

# Secular diffusion in discrete self-gravitating tepid discs

## II. Accounting for swing amplification via the matrix method<sup>★,★★</sup>

J. B. Fouvry<sup>1</sup>, C. Pichon<sup>1,2</sup>, J. Magorrian<sup>3,1</sup>, and P. H. Chavanis<sup>4</sup>

<sup>1</sup> Institut d'Astrophysique de Paris and UPMC, CNRS (UMR 7095), 98bis Boulevard Arago, 75014 Paris, France  
e-mail: fouvry@iap.fr

<sup>2</sup> Institute of Astronomy & KICC, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK

<sup>3</sup> Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

<sup>4</sup> Laboratoire de Physique Théorique (IRSAMC), CNRS and UPS, Univ. de Toulouse, 31062 Toulouse, France

Received 24 July 2015 / Accepted 25 August 2015

### ABSTRACT

The secular evolution of an infinitely thin tepid isolated galactic disc made of a finite number of particles is investigated using the inhomogeneous Balescu-Lenard equation expressed in terms of angle-action variables. The matrix method is implemented numerically in order to model the induced gravitational polarisation. Special care is taken to account for the amplification of potential fluctuations of mutually resonant orbits and the unwinding of the induced swing amplified transients. Quantitative comparisons with  $N$ -body simulations yield consistent scalings with the number of particles and with the self-gravity of the disc: the fewer the particles and the colder the disc, the faster the secular evolution. Secular evolution is driven by resonances, but does not depend on the initial phases of the disc. For a Mestel disc with  $Q \sim 1.5$ , the polarisation cloud around each star boosts its secular effect by a factor of a thousand or more, accordingly promoting the dynamical relevance of self-induced collisional secular evolution. The position and shape of the induced resonant ridge are found to be in very good agreement with the prediction of the Balescu-Lenard equation, which scales with the square of the susceptibility of the disc. In astrophysics, the inhomogeneous Balescu-Lenard equation may describe the secular diffusion of giant molecular clouds in galactic discs, the secular migration and segregation of planetesimals in proto-planetary discs, or even the long-term evolution of population of stars within the Galactic centre. It could be used as a valuable check of the accuracy of  $N$ -body integrators on secular timescales.

**Key words.** galaxies: evolution – galaxies: kinematics and dynamics – galaxies: spiral – diffusion – gravitation

### 1. Introduction

Because most of the stars are born in a stellar disc, galactic astronomy has striven to understand the dynamics and evolution of discs over cosmic time. Stars in cool discs are born on the almost closed orbits of their parent gas clouds, which means that they are confined to a very small volume of phase space. There is therefore a massive opportunity for increasing the entropy of the disc's phase-space distribution by spreading stars orbits to a wider range of eccentricities and inclinations while conserving the system's energy. Much empirical evidence of such heating has now been discovered: the random velocities of coeval cohorts of stars increases with the cohort's age (Wielen 1977; Aumer & Binney 2009); the velocity distribution at the Sun contains streams of stars (Dehnen 1998); and the density of stars in phase space exhibits elongated features (resonance ridges) consistent with a resonant condition (Sellwood 2010; McMillan 2013).

Such long-term evolution of the system's distribution function (DF) implies that the disc's gravitational potential cannot have remained static and axisymmetric. Axisymmetric discs support spiral waves that can move through the disc; leading

waves unwind into trailing waves and are amplified at corotation. Eventually each wave Landau damps by transferring its energy to stars that resonate with it: it heats the disc by increasing the eccentricity of resonating stars. In a disc made of a finite number of stars, shot noise is a natural excitation mechanism for launching such waves. When the disc is cool, fluctuations of the potential induced by these discrete encounters may be strongly amplified (Kalnajs 1972), so a weak stimulus can produce significant heat. Capturing this secular heating by shot noise requires an apparatus able to follow the dynamics of spiral waves from launch as leading waves through amplification and to thermalisation at resonances to determine their secular dynamical impact. In this paper we demonstrate how the Balescu-Lenard formalism provides an appropriate framework for taking such interactions into account.

The kinetic theory of stellar systems was initiated by Jeans (1929) and Chandrasekhar (1942) in the context of hot spherical stellar systems, such as elliptical galaxies and globular clusters for which the gravitational susceptibility can safely be neglected, as opposed to cold self-gravitating galactic discs. More generally, in astrophysics, the secular diffusion of giant molecular clouds in galactic discs, the secular migration and segregation of planetesimals in proto-planetary or debris discs, or even the long-term evolution of population of stars within the Galactic centre are all processes for which it is relevant to quantify the dynamical effect of gravitationally amplified potential fluctuations induced by the finite number of elements involved. More than 55 years ago, Balescu (1960) and Lenard (1960) developed

\* Appendices are available in electronic form at

<http://www.aanda.org>

\*\* A copy of the linear matrix response code is only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via

<http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/vol/page>

a rigorous kinetic theory taking collective effects into account and obtained the corresponding kinetic equation for plasmas, the Balescu-Lenard equation. More recently Heyvaerts (2010) and Chavanis (2012) have transposed the corresponding non-linear kinetic equation to the angle-action variables that are the appropriate variables to describe spatially inhomogeneous multi-periodic systems. The corresponding inhomogeneous Balescu-Lenard equation accounts for self-driven orbital secular diffusion of a self-gravitating system induced by the intrinsic shot noise due to its discreteness. The formal transposition from position-velocity to angle-action implies that the secular interactions need not be local in space: they only need to correspond to gravitationally amplified long-range correlations and resonances, which are indeed the driving mechanism for the secular evolution of isolated astrophysical discs via angular momentum redistribution (Lynden-Bell & Kalnajs 1972).

