were no perceptible benefits for our phase-spiral study, so we did not proceed further.[4]

Due to the presence of the simulation mesh, the force exerted by (and onto) any particle is necessarily softened, at least by a length equivalent to the side of cubic volume element (cell) equal to the limiting spatial resolution. We choose to soften the force exerted by the point-like perturber using a length equal to two such elements ($\epsilon_s \approx 72$ pc); for all other particles, ϵ_s is half this value. In addition, the force between the perturber and all the other particles is calculated by a direct N-body solver, rather than the particle-mesh (PM) method used otherwise.

From its starting point above the Galactic plane, the perturber reaches the stellar disc travelling at about 350 km s^{-1} after 95 Myr from an initial galactocentric distance of about 40 kpc above the disc. We provide a link to the movie in the footnote.[5] It crosses the disc at $R \approx 18$ kpc (i.e. $(x, y) \approx (-18, 0)$ kpc), and transits below the Galactic plane to a vertical distance of about $z \approx 35$ kpc some 250 Myr after the disc crossing, then falls back towards the Galaxy – see Fig. 5. At $t_\circ = 150$ Myr, i.e. shortly after it crosses the Galactic plane, we artificially decrease its mass over an e-folding timescale, $\tau_s = 30$ Myr (Sec. 4.1), such that by the time the intruder crosses the disc again about 450 Myr after the first transit, its mass is entirely negligible.

4.3 Model stability

We have studied the long-term stability of our models with a series of expensive simulations for an isolated galaxy run over 4 Gyr and at different resolutions. We made this investment because cold discs are susceptible to the emergence and evolution of spurious substructure (e.g. clumps, bars, spiral arms) as the disc evolves (D'Onghia et al. 2013; Fujii et al. 2018). To underscore the importance of this test, the reader is encouraged to compare two disc simulations with $N = 10^6$ and $N = 10^8$ particles (see footnote 5). By $T \sim 1$ Gyr, the low-resolution simulation excites complex bending modes in the outer disc that only get worse as time marches on, whereas no such effect is apparent in the high-resolution simulation. In the high-resolution simulation, the surface density profiles are reasonably constant over the full timespan (< 5 percent, rms), which amounts to about 20 disc rotations at the solar radius (Fig. 4, bottom right). The other panels show that deviations from the original rotation profiles are less than 1 percent at all radii.

We also examined the stellar dispersion profiles in both R and z. The vertical dispersion profile $\sigma_z(R)$ is essentially unchanged at all radii (< 5 percent), which shows that there is no disc heating induced by the coarseness of the finite resolution at any radius. In separate tests, we find that we can avoid internal disc heating when the number of disc particles exceeds about 10 million particles. The radial dispersion profile $\sigma_R(R)$ shows a slight excess (13 percent on average) over 4 Gyr; we deem this to be acceptable given the goals of the current work. From our perspective, the parametric profiles are essentially constant, such that the ultrathin disc retains the imposed ICs over many rotation periods, providing confirmation of the power of AGAMA in setting up robust models.

---

[3] To encourage more extensive use, our AGAMA setup files are available upon request.
[4] Our RAMSES setup files are available upon request.

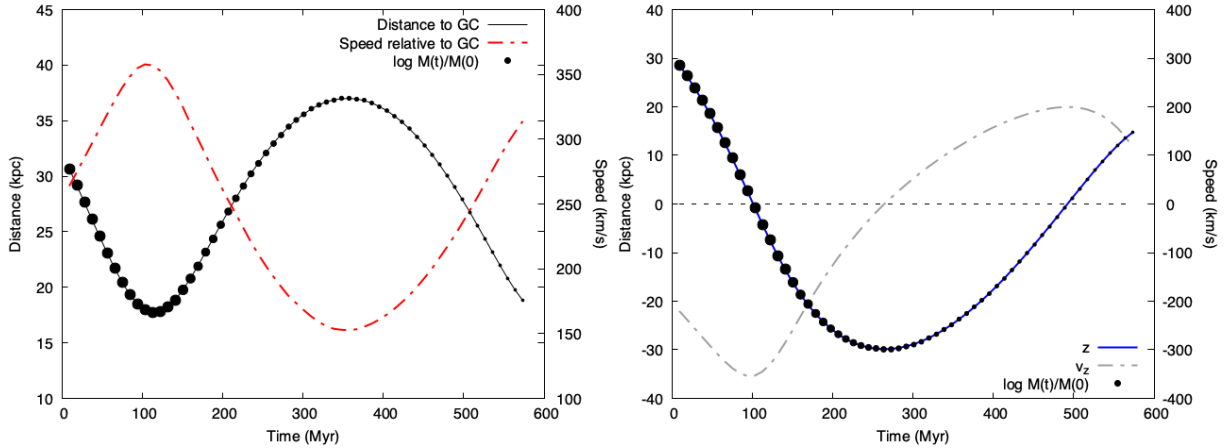[5] http://www.physics.usyd.edu.au/~tepper/proj_galseismo.html

Figure 5. Orbital history of a point mass perturber (Sgr) around the Galaxy in the N-body simulation. (Left) Total distance (black solid curve) and speed (red dot-dashed curve) relative to the Galactic Centre. (Right) Vertical distance z (blue solid curve) and vertical speed V_z (gray dot-dashed curve) relative to the Galactic plane. The disc crossings are identified by the intersection of the blue curve and $z = 0$ (dashed horizontal line) and are clearly visible at $T \approx 100$ Myr, and $T \approx 500$ Myr (bottom). The black dots on top of the black curve (top) and of the blue curve (bottom) are indicative of Sgr's mass along its orbit. The size of the symbol is directly proportional to $\log_{10}(\text{mass})$ relative to its initial value.

5 ANALYSIS AND RESULTS

5.1 Early stage evolution

In analysing the simulations, our first goal is to understand just how the cold disc experiences the impulsive force from the disc-crossing intruder. Our second goal is to determine how efficiently the perturber's impulse is able to trigger a response in the cold disc. This must depend both on the infall velocity ($\vec{v}_{\rm P}$) and the perturber mass ($M_{\rm P}$), with lower velocities and higher masses triggering a stronger response. Earlier simulations (Darling & Widrow 2019c) employ perfect coupling by inserting the bending modes by hand. But we seek to understand how the phase-spiral signal builds over time as a diagnostic for age-dating the phenomenon.

First, we investigate how well the BS18 picture is reproduced in our simulations. Movies of the disc-perturber interaction are available at our website; see footnote 5. We stress that, at each timestep, all results are shown in the reference frame of the simulated galaxy's centre of mass where the disc's angular momentum vector is aligned with the positive z-axis. The perturber moves parallel to the $+z$ axis towards the plane.

In Fig. 6, the response of the disc is presented in the $x - y$ plane at $T = 67$ Myr, $T = 95$ Myr (time of transit), $T = 143$ Myr and $T = 190$ Myr. Each time step shows three panels: (1) the projected stellar density; (2) the projected mean z-height $\langle z \rangle$ of the stars; (3) the projected mean z-motions $\langle V_z \rangle$ of the stars. In the $x - z$ projection, even before the disc crossing ($T = 67$ Myr), stars move in z towards the perturber along the impact trajectory. These are mostly stars that were moving *away* ($\phi - \phi_o > 0$) from the impact point ($\phi_o = 0$) in the disc plane (cf. BS18, Fig. 1) after transit ($T > 95$ Myr).

Conversely, after the perturber has transited, stars are observed to pursue the perturber along its impact trajectory. These are mostly stars that were moving *towards* ($\phi - \phi_o < 0$) the impact point in the disc plane at the time of transit. This

picture is in good agreement with the BS18 model, although it misses an important ingredient, as explained below.

The early phase of the disc's evolution is broadly consistent with the BS18 picture summarised in Sec. 2, except they do not consider the influence of the shearing disc due to differential rotation. We believe this to be crucial to how the phase spiral emerges. At $T = 143 - 190$ Myr, the first signs of a density wave become evident in the projected density map. Moreover, the impulse sets up a strong, low-order ($m = 1$) bending mode across the disc seen initially as the sine-wave warp of the outer disc in the $\langle z \rangle$ and $\langle V_z \rangle$ panels. Of particular interest is the transition from $m = 1$ in the outer disc to an $m = 2$ bending mode inside $R \approx 10$ kpc that wraps up with the differential rotation. (We explore this behaviour with a model in a later section.)

The early stages set up a sequence of events that unfold in the remaining time steps for which we show representative images sampled at intervals of 95 Myr (Fig. 7). The perturber triggers both a density wave (left panels) and a bending wave (middle and right panels) that wrap up as time passes. The panels are encoded in the same way as Fig. 6 but now zoomed to show the inner disc ($R < 11$ kpc) more clearly. Note that the vertical motion $\langle V_z \rangle$ is out of phase with the vertical displacement $\langle z \rangle$. This is expected for a corrugated wave that oscillates up and down as it wraps up. Observe also that the $\langle z \rangle$ displacement increases with radius as expected (B19, Figs. 20 and 21).

