

# The structural evolution of galaxies with both thin and thick discs

Michael Aumer <sup>★</sup> and James Binney

*Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, Oxford, OX1 3NP, UK*

Accepted 2017 May 22. Received 2017 May 22; in original form 2017 March 24

## ABSTRACT

We perform controlled  $N$ -body simulations of disc galaxies growing within live dark matter (DM) haloes to present-day galaxies that contain both thin and thick discs. We consider two types of models: a) thick disc initial conditions to which stars on near-circular orbits are continuously added over  $\sim 10$  Gyr, and b) models in which the birth velocity dispersion of stars decreases continuously over the same timescale. We show that both schemes produce double-exponential vertical profiles similar to that of the Milky Way (MW). We indicate how the spatial age structure of galaxies can be used to discriminate between scenarios. We show that the presence of a thick disc significantly alters and delays bar formation and thus makes possible models with a realistic bar *and* a high baryon-to-DM mass ratio in the central regions, as required by microlensing constraints. We examine how the radial mass distribution in stars and DM is affected by disc growth and non-axisymmetries. We discuss how bar buckling shapes the vertical age distribution of thin and thick disc stars in the bar region. The extent to which the combination of observationally motivated inside-out growth histories and cosmologically motivated dark halo properties leads to the spontaneous formation of non-axisymmetries which steer the models towards present-day MW-like galaxies is noteworthy.

**Key words:** methods: numerical - galaxies:evolution - galaxies:spiral - Galaxy: disc - Galaxy: kinematics and dynamics - Galaxy: structure;

## 1 INTRODUCTION

The vertical density profile of stars in the Milky Way (MW) is fitted well by a sum of two exponentials (Gilmore & Reid 1983). Jurić et al. (2008) find scaleheights of  $\sim 900$  pc for the *geometrical thick disc* and  $\sim 300$  pc for the *thin disc*. Similarly, the vertical surface brightness profiles of the majority of bright edge-on spiral galaxies show thin and thick components (Yoachim & Dalcanton 2006). Studies of the chemical abundances of Solar Neighbourhood (Snhd) stars reveal that populations with hotter vertical kinematics and thus larger scaleheights have abundances of  $\alpha$  elements relative to iron ( $[\alpha/\text{Fe}]$ ) that are larger than those of populations with small scaleheights and comparable iron abundances  $[\text{Fe}/\text{H}]$  (Fuhrmann 1998; Bensby et al. 2003).

In the  $[\alpha/\text{Fe}] - [\text{Fe}/\text{H}]$  plane, the low and high  $\alpha$  populations are generally found to separate into fairly distinct sequences. Higher  $[\alpha/\text{Fe}]$  indicates shorter chemical enrichment timescales and age determinations find systematically older ages for stars of the  $\alpha$  enhanced *chemical thick disc*

(Masseron & Gilmore 2015; Martig et al. 2016a): they are generally found to be older than 8 Gyr. The distinct sequences have motivated models in which the two components formed in two temporally separated phases in very different conditions (e.g. Chiappini et al. 1997), but can also be explained as the result of continuous star formation and chemical enrichment (Schönrich & Binney 2009).

It has now become clear that chemical and geometrical definitions of the thick disc yield different results. Whereas Jurić et al. (2008), who determined the density of all stars in the Snhd independent of chemistry, found that the geometrically thick disc has a longer radial scalelength than the thin disc, high  $[\alpha/\text{Fe}]$  stars are found to form a thicker, but more centrally concentrated component than low  $[\alpha/\text{Fe}]$  stars (Bovy et al. 2012; Hayden et al. 2015). A scenario in which the disc forms inside-out and each mono-age population *flares*, i.e. is thicker at outer than at inner radii, could potentially explain these observations (Minchev et al. 2015; Schönrich & McMillan 2017). Flaring can be caused by radial migration of stars (Schönrich & Binney 2012; Roškar et al. 2013) or by vertical heating of the outer disc through

<sup>★</sup> E-mail: Michael.Aumer@physics.ox.ac.uk (MA)

satellite interactions (Kazantzidis et al. 2008) or misaligned gas infall (Jiang & Binney 1999).

Aumer et al. (2016a,b) (hereafter Papers 1 and 2) presented  $\sim 100$  idealised  $N$ -body models of disc galaxies growing within live dark matter (DM) haloes over  $\sim 10$  Gyr. These models covered a large variety of star formation and radial growth histories and most of them followed the assumption that all stars are born on near-circular orbits as in the MW today. Structural and kinematical properties of the MW's thin disc, such as an exponential profile with scaleheight  $\sim 300$  pc or the local age-velocity dispersion relations could be reproduced if giant molecular clouds (GMCs) were included. Additionally, bars of comparable size and structure to that of the MW formed in these models. However, none of the models produced a realistic thick disc. The conclusion was that additional sources of heating were required early in the disc's life, prior to the onset of thin-disc formation.

In this paper, we create models similar to those of Paper 1 which, in addition to a realistic thin disc, also contain an appropriate  $\sim 10$  Gyr old thick disc. Demanding the presence of an old thick component makes it much harder to steer a model to a configuration consistent with current data, in part because the number of observations that need to be explained simultaneously increases roughly two-fold. Moreover, the thick disc must be steered into its present form by adjusting the conditions at the onset of disc formation, which will modify the subsequent formation of the thin disc component by altering spiral and bar structures and their interaction with the dark halo. These changes to the thin disc and dark halo will themselves modify the appearance of the current thick disc. Moreover, the old chemical thick disc of the MW is centrally concentrated and should be important in the central  $\sim 5$  kpc, a region dominated by the Galactic bar (Portail et al. 2017a), a structure that is supposed to have formed from a rather cold disc.

To create a thick disc we follow two approaches. In one scheme, we assume that the velocity dispersions of new-born stars decline with time. This is motivated by observations of redshift  $z_{\text{rs}} \sim 2$  galaxies, which show a high fraction of galaxies with  $H\alpha$  kinematics consistent with ordered rotation, but significantly higher velocity dispersions than today's disc galaxies (Förster Schreiber et al. 2009) and of galaxies at lower redshifts which indicate a continuous decline of gas velocity dispersion with decreasing redshift (Kassin et al. 2012; Wisnioski et al. 2015, but see Di Teodoro et al. 2016 for a different conclusion). Such a declining birth dispersion has also been found in hydrodynamical cosmological simulations of disc galaxy formation (Bird et al. 2013; Grand et al. 2016), and can be understood in terms of galaxies which gradually become less gas-rich and are characterised by gravitationally driven turbulence which decays (Forbes et al. 2012).

In an alternative scheme, we model thin+thick disc systems by creating thick initial conditions and growing thin discs within them. These models are thus representations of two-phase formation scenarios. We here do not model the formation of the thick disc prior to redshift  $z_{\text{rs}} \sim 2$ , but several overlapping scenarios envisage the production of such an object: heating of an initially thinner disc by a merger (e.g. Quinn et al. 1993); formation in early gas-rich mergers (Brook et al. 2004); the formation in a clumpy, turbulent disc (Bournaud et al. 2009). It should be noted that the

latter two scenarios could also be fitted into the picture of declining birth dispersions that constitutes our alternative modelling scheme.

In this paper we concentrate on the setup of our models and on their structural evolution. We explore how models with double-exponential vertical profiles, circular speed curves like that of the MW and bars of appropriate structure can be constructed. We examine how the presence of a thick disc changes the preferred density of the dark halo and influences the evolution of the thin disc and the formation of a bar. We show that different scenarios for the formation of the thick disc leave signatures in the current distribution of age with the  $(R, z)$  plane. A companion paper (Aumer et al. 2017, hereafter Paper 4) focuses on disc heating and radial migration in the models.

Our paper is structured as follows. Section 2 describes the setup and parameters of our simulations. Section 3 discusses the evolution of vertical density profiles and how this shapes the final age structure of the disc. Section 4 analyses the radial distribution of dark and baryonic mass components and Section 5 illustrates the evolution and structure of bars in the presence of thick discs. Section 6 discusses the successes and problems of our models. Section 7 concludes.

## 2 SIMULATIONS

The simulations analysed in this paper are similar to the models presented in Paper 1. These are simulations of growing disc galaxies within non-growing live DM haloes. They are run with the Tree code GADGET-3, last described in Springel (2005). We focus on standard-resolution models, which contain  $N = 5 \times 10^6$  particles in the DM halo and a similar number of particles in the final stellar system. We here rely on collisionless simulations; models which contain an isothermal gas component were discussed in Papers 1 and 2 and shown to differ only mildly from collisionless models in terms of the structure and kinematics of the stellar component. In addition, and crucially, all simulations contain a population of short-lived, massive particles representing GMCs. Papers 1 and 2 demonstrated the importance of GMC heating in creating thin disc components with realistic vertical structure as non-axisymmetric structure contributes little to vertical disc heating (see also Sellwood 2013).

### 2.1 Initial conditions

Table 1 gives an overview of the initial conditions (ICs) of our models. The ICs were created using the GALIC code (Yurin & Springel 2014). We refer the reader to Paper 1 for a full description of the IC creation process and here focus on how the ICs used in this paper differ from each other. All models discussed start with a spherical DM halo with a Hernquist (1990) profile

$$\rho_{\text{DM}}(r) = \frac{M_{\text{DM}}}{2\pi} \frac{a}{r(r+a)^3}. \quad (1)$$

The total mass of all ICs is  $M_{\text{tot}} = 10^{12} M_{\odot}$ , of which  $M_{\text{b},i}$  is  $M_{\text{disc},i} + M_{\text{bulge},i}$  is in a stellar component and the rest  $M_{\text{DM}}$  is in DM. The inner density profile is adjusted so that it is similar to an NFW profile with concentration  $c_{\text{halo}} = 6 - 9$  and thus scale radii are in the range  $a = 30 - 40$  kpc. The

**Table 1.** An Overview over the different ICs and their parameters: *1st Column:* IC Name; *2nd Column:* IC disc mass  $M_{\text{disc},i}$ ; *3rd Column:* IC bulge mass  $M_{\text{bulge},i}$ ; *4th Column:* Number of baryonic particles  $N_{b,i}$  in the ICs; *5th Column:* Concentration parameter for IC DM halo,  $c_{\text{halo}}$ ; *6th Column:* IC DM halo scalelength  $a_{\text{halo}}$ ; *7th Column:* IC radial disc scalelength  $h_{R,\text{disc}}$ ; *8th Column:* IC vertical disc scaleheight  $z_{0,\text{disc}}$ ; *9th Column:* IC bulge scalelength  $a_{\text{bulge}}$ ; *10th Column:* IC bulge flattening  $s$ ; *11th Column:* IC disc velocity ellipsoid ratio  $\sigma_R^2/\sigma_z^2$ ; *12th Column:* IC bulge rotation.

1st Name	2nd $M_{\text{disc},i}$ [ $10^9 M_\odot$ ]	3rd $M_{\text{bulge},i}$ [ $10^9 M_\odot$ ]	4th $N_{b,i}$	5th $c_{\text{halo}}$	6th $a_{\text{halo}}$ [kpc]	7th $h_{R,\text{disc}}$ [kpc]	8th $z_{0,\text{disc}}$ [kpc]	9th $a_{\text{bulge}}$ [kpc]	10th $s$	11th $\sigma_R^2/\sigma_z^2$ disc	12th rotation bulge
Y	5	—	500 000	9	30.2	1.5	0.10	—	—	2.0	—
P	15	—	1 500 000	9	30.2	2.5	1.75	—	—	1.0	—
Q	25	—	2 500 000	6.5	37.9	2.5	1.75	—	—	1.0	—
R	15	—	1 500 000	6.5	37.9	2.5	1.75	—	—	1.0	—
U	25	—	2 500 000	7.5	34.4	2.0	1.70	—	—	1.0	—
T	20	—	2 000 000	9	30.2	2.0	1.70	—	—	1.8	—
C	—	5	500 000	9	30.2	—	—	0.45	1.15	—	no
K	—	15	1 500 000	9	30.2	—	—	1.51	2.0	—	yes
M	—	5	500 000	9	30.2	—	—	1.51	2.0	—	yes
O	—	15	1 500 000	9	30.2	—	—	1.51	3.0	—	yes
V	—	5	500 000	6.5	37.9	—	—	1.51	2.0	—	yes
W	20	5	2 500 000	7.5	34.4	2.5	1.70	0.46	1.15	1.0	no
X	20	5	2 500 000	7.5	34.4	2.0	1.70	0.70	1.15	1.8	no

kinematics of the DM particles are initially isotropic (i.e. have equal velocity dispersions  $\sigma_r = \sigma_\phi = \sigma_\theta$ ). Each DM halo studied here is resolved with  $N_{\text{DM}} = 5\,000\,000$  particles.

ICs contain either a disc or a bulge component or both. IC discs have a mass in the range  $M_{\text{disc},i} = 5 - 25 \times 10^9 M_\odot$  and a density profile of the form

$$\rho_{\text{disc},i}(R, z) = \frac{M_{\text{disc},i}}{4\pi z_{0,\text{disc}} h_{R,\text{disc}}} \frac{1}{2} \text{sech}^2\left(\frac{z}{z_{0,\text{disc}}}\right) \exp\left(\frac{-R}{h_{R,\text{disc}}}\right), \quad (2)$$

where  $h_{R,\text{disc}} = 1.5 - 2.5 \text{ kpc}$  is the exponential disc scalelength. Radially constant isothermal vertical profiles with scaleheights  $z_{0,\text{disc}} = 0.1 - 1.7 \text{ kpc}$  are assumed. The ratios of the vertical to radial velocity dispersions are also radially constant in the IC discs and have values in the range  $\sigma_z^2/\sigma_R^2 = 1.0 - 2.0$ .

IC bulge components are set up with Hernquist density profiles with scalelengths  $a_{\text{bulge}} = 0.4 - 1.5 \text{ kpc}$  which are distorted to be oblate spheroids with axis ratios in the range  $s = 1 - 3$  as

$$\rho_{\text{bulge}}(R, z) = s \rho_{\text{Hernquist}}\left(\sqrt{R^2 + s^2 z^2}\right). \quad (3)$$

They are supposed to model either compact, non-rotating bulges or rotating spheroidal components. Following section 3.2 of Yurin & Springel (2014), the rotation of the axisymmetric bulge components is controlled via the Satoh (1980) parametrization

$$\langle v_\phi \rangle^2 = k^2 (\langle v_\phi^2 \rangle - \sigma_R^2). \quad (4)$$

Non-rotating bulges assume  $k = 0$ , whereas rotating bulges are modelled as isotropic rotators with  $k = 1$ .

Stellar particles in the ICs have a particle mass  $m = 1 \times 10^4 M_\odot$  and our ICs thus contain  $N_{b,i} = 5 - 25 \times 10^5$  stellar particles. As in Paper 1 the applied force softening lengths for the given resolutions are  $\epsilon_b = 30 \text{ pc}$  for baryonic particles (including GMCs) and  $\epsilon_{\text{DM}} = 134 \text{ pc}$  for DM particles.

The first letter of a model's name specifies its ICs according to the scheme laid out in Table 1. In summary:

- (i) Y ICs have a compact and thin baryonic disc;
- (ii) P, Q, R, U, and T ICs contain more extended and thicker discs, with varying size, thickness, mass and DM halo concentration;
- (iii) C ICs have a compact non-rotating bulge;
- (iv) W and X ICs combine a compact non-rotating bulge with a thick disc;
- (v) K, M, O and V ICs contain rotating oblate spheroids which crudely represent elliptical galaxies.

## 2.2 Growing the disc

To simulate the growth of galaxies over cosmological timescales, stellar particles with a particle mass  $m = 1 \times 10^4 M_\odot$  are continuously added to the simulations. As we are interested in the co-evolution of thin and thick discs, we pursue two different ideas, which were already touched upon in Paper 1: a) We assume that a thick and rotating stellar component was formed early on in a galaxy's history (e.g. through a merger), represent this component with a thick disc or rotating spheroid in our ICs and over the timespan of the simulation add star particles to the system on near-circular orbits. b) We assume that the birth velocity dispersions of stars have been continuously declining over the history of a galaxy and add stellar populations with continuously decreasing dispersions.

### 2.2.1 Evolution of input velocity dispersion

In case a), the young stellar populations are assigned low birth velocity dispersions in the range  $\sigma_0 = \sigma_1 = 6 - 10 \text{ km s}^{-1}$  in all three directions  $R$ ,  $\phi$  and  $z$  as observed in the MW (see blue stars in Aumer & Binney 2009). The mean rotation

velocity  $v_\phi(R)$  at radius  $R$  is set to the circular velocity  $v_{\text{circ}} = \sqrt{a_R(R)R}$ , where  $a_R(R)$  is the azimuthal average of the radial gravitational acceleration,  $\partial\Phi/\partial R$ . As was shown in Papers 1 and 2, the in-plane dispersions quickly adjust to higher values and an appropriate ratio  $\sigma_\phi/\sigma_R$  due to spiral and bar heating.

