Self-consistent modelling of the Milky Way structure using live potentials

Eva Durán-Camacho [ORCID],[1]★ Ana Duarte-Cabral [ORCID],[1] Alex R. Pettitt [ORCID],[2] Robin G. Treß,[3] Paul C. Clark [ORCID],[1] Ralf S. Klessen [ORCID],[4,5] Kamran R. J. Bogue,[6] Rowan J. Smith [ORCID][6,7] and Mattia C. Sormani[4]

[1]*School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK*
[2]*Department of Physics and Astronomy, California State University, Sacramento, 6000 J Street, Sacramento, CA 95819-6041, USA*
[3]*Institute of Physics, Laboratory for Galaxy Evolution and Spectral Modelling, EPFL, Observatoire de Sauverny, Chemin Pegais 51, CH-1290 Versoix, Switzerland*
[4]*Zentrum fur Astronomie, Institut für Theoretische Astrophysik, Universität Heidelberg, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany*
[5]*Interdisziplinäres Zentrum fur Wissenschaftliches Rechnen, Universität Heidelberg, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany*
[6]*School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, KY16 9SS, UK*
[7]*Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

ABSTRACT

To advance our understanding of the evolution of the interstellar medium (ISM) of our Galaxy, numerical models of Milky Way (MW) type galaxies are widely used. However, most models only vaguely resemble the MW (e.g. in total mass), and often use imposed analytic potentials (which cannot evolve dynamically). This poses a problem in asserting their applicability for the interpretation of observations of our own Galaxy. The goal of this work is to identify a numerical model that is not only an MW-type galaxy, but one that can mimic some of the main observed structures of our Galaxy, using dynamically evolving potentials, so that it can be used as a base model to study the ISM cycle in a galaxy like our own. This paper introduces a suite of 15 MW-type galaxy models developed using the AREPO numerical code, that are compared to Galactic observations of ^{12}CO and H I emission via longitude–velocity plots, from where we extract and compare the skeletons of major galactic features and the terminal gas velocities. We found that our best-fitting model to the overall structure, also reproduces some of the more specific observed features of the MW, including a bar with a pattern speed of $30.0 \pm 0.2 \ \mathrm{km \, s^{-1} \, kpc^{-1}}$, and a bar half-length of $3.2 \pm 0.8 \ \mathrm{kpc}$. Our model shows large streaming motions around spiral arms, and strong radial motions well beyond the inner bar. This model highlights the complex motions of a dynamic MW-type galaxy and has the potential to offer valuable insight into how our Galaxy regulates the ISM and star formation.

Key words: methods: numerical – ISM: structure – Galaxy: structure – galaxies: spiral.

1 INTRODUCTION

Star formation (SF) has historically been assumed to be universal, mainly due to the existence of a number of observational characteristics linked to the SF process that seem to be invariant. One such universal aspect is the Kennicutt–Schmidt law – the direct relation between the gas surface density and the surface density of the star formation rates (SFR) that holds for several orders of magnitude and many different types of galaxies (e.g. Schmidt 1959; Kennicutt 1989), suggesting that the rate of SF depends only on the amount of (dense) gas available, regardless of the properties of that dense gas itself. However, when resolving individual regions within a galaxy, such relations become less tight, with large intrinsic scatter appearing (e.g. Leroy et al. 2013; Pessa et al. 2022). The question then becomes what drives the scatter within a galaxy – is it the chaotic nature of the interstellar medium (ISM), or is some of that scatter controlled by the larger scale dynamics of the host galaxies?

While in nearby galaxies, resolving the SF process within clouds is still at the limit of observational capabilities, studies of SF in the Milky Way (MW) have been able to reveal the exquisite detail of star-forming regions (e.g. Arzoumanian et al. 2011), whose properties can be linked to the SF history of the clouds (Schneider et al. 2010; Duarte-Cabral et al. 2011; Nguyen Luong et al. 2011; Longmore et al. 2013; Peretto et al. 2013). However, linking how those variations in cloud properties (and subsequent SF) relate to the potential effect of the larger scale dynamics of the Galaxy is complicated by our inability to unequivocally associate clouds with their environment, due to the fact that we are located within the galactic disc and only can access an edge-on perspective.

Our understanding of the MW's structure remains uncertain, largely due to the challenges in creating a 3D map of the Galaxy (Elmegreen 1985). Nevertheless, there is consensus that the MW is a spiral-armed disc (Oort, Kerr & Westerhout 1958; Kerr 1962; Georgelin & Georgelin 1976; Dame et al. 1986; Kolpak et al. 2003; Xu et al. 2016), although the precise morphology of these arms remains debated (as highlighted in e.g. Xu, Hou & Wu 2018). Initial studies proposed a two-arm structure (Weaver 1970; Francis & Anderson 2012), but the limited evidence for stellar spiral arms,

★ E-mail: durancamachoe@cardiff.ac.uk

as indicated by data such as GLIMPSE (Galactic Legacy Infrared Midplane Survey Extraordinaire, Benjamin et al. 2005), has led to alternative models proposing a 2 + 2 spiral structure (e.g. Vallée 2008; Efremov 2011; Reid et al. 2019). A leading theory to reconcile these models suggests a dual approach: a four-arm structure predominantly of gas, alongside a two-arm structure composed mainly of stars (Drimmel 2000; Martos et al. 2004; Steiman-Cameron, Wolfire & Hollenbach 2010).

Modern observational efforts have brought exciting new data that allow the means to test galactic models, among which the data from the *Gaia* mission stands out (e.g Gaia Collaboration 2016). Gaia Collaboration (2018) provide a first look into the potential of the mission using data from the *Gaia* Data Release 2 (DR2), allowing for the creation of 3D velocity and velocity dispersion maps with unprecedented accuracy and spatial resolution. Xu et al. (2021) expand the results for the spiral arms obtained from the Very Long Baseline Interferometry maser data using the *Gaia* DR3. They find new OB-type stars and based on their clumped distribution they reach the conclusion that the Galaxy spiral structure is irregular. Therefore, the overall structure of the Galaxy remains uncertain as even more new features are being found with the help of these new DRs.

Recent data from the *Gaia* mission have significantly advanced our understanding of Galactic structures (e.g. Gaia Collaboration 2016). *Gaia* DR2 provided detailed 3D velocity maps, significantly enhancing spatial resolution and accuracy (Gaia Collaboration 2018). Further insights from *Gaia* DR3 continued to refine our understanding of the Galaxy's asymmetrical disc and spiral arms, suggesting an irregular and transient nature of the MW's spiral structure (Castro-Ginard et al. 2021; Xu et al. 2021), with complex dynamical patterns (Gaia Collaboration 2023; Kawata 2023; Khalil et al. 2023). These ongoing discoveries from each *Gaia* DR contribute to an ever-evolving picture of the Galaxy, revealing new complexities and challenging existing models.

The evaluation of the kinematics of the gas has been realized via longitude–velocity (*l-v*) maps using different tracers. The key tracers used to track the galactic morphology are H I (e.g. Kulkarni, Heiles & Blitz 1982; Kalberla & Kerp 2009) and ^{12}CO emission (e.g. Dame et al. 1986; Dame, Hartmann & Thaddeus 2001; Sawada et al. 2001), which tend to indicate the presence of four main spiral arms. In the advent of higher resolution surveys of the galactic plane in optically thinner molecular lines (e.g. Rigby et al. 2016; Colombo et al. 2021; Schuller et al. 2021), we begin to resolve and discern the gas emission with greater detail, but the spiral structure as seen in the gas, appears to be more flocculent in nature than proposed by earlier works (e.g. Duarte-Cabral et al. 2021; Colombo et al. 2022).

The existence of a Galactic bar towards the Galactic Centre is well supported by numerous observations, initially deduced via gas velocities (Cohen & Few 1976) and subsequently confirmed through photometry (Blitz & Spergel 1991) and star counts (Weinberg 1992). Peters (1975) proposed an elliptical bar model based on the H I velocities. Findings from *COBE* (Weiland et al. 1994) provided compelling evidence for the photometric asymmetry generated by the bar, and was subsequently followed by studies like Stanek et al. (1994), who first demonstrated a distance asymmetry with star counts. The recent study from Gaia Collaboration (2023) uses the latest DR3 to reveal a wealth of kinematic information, with prospects to enhance our knowledge of non-axisymmetric structures in the Galaxy.

Earlier studies indicated varying bar lengths and orientations (e.g. Binney, Gerhard & Spergel 1997; Bissantz & Gerhard 2002) suggesting a bar with a half-length of ~ 3.5 kpc and orientation of $\theta = 20°$. In contrast, López-Corredoira (2007) proposed a longer

bar of ~ 3.9 kpc half-length at $\theta \sim 43°$, while Hammersley et al. (2000) observed a bar with a radii extending ~ 4 kpc at $\theta \sim 43°$ using near-infrared data. Subsequently, Cabrera-Lavers et al. (2007) argue for two distinct bar structures, a long thin bar contained within the plane and orientation of $\theta = 43°$, and a triaxial boxy bulge oriented $\theta = 12.6°$. Wegg, Gerhard & Portail (2015) described a two-component long bar with estimated orientation angles of $28° - 33°$, emphasizing a 'thin' and 'superthin' structure, and a half-length of $4.6 - 5$ kpc. They contrast this with a 3D inner bulge/bar component at about $\sim 27°$ (Wegg & Gerhard 2013). More recent observational studies found values for the pattern speed between $37 - 41$ km s^{-1} kpc^{-1} and bar orientations between $20° - 28°$ (e.g. Bovy et al. 2019; Clarke et al. 2019; Queiroz et al. 2021). The study from Gaia Collaboration (2023) used data from the *Gaia* DR3 and found a bar orientation of $\sim 20°$ and pattern speed of ~ 38 km s^{-1} kpc^{-1}.

On the numerical side, Pettitt et al. (2014) in their hydrodynamical simulations of the ISM using smoothed particle hydrodynamics (SPH) conduct a parameter search with analytic prescriptions for the galactic bar potential, and find that their best combined model (arms + bar) occur with a pattern speed (Ω_p) of $\sim 50 - 60$ km s^{-1} kpc^{-1} and a bar orientation of $\sim 45°$. On the other hand, Li et al. (2022) use high-resolution simulations with a barred potential based on observations in agreement with the boxy/peanut bulge hypothesis. Their results provide a good fit to the observed bar, with a bar pattern speed that is in the range of $37.5 - 40$ km s^{-1} kpc^{-1}. However, numerical studies focusing on the inner MW, rather than encompassing the entire galactic disc, suggest pattern speed values as low as $30 - 40$ km s^{-1} kpc^{-1} (e.g. Weiner & Sellwood 1999; Sormani & Magorrian 2015; Wegg et al. 2015; Portail et al. 2017; Clarke et al. 2019; Clarke & Gerhard 2022; Sormani et al. 2022).

With the observational challenges of getting a clearer view of how the ISM in the MW is affected by its galactic structures, we approach this problem numerically. In order to model the spiral/bar features in the gas of the MW using numerical simulations, there are two possibilities regarding how the galactic (dark matter and stellar) potential is represented. The most common approach is to set up an analytical gravitational potential for the gas to follow (e.g Dobbs & Pringle 2013; Smith et al. 2020), which typically involves including spiral and bar patterns that rotate with a fixed angular velocity, that is, as a solid body. There have been many simulations that use different techniques, such as those based on Eulerian grids (e.g. Weiner & Sellwood 1999), particle-based Lagrangian methods where external potentials have been imposed (e.g. Lee et al. 1999; Pettitt et al. 2014), or moving-mesh codes with external potentials (e.g. Sormani et al. 2018; Hatchfield et al. 2021; Treß et al. 2021).

The other approach to model spiral galaxies is to use stellar and dark matter particles to recreate the galactic potential, and generate the galactic structure self-consistently through the dynamical evolution of the system (i.e. so-called live potentials). Large-scale overdensities in the galactic disc, like the spiral arms, or bars, should appear naturally from the gravitational and hydrodynamical interaction of the different components of the galaxy. There have been a number of studies that use live potentials to model spiral galaxies (e.g. Pettitt et al. 2015; Baba, Morokuma-Matsui & Saitoh 2017; Treß et al. 2021; Iles, Pettitt & Okamoto 2022), and those models show that the response of the gas to the spiral arms is different in live potentials compared to the imposed potentials, as the arms are more dynamic and usually transient, not restricted to solid-body-like rotations with a fixed pattern speed with radius (e.g. Dobbs & Baba 2014; Sellwood & Carlberg 2019; Pettitt, Ragan & Smith 2020). While analytical potentials have a clear benefit of being highly

tailored and controlled, the forces generated by the rigid external potential are imposed on the gas, whereas a live disc allows self gravity to interact both ways and induce a more realistic response of the gas.

Regardless of the type of potential used, most of the numerical works related to the MW, however, typically set out to produce an MW-type galaxy, rather than effectively attempting to fit the observed structure of the MW, which requires the creation and subsequent analysis of *l-v* maps (Fux 1999; Wada, Baba & Saitoh 2011; Li et al. 2016, e.g.). One exception is the work of Pettitt et al. (2014, 2015) who produce *l-v* maps from their models and quantitatively compare their results with observations via a χ^2-like statistic. Of particular relevance to our work, are the models with live potentials from Pettitt et al. (2015), which can reproduce the observed spiral arm features and the gas dynamics of the Galaxy, but fail to reproduce an inner bar.

In this study, we aim to reproduce the structure of the MW using numerical models with live potentials, by extending the parameter space probed by Pettitt et al. (2015). We explore the evolution of 15 different models with varying initial stellar masses and bulge fractions, in order to assess which configuration is able to better reproduce the observed MW structure. We compare with observations via *l-v* maps, utilizing a number of different statistical metrics. We investigate more specific features of our best model such as the spiral arms pattern, the galactic bar and its pattern speed, and the dynamics of the spiral arms at different radii. We then briefly compare with observations, assessing how good our best model is at reproducing these structures. Our best model will then serve in the future as the starting setup for a high-resolution simulation with chemistry, feedback, and SF. It will allow us to track the evolution of the ISM in detail, in a model that is not only MW-like, but that attempts to mimic the large-scale dynamics of the MW in a self-consistent way, key to understanding the effect of spiral structures of our own Galaxy on the SF within.

This paper is organized as follows. In Section 2, we describe the different methods used: the numerical code and the setup of the initial conditions. In Section 3, we review the observations used to constrain our simulations. In Section 4, we explain the techniques required to extract the different features of the model Galaxy, and in Section 5, we present the results from the different statistical tests. In Section 6, we examine the main galactic features of our best simulation compared to observations, and finally, we summarize our results in Section 7.

2 SIMULATION METHODOLOGY

2.1 Numerical method

For our simulations, we use the moving-mesh hydrodynamical code AREPO (Springel 2010) to simulate MW-type spiral galaxies. Here, we give only a brief description of the main features of the code – a more detailed explanation can be found in Springel (2010) and Pakmor et al. (2016), and the public code repository.[1]

AREPO combines the accuracy of grid-based methods with the adaptivity and Galilean invariance of SPH techniques, using a mesh defined by the Voronoi tessellation of a set of discrete points, which move with the velocity of the local flow. For the gravity treatment, AREPO benefits from an improved tree-based approach from the GADGET-2 code (Springel 2005). It has successfully been used in

the past on a variety of scales, being first designed for cosmological simulations at different epochs (e.g. Marinacci et al. 2015; Weinberger et al. 2017; Springel et al. 2018; Pillepich et al. 2018a, b, b; Nelson et al. 2019). AREPO has also been used to reproduce galaxies within a cosmological environment via the 'zoom-in' technique (e.g. Grand et al. 2016; Gómez et al. 2017; Garrison-Kimmel et al. 2018; Simpson et al. 2018; Grand et al. 2019; Liao et al. 2019; van de Voort et al. 2021), as well as individual isolated galaxies (e.g. Smith et al. 2014a; Glover & Smith 2016; Sormani et al. 2018, 2019; Reissl et al. 2020; Treß et al. 2021), and down to SF inside individual molecular clouds and filamentary structures (e.g. Smith, Glover & Klessen 2014b; Clark & Glover 2015; Bertram et al. 2016; Clark et al. 2019; Clarke, Williams & Walch 2020; Smith et al. 2020; Whitworth et al. 2022).

As the main interest of this study is the dynamical evolution of the stars in order to form spiral arms and bars, we do not include the gas self-gravity in our calculations. This means that while the gas only feels the gravitational potential from the stars and dark matter, the stars and dark matter halo particles feel the entire potential, including the gas contribution. This has therefore no impact on the large-scale potential felt by the stars and reduces the overall computational time required, and the gas effectively effectively traces the stellar disc potential gradients. Gas is still required in our simulations for several reasons, such as facilitating an easier transitions to modelling with a full set of stellar/ISM physics or the well-defined features the gas traces when in contrast with stars or DM for a latter comparison with observations. Nevertheless, the total gas mass is only ~ 20 per cent of the total stellar mass (and ~ 1 per cent of the total mass of the galaxy), so the contribution of the gas to the overall galactic potential is minimal, and removing the gas self-gravity does not have a great impact on the global dynamics of the gas at kpc scales. Tests including self-gravity showed some gas accumulating in artificial clumps, but the overall structure was consistent with the current simulations.

The models presented here are 3D and the gas cells move freely with no imposed analytical potential. Instead, the galactic potential is recreated by the inclusion of stellar and dark matter particles, which evolve dynamically and self-consistently. Depending on the stability of the initial conditions, this will allow the creation of spiral structures and bars, although bars can take significantly longer to manifest than arm features (on Gyr time-scales). Therefore, in order to allow the *N*-body part of the model to evolve dynamically for the duration needed for bar formation, while maintaining a reasonable computational time, we keep a relatively coarse resolution for the gas cells and impose an isothermal equation of state to avoid collapse associated with ISM cooling processes. The temperature, T, is controlled by the sound speed, c_s defined in equation (1) for the isothermal equation of state, where k_B is Boltzmann's constant, and m is the mass of a single hydrogen atom.

$$c_s = \sqrt{\frac{k_B T}{m}}. \tag{1}$$

We adopt $c_s = 10\,\mathrm{km\,s^{-1}}$, which corresponds to a temperature of $T \sim 10^4$ K. A further discussion of the equations solved by AREPO can be found in Treß et al. (2021).

The gas resolution of our simulations is defined by a refinement criteria based on mass and volume of the gas cells. For our models we use a gas target mass of $1000\,\mathrm{M_\odot}$, and we impose a limit to the volume of each cell. The minimum cell volume corresponds to $27\,\mathrm{pc^3}$ (equivalent to a cube of $3\,\mathrm{pc}$ in side) and the maximum volume to $1.25 \times 10^8\,\mathrm{pc^3}$, equivalent to a cube of $500\,\mathrm{pc}$ in side. The mass resolution used for each type of particle as well as the respective softening length are presented in Table 1. The gravitational

---

[1] AREPO and relevant documentation can be found at https://arepo-code.org/.

Table 1. Mass resolution and softening length of the different particle types. Gas cells follow AREPO's adaptive softening scheme, hence here are included the minimum and maximum softening values marked with ⋆.

	Mass resolution ($M_\odot$)	Softening (pc)
Dark matter	9×10^5	300
Stars	8×10^3	50
Gas	1×10^3	$(1-4)^\star$

softening of gas cells adheres to AREPO's adaptive softening scheme and changes throughout the simulation in accordance with changes in volume of the cells.

Our models are simulated within a box of 100 kpc side, and we use periodic boundary conditions for the gas. However, our box is big enough to avoid any boundary effects in the simulated galaxy. The total number of gas cells found in each of our simulations are in the range of $23 - 26$ million for the most evolved models.