The Balescu-Lenard equation is valid at the order  $1/N$  in an expansion of the dynamics in terms of this small parameter, where  $N \gg 1$  is the number of stars. Therefore, it takes finite- $N$  effects into account and describes the evolution of the system on a timescale of about  $Nt_D$ , where  $t_D$  is the dynamical time. For self-gravitating systems, the collective effects are responsible for an anti-shielding which tends to increase the effective mass of the stars, hence reducing the relaxation time. When the system is cold, each particle is dressed by the very strong gravitational polarisation it induces. Therefore secular effects can occur on much shorter timescales than one would naively expect, making the system behave as if it were made up of only  $N_{\text{eff}} \sim N/10^{\text{few}}$  particles. The purpose of this paper is to quantify this effect for stable but strongly susceptible galactic discs.

The Balescu-Lenard formalism has seldom been applied in its prime context, but only in various limits where it reduces to simpler kinetic equations (Landau 1936; Vlasov 1938; Chandrasekhar 1942; Rosenbluth et al. 1957). Weinberg (1993) presents an interesting first implementation, though in a somewhat over-simplified cartesian geometry. This formalism is quite unique in accounting for the non-linear evolution of discs and galaxies on secular timescales. While potentially probing similar processes,  $N$ -body simulations should be scrutinised carefully in this regime, as shadowing (Quinlan & Tremaine 1992) may, over many orbital times, impact resonant interactions. Nevertheless,  $N$ -body simulations have been shown to more or less reproduce growth rates of discs on dynamical timescales (see e.g. Sellwood & Evans 2001, and references therein, together with Appendix C); qualifying them quantitatively on secular timescales is now within reach of the Balescu-Lenard formalism.

A companion paper, (Fouvry et al. 2015a, hereafter Paper I) presented a simple and tractable quadrature for the Balescu-Lenard drift and diffusion coefficients under the assumption that the transient response of the disc can be described by tightly wound spirals via the WKB and epicyclic approximations. These simple expressions have provided insight into the physical processes at work during the secular diffusion of self-gravitating discrete discs. When applied to the secular evolution of an isolated stationary self-gravitating Mestel disc, it identified the importance of the corotation resonance in the inner regions of the disc leading to a regime with both radial migration and heating, in qualitative agreement with numerical simulations.

Yet, the tightly wound approximation is quantitatively questionable when transient spirals unwind. Indeed Paper I found a discrepancy between the predicted secular evolution timescale and the measured one, which might be driven by the incompleteness of the WKB basis, since such a basis can only correctly

represent tightly wound spirals. It considered only resonances between orbits that are close to one another in position space, and did not allow remote orbits to resonate, nor wave packets to propagate between such non-local resonances. Yet, the seminal works from Goldreich & Lynden-Bell (1965), Julian & Toomre (1966), Toomre (1981) show that any leading spiral wave undergoes significant amplification during its unwinding to a trailing wave. Because it involves the unwinding of spirals this mechanism is not captured by the WKB formalism of Paper I.

In this paper, we avoid such approximations by using the matrix method (Kalnajs 1976) to estimate the gravitational amplification of the secular response and the role of non-local resonances. From this we compute numerically the drift and diffusion coefficients that appear in the Balescu-Lenard equation. We then compare those predictions to crafted sets of numerical experiments, allowing us to estimate ensemble averaged secular responses of a sizeable number of simulations as a function of the total number of particles  $N$ . Such ensemble averaging allows us to make robust predictions for the  $N$ -scaling of the secular response and its dependence on halo to disc mass fraction, hence probing the secular importance of gravitational polarisation.

The paper is organised as follows. Section 2 briefly presents the content of the inhomogeneous Balescu-Lenard equation. Section 3 presents our implementation of the matrix method to compute the diffusion equation for an isolated self-gravitating tapered Mestel disc. Section 4 computes numerically the exact drift and diffusion coefficient in action space for such a truncated Mestel disc, and compares the divergence of the corresponding flux density to the initial measured rate of change of the distribution function. Section 5 presents our  $N$ -body simulations and compares scaling of the flux with the number of particles and the fraction of mass in the disc. Finally, Sect. 6 wraps up. Appendix A presents the relevant bi-orthogonal basis function. Appendix C validates the response matrix method and the  $N$ -body integrator while matching growth rates and pattern speeds of unstable Mestel discs. Appendix D investigates the roles of self-gravity and basis completeness. Appendix E describes the sampling strategy for the initial distribution. Appendix G presents briefly the available online codes.

## 2. The inhomogeneous Balescu-Lenard equation

We intend to describe the long-term evolution of a system made of  $N$  particles. We assume that the gravitational background  $\psi_0$  of the system is stationary and integrable, and associated with the Hamiltonian  $H_0$ . As a consequence, one can always remap the physical space-coordinates  $(\mathbf{x}, \mathbf{v})$  to the angle-action coordinates  $(\boldsymbol{\theta}, \mathbf{J})$  (Goldstein 1950; Born 1960; Binney & Tremaine 2008). We define the intrinsic frequencies of motions along the action torus as

$$\boldsymbol{\Omega}(\mathbf{J}) = \dot{\boldsymbol{\theta}} = \frac{\partial H_0}{\partial \mathbf{J}}. \quad (1)$$

Within these new coordinates, one has that along the unperturbed trajectories the angles  $\boldsymbol{\theta}$  are  $2\pi$ -periodic, evolving with the frequencies  $\boldsymbol{\Omega}$ , whereas the actions  $\mathbf{J}$  are conserved. We assume that the system is always in a virialised state, so that its distribution function (DF) can be written as a quasi-stationary DF of the form  $F = F(\mathbf{J}, t)$ , satisfying the normalisation constraint  $\int d\mathbf{x} d\mathbf{v} F = M_{\text{tot}}$ , where  $M_{\text{tot}}$  is the total mass of the system. On secular timescales, this isolated DF evolves under the effect of stellar “encounters” (finite- $N$  effects). Such

a collisional long-term evolution is described by the inhomogeneous Balescu-Lenard equation (Heyvaerts 2010; Chavanis 2012) which reads