In case b) we choose a value of  $\sigma_0$  that is declining with time. Paper 1 had applied

$$\sigma_0(t) = (6 + 30e^{-t/1.5\text{Gyr}}) \text{ km s}^{-1}, \quad (5)$$

which provided significantly too few hot stars. So, here we consider different functional forms:

$$\sigma_0(t) = \sigma_1 \left[ \arctan\left(\frac{t_1 - t}{1 \text{ Gyr}}\right) + \frac{\pi}{2} \right] + \sigma_2 \quad (\text{type atan}) \quad (6)$$

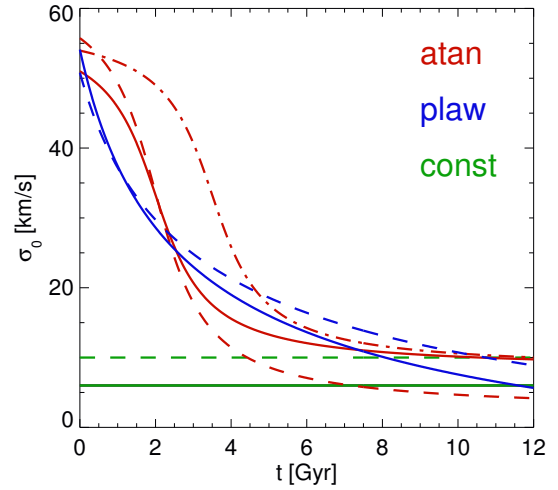
or

$$\sigma_0(t) = \sigma_1 \left( \frac{t + t_1}{t_2} \right)^{-t} - \sigma_2 \quad (\text{type plaw}) \quad (7)$$

In observations of high redshift galaxies, it is common to study kinematics of the  $\text{H}\alpha$  emission line and to assign a single characteristic velocity dispersion  $\sigma_{\text{H}\alpha}$  to each galaxy. Applying such a procedure, Wisnioski et al. (2015) find a dependence of observed  $\text{H}\alpha$  dispersions  $\sigma_{\text{H}\alpha} \propto (1 + z_{\text{rs}})$ , where  $z_{\text{rs}}$  is redshift (see also Kassin et al. 2012). The decline of our input velocity dispersion parameter  $\sigma_0(t)$  for simulations of type ‘plaw’ (Equation 7) resulted from a rough approximation to this proportionality under the crude assumption that the kinematics of young stars follow  $\text{H}\alpha$  kinematics. Note that this trend is observed for populations of galaxies at varying redshifts  $z_{\text{rs}}$ . The spread in  $\sigma_{\text{H}\alpha}$  for each given redshift range in Wisnioski et al. (2015) is significant and the expected formation histories for galaxies are diverse, so that it is reasonable to assume that individual galaxies can show evolutions of velocity dispersions very different from  $\sigma_{\text{H}\alpha} \propto (1 + z_{\text{rs}})$ . We thus also test decline histories of type ‘atan’ (Equation 6), which represents a scenario in which the transition from hot formation to cold formation is faster, yielding rather distinct hot and cold phases. Figure 1 visualises the difference between assumed  $\sigma_0$  histories.

We assume that  $\sigma_0$  is radially constant and always set  $\sigma_z(R, t) = \sigma_0(t)$ . Resolution effects prevent high  $z_{\text{rs}}$  observations from constraining the radial dependence of  $\sigma_{\text{H}\alpha}$  reliably, but the following two observational findings motivate this simple modelling approach. a) Newman et al. (2013) find that, if one characteristic value is assigned to each galaxy,  $\sigma_{\text{H}\alpha}$  depends little on the size of the galaxy; b) Jones et al. (2013) show radial  $\sigma_{\text{H}\alpha}(R)$  profiles of lensed galaxies and these show various shapes and high dispersions at outer radii.

In the case of high input dispersions, it is not desirable to use  $\sigma_\phi/\sigma_R = 1$  for an input, as this creates velocity distributions out of equilibrium. When  $\sigma_0 < 10 \text{ km s}^{-1}$ , heating by non-axisymmetries is very efficient for young stars (see Paper 2) and the dispersions will quickly adjust to an appropriate ratio  $\sigma_\phi/\sigma_R$ , but heating after birth will play a minor role for high  $\sigma_0$ . Since  $\sigma_\phi/\sigma_R$  depends on the shape of the rotation curve and the radial profile  $\sigma_R(R)$  (see e.g. Binney & Tremaine 2008, Section 4.4.3), we apply a simple approach: We assume  $\sigma_\phi = \sqrt{0.5}\sigma_R$  and an asymmetric drift correction for  $\langle v_\phi \rangle$  derived from Equation (4.228) in Binney & Tremaine (2008). We find that  $\sigma_\phi/\sigma_R$  for these assumptions only adjusts mildly after insertion.



**Figure 1.** Evolution of the input velocity dispersion parameter  $\sigma_0$  with simulation time  $t$  for various scenarios. Green lines are for type ‘const’ (solid for  $\sigma_0 = 6 \text{ km s}^{-1}$ , dashed for  $\sigma_0 = 10 \text{ km s}^{-1}$ ), blue lines are for type ‘plaw’ (solid for  $\iota = 0.56$ , dashed for  $\iota = 0.47$ , other parameters as in Table 2), and red lines are for type ‘atan’ (solid for  $\sigma_1 = 16 \text{ km s}^{-1}$ ,  $\sigma_2 = 8 \text{ km s}^{-1}$ ,  $t_1 = 2 \text{ Gyr}$ ; dashed for  $\sigma_1 = 20 \text{ km s}^{-1}$ ,  $\sigma_2 = 2 \text{ km s}^{-1}$ ,  $t_1 = 2 \text{ Gyr}$ ; dot-dashed for  $\sigma_1 = 16 \text{ km s}^{-1}$ ,  $\sigma_2 = 8 \text{ km s}^{-1}$ ,  $t_1 = 3.5 \text{ Gyr}$ ).

The choice for the ratio  $\sigma_z/\sigma_R$  is also unclear. For high  $z_{\text{rs}}$  discs, Genzel et al. (2011) studied  $\sigma_{\text{H}\alpha}$  as a function of disc inclination. Edge-on galaxies would thus be dominated by in-plane dispersions, whereas face-on galaxies would be dominated by vertical dispersions. They found mild indications for  $\sigma_R > \sigma_z$ , but Wisnioski et al. (2015) could not confirm this trend with a larger sample of galaxies. From the perspective of the thick-disc stars in the MW, Piffl et al. (2014) find  $\sigma_z \approx \sigma_R$ . We test values  $\sigma_R = \lambda\sigma_0$  with  $\lambda = 1 - 1.3$ .

As we place all newly added star particles in the mid-plane  $z = 0$ , the measured vertical dispersion  $\sigma_z$  of a young stellar component is smaller than  $\sigma_0$ , as the midplane is at the bottom of the vertical potential well and particles quickly lose kinetic energy moving away from it.

### 2.2.2 Star formation history

The star-formation rate is either constant (type 0), or

$$\text{SFR}(t) = \text{SFR}_0 \times \exp(-t/t_{\text{SFR}}) \quad (\text{type 1}), \quad (8)$$

or

$$\text{SFR}(t) = \text{SFR}_0 \times \exp(-t/t_{\text{SFR}} - 0.5 \text{ Gyr}/t) \quad (\text{type 2}), \quad (9)$$

with  $t_{\text{SFR}} = 6 - 12 \text{ Gyr}$ . Our simulations run for a total time of  $t_f = 10 - 12 \text{ Gyr}$  and the constant  $\text{SFR}_0$  is adjusted to produce at  $t_f$  a target final baryonic mass  $M_f$  in the range  $5 - 6 \times 10^{10} M_\odot$ , including the mass of the stars in the ICs. Paper 1 showed that for the range of DM density profiles adopted for our models, galaxy masses  $M_f$  in this range provide the right level of self-gravity to explain the age-velocity dispersion relation of the Sphd. McMillan (2017) favours similar galaxy masses for his MW mass models.

**Table 2.** List of models analysed in this paper. *1st Column:* Model Name; *2nd Column:* Initial Conditions; *3rd Column:* Final total baryonic mass  $M_f$  (including initial baryonic mass); *4th Column:* Final time  $t_f$ ; *5th Column:* Initial disc scalelength  $h_{R,i}$ ; *6th Column:* Final disc scalelength  $h_{R,f}$ ; *7th Column:* Scalelength growth parameter  $\xi$ ; *8th Column:* Type of SFH law; *9th Column:* Exponential decay timescale  $t_{\text{SFR}}$  for the star formation rate; *10th Column:* Radial to vertical dispersion ratio for inserted particles  $\lambda$ ; *11th Column:* Prescription for initial velocity dispersion for inserted stellar particles,  $\sigma_0(t)$ ; *12th-16th Column:* Parameters  $\sigma_1$ ,  $\sigma_2$ ,  $t_1$ ,  $t_2$  and  $\iota$ , which determine  $\sigma_0(t)$ ; *17th Column:* GMC star formation efficiency  $\zeta$ .

1st Name	2nd ICs	3rd $M_f$ [ $10^{10} M_\odot$ ]	4th $t_f$ [Gyr]	5th $h_{R,i}$ [kpc]	6th $h_{R,f}$ [kpc]	7th $\xi$	8th SF Type	9th $t_{\text{SFR}}$ [Gyr]	10th $\lambda$	11th $\sigma_0$ Type	12th $\sigma_1$ [ $\text{km s}^{-1}$ ]	13th $\sigma_2$ [ $\text{km s}^{-1}$ ]	14th $t_1$ [Gyr]	15th $t_2$ [Gyr]	16th $\iota$	17th $\zeta$
Y1	Y	5	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.08
Y1 $\zeta$ -	Y	5	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.04
Y2	Y	5	10	2.5	2.5	0.0	1	8.0	1.0	const	6	–	–	–	–	0.08
P1 $\zeta$ -	P	5	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.04
P1s6	P	5	10	1.5	4.3	0.5	0	–	1.0	const	6	–	–	–	–	0.08
P1 $\sigma$	P	5	10	1.5	4.3	0.5	1	8.0	1.0	const	10	–	–	–	–	0.08
P2	P	5	10	2.5	2.5	0.0	1	8.0	1.0	const	6	–	–	–	–	0.08
Q1	Q	6	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.08
Q1 $\zeta$ -	Q	6	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.04
R2 $\sigma$	R	5	10	2.5	2.5	0.0	1	8.0	1.0	const	10	–	–	–	–	0.08
U1	U	6	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.08
U1 $\sigma\zeta^*$	U	6	10	1.5	4.3	0.5	1	8.0	1.0	const	10	–	–	–	–	0.06
W1	W	6	10	1.5	4.3	0.5	1	8.0	1.0	const	6	–	–	–	–	0.08
X5 $\zeta^*$	X	6	10	1.5	3.5	0.5	1	8.0	1.0	const	6	–	–	–	–	0.06
T5 $\zeta^*$	T	5.5	10	1.5	3.5	0.5	1	8.0	1.0	const	6	–	–	–	–	0.06
K2	K	5	10	2.5	2.5	0.0	1	8.0	1.0	const	6	–	–	–	–	0.08
O2	O	5	10	2.5	2.5	0.0	1	8.0	1.0	const	6	–	–	–	–	0.08
C $\alpha$ 2	C	5	10	2.5	2.5	0.0	1	8.0	1.0	atan	20	2	2.0	–	–	0.08
C $\beta$ 2s5	C	5	12	2.5	2.5	0.0	1	8.0	1.0	plaw	51	15	1.57	2.7	0.56	0.08
M $\alpha$ 1	M	5	10	1.5	4.3	0.5	1	8.0	1.0	atan	16	8	2.0	–	–	0.08
M $\alpha$ 1 $\zeta^*$	M	5	10	1.5	4.3	0.5	1	8.0	1.0	atan	16	8	2.0	–	–	0.06
M $\beta$ 1s5	M	5	12	1.5	4.3	0.5	1	8.0	1.0	plaw	51	15	1.57	2.7	0.47	0.08
M $\alpha$ 8s7	M	5	12	1.0	4.3	0.6	2	12.0	1.0	atan	16	8	3.5	–	–	0.08
V $\alpha$ 1	V	5	10	1.5	4.3	0.5	1	8.0	1.0	atan	16	8	2.0	–	–	0.08
V $\alpha$ 5 $\lambda$	V	5	10	1.5	3.5	0.5	1	8.0	1.25	atan	16	8	2.0	–	–	0.08
V $\alpha$ 8s5	V	6	12	1.0	4.3	0.6	1	8.0	1.0	atan	16	8	2.0	–	–	0.08
V $\alpha$ 8s7	V	6	12	1.0	4.3	0.6	2	12.0	1.0	atan	16	8	3.5	–	–	0.08
V $\beta$ 8s5	V	6	12	1.0	4.3	0.6	1	8.0	1.0	plaw	51	15	1.57	2.7	0.47	0.08
V $\alpha$ 9s7 $\lambda\zeta^*$	V	6	12	1.0	3.5	0.6	2	12.0	1.25	atan	16	8	3.5	–	–	0.06
V $\alpha$ 9s8 $\lambda\zeta^*$	V	6	12	1.0	3.5	0.6	1	6.0	1.25	atan	16	8	2.0	–	–	0.06

### 2.2.3 Radial growth history

Particles are added randomly distributed in azimuth every five Myr with an exponential radial density profile  $\Sigma_{\text{SF}}(R) \propto \exp(-R/h_R(t))$ . The scalelength  $h_R(t)$  of the newly added particles grows in time as

$$h_R(t) = h_{R,i} + (h_{R,f} - h_{R,i})(t/t_f)^\xi. \quad (10)$$

To avoid inserting particles in the bar region, where near-circular orbits do not exist, particles are not added inside the cutoff radius  $R_{\text{cut}}$ , which is determined by the current bar length (‘adaptive cutoff’, see Paper 1 for details).

## 2.3 GMCs

GMCs are modelled as a population of massive collisionless particles drawn from a mass function of the form  $dN/dM \propto M^\gamma$  with lower and upper mass limits  $M_{\text{low}} = 10^5 M_\odot$  and  $M_{\text{up}} = 10^7 M_\odot$  and an exponent  $\gamma = -1.6$ . Their radial density is proportional to the star formation surface density

$\Sigma_{\text{SF}}(R)$ , and their azimuthal surface density is given by

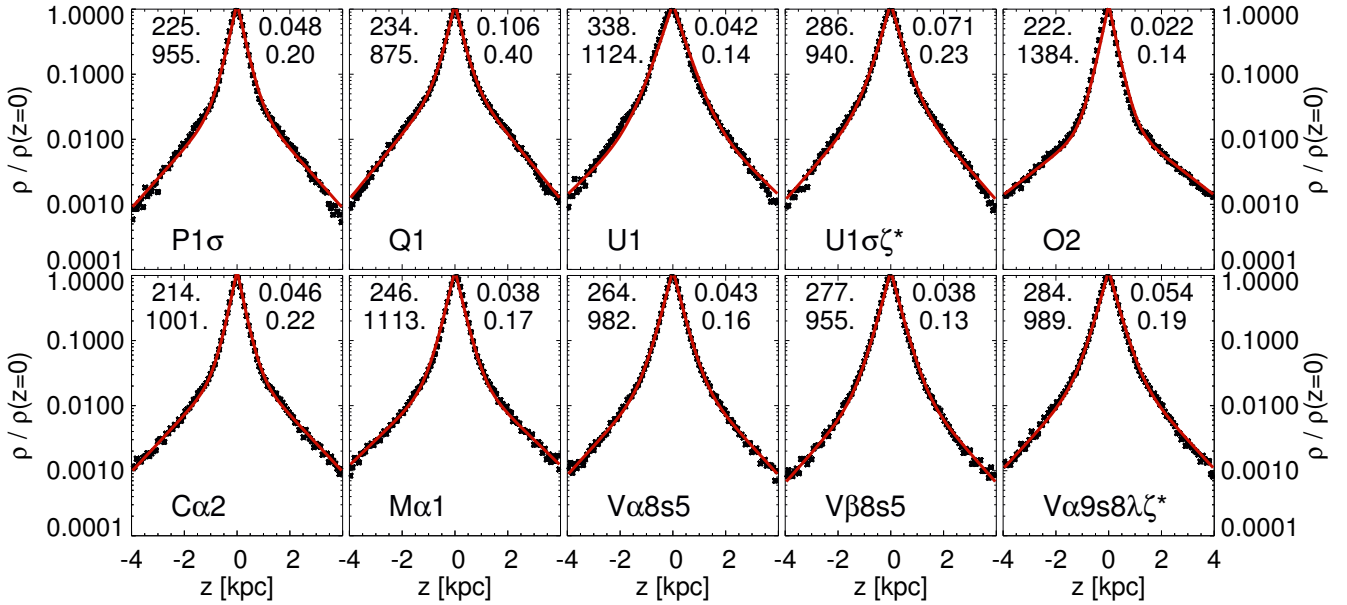
$$\Sigma_{\text{GMC}}(\phi) \propto [\Sigma_{\text{ys}}(\phi)]^\alpha, \quad (11)$$

where  $\Sigma_{\text{ys}}(\phi, R)$  is the surface density of young stars with ages 200–400 Myr and  $\alpha = 1$ . The mass in GMCs is determined by the SFR efficiency  $\zeta$ . Specifically, for each  $\Delta m_{\text{stars}}$  of stars formed, a total GMC mass  $\Delta m_{\text{GMC}} = \Delta m_{\text{stars}}/\zeta$  is created. GMC particles live for only 50 Myr: for 25 Myr their masses grow with time as  $m \propto t^2$ , and for the final 25 Myr of their lives their masses are constant, before they disappear instantaneously. GMCs are added on orbits with  $\sigma_0 = 6 \text{ km s}^{-1}$ . See Paper 1 for more details.

## 2.4 Overview of Models and their Naming

Table 2 gives an overview of the models discussed. We use an extended version of the naming convention described in Paper 1. All model names start with a capital letter identifying their IC according to the scheme defined by Table 1.





**Figure 2.** Vertical profiles of models at  $t = t_f$  and  $R = 8 \pm 0.5$  kpc are shown as black points. Overplotted are fits of Equation (12) to these profiles. The numbers in the upper left corners are the values of the scaleheights  $h_{\text{thin}}$  and  $h_{\text{thick}}$  in pc. The numbers in the upper right corners are the values of the density ratio  $f$  and the surface density ratio  $f_{\Sigma}$ .

A Greek letter  $\alpha$  or  $\beta$  following the initial capital letter indicates the model has declining birth velocity dispersion  $\sigma_0$  rather than a thick disc in its ICs. In  $\alpha$  models the decline of  $\sigma_0$  follows an ‘atan’ shape (Equation 6) while in  $\beta$  models it follows a ‘plaw’ shape (Equation 7). For simplicity the model names do not reflect the specific choices for the parameters of Equations (6) and (7), as they are of minor importance for the analyses in this paper.

These capital and, if present, Greek letters are followed by a number between 1 and 9 describing the radial growth history of the model, determined by parameters  $h_{R,i}$ ,  $h_{R,f}$  and  $\xi$ . Growth histories ‘1’, ‘2’, and ‘5’ were already used in Paper 1, ‘8’ ( $h_{R,i} = 1.0$  kpc,  $h_{R,f} = 4.3$  kpc and  $\xi = 0.6$ ) and ‘9’ ( $h_{R,i} = 1.0$  kpc,  $h_{R,f} = 3.5$  kpc and  $\xi = 0.6$ ) are new and represent inside out growth from a very compact disc into an extended disc.