2.2 Initial conditions

The initial conditions for our MW models were set up using the moment-based code MAKENEWDISK, described in Springel, Di Matteo & Hernquist (2005), and they consist of a spherically symmetric dark matter halo, a rotationally supported disc of gas and stars, and a central stellar bulge. Motivated by cosmological simulations, the dark matter halo has a mass distribution that peaks in the centre and drops at larger radii, following the Hernquist (1990) profile, and whose analytical representation follows:

$$\rho_{\rm h}(r) = \frac{M_{\rm h}}{2\pi} \frac{a}{r(r+a)^3}, \qquad (2)$$

where, $M_{\rm h}$ refers to the total halo mass and a is the scale length of the Hernquist profile for the halo. This distribution is in agreement with the well-known Navarro–Frank–White (NFW) profile (Navarro, Frenk & White 1996) in its inner parts, but has a faster decline outwards. The scale length is often given in terms of the halo concentration c, which allows a direct comparison between profiles, and it is defined as $c = r_{200}/r_s$, where r_{200} is the radius at which the mean enclosed DM density is 200 times the critical density over the scale length of the NFW profile. The relation between a and c is given by

$$a = r_s \sqrt{2[\ln(1+c) - c/(1+c)]}. \qquad (3)$$

A more in depth comparison of the NFW and Hernquist profiles can be found in the original reference. To set up each of our models, we explore a range of concentration indices, c, between $3 - 11$, and decide the final concentration from the global rotation curve generated, in comparison with the observational values from the works of Sofue (2012) and Eilers et al. (2019).

Note that AREPO requires all space within the defined box to be populated with cells, that is, no empty space is allowed as in Lagrangian codes. Therefore, some cells with very low-density material are needed in order to facilitate the creation of the Voronoi tessellation. Hence, the total mass of the halo is in fact a combination of dark matter particles (with total mass $M_{\rm h}^{\rm dm}$) and gas (with mass $M_{\rm h}^{\rm gas}$). For all models, we assume the gaseous halo mass to be 40 per cent of the gas mass in the disc, $M_{\rm d}^{\rm gas}$.

The stellar bulge is also setup assuming spherical symmetry and a Hernquist profile with a scale length b, and total mass $M_{\rm b}$.

The disc component (of both gas and stars), follows an exponential surface density ($\Sigma_{\rm d}$) profile as:

$$\Sigma_{\rm d}(r) = \frac{M_{\rm d}}{2\pi h^2} \exp(-r/h), \qquad (4)$$

where h is the disc scale length, and $M_{\rm d}$ is the total mass in the disc, that is, the sum of both gaseous ($M_{\rm d}^{\rm gas}$) and stellar ($M_{\rm d}^*$) components.[2] In the z-direction it follows the same profile as in Springel et al. (2005) with a scale height given by z_0.

The disc scale length is directly related to the spin parameter λ and the fraction of angular momentum in the disc J_d, which assumes conservation of specific angular momentum of the material that forms the disc. Both parameters are inputs to the MAKENEWDISK code and are directly correlated with $M_{\rm d}$ and h (see Springel et al. 2005 for details). The spin parameter λ is also dependent on the concentration of the dark matter halo, c. For our simulations, we adopt $h = 3$ kpc as per the work of Pettitt et al. (2015). Hence, λ is chosen so as to obtain the desired scale length for that specific c.

In terms of velocities, as explained in full detail in section 2.3 of Springel et al. (2005), the initial rotation of the disc is determined from the potential, and the vertical velocity dispersion of the stars in the disc is set by the specified disc scale height. We adopt the default ratio as in Springel et al. (2005). For our models, the resulting Toomre Q parameters of the discs are always above a value of 6, thus they are largely Toomre stable throughout. For the bulge and halo, the velocities are setup such that the bulge will have no net rotation, and the halo has the same specific angular momentum as the disc.

A detailed description of all the parameters can be found in Springel et al. (2005). Most of the parameters we use for setting up our MW models were based on the work by Pettitt et al. (2015), given that they use hydrodynamical $+ N$-body simulations to represent the disc and bulge of the MW with live potentials as well. We therefore take their best model (i.e. Bc/Normal model) as our starting point and explore a few variations on the stellar mass distribution, based on observational constraints of the total stellar mass of the MW.

For all models, we assume a total gas mass in the disc of $M_{\rm d}^{\rm gas} = 8.6 \times 10^9\,{\rm M_\odot}$, so the exponential surface density integrated within 13 kpc reaches $\sim 8 \times 10^9\,{\rm M_\odot}$ as in Pettitt et al. (2015). Their Bc/Normal model has a total stellar mass (disc + bulge) of $M_{\rm stars} = 4.25 \times 10^{10}\,{\rm M_\odot}$, which is on the lower end of more recent observationally derived values. For instance, the work by Cautun et al. (2020) infers the MW stellar mass profile by fitting observational data from the *Gaia* DR2, and due to the large uncertainties in the observations, they find their best fit with a total stellar mass of $M_{\rm stars} \sim 5 - 6 \times 10^{10}\,{\rm M_\odot}$, depending on the assumed halo profile. Therefore, in this work we explore three initial stellar masses, $4.25 \times 10^{10}, 5.25 \times 10^{10}$, and $6.25 \times 10^{10}\,{\rm M_\odot}$, covering the range of observed values for the MW. These values are also in agreement with the mass models found by McMillan (2017).

In addition, we vary the stellar bulge fraction (BF $= M_{\rm b}^\star/M_\star$) to explore the effects on the overall galactic structure. We start with 25 per cent similar to Pettitt et al. (2015), and decrease in steps of

---

[2]The assumption that the gas follows the same exponential profile as the stellar disc is likely not appropriate for the MW. Nevertheless, we have tested the impact of redistributing the gas using a different surface density profile – such as a flat one – and found that the overall resulting structures (bar/spiral arms) were identical. Therefore, for the purpose of this paper, we have kept the original surface density profiles as per MAKENEWDISK. However, in future work – when including SF and feedback – the gas surface density profiles will need to be adjusted in order to be a better representation of the MW's gas surface densities.

Table 2. Parameters used to build the initial conditions for our different models with the MAKENEWDISK code: $M_\star$ is the total stellar mass and values are based of Pettitt et al. (2015) best model and the mass profile from Cautun et al. (2020); BF is the stellar bulge mass fraction; λ is the spin parameter (see Springel et al. 2005); J_d is the disc spin fraction; $M_d^\star$ is the stellar mass in the disc; $N_d^\star$ is the number of stellar particles in the disc; $M_b^\star$ is the stellar mass in the bulge; $N_b^\star$ is the number of stellar particles in the bulge; c is the halo concentration; M_h^{dm} is the dark matter mass in the halo; and N_h^{dm} is the number of dark matter particles in the halo. The number of particles are set by the imposed mass resolution and total masses. The values for the disc scale height and bulge scale length are common to all values and are 0.29 and 0.35 kpc, respectively.

Model	$M_\star$ $[10^{10}\,M_\odot]$	BF [per cent]	λ	J_d	$M_d^\star$ $[10^{10}\,M_\odot]$	$N_d^\star$ $[10^6]$	$M_b^\star$ $[10^{10}\,M_\odot]$	$N_b^\star$ $[10^6]$	c	M_h^{dm} $[10^{10}\,M_\odot]$	N_h^{dm} $[10^6]$
1	4.25	25	0.037	0.040	3.19	3.98	1.06	1.33	8	89.75	0.99
2	4.25	20	0.037	0.042	3.40	4.25	0.85	1.06	9	89.75	0.99
3	4.25	15	0.038	0.048	3.61	4.52	0.64	0.80	9	89.75	0.99
4	4.25	10	0.039	0.049	3.83	4.78	0.43	0.53	9	89.75	0.99
5	4.25	5	0.040	0.051	4.04	5.05	0.21	0.27	10	89.75	0.99
6	5.25	25	0.033	0.050	3.94	4.92	1.31	1.64	5	88.75	0.98
7	5.25	20	0.033	0.053	4.20	5.25	1.05	1.31	5	88.75	0.98
8	5.25	15	0.035	0.055	4.46	5.58	0.79	0.98	6	88.75	0.98
9	5.25	10	0.035	0.059	4.73	5.90	0.53	0.66	6	88.75	0.98
10	5.25	5	0.034	0.061	4.99	6.23	0.26	0.33	6	88.75	0.98
11	6.25	25	0.031	0.058	4.69	5.86	1.56	1.95	3	87.75	0.97
12	6.25	20	0.031	0.062	5.00	6.25	1.25	1.56	3	87.75	0.97
13	6.25	15	0.031	0.065	5.31	6.64	0.94	1.17	3	87.75	0.97
14	6.25	10	0.033	0.068	5.63	7.03	0.63	0.78	4	87.75	0.97
15	6.25	5	0.033	0.071	5.94	7.42	0.31	0.39	4	87.75	0.97

5 per cent until we reach the extreme 5 per cent, where there is almost no bulge. Altogether, we generate a range of 15 different models that extend over a wide range of parameters to explore the combination that best reproduces the observations of our Galaxy. We enumerate our models from 1 to 15 with increasing initial stellar mass, and from larger to smaller bulge fraction (see Table 2).

Finally, the total mass of the galaxy (i.e. dark matter, gas, and stars) is constant throughout all models and determined using

$$M_{total} = \frac{V_{200}^3}{(10 G H_0)}, \tag{5}$$

where G is Newton's gravitational constant and H_0 Hubble's constant, and the virial velocity $V_{200} = 160 \,\mathrm{km\,s^{-1}}$ is taken from Springel et al. (2005).

Fig. 1 shows an example rotation curve for Model 4. For context, we compare this model with observed values from Eilers et al. (2019), determined from 6D data from giant stars and for which the circular velocity at the Sun's location in the local standard of rest (LSR) is assumed to be $V_0 = 229 \,\mathrm{km}\,s^{-1}$. Additionally, we include the rotation curve form Reid et al. (2019), derived from maser observations and independent of the LSR velocity assumption. We have also included the data from Sofue (2012), adjusting their initial assumption of $V_0 = 200 \,\mathrm{km}\ s^{-1}$ to align with the higher circular velocities of the LSR reported by Eilers et al. (2019). This adjustment uses a distance of 8.2 kpc for the LSR, based on the results from GRAVITY Collaboration (2019). The initial rotation curve of Model 4, depicted in blue in Fig. 1, is generated using the MAKENEWDISK code. The dashed lines represent the contribution from the different galactic components: the disc in orange, the halo in green, and the bulge in red. The solid yellow line illustrates the model's rotation curve computed at a later time $t = 2.6 \,\mathrm{Gyr}$. The shaded area constitutes the 1σ standard deviation (in velocity) of the data at each radial bin, for that later time. Note that the effect of changing the bulge fraction affects the inner part of the galaxy, whereas the disc dominates towards larger radii. Therefore, for a given total stellar mass, the larger the BF, the stronger the peak in the rotation curve at $R < 1$ kpc. On the other hand, increasing the initial stellar mass reduces the halo
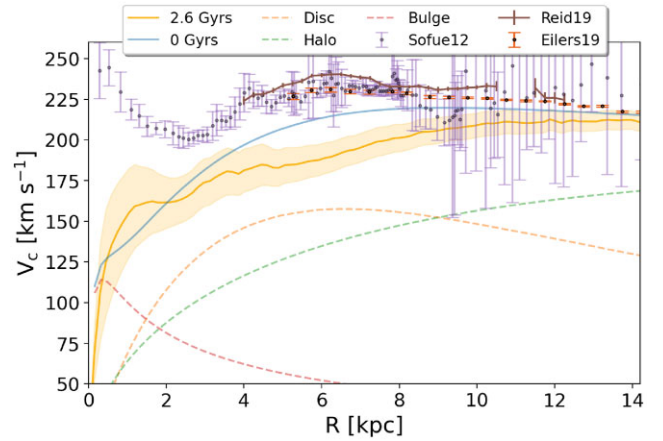

Figure 1. Example of the rotation curve for Model 4. The blue solid line indicates the initial rotation curve generated by the MAKENEWDISK code, with the contribution from the different galactic components shown in dashed lines: orange for the disc, green for the halo, and red for the bulge. The solid yellow line represents the rotation curve of the same model at a later time, where the shaded area constitutes the 1σ standard deviation (in velocity) of the data at each radial bin, at that later time. The observations from Sofue (2012), Eilers et al. (2019), and Reid et al. (2019) all shown as black dots, but with light purple, brown and orange error bars, respectively. Data for Sofue (2012) have been corrected so that the circular velocity of the LSR matches that of Eilers et al. (2019) at 229 km s^{-1}.

mass, given that the disc gas mass and total mass of the galaxy are fixed parameters. Hence, for a constant BF, the larger the stellar mass, the faster the rotation curve decays at larger distances from the galactic centre, where it is dominated by the halo contribution. When compared to the observations (see in Fig. 1 for reference), our models tend to underestimate the velocities, especially at radii below 8 kpc. We note, however, while the initial rotation curve is only slightly underestimated at those intermediate radii (by up to ~10 per cent), the difference between the rotation curves of our models and the

observations becomes more severe at later times, after the disc has settled – likely due to radial migration of mass, and perhaps linked to the lack of gas self-gravity (which promotes an initial redistribution of the gas outwards). Another notable difference is towards the galactic centre (i.e. $R < 2\,$kpc) in the initial conditions, but this aids to destabilize the central region and create the right conditions to generate a bar, which tends to be hampered by the presence of strong spherically symmetric and non-rotational components such as bulges and centrally concentrated haloes (see e.g. Kataria & Das 2018, and references in the Introduction). With the evolution of the galaxy and the formation of a central bar, the rotation curve changes and we can see that the central region does reach higher velocities than in the original curve. However, as we will see in Section 6.4, even with the bars, we still underestimate the velocities in the most central region ($R < 1\,$kpc), in which the motions of the MW are mostly dominated by x_2 orbits (see e.g. Binney et al. 1991; Hatchfield et al. 2021), which we do not resolve properly in all of our models. We note, however, that the observed velocities inside a radius of $\sim 4\,$kpc are unlikely to be correct due to the assumption of circular orbits in the bar region.

Summarizing, we have 15 different models and vary their initial stellar mass and bulge fractions to account for a diversity of possible large-scale configurations. Each model has a fixed halo concentration, and spin parameter tailored to provide a disc scale length of $h = 3\,$kpc. Disc scale height (z_0) and bulge scale length (b) are calculated in terms of h, and have final values for all models of 0.29 and 0.35 kpc, respectively. The total mass of the galaxy is constant throughout the sample, $M_{\rm total} = 9.52 \times 10^{11}\,{\rm M_\odot}$, as well as the disc gas mass, $M_{\rm d}^{\rm gas} = 8.6 \times 10^9\,{\rm M_\odot}$. We assume that the gas mass in the halo corresponds to 40 per cent of the gaseous component of the disc, and thus it is a fixed quantity, $M_{\rm h}^{\rm gas} = 3.44 \times 10^9\,{\rm M_\odot}$. Table 2 collects all the other different parameters for our range of models that are not pre-determined. Example top-view plots of our different models can be seen in Appendix B.

3 OBSERVATIONS

We make use of the ^{12}CO emission cubes from Dame et al. (2001),[3] as the primary data set of comparison with our models. We use them to obtain the main galactic features as seen in longitude–velocity space, as this is the gas tracer that sees the galactic spiral features with highest contrast. The Dame et al. (2001) CO survey covers the entirety of the galactic plane, and the measurements were obtained with the CfA 1.2 m Millimeter-Wave Telescope at the Harvard-Smithsonian Center for Astrophysics and its counterpart instrument at CTIO in Chile. This survey is formed by 488 000 spectra that cover the MW over a strip wide in latitude of $4°$–$10°$. For our study, we use the *l-v* map that can be seen in Fig. 2, which was smoothed to a resolution of $\Delta v = 2\,{\rm km\,s^{-1}}$ in velocity and $\Delta l = 12\,$arcmin in longitude.

In addition to the ^{12}CO data, we also make use of H I data to determine the terminal velocities of the MW in *l-v* space (see Section. 4.4), as it traces the more diffuse gas of the outer Galaxy better than CO. For the H I emission we use the HI-4PI all-sky H I Survey (HI4PI Collaboration 2016), which contains data for $N_{\rm H_I}$ column density as well as latitude, longitude, and brightness temperature. The obtained *l-v* map can be seen in Fig. 3, at the same resolution of Fig. 2. The HI-4PI Survey is a combination of the first coverage of the Effelsberg–
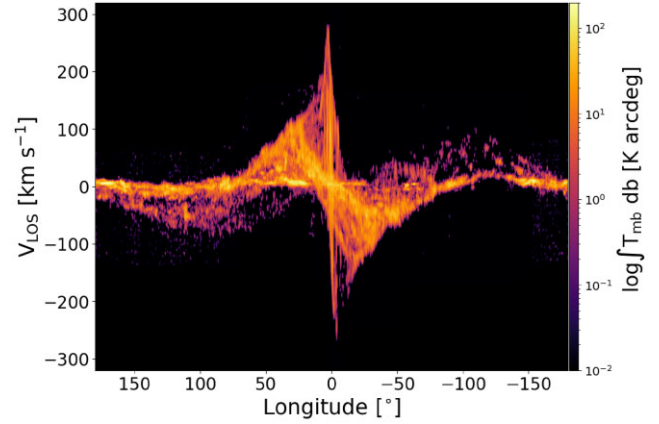
Figure 2. All-sky longitude–velocity map of the MW extracted from the work by Dame et al. (2001). This map was smoothed to a resolution of $\Delta v = 2\,{\rm km\,s^{-1}}$ in velocity and $\Delta l = 12\,$arcmin in longitude. The colour bar refers to the logged integrated brightness temperature of the CO ($J = 0 - 1$) transition.
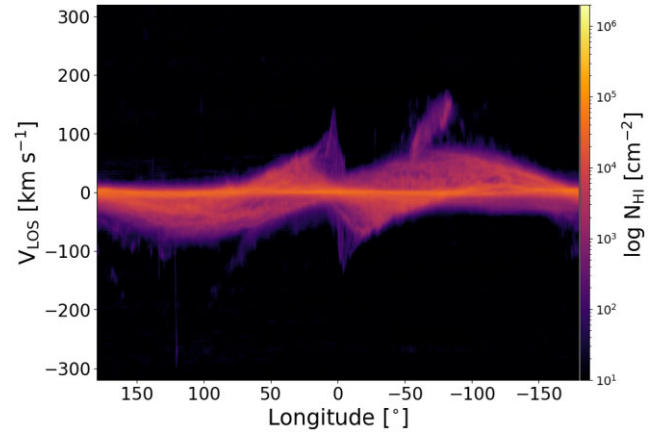

Figure 3. All-sky longitude–velocity map of H I column density of the MW. The data were extracted from the HI-4PI Survey (HI4PI Collaboration 2016) and plotted in a logarithm scale.

Bonn H I Survey and the third version of the Galactic All-Sky Survey. It has an angular resolution of $\vartheta_{\rm fwhm} = 16.2$, and covers the complete MW in longitude. All the data are publicly available.[4]

4 EXTRACTION OF THE MAIN GALACTIC FEATURES

In this section, we present the methods that we use to quantify how well our numerical models reproduce the observed galactic structures. In particular, we explain how we place ourselves inside the models in order to mimic the observer's perspective of the MW (i.e. through *l-v* plots), the techniques we use to identify the key galactic features, and the metrics used to quantify the how the models compare with observations.