$$\frac{\partial F}{\partial t} = \pi(2\pi)^d \mu \frac{\partial}{\partial J_1} \cdot \left[ \sum_{m_1, m_2} m_1 \int dJ_2 \frac{\delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2)}{|\mathcal{D}_{m_1, m_2}(J_1, J_2, m_1 \cdot \Omega_1)|^2} \times \left( m_1 \cdot \frac{\partial}{\partial J_1} - m_2 \cdot \frac{\partial}{\partial J_2} \right) F(J_1, t) F(J_2, t) \right], \quad (2)$$

where  $1/\mathcal{D}_{m_1, m_2}(J_1, J_2, \omega)$  are the dressed susceptibility coefficients,  $d$  is the dimension of the physical space,  $\mu = M_{\text{tot}}/N$  is the mass of the individual particles, and where we used the shortened notation  $\Omega_i = \Omega(J_i, t)$ . Since the r.h.s. is the divergence of a flux, this diffusion equation conserves the number of stars. One should also note the resonance condition encapsulated in the Dirac delta  $\delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2)$ , with the integration over the dummy variable  $J_2$  scanning for points where the resonance condition is satisfied. Note importantly right away that Eq. (2) scales like  $1/(N\mathcal{D}^2)$  (since  $\mu \propto 1/N$ ), so that increasing  $N$  or increasing the heat content of the disc have the same effect. For a more detailed discussion on the content of the Balescu-Lenard equation, see Paper I.

In order to solve the non-local Poisson equation, we follow Kalnajs' matrix method (Kalnajs 1976), so that we introduce a complete biorthonormal basis of potentials and densities  $\psi^{(p)}(\mathbf{x})$  and  $\rho^{(p)}(\mathbf{x})$  such that

$$\Delta \psi^{(p)} = 4\pi G \rho^{(p)}, \quad \int d\mathbf{x} [\psi^{(p)}(\mathbf{x})]^* \rho^{(q)}(\mathbf{x}) = -\delta_p^q. \quad (3)$$

The dressed susceptibility coefficients appearing in Eq. (2) are then given by

$$\frac{1}{\mathcal{D}_{m_1, m_2}(J_1, J_2, \omega)} = \sum_{p, q} \psi_m^{(p)}(J_1) [\mathbf{I} - \widehat{\mathbf{M}}(\omega)]_{pq}^{-1} [\psi_m^{(q)}(J_2)]^*, \quad (4)$$

where  $\mathbf{I}$  is the identity matrix and  $\widehat{\mathbf{M}}$  is the response matrix defined as

$$\widehat{\mathbf{M}}_{pq}(\omega) = (2\pi)^d \sum_m \int dJ \frac{m \cdot \partial F / \partial J}{\omega - m \cdot \Omega} [\psi_m^{(p)}(J)]^* \psi_m^{(q)}(J). \quad (5)$$

In the previous expression, we introduced as  $\psi_m^{(p)}(J)$  the Fourier transform in angles of the basis elements  $\psi^{(p)}(\mathbf{x})$  using the convention that the Fourier transform of a function  $X(\theta, \mathbf{J})$  is given by

$$\begin{cases} X(\theta, \mathbf{J}) = \sum_{m \in \mathbb{Z}^d} X_m(\mathbf{J}) e^{im \cdot \theta}, \\ X_m(\mathbf{J}) = \frac{1}{(2\pi)^d} \int d\theta X(\theta, \mathbf{J}) e^{-im \cdot \theta}. \end{cases} \quad (6)$$

In order to ease the understanding of the Balescu-Lenard Eq. (2), one may rewrite it as an anisotropic Fokker-Planck equation by introducing the relevant drift and diffusion coefficients. Indeed, Eq. (2) can be put into the form

$$\frac{\partial F}{\partial t} = \sum_{m_1} \frac{\partial}{\partial J_1} \cdot \left[ m_1 \left( A_{m_1}(J_1) F(J_1) + D_{m_1}(J_1) m_1 \cdot \frac{\partial F}{\partial J_1} \right) \right], \quad (7)$$

where  $A_{m_1}(J_1)$  and  $D_{m_1}(J_1)$  are respectively the drift and diffusion coefficients associated with a given resonance  $m$ . We note that they both depend secularly on the DF  $F$ , but we omit this dependence to simplify our notations. The drift coefficients are given by

$$A_{m_1}(\mathbf{J}) = -\pi(2\pi)^d \mu \sum_{m_2} \int dJ_2 \frac{\delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2)}{|\mathcal{D}_{m_1, m_2}(J_1, J_2, m_1 \cdot \Omega_1)|^2} m_2 \cdot \frac{\partial F}{\partial J_2}, \quad (8)$$

and the diffusion coefficients are given by

$$D_{m_1}(J_1) = \pi(2\pi)^d \mu \sum_{m_2} \int dJ_2 \frac{\delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2)}{|\mathcal{D}_{m_1, m_2}(J_1, J_2, m_1 \cdot \Omega_1)|^2} F(J_2). \quad (9)$$

One can also introduce the total flux of diffusion  $\mathcal{F}_{\text{tot}}$  as

$$\mathcal{F}_{\text{tot}} = \sum_m m \left( A_m(\mathbf{J}) F(\mathbf{J}) + D_m(\mathbf{J}) m \cdot \frac{\partial F}{\partial \mathbf{J}} \right), \quad (10)$$

so that the Balescu-Lenard equation from Eqs. (2) and (7) takes the explicitly conservative form

$$\frac{\partial F}{\partial t} = \text{div}(\mathcal{F}_{\text{tot}}). \quad (11)$$

### 3. The Matrix diffusion equation

When computing the Balescu-Lenard diffusion and drift coefficients, three main difficulties have to be addressed. First, one must build the mapping  $(\mathbf{x}, \mathbf{v}) \mapsto (\theta, \mathbf{J})$ , because the drift and diffusion coefficients are associated with a diffusion in action space. The second difficulty follows from the non-locality of Poisson's equation and the estimation of the response matrix  $\widehat{\mathbf{M}}$  from Eq. (5). Indeed, as noted in Eq. (3), the matrix relies on potential basis elements  $\psi^{(p)}$  which must be integrated over the whole action space with functions possessing a pole  $1/(\omega - m \cdot \Omega)$ . This cumbersome and difficult evaluation has to be performed numerically, along with the matrix inversion needed to estimate the susceptibility coefficients from Eq. (4). Finally, the third difficulty arises from the resonance condition  $\delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2)$ , which involves determining how orbits may resonate one with another. In Paper I we relied on the epicyclic approximation to build the angle-action mapping and on a WKB basis to treat gravity locally in order to solve these issues while obtaining tractable though approximate expressions.