The final baryonic masses  $M_f$  of all models lie in the rather narrow range  $5 - 6 \times 10^{10} M_{\odot}$  and we do not include the variations in the naming convention. For all other parameters we define standard values and additional digits added to the model name only when a model deviates in one or more parameters from the standard. The meanings of the additional digits are:

- The overall star formation history (SFH) of a model is described by the SFR type, the final time  $t_f$  and the SF timescale  $t_{\text{SFR}}$ . Our standard choice is a type 1 SFR with  $t_f = 10$  Gyr and  $t_{\text{SFR}} = 8$  Gyr. We have applied four additional SFHs, which are labelled by ‘s5’, ‘s6’, ‘s7’, ‘s8’.
- The standard input velocity dispersion in models without declining  $\sigma_0$  is  $\sigma_0 = 6 \text{ km s}^{-1}$ . Models with  $\sigma_0 = 10 \text{ km s}^{-1}$  are labelled ‘σ’.
- The GMC star formation efficiency  $\zeta$  has a standard value of 0.08. Models with  $\zeta = 0.04$  are labelled as ‘ζ-’, and models with  $\zeta = 0.06$  are labelled as ‘ζ\*’.
- The standard choice for the radial to vertical dispersion

ratio for inserted particles is  $\lambda = 1$ . Models with  $\lambda = 1.25$  are labelled ‘λ’.

### 3 VERTICAL PROFILES

Jurić et al. (2008) studied the vertical stellar density profile in the Sphd and presented a bias-corrected model-fit to Sloan Digital Sky Survey data of the form

$$\rho(z, R = 8 \text{ kpc}) = \rho_0 [\exp(-|z|/h_{\text{thin}}) + f \exp(-|z|/h_{\text{thick}})] . \quad (12)$$

They found  $h_{\text{thin}} = 300$  pc and  $h_{\text{thick}} = 900$  pc with 20 per cent uncertainty each and  $f = 0.12$  with 10 per cent uncertainty.  $f$  here is the ratio of local densities, whereas  $f_{\Sigma} = f h_{\text{thick}}/h_{\text{thin}} = 0.36$  is the ratio of local surface densities. Bland-Hawthorn & Gerhard (2016) show that while most studies in the literature find similar scaleheights  $h_{\text{thick}}$  and  $h_{\text{thin}}$  to Jurić et al. (2008), this determination ranks at the upper end in terms of  $f_{\Sigma}$  for photometric surveys, the average literature value of which is significantly lower,  $f_{\Sigma} = 0.12$ . This strong variation in  $f_{\Sigma}$  between different studies is likely caused by significant differences in the survey selection functions and degeneracies between  $f_{\Sigma}$  and the scaleheights.

#### 3.1 Vertical profile shape

In the models of Paper 1, GMC heating created remarkably exponential vertical profiles with scaleheights  $h_{\text{thin}} = 200 - 350$  pc. Thus these models can reproduce the scaleheight of the MW’s thin disc but fail to create a realistic thick disc. In Figure 2 we examine the vertical profiles of the present models at  $t = t_f$  and  $R = 8 \pm 0.5$  kpc as black symbols. We overplot in red fits of Equation (12) to these profiles with the scaleheights given in the top left corner of each panel.

The values of  $f$  and  $f_\Sigma$  are given in the top right corner of each panel. The figure illustrates that both approaches to producing thin+thick disc systems presented here produce double-exponential profiles similar to the one observed in the MW.

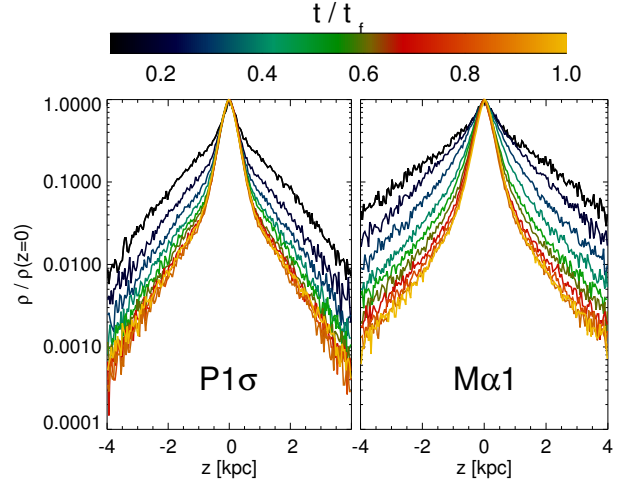
In the upper row of Figure 2 we show profiles for models with thick-disc ICs. We find that using disc-like ICs with  $z_0 \sim 1.7$  kpc yields values of  $h_{\text{thick}}$  in the range 815 – 1133 pc at  $t_f$  similar to the one inferred for the MW. As was already shown in Paper 1,  $z_0 \sim 1$  kpc yields final thick discs which are too thin. For elliptical ICs we find that setting the axis ratio  $s = 2$  in model K2 yields  $h_{\text{thick}} \approx 2.4$  kpc and  $s = 3$  in model O2 yields  $h_{\text{thick}} \approx 1.4$  kpc, which are both too thick. We thus decided to focus on disc-like thick ICs.

In terms of thin disc scaleheights, we find values  $h_{\text{thin}} = 200\text{--}339$  pc. As was already discussed in Paper 1, for a given GMC mass function, lower values of the star formation efficiency  $\zeta \sim 0.05$  and thus a higher total mass in GMCs per mass of formed stars are required to obtain  $h_{\text{thin}} \sim 300$  pc as in the MW. Lowering the DM halo concentration reduces the vertical force contribution from the halo and thus also mildly increases  $h_{\text{thin}}$ . The highest values for both  $h_{\text{thick}}$  and  $h_{\text{thin}}$  for disc IC models are found for model U1 which at  $t_f$  features an overly extended and thick bar as is discussed in Section 5. As was discussed in Paper 1, bars which extend beyond  $R = 5$  kpc can significantly thicken vertical profiles at  $R = 8$  kpc.

For the density and surface density ratios in models with thick-disc ICs, we find ranges of  $f = 0.04\text{--}0.167$  and  $f_\Sigma = 0.14\text{--}0.53$ , which include the Jurić et al. (2008) values and are in the upper half of the values of the Bland-Hawthorn & Gerhard (2016) literature compilation. Naturally, increasing the thick disc mass  $M_{\text{disc},i}$  at fixed thick disc scalelength  $h_{R,\text{disc}}$  increases these ratios (Q vs P models) and decreasing  $h_{R,\text{disc}}$  at fixed  $M_{\text{disc},i}$  lowers them (U vs Q models). As far as other model parameters are concerned there are no clear patterns apparent. This is likely connected to competing effects: E.g. few GMCs produce less vertical heating, but lead to stronger bars, which, if long enough, can thicken the vertical profile.

The lower row of Figure 2 shows models with declining birth dispersions. We find values for the scaleheights in the ranges  $h_{\text{thin}} = 214\text{--}284$  pc and  $h_{\text{thick}} = 863\text{--}1234$  pc, very similar to the ranges found for thick IC disc models. Model Ca2 has the thinnest thin disc as at late times it has a very low value of  $\sigma_0 < 5 \text{ km s}^{-1}$  and it also has a high star formation efficiency  $\zeta = 0.08$ . Va9s8 $\lambda\zeta^*$  has  $\zeta = 0.06$  and  $\sigma_0 \sim 10 \text{ km s}^{-1}$  at late times and thus the highest  $h_{\text{thin}}$  among these models. Va8s5 and Vβ8s5 only differ in the shape of the declining  $\sigma_0$  curve: Va8s5 is of type ‘atan’, whereas Vβ8s5 is of type ‘plaw’ (see Equations 6 and 7). Their final vertical profiles are rather similar, so both types of decline are acceptable. The value of  $h_{\text{thick}}$  is also mildly influenced by the ICs, which for V and M models is a low-mass  $s = 2$  elliptical. As models Va8s5/Vβ8s5 have a higher final mass than Mα1 and thus a lower fraction of IC stars, their thick discs appear mildly thinner.

The corresponding thin-to-thick disc ratios are  $f = 0.025\text{--}0.055$  and  $f_\Sigma = 0.11\text{--}0.22$ , which are lower than the Jurić et al. (2008) values, but comfortably within the range found in the literature. For models with declining  $\sigma_0(t)$  the thick disc mass fraction depends on the fraction of mass



**Figure 3.** Vertical profiles of models at  $R = 8 \pm 0.5$  kpc and various times  $t$  as indicated by the colour bar.

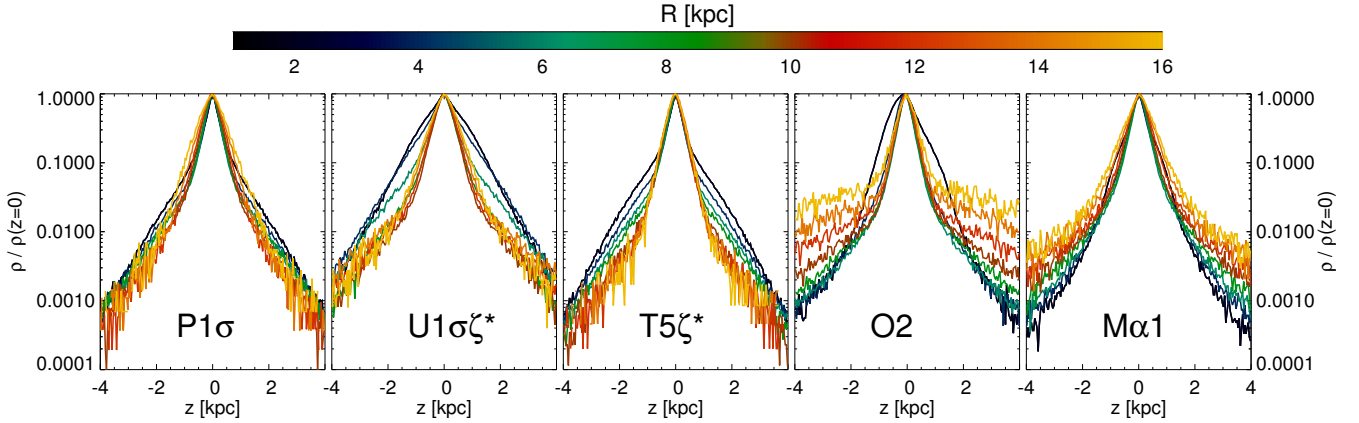
formed during early formation stages with high  $\sigma_0$  and thus on the detailed forms of  $\sigma_0(t)$  and the SFH. Moreover, the radial growth history determines how many stars are formed at a certain radius during this period. Dynamical heating and migration processes also influence the number of old stars found at  $t_f$ , so the final value of  $f$  is not easily predicted.

### 3.2 Vertical profile evolution with time

In Figure 3 we analyse the temporal evolution of the shapes of vertical profiles in two different types of simulations. In the models of Paper 1 that included GMC heating, thin disc vertical profiles are at all times exponential and their scaleheights change very little with time. This finding reflects balance between mass growth, which continuously supplies cold particles and deepens the vertical potential well, and GMC heating which efficiently increases the vertical velocity dispersions of young stars.

The thick-disc IC model P1σ shows a double exponential profile from early-on. The thin disc part of its profile is very constant, just like the thin-disc-only models of Paper 1. What changes are the surface density ratio  $f_\Sigma$ , which by construction becomes more and more thin-disc-dominated, and the scaleheight of the thick disc, which becomes smaller with time, because the growth in the thin disc’s mass deepens the vertical potential well and the thick disc is not heated significantly.

Mα1 represents models with declining  $\sigma_0(t)$ . Its profile at early times is closer to a single- than a double-exponential, as there are no cold, thin-disc populations present. Only as  $\sigma_0(t)$  falls below  $\sim 20 \text{ km s}^{-1}$  does a thin disc build up. The scaleheight of the thin disc  $h_{\text{thin}}$  becomes smaller with time as the decline in  $\sigma_0(t)$  cools the thin-disc population as a whole despite the vertical heating due to the GMCs. The deepening of the vertical potential well adds to that effect and reduces  $h_{\text{thick}}$  in the same way as in P1σ.



**Figure 4.** Vertical profiles of models at  $t = t_f$  and various radii  $R$  as indicated by the colour bar.

### 3.3 Radial dependence of vertical profiles

The vertical profiles of observed disc galaxies are very constant radially (van der Kruit & Searle 1982). Paper 1 showed that in thin-disc-only models that include GMC heating, vertical profiles are almost independent of radius  $R$ , unless there is a buckled bar, which thickens only the central region. Figure 4 shows the radial variations of the vertical profiles in five of our thin+thick models. The inner parts of models U1 $\sigma\zeta^*$  and O2 show thickening by a bar. U1 $\sigma\zeta^*$  otherwise shows a profile which is almost constant in radius.

Model P1 $\sigma$  shows a thin disc that becomes mildly thicker and has a higher mass fraction towards larger  $R$ , whereas  $h_{\text{thick}}$  stays roughly constant. This is caused by inside-out formation, on account of which the young thin disc has a longer scalelength than the thick disc, and higher  $\sigma_0 = 10 \text{ km s}^{-1}$  than the models studied in Paper 1. Combined with a shallower potential well at outer radii, where heating is limited, this value of  $\sigma_0$  yields thicker thin discs.

T5 $\zeta^*$  shows a stronger fading and a mild thickening of the thick populations towards the outskirts. Compared to P ICs, the T ICs have a more compact and more massive thick disc and a also bulge component. The effective vertical profile of the two components thus varies with radius already in the ICs. O2 has elliptical ICs, which lead to a thick population that becomes thicker with increasing  $R$  and also attains a higher mass fraction in the outskirts.

Model M $\alpha$ 1 with declining  $\sigma_0$  shows a thin disc that thickens mildly with  $R$ , because at late times it has  $\sigma_0 \sim 10 \text{ km s}^{-1}$  and thus behaves similarly to model P1 $\sigma$  discussed above. The thick disc becomes thicker with  $R$ . This is characteristic for models of this type. It is a consequence of our assumption that  $\sigma_0$  is constant with radius. This leads to a flaring of the hot component, which is hardly influenced by vertical heating.

Note that in contrast to the thick discs in models with declining  $\sigma_0$ , the thick-disc ICs of our alternative modelling scheme were set up with a radially constant scaleheight and thus a vertical velocity dispersion which declines with  $R$ .

### 3.4 Flaring of mono-age components

In Figure 5 we examine for a selection of models how the median distance from the midplane  $|z|_{\text{med}}$  varies with radius

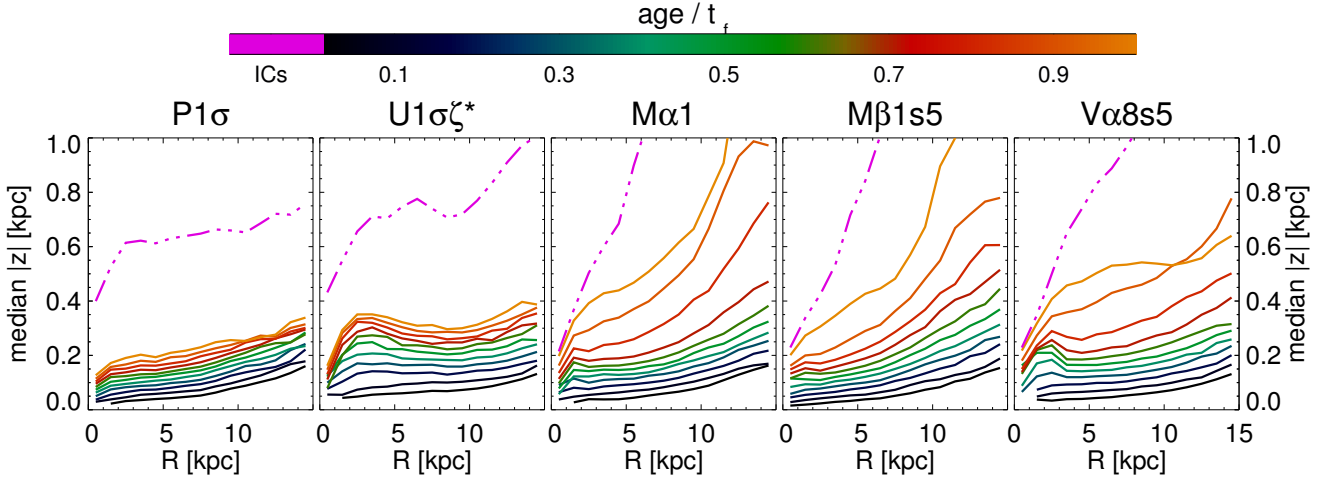
$R$  for populations of different ages. It has been suggested that the lack of significant changes with disc radius  $R$  in the double-exponential vertical mass profiles of disc galaxies is a consequence of inside-out growth combined with  $|z|_{\text{med}}$  being an increasing function of  $R$  for all mono-age populations (‘flaring’, Minchev et al. 2015).

Flaring can be caused by satellite interactions or misaligned infall of gas, which are not present in our models. However, Paper 1 showed that disc galaxies formed in isolation also have flaring mono-age components. The amount of flaring is determined by the radial mass profile of the disc and the radial profile of the vertical velocity dispersion of mono-age components  $\sigma_z(R, \tau)$ , which is determined by the birth dispersions of stars and the vertical heating mechanism(s) at work. Radial migration of stars can also influence  $\sigma_z(R, \tau)$  (Schönrich & Binney 2012).

The left-most panel of Figure 5 depicts  $|z|_{\text{med}}(R)$  for the mono-age populations of Model P1 $\sigma$ , which has thick-disc ICs. The thin-disc populations are at all radii substantially thinner than the old thick-disc stars. These mono-age populations all show flaring, and for the youngest populations  $|z|_{\text{med}}$  increases from  $R = 1$  to  $15 \text{ kpc}$  by up to a factor of  $\eta \equiv |z|_{\text{med}}(15 \text{ kpc})/|z|_{\text{med}}(1 \text{ kpc}) \sim 5$ . In fact, in P1 $\sigma$  the structure of the thin-disc populations is similar to that in thin-disc only model Y1 examined in Paper 1. The relative increase in  $|z|_{\text{med}}$  with  $R$  becomes smaller with increasing age, and we find  $\eta \sim 2.5$  for the oldest inserted stars. In Model U1 $\sigma\zeta^*$  (second panel in Figure 5) the situation is altered by a vertically extended bar. On account of bar buckling,  $|z|_{\text{med}}$  for intermediate-age populations now peaks around  $R \sim 3 \text{ kpc}$ , then declines slightly to  $R \sim 8 \text{ kpc}$ , and gradually increases further outward. The youngest populations have not been affected by bar buckling and have  $\eta \sim 3$ .