4.1 Construction of *l-v* plots

The spiral patterns of galaxies are generated by the dynamical evolution of the different galactic components, and the larger stellar potential well inside a spiral arm aids the gas to concentrate, favouring the formation of larger complexes of molecular gas (giant molecular clouds). As such, molecular gas is typically a good tracer of the spiral pattern of galaxies (e.g. Young & Scoville 1991; Wright et al. 2001; Combes, Young & Bureau 2007), as it presents a higher contrast between arm and interarms regions than if observing the atomic gas (e.g. Querejeta et al. 2021 where they used H I). This spiral pattern can be seen as clear lines (and loops) in *l-v* maps of gas tracers, such as those in Dame et al. (2001). For our models, given that we are simply interested in the position of the spiral arm tracks in *l-v* space, and that we do not include chemistry or cooling in the models, we instead simply generate synthetic *l-v* maps from the total gas mass along each line of sight.

The framework for the construction of *l-v* plots from a 3D galaxy is explained in detail in Binney & Tremaine (2008) and Binney & Merrifield (1998), as well as in Pettitt et al. (2014). However, here we will give a brief summary of the process, as we adopt a more graphical approach.

The first step in constructing an *l-v* map is to adopt a position for the observer (i.e. the Sun's position). These coordinates can be defined by the distance from the galactic centre, or R_{obs}, the angular position in the box within the disc with respect to the Sun–Galactic centre line ϕ_{obs}, and the tangential velocity at the Solar position V_{obs}. The position of the galactic centre is determined for each time, as the centre of mass of the entire system (gas, stars, and dark matter). For R_{obs}, we adopt an approximate value of $R_{obs} \sim 8.2$ kpc, based on the study by GRAVITY Collaboration (2019). The azimuthal position of the observer is then essentially a free parameter. We first perform a weighted principal component analysis (WPCA, Bailey 2012) on the top-view column density of each model in our sample to determine the orientation of the Galactic bar. Afterwards, we rotate the model to align the bar along the *x*-axis and examine six different angles. Half of these angles are selected to ensure that the bar's orientation falls within the observed range of 20°–45°. Therefore, we select angles spaced by 12.5°, and the remaining three angles are their mirrored counterparts, effectively adding 180° to each of the initial three angles. We then produce an *l-v* map for each observer position. An illustration of the different angles used can be seen in Fig. 4.

Using the galactic centre as the origin, the Sun position can be calculated in polar coordinates via R_{obs} and ϕ_{obs}. Given the gas position and velocity from AREPO in Cartesian coordinates, we then rotate the galaxy clockwise with respect to the galactic centre, such that the Sun is positioned at $y = 0$ and $x = R_{obs}$. We then obtain the longitude l_{gas} of each gas cell using equation (6):

$$l_{gas} = \frac{3\pi}{2} - \mathrm{atan2}(y, x), \tag{6}$$

where *x* and *y* are the gas coordinates in the rotated frame with origin at the Sun's position. We then obtain the velocity of the gas at the observer's position, V_{obs}, as the mean velocity of all particles that fall within a range of radius of $R_{obs} \pm 0.1$ kpc of the observer. We assume that the observer's motion is purely circular (i.e. we neglect any radial motions), and therefore we assume that its velocity has only a *y*-component in this frame. Finally, we project the cell velocity in the line-of-sight direction via a dot product of the cell velocity vector and the unity vector in the line of sight direction of the observer, obtaining v_{los}.
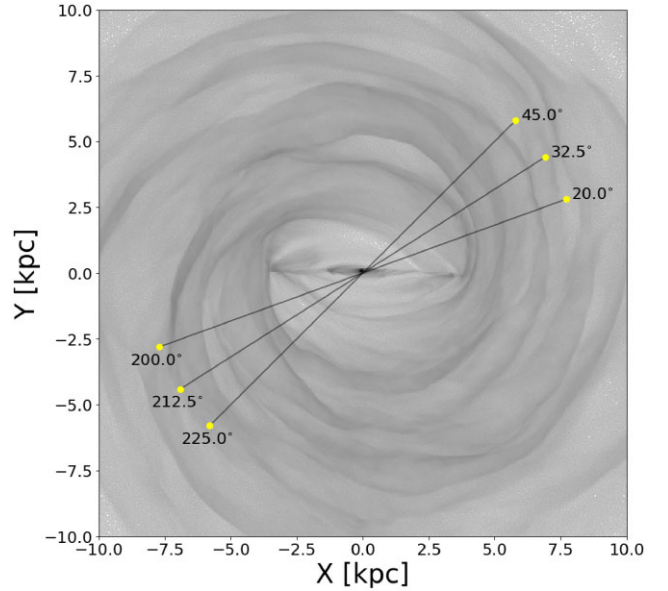

Figure 4. Example of one of the top-down views of Model 4 at a time of ~ 2.6 Gyr rotated so the bar is along the *x*-axis, and illustrating the range of angular positions in the box, ϕ_{obs}, assumed for the Sun's position when creating the *l-v* maps.

4.2 Skeleton extraction

We aim to trace the position of the spiral arms in the gas and carry out a direct comparison with observations, where the denser ridges of the spiral arms are best traced by the molecular gas, and in particular we select CO emission. A more in depth comparison of numerical and observational data, would ideally need a post-processing analysis of the numerical data with radiative transfer codes, in order to simulate the observed emission (e.g. Duarte-Cabral et al. 2015). Nevertheless, given that the large-scale galactic features are mainly dominated by the large-scale dynamics of the stellar population rather than the local physics of the gas (Pettitt et al. 2014, e.g.), it is possible to do a feature comparison based on matching the fingerprints (or 'skeletons') (Fux 1999; Rodriguez-Fernandez & Combes 2008), of our simulations to those in observations. As such, we extract the 'skeletons' of all *l-v* maps in order to highlight the position of the key structures as seen in the *l-v* space in both the models, and in the observations.

Essentially, we take the following steps to make a skeleton *l-v* map, also illustrated in Fig. 5 (for the observed CO emission on the top, and for one of our models on the bottom):

(i) We regrid both simulations and observations to a common grid, so that the comparison can be done on a 1-to-1 basis (first column in Fig. 5). We used a pixel size with $\Delta l = 0.125°$ in the longitude axis, and $\Delta v = 1.3$ km s^{-1} in velocity, based on the resolution of the original observations.

(ii) We smooth these *l-v* images by convolving them with a Gaussian 2D Kernel to reduce the noise, mostly to aid in highlighting the contiguous emission in the observational data along the spiral arms (rather than the 'broken' emission from individual molecular clouds). For this convolution, we used a standard deviation of $\sigma_l = 1.0°$ and $\sigma_v = 5$ km s^{-1}, and we do the same convolution to the modelled *l-v* to keep them comparable (second column in Fig. 5).

(iii) We compute, for each pixel, the Hessian matrix and its respective eigenvalues and eigenvectors. The eigenvalues of the Hessian matrix describe the magnitude of the curvature along the direction of their respective eigenvector. Negative eigenvalues
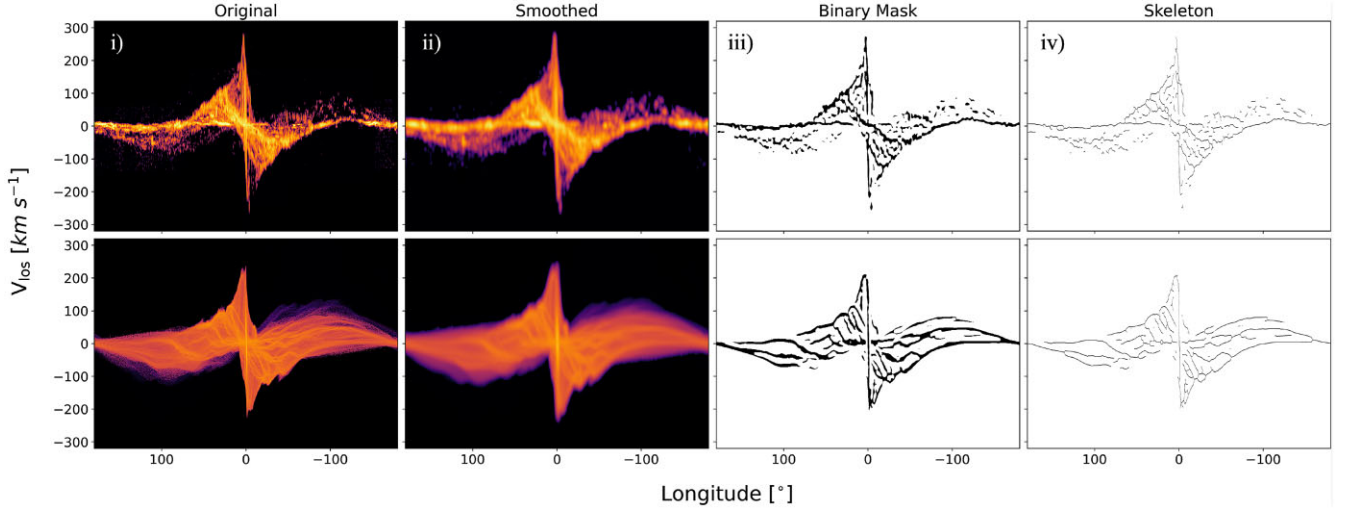
Figure 5. Different steps of the extraction of the skeletons of the observed and simulated *l-v* maps. The top four panels illustrate the C0 emission map from Dame et al. ([2001](#)), whilst the bottom panels display the example *l-v* map for Model 3 at a time ~ 2.4 Gyr and angle $\phi_{\mathrm{obs}} = 32.5°$. (i) Original image. (ii) Smoothed image with a Gaussian 2D kernel of $\sigma_1 = 1.0°$ and $\sigma_v = 5$ km s^{-1}. (iii) Binary mask obtained after finding the peaks via a Hessian matrix and applying a binary erosion of 2 pixels. (iv) Final skeleton image obtained by computing the medial axis of the previous image.

correspond to convex curvature, positive eigenvalues correspond to concave curvature, and a value of 0 is flat (e.g. Liang & Yuen [2013](#)). Given that the spiral arms in an *l-v* map are apparent as long and thin tracks, they can be approximated to 'filamentary' structures. We define a filament in 2D as a convex structure along the shortest axis. Therefore, for a 2D filamentary structure, at least one of the eigenvalues must be always negative, $\lambda_1 < 0$. The second eigenvalue λ_2 can be either concave or convex. However, λ_1 should always be more strongly curved than λ_2. Therefore, the second requirement is that $| \lambda_1 | > | \lambda_2 |$. In order to avoid faint filamentary structures, we add a third requirement on the minimum amount of curvature. We set a λ_{lim} so that only regions of the map that have $\lambda_1 < \lambda_{\mathrm{lim}}$ are classified as filaments. This λ_{lim} is always negative and has to be chosen manually for each different case. With the pixels that satisfy these conditions, we create a binary mask, to which we apply an extra binary erosion of 2 pixels to thinner these wider areas (third column in Fig. [5](#)).

(iv) We compute the medial axis of the eroded mask of the maps from (iii). This returns a structure with a width of 1 pixel, that collects all bins containing more than one point closer to them than the boundary of the object. This is the final binary map that collects the skeleton information. A value of 1 is given to those pixels containing a skeleton, and value of 0 is given elsewhere (fourth column in Fig. [5](#)).

We perform this process for the observations as well as for our entire set of simulations at all studied times (see Section [5.1](#)) and angles (see Section [4.1](#)).

4.3 Comparison of skeletons between simulations and observations: SMHD

We use the symmetrized modified Hausdorff distance (SMHD) statistical tool (presented in Sormani & Magorrian [2015](#)) to evaluate how well our models reproduce the large-scale morphology of the MW. This is a symmetric version of the metric first introduced in Dubuisson & Jain ([1994](#)) and works well even with few-defined parameters or contaminated data. As stated in the work by Sormani & Magorrian ([2015](#)), this method is found to be more powerful than some of the other methods used in the literature for similar purposes,

as it works well even when fewer parameters are defined or the data are contaminated.

Essentially, this technique is a comparison of pixel distances between binary images. The distance d in *l-v* space is defined in equation ([7](#)). Given that in our case, our 2D image is not composed of two spacial axis, we need to weight the distance in each direction differently. Therefore, we set the 'uncertainty' in the longitude axis to $\Delta l = 1°$ (which corresponds to the angular resolution of the smoothed image we use to construct the skeletons) and in the velocity axis to $\Delta v = 10$ km s^{-1} (which is the typical velocity width of a spiral arm).

$$d(a, b) = \frac{| l_a - l_b |}{\Delta l} + \frac{| v_a - v_b |}{\Delta v}. \tag{7}$$

The modified Hausdorff distance (MHD) is defined as the sum of the distances between each pixel i containing 1 in the binary image a, and the closest pixel j containing 1 in the binary image b, and the expression can be seen in equation ([8](#)):

$$\mathrm{MHD}(a, b) \equiv \sum_i \min_j (d(a_i, b_j)). \tag{8}$$

If the set a has N pixels, and the set b M pixels, then the symmetric version is defined as seen in equation ([9](#)):

$$\mathrm{SMHD}(a, b) \equiv \frac{1}{2N} \mathrm{MHD}(a, b) + \frac{1}{2M} \mathrm{MHD}(b, a). \tag{9}$$

Hence, lower values of this metric indicate a better fit: distances from structures in the simulations are shorter to structures in the observations (and vice versa).

Similarly to Sormani & Magorrian ([2015](#)), we apply the SMHD metric on the skeletonized *l-v*-maps, to compare the observed and model skeletons. However, note that although our procedure to extract skeletons is similar to that used by Sormani & Magorrian ([2015](#)), it is not identical. Our method differs to that of Sormani & Magorrian ([2015](#)) in that we do not do an envelope enhancement, because we want this method to be mostly sensitive to the internal features of the galaxy (such as spiral arm tracks), rather than including the terminal velocity into the same metric. However, we then analyse the terminal velocity separately with another metric

(see Section 4.4). We also do not do a non-maximum suppression and hysteresis thresholding, but do a medial-axis calculation instead. This, in addition to the fact that the pixel sizes and areas covered by our calculations are not the same as those in Sormani & Magorrian (2015, therefore changing the number of 'ridge' pixels in which to estimate the metric), means the values of the metric itself, as applied to our skeletonized data set, can only be used to inter-compare our models, but cannot be directly compared to the metric values from Sormani & Magorrian (2015).

After some experimentation, we found that this metric was prone to favouring simulations that were more crowded (i.e. had more structures/skeletons), even if such structures did not exist in the observed data set. To quantify this effect, we include an extra metric based on the fraction of skeleton pixels that are mismatched between the two data sets. For this, we store the information of which skeleton pixel of image *a* is matched with in image *b* and vice versa, and then find the pairs that are uniquely matched to each other (i.e. a perfect match). All remaining pixels are considered mismatches – i.e. they represent structures that are not present on both comparison images.

We then look at each of the individual sets: simulations (sim) and observations (obs). For each of them, we define the fraction of mismatches, f_{sim} or f_{obs}, as the number of mismatched skeleton pixels (S_{sim} or S_{obs}) over the total number of pixels in that skeleton (A_{sim} or A_{obs}), as

$$f_{sim,obs} = \frac{S_{sim,obs}}{A_{sim,obs}} \tag{10}$$

We then add the two fractions to create a global value for the fraction of mismatches of each model, f_{mis}, as

$$f_{mis} = f_{sim} + f_{obs} \tag{11}$$

that can vary between 0 if it is a perfect match, and 2 if there are no uniquely paired pixels. This will only act as an extra layer to compare models between each other and it is not accurate enough to rely on to predict the best model on its own. A further discussion can be seen in Appendix A.

4.4 Terminal velocity

One other way of gauging whether the models are able to reproduce the observations, is to compare the area occupied by the gas in *l-v* space. To do this we need to define the outer edge of this area, that is, the terminal velocities along a given line of sight. In order to do so for the observational data, since the edge is not nicely defined as a continuous line due to noise, we extract the polygon that roughly outlines the edges of the images by hand using the Cube Analysis and Rendering Tool for Astronomy (CARTA).[5] We create the polygon mask image for the observations in ^{12}CO from Dame et al. (2001) and the H I from HI4PI Collaboration (2016) separately, we regrid the H I one to match the resolution of the ^{12}CO map, and then we add them both to create a single mask from observations. As the *l-v*-maps of our simulations are noise-free and their extent is well defined, the terminal velocity edge is defined from a simple contour of the data in *l-v* space. These masks contain 1 for all pixels inside the outer edge, and 0 elsewhere. We proceed to subtract the mask of the simulations from the mask of the observations, so as to obtain an *l-v* map where: matching pixels receive a value of 0, pixels only covered in observations have a value of 1, and pixels solely in the models

---

[5] https://cartavis.org/

have a value of −1. We can use this information to further infer how good a model is at reproducing the observed terminal velocities by quantifying this difference. Thus, we estimate the fraction of non-zero pixels, *m*, and average by the longitude range of the region used for the statistics, Δl, such that we obtain the average fraction of non-matched pixels per unit angle as $\langle m \rangle = m/\Delta l$.

5 FINDING THE BEST FIT FOR THE MILKY WAY STRUCTURE

In order to compare our suite of numerical models to the observations of the MW, we apply the techniques described in Section 4 to our simulations at a number of different timestamps. In this section, we describe how we select the simulation snapshots to analyse, and the results of that analysis in terms of the best snapshot per model, as well as the overall best model at reproducing the observed galactic structure.

5.1 Numerical sampling

We analyse the galactic structure of the 15 different models with varying stellar distributions that compose our numerical sample. We only investigate the models from a time of ~ 2 Gyr onwards, so that we avoid any potential structural artefacts arising from the initial settling of the disc. We study a total of 20 snapshots per model, in intervals of ~ 50 Myr up until the simulation reaches a total time of ~ 3 Gyr. For each of these snapshots, we select a sample of six different angles ϕ_{obs} around the galactic centre for the Sun position and produce the *l-v* maps as explained in Section 4.1. In total, we have a sample of 1800 different simulated *l-v* maps to compare to the observed ^{12}CO emission.

5.2 Results from SMHD analysis

As detailed in Section 4.3, we use the SMHD metric plus the global fraction of mismatches, f_{mis}, for a first comparison of the galactic structures formed in the models to the CO observations from Dame et al. (2001). In this section, we will find the overall lowest value of SMHD in time as well as angle for each case, also dividing the galaxy in an inner region of longitude $| l | < 20°$, and an outer region with $| l | \geq 20°$. In doing so, we explore the effect that the spiral structures as well as the bar (if any) have on our metric and on the decision of which model provides the best fit of the MW. We then proceed to perform a direct comparison of the results of the metric between the models in our numerical sample.

5.2.1 Optimal time and angle for each model

We start by investigating how the average value of SMHD at each snapshot (averaged over all viewing angles ϕ_{obs}) varied with time. We do not find any specific trends, which suggests that, within the times studied, all models reached a relatively stable state in terms of the type of galactic structures formed. Therefore, we select our optimal snapshot for each model, as the time at which that model has the lowest averaged SMHD, and that is the only snapshot used for the remainder of the analysis. In order to ensure that the overall best fit was not dominated by either just the bar or the spiral arms, we further analyse this behaviour for the inner and outer Galaxy. Neither of these cases show a clear structure to the time evolution of the metric, and hence do not contribute any clear results.