The twelve circles shown in each of the panels of Fig. 7 are the sampled volumes in our later discussion of phase-spiral evolution. Note three things: (i) our solar neighbourhood is "volume 1" ($x = -8.2$ kpc) and the intruder crossing point is at $x = -18.0$ kpc; (ii) at $T = 95$ Myr, these lie along the same radius vector at disc transit; (iii) at $T \approx 200$ Myr, volume 1 has moved to the other side of the disc since the impact, consistent with the orbital period at the Sun's radius ($T_\odot \approx 210$ Myr), while the crossing point ($T_C \approx 460$ Myr orbital period) has only moved by about 80° bringing it in line with volume 10. After a full orbit of volume 1 ($T = 305$
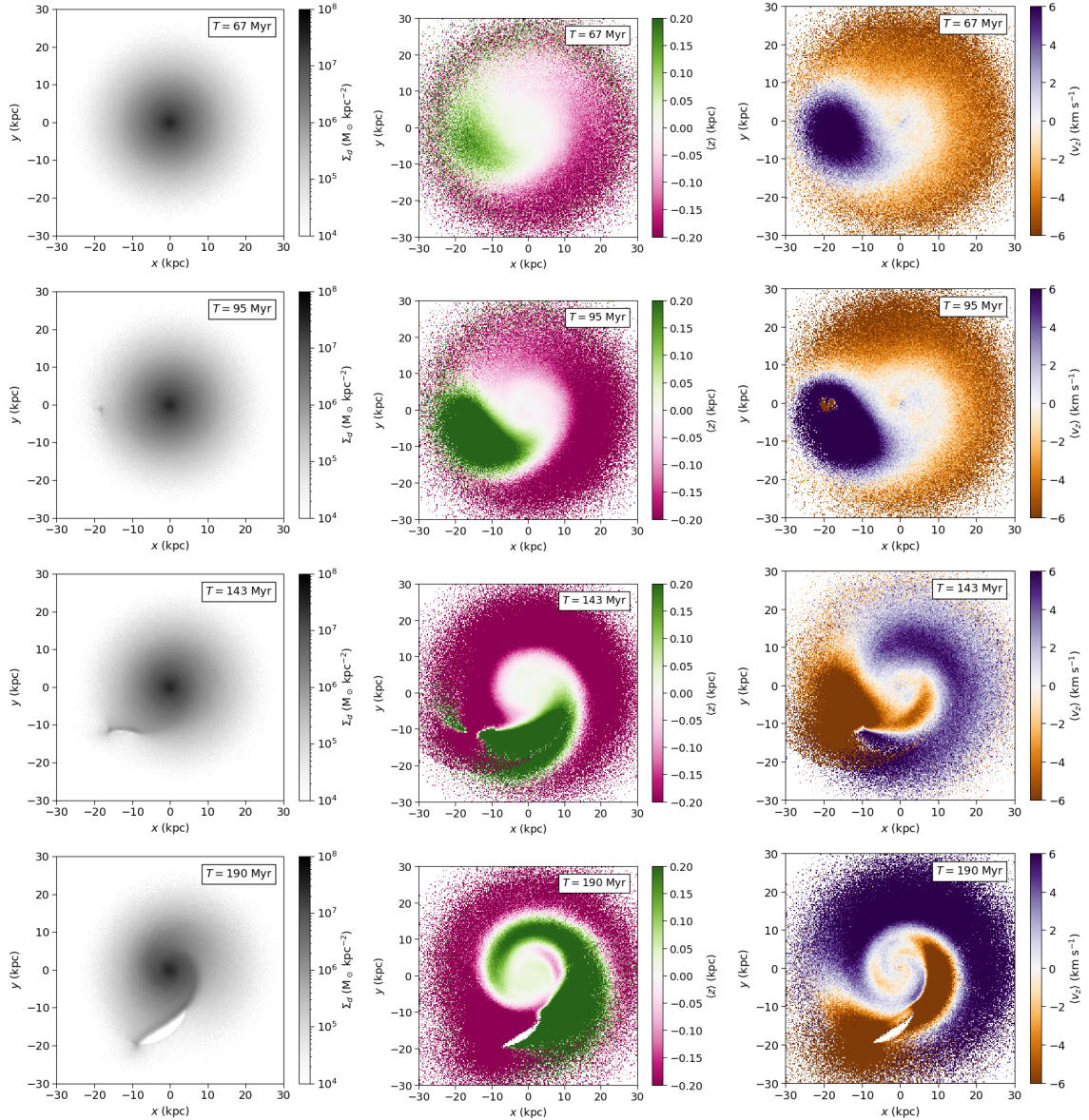
Figure 6. The projected stellar density distribution of the disc (left), the average height of the stars (middle), and the average vertical speed of the stars (right) at four time steps ($T = 67, 95, 143, 190$ Myr). The disc's rotation is counter-clockwise in this projection. The point mass crosses the disc at $(x, y) \approx (-18, 0)$ kpc ($T = 95$ Myr). The disturbance is already apparent at $T = 67$ Myr when the perturber is about 20 kpc from the impact point. Stars in the vicinity of the impact point are initially lifted towards the intruder. A bending wave is triggered by the impulse and it immediately begins to wrap up with the differential rotation.

Myr in our time frame), the crossing point now lies along the radius vector with volume 7.

5.2 Middle to late stage evolution

5.2.1 Kinematic density wave

For eighty years, we have known that galaxy interactions excite trailing spiral arms in disc galaxies (Holmberg 1941). Thereafter, Lindblad discovered that the force fields of disc galaxies preserve twofold symmetric (bisymmetric) structures against dissolution from the differential rotation (Lindblad 1956, 1960). In an isolated galaxy, stellar orbits are mostly elliptic and precess about the centre. An external

perturber can organize the elliptic orbits into a coordinated precession (2-arm spiral pattern) that depends on the local angular frequency $\Omega(R)$ and radial frequency $\kappa(R)$, subject to the shearing disc's lowest order resonances, $\Omega(R) \pm \kappa(R)/2$ (Kalnajs 1973). These features are observed in numerical simulations to be associated with density waves that wind up slowly, rather than material arms that wind up rapidly with the disc's differential rotation (Sundelius et al. 1987; Oh et al. 2008; Dobbs & Baba 2014).

The spiral pattern is the natural response of a cold, differentially rotating disc subjected to a strong impulsive force (Binney & Tremaine 2008, §6.2). For an $m = 2$ density wave with trailing arms ($\Omega > \kappa/2$), as seen from the positive z axis, the spiral pattern $\phi_D(R)$ (R in units of kpc) winds up
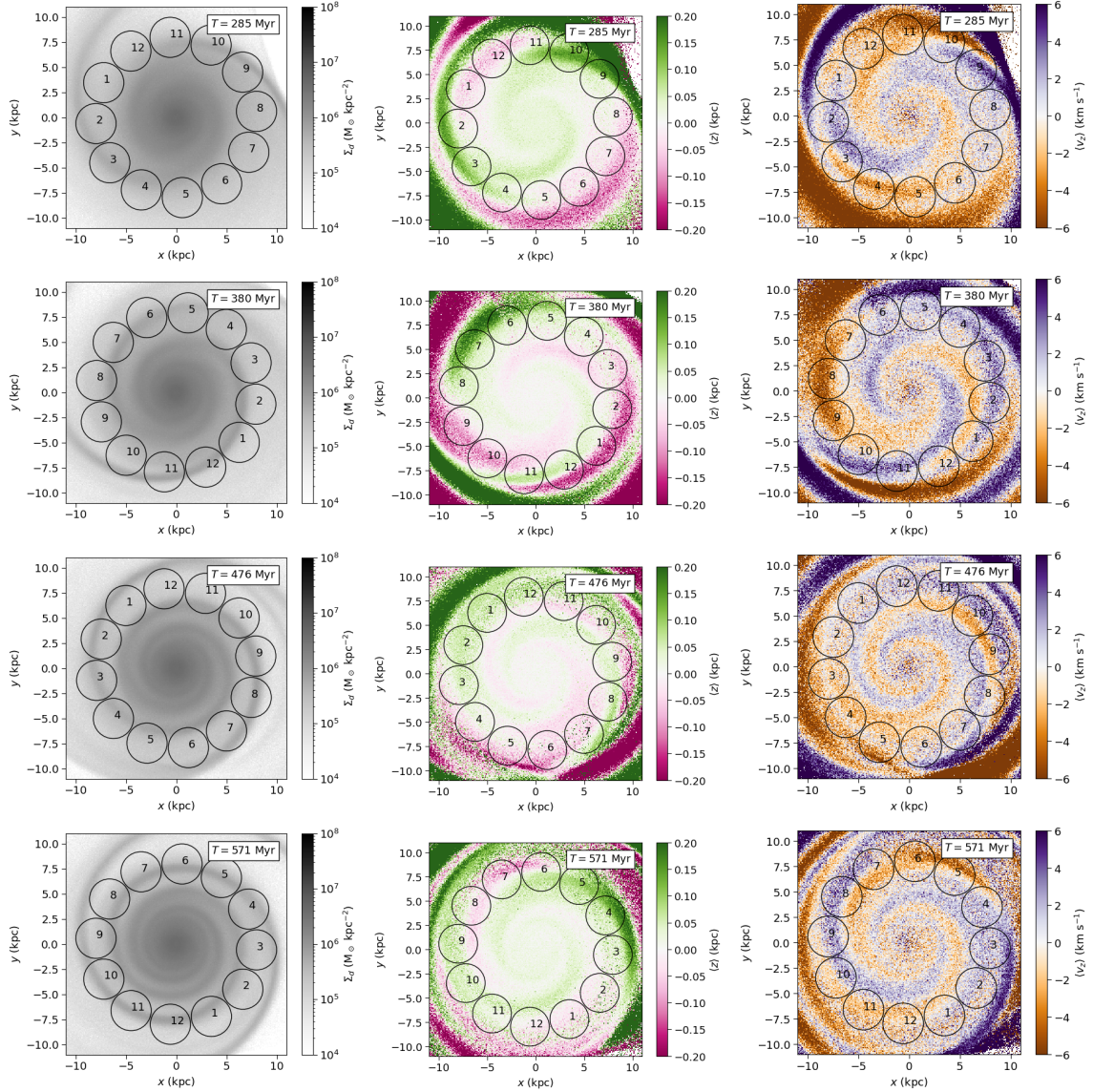
Figure 7. Early to middle stage evolution of the disc seen from the $z > 0$ axis; the time sequence continues in the next figure. The first panel is the projected stellar density field in the $x - y$ plane. The second and third panels are the average vertical distance $\langle z \rangle$ and the average vertical velocity $\langle V_z \rangle$ respectively in the $x - y$ plane. The first panel shows the $m = 2$ spiral density wave; the second and third panels reveal the $m = 2$ bending mode. The numbered circles are 4 kpc in diameter and spread along the Solar Circle ($R = 8.2$ kpc). These relate to our discussion of the phase spiral and show the rotation of the disc.

following

$$\phi_D(R, T) = \left(\frac{\Omega_D(R) + \Omega_o}{1 \text{ km s}^{-1} \text{ kpc}^{-1}} \right) \left(\frac{T}{978 \text{ Myr}} \right) + \phi_{D,o} \quad (2)$$

where

$$\Omega_D(R) = \Omega(R) - \frac{1}{2}\kappa(R) \quad (3)$$

and Ω_o is included in the fitting process to account for any pattern speed due to figure rotation. ($\phi_{D,o}$ is an arbitrary offset.) All frequencies are in units of km s^{-1} kpc^{-1} (equivalent to $(978 \text{ Myr})^{-1}$), T is in units of Myr, and angles are in radians.