In all our models, non-IC stars at any given time  $t$  are inserted with a radially constant birth dispersion  $\sigma_0(t)$ . On account of the outwards decrease in surface density, this results in flaring. GMC heating increases  $\sigma_z$  more strongly in the centre than in the outskirts and thus flattens the increase in  $|z|_{\text{med}}(R)$ . For models without buckled bars and with  $\zeta = 0.04$  and thus more GMCs per unit mass of inserted stars,  $\eta$  can be as low as  $\sim 1.5$  for old inserted stars. Bars can additionally heat the central regions of galaxies (see e.g. Grand et al. 2016) and thus cause even flatter runs of  $|z|_{\text{med}}(R)$ .





**Figure 5.** The median vertical distance from the midplane  $|z|_{\text{med}}$  as a function of radius  $R$  for various mono-age populations. Solid lines are for inserted particles of various ages (see colour bar). The pink dashed line is for IC stars.

The three panels on the right of Figure 5 show models with declining  $\sigma_0$ , and in these models the structure of  $|z|_{\text{med}}(R)$  is quite different from what it is in the models with thick IC discs. Now the low-mass elliptical ICs are unimportant because the thick disc is formed mainly by old added stars. Due to the continuous decline in  $\sigma_0(t)$ , the curves form a continuum rather than a bi-modal grouping. The young, thin-disc populations behave very similarly to those in the models with thick-disc ICs, the local maxima in  $|z|_{\text{med}}$  at  $R \sim 2$  kpc in model  $V\alpha 8s5$  being caused by a buckled bar.

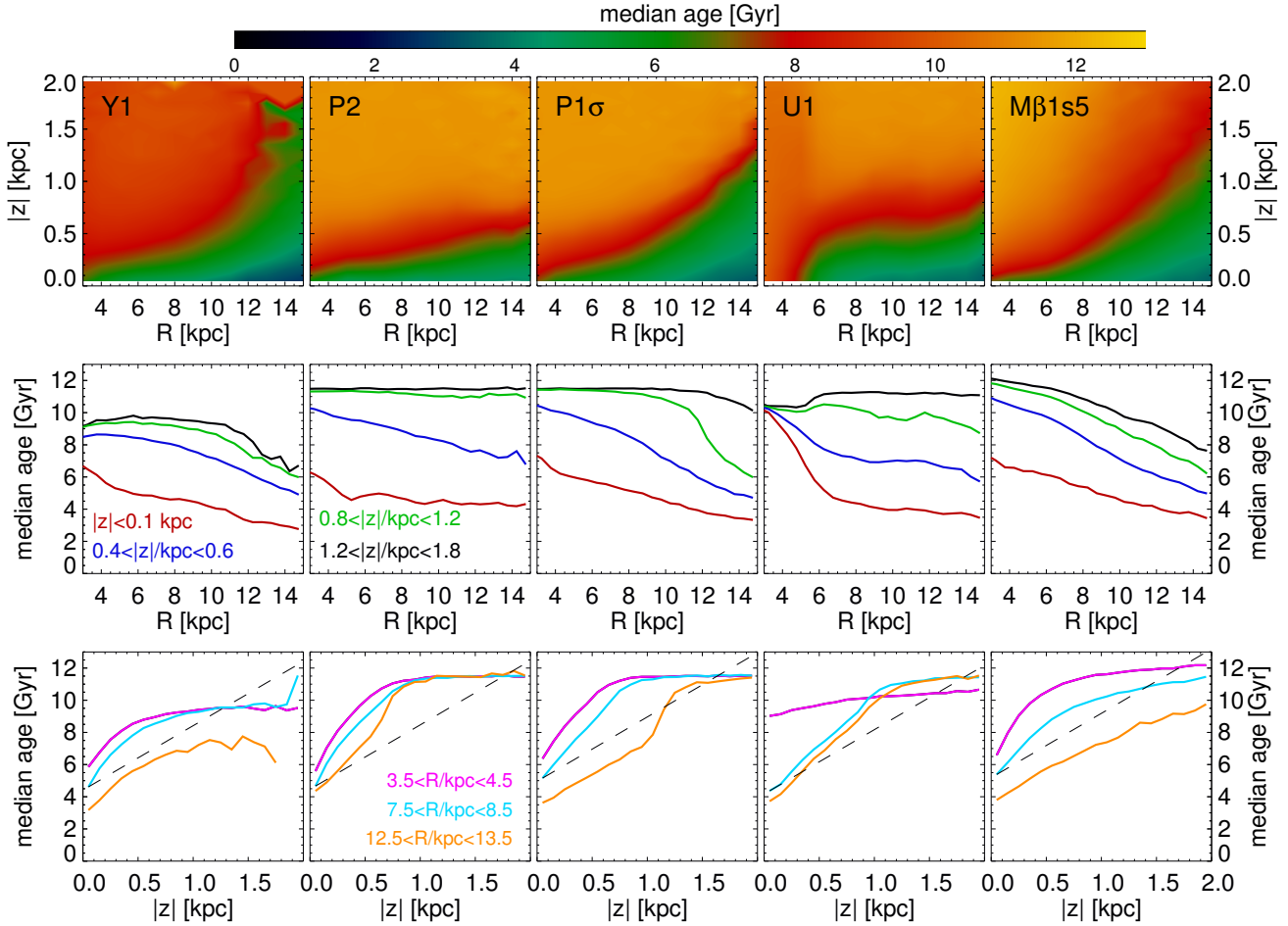
The thick-disc stars in models with declining  $\sigma_0$  (red and orange curves) yield a similar value of  $\eta$  to the youngest thin-disc stars in all models. Again this reflects our decision to make the declining birth dispersion  $\sigma_0$  independent of  $R$ . Given that vertical GMC heating has little influence on the thick-disc populations, and that the depth of the vertical potential well declines with  $R$ , strong flaring is an inevitable consequence. The flattest curves are thus found for the intermediate age populations, which were already born on relatively cold orbits and have been significantly affected by GMC and bar heating. The differences between models  $M\alpha 1$  and  $M\beta 1s5$  can be explained by the different shapes of  $\sigma_0(t)$  applied (see Figure 1). Whereas the oldest component of each model flares in a similar way, the two next-oldest populations show stronger flaring in the atan model  $M\alpha 1$  than in the plaw model  $M\beta 1s5$ , in which  $\sigma_0$  declines more gradually. Also, in the atan model  $|z|_{\text{med}}$  declines with decreasing age faster than in the plaw model.

The flaring of all disc populations discussed so far can be qualitatively explained by birth-dispersion profiles and disc heating mechanisms. A comparison between the thick-disc IC stars in models  $P1\sigma$  and  $U1\sigma\zeta^*$  indicates that radial migration plays a role as well. The IC stars in  $P1\sigma$  only show a strong outwards increase in  $|z|_{\text{med}}$  in the centre and hardly any flaring at larger radii, whereas in  $U1\sigma\zeta^*$ , the flaring of the IC stars in the outer disc is stronger. The thick-disc ICs were created with a radially constant scaleheight and are hardly affected by vertical heating mechanisms, so their curve of  $|z|_{\text{med}}(R)$  at  $t_f$  is determined by the change in the vertical potential well together with extent to which their stars migrate radially. Bar formation funnels a lot of mass

to the centre and consequently the disc's thickness decreases there. The stronger flaring in  $U1\sigma\zeta^*$  is likely explained by higher levels of radial migration.

When stars migrate radially, their vertical actions  $J_z$  are conserved (Solway et al. 2012). Populations of stars that are born with a radially constant scaleheight in a MW-like disc have their mean  $J_z$  decreasing with  $R$ . Consequently, outwards migrators at a given  $R$  have higher  $J_z$  (and thus higher  $\sigma_z$ ) than non- or inwards migrators. If there are more outwards than inwards migrators, as expected for the outer disc regions, such populations are expected to flare (Schönrich & Binney 2012). This principle is complicated by the finding that stars migrate less if they have high  $J_z$  (Vera-Ciro et al. 2014), but Paper 4 shows that for our thin-disc populations and for stars from thick-disc ICs at  $t = t_f$ , outwards migrators are indeed more numerous in the outer disc and have higher  $\sigma_z$  than inwards migrators (see also Roškar et al. 2013). Model  $U1\sigma\zeta^*$  has both a more compact IC disc and a lower-density dark halo than Model  $P1\sigma$ , which leads to stronger non-axisymmetries and a higher fraction of outwards migrators in the outer disc (Paper 4), which in turn explains the stronger flaring of IC stars in  $U1\sigma\zeta^*$ .

Compared to the oldest inserted stars in Models  $M\alpha 1$  and  $M\beta 1s5$ , in Model  $V\alpha 8s5$  this population shows a much flatter curve  $|z|_{\text{med}}(R)$  at  $R > 5$  kpc. Model  $V\alpha 8s5$  has a lower-density dark halo and a more compact disc at early times than the M models. Hence in this model the oldest inserted stars in the outer disc have a higher fraction of outwards migrators. For models with declining  $\sigma_0$ , thick-disc stars are born with radially constant  $\sigma_z$  and thus their mean  $J_z$  increases with  $R$ . As they are hardly affected by disc heating, in these models old outwards migrators at a given  $R$  have lower  $J_z$  and thus lower  $\sigma_z$  than inwards migrators. This likely explains the flatter curve  $|z|_{\text{med}}(R)$  for the oldest inserted stars in Model  $V\alpha 8s5$ . So depending on the shape of  $\sigma_z(R)$  at birth, radial migration can both strengthen and weaken the flaring of a mono-age population. A detailed analysis of radial migration in our models is presented in Paper 4.



**Figure 6.** Top: median age  $\tau_{\text{med}}$  maps in the  $R$ - $|z|$ -plane. The colour bar shows the age encoding. Middle: Median ages  $\tau_{\text{med}}$  as a function of radius  $R$  for different vertical distances from the midplane: red is  $|z| < 0.1$  kpc, green is  $|z| = 0.5 \pm 0.1$  kpc, blue is  $|z| = 1.0 \pm 0.2$  kpc and black is  $|z| = 1.5 \pm 0.3$  kpc. Lower: Median ages  $\tau_{\text{med}}$  as a function of vertical distance from the midplane  $|z|$  at different radii: pink is  $R = 4 \pm 0.5$  kpc, cyan is  $R = 8 \pm 0.5$  kpc, orange is  $R = 13 \pm 0.5$  kpc. Each column presents one model.

### 3.5 Radial and vertical age structure

The combination of recent and ongoing astrometric and spectroscopic surveys of MW stars is about to increase vastly our knowledge of the age structure in the Galactic disc(s) (e.g. [Martig et al. 2016a](#)). In the top row of Figure 6 we therefore show maps of median age  $\tau_{\text{med}}$  as a function of  $R$  and  $|z|$  for various models. To do so we assume that the oldest stars in all models are 13 Gyr old and thus randomly assign ages in the range  $[t_f, 13 \text{ Gyr}]$  to IC star particles.

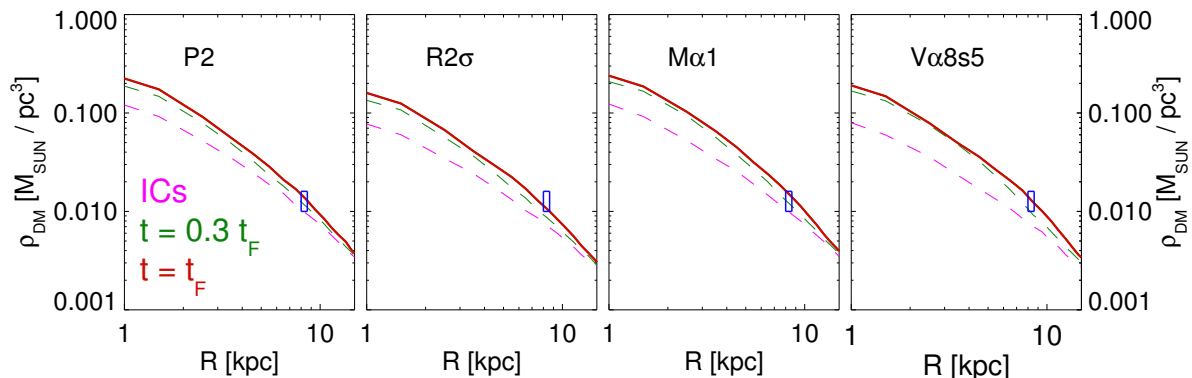
The leftmost panel shows the thin-disc-only model Y1 from Paper 1. The innermost region  $R < 4$  kpc at all altitudes and high altitudes  $|z| > 1.5$  kpc at all radii are dominated by old stars. The youngest  $\tau_{\text{med}}$  are found at high  $R$  and low  $|z|$ . Intermediate  $\tau_{\text{med}}$  are confined to  $|z| < 200$  pc at  $R < 6$  kpc, to  $|z| < 500$  pc at  $R < 10$  kpc, but can populate regions up to  $|z| \sim 1500$  pc at  $R \sim 15$  kpc. This characteristic  $\tau_{\text{med}}$  pattern is caused by a combination of inside-out formation and disc heating. The oldest population is more compact and thicker, whereas the younger population lives closer to the midplane and preferentially at larger radii, where it extends to higher  $|z|$ .

Model P1 $\sigma$  has an old thick disc, that is more massive,

much thicker and more extended than the oldest disc population in Y1. This changes the  $\tau_{\text{med}}$  map only marginally. The high  $|z|$  populations are by construction older than in Y1 and the young, outermost populations are also mildly older than their counterparts in Y1. The  $\tau_{\text{med}}$  pattern is however, rather similar.

P2 differs from P1 $\sigma$  in lacking inside-out formation, which makes the outer populations older. It also has lower  $\sigma_0$ , which reduces the flaring of the outer populations. Consequently, at all radii  $R$ , all altitudes  $|z| > 500$  pc are dominated by old stars and the zone populated by intermediate  $\tau_{\text{med}}$  is confined to  $|z| > 300$  pc at  $R < 10$  kpc and to  $|z| > 500$  pc at all  $R$ .

Model U1 is an inside-out model, but compared to P1 $\sigma$  has a lower concentration halo, a more massive and more compact IC thick disc and lower  $\sigma_0$ . It has a thicker and longer bar than the other depicted models. The buckled bar causes an area at  $R < 5$  kpc, which is populated by stars with old  $\tau_{\text{med}}$  and shows no noticeable vertical age gradient. The bar also heats the disc at  $R \sim 5 - 10$  kpc vertically and thus increases the altitudes at which younger stars are found, so the radial increase of the maximum  $|z|$  at which intermediate  $\tau_{\text{med}}$  are found is flatter in this model.



**Figure 7.** DM density profiles  $\rho_{\text{DM}}(R)$  as measured in the midplane of the galaxies P2, R2 $\sigma$ , M $\alpha$ 1 and V $\alpha$ 8s5. The solid red lines show the models at  $t = t_f$ , whereas the dashed lines show  $t = 0.3t_f$  (green) and the ICs (pink). The blue boxes mark the constraints for today's Snhd from McKee et al. (2015).

Model M $\beta$ 1s5 is a model with declining  $\sigma_0$ . It shows a  $\tau_{\text{med}}$  pattern that differs from those of the other models shown in Figure 6. At  $R < 5$  kpc, its vertical age structure is similar to that in the P models, but at  $R > 5$  kpc the intermediate age stars reach higher altitudes, which leads to declining radial age gradients at all  $|z|$ . This is caused by our assumption of radially constant  $\sigma_0(t)$ , as was also discussed in relation to the flaring of mono-age populations shown in Figure 5. Moreover, vertical heating is inefficient for stars with high  $\sigma_0$  and for stars at large radii.

Martig et al. (2016b) have recently presented measurements of radial gradients in  $\tau_{\text{med}}$  at various altitudes  $|z|$  in the Snhd. At  $|z| > 500$  pc, they find significant declines in  $\tau_{\text{med}}$  with radius  $R$  at all radii. However, the value of  $\partial\tau_{\text{med}}/\partial R$  is still very uncertain. Close to the plane near the Solar radius  $R_0$ ,  $\tau_{\text{med}}(R)$  is rather flat. In the middle row of Figure 6 we show  $\tau_{\text{med}}$  vs  $R$  at four ranges in  $|z|$ .

As model Y1 lacks a thick disc, we will not discuss it in detail. The P and U models with thick IC discs clearly show no decline in  $\tau_{\text{med}}$  with  $R$  at  $|z| \sim 1.5$  kpc and only P1 $\sigma$  shows a negative  $d\tau_{\text{med}}/dR$  at  $|z| \sim 1.0$  kpc and  $R > 10$  kpc due to inside-out formation and a stronger flaring in the young disc due to higher  $\sigma_0$ . By contrast, M $\beta$ 1s5 shows clear negative gradients throughout the whole disc at all latitudes and is thus more in agreement with the age determinations of Martig et al. (2016b).

At  $|z| \sim 0.5$  kpc, all models show a negative radial  $\tau_{\text{med}}$  gradient. This gradient is weaker in U1 due to the influence of the unrealistically long and thick bar. In the midplane, the inside-out models also show clear negative age gradients, whereas P2, which has a constant radial feeding scalelength, shows a rather flat age profile, as does U1, again strongly affected by the long bar. As the age gradients in M $\beta$ 1s5 and P1 $\sigma$  are flatter closer to the midplane and the observations are still very uncertain, little can be deduced yet about inside-out growth.

Casagrande et al. (2016) have presented evidence for a vertical age gradient in the Snhd from asteroseismology. They find a decline by  $\sim 4 \pm 2$  Gyr/kpc, but have little knowledge of the shape of the decline. In the lower row of Figure 6 we present  $\tau_{\text{med}}$  as a function of  $|z|$  at three radii, the cyan line representing a Solar-like radius and the dashed line showing a constant vertical gradient of 4 Gyr/kpc.