For a 2D axisymmetric potential, one can define explicitly the actions of the system. Following Lynden-Bell & Kalnajs (1972), Tremaine & Weinberg (1984), the two natural actions of the system are given by a quadrature and an identity

$$\begin{cases} J_1 = J_r = \frac{1}{\pi} \int_{r_p}^{r_a} dr \sqrt{2(E - \psi_0(r)) - L^2/r^2}, \\ J_2 = J_\phi = L, \end{cases} \quad (12)$$

where  $r_p$  and  $r_a$  are respectively the pericentre and the apocentre of the trajectory, while  $E$  and  $L$  are the energy and angular momentum of the star. The first action  $J_1$  encodes the amount of radial energy of the star, so that  $J_r = 0$  corresponds to circular orbits. The second action  $J_2$  is the angular momentum  $L$  of the star. One can then define the two intrinsic frequencies of motion  $\Omega_1 = \kappa$  associated with the radial oscillations and  $\Omega_2 = \Omega_\phi$  associated with the azimuthal oscillations. Indeed, one has

$$\frac{2\pi}{\Omega_1} = 2 \int_{r_p}^{r_a} \frac{dr}{\sqrt{2(E - \psi_0(r)) - J_2^2/r^2}}, \quad (13)$$

while the azimuthal frequency  $\Omega_2$  can then be determined via the relation

$$\frac{\Omega_2}{\Omega_1} = \frac{J_2}{\pi} \int_{r_p}^{r_a} \frac{dr}{r^2 \sqrt{2(E - \psi_0(r)) - J_2^2/r^2}}. \quad (14)$$



At this stage, one should note that various coordinates can be used to represent the 2D action space. Indeed, once the background potential  $\psi_0$  is known, one has the bijections  $(r_p, r_a) \leftrightarrow (E, L) \leftrightarrow (J_r, J_\phi)$ . As a consequence, any orbit can equivalently be represented by the set of the pericentre and apocentre  $(r_p, r_a)$  or by its actions  $(J_1, J_2)$ . However, determining the actions associated with one set  $(r_p, r_a)$  only requires the computation of a 1D integral as in Eq. (12), whereas determining the pericentre and apocentre associated with a set of actions  $(J_1, J_2)$  requires the inversion of the same non-trivial relation. Because the peri/apocentres are the two roots of the equation  $2(E - \psi_0(r)) - L^2/r^2 = 0$ , one also immediately obtains that for a given value of  $r_p$  and  $r_a$ , the energy  $E$  and the angular momentum  $L$  of the orbit are given by

$$E = \frac{r_a^2 \psi_a - r_p^2 \psi_p}{r_a^2 - r_p^2}; \quad L = \sqrt{\frac{2(\psi_a - \psi_p)}{r_p^{-2} - r_a^{-2}}}, \quad (15)$$

where we used the shortening notations  $\psi_{p/a} = \psi_0(r_{p/a})$ . Therefore, in the upcoming calculations, we use  $(r_p, r_a)$  as the representative variables of the 2D action space.

### 3.1. The basis elements

The expressions (4) and (5) of the susceptibility coefficients and the response matrix require the introduction of 2D potential-density basis elements. The 2D potential basis elements  $\psi^{(p)}$  that we consider depend on two indices spanning the two degrees of freedom so that one has

$$\psi^{(p)}(R, \phi) = \psi_n^\ell(R, \phi) = e^{i\ell\phi} \mathcal{U}_n^\ell(R, \phi), \quad (16)$$

where  $\mathcal{U}_n^\ell$  is a real radial function and  $(R, \phi)$  are the usual polar coordinates. The associated surface densities elements are of the form

$$\Sigma^{(p)}(R, \phi) = \Sigma_n^\ell(R, \phi) = e^{i\ell\phi} \mathcal{D}_n^\ell(R, \phi), \quad (17)$$

where  $\mathcal{D}_n^\ell$  is a real radial function. The basis elements therefore depend on two indices  $\ell \geq 0$  and  $n \geq 0$ . In all the numerical calculations, we used the radial functions from Kalnajs (1976), which are recalled in Appendix A.

The next step is then to determine the Fourier transform with respect to the angles  $\theta$  of the basis elements. Indeed, to compute the response matrix  $\widehat{\mathbf{M}}$  from Eq. (5), one has to compute  $\psi_m^{(p)}(\mathbf{J})$ , where the resonance vector is given by  $\mathbf{m} = (m_1, m_2)$ . Following Eq. (6), it is given by

$$\psi_m^{(p)}(\mathbf{J}) = \frac{1}{(2\pi)^2} \int d\theta_1 d\theta_2 \psi^{(p)}(R, \phi) e^{-im_1\theta_1} e^{-im_2\theta_2}. \quad (18)$$

From Lynden-Bell & Kalnajs (1972), the angles  $\theta_1$  and  $\theta_2$  associated with the actions from Eq. (12) are given by

$$\begin{cases} \theta_1 = \Omega_1 \int_{C_1} dr \frac{1}{\sqrt{2(E - \psi_0(r)) - J_2^2/r^2}}, \\ \theta_2 = \phi + \int_{C_1} dr \frac{\Omega_2 - J_2/r^2}{\sqrt{2(E - \psi_0(r)) - J_2^2/r^2}}, \end{cases} \quad (19)$$

where  $C_1$  is a contour starting from the pericentre  $r_p$  and going up to the current position  $r = r(\theta_1)$  along the radial oscillation. Following the notations from Tremaine & Weinberg (1984), one can straightforwardly show that Eq. (18) takes the form

$$\psi_m^{(p)}(\mathbf{J}) = \delta_{m_2}^{\ell p} \mathcal{W}_{\ell p m_2 n^p}^{m_1}(\mathbf{J}), \quad (20)$$

where  $\mathcal{W}_{\ell p m_2 n^p}^{m_1}(\mathbf{J})$  is given by

$$\mathcal{W}_{\ell p m_2 n^p}^{m_1}(\mathbf{J}) = \frac{1}{\pi} \int_{r_p}^{r_a} dr \frac{d\theta_1}{dr} \mathcal{U}_{n^p}^{\ell p}(r) \cos[m_1\theta_1[r] + m_2(\theta_2 - \phi)[r]]. \quad (21)$$