To determine the most suitable viewing angle ϕ_{obs} for each model at its optimal time, we look for the angle where the SMHD value is the

Table 3. Table containing results for the best time and viewing angle for all different measurements and metrics performed on the 15 different models that compose our numerical sample. From left to right: model number; optimal time; SMHD metric for global, inner and outer galaxy (SMHD, SMHD$_{inn}$, and SMHD$_{out}$); global, inner, and outer fraction of mismatches (f_{global}, f_{inn}, and f_{out}); global, inner, and outer angle averaged terminal velocity metric ($\langle m_{global}\rangle$, $\langle m_{inn}\rangle$, and $\langle m_{out}\rangle$); bar half-length ($L/2$); bar inclination (α); and bar pattern speed Ω.

Model	Time	SMHD	SMHD$_{inn}$	SMHD$_{out}$	f_{global}	f_{inn}	f_{out}	$\langle m_{global}\rangle$	$\langle m_{inn}\rangle$	$\langle m_{out}\rangle$	$L/2$	α	Ω
1	2400	[0.52, 0.54]	0.79	0.44	[0.62, 0.65]	0.99	0.54	[292, 299]	680	249	nan	32	21.0 ± 2.0
2	2700	[0.52, 0.53]	0.71	0.46	[0.59, 0.60]	0.97	0.48	[250, 252]	518	217	3.8 ± 0.8	45	32.0 ± 2.0
3	2600	[0.56, 0.58]	0.83	0.48	[0.64, 0.66]	1.01	0.54	[261, 265]	503	235	3.1 ± 0.8	20	38.2 ± 0.4
4	2650	[0.48, 0.50]	0.61	0.44	[0.62, 0.66]	0.88	0.53	[223, 255]	457	219	3.2 ± 0.8	45	30.0 ± 0.2
5	2050	[0.55, 0.56]	0.77	0.49	[0.64, 0.66]	0.99	0.53	[228, 230]	548	189	3.4 ± 0.8	32	31.2 ± 0.7
6	2100	[0.59, 0.64]	0.9	0.49	[0.67, 0.77]	1.11	0.54	[274, 321]	481	248	4.0 ± 0.8	20	29.2 ± 0.1
7	2450	[0.64, 0.70]	1.04	0.54	[0.69, 0.70]	1.13	0.57	[216, 217]	543	175	2.8 ± 0.8	20	27.3 ± 0.3
8	2150	[0.70, 0.72]	1.17	0.56	[0.64, 0.67]	1.08	0.53	[246, 254]	401	227	4.9 ± 0.8	45	26.8 ± 0.2
9	2100	[0.63, 0.65]	0.96	0.54	[0.60, 0.63]	0.98	0.51	[221, 223]	272	215	3.1 ± 0.8	20	26.1 ± 0.2
10	2200	[0.62, 0.66]	1.07	0.48	[0.70, 0.72]	1.09	0.58	[216, 248]	445	199	3.7 ± 0.8	45	26.7 ± 0.2
11	2950	[0.83, 0.84]	1.4	0.66	[0.79, 0.83]	1.23	0.67	[273, 282]	822	206	4.4 ± 0.8	45	29.0 ± 0.2
12	2050	[0.76, 0.77]	1.44	0.55	[0.70, 0.72]	1.11	0.58	[307, 319]	728	260	3.6 ± 0.8	32	29.0 ± 0.2
13	2100	[0.75, 0.77]	1.16	0.62	[0.77, 0.78]	1.12	0.67	[264, 288]	663	234	2.8 ± 0.8	45	26.5 ± 0.4
14	2050	[0.75, 0.76]	1.22	0.61	[0.73, 0.74]	1.17	0.6	[271, 284]	448	258	4.0 ± 0.8	20	20.8 ± 0.8
15	2350	[0.69, 0.73]	1.22	0.54	[0.67, 0.72]	1.09	0.55	[272, 855]	684	221	4.1 ± 0.8	20	26.8 ± 0.2

lowest. This choice is made after ensuring that the bar's inclination relative to the observer aligns with known observations (as detailed in Section 4). In order to quantify the possible variations of the SMHD metric value due to small changes in ϕ_{obs}, we also produce the *l-v* plots and estimate the respective SMHD value, for two additional angles, $5°$ apart from the original optimal ϕ_{obs}. Our analysis confirmed that the initially selected angle was indeed the most suitable within this expanded range for all models, showing that our results are robust and not just a consequence of the limited number of angles initially considered.

Table 3 presents the range of SMHD values for these three angles. The lowest SMHD value corresponds to the optimal ϕ_{obs}. Additionally, Table 3 also includes the SMHD values at the optimal time and viewing angle, as measured for the inner ($|l| < 20°$) and outer ($|l| > 20°$) regions of the galaxy separately.

In summary, we select the optimal time for each model based on the minimum of the angle-averaged SMHD metric, and the optimal viewing angle where the bar inclination is within observed ranges and where the SMHD metric was lowest.

5.2.2 Optimal model

We now proceed to carry out a direct comparison of our models between each other, in terms of their SMHD metric. Table 3 is a compilation of all the metrics we use to intercompare models. We include the range of values for the global SMHD metric, fraction of mismatches and terminal velocity metric at the optimal time for the optimal viewing angle $\phi_{obs} \pm 5°$. For all cases, the lowest value correspond to the final chosen viewing angle. We also include the inner and outer values of this metrics, as well as the bar parameters (half-length $L/2$, inclination α, and pattern speed Ω, as determined in Section 6.1).

In order to see how the models compare to each other, we represent the variation of the final SMHD against the models in Fig. 6. The data points have been colour-coded so that for a fixed initial stellar mass, lighter to darker colours indicate a decrease in the initial bulge fraction: 25 per cent, 20 per cent, 15 per cent, 10 per cent, and 5 per cent. Blue data points indicate models with an initial stellar mass of $M_\star = 4.25 \times 10^{10}$ M$_\odot$, whereas red indicates $M_\star = 5.25 \times 10^{10}$ M$_\odot$ and green $M_\star = 6.25 \times 10^{10}$ M$_\odot$. We have added an arrow

to indicate the variation of the SMHD with the range of $\pm 5°$, as well as a horizontal line at the median of the distribution, for easier comparison of the variability with viewing angle between models. It is clearly visible that the lower values for the metric are achieved with the lower initial stellar mass. Therefore, we exclude all models above 5 as the SMHD metric keeps increasing. Model 1 can also be dismissed, as it does not generate a galactic bar.

In order to assess how good a match any given model is at its optimal time and viewing angle, we not only check that the SMHD metric is low, but also that the number of mismatches is low. Hence, we then investigate the fraction of mismatches and check whether they concur with the SMHD metric. For all models f_{inn} is always higher than f_{out} (as seen in Table 3), suggesting that there are more unmatched structures towards the inner galaxy. The global fraction of mismatches for each model can be seen in the right panel of Fig. 6. Similarly to the SMHD, the value of f_{global} generally increases with increasing initial stellar masses. There is one clear exception to this trend, for model 9. This model has values of f_{global}, f_{inn}, and f_{out} similar to those of Models 2–5. However, given that its global SMHD is significantly worse than Models 2–5, we do not consider this model further.

In our comparison of models with a bar and the lowest global SMHD values (i.e. Models 2, 3, 4 and 5), we find that the fraction of mismatches is relatively similar across these models, both globally and within the inner/outer regions of the Galaxy. Notably, Models 2 and 4 exhibit the lowest mismatch fractions. Among them, Model 2 has a marginally lower overall fraction of mismatches. However, Model 4 demonstrates a reduced fraction of mismatches in the inner Galaxy (f_{inn}), suggesting it more accurately represents the inner galactic structures. Furthermore, the global SMHD metric for Model 4 is distinctly lower than that of Model 2, including both the inner and outer values. These findings indicate that Model 4 is a better overall match of the structures in $l-v$ space.

Fig. 7 illustrates this point further. It displays the extracted skeletons from the $l-v$ maps of the ^{12}CO observations by Dame et al. (1986) at the top, and our Model 4 at approximately ~ 2.6 Gyr with an observer angle $\phi_{obs} = 45°$ at the bottom. The pixels are color-coded, with blue representing uniquely associated pixels and red indicating those defined as mismatches. This figure reveals that, while most structures in the simulation's inner and outer
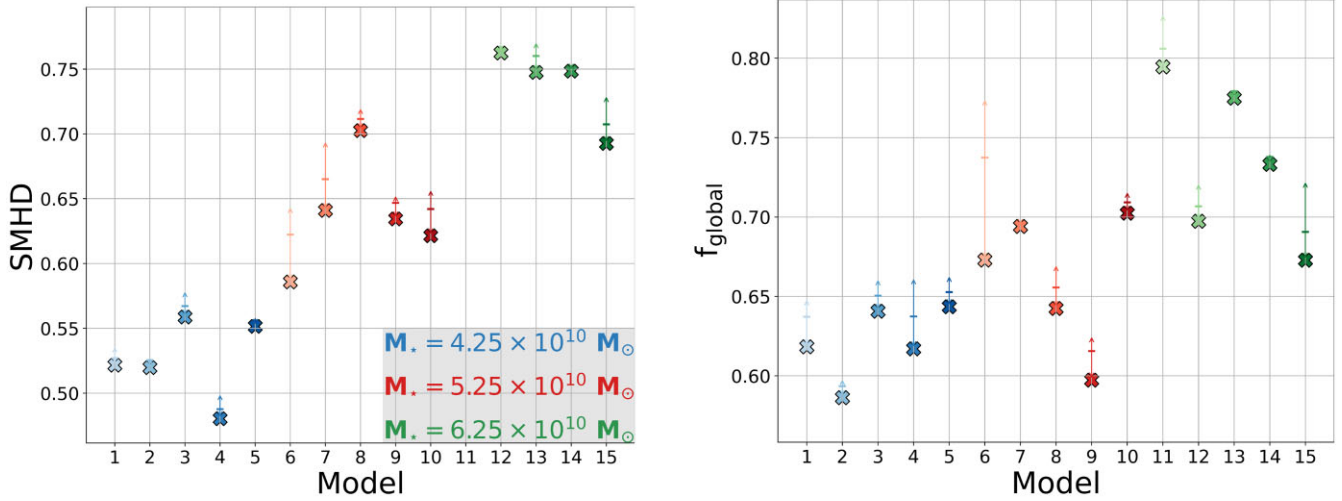
Figure 6. Variation of the final SMHD value (left) and global fraction of mismatches (right) for the 15 models explored, at their selected time and viewing angle (from Section. 5.2.1). For both panels, the data points have been colour-coded so that for a fixed initial stellar mass, lighter to darker colours indicate a decrease in the initial bulge fraction: 25 per cent, 20 per cent, 15 per cent, 10 per cent, and 5 per cent. Blue data points indicate models with an initial stellar mass of $M_\star = 4.25 \times 10^{10}$ $M_\odot$, whilst red indicates $M_\star = 5.25 \times 10^{10}$ $M_\odot$ and green $M_\star = 6.25 \times 10^{10}$ $M_\odot$. The arrows indicate the variation of SMHD within each model for the optimal viewing angle $\phi_{obs} \pm 5°$, with the median marked as a horizontal line.
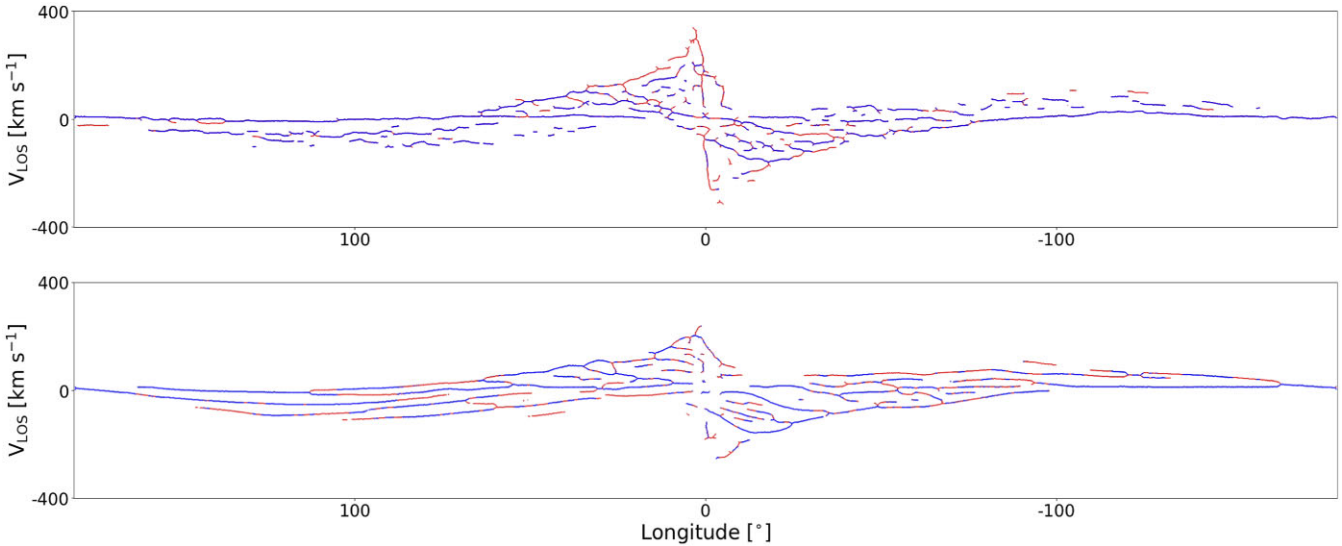
Figure 7. Skeletons extracted from the *l-v* maps of the ^{12}CO observations by Dame et al. (1986) (top) and our Model 4 at a time ~ 2.6 Gyr and angle $\phi_{obs} = 45°$ (bottom). Blue pixels are uniquely associated with each other, whilst red-coloured ones display those defined as mismatches.

regions correspond to those found in the observations, Model 4 struggles to replicate the highest velocities observed in the inner Galaxy.

Although no model is a perfect match to the observations, the SMHD metric and the fraction of mismatches collectively suggest Model 4 as the most promising candidate, with Model 2 closely following. The variability of SMHD values with small changes in viewing angle is minimal, as indicated by the arrows in the left plot of Fig. 6. We further investigate if the selection of a different bar inclination would affect the final best model chosen by the SMHD metric, and we find that Model 4 still has an overall lower value when exclusively looking at inclinations of 20° and 32.5°. Similarly, the variation in the fraction of mismatches with viewing angle is not significantly large, as observed in the right plot of the same figure. Nevertheless, to further refine our assessment of which model best

represents the structure of the MW, we will also incorporate the terminal velocity as an additional evaluative measure.

5.3 Comparison of the terminal velocity

Besides the imprints of the spiral/bar pattern in the *l-v* maps, one other test to deduce how well our models are able to reproduce the observed velocity pattern of the Galaxy, is to compare the terminal velocities that our models reach, versus the observed ones. We do this test solely based on the selected best overall time and viewing angle for each of our models as per the results from the SMHD metric (from Section 5.2.1). As explained in Section 4.4, we obtain the outer edge of the area occupied by the gas in the *l-v* maps, and calculate the difference between the area covered by the models, and that of the
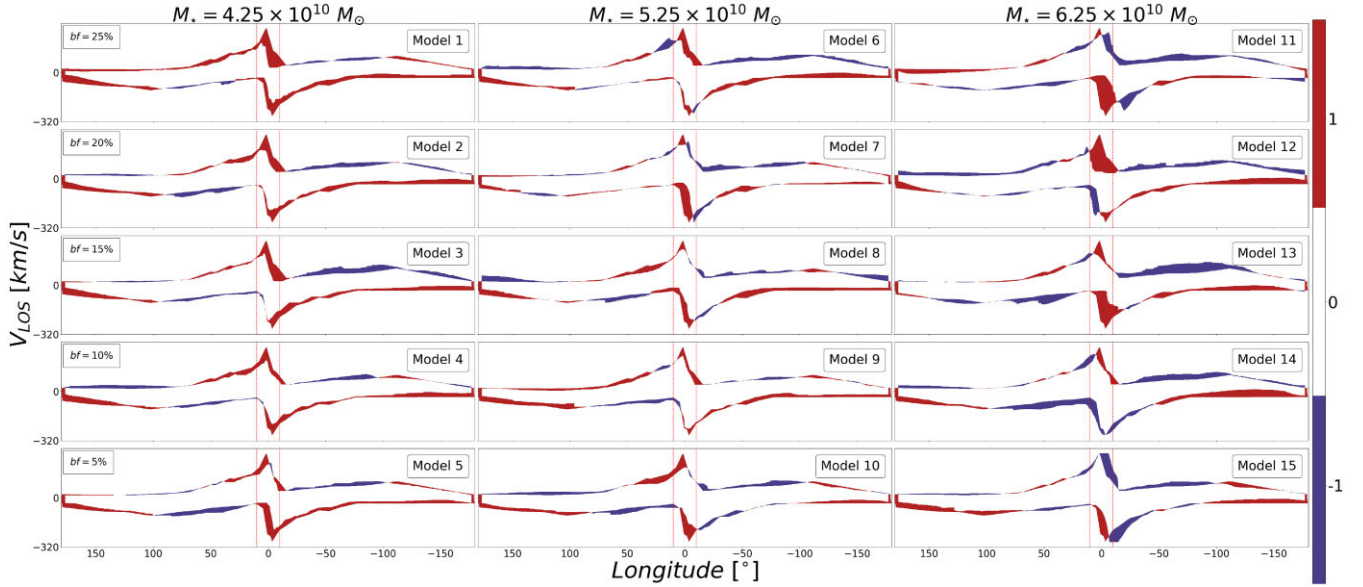
Figure 8. Longitude–velocity diagrams showing the comparison of the terminal velocity of the Dame et al. (2001) ^{12}CO observations combined with H I data from HI4PI Collaboration (2016) with our 15 different models. Columns increase in initial stellar mass from left to right, and rows decrease in bulge fraction from top to bottom. Red indicates pixels present in the observations but not in the simulations, where blue refers to those present in the simulations, but not on the observations. The vertical dashed red lines divide the plots in the inner ($|\,l\,|< 20°$) and outer ($|\,l\,|< 20°$) regions.

observations. Once more, we divide this analysis in overall, inner, and outer galaxy to gauge the weight of the different areas on the overall result.

The results from this analysis are shown in Fig. 8. Here, red pixels indicate those regions from the observed ^{12}CO and H I emission in the *l-v* map that are not present in the *l-v* space of the simulations. On the other side, blue pixels specify the regions in *l-v* of our models that do not appear on the observations. White pixels inside the contours represent the velocity–longitude pairs that appear in both observations and simulations. The division between inner and outer galaxy is shown with the two vertical red dashed lines. A visual comparison of the results can be achieved by looking at those non-white pixels in Fig. 8. The bigger the red/blue areas in the individual images, the more discrepancies that model exhibits with the observations. It is possible to see that models with larger stellar masses tend to have a larger number of non-zero pixels, making these candidates less suitable to reproduce the observations. This result is in accordance with the outcomes from the SMHD metric. A further discussion can be seen in Appendix C.

As explained in Section 4.4, to quantify this difference, we estimate the angle-averaged number of non-zero pixels, $\langle m \rangle = m/\Delta l$, where Δl is the longitude range of the region used for the statistics. The angle-averaged number of non-zero pixels for the overall, inner, and outer galaxy, $\langle m_{\mathrm{global}} \rangle$, $\langle m_{\mathrm{inn}} \rangle$, and $\langle m_{\mathrm{out}} \rangle$, are collated in Table 3. Similarly to the results from the SMHD metric, the models with the lowest $\langle m_{\mathrm{global}} \rangle$ are Models 2, 3, 4, and 5. Focusing on the preferred models from the SMHD analysis in Section 5.2.2 (Models 2 and 4), Model 4 presents the lowest values in the overall $\langle m \rangle$ metric, as well as in the inner and outer regions of the galaxy, meaning that it is a better reproduction of the Galaxy velocity profile. Therefore, we conclude that out of our suite of 15 models, Model 4 is consistently the best at reproducing the studied observations of the MW, showing the lowest values of the terminal velocity and SMHD metric. We adopt this as our best model and proceed to further analyse how some specific galactic features compare to those of the MW.