If $\Omega_o = 0$, this form describes a "kinematic" density wave because it depends only on the kinematics of elliptic orbit crowding at aphelia. In this instance, the spiral pattern is not

a propagating wave mode, it has no group velocity, and the complicating effects of self-gravity and swing amplification are negligible. For a flat rotation curve, a kinematic density wave wraps up at about 30% of the angular rate of a material arm carried by the shearing disc ($\Omega_D = \Omega$).

Interestingly, this simple form is not exactly what is observed in simulations of disc mergers and perturbed discs to date (Struck et al. 2011). For example, the "grand design" spiral in M51 is a common target for modellers. Numerical simulations of the 2-arm spiral pattern are found to wind up at a higher rate than indicated by Eq. 2, with a non-linear dependence on radius (Oh et al. 2008; Dobbs et al. 2010; Salo & Laurikainen 2000).

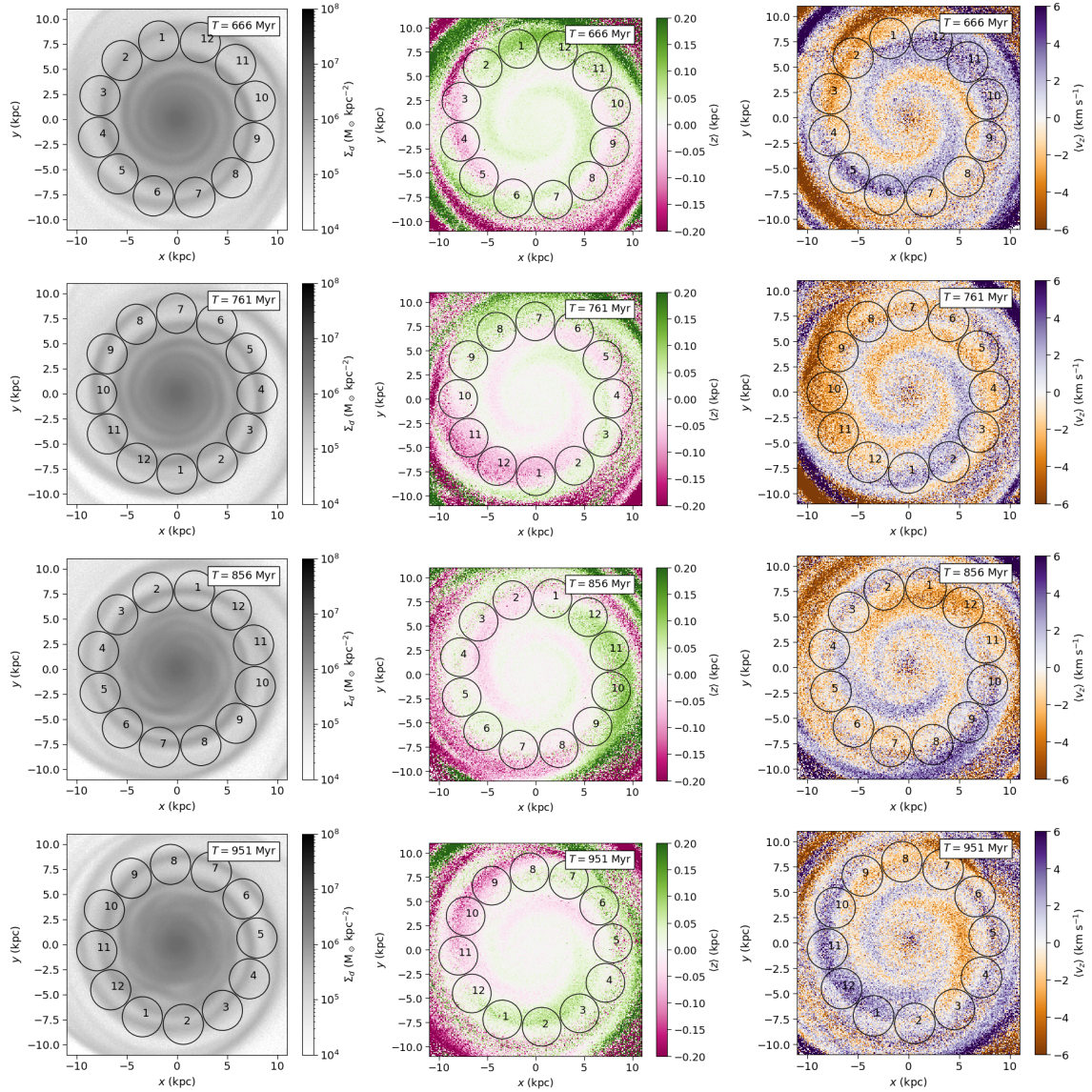So what do we find? In the last section, we saw how the density wave developed in the projected stellar density

Figure 7. Middle stage to late stage evolution of the disc seen from the $z > 0$ axis (continues from previous figure).

map (Fig. 7). For the wave to be kinematic, it must simply wind up with no additional contribution from, say, figure rotation with an associated pattern speed. In Fig. 8, the model described by Eq. 2 is overlaid on the density map where both the arm (blue: ϕ_D) and counter arm (red: $\phi_D + \pi$) are shown. The parameters $\Omega(R)$ and $\kappa(R)$ are measured from the isolated disc simulation and presented in Fig. 9 (left panel). After determining $\Omega = V_\phi/R$, we calculate

$$\kappa^2(R) = \left(\frac{\partial^2 \Phi_{\text{eff}}}{\partial R^2}\right)_{z=0} = \left(R\frac{d\Omega^2}{dR} + 4\Omega^2\right)_{z=0} \quad (4)$$

where $\Phi(R)_{\text{eff}}$ is the effective gravitational potential taken from the simulation, and the subscript $z = 0$ indicates that all measurements are made in the plane. In the early stages, at the time of the disc transit, the spiral pattern is highly asymmetric, but the pattern clearly settles down after $T = 190$ Myr and starts to exhibit 2-fold symmetry. After $T = 450$ Myr, the Lindblad-Kalnajs pattern (Fig. 8) reveals itself and

tracks the evolution of the spiral pattern for the remainder of the simulation.