In Eq. (21), the boundaries of the integral are given by the pericentre  $r_p$  and apocentre  $r_a$  associated with the action  $\mathbf{J}$ . Such an expression underlines the reason why  $(r_p, r_a)$  appear naturally as good coordinates to describe the 2D action space. One can note that Eq. (21) involves an integral over  $r$  thanks to the change of variables  $\theta_1 \rightarrow r$ , which satisfies

$$\frac{d\theta_1}{dr} = \frac{\Omega_1}{\sqrt{2(E - \psi_0(r)) - J_2^2/r^2}}. \quad (22)$$

In Eq. (21),  $\theta_1[r]$  and  $(\theta_2 - \phi)[r]$  only depend on  $r$  via the mappings from Eq. (19). Provided that  $\mathcal{U}_{n^p}^{\ell p}$  is a real function, the coefficients  $\mathcal{W}_{\ell p m_2 n^p}^{m_1}$  are always real. Because these coefficients involve two intricate integrals, they are numerically expensive to compute. However by parity, they obey  $\mathcal{W}_{\ell p (-m_2) n^p}^{(-m_1)} = \mathcal{W}_{\ell p m_2 n^p}^{m_1}$ , which offers a significant reduction of the number of coefficients to compute.

### 3.2. Computation of the response matrix

We now have all the elements required to compute the response matrix from Eq. (5). In its definition, one should note the presence of an integral over the dummy variable  $\mathbf{J}$ , which, as discussed previously, is performed in the 2D  $(r_p, r_a)$ -space. The first step is to go from  $\mathbf{J} = (J_1, J_2)$  to  $(E, L)$ . The Jacobian of this transformation is given by

$$\frac{\partial(E, L)}{\partial(J_1, J_2)} = \begin{vmatrix} \frac{\partial E}{\partial J_1} & \frac{\partial E}{\partial J_2} \\ \frac{\partial L}{\partial J_1} & \frac{\partial L}{\partial J_2} \end{vmatrix} = \begin{vmatrix} \Omega_1 & \Omega_2 \\ 0 & 1 \end{vmatrix} = \Omega_1, \quad (23)$$

so that one immediately has  $dJ_1 dJ_2 = dE dL / \Omega_1$ . Given the expression (20) of the 2D Fourier transformed basis elements, the response matrix may be written as

$$\widehat{\mathbf{M}}_{pq}(\omega) = (2\pi)^2 \delta_{\ell^q}^{\ell^p} \sum_{m_1} \int dE dL \frac{1}{\Omega_1} \frac{(m_1, \ell^p) \cdot \partial F_0 / \partial \mathbf{J}}{\omega - (m_1, \ell^p) \cdot \boldsymbol{\Omega}} \times \mathcal{W}_{\ell^p \ell^p n^p}^{m_1}(\mathbf{J}) \mathcal{W}_{\ell^p \ell^p n^q}^{m_1}(\mathbf{J}), \quad (24)$$

where the sum on  $m_2$  has been executed using the Kronecker delta from Eq. (20). Moreover, we dropped the conjugate over  $\mathcal{W}_{\ell^p \ell^p n^p}^{m_1}$  since the latter is always real. We may now perform the change of variables  $(E, L) \rightarrow (r_p, r_a)$ , so as to rewrite Eq. (24) as

$$\widehat{\mathbf{M}}_{pq}(\omega) = \delta_{\ell^q}^{\ell^p} \sum_{m_1} \int dr_p dr_a \frac{g_{m_1}^{\ell^p n^p n^q}(r_p, r_a)}{h_{m_1, \ell^p}^\omega(r_p, r_a)}, \quad (25)$$

where the functions  $g_{m_1}^{\ell^p n^p n^q}(r_p, r_a)$  and  $h_{m_1, \ell^p}^\omega(r_p, r_a)$  are respectively given by

$$g_{m_1}^{\ell^p n^p n^q}(r_p, r_a) = (2\pi)^2 \left| \frac{\partial(E, L)}{\partial(r_p, r_a)} \right| \frac{1}{\Omega_1} \left[ (m_1, \ell^p) \cdot \frac{\partial F_0}{\partial \mathbf{J}} \right] \times \mathcal{W}_{\ell^p \ell^p n^p}^{m_1}(\mathbf{J}) \mathcal{W}_{\ell^p \ell^p n^q}^{m_1}(\mathbf{J}), \quad (26)$$

and

$$h_{m_1, \ell^p}^\omega(r_p, r_a) = \omega - (m_1, \ell^p) \cdot \boldsymbol{\Omega}. \quad (27)$$

It is important to note here that the response matrix is diagonal with respect to the  $\ell^p$  and  $\ell^q$  indices so that each  $\ell$  may be treated independently. The definition of the function  $g$  from Eq. (26) involves the Jacobian  $\partial(E, L)/\partial(r_p, r_a)$  of the transformation  $(E, L) \rightarrow (r_p, r_a)$  which can be immediately computed from the expressions (15) of  $E = E(r_p, r_a)$  and  $L = L(r_p, r_a)$ . Moreover, in some situations, the DF  $F = F(\mathbf{J})$ , may also rather be defined as  $F = F(E, L)$ . It is straightforward to show that one has

$$\mathbf{m} \cdot \frac{\partial F}{\partial \mathbf{J}} = m_1 \Omega_1 \left( \frac{\partial F}{\partial E} \right)_L + m_2 \left[ \Omega_2 \left( \frac{\partial F}{\partial E} \right)_L + \left( \frac{\partial F}{\partial L} \right)_E \right]. \quad (28)$$