6 CHARACTERIZATION OF THE BEST MODEL

6.1 The Galactic bar

As mentioned in Introduction, there is some debate on the different observational traits of the galactic bar. In this section, we investigate the bar length measured for the models in our sample, with particular focus on the values obtained for our best model. Note, however, that we do this in a simplistic way, as our intent is purely to have an estimate of the overall length, orientation, and pattern speed of the bars that are formed in the models as seen in the gas. Fitting the stellar bar more accurately is beyond the scope of this paper, but on Appendix. D, we show the 3D distribution of the stellar bar, which suggests that it has (at least) two components – a longer thinner bar, and a boxy/peanut bulge-like structure at the centre. In Appendix D, we also show the radial and circular components of the velocity of the stars, showing a very similar pattern to that seen in the gas (see Section 6.3). For consistency, from here onwards we will use the half-length of the bar (3–5 kpc) as a reference for our analyses (see Introduction for further information).

For the purpose of this paper, we will simply use the gas top-down surface density distribution (where the transitions are sharper, see figures from Appendix A), and we use the CARTA software to manually define the bar ends (i.e. the position where the linear bar meets the beginning of the spiral arms) from those top-down gas surface density maps. Table 3 summarizes the results, where errors on the length were obtained via a Monte Carlo error propagation, assuming an initial uncertainty of 500 pc in the definition of the bar ends. The bar half-length ($L/2$) is defined as half the distance between the extremities, and ranges between 2.8 and 4.9 kpc in our barred models. Model 4, our best model, generates a bar with a half-length of 3.1 ± 0.8 kpc, meaning that the bar extends to ~ 3 kpc from the galactic centre. This result is in agreement with recent studies that use data from *Gaia* EDR3, such as Queiroz et al. (2021) or Lucey et al. (2023), which find, using 6D phase-space information, that the Galactic bar extends to about ~ 4 kpc from the Galactic Centre.
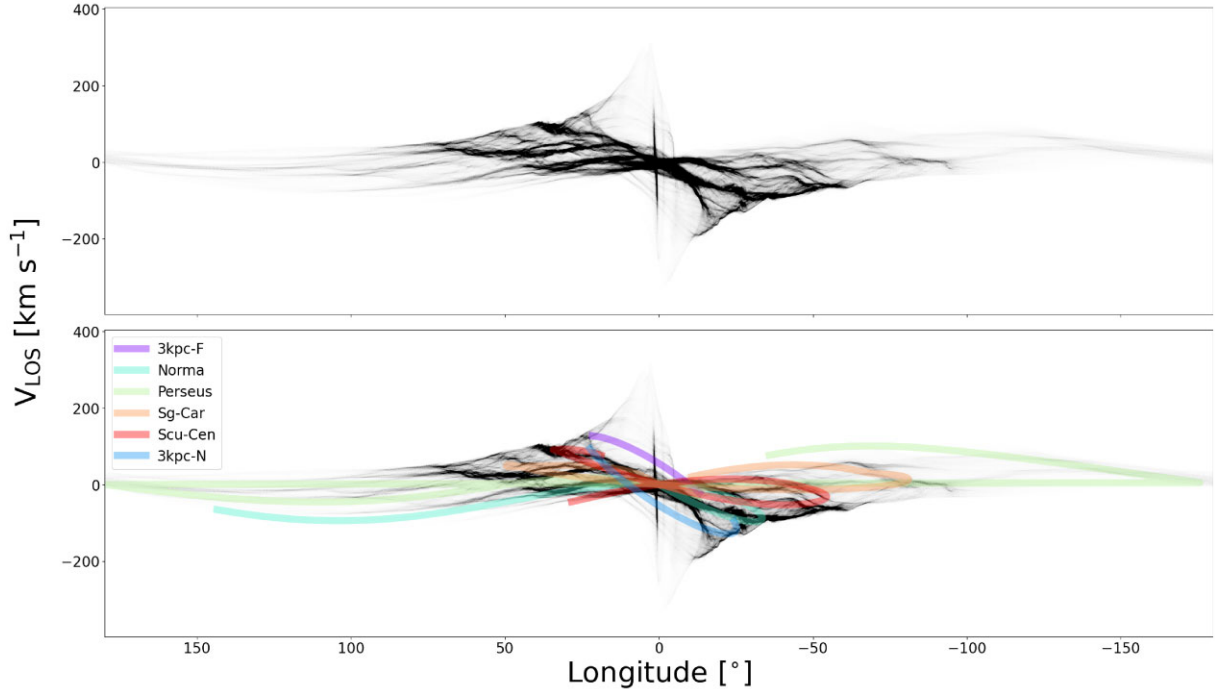
Figure 9. Top: longitude–velocity map for our best model, Model 4, at a viewing angle $\phi_{obs} = 45°$. Bottom: same image, where the spiral arms tracks from Taylor & Cordes (1993) are superimposed. These tracks are displayed with a width of $10\,\mathrm{km\,s^{-1}}$ and the different colours represent different spiral arms: the 3 kpc-arm is represented in blue for the near and purple for the far counterparts; and light blue, light green, yellow, and orange represent the Norma-Outer, Perseus, Sagittarius-Carina, and Scutum-Centaurus arms.

To determine the pattern speed of the bars in our simulations, we analyse the change in angles derived from the WPCA analysis (see Section. 4.1) at three distinct times: the optimal time identified via the SMHD metric, and two additional times, at 50 Myr before and 50 Myr after the optimal time. This approach provides us with three measurements for the pattern speed. The final pattern speed is calculated as the mean of these values. The associated error is estimated based on the standard deviation of these measurements. The results for the pattern speed (Ω) can be seen in Table 3, where these values range between $20.8 \pm 0.8\,\mathrm{km\ s^{-1}\ kpc^{-1}}$ to $38.2 \pm 0.4\,\mathrm{km\ s^{-1}\ kpc^{-1}}$, with a value of $30.0 \pm 0.2\,\mathrm{km\ s^{-1}\ kpc^{-1}}$ for our best model (Model 4). Values deduced from observations are in between $30 - 40\,\mathrm{km\ s^{-1}\ kpc^{-1}}$ (e.g. Clarke & Gerhard 2022; Gaia Collaboration 2023), thus our best model is within the observed range.

6.2 The spiral pattern

As a means to visually compare the spiral pattern formed in our best model, to some of the most commonly used spiral arm tracks of the MW, we overplot on our best model, the spiral track models by Taylor & Cordes (1993) updated by McClure-Griffiths et al. (2004), which are used by several MW molecular gas surveys (e.g. Colombo et al. 2022). We do so both in the *l-v* diagram as well as the top-down view of the Galaxy. The *l-v* map at the best viewing angle of $\phi_{obs} = 45°$ and the $20 \times 20\,\mathrm{kpc^2}$ face-on view of Model 4 can be seen in Figs 9 and 10, both with and without the spiral tracks from Taylor & Cordes (1993) superimposed in colour. From these plots, we can see that the spiral arm loops (in *l-v* space) not always match exactly the position of the loops in the model which could be partly a consequence of us sampling the viewing angle only every 12.5°.

The spiral tracks from Taylor & Cordes (1993) are also idealized log spirals with specific pitch angles (with the exception of the kinks closer to the Sun), which is also not always a perfect match to the pattern seen in molecular gas tracers. Indeed, perfect log-spirals are not reproduced in our live models, as our spiral pattern is more transient and dynamic in nature. Furthermore, the specific choice of spiral arms from different observational works would place the tracks in slightly different positions (e.g. Reid et al. 2019). Similarly to the *l-v* space, we can see that some spiral arm tracks in the top-down view are well matched, while others have some small shifts with respect to the position of our spiral arms. Nevertheless, we note that the top-down projection of these tracks from Taylor & Cordes (1993) are highly uncertain, thus should only be taken as purely illustrative.

Nevertheless, even though we do not try to match our models to these idealized tracks (as we do the comparison directly on the observational data, without assuming any specific model of the spiral pattern), we can see that in general our best model resembles the idealized global structure of the MW.

6.3 Galactic dynamics

We investigate the dynamics generated by our best model by looking at the radial and tangential (circular) velocity components of the gas in our Model 4, V_{rad} and V_{tan}. Note that in Appendix D we also show the radial and tangential velocity components for the stars, showing a very similar pattern to that seen in the gas, albeit with less sharp changes. The top-down view of the simulation with the respective values of the gas velocities can be seen in Fig. 11, where the surface density of the gas has been faded in the background for a visual comparison with the positions of the spiral arms and bar. The solid black lines with numbers represent the cross-section of the spiral arms in the sample that we investigate further down.
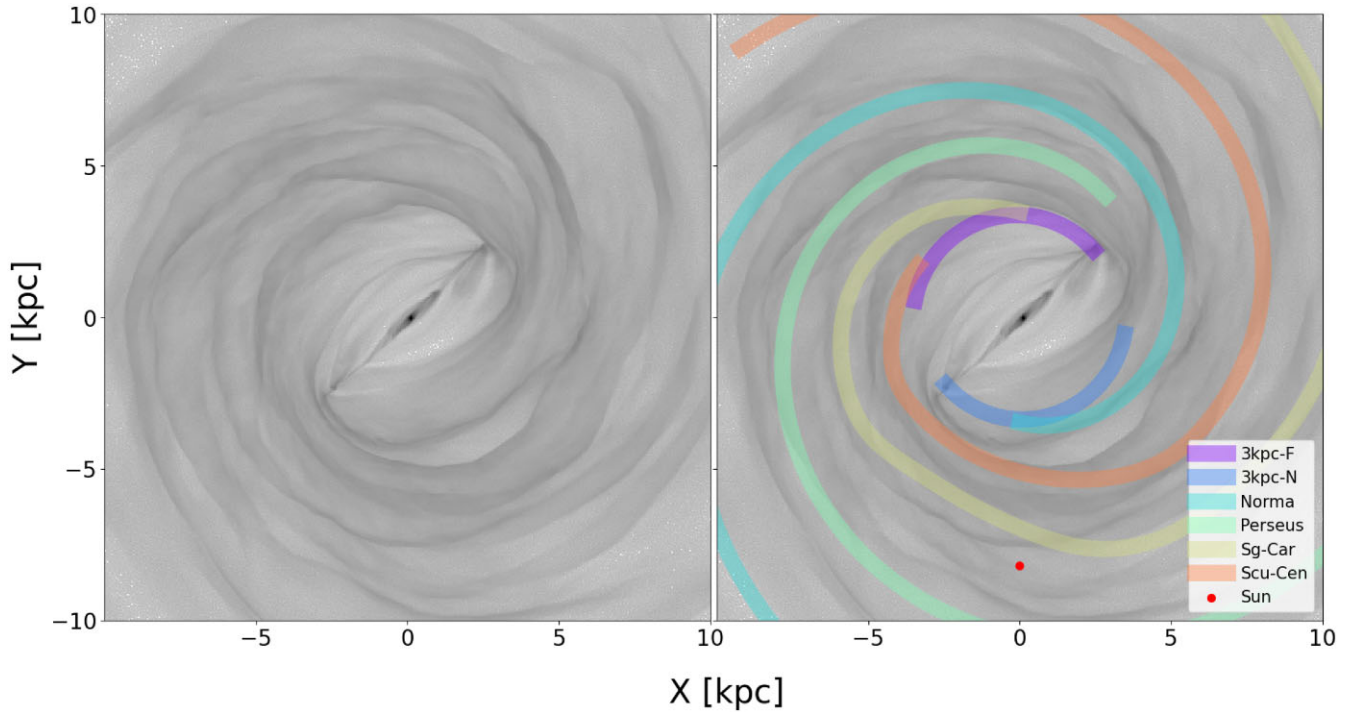
Figure 10. Same as Fig. 9 for the 20 × 20 kpc^2 face-on view of our best model, Model 4. Original image is shown on the left, whilst the spiral tracks from Taylor & Cordes (1993) are superimposed on the right panel. The Sun was positioned at a distance from the Galactic Centre of 8.2 kpc and represented with a red dot.
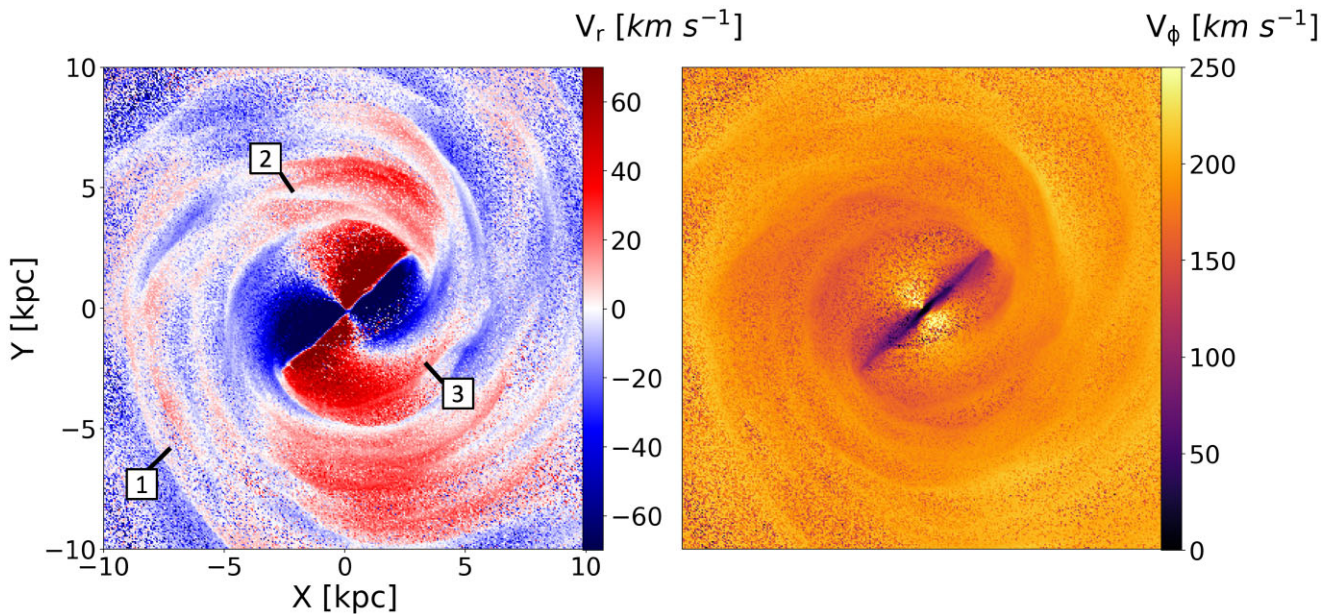

Figure 11. Top-down view of the 20 × 20 kpc^2 faded surface density of Model 4. The left panel is colour-coded with the radial velocity, $V_{\rm rad}$, of the gas cells, whilst the right panel shows the tangential component, V_ϕ. Black solid lines in the former represent the cross-section of the spiral arms analysed in this model.

From the left panel of Fig. 11, we can see that the galactic centre presents a 'butterfly' or quadrupole pattern in radial velocity, typical of barred-galaxies (e.g. Buta & Block 2001; Buta, Laurikainen & Salo 2004; Querejeta et al. 2016; Cuomo et al. 2021). There is a four-quadrant section delimited by the symmetry axes of the galactic bar, where the radial velocities of the gas cells shift from positive to negative values. This feature has directly been observed in the MW using *Gaia* measurements (e.g. Queiroz et al. 2021; Leung et al. 2023). It is also possible to see that the spiral arms are mostly located in regions where there is a drastic change in $V_{\rm rad}$, creating shock fronts where the gas accumulates and creates the spiral arm overdensities, and the gas in the interarm regions is mostly radially static, with very low $V_{\rm rad}$.
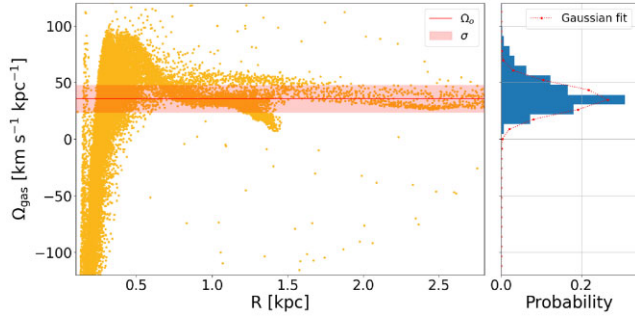
Figure 12. Left: Ω_{gas} against distance from the Galactic Centre (R) for those gas cells that fall within 5 pixels (or 200 pc) from the bar major axis. Here, Ω_{gas} is defined as $V_{tan}(R)/R$. Solid red line indicates the fitted peak of the Ω_{gas} distribution using a Gaussian least-squares fitting for gas cells located at $R > 0.5$ kpc (Ω_o). The red-shaded area represents the dispersion of the data (σ). Right: vertical histogram of the data for Ω_{gas}, with the respective Gaussian fit overlayed as the dotted red line.

From the right panel of Fig. 11, we can see how the tangential velocity of the gas varies in our simulation. The galactic bar in this case produces very low values of V_ϕ along its major axis, as this is where the gas is mostly moving radially. The well-defined spiral arms also show lower values of V_ϕ (thus contributing to reduced shear), whilst the gas in the interarm regions have higher V_ϕ and thus induce larger shear motions from the differential rotation.

Furthermore, we look at how the tangential (circular) velocity of the gas along the bar major axis vary with radial distance from the Galactic Centre (R), by taking the velocities of all gas cells within 200 pc from the bar major axis. Therefore, we define a new quantity Ω_{gas} as $V_{tan}(R)/R$ with dimensions of pattern speed, and we represent it against R in the left panel of Fig. 12. We can see that outside the central-most region (beyond $R > 0.5$ kpc), the circular velocity of the gas becomes constant and with extremely low dispersion, suggesting that the gas along the bar has near-constant angular velocity, thus mimicking a solid body- type of rotation. We can estimate the mean value and its dispersion (σ) by creating a histogram of the data and performing a Gaussian least-squaresfitting. The resulting value for distances larger than $R > 0.5$ kpc indicate $\Omega_{gas} = 35.8 \pm 11.8$ km s^{-1} kpc^{-1}. The solid red line on the left panel of Fig. 12 indicates the fitted peak of the Ω_{gas} distribution, whilst the red-shaded area represents the dispersion of the data (σ). The right panel of Fig. 12 shows the vertical histogram of the data for Ω_{gas}, with the respective Gaussian fit overlayed as the dotted red line. The resulted value for Ω_o is comparable with the values found for the bar pattern speed (see Table 3), suggesting that as the gas travels across the bar major axis, its trajectory changes, affected by the potential generated by the bar, such that it nearly becomes co-rotating with the bar, and travels mostly radially, in x_1-type orbits, along the length of the bar.