Due to the dominance of the thick disc at  $|z| > 1$  kpc and the weak radial age gradient, at all radii in P2, age increases more strongly than at 4 Gyr/kpc up to  $|z| \sim 1$  kpc and then flattens out. In P1 $\sigma$  at  $R = 13$  kpc, the enhanced presence of younger stars away from the plane results in a flatter increase of  $\tau_{\text{med}}$  with  $|z|$ , but at  $R = 8$  kpc the situation is similar to that in P2. In U1 the long bar causes an almost flat  $\tau_{\text{med}}$  vs.  $|z|$  plot in the bar region and a flatter gradient at  $R = 8$  kpc. In M $\beta$ 1s5,  $\tau_{\text{med}}(|z|)$  at  $R = 8$  kpc flattens more gradually and at  $R = 13$  kpc has an almost constant slope due to the stronger flaring of mono-age populations.

Averaged over the studied vertical extent of 2 kpc, all models show a vertical gradient consistent with the findings of Casagrande et al. (2016). As they observe very few stars above  $|z| > 1.2$  kpc, model M $\beta$ 1s5 shows the best agreement with the still very uncertain data.

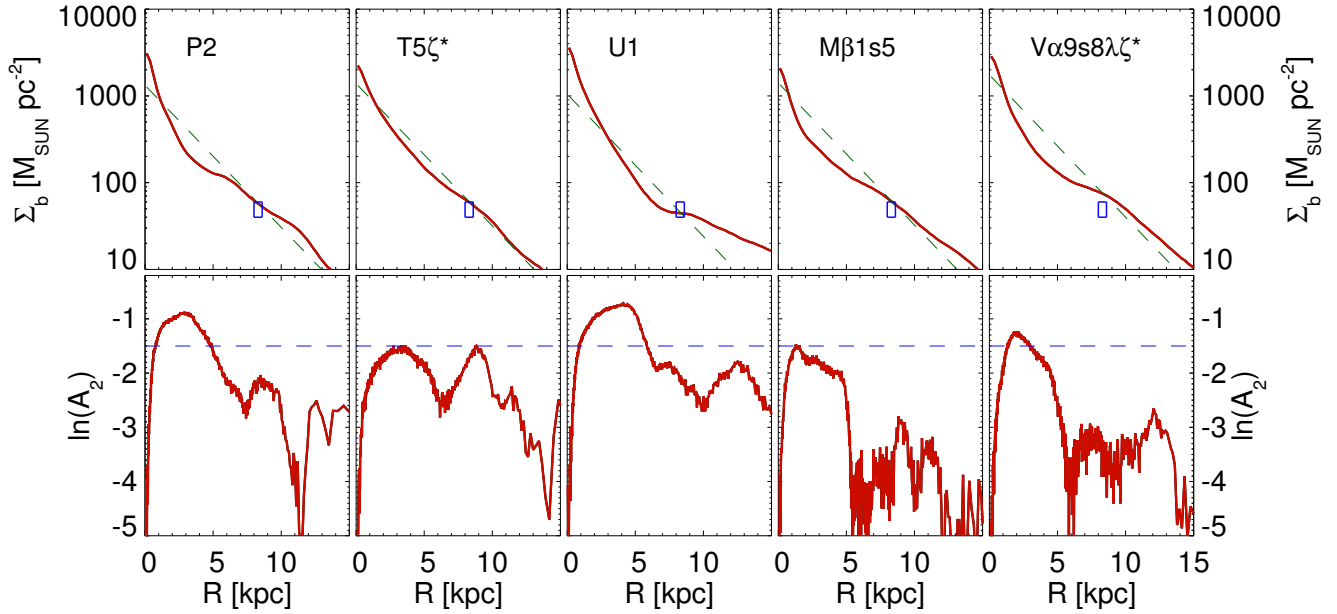
## 4 RADIAL MASS DISTRIBUTION

In this section we investigate the radial distribution of baryonic and dark matter in our models and the circular speed curves  $v_{\text{circ}}(R)$  which result from them.

### 4.1 Dark matter density

As discussed in Paper 1 the parameters for our DM haloes as set up in the ICs are motivated by what  $\Lambda$ CDM predicts for haloes associated with MW mass galaxies. The DM profile  $\rho_{\text{DM}}(r)$  will be modified by growing a massive baryonic disc within the DM halo and by interaction with non-axisymmetric disc structures such as the bar and spirals. As the halo is always spherical in the ICs, but the disc mass fraction in the ICs varies strongly between ICs and the various galaxy models evolve differently, the final haloes differ even if two models share the same DM IC parameters.

Our best constraints on the DM content in the MW come from dynamical measurements of the total matter surface density in the Snhd. Subtracting the baryonic components, McKee et al. (2015) find a local DM density of  $\rho_{\text{DM}} = 0.013 \pm 0.003 M_{\odot} \text{pc}^{-3}$ . In Figure 7 we plot in red the DM density profiles  $\rho_{\text{DM}}(R)$  as measured in the midplane of the galaxy at  $t = t_f$  and compare them to the Snhd constraints



**Figure 8.** Top: Baryonic surface density profiles  $\Sigma_b(R)$  for various models. The blue boxes mark the constraints for the Snhd from McKee et al. (2015). The dashed line marks an exponential with scalelength  $h_R = 2.65$  kpc. Bottom: Radial profiles of the  $m = 2$  Fourier amplitude  $A_2$ . The dashed lines mark  $\ln(A_2) = -1.5$ .

assuming the Solar Galactic radius is  $R_0 = 8.3 \pm 0.3$  kpc (Schönrich 2012).

We find that at  $t_f$  all models fall within the constraints of McKee et al. (2015). The four shown in Figure 7 are a representative selection. The main drivers for  $\rho_{DM}(R_0)$  at  $t_f$  are, as expected, initial halo concentration  $c$  and the added disc mass  $M_{add} = M_f - M_{disc,i} - M_{bulge,i}$ . Consequently R2 $\sigma$ , which has  $c = 6.5$  and  $M_{add} = 3.5 \times 10^{10} M_\odot$  has the lowest  $\rho_{DM}(R_0)$ , whereas M $\alpha$ 1 with  $c = 9$  and  $M_{add} = 4.5 \times 10^{10} M_\odot$  has the highest  $\rho_{DM}(R_0)$  among the models shown. P2 with  $c = 9$  and  $M_{add} = 3.5 \times 10^{10} M_\odot$  and V $\alpha$ 8s5 with  $c = 6.5$  and  $M_{add} = 5.5 \times 10^{10} M_\odot$  show intermediate  $\rho_{DM}(R_0)$ . It is worth noting that the models with  $c = 4$  presented in Paper 1 and discarded because of overly strong bars indeed show too low  $\rho_{DM}(R_0)$ .

We also note that at  $t_f$  none of our models shows a cored DM profile in the centre, as was recently favoured by Cole & Binney (2017). Our IC DM profiles do not contain a core as is indicated by the pink dashed lines. The DM densities  $\rho_{DM}(R_0)$  of the ICs are significantly lower than in the final models. Initial profiles with  $c = 9$  lie at the lower allowed limit for today’s Snhd and models with  $c = 6.5$  are clearly below this limit. During the simulations they are altered by compression due to the added mass in stars and by angular momentum transfer from stars to DM due to bars and spirals. As all models have declining SFR( $t$ ) and bars form in the later evolution stages as shown in Section 5, the increase in the DM density at  $R < 15$  kpc, where the disc grows, is strong up to  $t = 0.3t_f$  as indicated by the green dashed lines and rather weak afterwards. In the four models shown, the relative increase in  $\rho_{DM}(R < 15$  kpc) is strongest in models V $\alpha$ 8s5 as it has the largest baryonic mass fraction and the largest  $M_{add}$ , and the increase is weakest in P2. Angular momentum transfer to the halo by spirals and the bar is not strong enough to create cores, as was also shown by Sellwood (2008).

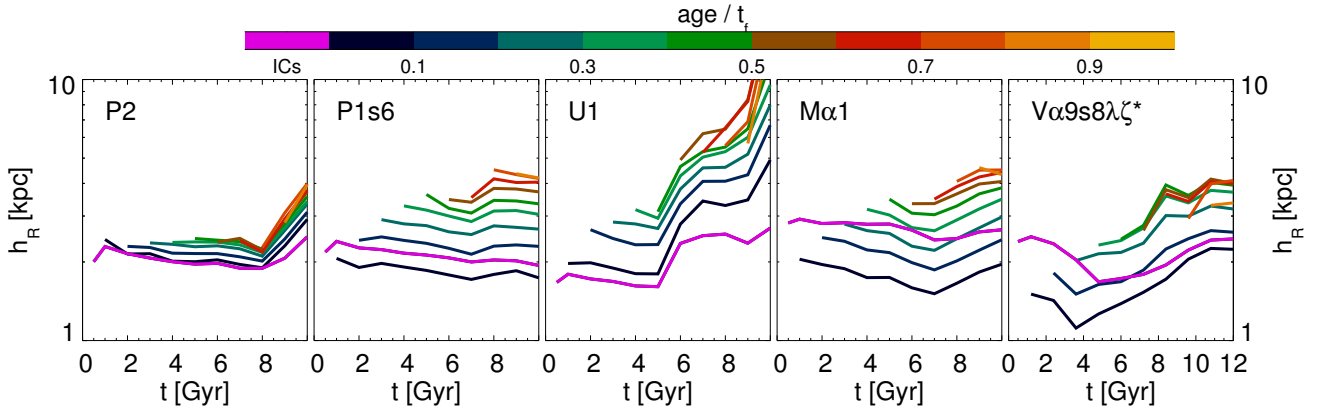
## 4.2 Solar Neighbourhood surface density

Paper 1 showed that despite having control over the evolution of the input scalelength  $h_R(t)$ , there was little control over the final surface density profile of the models. The more compactly a disc was fed, the earlier it grew a bar, which redistributed matter and, as we avoid inserting particles into the bar region, shifted the inner cutoff radius outwards. In the end, the surface density profiles of a range of models with different radial growth histories were thus rather similar. Our surface density was thus decided by the total mass of the final model, which we justified from the reasonable agreement of our models with a) the vertical scaleheight of the thin disc, b) an appropriate local circular speed, c) an appropriate amount of radial migration to  $R_0$  and d) appropriate vertical and radial velocity dispersions.

In Figure 8 we examine, how well our models fulfil constraints on the Snhd baryonic surface density  $\Sigma_b(R_0)$ . Table 3 of McKee et al. (2015) gives an overview of determinations of  $\Sigma_b(R_0)$  including gas of all phases, stars and stellar remnants.  $\Sigma_b$  is consistently found to be in the range  $\Sigma_b(R_0) = 40 - 60 M_\odot \text{pc}^{-2}$ , where the given errors are included in the interval. Gas is found to contribute 25 – 30 per cent. As in our simulations gas is only represented by GMCs, our gas fractions are much lower.

We choose  $\Sigma_b(R)$  over the stellar surface density  $\Sigma_\star(R)$ , as we are interested in the connection between disc structure and kinematics, and the strength of non-axisymmetries and thus the levels of radial disc heating and radial migration are determined by  $\Sigma_b(R)$ . Moreover, the interplay between vertical profiles and vertical velocity dispersions depends on the total mass surface density and not only on  $\Sigma_\star(R)$  and it is therefore appropriate to compare  $\Sigma_b$  in models and observations, although the division of mass between gas and stars is very different. It should be noted, that the neutral hydrogen component, which is missing in our models, will





**Figure 9.** Evolution with time of exponential scalelength  $h_R$  at  $R = 8$  kpc for various mono-age components.

have a smaller scaleheight than the stars and thus a model which has the right vertical profile, kinematics and DM halo, is not expected to agree with the Sphd  $\Sigma_b(R_0)$ .

The local radial exponential scalelength  $h_R(R_0)$  of the MW is rather uncertain. Licquia & Newman (2016) recently compiled a variety of measurements in the optical and infrared, the vast majority of which fall in the range 2–4 kpc. Their meta-analysis of 29 previous measurements yields an estimate of  $h_R(R_0) \sim 2.65$  kpc. Bovy et al. (2012) showed that populations of stars with different chemical abundances show widely varying scalelength, the most compact of which have  $h_R \sim 1.5$  kpc and the most extended of which are consistent with locally flat profiles. As we are plotting  $\Sigma_b(R_0)$  which includes GMCs, the comparison is not exact, but because at  $t = t_f$  the GMC mass fractions are 2–3 per cent as for molecular gas in the MW today, the correction is negligible for our purposes. A more relevant question is whether the missing neutral gas mass, which is a highly relevant mass component in the outer MW disc, is properly represented in our models.

In the upper row of Figure 8 we plot  $\Sigma_b(R)$  for various models and overplot a blue box indicating  $\Sigma_b(R_0) = 40 - 60 M_\odot \text{pc}^{-2}$  at  $R_0 = 8.3 \pm 0.3$  kpc and a dashed line indicating an exponential with  $h_R(R_0) = 2.65$  kpc. Due to the connection of  $\Sigma_b(R)$  to bars, we also plot the  $m = 2$  Fourier amplitude

$$A_2(R) \equiv \frac{1}{N(R)} \sum_{j=1}^{N(R)} e^{2i\phi_j} \quad (13)$$

in the lower row of Figure 8. The dashed line marks  $\ln A_2 = -1.5$ , which is used for determining the adaptive cutoff region, within which no particles are inserted in our models.

Model P2 has a constant feeding scalelength  $h_R = 2.5$  kpc, an IC disc scalelength  $h_{R,\text{disc}} = 2.5$  kpc and lives in a  $c = 9$  halo. Despite the input scalelengths the final profile is flatter than the dashed  $h_R(R_0) = 2.65$  kpc line at all but the outermost and the innermost radii. This is connected to the bar, which at  $t_f$  has a length of  $\sim 5$  kpc, similar to that of the MW bar, and lengthens scalelengths around  $R_0$  as already discussed in Paper 1. The Sphd surface density is close to the upper limit of the observed range  $= 40 - 60 M_\odot \text{pc}^{-2}$ .

Model T5 $\zeta^*$  has a more compact and more massive IC thick disc than P2, inside-out formation in the range 1.5–3.5 kpc, a mildly higher final mass and a higher GMC mass

fraction. Its thick disc also has a higher-than-average ratio of vertical to radial velocity dispersions  $\sigma_R^2/\sigma_z^2 = 1.8$ . The outcome is a model with a weaker bar and thus a  $\Sigma_b(R)$  profile closer to an exponential with a similar  $\Sigma_b(R_0)$  as P2. This is one of the models which comes closest to the inferred local profile of the MW.

Model M $\beta$ 1s5 has declining  $\sigma_0$ . It lives in a  $c = 9$  halo and has inside-out growth in the range 1.5–4.5 kpc, which generates a final exponential flatter than  $h_R(R_0) = 2.65$  kpc. Its value of  $\Sigma_b(R_0)$  is at the upper limit allowed for the MW. It has a rather weak bar, which influences the profile only at the inner radii.

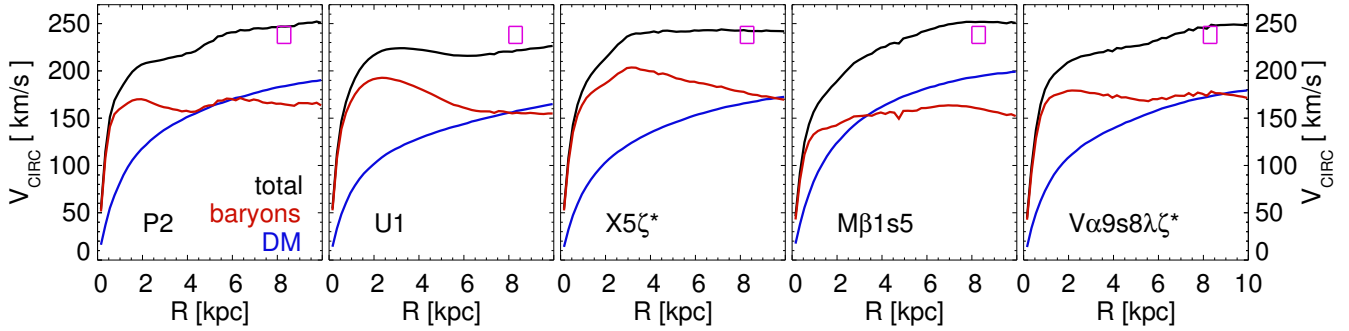
Model U1 lives in a  $c = 7.5$  halo and has a massive and compact thick IC disc and inside out formation in the range 1.5–4.5 kpc. It has a stronger and longer bar compared to the two previous models with  $c = 9$  haloes due to a higher baryon fraction as discussed in Paper 1. Its profile shows a steep bar region out to  $R \sim 6$  kpc, a flat region at  $R \sim 6 - 10$  kpc, which causes a rather low, but acceptable value of  $\Sigma_b(R_0)$  and a shallow exponential decline in the outer parts.

The declining  $\sigma_0$  model V $\alpha$ 9s8 $\lambda\zeta^*$  has a  $c = 6.5$  halo, a high final mass  $M_f = 6 \times 10^{10} M_\odot$  and grows inside out in the range 1.0–3.5 kpc. Its  $\Sigma_b(R_0)$  is too high for MW constraints and although its bar is weaker than in U1 and also P2, its  $\Sigma_b(R)$  profile is significantly flattened at  $R = 8$  kpc.

There is a clear tendency of  $A_2$  being lower in models with declining  $\sigma_0$  compared to models with thick IC discs. We will discuss this further in Section 5.

### 4.3 Radial profile evolution of mono-age components

Paper 1 and Section 4.2 have shown that the output scalelengths are somewhat independent of the input scalelengths, as bars and spirals redistribute matter. We know from observations of stars in the Sphd that the old thick populations are more compact than the young thin ones (Bovy et al. 2012). To test, how different populations of stars are affected by changes in the radial distribution, in Figure 9 we plot local scalelengths  $h_R(R_0)$  as a function of time for populations of different ages. The pink line is for IC stars and the other colours are for ten equally spaced age bins of all stars inserted during the simulations.  $h_R(R_0)$  is determined



**Figure 10.** Circular speed curves  $v_{\text{circ}}(R)$  measured in the midplane of the disc and averaged azimuthally. Blue lines mark the contributions from DM and red lines the baryonic contribution. Pink boxes mark the constraints on  $R_0$  and  $v_{\text{circ}}(R_0)$  determined by Schörrich (2012).

by a single-exponential fit to the surface density profile  $\Sigma(R)$  at  $R = 6 - 10$  kpc, irrespective of how good the fit is.