### 3.3. Sub-region integration

The next step in the calculation is to perform the remaining integration over  $(r_p, r_a)$  from Eq. (25). Because of the presence of the resonant pole  $1/h_{m_1, \ell^p}^\omega$ , such a numerical integration has to be performed carefully. We cut the integration domain  $(r_p, r_a)$  into various subregions indexed by  $i$ . The  $i$ th-region is centred around the position  $(r_p^i, r_a^i)$  and corresponds to the square domain such that  $r_p \in [r_p^i - \Delta r/2; r_p^i + \Delta r/2]$  and  $r_a \in [r_a^i - \Delta r/2; r_a^i + \Delta r/2]$ , where  $\Delta r$  corresponds to the size of the subregions. The smaller  $\Delta r$ , the better the approximated estimations of the response matrix. Within the  $i$ th-region, one can write first-order Taylor expansions of the functions  $g$  and  $h$  from Eqs. (26) and (27) around the centre  $(r_p^i, r_a^i)$  of the region such that

$$\begin{cases} g(r_p^i + \Delta r_p, r_a^i + \Delta r_a) \simeq a_g^i + b_g^i \Delta r_p + c_g^i \Delta r_a, \\ h(r_p^i + \Delta r_p, r_a^i + \Delta r_a) \simeq a_h^i + b_h^i \Delta r_p + c_h^i \Delta r_a, \end{cases} \quad (29)$$

where for convenience we shortened the index dependences from Eqs. (26) and (27). The coefficients  $a_g^i, b_g^i$  and  $c_g^i$  (similarly for  $h$ ) are given by

$$a_g^i = g(r_p^i, r_a^i); \quad b_g^i = \left. \frac{\partial g}{\partial r_p} \right|_{(r_p^i, r_a^i)}; \quad c_g^i = \left. \frac{\partial g}{\partial r_a} \right|_{(r_p^i, r_a^i)}. \quad (30)$$

In the numerical implementation, the coefficients involving partial derivatives are estimated by finite differences, so that one has for instance

$$b_g(r_p^i, r_a^i) = \frac{g(r_p^i + \Delta r, r_a^i) - g(r_p^i - \Delta r, r_a^i)}{2\Delta r}, \quad (31)$$

which allows us to minimise the number of evaluations of  $g$  required. The approximated integration on each sub-region can then be performed and takes the form

$$\begin{aligned} \int_i \int dr_p dr_a \frac{g(r_p, r_a)}{h(r_p, r_a)} &\simeq \int_{-\frac{\Delta r}{2}}^{\frac{\Delta r}{2}} \int_{-\frac{\Delta r}{2}}^{\frac{\Delta r}{2}} dx_p dx_a \frac{a_g^i + b_g^i x_p + c_g^i x_a}{a_h^i + b_h^i x_p + c_h^i x_a + i\eta} \\ &= \mathfrak{N}(a_g^i, b_g^i, c_g^i, a_h^i, b_h^i, c_h^i, \eta, \Delta r), \end{aligned} \quad (32)$$

where  $\mathfrak{N}$  is an analytical function which only depends on the coefficients obtained in the limited developments from Eq. (29). In order to have a well-defined integral, we added an imaginary part  $\eta > 0$  to the temporal frequency  $\omega$ , so that  $\omega = \omega_0 + i\eta$ . When looking for unstable modes in a disc, this imaginary part  $\eta$  corresponds to the growth rate of the mode. It is also crucial to note here that one always has  $a_g^i, b_g^i, c_g^i \in \mathbb{R}$ , and similarly  $a_h^i, b_h^i, c_h^i \in \mathbb{R}$ . The effective computation of the function  $\mathfrak{N}$  is presented in Appendix B. Thanks to Eq. (32), the expression (25) becomes

$$\widehat{\mathbf{M}}_{pq}(\omega) = \delta_{\ell^p \ell^q} \sum_{m_1} \sum_i \mathfrak{N}(a_g^i, b_g^i, c_g^i, a_h^i, b_h^i, c_h^i, \eta, \Delta r). \quad (33)$$

In the previous expression, in order to effectively compute numerically the sum on  $m_1$ , we introduce a bound  $m_1^{\max}$ , so that the sum is only reduced to  $|m_1| \leq m_1^{\max}$ . Because of the requirement to truncate the action space in various subregions as in Eq. (33), the computation of the response matrix still remains a daunting task, to ensure appropriate numerical convergence. In Appendix C, we detail the validation of our implementation of the response matrix calculation by recovering known unstable modes of truncated Mestel discs (Zang 1976; Evans & Read 1998b; Sellwood & Evans 2001). Once the response matrix  $\widehat{\mathbf{M}}$  is known, the determination of the dressed susceptibility coefficients  $1/|\mathcal{D}|^2$  from Eq. (4) involves a straightforward summation<sup>1</sup>.

### 3.4. Critical resonant line

The resonance condition encapsulated in the Dirac delta  $\delta_D(\mathbf{m}_1 \cdot \boldsymbol{\Omega}_1 - \mathbf{m}_2 \cdot \boldsymbol{\Omega}_2)$  generates an additional difficulty in the calculation of the Balescu-Lenard drift and diffusion coefficients from Eqs. (8) and (9). Recall the definition of the composition of a Dirac delta, and a function (Hörmander 2003), which in a  $d$ -dimensional setup takes the form

$$\int_{\mathbb{R}^d} d\mathbf{x} f(\mathbf{x}) \delta_D(g(\mathbf{x})) = \int_{g^{-1}(0)} d\sigma(\mathbf{x}) \frac{f(\mathbf{x})}{|\nabla g(\mathbf{x})|}, \quad (34)$$