Once the functional form of the pattern is fixed, there are no free parameters for any of the other time steps shown. But this statement requires two qualifications: (i) The time variable T has an arbitrary offset T' that determines the degree of winding when the spiral pattern first appears. An initial line of points is assumed to lie along $y > 0$ and $x = 0$; the points are then advanced in time to match the spiral pattern at some epoch (Binney & Tremaine 2008). For our model to match, we set $T' = 550$ Myr. (ii) Intriguingly, we find that the spiral pattern does have some figure rotation. We measure this to be $\pi/24$ rad every 100 Myr once the density wave has settled down from the initial impulse. Using the conversion factor above, the rigid pattern speed is equivalent to $\Omega_o = 1.3$ km s^{-1} kpc^{-1}. While Ω_o is very slow, it is non-zero and indicates that the spiral arms are propagating as a dynamical wave mode, rather than a kinematic anomaly (cf. Lin & Shu 1966).
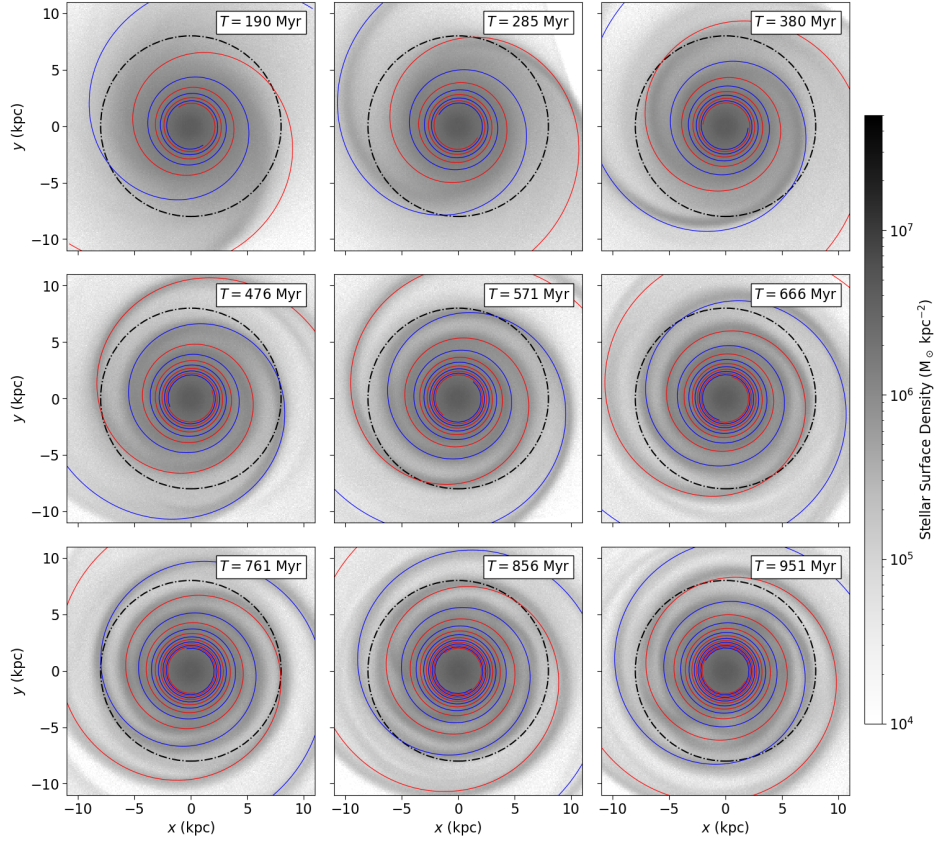
Figure 8. Middle and late stage evolution of the kinematic density wave in the N-body simulation. The rotating galaxy is viewed at each time step from an inertial (non-rotating) frame. The solar circle is indicated with the dot-dashed line. After the phase of the pattern is fixed at $t = 571$ Myr, there are no free parameters for the functional form of the spiral model in the other time steps. For 200 Myr after the impulse, the non-symmetric spiral pattern is evolving towards 2-fold symmetry. After $t = 300$ Myr, the bisymmetry is revealed and the correspondence is very good, although a vestigial non-bisymmetry exists at all time steps. The pattern is winding up as $\phi(R, t) \propto (\Omega(R) - \frac{1}{2}\kappa(R)) \, t$. This is the natural response of a cold disc perturbed by a strong impulse.
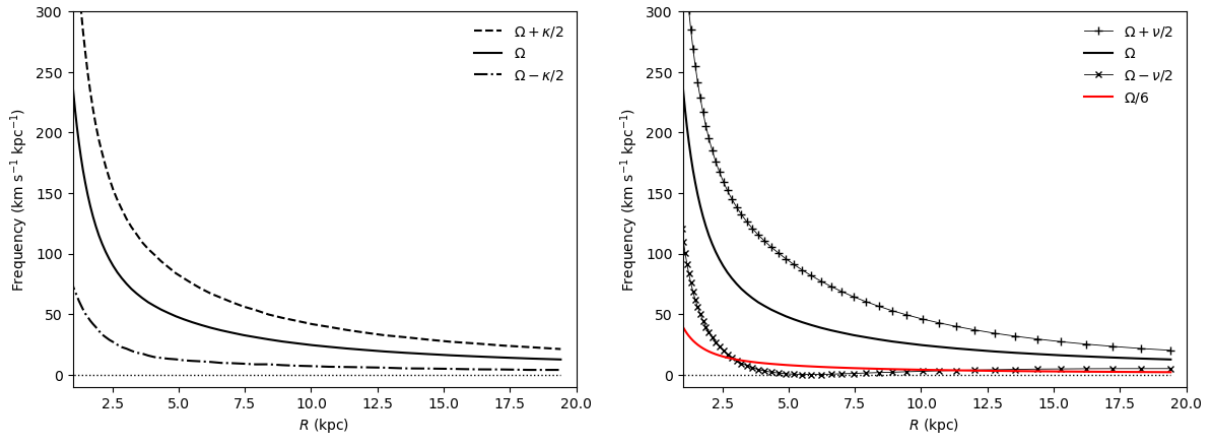
Figure 9. (Left) The angular frequency (Ω) and the planar Lindblad resonance ($\Omega - \kappa/2$) as a function of radius R measured from the isolated disc simulation in the mid-plane ($z = 0$). (Right) The angular frequency and vertical Lindblad resonance ($\Omega - \nu/2$) measured from the same simulation in the mid-plane ($z = 0$). Note that Ω (solid curve) is identical in both panels.

In summary, the evolution of the $x - y$ spiral has two components: (a) a wind-up rate given by the simple Lindblad-Kalnajs formula (eq. 3); (b) an additional slow figure rotation Ω_o that rotates the fixed pattern in a prograde (anticlockwise) direction as one moves forward in time. The combined effect is much slower than the wind-up rate of material arms.

As far as we can establish, the most basic form for the spiral evolution described in Eq. 2 has not strictly been observed before in simulations, certainly not over the long time frames considered here. This may reflect the high level of self-consistency made possible by the AGAMA initialisation of our cold disc simulations. Furthermore, it appears to be a *travelling* density wave – the 'Lindblad-Kalnajs' propagating wave mode.

5.2.2 Kinematic bending wave

In addition to in-plane spiral density waves, corrugated waves are now well established in the Galaxy (Sec. 1). How these work together, if at all, is entirely unclear. A bending wave is observed in our new simulation, and appears to wrap up like the density wave discussed in the last section. They appear to share the same common origin, but are these a superposition that propagate without interference from the other? Like spiral density waves, bending modes may be transient on a timescale of a few rotation periods at the Sun's radius (Hunter & Toomre 1969a; Sparke 1984; Binney et al. 1998).

The progress that has been made in understanding spiral density waves is restricted to 2D razor-thin discs. Just how this action extends to 3D discs is largely unknown (Fouvry et al. 2017). It is clear that the fundamentally in-plane mode must involve V_z in addition to V_ϕ because stars will be drawn down to regions of high density (Masset & Tagger 1997). These motions will remain conjectural until the theory of spiral structure has been extended from razor-thin discs, in which vertical motion is impossible, to discs of finite thickness, a very difficult proposition (Binney & Tremaine 2008).

The available formalism relating to the second kind of mode, the corrugation or bending wave, is even more primitive than the current theory of spiral structure because it involves neglecting epicyclic oscillations in addition to taking the disc to be razor thin (Hunter & Toomre 1969b). Hence we really have very little idea what a proper theory of corrugation waves would look like. It is plausible that the V_ϕ motions are coupled with V_z motions because warps (that can be considered as concentric rings) arise from torques exerted by one ring on another.

Both wave phenomena are defined by natural oscillation frequencies and are described by dispersion relations that look similar, although the dependence on the disc self gravity has the opposite sign (Binney & Tremaine 2008). The vertical frequency has a similar form to the epicyclic frequency in Eq. 4, viz.

$$\nu^2(R) = \left(\frac{\partial^2 \Phi_{\rm eff}}{\partial z^2}\right)_{z=0} \tag{5}$$

The vertical height of any bending mode with parameter m is described by (e.g. Binney & Tremaine 2008)

$$z_B(R, \phi, T) = z_o(R) \cos[m(\phi - \Omega_B T)] \cos \nu T \tag{6}$$

The kinematic bending modes are subject to winding up much like the density waves. The pattern speed of the bending wave is

$$\Omega_B = \Omega(R) \pm \nu(R)/m \tag{7}$$

We now investigate the bending mode in our simulation because of its role in the phase-spiral phenomenon discussed below. First, we measure $\nu(R)$ directly from the isolated disc simulation using Eq. 5; the results are presented in Fig. 9.

Our expectation is that the winding-up of the $m = 2$ bending model will lag behind the $m = 2$ spiral density wave. In a spherical gravitational potential with average density ρ_s, $\kappa = \nu$; for a flattened disc potential with density ρ_d, $\nu^2/\kappa^2 \approx (3/2)\,(\rho_d/\rho_s)$. For example, along the solar circle, $\nu/\kappa \approx 2$ (Binney & Tremaine 2008). Thus we expect $\Omega - \nu/2 < \Omega - \kappa/2$ and this is indeed what we find when comparing both panels in Fig. 9. Fig. 7 presents a comparison of the projected stellar density, the average vertical height $\langle z \rangle$ and the average vertical velocity $\langle V_z \rangle$ in the $x - y$ plane. The slow wrapping up of the bending wave through the oscillatory behaviour in $\langle z \rangle$ and $\langle V_z \rangle$ is clearly evident in the last two columns.

We now model these patterns using Eq. 6 and the measured properties of the disc (Fig. 9). The spiral patterns in $\langle z \rangle$ and $\langle V_z \rangle$ are out of phase by $\pi/2$ consistent with their oscillatory motion. We present our toy model for $\langle V_z \rangle$ in Fig. 10 to be compared to the last column in Fig. 7. The early time step is included to illustrate the important transition that was discussed in Sec. 5.1. The $m = 1$ integral sign warp is seen early on, but a stronger $m = 2$ mode develops across the inner disc by $T = 190$ Myr and dominates the later evolution. After this time, the outer disc behaviour is complex, but the inner disc settles down to a well organized $m = 2$ bending mode that wraps up slowly.