Model P1s6 has little bar activity at any stage of its evolution and is thus well suited to understand the plots. As it has inside-out growth from 1.5 to 4.3 kpc paired with a thick IC disc with  $h_{R,\text{disc}} = 2.5$  kpc, different age components are rather well separated in size. The IC component has a shorter scalelength than in the setup, because it is compressed by the disc’s gravitational field. Apart from a mild shrinking of all populations due to compression and a mild level of noise, which is likely caused by spiral activity, the output scalelengths are essentially set by the input scalelengths.

P2 has a constant input scalelength  $h_R = 2.5$  kpc and the same ICs as P1s6 and thus the populations of different ages are only mildly separated in size due to continuous compression. At  $t \sim 8$  Gyr bar formation causes an increase in  $h_R(R_0)$ , which is stronger for younger populations, so the measured scalelengths increase to  $h_R(R_0) \sim 2.5$  kpc for the oldest and  $\sim 4$  kpc for the youngest components. As already shown in Figure 8, U1 is more strongly affected by a bar. U1 has inside out growth as in P1s6 and a shorter IC disc scalelength  $h_{R,\text{disc}} = 2.0$  kpc. Bar formation at  $t \sim 4$  Gyr causes a strong increase in  $h_R(R_0)$  for all age groups and at  $t \sim 9$  Gyr bar growth causes another increase, so at  $t = t_f$ , the oldest population has  $h_R(R_0) \sim 2.5$  kpc and the youngest populations have an essentially flat profile.

The declining  $\sigma_0$  models  $M\alpha 1$  and  $V\alpha 9s8l\zeta^*$  have low-mass elliptical ICs, for which we find  $h_R(R_0) \sim 2.5$  kpc fits at early times. They both grow inside-out,  $M\alpha 1$  from 1.5 to 4.3 kpc and  $V\alpha 9s8l\zeta^*$  from 1.0 to 3.5 kpc. As the initial mass of the baryonic ICs is much lower than in P and U models, the amount of compression for the oldest components is stronger. This is especially true for  $V\alpha 9s8l\zeta^*$ , which has more compact feeding scalelengths at early times and also a shorter-than-average SFR timescale  $t_{\text{SFR}} = 6$  Gyr. The latter increases the mass in stars added at early times and thus also the mass in GMCs present at these formation stages.  $M\alpha 1$  has a bar from  $t \sim 6$  Gyr, which causes a mild increase for all  $h_R(R_0)$ , whereas bar activity is measurable from  $t \sim 3$  Gyr onwards in  $V\alpha 9s8l\zeta^*$ .

Irrespective of how long a bar is and how strongly it affects disc evolution and how high is the level of compression, the final ordering of  $h_R(R_0)$  always reflects the ordering of scalelengths at input. Additionally, in combination with the

results of Section 5, it is clear that all models with a bar similar to that of the MW show an increase for  $h_R(R_0)$  of all age components with time due to bar formation and growth.

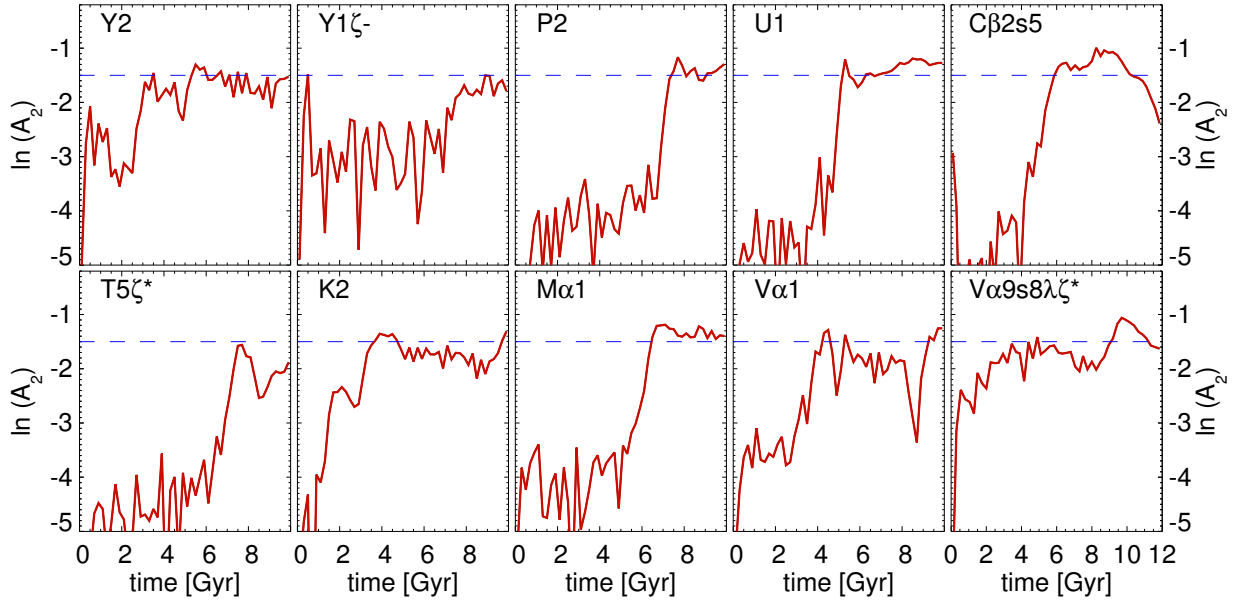
#### 4.4 Circular speed curves

Recently, various surveys of bulge/bar stars and microlensing data have enabled more detailed mass models of the centre of the MW (Wegg et al. 2016; Portail et al. 2017a; Cole & Binney 2017). These models agree in the following points: 1) The centre of the MW is baryon dominated; 2) The baryonic contribution to the rotation curve at  $R \sim 3$  kpc is  $v_{\text{circ,b}} \sim 185 \text{ km s}^{-1}$  (although Wegg et al. 2016 find an uncertainty  $\sim \pm 25 \text{ km s}^{-1}$ ); 3) The contributions of DM and baryons to  $v_{\text{circ}}$  are roughly equal at  $R_0$ . Further constraints on  $v_{\text{circ}}(R)$  come from the motion of stars in the Snhd: Schörrich (2012) finds  $v_{\text{circ}}(R_0) = 238 \pm 9 \text{ km s}^{-1}$  and  $R_0 = 8.3 \pm 0.3$  kpc.

Aumer & Schörrich (2015) (hereafter AS15) presented an inside-out growing model in a  $c = 9$  halo. Its rotation curve fulfilled the Schörrich (2012) constraints, but has too few baryons in the centre to match any of the constraints from microlensing. In Figure 10 we present circular speed curves for a selection of our models: black is total  $v_{\text{circ}}(R)$ , red is the baryonic contribution  $v_{\text{circ,b}}(R)$  and blue is the DM contribution  $v_{\text{circ,DM}}(R)$ . The pink boxes mark the Schörrich (2012) constraints.  $v_{\text{circ}}(R)$  is measured in the midplane of the disc and averaged azimuthally.

Figure 10 shows two  $c = 9$  models: P2 and  $M\beta 1s5$ . P2 has  $h_R = 2.5$  kpc both in the ICs and at all times through the simulation. Its value of  $v_{\text{circ}}(R_0)$  is at the upper end of allowed values. Its central baryonic contribution is higher than that of AS15 but still too low for the microlensing constraints.  $M\beta 1s5$  grows inside out in the range 1.5–4.3 kpc and has a higher added mass  $M_{\text{add}}$ . The former leads to a weaker central baryonic contribution than in P2 and the latter causes a stronger compression of the halo and thus an unacceptably high  $v_{\text{circ}}(R_0)$ .

The microlensing constraints suggest shifting mass from the halo to the discs. As was shown above, the Snhd DM density allows concentrations as low as  $c = 6.5$  given a constant IC halo mass of  $M_{\text{tot}} = 10^{12} M_{\odot}$ . Models U1 and  $X5\zeta^*$  have haloes with  $c = 7.5$  and  $V\alpha 9s8l\zeta^*$  has a  $c = 6.5$  halo. U1 has a massive thick IC disc with  $M_{\text{disc,i}} = 2.5 \times 10^{10} M_{\odot}$  and  $h_{R,\text{disc}} = 2.0$  kpc. It grows inside out from 1.5 to 4.3 kpc



**Figure 11.** Evolution with time of the  $m = 2$  Fourier amplitude  $A_2$  measured within  $R = 3$  kpc. The blue dashed lines mark  $\ln(A_2) = -1.5$ . Note that of the shown models only  $V\alpha 9s8\lambda\zeta^*$  and  $C\beta 2s5$  have  $t_f = 12$  Gyr and all others have  $t_f = 10$  Gyr.

reaching a final mass of  $M_f = 6 \times 10^{10} M_\odot$ . Its baryonic contribution to the rotation curve peaks at  $v_{\text{circ},b} \sim 190 \text{ km s}^{-1}$  and falls below the DM contribution at  $R \sim 8 \text{ kpc}$  and thus fulfils all microlensing constraints. Its Sphd  $v_{\text{circ}}(R_0)$  is lower than that found by Schönrich (2012).

Model  $X5\zeta^*$  has a thick IC disc with  $M_{\text{disc},i} = 2.0 \times 10^{10} M_\odot$  and  $h_{R,\text{disc}} = 2.0 \text{ kpc}$  and in addition an IC bulge with  $a_{\text{bulge}} = 0.7 \text{ kpc}$  and  $M_{\text{bulge},i} = 0.5 \times 10^{10} M_\odot$ . It has inside out growth in the range  $1.5 - 3.5 \text{ kpc}$  and the same final mass as U1. Consequently, its peak  $v_{\text{circ},b}$  is higher at  $\sim 200 \text{ km s}^{-1}$ , but it still fulfils all constraints from microlensing, as well as the Sphd  $v_{\text{circ}}(R_0)$  constraints. The declining  $\sigma_0$  model  $V\alpha 9s8\lambda\zeta^*$  also fulfils all constraints but in a different way. It starts from a low mass elliptical IC and grows a disc with  $M_f = 6 \times 10^{10} M_\odot$  like those of two previous models. Its inside-out growth is from  $1.0$  to  $3.5 \text{ kpc}$  and its final  $v_{\text{circ},b}(R)$  is rather constant at  $170 - 180 \text{ km s}^{-1}$  in the range  $R = 2 - 10 \text{ kpc}$ . Unlike  $X5\zeta^*$ , which has a flat total circular speed with  $240 \text{ km s}^{-1}$  for  $R = 3 - 10 \text{ kpc}$ ,  $V\alpha 9s8\lambda\zeta^*$  has  $v_{\text{circ}}(R)$  increasing in this radial range from  $220 \text{ km s}^{-1}$  to  $250 \text{ km s}^{-1}$ .

## 5 BAR FORMATION AND EVOLUTION

Paper 1 showed that models starting with a thin-disc IC and having no GMC heating undergo strong bar activity from early times. GMC heating can delay and weaken bar formation and evolution and in extreme cases prevent the formation of a strong bar over cosmological timescales. In this Section, we examine how this picture is modified by an old thick disc component.

### 5.1 Bar strengths

Figure 11 displays for ten models the evolution of the  $m = 2$  Fourier amplitude  $A_2$  (see Equation 13) for all the stars within  $R = 3 \text{ kpc}$ . Model Y2 represents thin-disc-only models

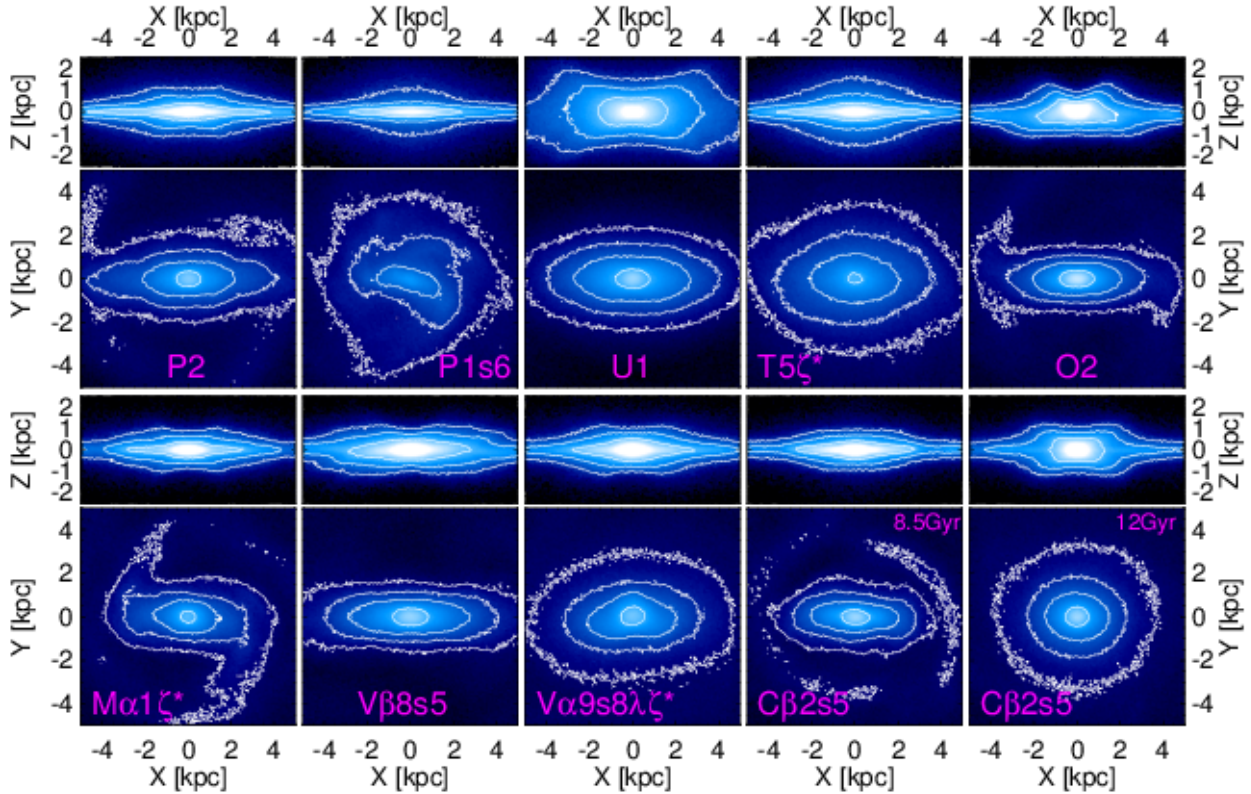
and we see that  $\ln(A_2)$  instantaneously increases to  $-2.5$  as the addition of mass to the thin and compact IC disc makes the system develop non-axisymmetries. As discussed in Paper 2, the radial heating by GMCs at low disc mass and high SFR is important and delays the formation of a strong bar during the first 3 Gyr of evolution in Y2. This effect is enhanced in model Y1 $\zeta^-$ , which has  $\zeta = 0.04$  and thus twice as many GMCs per unit mass of added stars:  $\ln(A_2)$  is kept at  $\sim -3$  until  $t \sim 7 \text{ Gyr}$ , when it increases to  $\sim -1.7$ , indicating a rather weak bar.

The curve for model P2 is very different. P2 shares with Y2, the constant radial growth history  $h_R(t) = 2.5 \text{ kpc}$ , the final mass  $M_f = 5 \times 10^{10} M_\odot$  and the shape of the SFH. Its IC disc is thicker, more extended and more massive than the one in Y2 and the normalisation of its SFH is thus lower. The existence of a hot disc, which is stable against bar formation prevents the growth of  $A_2$ , although stars are continuously added on cold orbits throughout the simulation. Only at  $t \sim 7 \text{ Gyr}$  has enough thin disc been accumulated to make the composite system unstable to bar formation.

Paper 1 showed that, in the absence of a thick disc, lowering the halo concentration from  $c = 9$  is problematic, as the system becomes more self-gravitating and bar unstable. U1 has a more massive and more compact IC thick disc than P2, lives in a  $c = 7.5$  halo and has an inside-out growth history. Despite the higher thick disc central surface density and the lower halo concentration than P2, the thick disc still suppresses bar formation for 4 Gyr.

Another ingredient here is the radial to vertical velocity dispersion ratio in the IC thick disc. Model T5 $\zeta^*$  has a more massive and more compact IC thick disc and a slightly higher final mass than P2 and inside out formation from  $1.5$  to  $3.5 \text{ kpc}$ . Despite the higher surface densities,  $\sigma_R^2/\sigma_z^2 = 1.8$  for the ICs in T5 $\zeta^*$  compared to a value of 1 in P2 makes the system more stable against bar formation and thus weakens bar formation more strongly than in P2.

To understand bar evolution in models with declining



**Figure 12.** Edge-on and face-on surface density maps of the bar regions in several of our models. Most models are shown at  $t = t_f$ , only Cβ2s5 appears additionally at  $t = 8.5$  Gyr.

$\sigma_0$ , which start with low mass elliptical ICs, we first examine model K2 which grows a thin disc inside a higher mass elliptical IC. K2 and P2 differ only in that P2 has a thick-disc IC of the same mass. During the early evolution phases of K2,  $A_2$  is suppressed due to the elliptical ICs. However, in K2  $A_2$  increases at an earlier time than in P2 and from  $t \sim 3$  Gyr on shows a significant bar. By construction, the in-plane velocity dispersions of the IC stars at  $R \lesssim 2$  kpc are similar in these models. However, the surface densities are higher and thus the rotation velocities of the IC stars are faster in K2. Consequently, K2 is more unstable to bar formation than P2.

Compared to the K ICs, the M ICs contain an elliptical, which is three times less massive. So a cold disc model in M would have high  $A_2$  at an earlier time than K2. The hot input dispersions for the old populations in Mα1, however, act in the same way as the thick IC disc in the P models and delay bar formation until  $t \sim 6$  Gyr. We find that the specific shape of declining  $\sigma_0(t)$  does not significantly influence bar formation.

Vα1 is the equivalent model to Mα1, but it lives in a lower concentration  $c = 6.5$  halo. Still, bar formation is delayed until  $t \sim 4$  Gyr. Vα9s8λζ\* has a higher mass, a more compact feeding history and a shorter SFR timescale  $t_{\text{SFR}}$  than Vα1. All three factors lead to a much faster increase in surface density at early times, which outweighs the fact that at feeding the radial-to-vertical input dispersion ratio  $\sigma_R/\sigma_z = 1.25$  is higher than in Vα1. Consequently,  $A_2$  is higher at early times.