where  $g^{-1}(0) = \{\mathbf{x} | g(\mathbf{x}) = 0\}$  is the hyper-surface of dimension (generically)  $(d-1)$  defined by the constraint  $g(\mathbf{x}) = 0$ , and  $d\sigma(\mathbf{x})$  is the surface measure on  $g^{-1}(0)$ . We have also naturally defined the euclidean norm of the gradient of  $g$  as  $|\nabla g(\mathbf{x})| = \sqrt{|\partial g/\partial x_1|^2 + \dots + |\partial g/\partial x_d|^2}$ . Here we assumed that the resonance condition associated with the function  $g(\mathbf{J}_2) = \mathbf{m}_1 \cdot \boldsymbol{\Omega}_1 - \mathbf{m}_2 \cdot \boldsymbol{\Omega}_2$  is non-degenerate, so that  $\forall \mathbf{x} \in g^{-1}(0), |\nabla g(\mathbf{x})| > 0$ , which also ensures that the dimension of the set  $g^{-1}(0)$  is  $(d-1)$ . One should note that this degeneracy condition is not satisfied by the Keplerian or harmonic potentials. Because we are considering an infinitely thin disc, the dimension of the physical space is given by  $d = 2$ , so that the set  $g^{-1}(0)$  takes the form of a curve  $\gamma$ , that we call the critical resonant curve. Generically, it takes the form of an application of the type

$$\gamma: u \mapsto \gamma(u) = (\gamma_1(u), \gamma_2(u)). \quad (35)$$

One can then immediately compute the r.h.s. of Eq. (34) via

$$\int_{\gamma} d\sigma(\mathbf{x}) \frac{f(\mathbf{x})}{|\nabla g(\mathbf{x})|} = \int du \frac{f(\gamma(u))}{|\nabla g(\gamma(u))|} |\gamma'(u)|, \quad (36)$$

where we have naturally defined  $|\gamma'(u)| = \sqrt{|\dot{\gamma}_1/du|^2 + |\dot{\gamma}_2/du|^2}$ .

As noted in Eq. (15), using the pericentres and apocentres  $(r_p, r_a)$ , given the Jacobian from Eq. (23) and proceeding in the same way as in Eq. (25) for the response matrix, one may rewrite the drift and diffusion coefficients from Eqs. (8) and (9) as

$$A_{m_1}(\mathbf{J}_1) = \sum_{m_2} \int dr_p dr_a \delta_D(\mathbf{m}_1 \cdot \boldsymbol{\Omega}_1 - \mathbf{m}_2 \cdot \boldsymbol{\Omega}_2) G_{m_1, m_2}^A(r_p, r_a), \quad (37)$$

and

$$D_{m_1}(\mathbf{J}_1) = \sum_{m_2} \int dr_p dr_a \delta_D(\mathbf{m}_1 \cdot \boldsymbol{\Omega}_1 - \mathbf{m}_2 \cdot \boldsymbol{\Omega}_2) G_{m_1, m_2}^D(r_p, r_a), \quad (38)$$

<sup>1</sup> One could if needed regularise the inversion of  $\mathbf{I} - \widehat{\mathbf{M}}$  to avoid Gibbs ringing since our basis is significantly truncated; this has proven not necessary here.

where the functions  $G_{\mathbf{m}_1, \mathbf{m}_2}^A$  and  $G_{\mathbf{m}_1, \mathbf{m}_2}^D$  are respectively defined as

$$G_{\mathbf{m}_1, \mathbf{m}_2}^A(r_p, r_a) = -\frac{1}{\Omega_1} \left| \frac{\partial(E, L)}{\partial(r_p, r_a)} \right| \frac{4\pi^3 \mu \mathbf{m}_2 \cdot \partial F / \partial \mathbf{J}_2}{|\mathcal{D}_{\mathbf{m}_1, \mathbf{m}_2}(\mathbf{J}_1, \mathbf{J}_2, \mathbf{m}_1 \cdot \boldsymbol{\Omega}_1)|^2}, \quad (39)$$

and

$$G_{\mathbf{m}_1, \mathbf{m}_2}^D(r_p, r_a) = \frac{1}{\Omega_1} \left| \frac{\partial(E, L)}{\partial(r_p, r_a)} \right| \frac{4\pi^3 \mu F(\mathbf{J}_2)}{|\mathcal{D}_{\mathbf{m}_1, \mathbf{m}_2}(\mathbf{J}_1, \mathbf{J}_2, \mathbf{m}_1 \cdot \boldsymbol{\Omega}_1)|^2}. \quad (40)$$

For a given value of  $\mathbf{J}_1$ ,  $\mathbf{m}_1$  and  $\mathbf{m}_2$ , and introducing  $\omega_1 = \mathbf{m}_1 \cdot \boldsymbol{\Omega}_1$ , we define the critical curve  $\gamma_{\mathbf{m}_2}(\omega_1)$  as

$$\gamma_{\mathbf{m}_2}(\omega_1) = \left\{ (r_p, r_a) \mid \mathbf{m}_2 \cdot \boldsymbol{\Omega}(r_p, r_a) = \omega_1 \right\}. \quad (41)$$

The expressions (37) and (38) of the drift and diffusion coefficients immediately become

$$\begin{cases} A_{\mathbf{m}_1}(\mathbf{J}_1) = \sum_{\mathbf{m}_2} \int_{\gamma_{\mathbf{m}_2}(\omega_1)} d\sigma \frac{G_{\mathbf{m}_1, \mathbf{m}_2}^A}{|\nabla(\mathbf{m}_2 \cdot \boldsymbol{\Omega}_2)|}, \\ D_{\mathbf{m}_1}(\mathbf{J}_1) = \sum_{\mathbf{m}_2} \int_{\gamma_{\mathbf{m}_2}(\omega_1)} d\sigma \frac{G_{\mathbf{m}_1, \mathbf{m}_2}^D}{|\nabla(\mathbf{m}_2 \cdot \boldsymbol{\Omega}_2)|}, \end{cases} \quad (42)$$

where the resonant contribution  $|\nabla(\mathbf{m}_2 \cdot \boldsymbol{\Omega}_2)|$  is defined as

$$|\nabla(\mathbf{m}_2 \cdot \boldsymbol{\Omega}_2)| = \sqrt{\left[ \mathbf{m}_2 \cdot \frac{\partial \boldsymbol{\Omega}_2}{\partial r_p} \right]^2 + \left[ \mathbf{m}_2 \cdot \frac{\partial \boldsymbol{\Omega}_2}{\partial r_a} \right]^2}. \quad (43)$$