Lastly, we investigate the streaming motions around spiral arms from analysing the radial velocity shown in the left panel of Fig. 11. We examine how the density of the gas and the radial velocity changes along a radial line that crosses three spiral arms, at various distances from the galactic centre. The top-left panel of Fig. 13 shows the top-view distribution of the radial velocity of Model 4 with black lines indicating the position of the three different spiral arms we look into. Panels enumerated from 1 to 3 show the variation of gas density of all cells (within a height of ± 300 pc from the galactic plane and a radius of ± 15 kpc) along those lines as a function of distance to the galactic centre in dark blue, and the variation of V_{rad} is
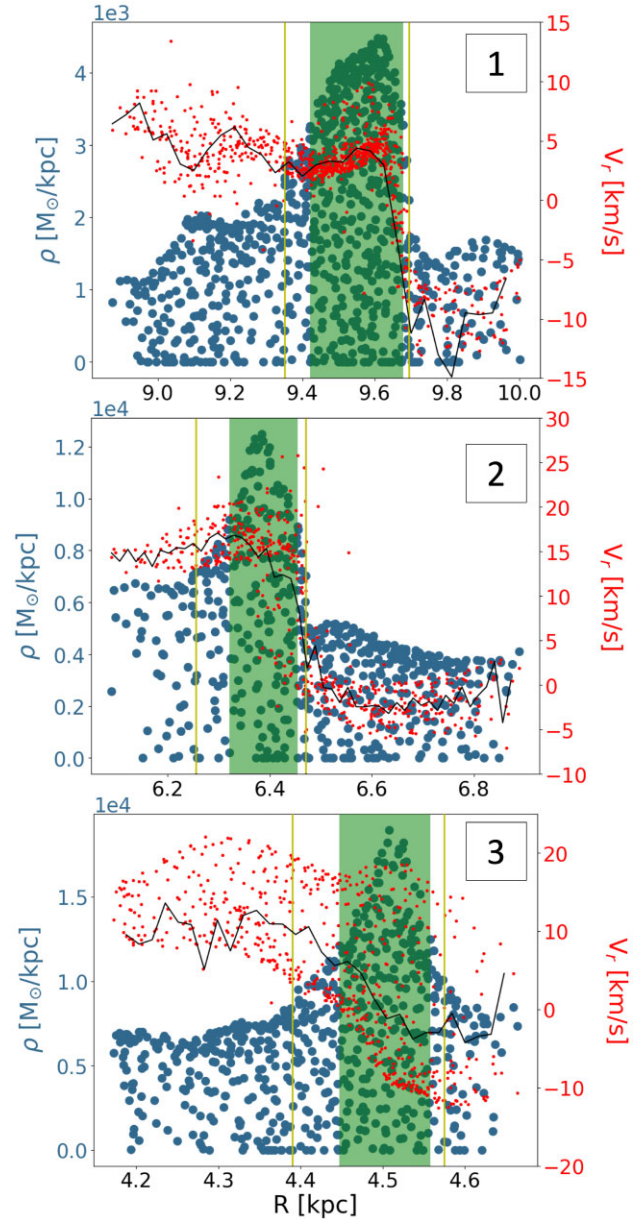

Figure 13. Spiral arm study across three different regions of Model 4 indicated in Fig. 11, decreasing in galactocentric radius from top to bottom. Each individual image shows the variation of gaseous density along the respective lines versus the distance to the Galactic Centre in dark blue, and the variation of V_{rad} in red. The continuous black line represents the running median of the velocity. The yellow lines are a visual indication of the width of the arm, showing where the densities start to rise substantially above the background, whilst the green-shaded area corresponds to the FWHM of the peak in density once subtracted the background.

shown as red dots for gas. The solid black line represents the running median of the velocities. The vertical yellow lines are positioned at the start and end points of the overdensities, corresponding to the width of the arms. The typical change in velocity across the arms can then be calculated as the difference in V_{rad} where the black meets the yellow lines. The green-shaded area corresponds to the full-width half maximum (FWHM) of the peak of gaseous density once substracted the background. The measurements for the amplitude of these streaming motions for the three studied regions, alongside the

Table 4. Spiral arm properties obtained from the density and rotational velocity of gaseous cells crossing three different arms defined in the top-left panel of Fig. 13: R_i denotes the ridge position from the galactic centre of the corresponding arm, ΔV is the change in velocity across the arm, ΔW the width of the arm, and FWHM denotes the full-width half maximum of the peak in density.

Arm slice	R_i (kpc)	ΔV (km s^{-1})	Width (kpc)	FWHM (kpc)
1	9.61	~ 14.4	~ 0.35	0.33
2	6.37	~ 14.0	~ 0.21	0.18
3	4.57	~ 12.5	~ 0.18	0.17

galactocentric distance of each arm segment studied, are collected in Table 4. We can see that the typical change in radial velocity crossing a spiral arm for all three cases ranges in between $\sim 12 - 15$km s^{-1}. However, the width and FWHM of these arms is larger for the larger galactocentric distances. The similar velocity shift but across larger cross-sections, results in effectively a lower velocity gradient across the arm, and thus resulting in weaker shocks and spiral arms that are less well defined at larger galactocentric distances. This result is only tentative, given the very low number statistics and possibly biased to our specific selection of spiral arm segments. However, it does suggest that the strength of the spiral arm shocks, from the streaming motions, varies as a function of position in the galaxy, and this could potentially go on to regulate the amount of dense gas (and thus regulate the conversion of atomic to molecular gas) that takes place in the arms. In order to determine what defines the strength of the spiral shock, we would need to do this across entire spiral arms. Indeed, from Fig. 13, we can see that some spiral arms segments have strong converging motions with gas coming into the spiral arms from both sides (i.e. when we have the convergence of radially inflowing and radially outflowing gas), while other arm segments have gas entering the arm only from one direction and crossing it (i.e. either always radially moving outwards, or inwards). It would thus also be interesting to investigate if and how these dynamics may influence the ability of molecular gas to accumulate and form stars within the arms, as some extragalactic studies seem to suggest that regions with negative or positive torques (i.e. radially inflowing or outflowing gas) are not equally efficient at forming stars (Meidt et al. 2013). Doing this more comprehensive study is well beyond the scope of this paper, but is an interesting avenue of follow-up work. Nevertheless, this tentative result is consistent with observations that the MW's spiral arms are better defined, with more dense and molecular gas, at lower galactocentric radii, compared to the outer arms (e.g. Clemens 1985; Bovy et al. 2015; Khanna et al. 2023).

6.4 The Galactic centre

We now proceed to look at the features present in the innermost parts of our best model and compare them to the ^{12}CO observations from Dame et al. (2001), as well as different theoretical models. We show a zoom-in view of the central region of our Model 4 in Fig. 14. The two left panels show a zoom-in of the central 6×6 kpc^2 box of the top-down view of the simulation on top, and the 2×2 kpc^2 box on the bottom. The Sun is positioned at $(0, -8.2)$ kpc. The middle two panels show the *l-v* maps of Model 4 at a viewing angle of $\phi_{\rm obs} = 45°$. Top panel shows the entire range, whilst we zoom-in to the inner $|l| = 15°$ on the bottom panel. The right two panels show the ^{12}CO observations in the same range as the middle panels.

As mentioned by Binney et al. (1991), in our MW, gas following x_1 orbits (parallel to the bar major axis) in the outer parts of the

bar can transfer into x_2 types of orbits (perpendicular to the bar major axis) when approaching the Galactic Centre. Dust lanes are dark streaks or filaments that have been observed in external barred galaxies (e.g. Athanassoula 1992b), and also identified in the ^{12}CO emission in the MW (e.g. Cohen & Davies 1976; Fux 1999; Marshall et al. 2008). The study by Pfenniger & Friedli (1991) proposed that the formation of dust lanes in barred galaxies like the MW could be attributed to the intersection of x_1 and x_2 orbits. According to this model, the collision of gas clouds in these orbits creates shock waves that compress the gas and dust into a thin lane along the leading edge of the bar. More recent simulations by Athanassoula (2016) suggest that while x_1 and x_2 orbits can play a role in shaping the gas distribution in barred galaxies, the formation of dust lanes may be more complex and involve other physical processes, such as gas inflow and turbulence.

In our case, we are able to see shock fronts that produce higher densities both in the top view and the *l-v* image of our models in the two middle panels (marked with arrows). We can also distinguish the interbar material with mostly non-circular motions, with gas inflowing and outflowing with respect to the Galactic Centre, influenced by the galactic bar (see also left panel of Fig. 11). We can also distinguish a few 'connecting arms' that link the two extremities of the bar, and that produce an *l-v* signature similar to the observed expanding 3-kpc arms (far and near components). However, we cannot distinguish x_2-type orbits in the inner parts of our Model 4 (see Fig. 14). These type of orbits are a result of the bar dynamics and the bulge size and mass (see e.g. Bureau & Freeman 1999). Some of our heavier-disc models like 6, 11, and 12, present these features in the top-view images (see Appendix A). This suggests that we may be underestimating the total potential felt by the gas in the inner parts of the galaxy model, perhaps simply as a consequence of not including gas self-gravity (thus the gas does not feel its own potential), or maybe due to our usage of a gas surface density profile that purely follows the stellar one, rather than one that matches the observed gas profile. Another possible factor could be the slight underestimation of stellar mass in the centre, although we do not think this should be the main cause, given that we do explore models with larger central stellar masses, and they are worse matches to the overall structure (see also Appendix D for a more detailed discussion of the stellar distribution).

These type of orbits may appear once we have higher resolution simulations which include gas self-gravity and feedback added to the simulation, or with small adjustments to the gas surface density profile, in order to incorporate more material in the galactic centre – this will be tested in future work.

Indeed, Ridley et al. (2017) studied the gas flow in the inner Galaxy using hydrodynamical isothermal simulations under an imposed potential, and they found a central disc of gas following x_2 orbits. Following up on their work, Sormani et al. (2018) generated higher resolution (target mass of 100M$_\odot$ compared to our 1000M$_\odot$) non-isothermal simulations under the same potential, where they included chemistry, and those simulations better reproduced these orbits, as well as explained the observed asymmetry of the Central Molecular Zone (CMZ).

Another option that could explain the lack of x_2-type orbits would be the position of the Lindblad resonances. Indeed, x_2-type orbits' size and extent are known to be influenced by the location of the inner Lindblad resonance (ILR) in the gravitational potential, as per the epicyclic approximation (e.g. Contopoulos & Grosbol 1989; Athanassoula 1992a). To further investigate this issue, the top panel of Fig. 15 illustrates the rotation curve derived from the gravitational potential (Φ_o) using the epicyclic approximation:
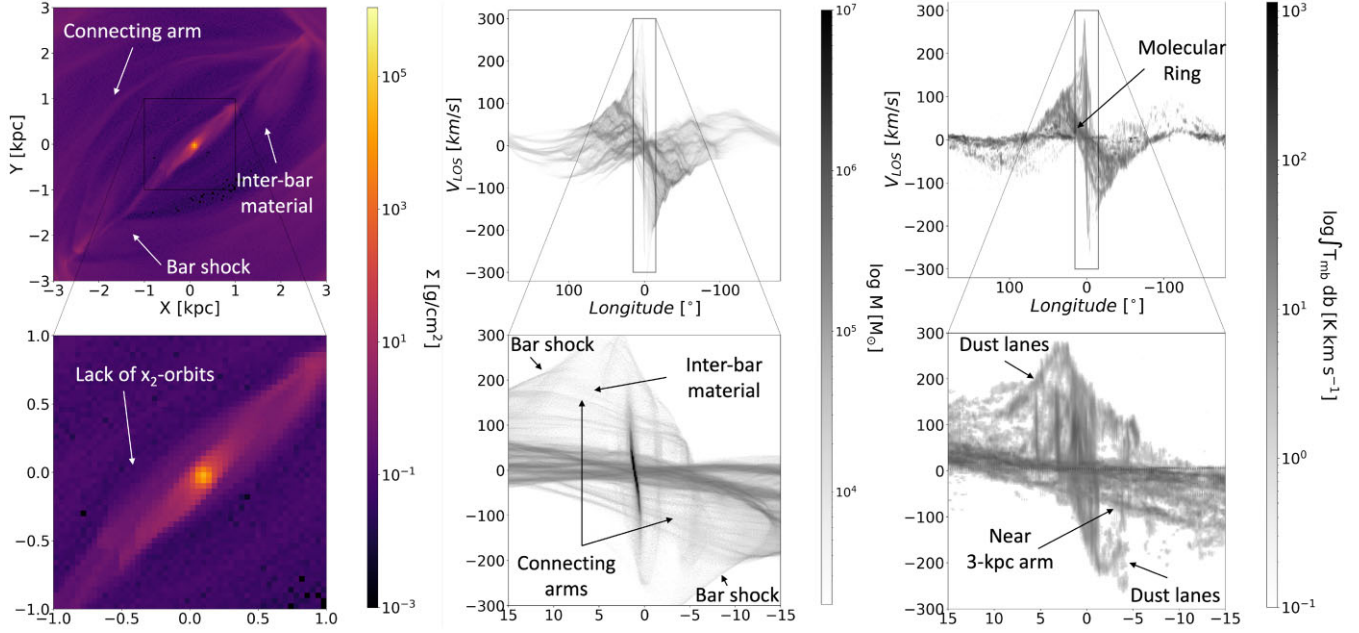
Figure 14. Top left: zoom-in to the central 6×6 kpc^2 box of the top-down view of Model 4 at 2.6 Gyr, where the Sun is positioned at (0,8.2) kpc. Bottom left: zoom-in of the 2×2 kpc^2 box. Top middle: longitude–velocity map of Model 4 at a viewing angle of $\phi_{obs} = 45°$. Bottom middle: zoom-in of the *l-v* map of the inner $|l| = 15°$. The proposed 'dust lane' positions as well as the ring-like structure are marked with arrows. Top right: ^{12}CO *l-v* map from Dame et al. (2001). Bottom right: zoom-in to the observed $|l| = 15°$.
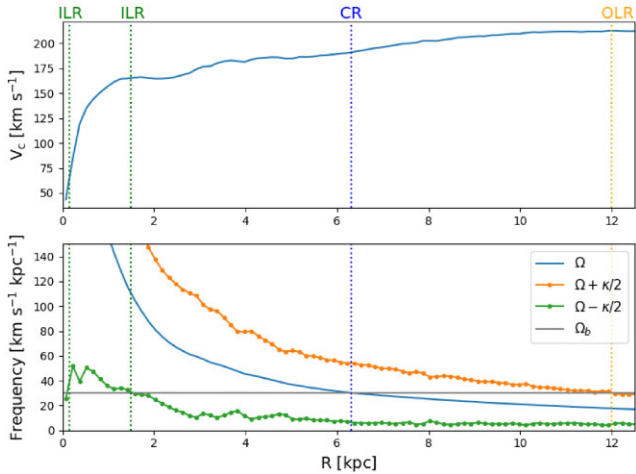
Figure 15. Top panel: the solid blue line represents the circular velocity curve as obtained from the gravitational potential in the epicyclic approximation: $V_c(R) = \sqrt{R d\Phi_o/dR}$. Bottom panel: curves for Ω in blue, $\Omega \pm \kappa/2$ in orange and green respectively. κ represents the epicyclic frequency, $\kappa = (2\Omega/R)(d(\Omega R^2)/dR)$. The black solid line represents the value of the bar pattern speed. The ILRs occur at the intersections of the horizontal value of Ω_p with the $\Omega - \kappa/2$ curve, CR at the intersection with Ω, and OLR with the $\Omega + \kappa/2$ curve. They have been indicated with a horizontal line across the two subplots and colour-coded in accordance with the intersected curve.

$V_c(R) = \sqrt{R d\Phi_o/dR}$. The pattern speed together with the bar length and rotation curve suggests a slow bar in Model 4 with CR radius at ~ 6.5 kpc, which we denote in Fig. 15 with a blue vertical dashed line.

In the bottom panel, we present a resonance diagram where Ω, the rotation velocity in the inertial frame, is shown in blue, and κ is the epicyclic frequency calculated as $\kappa = (2\Omega/R)(d(\Omega R^2)/dR)$.

The ILRs are identified at the intersections of the horizontal black line, representing the bar's pattern speed Ω_p, with the green $\Omega - \kappa/2$ line. Co-rotation (CR) and outer Lindblad resonance (OLR) are located at the intersections of Ω_p with the blue Ω and the orange $\Omega + \kappa/2$ curves, respectively. This diagram reveals two inner ILRs, the first very close to the galactic centre at approximately ~ 0.2 kpc, and the second at about ~ 1.1 kpc. The existence of ILRs then suggests we ought to expect x_2-type orbits. However, Athanassoula (1992a), in their fig. 8, illustrate that the presence of x_2-type orbits also diminishes with an increase in bar strength. Therefore, it is plausible that in Model 4, the bar's strength is sufficiently high to significantly reduce or even eliminate the extent of x_2-type orbits, which might explain their absence in the gas flow patterns observed. Again, in future work, with simulations with higher resolution and SN feedback, it is possible that the feedback might potentially help break out the strong bar potential, and allow these orbits to form.

The high-velocity peaks that we can see in our *l-v* maps towards $l = 0°$ can be explained as these non-circular motions that increase towards the galactic centre due to the effect of the bar on the kinematics of the gas (e.g. Fux 1999; Pettitt et al. 2014, 2015; Sormani & Magorrian 2015; Li et al. 2016). The strong vertical features that appear in the observations near the inner $|l| = 5°$ are very weak to non-existent in our models. These are directly related to the high-velocity gas along the dust lane that overshoots the CMZ and is a accreted later in time (see e.g. Regan, Vogel & Teuben 1997 for observations in NGC 1530 or Sormani et al. 2018; Hatchfield et al. 2021 for theoretical modelling). When comparing to the observations in Fig. 14, we also note that our model does not reach the ~ 300km s^{-1} peak around $l = 0°$. As our simulation does not yet properly resolve the dynamics of the central part, it cannot reproduce these features.

The bright band of molecular gas crossing the Galactic centre in $l - v$ maps, is sometimes referred to as the 'molecular ring' (e.g. Dame et al. 2001), which was first thought to be due to a resonance

(e.g. Binney et al. 1991; Combes 1996). However, whether this feature actually corresponds to a physical ring is debated, and some works suggest instead that this dynamical feature is composed of the inner parts of spiral arms (e.g. Bissantz, Englmaier & Gerhard 2003; Nakanishi & Sofue 2006; Rodriguez-Fernandez & Combes 2008). In Model 4, we observe a significant amount of material in the inner *l-v* image from Fig. 14, which resembles the observed structure marked in the *l-v* map by Dame et al. (2001). However, this gas is not actually distributed in a ring-like structure in the top-down view of the galaxy, and the $l - v$ plot does show that this band is composed of multiple tracks overlaid. This supports the idea that the observed $l - v$ feature might be, in fact, arising from the inner segments of spiral arms, connected to the bar tips (e.g. Dobbs & Burkert 2012).

Given that our models are not specifically tailored to mimic the Galactic Centre of the MW, it is reassuring that our best overall model is in fact able to reproduce most of the observable features of the inner Galaxy, even though it inherits the more dynamic/uncontrolled nature of a live/dynamic stellar potential. In addition, comparing the gas flow patterns in Model 4 with those in the study by Li et al. (2022) reveals striking similarities. The close resemblance with our best model provides a reassuring validation of our results. The consistency between the two models underscores the robustness of our approach in simulating the gas dynamics of the MW. This model will therefore serve as our base model of the MW for future work, which will involve the inclusion of more sophisticated physics, including gas self-gravity, cooling and heating of the ISM, chemistry, SF, supernova and stellar feedback, and ultimately magnetic fields, all at higher resolution.

7 SUMMARY AND CONCLUSIONS

We have used the moving-mesh AREPO code with live stellar and dark matter potentials and isothermal conditions for the gas, in order to produce a suite of numerical simulations whose aim was to recover the large-scale structure of the MW. The initial conditions have been built using the MAKENEWDISK code, whose parameters were largely based on the best-fitting model of Pettitt et al. (2015). However, as in their study they do not obtain a bar in the galactic centre, we create a set of 15 different models varying in initial stellar mass and stellar bulge fraction.