The specifics of the SFH, the radial growth history, the dispersions of the old components and GMC heating thus determine the bar formation history of an individual model. Figures 8 and 11 however show that halo concentrations  $c = 6 - 7$  allow models with reasonable final bars in the presence of hot disc components. The fact that in the lower row of Figure 8, the  $m = 2$  amplitudes at final times and radii  $R > 5$  kpc are lower in models with declining  $\sigma_0$  than in the thick IC disc models are connected to the gradients of  $\sigma_R$ . In models with thick-disc ICs, the oldest stars have radially constant scaleheights and  $\sigma_R/\sigma_z$  and thus declining  $\sigma_R(R)$ , whereas models with declining  $\sigma_0$  assume a radially constant input dispersion  $\sigma_R$ . Thus in the end, the outer thick components are radially hotter in models with declining  $\sigma_0$  and thus less unstable to  $m = 2$  modes.

## 5.2 Bar morphology

The central region of our Galaxy is dominated by a bar, the inner part of which consists of a boxy/peanut-shaped bulge at  $R < 2$  kpc with an X-shape at  $|z| > 500$  pc (Wegg & Gerhard 2013) surrounded by a vertically thin part, the *long bar*, extending to  $R \sim 4 - 5$  kpc (Wegg et al. 2015). AS15 demonstrated that an inside-out growing model without GMCs in a  $c = 9$  halo produces a bar with X-shaped structure with the tips of this structure at  $(x, z) \sim (2, 1.3)$  kpc, very similar to the structure of the MW bulge/bar region inferred by Wegg & Gerhard (2013). Paper 1 showed that some thin-disc-only models with GMCs also displayed bars very similar to the



one in the MW, but also noted that not all of these bars are buckled and that lower concentration haloes favoured unrealistically long bars.

Here we test how well our thick disc models can reproduce the MW bar. This is interesting as the chemically defined thick disc is concentrated and should thus have a high mass fraction in the bar region, but can only form a bar if the thin disc fraction is high enough, as discussed above. Figure 12 shows edge-on and face-on surface density maps of several galaxies at  $t_f$ .

Like the standard Y models in Paper 1 and the model in AS15, P2 lives in  $c = 9$  halo and shows a bar which is  $\sim 5$  kpc long. Model P1s6, which has a more radially extended feeding history shows only a small, weak bar in the central  $R < 2$  kpc. The vertical structure of P1s6 is indistinguishable from a pure disc galaxy, whereas P2 has a boxy shape with a lateral extent of  $\pm \sim 2$  kpc and a vertical extent of  $\pm \sim 1$  kpc and a mild hint of an X structure. As in the MW bar, the outer regions are thinner. T5 $\zeta^*$  has a more compact and more massive, but radially hotter thick disc and does not show a bar, just mildly elliptical surface density contours.

The galaxy in model O2 evolves in a  $c = 9$  halo from an elliptical IC. At  $t_f$ , its bar is  $\sim 4$  kpc long and is currently buckling as indicated by the broken mirror symmetry relative to the  $x$ -axis as first observed in a simulation by Raha et al. (1991). This event will eventually produce an edge-on peanut bulge with a characteristic X-shape (see Combes & Sanders 1981). U1 has a compact, massive IC thick disc, a higher than average final mass and lives in a  $c = 7.5$  halo. Its bar grows to a length  $\sim 6$  kpc and is thus only mildly longer than the P2 bar. However, bar buckling has created a significantly more extended X-shape with a lateral extent of  $\pm \sim 4$  kpc and a vertical extent of  $\pm \sim 2$  kpc.

The lower row displays bars in models with declining  $\sigma_0$ . Model C $\beta$ 2s5 grows a disc around a compact bulge IC in a  $c = 9$  halo. At  $t_f$ , it clearly shows an X-shaped edge-on structure in the central  $\sim 2$  kpc, stronger than the similarly sized one in P2. However, its face-on image reveals an almost axisymmetric image. Going back in time, we find that at  $t = 8.5$  Gyr C $\beta$ 2s5 exhibited a strong bar that had not yet buckled. From the evolution of  $A_2(R < 3$  kpc) depicted in Figure 11, we learn that the bar formed around  $t = 6$  Gyr. Between  $t = 8.5$  Gyr and  $t_f$  C $\beta$ 2s5 undergoes buckling, but in the final  $\sim 2$  Gyr of the simulation its bar becomes continuously weaker.

Models M $\alpha$ 1 $\zeta^*$  ( $c = 9$  halo) and V $\beta$ 8s5 and V $\alpha$ 9s8 $\lambda\zeta^*$  ( $c = 6.5$  halo) are models with declining  $\sigma_0$  starting from low-mass elliptical ICs. They display bars of reasonable sizes (3.5 – 5.5 kpc) and varying strengths. Their vertical profiles all show boxy edge-on shapes with vertical extents  $\pm \sim 1.0$  – 1.5 kpc and lateral extents which are wider for longer bars.

In summary, it is possible to create models which show reasonable agreement with the vertical profile of the MW, its circular speed curve and at the same time contain a bar, which in length, strength and vertical extent agrees reasonably with that of the MW. The problem is that the details depend on mass and size growth history, DM halo density and GMC heating and that, additionally, the evolution of bar length and strength is to some degree stochastic (Sellwood & Debattista 2009), so that, not even within the limits of our methods, is it possible to determine which model best represents our Galaxy.

### 5.3 Bar age structure

In Figure 13, we examine the age structure of the bar. For several models we plot a map of median age  $\tau_{\text{med}}$  in the  $x$ - $z$  plane, where  $x$  is measured along the major axis of the bar and  $z$  is perpendicular to the disc. All stars with  $|y| < 1.5$  kpc are considered for the map. If  $\tau_{\text{med}}$  at a certain position corresponds to a star particle from the ICs, we apply a distinct yellow colour, whereas inserted star ages range from black (young) to orange (old). We overplot in white edge-on density contours.

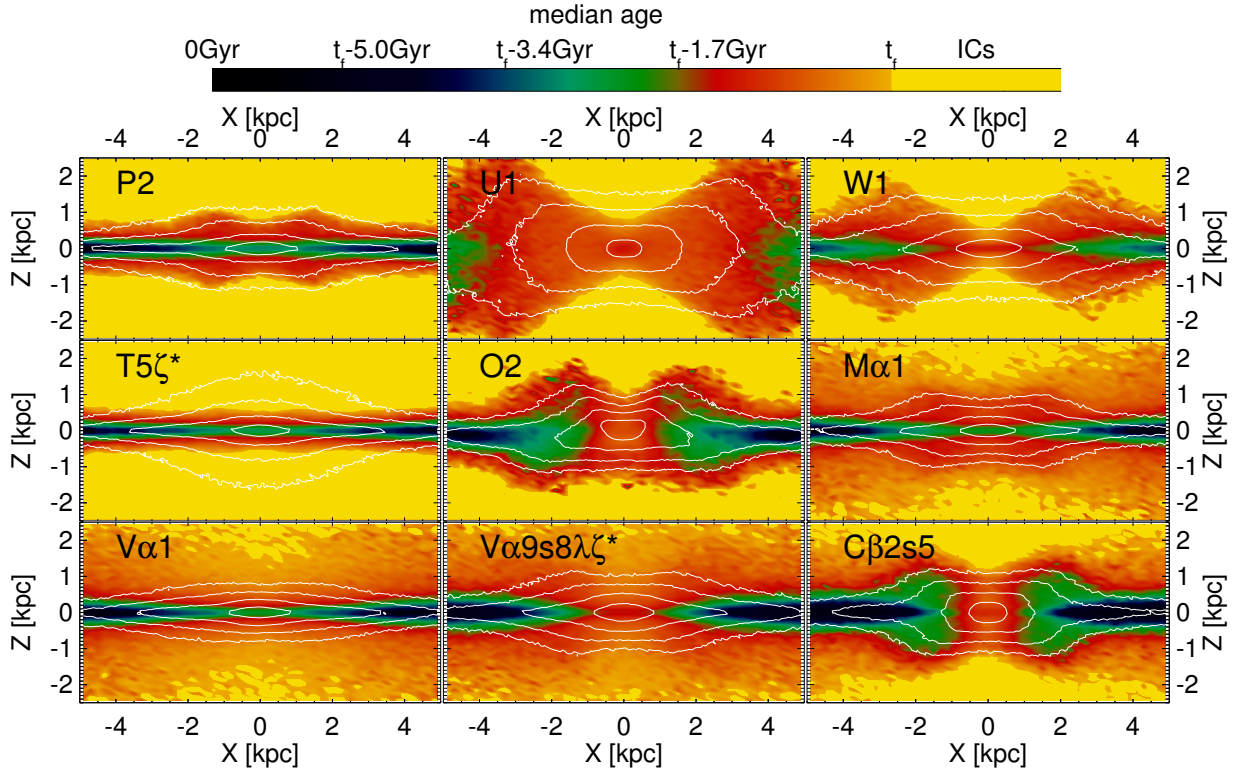
We start by analysing model T5 $\zeta^*$ , which does not show a noticeable bar, in the left-hand panel of the middle row. Clearly, above  $|z| \sim 600$  pc the IC thick disc dominates at all  $x$ . The thin disc is youngest in the plane and has a vertical age gradient, as an effect of stars being born cold and being vertically heated by GMCs. It also has a radial age gradient, due to inside-out formation.

In contrast to model T5 $\zeta^*$ , the three models displayed in the top row of Figure 13 at  $t_f$  show X-shaped edge-on structures. These three models all grow thin discs within thick IC discs. The X-shape is beautifully visible in the  $\tau_{\text{med}}$  maps. The cones of relatively lower surface density above and below the galactic centres are dominated by old, thick-disc stars, whereas younger stars fill diamonds which lie at both sides of the cones in the  $x$ -direction. These diamond shaped regions have vertical age gradients of different strengths. This is connected to the fact that we do not feed new stars into regions of strong bars. The bar in U1 forms already at  $t = 4$  Gyr, whereas the bar in P2 only forms at  $t = 7.5$  Gyr. As shown in AS15, stars can be captured by the bar, but the rates of this process are low, so the age structure is only mildly affected. Consequently, the number of relatively young stars present in P2 is high and vertical and radial age gradients are present in the bar region. In U1, the region at  $|x| < 4$  kpc, which is part of the diamonds, is well mixed in  $\tau_{\text{med}}$ . The bar in W1 forms at  $t = 6$  Gyr and the model is thus intermediate to P2 and U1 in terms of age gradients.

The vertical mixing of stars of various ages happens during bar buckling, as is illustrated by model O2, which is undergoing the process at the depicted moment. We see how the younger stars are spread vertically at around  $|x| \sim 2$  kpc and the model thus transitions from a disc-like to an X-shaped age structure. Note that, as bars can vary in length and strength with time, cold, young stars can be added in the plane after buckling events and there can be multiple buckling events, which complicates the age structure.

The remaining four models in Figure 13 feature declining  $\sigma_0$ . As the thick discs of these models comprise fed-in stars, which have radial and vertical density distributions that vary continuously with age, the age structure is more complicated to interpret, as in real galaxies. At  $t_f$ , V $\alpha$ 1, like T5 $\zeta^*$ , has no X-shaped edge-on structure. The thin discs of these galaxies exhibit qualitatively similar vertical and radial age gradients. The thick disc of V $\alpha$ 1 shows a continuation of these gradients at higher  $|z|$ , the gradients in these regions are, however, much shallower.

M $\alpha$ 1 and V $\alpha$ 9s8 $\lambda\zeta^*$  have edge-on peanut-shaped density structures similar to that of P2 and, indeed, the age structures of the three bars are qualitatively similar. As the thick disc background, however, is, again, much less distinct,



**Figure 13.** Median ages  $\tau_{\text{med}}$  as a function of positions  $(x, z)$  in edge-on views of the central regions of several models.  $x$  is along the major axis of the bar. Surface density contours are overplotted in white. Models are shown at  $t = t_f$ .

the structure combining old cones plus younger diamonds is much less evident and might be hard to measure in a real galaxy.

Model C $\beta$ 2s5 differs from the other models with declining  $\sigma_0$  in having compact bulge ICs and thus the density of IC stars near the galactic centre is higher and the  $\tau_{\text{med}}$  gradients away from the centre are steeper than in models with low-mass elliptical ICs. A relatively recent and strong buckling event before the disappearance of the bar and a radial growth history with constant  $h_R = 2.5$  kpc has led to relatively young areas at  $|x| \sim 2$  kpc.

## 6 DISCUSSION

### 6.1 Vertical profiles and age gradients

In all final models the vertical profile of the Sphd is fitted well by the sum of two exponentials similar to that of the MW. When the IC contains a thick disc, the vertical profile is double-exponential from the outset, with a growing thin-to-thick density ratio, whereas when  $\sigma_0$  declines, the vertical profile gradually evolves a double-exponential structure as a sufficiently massive, cold and thin population forms, and the scaleheights of both components change continuously.

We cannot directly measure the evolution of scaleheights over cosmic time, but in nearby galaxies we can probe this evolution through observations of the radial variation of the vertical profile and age structure of the disc. The radial variation of the vertical profile is dominated by the evolution of the thick disc because our thin discs have scaleheights that are almost independent of both  $R$  and  $t$ .

Our thick-disc ICs are set up with declining  $\sigma_z(R)$  and radially constant scaleheights, and the latter property is roughly conserved to the present epoch, as radial migration causes only mild levels of flaring. Models with declining  $\sigma_0$  have, by construction, radially constant  $\sigma_z$  at birth. As the oldest and hottest stars are barely affected by vertical heating, their thick discs flare strongly and their scaleheights increase with radius. In these models, radial migration weakens flaring, but the effect is not strong enough to balance the outwards increase in scaleheight imprinted at birth.

This has important consequences on the age structure of the disc. At radii  $R = 5\text{--}12$  kpc and altitudes  $|z| > 1$  kpc models with thick-disc ICs are dominated by old IC disc stars, they thus show no radial variation in median age  $\tau_{\text{med}}$ . At lower altitudes these discs become younger with increasing  $R$ . The age structure is markedly different in models with declining  $\sigma_0$ . On account of the strongly flaring old and intermediate age components, they show negative  $\tau_{\text{med}}$  gradients at all  $|z|$ . Models with declining  $\sigma_0$  thus agree better with recent measurements of the radial age structure at various altitudes in the MW by Martig et al. (2016b). They also show somewhat flatter vertical age gradients at  $R_0$ , in rough agreement with measurements by Casagrande et al. (2016). For thick IC disc models, strong inside-out growth improves the agreement with these measurements.

Our discs evolve in isolation, whereas at least the low-density outskirts of discs are likely to be affected by processes capable of significant vertical thickening, such as disc-satellite interactions (Kazantzidis et al. 2008) or infall of gas with misaligned angular momentum (Jiang & Binney 1999). Including minor mergers or adding stars in tilted outer discs would thus be a valuable extension of our models.

## 6.2 The interplay between the dark halo, the thick disc and the bar

The evolution of a disc depends on the local density of the dark halo because increasing the latter reduces the extent to which the disc controls the gravitational field in which it moves. In particular, decreasing the local DM density increases the amplitude of non-axisymmetric structure. We varied the local DM density by varying the initial concentration parameter  $c$  at a fixed halo mass,  $M_{\text{tot}} = 10^{12} M_{\odot}$ . Concentrations in the range  $c = 6-9$  work well. Indeed, after the DM has been compressed by the disc, the final DM density in the Sphd then agrees with observational constraints, and in many models a bar similar to that of the MW emerges before the current epoch. Although non-axisymmetric structures in the disc transfer angular momentum to the DM, the dark halo does not acquire a core like that favoured in the MW (Cole & Binney 2017). If such a core exists in a  $\Lambda$ CDM context, it thus probably formed in the very early evolution stages of the galaxy that are not modelled here.

Observational constraints on the circular speed,  $v_{\text{circ}}(R_0)$ , near the Sun (Schönrich 2012) and on the microlensing optical depth towards the MW bar/bulge (Wegg et al. 2016; Cole & Binney 2017) indicate a baryon dominated central MW and roughly equal contributions of DM and baryons to  $v_{\text{circ}}(R_0)$ . To achieve this in our models, initial concentrations  $c \sim 6-7$  are favoured as higher values do not allow for enough baryonic mass in the central regions.

Paper 1 favoured  $c \sim 9$  because thin-disc-only models with lower values of  $c$  showed unrealistically long and strong bars. The presence of old thick, and thus kinematically hot, disc components alleviates this problem. By shifting mass from the cold thin disc to a thick, radially hot component, the formation of a bar can be delayed by several Gyrs. It is immaterial whether the hot component is included in the ICs or arises from declining  $\sigma_0$ , and the ratio of radial to vertical dispersions can be  $\sigma_R/\sigma_z \sim 1$ , so smaller than the values  $\sigma_R/\sigma_z \sim 2$  in thin discs. As a consequence, the time available for the bar to grow in strength and length is limited. Indeed, models with  $c = 9$  haloes tend to show an unrealistically weak final bar. Models with  $c \sim 6-7$  generally have bars with lengths similar to that of the MW's bar. The delaying of bar formation by old thick discs also explains the observation that the fraction of barred disc galaxies decreases with increasing redshift (Sheth et al. 2008).

## 6.3 The edge-on structure of bars

Several models have bars that are morphologically similar to that of the MW: a boxy bulge in the central  $R \lesssim 2$  kpc with an X-shape up to  $|z| \sim 1-1.5$  kpc at the centre of a thinner outer bar that extends to  $R \sim 5$  kpc (Wegg & Gerhard 2013; Wegg et al. 2015). In one of these models the bar dissolved at the very end of the simulation. However, notwithstanding the face-on surface density being almost axisymmetric, the edge-on peanut shape survived. Thus not all observed boxy edge-on bulges need be bars.

In edge-on density projections of these models, the X-shapes are not as striking as in the model of AS15, which lacks a thick disc and GMCs, but features an isothermal gas component. This is not surprising as the MW bulge does not show an X-shape in all stellar components. Dékány et

al. (2013) found that old and metal-poor RR Lyrae stars appear to have a more spheroidal shape and Portail et al. (2017b) found that low-metallicity stars ( $[\text{Fe}/\text{H}] < -0.5$ ) in the bar/bulge contain a much lower fraction of stars on bar-supporting orbits than stars with higher metallicities. Moreover, the X is possibly absent in younger populations as well (López-Corredoira 2016). Analysis on the edge-on age structure of our bars offers insight into why the shape should vary with stellar population.