Our model for the $m = 2$ bending wave is not as trivial as the form used for the spiral density wave. The equivalent form would be $\Omega_B = \Omega(R) - \nu(R)/2$ but this is at odds with what we find. In Fig. 9, note that $\Omega(R) - \nu(R)/2 \approx 0$ in the range $4 \lesssim R \lesssim 6$ kpc. But the simulated disc continually winds up for all time over this radial range. Other values of $m > 2$ wrap up too rapidly. In its place, we use a slightly modified form, i.e.

$$\Omega_B = \max[\,\Omega(R) - \nu(R)/2,\ \Omega(R)/6\,] \tag{8}$$

such that the measured trend is used in the inner disc, but beyond $R > 3$ kpc, the angular frequency transitions to $\Omega_B = \Omega/6$ to ensure the 'drop out' near $R = 5$ kpc is never reached (red curve in Fig. 9). Our choice of this minimum threshold can be defended. The transient density wave model is defined by $\Omega_D \approx \Omega/3$; the bending mode must be slower and therefore we adopt $\Omega_B \approx \Omega_D/2$ after some experimentation. This approximation tracks the observed wrapping up of the bending mode quite well. Because of our assumption, we are not able to reliably detect any fixed pattern speed by analogy with the density wave, and therefore we do not fit for it.

In Fig. 10, the model for the evolving spiral density wave is superimposed on the bending wave in each time step. There is clearly a complex interaction between both wave modes as they wrap up at different rates. *We return to this behaviour below as it is central to our explanation for the phase spiral phenomenon.*
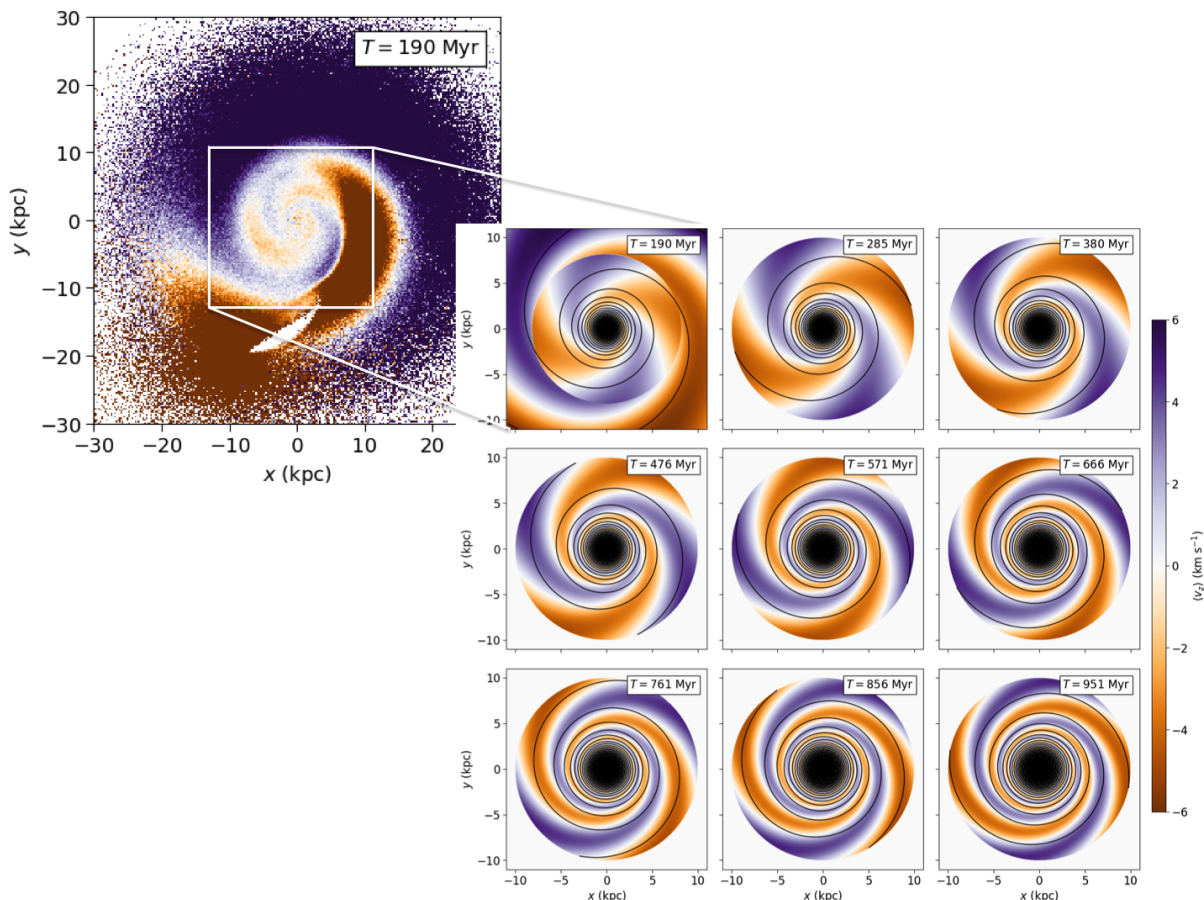
Figure 10. A comparison of the evolution of the bending wave (colour image) with the wrapping up of the kinematic density wave (spiral lines) for $R \leqslant 10$ kpc. The colour image represents the mean vertical velocity field $\langle v_z \rangle$ from −6 km s^{-1} (blue) to +6 km s^{-1}(red). The interplay of the two distinct wave modes wrapping up at different rates is the origin of the phase spiral, as we discuss. Both are model fits to the disc simulations. At the first time step ($T = 190$ Myr), we include a composite model that is to be compared to the inset. This shows how the $m = 1$ bending mode ($R > 8$ kpc) relates to the inner $m = 2$ bending mode ($R < 8$ kpc).

5.3 Phase spiral evolution

We now describe the origin and evolution of the phase spiral as observed in our impulse-driven simulation. In Fig. 11, the $z - V_z$ plane is presented for 12 spherical volumes (4 kpc diameter) spread along the solar circle (Fig. 7). The volumes are numbered 1 to 12 in an anti-clockwise direction, which serves to identify the column in Fig. 11. After averaging over the phase-space volume, we form maps of $\langle V_\phi \rangle$ and $\langle V_R \rangle$. The numbered volumes are much larger than the original discovery volume ($100 \times 100 \times 1000$ pc).

Each row is a time step separated by 47.5 Myr intervals for the entire duration of the simulation (951 Myr), i.e. 21 time steps where the reference frame is the Rotational Standard of Rest (RSR). The simulated solar neighbourhood (volume 1) is in column 1. Our focus on the solar circle is important for several reasons: (i) the local volume retains the clearest signal in the *Gaia* radial velocity data; (ii) to illustrate the phase spiral evolution, we remove the Galactic rotation that can only be carried out at one radius if we are to avoid complex resampling of the simulation.

The solar circle at the time of the impact is shown in Region A (Fig. 11). As already seen (Fig. 6), the early stage evolution reveals that the disc is highly disturbed up to at

least $T = 380$ Myr. At $T = 238$ Myr, we observe a thick, red outer band (Region B) from a 'feather' (in the language of L19), i.e. a part of the disc that separates and lifts away from the disc during the settling process in the outer disc. A fixed amplitude, bending wave propagates along the filament; this leads to an outer ring or annulus in phase space as illustrated in Fig. 12.

The phase spiral starts to emerge in both $\langle V_\phi \rangle$ and $\langle V_R \rangle$ simultaneously around $T = 476$ Myr. The phenomenon comes and goes at all radii from $R \gtrsim 3$ kpc to the outer disc. For the first time, we detect the phase spiral over the same extent in the $z - V_z$ plane and at the same intrinsic resolution as the discovery paper; earlier work (e.g. L19, B19) had intrinsically lower resolution and a z−range twice as large compared to observations. The inversion of the relative z and V_z extent as a function of radius is precisely as predicted in B19 (Fig. 20). But here the focus is on evolution along the solar circle, which is sufficient for our purposes.