We compare our suite of simulations with ^{12}CO observations from Dame et al. (2001) using *l-v* maps, where we explore the galactic features that would be observed at different times, by placing the observer inside the model galaxy, and varying the Sun position using a range of angles around the galactic centre. From the *l-v* diagrams, we obtain the skeletons of the main structures (Section 4.2), and use the SMHD metric and fraction of mismatches, to obtain a best fit in time and ϕ_{obs} for each of the models. This analysis suggest that the set of models with lower total stellar mass are generally better fits, with Models 2, 3, 4, and 5 providing the best fits, with a slight preference for Models 2 and 4. We then analyse the terminal velocity of our numerical simulations compared to the combined ^{12}CO and H I observations of the MW. This analysis suggests that Model 4, is our global best-fitting model, at a time ~ 2.6 Gyr, when viewed through an observer placed at $\phi_{obs} = 20°-45°$ inclination with respect to the galactic bar.

We then proceed to characterize the details of some of the features of our best model in order to verify its ability to reproduce specific features of the MW. We first look at the galactic bar and find a bar with a half-length of 3.1 ± 0.8 kpc, comparable to the results from recent studies using the *Gaia* EDR3 data, where they find a bar with a half-length of ~ 4 kpc (see e.g. Queiroz et al. 2021). The pattern

speed for our best model is found to be $\Omega = 30.0 \pm 0.2$ km s^{-1} kpc^{-1}, which is also within the observed range of 30–40 km s^{-1} kpc^{-1}.

We also looked at the dynamics of our models by generating the top-down view of the radial and tangential velocities of the gas cells. The radial velocities show a 'butterfly' or quadrupole pattern in the central region, typical of barred galaxies. We then analyse the amplitude of the streaming motions around the spiral arms of the model, by looking at the change in radial velocity across sections of a selection of spiral arms at different positions. This analysis reveals that the change in velocity that gas cells experience across the arms is around $\sim 12 - 15$ km s^{-1}, and the arm widths are larger for the arm segments at larger radial distance. Although from the low number statistics we cannot conclude that this galactocentric trend holds throughout the entire galaxy, this result does suggest that the effective velocity gradient in the arms (and thus the strength of the spiral arm shocks) varies across the galaxy, and that the enhancement in the densities in the arm is correlated with the strength of the velocity gradient. This effect could potentially go on to regulate the conversion of atomic to molecular amount in the arms, in different parts of the galaxy, but we do not include chemistry in these models to test this. The tentative trend of weaker spiral shocks at larger galactocentric distances that we see in our model is consistent with observations that find that the spiral arms in the MW are denser and more molecular in the inner portions of the Galaxy, compared to the outer arms.

Finally, we look at the main structures of the galactic centre of our best model and compare them to the ^{12}CO observations, as well as a number of theoretical models. We look at the inner 6×6 and 2×2 kpc^2 inner top-down views of our simulation, as well as the inner $|l| = 15°$ of both observed and numerical *l-v* maps. We are able to identify the bar shocks, the interbar material with highly non-circular orbits, a structure in the *l-v* map that resembles that of the observed so-called molecular ring that in fact correspond to the inner sections of the spiral arms starting at the tip of the bar, and lastly we can also detect a few 'connecting arms' which produce *l-v*-tracks comparable to the 3-kpc expanding arm from observations. We do not reproduce with clarity the x_2 orbits on which gas in the CMZ in the inner few hundred parsecs of the Galaxy flows. Even though there are two ILRs in the inner $0.1 - 1.1$ kpc, the resonance is too weak to support x_2 orbits in the non-axisymmetric bar potential, suggesting that we might lack mass in the inner disc (see e.g. Athanassoula 1992a; Sormani et al. 2018).

Given that our models are not specifically tailored to mimic the exact potential of the MW, nor do they include the evolution history of the MW and its past or present interactions (which could be responsible for some of the large-scale galactic features), it is remarkable that from our best overall model, which inherits the more dynamic/uncontrolled nature of a live stellar potential, is capable of reproducing many of the observable features of the Galaxy, including the global spiral arm pattern, terminal velocity, and bar properties, as well as more specific features such as the amplitude of the streaming motions around spiral arms, and the dynamics of the connecting arms and bar shock material in the centre of the galaxy. The two main caveats of our model in terms of being a good reproduction of the MW are the slight underestimation of the overall rotation curve, and the lack of x_2-type orbits in the very centre of the galaxy (which might be solved with the inclusion of self-gravity, ISM chemistry, and SN feedback in future work). This model will thus serve as the base configuration for more sophisticated modelling of the ISM in an MW-like galaxy – by including more physical processes such as SNe and stellar feedback, chemistry, sink particles, and magnetic fields. A model that is capable of reproducing most of the MW observable structure self-consistently, is therefore a powerful tool

to gain insight into what might be regulating SF in our Galaxy, giving the opportunity to increase the resolution and zoom-in to individual molecular clouds and star-forming cores in a different range of environments within an MW-like model.

DATA AVAILABILITY

The maps for the top-down surface density, and *l-v* projection of our best model (Model 4) are publicly available in https://dx.doi.org/10.11570/24.0088, as well as in the FFOGG (Following the Flow of Gas in Galaxies) project website (https://ffogg.github.io/)

REFERENCES

Arzoumanian D. et al., 2011, A&A, 529, L6
Athanassoula E., 1992a, MNRAS, 259, 328
Athanassoula E., 1992b, MNRAS, 259, 345
Athanassoula E., 2016, in Astrophysics and Space Science Library, Vol. 418, Galactic Bulges. Springer International Publishing, Switzerland, p. 391
Baba J., Morokuma-Matsui K., Saitoh T. R., 2017, MNRAS, 464, 246
Bailey S., 2012, PASP, 124, 1015
Benjamin R. A. et al., 2005, ApJ, 630, L149
Bertram E., Glover S. C. O., Clark P. C., Ragan S. E., Klessen R. S., 2016, MNRAS, 455, 3763
Binney J., Merrifield M., 1998, Galactic Astronomy. Princeton University Press, Princeton, NJ
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press, Princeton, NJ
Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, MNRAS, 252, 210
Binney J., Gerhard O., Spergel D., 1997, MNRAS, 288, 365
Bissantz N., Gerhard O., 2002, MNRAS, 330, 591
Bissantz N., Englmaier P., Gerhard O., 2003, MNRAS, 340, 949
Blitz L., Spergel D. N., 1991, ApJ, 379, 631
Bovy J., Bird J. C., García Pérez A. E., Majewski S. R., Nidever D. L., Zasowski G., 2015, ApJ, 800, 83
Bovy J., Leung H. W., Hunt J. A. S., Mackereth J. T., García-Hernández D. A., Roman-Lopes A., 2019, MNRAS, 490, 4740
Bureau M., Freeman K. C., 1999, AJ, 118, 126
Buta R., Block D. L., 2001, ApJ, 550, 243
Buta R., Laurikainen E., Salo H., 2004, AJ, 127, 279
Cabrera-Lavers A., Hammersley P. L., González-Fernández C., López-Corredoira M., Garzón F., Mahoney T. J., 2007, A&A, 465, 825
Castro-Ginard A. et al., 2021, A&A, 652, A162
Cautun M. et al., 2020, MNRAS, 494, 4291
Clark P. C., Glover S. C. O., 2015, MNRAS, 452, 2057
Clark P. C., Glover S. C. O., Ragan S. E., Duarte-Cabral A., 2019, MNRAS, 486, 4622
Clarke J. P., Gerhard O., 2022, MNRAS, 512, 2171
Clarke J. P., Wegg C., Gerhard O., Smith L. C., Lucas P. W., Wylie S. M., 2019, MNRAS, 489, 3519
Clarke S. D., Williams G. M., Walch S., 2020, MNRAS, 497, 4390
Clemens D. P., 1985, ApJ, 295, 422
Cohen R. J., Davies R. D., 1976, MNRAS, 175, 1
Cohen R. J., Few R. W., 1976, MNRAS, 176, 495
Colombo D. et al., 2021, A&A, 655, L2
Colombo D. et al., 2022, A&A, 658, A54
Combes F., 1996, in Buta R., Crocker D. A., Elmegreen B. G., eds, ASP Conf. Ser., Vol. 91: Barred Galaxies. Astron. Soc. Pac., San Francisco, p. 286
Combes F., Young L. M., Bureau M., 2007, MNRAS, 377, 1795
Contopoulos G., Grosbol P., 1989, A&AR, 1, 261
Cuomo V., Lee Y. H., Buttitta C., Aguerri J. A. L., Corsini E. M., Morelli L., 2021, A&A, 649, A30
Dame T. M., Elmegreen B. G., Cohen R. S., Thaddeus P., 1986, ApJ, 305, 892
Dame T. M., Hartmann D., Thaddeus P., 2001, ApJ, 547, 792
Dobbs C., Baba J., 2014, Publ. Astron. Soc. Aust., 31, e035
Dobbs C. L., Burkert A., 2012, MNRAS, 421, 2940
Dobbs C. L., Pringle J. E., 2013, MNRAS, 432, 653
Drimmel R., 2000, A&A, 358, L13
Duarte-Cabral A., Dobbs C. L., Peretto N., Fuller G. A., 2011, A&A, 528, A50
Duarte-Cabral A., Acreman D. M., Dobbs C. L., Mottram J. C., Gibson S. J., Brunt C. M., Douglas K. A., 2015, MNRAS, 447, 2144
Duarte-Cabral A. et al., 2021, MNRAS, 500, 3027
Dubuisson M.-P., Jain A. K., 1994, in 12th International Conference on Pattern Recognition, vol. 1. IEEE, Jerusalem, Israel, p. 566
Efremov Y. N., 2011, Astron. Rep., 55, 108
Eilers A.-C., Hogg D. W., Rix H.-W., Ness M. K., 2019, ApJ, 871, 120
Elmegreen D. M., 1985, in Proc. IAU Symp., The Milky Way Galaxy. D. Reidel Publishing Co., Dordrecht, p. 255
Francis C., Anderson E., 2012, MNRAS, 422, 1283
Fux R., 1999, A&A, 345, 787
GRAVITY Collaboration, 2019, A&A, 625, L10
Gaia Collaboration, 2016, A&A, 595, A1
Gaia Collaboration, 2018, A&A, 616, A11
Gaia Collaboration, 2023, A&A, 674, A37
Garrison-Kimmel S. et al., 2018, MNRAS, 481, 4133
Georgelin Y. M., Georgelin Y. P., 1976, A&A, 49, 57
Glover S. C. O., Smith R. J., 2016, MNRAS, 462, 3011
Gómez F. A. et al., 2017, MNRAS, 472, 3722
Grand R. J. J., Springel V., Gómez F. A., Marinacci F., Pakmor R., Campbell D. J. R., Jenkins A., 2016, MNRAS, 459, 199
Grand R. J. J. et al., 2019, MNRAS, 490, 4786
HI4PI Collaboration, 2016, A&A, 594, A116
Hammersley P. L., Garzón F., Mahoney T. J., López-Corredoira M., Torres M. A. P., 2000, MNRAS, 317, L45
Hatchfield H. P., Sormani M. C., Tress R. G., Battersby C., Smith R. J., Glover S. C. O., Klessen R. S., 2021, ApJ, 922, 79
Hernquist L., 1990, ApJ, 356, 359
Iles E. J., Pettitt A. R., Okamoto T., 2022, MNRAS, 510, 3899
Kalberla P. M. W., Kerp J., 2009, ARA&A, 47, 27
Kataria S. K., Das M., 2018, MNRAS, 475, 1653
Kawata D., 2023, in ZenodoGranadaSpain Galactic Bars: Driving and Decoding Galaxy Evolution. p. 26
Kennicutt Robert C. J., 1989, ApJ, 344, 685
Kerr F. J., 1962, MNRAS, 123, 327
Khalil Y., Famaey B., Monari G., Siebert A., 2023, in N'Diaye M., et al., eds, SF2A-2023: Proc. Annual meeting of the French Society of Astronomy and Astrophysics, Strasbourg, France. p. 91
Khanna S., Sharma S., Bland-Hawthorn J., Hayden M., 2023, MNRAS, 520, 5002

Kolpak M. A., Jackson J. M., Bania T. M., Clemens D. P., Dickey J. M., 2003, ApJ, 582, 756

Kulkarni S. R., Heiles C., Blitz L., 1982, ApJ, 259, L63

Lee C. W., Lee H. M., Ann H. B., Kwon K. H., 1999, ApJ, 513, 242

Leroy A. K. et al., 2013, AJ, 146, 19

Leung H. W., Bovy J., Mackereth J. T., Hunt J. A. S., Lane R. R., Wilson J. C., 2023, MNRAS, 519, 948

Li Z., Gerhard O., Shen J., Portail M., Wegg C., 2016, ApJ, 824, 13

Li Z., Shen J., Gerhard O., Clarke J. P., 2022, ApJ, 925, 71

Liang J., Yuen S. Y., 2013, Pattern Recog. Lett., 34, 1519

Liao S. et al., 2019, MNRAS, 490, 5182

Longmore S. N. et al., 2013, MNRAS, 429, 987

López-Corredoira M., 2007, A&A, 469, 471

Lucey M., Pearson S., Hunt J. A. S., Hawkins K., Ness M., Petersen M. S., Price-Whelan A. M., Weinberg M. D., 2023, MNRAS, 520, 4779

Marinacci F., Vogelsberger M., Mocz P., Pakmor R., 2015, MNRAS, 453, 3999

Marshall D. J., Fux R., Robin A. C., Reylé C., 2008, A&A, 477, L21

Martos M., Yañez M., Hernandez X., Moreno E., Pichardo B., 2004, J. Korean Astron. Soc., 37, 199

McClure-Griffiths N. M., Dickey J. M., Gaensler B. M., Green A. J., 2004, ApJ, 611, L145

McMillan P. J., 2017, MNRAS, 465, 76

Meidt S. E. et al., 2013, ApJ, 779, 45

Nakanishi H., Sofue Y., 2006, PASJ, 58, 847

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Nelson D. et al., 2019, MNRAS, 490, 3234

Nguyen Luong Q. et al., 2011, A&A, 529, A41

Oort J. H., Kerr F. J., Westerhout G., 1958, MNRAS, 118, 379

Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaal K., Zhu C., 2016, MNRAS, 455, 1134

Peretto N. et al., 2013, A&A, 555, A112

Pessa I. et al., 2022, A&A, 663, A61

Peters W. L. I., 1975, ApJ, 195, 617

Pettitt A. R., Dobbs C. L., Acreman D. M., Price D. J., 2014, MNRAS, 444, 919

Pettitt A. R., Dobbs C. L., Acreman D. M., Bate M. R., 2015, MNRAS, 449, 3911

Pettitt A. R., Ragan S. E., Smith M. C., 2020, MNRAS, 491, 2162

Pfenniger D., Friedli D., 1991, A&A, 252, 75

Pillepich A. et al., 2018a, MNRAS, 473, 4077

Pillepich A. et al., 2018b, MNRAS, 475, 648

Portail M., Gerhard O., Wegg C., Ness M., 2017, MNRAS, 465, 1621

Queiroz A. B. A. et al., 2021, A&A, 656, A156

Querejeta M. et al., 2016, A&A, 588, A33

Querejeta M. et al., 2021, A&A, 656, A133

Regan M. W., Vogel S. N., Teuben P. J., 1997, ApJ, 482, L143

Reid M. J. et al., 2019, ApJ, 885, 131

Reissl S. et al., 2020, A&A, 642, A201

Ridley M. G. L., Sormani M. C., Treß R. G., Magorrian J., Klessen R. S., 2017, MNRAS, 469, 2251

Rigby A. J. et al., 2016, MNRAS, 456, 2885

Rodriguez-Fernandez N. J., Combes F., 2008, A&A, 489, 115

Sawada T. et al., 2001, ApJS, 136, 189

Schmidt M., 1959, ApJ, 129, 243

Schneider N., Csengeri T., Bontemps S., Motte F., Simon R., Hennebelle P., Federrath C., Klessen R., 2010, A&A, 520, A49

Schuller F. et al., 2021, MNRAS, 500, 3064

Sellwood J. A., Carlberg R. G., 2019, MNRAS, 489, 116

Simpson C. M., Grand R. J. J., Gómez F. A., Marinacci F., Pakmor R., Springel V., Campbell D. J. R., Frenk C. S., 2018, MNRAS, 478, 548

Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014a, MNRAS, 441, 1628

Smith R. J., Glover S. C. O., Klessen R. S., 2014b, MNRAS, 445, 2900

Smith R. J. et al., 2020, MNRAS, 492, 1594

Sofue Y., 2012, PASJ, 64, 75

Sormani M. C., Magorrian J., 2015, MNRAS, 446, 4186

Sormani M. C., Treß R. G., Ridley M., Glover S. C. O., Klessen R. S., Binney J., Magorrian J., Smith R., 2018, MNRAS, 475, 2383

Sormani M. C. et al., 2019, MNRAS, 488, 4663

Sormani M. C., Gerhard O., Portail M., Vasiliev E., Clarke J., 2022, MNRAS, 514, L1

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Springel V. et al., 2018, MNRAS, 475, 676

Stanek K. Z., Mateo M., Udalski A., Szymanski M., Kaluzny J., Kubiak M., Krzeminski W., 1994, ApJ, 429, L73

Steiman-Cameron T. Y., Wolfire M., Hollenbach D., 2010, ApJ, 722, 1460

Taylor J. H., Cordes J. M., 1993, ApJ, 411, 674

Treß R. G., Sormani M. C., Smith R. J., Glover S. C. O., Klessen R. S., Mac Low M.-M., Clark P., Duarte-Cabral A., 2021, MNRAS, 505, 5438

Valenti E. et al., 2016, A&A, 587, L6

Vallée J. P., 2008, AJ, 135, 1301

Wada K., Baba J., Saitoh T. R., 2011, ApJ, 735, 1

Weaver H., 1970, in Becker W., Kontopoulos G. I., eds, Proc. IAU Symp. 38, The Spiral Structure of our Galaxy. Reidel, Dordrecht, p. 126

Wegg C., Gerhard O., 2013, MNRAS, 435, 1874

Wegg C., Gerhard O., Portail M., 2015, MNRAS, 450, 4050

Weiland J. L. et al., 1994, ApJ, 425, L81

Weinberg M. D., 1992, ApJ, 384, 81

Weinberger R. et al., 2017, MNRAS, 465, 3291

Weiner B. J., Sellwood J. A., 1999, ApJ, 524, 112

Whitworth D. J., Smith R. J., Tress R., Kay S. T., Glover S. C. O., Sormani M. C., Klessen R. S., 2022, MNRAS, 510, 4146

Wright M. C. H., Coil A. L., McGary R. S., Ho P. T. P., Harris A. I., 2001, ApJ, 551, 254

Xu Y. et al., 2016, Sci. Adv., 2, e1600878

Xu Y., Hou L.-G., Wu Y.-W., 2018, Res. Astron. Astrophys., 18, 146

Xu Y., Hou L. G., Bian S. B., Hao C. J., Liu D. J., Li J. J., Li Y. J., 2021, A&A, 645, L8

Young J. S., Scoville N. Z., 1991, ARA&A, 29, 581

van de Voort F., Bieri R., Pakmor R., Gómez F. A., Grand R. J. J., Marinacci F., 2021, MNRAS, 501, 4888