The X structures are particularly pronounced in models with thick-disc ICs that have final boxy bulge/bar regions which are more extended than in the MW. Here bar buckling spreads stars from the thin disc vertically, and because they mainly populate the 2:2:1 resonant orbit family (Pfeniger & Friedli 1991), these stars form a structure that, seen edge-on, resembles two diamonds overlapping at the galactic centre. The cones above and below the centre, which are not populated by the 2:2:1 orbits, are dominated by thick-disc stars. Young stars that were captured by the bar after the buckling event will be found in the plane. Consequently, a distinct age pattern should be observed in buckled edge-on bars.

A similar separation in edge-on morphology between the oldest, the intermediate age and the youngest stars in a barred galaxy has also been found in the simulations of Athanassoula et al. (2016). Debattista et al. (2016) recently studied bar formation in galaxies that contain disc populations with differing random motions. They showed that radially cooler populations form stronger bars, the edge-on profiles of which are vertically thinner and peanut-shaped, whereas the hotter populations form a weaker bar with a vertically thicker edge-on box shape (see also Fragkoudi et al. 2017).

The thick and thin discs of models with declining  $\sigma_0$  are not strictly separated in age, as they are in models with thick-disc ICs, and in consequence their characteristic age structure is less clear in an edge-on age map. Thus while observations of the edge-on age structures of bars have the potential to betray the formation history of the bulge, central thick and thin discs and the timing of the bar buckling event, the constraints will be less tight if the scenario with declining  $\sigma_0$  is more appropriate than that in which the thick disc is included in the ICs.

## 6.4 Radial redistribution and inside-out growth

Bars and spiral structure make it hard to steer the disc's radial scalelength  $h_R$  at the Solar radius  $R_0$  to a preferred value,  $h_R(R_0) \sim 2.6$  kpc. As was already discussed in Paper 1, making the disc more compact results in stronger non-axisymmetries, which in turn leads to more mass redistribution and larger  $h_R(R_0)$ . We have shown that this is the case for all age components and for all models with appropriate bars, so in the past  $h_R(R_0)$  would likely have been smaller than it is today. Assigning higher values of  $\sigma_R/\sigma_z$  to the old thick disc can yield steeper-than-average profiles, but doing so weakens bars inappropriately.

Still, for models that grow inside-out,  $h_R(R_0)$  always increases with decreasing age, just as observations of the Sphd suggest (Bovy et al. 2012). As these observations show scalelengths  $h_R(R_0) \sim 1.5-2.0$  kpc for the most compact and oldest mono-abundance populations (see also Cheng et



al. 2012),  $h_R = 2$  kpc can be regarded as an upper limit on the birth scalelength at the earliest times, but Model  $V\alpha 9s8\lambda\zeta^*$  demonstrates that the scalelength at birth could have been as small as  $h_R = 1$  kpc, a conclusion similar to that of Schönrich & McMillan (2017). The flat age metallicity relation of the Sph and the radial metallicity gradient in the MW make it hard to infer the scalelength of current star formation by studying mono-abundance populations. In our models input scalelength at late times in the range  $h_R \sim 3 - 4$  kpc give reasonable results.

### 6.5 Thick disc formation scenarios

The age structure of the MW disc points towards a model with declining  $\sigma_0$ . Most hydrodynamical cosmological simulations of disc galaxies support a picture in which birth dispersions and gas fractions decline continuously with time (e.g. Bird et al. 2013; Stinson et al. 2013; Ma et al. 2017). Such a scenario has also been inferred from observations of H $\alpha$  kinematics (e.g. Kassin et al. 2012; Wisnioski et al. 2015). However, Di Teodoro et al. (2016) argue that there is no substantial difference between the gas kinematics of galaxies at redshifts  $z_{rs} \sim 1$  and today. Moreover, Martig et al. (2014) find that models that have kinematics in line with those found in the Sph favour a two-phase formation scenario, in which the thick-disc stars are born in a turbulent, merger-dominated phase and the thin-disc stars are born cold and heated subsequently.

Such a two-phase scenario motivated our models with thick-disc ICs, but in setting up an equilibrium thick stellar disc with a radially constant scaleheight we have ignored correlations between stellar ages, kinematics and density profiles that would naturally arise during formation of the proto-thick disc. Declining  $\sigma_0$  inherently produces such correlations. However, an inappropriate thick disc is still liable to emerge through poor choices for  $\sigma_0(t)$  or  $h_R(t)$ , or the choice of a radially constant  $\sigma_0$ .

The main limiting factor of our models is thus the lack of a self-consistent heating mechanism for thick-disc stars: in both scenarios thick-disc stars are created ad-hoc. The heating mechanism will affect non-axisymmetries and the radial distribution of matter, which are crucial for disc evolution. As measurements of gas fractions at redshifts  $z_{rs} \sim 2$ , a time consistent with the formation of the chemically defined thick disc of the MW, indicate that molecular gas makes up 50 per cent or more of the baryonic masses of galaxies (e.g. Genzel et al. 2015), the lack of a realistic gas component is a connected problem. Although at early times models such as  $V\alpha 9s8\lambda\zeta^*$  have as much as 45 per cent of their baryonic mass in GMCs, the GMCs have the same mass function as a present-day spiral galaxy, in which the gas fraction is lower and stars form cold. In a picture in which turbulence driven by gravitational disc instabilities causes stars to form with large dispersions (Forbes et al. 2012), molecular complexes would be expected to be more massive. Scattering of stars by massive clumps could contribute to thick disc formation (Bournaud et al. 2009).

## 7 CONCLUSIONS

We have presented a new set of idealised  $N$ -body simulations of disc galaxies with both thin and thick discs within live dark haloes. These models are grown over 10 – 12 Gyr by continuously adding new stellar particles with specified age-dependent velocity dispersions. Short-lived massive particles represent GMCs. Thin discs grow by the addition of stars on near-circular orbits, whereas for thick disc components we rely on two different concepts: a) create an appropriate thick disc in the ICs and only add thin-disc stars during the simulation, or b) start with low-mass, diffuse elliptical or compact bulge ICs and add stars with continuously declining input velocity dispersion  $\sigma_0(t)$ . Hence in scenario b) we form kinematically hot thick disc stars at early times and cold thin populations at late times, whereas in scenario a) the simulation starts after the structure that will morph into the thick disc is fully formed. To understand the evolution of our models, we simulate a variety of histories of star formation, dark halo densities and thick disc properties.

Both types of models can produce at final time  $t_f$  models which are similar in structure to the MW. We find:

- Both scenarios create double-exponential vertical profiles. The scaleheight of the thin disc is governed by GMC heating. To achieve a MW-like thick disc exponential scaleheight  $h_{\text{thick}} \sim 1$  kpc at  $t_f$ , thick-disc ICs with isothermal vertical scaleheights  $z_0 \sim 1.7$  kpc are suitable. For declining  $\sigma_0$  models, the input velocity dispersions should be  $\sigma_0 \sim 40 - 50$  km s $^{-1}$  at the earliest formation stages. Thick disc scaleheights are not affected by GMC heating.
- Models need to undergo inside-out growth to reproduce the observed dependence of radial scale-length  $h_R$  on chemical composition of disc stars. We find that models which grow from  $h_R \sim 1 - 2$  kpc at early times to  $h_R \sim 3 - 4$  kpc today are suitable.
- To explain the baryon dominance of the Galactic centre, the circular speed curve of the MW and the structure of the bar in the presence of a thick disc, we favour DM haloes which at mass  $M_{\text{DM}} = 10^{12} M_\odot$  have an initial concentration parameter  $c \sim 7$ .

It is essential that thick disc stars are already hot when the thin disc starts forming because Paper 1 showed that heating by GMCs and non-axisymmetries is incapable of producing the thick disc, although it explains the properties of the thin disc. The presence of the thick disc modifies the evolution of the thin disc, but the final properties of the thin discs in our thin+thick disc models are similar to those of thin-disc-only models in slightly more concentrated dark haloes. Crucially, this change in halo density and the presence of a hot and thick disc make it possible to bring models with appropriate bars into agreement with the baryon fractions inferred for the central MW.

Regarding the non-axisymmetric structures of the disc, we find:

- Bars with a structure similar to that of the MW bar, i.e. a boxy/peanut-shaped bulge at  $R < 2$  kpc with an X-shape surrounded by a vertically thin part extending to  $R \sim 4 - 5$  kpc, can be found in some of our viable models. Stochasticity in the evolution of bar lengths and strengths complicates the comparison.
- The presence of a hot, thick-disc stellar popula-



tion at the start of thin-disc formation suppresses non-axisymmetries and delays the formation of the bar.

- In models with an appropriate bar, the local exponential scalelengths  $h_R(R_0)$  of all mono-age populations are increased by the radial redistribution of matter that the bar and spirals generate. Populations measured in the Sphd today have thus likely had lower  $h_R(R_0)$  in the past.
- The dark halo's density profile is modified by the growth of the disc and the non-axisymmetric structures that form in the disc, but the profile does not develop a central core as is currently favoured for the centre of the MW.

To distinguish between formation scenarios, it is helpful to study the radial and vertical age structure of disc galaxies. We find:

- Our two types of models for the creation of the thick disc differ significantly in the predicted age maps of the discs. The observed Sphd radial age gradient at  $|z| > 1$  kpc and the vertical age gradient both favour models with declining  $\sigma_0$ . However, the measurements are still rather uncertain and models with thick IC discs by construction ignore any internal structure of the thick disc. Such structure could be added to these models.
- Bar buckling in thin+thick disc systems creates characteristic age patterns in edge-on views of the bar region. Buckling predominantly affects thin-disc stars and causes them to form a structure in the  $(R, z)$  plane resembling two diamonds overlapping at the galactic centre. The cones above and below the centre are dominated by thick-disc stars.

Considering the wealth of data on the structure of the MW that will soon become available from surveys such as Gaia (Gaia Collaboration 2016), evolutionary models of disc galaxies that grow over cosmological timescales and contain both thick and thin discs will be essential to connect the data to the formation history of the MW. Our models allow for a relatively controlled and flexible setup, can be produced in large numbers and capture a wealth of important dynamical processes.

We have demonstrated that our models can reasonably reproduce a variety of observations of the structure of the MW. No model sticks out as particularly similar to the MW in all aspects, but this is to be expected given the remaining shortcomings in modelling. In a companion paper (Paper 4) we examine the models presented here in light of Sphd kinematics and constraints on radial migration.

## ACKNOWLEDGEMENTS

We thank the referee for comments that helped improve the paper. It is a pleasure to thank Ralph Schönrich for valuable discussions and comments on the manuscript.

This work was supported by the UK Science and Technology Facilities Council (STFC) through grant ST/K00106X/1 and by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 321067. This work used the following compute clusters of the STFC DiRAC HPC Facility (www.dirac.ac.uk): i) The COSMA Data Centric system at Durham University, operated by the Institute for Computational Cosmology. This equipment was funded by a BIS National E-infrastructure capital

grant ST/K00042X/1, STFC capital grant ST/K00087X/1, DiRAC Operations grant ST/K003267/1 and Durham University. ii) The DiRAC Complexity system, operated by the University of Leicester IT Services. This equipment is funded by BIS National E-Infrastructure capital grant ST/K000373/1 and STFC DiRAC Operations grant ST/K0003259/1. iii) The Oxford University Berg Cluster jointly funded by STFC, the Large Facilities Capital Fund of BIS and the University of Oxford. DiRAC is part of the National E-Infrastructure.

## REFERENCES

- Athanassoula E., Rodionov S. A., Peschken N., Lambert J. C., 2016, *ApJ*, 821, 90
- Aumer M., Binney J. J., 2009, *MNRAS*, 397, 1286
- Aumer M., Schönrich R., 2015, *MNRAS*, 454, 3166 (AS15)
- Aumer M., Binney J., Schönrich R., 2016a, *MNRAS*, 459, 3326 (Paper 1)
- Aumer M., Binney J., Schönrich R., 2016b, *MNRAS*, 462, 1697 (Paper 2)
- Aumer M., Binney J., Schönrich R., 2017, submitted to *MNRAS* (Paper 4)
- Bensby T., Feltzing S., Lundström I., 2003, *A&A*, 410, 527
- Binney J., Tremaine S., 2008, *Galactic Dynamics: Second Edition*, Princeton University Press, Princeton
- Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, *ApJ*, 773, 43
- Bland-Hawthorn J., Gerhard O., 2016, *ARA&A*, 54, 529
- Bournaud F., Elmegreen B. G., Martig M., 2009, *ApJ*, 707L, 1
- Bovy J., Rix H.-W., Liu C., Hogg D. W., Beers T. C., Lee Y. S., 2012, *ApJ*, 753, 148
- Brook C. B., Kawata D., Gibson B. K., Freeman K. C., 2004, *ApJ*, 612, 894
- Casagrande L. et al., 2016, *MNRAS*, 455, 987
- Cheng J. Y. et al., 2012, *ApJ*, 752, 51
- Chiappini C., Matteucci F., Gratton R., 1997, *ApJ*, 477, 765
- Cole D. R., Binney J., 2017, *MNRAS*, 465, 798
- Combes F., Sanders R. H., 1981, *A&A*, 96, 164
- Debattista V. P., Ness M., Gonzalez O. A., Freeman K., Zoccali M., Minniti D., 2016, submitted to *MNRAS*, arXiv:1611.09023
- Dékány I., Minniti D., Catelan M., Zoccali M., Saito R. K., Hempel M., Gonzalez O. A., 2013, *ApJ*, 776, L19
- Di Teodoro E. M., Fraternali F., Miller S. H., 2016, *A&A*, 594, 77
- Forbes J., Krumholz M., Burkert A., 2012, *ApJ*, 754, 48
- Förster Schreiber N. M. et al., 2009, *ApJ*, 706, 1364
- Fragkoudi F., Di Matteo P., Haywood M., Gómez A., Combes F., Katz D., Semelin B., 2017, submitted to *A&A*, arXiv:1704.00734
- Fuhrmann K., 1998, *A&A*, 338, 161
- Gaia Collaboration, Prusti T. et al., 2016, *A&A*, 595, A1
- Genzel R. et al., 2011, *ApJ*, 733, 101
- Genzel R. et al., 2015, *ApJ*, 800, 20
- Gilmore G., Reid N., 1983, *MNRAS*, 202, 1025
- Grand R. J. J., Springel V., Gómez F. A., Marinacci F., Pakmor R., Campbell D. J. R., Jenkins A., 2016, *MNRAS*, 459, 199
- Hayden M. R. et al., 2015, *ApJ*, 808, 132
- Hernquist L., 1990, *ApJ*, 356, 359
- Jiang I.-G., Binney J., 1999, *MNRAS*, 303, L7
- Jones T. A., Ellis R. S., Schenker M. A., Stark D. P., 2013, *ApJ*, 779, 52
- Jurić M. et al., 2008, *ApJ*, 673, 864
- Kassin S. A. et al., 2012, *ApJ*, 758, 106

- Kazantzidis S., Bullock J. S., Zentner A. R., Kravtsov A. V., Moustakas L. A., 2008, *ApJ*, 688, 254
- Licquia T. C., Newman J. A., 2016, *ApJ*, 831, 71
- López-Corredoira M., 2016, *A&A*, 593, 66
- Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Angles-Alcazar D., Faucher-Giguere C.-A., Keres D., Quataert E., 2017, *MNRAS*, 467, 2430
- Martig M., Minchev I., Flynn C., 2014, *MNRAS*, 443, 2452
- Martig M. et al., 2016a, *MNRAS*, 456, 3655
- Martig M., Minchev I., Ness M., Fouesneau M., Rix H.-W., 2016b, *ApJ*, 831, 139
- Masseron T., Gilmore G., 2015, *MNRAS*, 453, 1855
- McKee C. F., Parravano A., Hollenbach D. J., 2015, *ApJ*, 814, 13
- McMillan P. J., 2017, *MNRAS*, 465, 76
- Minchev I., Martig M., Streich D., Scannapieco C., de Jong R. S., Steinmetz M., 2015, *ApJ*, 804, L9
- Newman S. F. et al., 2013, *ApJ*, 767, 104
- Pfenniger D., Friedli D., 1991, *A&A*, 252, 75
- Piffl T. et al., 2014, *MNRAS*, 445, 3133
- Portail M., Gerhard O., Wegg C., Ness M., 2017a, *MNRAS*, 465, 1621
- Portail M., Wegg C., Gerhard O., Ness M., 2017b, submitted to *MNRAS*, arXiv:1704.07821
- Quinn P. J., Hernquist L., Fullagar D. P., 1993, *ApJ*, 403, 74
- Raha N., Sellwood J. A., James R. A., Kahn F. D., 1991, *Nature*, 352, 411
- Roškar R., Debattista V. P., Loebman S. R., 2013, *MNRAS*, 433, 976
- Sato C., 1980, *PASJ*, 32, 41
- Schönrich R., Binney J., 2009, *MNRAS*, 399, 1145
- Schönrich R., Binney J., 2012, *MNRAS*, 419, 1546
- Schönrich R., 2012, *MNRAS*, 427, 274
- Schönrich R., McMillan P. J., 2017, *MNRAS*, 467, 1154
- Sellwood J. A., 2008, *ApJ*, 679, 379
- Sellwood J. A., Debattista V. P., 2009, *MNRAS*, 398, 1279
- Sellwood J. A., 2013, *ApJ*, 769L, 24
- Sheth K. et al., 2008, *ApJ*, 675, 1141
- Solway M., Sellwood J. A., Schönrich R., 2012, *MNRAS*, 422, 1363
- Springel V., 2005, *MNRAS*, 364, 1105
- Stinson G. S. et al., 2013, *MNRAS*, 436, 625
- van der Kruit P. C., Searle L., 1982, *A&A*, 110, 61
- Vera-Ciro C., D’Onghia E., Navarro J., Abadi M., 2014, *ApJ*, 794, 173
- Wegg C., Gerhard O., 2013, *MNRAS*, 435, 1874
- Wegg C., Gerhard O., Portail M., 2015, *MNRAS*, 450, 4050
- Wegg C., Gerhard O., Portail M., 2016, *MNRAS*, 463, 557
- Wisnioski E. et al., 2015, *ApJ*, 799, 209
- Yoachim P., Dalcanton J. J., 2006, *AJ*, 131, 226
- Yurin D., Springel V., 2014, *MNRAS*, 444, 62