In Fig. 11, diagonal bands from upper right to lower left are apparent in both mosaics. These track along the same line as the two sequences of yellow (left) and black (right) boxes. The simulated solar neighbourhood (volume 1) lies along the same radius vector as the impact site at $T = 95$ Myr. But this site increasingly lags behind volume
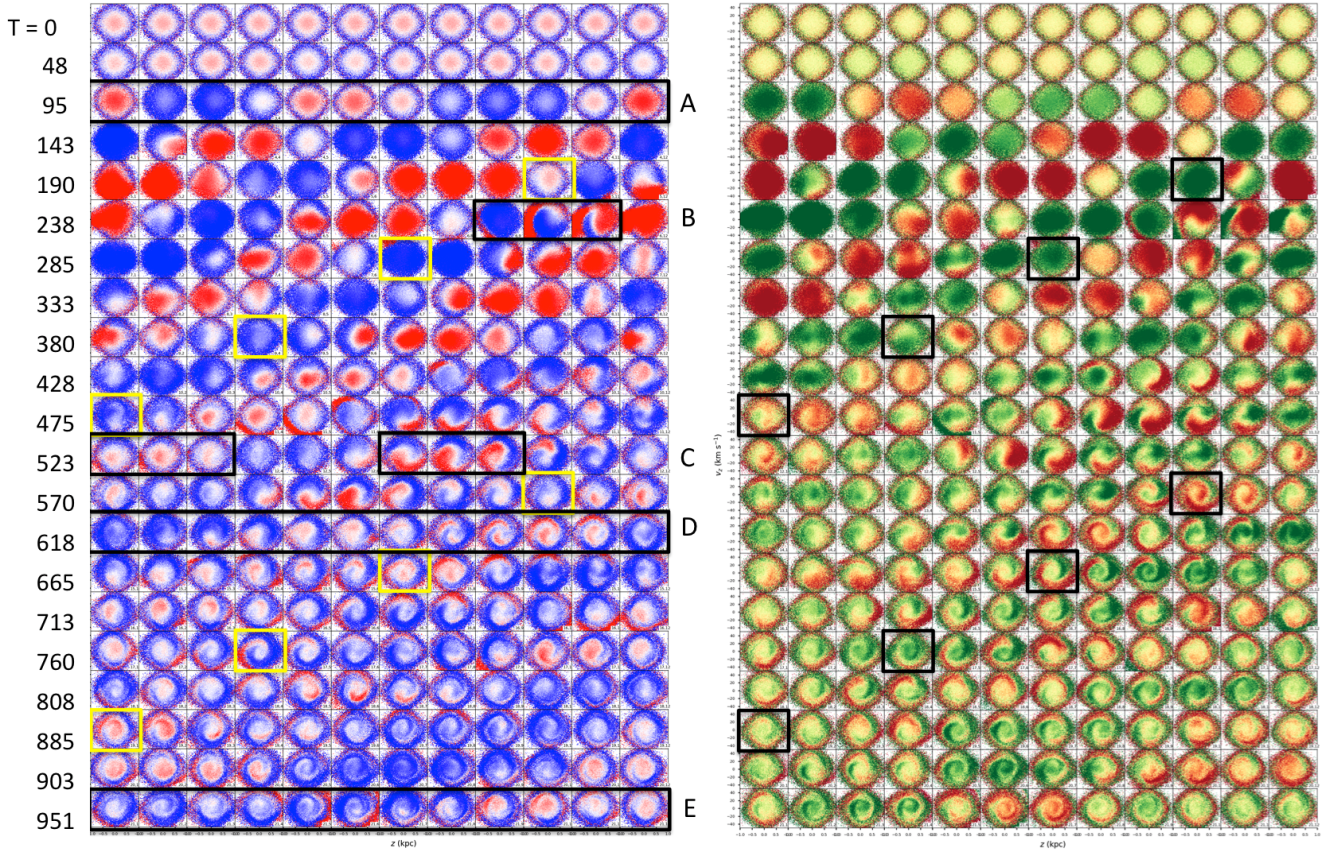
Figure 11. Along each row, the *i*th volume in Fig. 7 corresponds to the *i*th column. Each column corresponds to a fixed volume moving with the rotation: the solar neighbourhood is in column 1; its antipode is in column 7. Each row corresponds to a time step separated by 47.5 Myr as indicated. Moving to the right along a row corresponds to moving counter-clockwise by 30° per column along the solar circle, and the figure wraps around from the last to the first column on the left. The colour coding indicates the mean value of $\langle V_\Phi \rangle$ (left) and $\langle V_R \rangle$ (right); the ranges are (230, 250) km s^{-1} and (-10,+10) km s^{-1} respectively. Regions A-E are discussed in the paper. The yellow (left) and black (right) boxes indicate the volumes that align with the original impact site, which lags further behind the solar neighbourhood as time passes. Higher resolution images are available at the website (footnote 5).

1 (period $T_\odot = 210$ Myr) as the disc rotates, but realigns at a later time. The boxes indicate which volumes at specific timesteps are precisely aligned with the radius vector to the original impact site.

The period at the impactor radius ($R = 18$ kpc) is about $T_C = 460$ Myr, such that volume 1 realigns with the impact site roughly every $T_{\odot,C} = (T_\odot T_C)/(T_C - T_\odot) \approx 380$ Myr. Interestingly, this is precisely the time delay for a phase spiral to emerge along the solar circle. The continued interaction of the solar neighbourhood with the region of the disc around the impact site is a symptom of the wrapping up of the bending mode and the density wave.

In Fig. 11, we observe that the phase spiral first emerges at $T = 476$ Myr. In Region C at $T = 523$ Myr, we see the phase spiral clearly in the antipode region (cols. 7-9), something that is not seen diametrically opposite in the solar neighbourhood (cols. 1-3). The subsequent evolution tracks along the diagonal bands, and so is tied to the original impact site. Even though the density wave and bending mode appear bisymmetric by nature, the impulsive event is intrinsically one-sided.

Fig. 13 shows how difficult it is to associate the phase spiral with the observed properties of the disc in the $x - y$ plane at any given epoch. The numbered $z - V_z$ diagrams

around the outside are taken from row D (Fig. 11). The $x - y$ plane at $T = 618$ Myr looks highly bisymmetric in projected stellar density, and in $\langle z \rangle$ and $\langle V_z \rangle$. But the phase spiral becomes stronger in both $\langle V_R \rangle$ and $\langle V_\phi \rangle$ as one moves around the clock (volumes 1 to 11).

There is an inherent *asymmetry* in the strength of the disturbance, as we saw in Sec. 5.1, in addition to the underlying $m = 1$ mode. The phase spiral in each volume, to borrow from R.P. Feynman, is a sum over history. The disc crossing occurred in the distant past but the transit region (albeit stretched by the disc's shear) lies immediately outside of volumes 8-9 at this time (cf. Fig. 11). *The imprint of the one-sided disturbance is preserved for at least 500 Myr!*

At later times, with respect to the solar circle, the phase spiral develops both in the upstream (against rotation) and downstream direction (with rotation) due to the disc's shear. Once triggered, an increasing number of panels in Fig. 11 show the phase spiral with the passage of time, and its fidelity also improves as more wraps emerge. By the end of the simulation, the phase spiral has taken hold of the entire solar circle (Row E). In this respect, we can use the ubiquity of the phase spiral to age-date the impulsive event, and even to locate the earlier crossing point in the outer disc. These
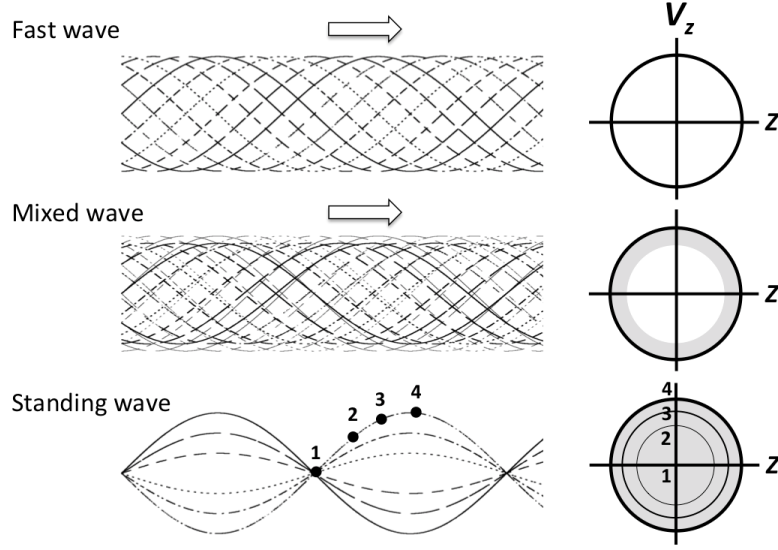
Figure 12. Examples of how the $z - V_z$ plane can be filled by three distinct bending wave modes: a fast wave with constant amplitude, a composite wave with mixed amplitudes, and a standing wave. To the right, the idealized $z - V_z$ phase plane is shown. The set of wave modes operating within a volume determine how the $z - V_z$ plane is filled.
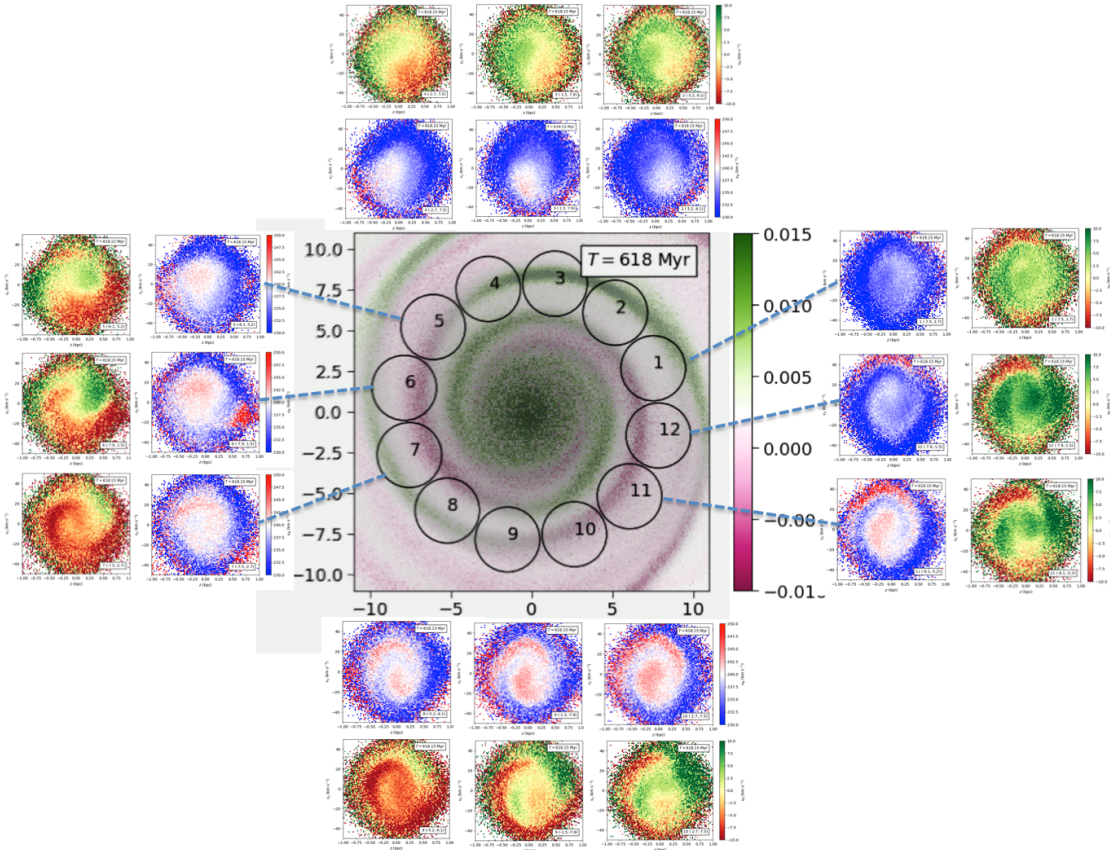