APPENDIX A: TOP-VIEW MAPS

Fig. A1 shows the top-down view plots of gas (left) and stars (right) for the inner 10×10 kpc^2 box of our different 15 models at their optimal time, as selected from the SMHD metric (see Section 5.2). The initial stellar mass increases from left to right and the bulge fraction decreases from top to bottom. The colour scale is the same for all panels and represents the surface density in g cm^{-2}. All models, except Model 1, produce a bar in the galactic centre. Bars can be seen in both stellar and gas distributions. The observer's viewing angle differs between models and is explained in more detailed in Section 4.1.
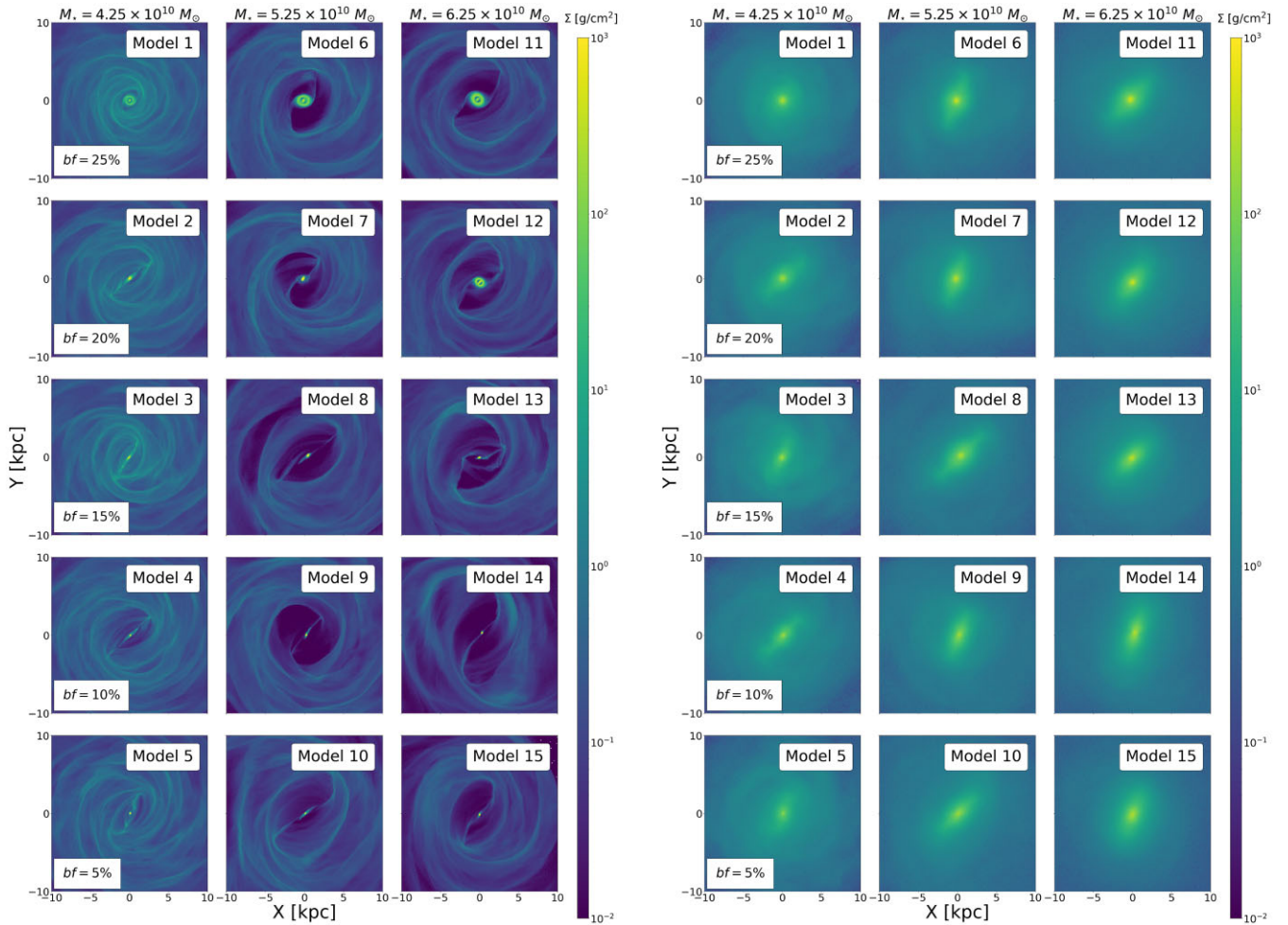
Figure A1. Top-view plots of the inner 10×10 kpc^2 box for the gas distribution (left) and stars (right) of all 15 models that make up our sample, at the optimal time for each model and with the Sun positioned at $(0, -8.2)$ kpc. The initial stellar mass increases from left to right and the bulge fraction decreases from top to bottom. The colour bar represents the surface density in $\mathrm{g\,cm}^{-2}$.

APPENDIX B: FRACTION OF MISMATCHES

The fraction of mismatches that we introduce in Section 4.3 is obtained using equations (10) and (11). The goal of this metric is to quantify the number of structures in the *l-v* maps that our simulations create in excess (or in deficit) when compared to observations. For each pixel in our simulations, a distance to the closest point in the observations is calculated, and vice versa. However, if, for example, the *l-v* map of a given model contains most structures, by construction, it will mean that when looking for the closest point from the observations to that model, the closest points will always have small distances (and therefore giving a low SHMD metric value), while the model is not actually a good representation of the observed structures. Therefore, values for our metric SMHD can be underestimated and favour models that do not necessarily reproduce the structures of the observations.

This effect is illustrated in Fig. B1. Both panels show a zoom-in of the inner $\sim 20°$ region of the Galaxy. The figure shows the skeletons of the *l-v* maps of the ^{12}CO observations by Dame et al. (1986) (left) and our Model 5 at a time ~ 2.3 Gyr and angle $\phi_{\mathrm{obs}} = 150°$ (right). Pixels are colour-coded depending on whether they are unique pairs in blue (i.e. pairs which are common on both directions: sims-to-obs and obs-to-sims), and pixels classified as mismatches in red. We can see that a large number of skeletons in red in the model do not
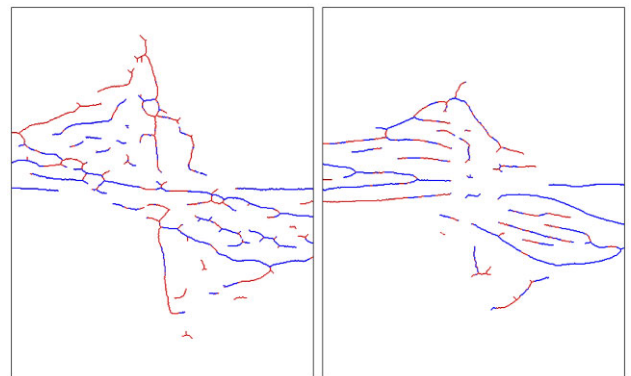

Figure B1. Skeletons extracted from the *l-v* maps of the ^{12}CO observations by Dame et al. (2001) (left) and our Model 4 at a time ~ 2.6 Gyr and angle $\phi_{\mathrm{obs}} = 45°$ (right). Both images show the inner $|l| = 20°$ of the Galaxy. Blue pixels are uniquely associated with each other, whilst red-coloured ones display those defined as mismatches.

correspond to any structures in the observations. These contribute to lowering the SMHD metric, but the model does not represent a good match of the observations. This is the reason why we introduce the extra metric of the fraction of mismatches, where only the pairs of

pixels which are common on both directions (i.e. that appear in both sims-to-obs and obs-to-sims), are considered to be a good match of the underlying structure. The lower the fraction of mismatches, the better the match. We use this metric in conjunction with the SMHD metric (i.e. the best overall fit of the observed structure are the models with low values on both metrics).

APPENDIX C: OUTLINE OF THE *l-v* MAPS

Fig. C1 shows the *l-v* maps of the ^{12}CO and H I emission (top and middle panels, respectively) and our Model 4 at a time of ~ 2.6 Gyr and viewing angle $\phi_{obs} = 45°$ (bottom panel). The selected outline for the terminal velocity has been superimposed on each of the images: red solid line corresponds to the combined outline from both observations, and blue contour shows the silhouette of the *l-v* map of our model (see Section 4.4 for more details on the methodology). Colours have been chosen to match those of Fig. 8 for consistency. Note that in the middle panel of Fig. C1, there is some emission not included within the observational contour at longitude $\sim -80°$ and velocity ~ 100 km s^{-1}. This is H I emission from the Magellanic Clouds and therefore it is excluded for the study of the terminal velocities of the MW.
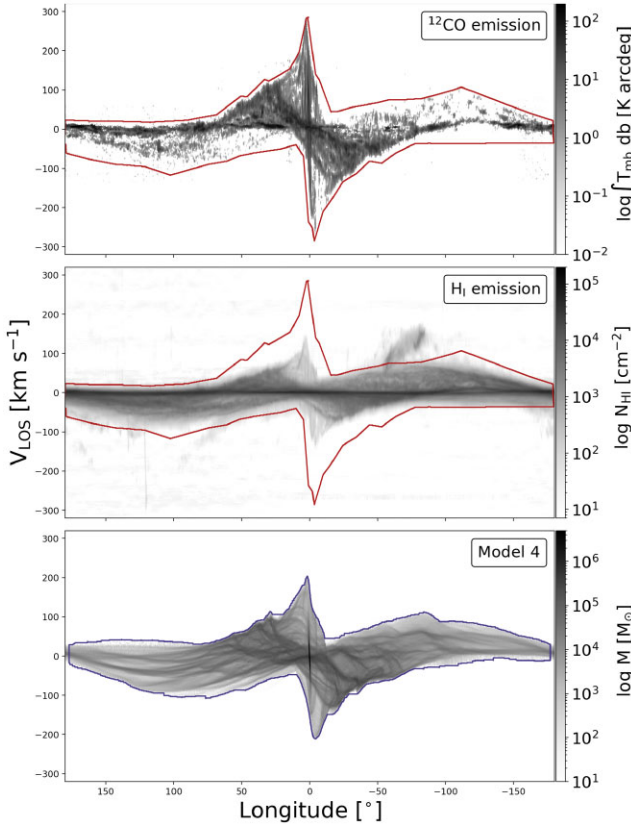

Figure C1. Top: *l-v* map of the ^{12}CO emission from the work by Dame et al. (2001). Middle: *l-v* map of the H I emission extracted from the HI-4PI Survey (HI4PI Collaboration 2016). Bottom: *l-v* map from our Model 4 at a time of ~ 2.6 Gyr. All images have been superimposed with their respective outlines (red continuous line for observations and blue for simulation) used for the analysis of the terminal velocity (see Section 4.4 for methodology and Section. 5.3 for results and discussion).

APPENDIX D: STELLAR DISTRIBUTION

This section presents a brief overview of the main properties of the stellar distribution of our best model. A more in-depth analysis of these and their direct comparison to observational studies will be subject of a follow up study.

Fig. D1 shows the stellar velocity fields in terms of radial and tangential velocity components. We can see that these mimic the pattern seen for the gas (from Fig. 11), albeit with less sharp transitions. We can see that the radial velocity field exhibits a quadrupole pattern, comparable to results found in observations of the MW with APOGEE-Gaia (Apache Point Observatory Galactic Evolution Experiment, Leung et al. 2023) or *Gaia*-DR3 (Gaia Collaboration
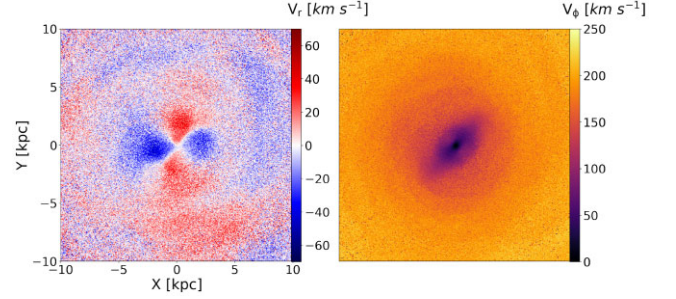

Figure D1. Top-down view of the 20×20 kpc^2 stellar distribution of Model 4 (same as Fig. 11, but for the stars). The left panel is colour-coded with the radial velocity, V_{rad}, of stars, while the right panel shows the tangential component, V_ϕ.
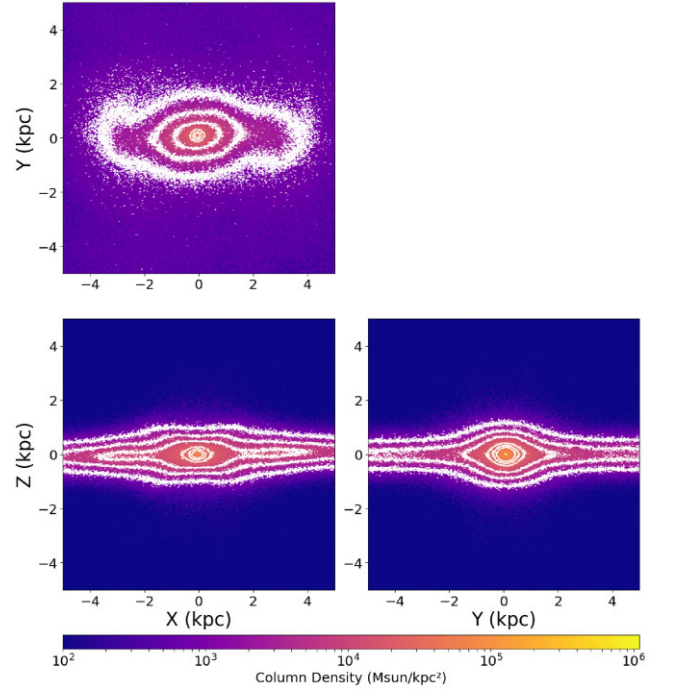

Figure D2. Stellar distribution within a $10 \times 10 \times 10$ kpc^3 region of Model 4 at approximately ~ 2.6 Gyr, showcased in xy (top), xz (bottom-left), and yz (bottom-right) panels. White lines represent iso-density contours at 10 per cent, 20 per cent, 50 per cent, 70 per cent, and 90 per cent of the peak. This shows that the central area exhibits a longer fainter stellar bar (extending to ~ 3 kpc radius), and a more compact ellipsoidal towards the central region, with a flattened profile in the z-direction.

2023), albeit perhaps the pattern in our model has a slightly smaller extent.

Fig. D2 shows the stellar distribution within a $10 \times 10 \times 10 \, \text{kpc}^3$ region of Model 4 at approximately ~ 2.6 Gyr, showcased in xy (top), xz (bottom-left), and yz (bottom-right) panels. White lines represent iso-density contours, aiding in visualizing the distribution. We can see that this model develops a long bar and a more 'boxy-peanut' shape of the inner stellar distribution ($R < 1 - 2 \, \text{kpc}$), in line with some observational studies of the MW stellar bar (e.g. Benjamin et al. 2005; Wegg et al. 2015).

We also estimate the total stellar mass enclosed within a 5 kpc distance from the galactic centre, and up to 1.5 kpc off the mid-plane (equivalent to the observations), and obtain a total enclosed stellar mass of $2.65 \times 10^{10} \, \text{M}_\odot$ for the initial time of Model 4, and $1.98 \times 10^{10} \, \text{M}_\odot$ for the best time. These values are notably close to the observed MW estimates provided by Valenti et al. (2016) of $2.0 \pm 0.3 \times 10^{10} \, \text{M}_\odot$, but slightly lower than those estimated by Portail et al. (2017) of $3.17 \pm 0.24 \times 10^{10} \, \text{M}_\odot$.

APPENDIX E: *l-v* TRACKS ON OBSERVATIONS

Fig. E1 presents the *l-v* map based on the ^{12}CO observations by Dame et al. (2001) arranged similarly to Fig. 9 for visual comparison. Additionally, the bottom panel of Fig. E1 incorporates the same spiral arm tracks obtained from Taylor & Cordes (1993).
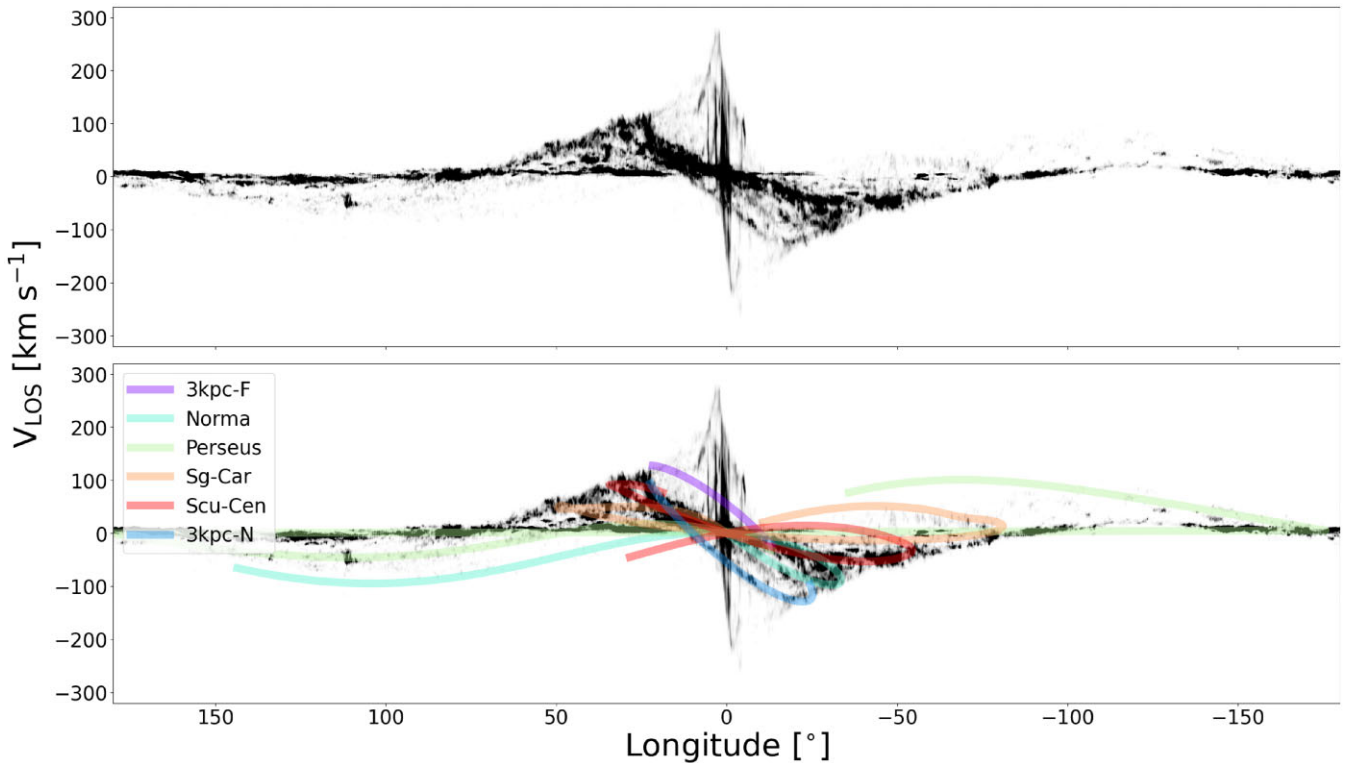

Figure E1. Same as Fig. 9 for the observations. Top: $l - v$ map of the ^{12}CO observations by Dame et al. (2001). Bottom: same image, with the spiral arms tracks from Taylor & Cordes (1993) superimposed. These tracks are displayed with a width of $10 \, \text{km s}^{-1}$ and the different colours represent different spiral arms: the 3 kpc arm is represented in blue for the near and purple for the far counterparts; and light blue, light green, yellow, and orange represent the Norma-Outer, Perseus, Sagittarius-Carina, and Scutum-Centaurus arms, respectively.

This paper has been typeset from a TEX/LATEX file prepared by the author.