Figure 13. A density-weighted $\langle z \rangle$ map at time step $T = 618$ Myr. In the central 11×11 kpc panel, the colour coding shows $\langle z \rangle$, the mean height (kpc) of the stars above and below the plane. The vertical undulations arising from the bending mode are clearly seen. Individual spiral arms roll up and down as the bending mode moves underneath them. The pink-coloured inner arms passing through volumes 7 and 12 curve inwards avoiding the next volume; note how the bending mode under both inner arms crosses over to the counter arm. The outer collages for $\langle V_\phi \rangle(z, V_z)$ and $\langle V_R \rangle(z, V_z)$ are taken from Fig. 11; the ranges are (230, 250) km s^{-1} and (-10,+10) km s^{-1} respectively.
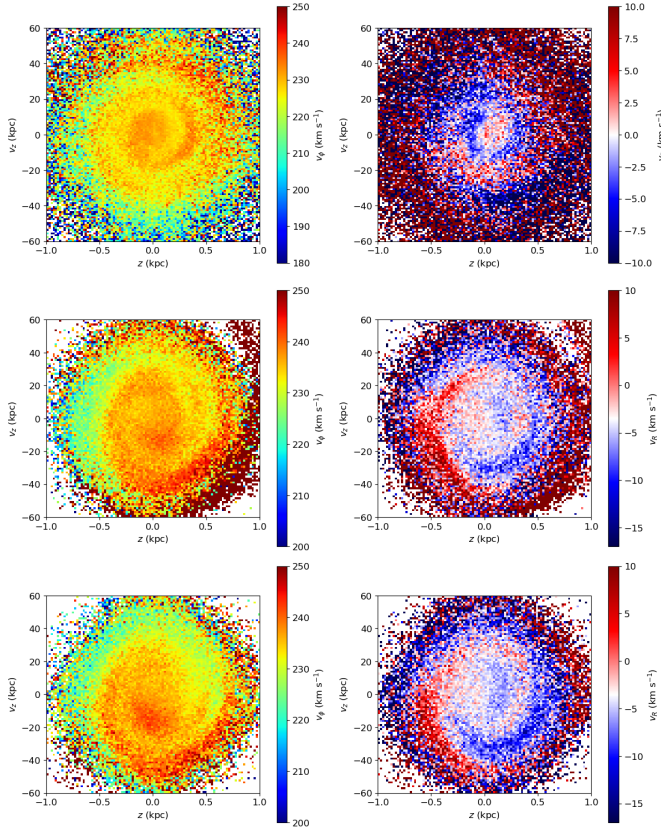
Figure 14. A comparison of the $z - V_z$ plane presented in the discovery paper (Antoja et al. 2018, top) with volumes 5 and 6 in our last time frame, $T = 951$ Myr (middle & bottom); we have preserved the colour schemes from the earlier paper. Our volumes are much larger than the *Gaia* volume (see text) but we match the extent in the $z - V_z$ and the range in values for the first time; the fine structure of the phase spiral is evident in both also. The simulation produces fewer wraps in $\langle V_\phi \rangle$ (left) than in $\langle V_R \rangle$ (right), which may reflect the need for more disc particles and/or longer timespans in future simulations.

are themes that we anticipate will be addressed in future papers on Galactic seismology.

6 DISCUSSION

The *Gaia* discovery of the phase spiral was entirely unanticipated (Antoja et al. 2018). As described in Sec. 4.1, the most natural interpretation is incomplete phase mixing after a major disturbance of the disc. But the remarkable signature raises important questions: What is the precise mechanism that explains it? Is the signature fragile or robust? How widespread is it and can we learn from it?

On the balance of evidence presented, Galactic seismology, a subset of Galactic archaeology, has a rich future. But our work raises more questions than answers. In the closing comments of the last section, we make the case for follow-up studies. Within the broad remit of Galactic seismology, an important goal is to age-date the origin of the phase spiral. Using two distinct mechanisms, B19 and Khoperskov et al. (2019) show that once the corrugation is excited, it can last at least 2 Gyr. But repeated disc crossings from a massive

perturber may wipe out the phase coherence imposed by the previous crossing (L19, B19).

Over a timeline as long as 1 Gyr, the phase-spiral signal appears to be quite fragile, coming into view in one time step (e.g. Fig. 11) and then either evolving *in situ*, migrating in azimuth (co-moving RSR frame), or fading altogether in the next time step. Although no panel in Fig. 11 shows a close match in both $\langle V_\phi \rangle$ and $\langle V_R \rangle$ simultaneously, we find features that are ubiquitous and reminiscent of the discovery paper (see Fig. 14). Note here that the simulated panels show weak evidence of a second phase spiral emerging out of sync with the dominant one. This can arise from having two spiral arms moving through the same volume, which becomes increasingly likely as the density wave becomes highly wrapped at late times. There is a case for repeating the simulations with more particles allowing for smaller sample volumes.

The simulated phase spiral encoded by $\langle V_R \rangle$ is acceptable, but the phase spiral encoded by $\langle V_\phi \rangle$ shows up to one less wrap than observed. Khanna et al. (2019) found that the *Gaia* phase spiral is particularly prominent and well behaved in $\langle L_z \rangle$, but we were unable to see any improvement in our simulated signal when encoded in the same way. Given the central role of orbital angular momentum in generating the phase spiral (Sec. 2), it is not clear why this should be. While our simulations achieve the recommended particle count of $N \sim 10^8$ (Binney & Schönrich 2018), our test simulations with smaller intrinsic spatial resolution (Sec. 4) did not alter the outcome significantly. But more particles would allow us to better match the *Gaia* RVS volume.

There is one important consequence of our "rollercoaster" model that has been glossed over. The interaction of density waves and bending modes implies that the younger stellar populations are those that are most likely to be caught up in the phase-spiral action. It is probable that older stars also contribute to the signal, but the feature is likely to be more enhanced for the younger stars. While there is existing evidence that this is true (B19, Li & Shen 2020), the dominant population of the phase spiral needs to be addressed more carefully (e.g. better stellar ages) in future studies (Sharma et al. 2016; Miglio et al. 2020).

This brings us full circle to the original prediction of the phase spiral (with two arms rather than the singular arm observed by *Gaia*) arising from dissolving star clusters (Candlish et al. 2014) that predated the *Gaia* era. Spiral arms in the Milky Way can be identified from star-forming regions (e.g. Lépine et al. 2001), as they can be in external galaxies, but this is no guarantee that young star clusters ($\lesssim 100$ Myr) are entirely restricted to the spiral arms. An inventory of almost 2000 star clusters confirmed by *Gaia* (Cantat-Gaudin et al. 2020, Fig. 8) shows that star clusters of all ages inhabit our neighbourhood within a few kiloparsecs, with the youngest star clusters confined to broad bands rather than ridge lines. Most clusters that form appear to dissolve within 100 Myr (Krumholz et al. 2019; Adamo et al. 2020) such that cluster dissolution can conceivably contribute to the innermost wraps of the phase spiral (cf. Li & Shen 2020). Cluster dissolution is unlikely account for the signature *in toto*, not least because the models to date predict a 2-arm phase spiral (Candlish et al. 2014), and that has not been observed so far.

Intriguingly, Michtchenko et al. (2019) identify parts of the local phase spiral as arising from several moving groups,

including Pleaides, Hyades, Sirius (Ursa Major) and Coma Berenices. These are four of the closest clusters, all of which appear to be unbound. What is particularly striking is that their ages range from 100 to 600 Myr (Famaey et al. 2008) with the youth carrying the stronger signal, in agreement with B19. More work is needed to assess the detailed structure of the local corrugation (e.g. Friske & Schönrich 2019; Beane et al. 2019) and its relationship to the local spiral arms (e.g. Miyachi et al. 2019; Griv et al. 2020). How these relate to the local young clusters and star-forming regions also becomes important (Quillen et al. 2020; Cantat-Gaudin et al. 2020).

It is likely that our cold-disc simulations, while very carefully constructed, are too simplistic. In the next phase, our simulations introduce a central bar, giant molecular clouds, and allow for internally-driven density waves, with the overarching goal of producing a more realistic framework for future study. This work will inform the nature of future studies, leading to better targetted surveys rather than all-sky surveys to test different non-equilibrium dynamical models.

To date, we have not considered the contribution of the cool gas. Another prediction from B19 (Sec. 8.2) is that the corrugation may be evident in the Galactic H I or the molecular hydrogen distribution, and may even explain Gould's belt and other local phenomena. Connecting this information with star clusters at different stages of their evolution may also provide important insights (e.g. Quillen et al. 2020).

In the spirit of Widrow et al. (2012), there is a case for extending galactoseismology to the study of nearby galaxies. Recent arguments have been made for undulations in the Galactic rotation curve (Martinez-Medina et al. 2019) being correlated with the spectacular $R - V_\phi$ ridges seen by *Gaia* (Antoja et al. 2018; Khanna et al. 2019). With careful disc modelling and subtraction, it may be possible to identify the signatures of bending modes in substantial numbers of nearby disc galaxies (cf. Matthews & Uson 2008). In such instances, it will be important to establish whether the bending mode wraps up with the disc rotation, as we show here. Morphologically, the wrapping of the spiral density wave is more easily determined and thus it may even be possible to establish whether the interaction of these two wave modes is a regular phenomenon.

REFERENCES

Adamo A., et al., 2020, Space Sci. Rev., 216, 69
Aguilar L. A., White S. D. M., 1985, ApJ, 295, 374
Antoja T., et al., 2018, Nature, 561, 360
Beane A., et al., 2019, ApJ, 883, 103
Binney J., Schönrich R., 2018, MNRAS, 481, 1501
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
Binney J., Jiang I.-G., Dutta S., 1998, MNRAS, 297, 1237
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bland-Hawthorn J., et al., 2019, MNRAS, 486, 1167
Candlish G. N., Smith R., Fellhauer M., Gibson B. K., Kroupa P., Assmann P., 2014, MNRAS, 437, 3702
Cantat-Gaudin T., et al., 2020, A&A, 640, A1
Carlin J. L., et al., 2013, ApJ, 777, L5
D'Onghia E., Vogelsberger M., Hernquist L., 2013, ApJ, 766, 34
D'Onghia E., Madau P., Vera-Ciro C., Quillen A., Hernquist L., 2016, ApJ, 823, 4

Darling K., Widrow L. M., 2019a, MNRAS, 484, 1050
Darling K., Widrow L. M., 2019b, MNRAS, 490, 114
Darling K., Widrow L. M., 2019c, MNRAS, 490, 114
De Silva G. M., et al., 2015, MNRAS, 449, 2604
Deng L.-C., et al., 2012, Research in Astronomy and Astrophysics, 12, 735
Dobbs C., Baba J., 2014, Publ. Astron. Soc. Australia, 31, e035
Dobbs C. L., Theis C., Pringle J. E., Bate M. R., 2010, MNRAS, 403, 625
Famaey B., Siebert A., Jorissen A., 2008, A&A, 483, 453
Fouvry J.-B., Pichon C., Chavanis P.-H., Monk L., 2017, MNRAS, 471, 2642
Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487
Friske J. K. S., Schönrich R., 2019, MNRAS, 490, 5414
Fujii M. S., Bédorf J., Baba J., Portegies Zwart S., 2018, MNRAS, 477, 1451
Gaia Collaboration et al., 2018, A&A, 616, A1
Gilmore G., Reid N., 1983, MNRAS, 202, 1025
Gilmore G., et al., 2012, The Messenger, 147, 25
Griv E., Gedalin M., Shih I. C., Hou L.-G., Jiang I.-G., 2020, MNRAS, 493, 2111
Hernquist L., 1990, ApJ, 356, 359
Holmberg E., 1941, ApJ, 94, 385
Hui L., Ostriker J. P., Tremaine S., Witten E., 2017, Phys. Rev. D, 95, 043541
Hunt J. A. S., Bub M. W., Bovy J., Mackereth J. T., Trick W. H., Kawata D., 2019, MNRAS, 490, 1026
Hunter J. D., 2007, Computing in Science & Engineering, 9, 90
Hunter C., Toomre A., 1969a, ApJ, 155, 747
Hunter C., Toomre A., 1969b, ApJ, 155, 747
Kalnajs A. J., 1973, Proceedings of the Astronomical Society of Australia, 2, 174
Khanna S., et al., 2019, MNRAS, 489, 4962
Khoperskov S., Di Matteo P., Gerhard O., Katz D., Haywood M., Combes F., Berczik P., Gomez A., 2019, A&A, 622, L6
Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227
Laporte C. F. P., Minchev I., Johnston K. V., Gómez F. A., 2019, MNRAS, 485, 3134
Law D. R., Johnston K. V., Majewski S. R., 2005, ApJ, 619, 807
Lépine J. R. D., Mishurov Y. N., Dedikov S. Y., 2001, ApJ, 546, 234
Li Z.-Y., Shen J., 2020, ApJ, 890, 85
Lin C. C., Shu F. H., 1966, Proceedings of the National Academy of Science, 55, 229
Lindblad B., 1956, Stockholms Observatoriums Annaler, 7, 7
Lindblad P. O., 1960, Stockholms Observatoriums Annaler, 4, 4
Lynden-Bell D., 1967, MNRAS, 136, 101
Mackereth J. T., et al., 2019, MNRAS, 489, 176
Majewski S. R., et al., 2017, AJ, 154, 94
Martinez-Medina L., Pichardo B., Peimbert A., Valenzuela O., 2019, MNRAS, 485, L104
Masset F., Tagger M., 1997, A&A, 322, 442
Matthews L. D., Uson J. M., 2008, AJ, 135, 291
Michtchenko T. A., Barros D. A., Pérez-Villegas A., Lépine J. R. D., 2019, ApJ, 876, 36
Miglio A., et al., 2020, arXiv e-prints, p. arXiv:2004.14806
Milgrom M., Sanders R. H., 2008, ApJ, 678, 131
Miyachi Y., Sakai N., Kawata D., Baba J., Honma M., Matsunaga N., Fujisawa K., 2019, ApJ, 882, 48
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nichols M., Mirabal N., Agertz O., Lockman F. J., Bland-Hawthorn J., 2014, MNRAS, 442, 2883
Oh S. H., Kim W.-T., Lee H. M., Kim J., 2008, ApJ, 683, 94
Perryman M. A. C., et al., 2001, A&A, 369, 339
Pontzen A., Roškar R., Stinson G. S., Woods R., Reed D. M., Coles J., Quinn T. R., 2013, pynbody: Astrophysics Simulation Analysis for Python

Prusti T., et al., 2016, A&A, 595, A1

Quillen A. C., Pettitt A. R., Chakrabarti S., Zhang Y., Gagné J., Minchev I., 2020, arXiv e-prints, p. arXiv:2006.01723

Salo H., Laurikainen E., 2000, MNRAS, 319, 393

Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, ApJ, 822, 15

Slater C. T., et al., 2014, ApJ, 791, 9

Soubiran C., et al., 2018, A&A, 616, A7

Sparke L. S., 1984, ApJ, 280, 117

Steinmetz M., et al., 2006, AJ, 132, 1645

Struck C., Dobbs C. L., Hwang J.-S., 2011, MNRAS, 414, 2498

Sundelius B., Thomasson M., Valtonen M. J., Byrd G. G., 1987, A&A, 174, 67

Tepper-García T., Bland-Hawthorn J., 2018, MNRAS, 478, 5263

Tepper-García T., Bland-Hawthorn J., Li D., 2020, MNRAS, 493, 5636

Teyssier R., 2002, A&A, 385, 337

Toomre A., 1964, ApJ, 139, 1217

Vasiliev E., 2019, MNRAS, 482, 1525

Widrow L. M., Dubinski J., 2005, ApJ, 631, 838

Widrow L. M., Gardner S., Yanny B., Dodelson S., Chen H.-Y., 2012, ApJ, 750, L41

Williams M. E. K., et al., 2013, MNRAS, 436, 101

Xu Y., Newberg H. J., Carlin J. L., Liu C., Deng L., Li J., Schönrich R., Yanny B., 2015, ApJ, 801, 105

Yanny B., Gardner S., 2013, ApJ, 777, 91

APPENDIX A: IMPULSE APPROXIMATION

Our set-up ensures that we are working entirely within the impulse approximation. In Fig. A1, we show the force experienced by a star in the immediate vicinity of Sgr's crossing point ($R \approx 18$ kpc) as a function of time before and after Sgr's disc transit (solid curve). The top x-axis displays Sgr's vertical distance to the disc plane along its orbit. The gray-shaded area indicates the width of the stellar disc, here given by twice its scale height ($2z_t \approx 500$ pc). The interaction between stars at the crossing point is clearly impulsive. The apparent asymmetry in the impulse is the result of the reduced mass after transit (cf. Sec. 4.1). The impulsive force
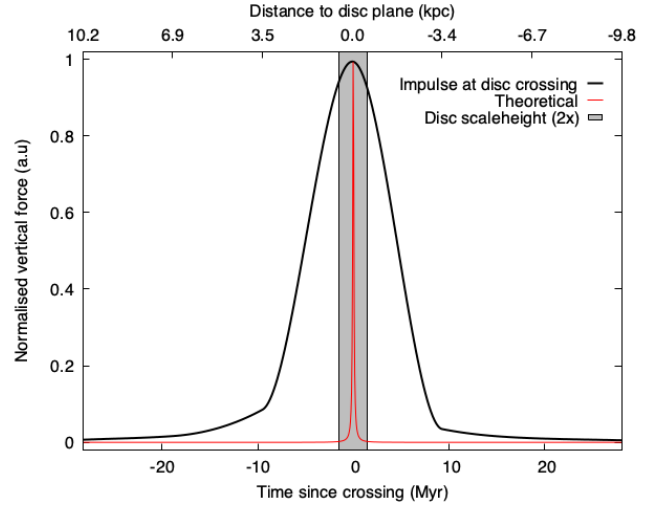

Figure A1. The force experienced by a star close to Sgr's crossing point ($R \approx 18$ kpc) as a function of time (solid curve). This is taken from the N-body simulation data dumped to storage roughly every 10 Myr. Since the time step in the simulation is orders of magnitude smaller, the impulsive response to the intruder is vastly narrower in the simulation. This is illustrated by the red curve, which indicates the theoretical result. The small asymmetry imposed by time-dependent intruder mass is clear. Note that the curves have been normalised to a peak height of unity. The shaded area indicates the width of the stellar disc defined as twice its scale height ($2z_t \approx 500$ pc).

in the simulation is much narrower in practice, but appears broadened due to the coarse time resolution ($\Delta t \approx 10$ Myr) adopted for storing the timesteps. To emphasize this point, we include in the figure the theoretical result (red curve). The latter corresponds to the force experienced by a body in the vicinity of a point mass, softened by ϵ_s, i.e. $\sim r/(r^2 + \epsilon_s^2)^{3/2}$ where r is the radial separation.