The derivatives of the intrinsic frequencies with respect to  $r_p$  and  $r_a$  appearing in Eq. (43) are computed as in Eq. (31) using finite differences. Once the critical lines of resonances have been determined, the computation of the drift and diffusion coefficients from Eq. (42) is straightforward, so that the full secular diffusion flux  $\mathcal{F}_{\text{tot}}$  from Eq. (10) may be determined.

#### 4. Balescu-Lenard flux divergences

We may now illustrate how the previous computations of the response matrix and the Balescu-Lenard drift and diffusion coefficients can be used to recover some results obtained in well-crafted numerical simulations of galactic discs. Indeed, Sellwood (2012; hereafter S12), studied the long-term evolution of an isolated stable truncated Mestel disc (Mestel 1963). After letting the disc evolve for hundreds of dynamical times, S12 observed a secular diffusion of the disc's DF in action space, through the spontaneous generation of transient spiral waves. The most striking result of this evolution is given in Fig. 7 of S12, which exhibits the late time formation of a resonant ridge in action space along a specific resonant direction. Such diffusion features observed in the late evolution of an isolated, stable, and discrete system are thought to be signatures of a secular evolution induced by finite- $N$  effects, as described by the Balescu-Lenard formalism. Because the system is made of a finite number  $N$  of pointwise particles, it undergoes (long-range) resonant encounters leading to an irreversible secular evolution. In order to investigate such a collisional evolution, Paper I applied the WKB limit of the Balescu-Lenard formalism to S12's simulation. While most of the secular diffusion was qualitatively recovered, there remained a significant timescale discrepancy, since the typical timescale diffusion predicted by this approach was typically  $10^3$  times too slow compared to the observations made in S12. The use of a non-local basis such as Eq. (16) and the

numerical computation of the response matrix from Eq. (24) allows us to incorporate in the present paper these previously ignored contributions from the WKB approach. In the upcoming sections, we therefore present briefly the disc considered by S12 and our determination of the secular diffusion flux predicted by the Balescu-Lenard formalism.

##### 4.1. Initial setup

We consider the same disc as presented in Sellwood (2012). It is an infinitely thin Mestel disc for which the circular speed  $v_\phi$  is a constant  $V_0$  independent of the radius. The stationary background potential  $\psi_M$  and its associated surface density  $\Sigma_M$  are given by

$$\psi_M(R) = V_0^2 \log \left[ \frac{R}{R_{\text{max}}} \right]; \quad \Sigma_M(R) = \frac{V_0^2}{2\pi G R}, \quad (44)$$

where  $R_{\text{max}}$  is a scale parameter of the disc. Following Toomre (1977), Binney & Tremaine (2008), a self-consistent DF for this system is given by

$$F_M(E, J_\phi) = C_M J_\phi^q \exp \left[ -E/\sigma_r^2 \right], \quad (45)$$

where the exponent  $q$  is given by

$$q = \frac{V_0^2}{\sigma_r^2} - 1, \quad (46)$$

with  $\sigma_r$  being the constant radial velocities spread within the disc. In Eq. (45),  $C_M$  is a normalisation constant given by

$$C_M = \frac{V_0^2}{2^{1+q/2} \pi^{3/2} G \sigma_r^{q+2} \Gamma \left[ \frac{1}{2} + \frac{q}{2} \right] R_{\text{max}}^{q+1}}. \quad (47)$$

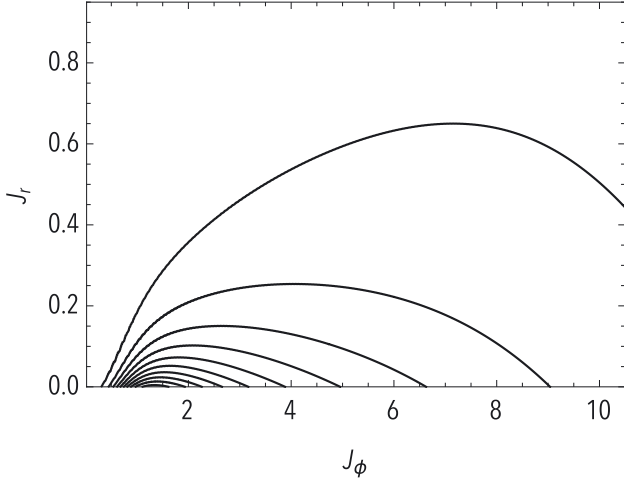
In order to deal with the central singularity of the Mestel disc along with its infinite extent, we introduce two tapering functions

$$\begin{cases} T_{\text{inner}}(J_\phi) = \frac{J_\phi^{\nu_t}}{(R_i V_0)^{\nu_t} + J_\phi^{\nu_t}}, \\ T_{\text{outer}}(J_\phi) = \left[ 1 + \left[ \frac{J_\phi}{R_0 V_0} \right]^{\mu_t} \right]^{-1}, \end{cases} \quad (48)$$

where the indices  $\nu_t$  and  $\mu_t$  control the sharpness of the two tapers, while the radii  $R_i$  and  $R_0$  are two scale parameters. These tapers  $T_{\text{inner}}$  and  $T_{\text{outer}}$  respectively represent the bulge and the outer truncation of the disc. In addition to these taperings, we also suppose that only a fraction  $\xi$  of the stellar disc is self-gravitating, with  $0 \leq \xi \leq 1$ , while the rest of the gravitational potential is provided by the static halo. As a consequence, the active distribution function  $F_{\text{star}}$  is given by

$$F_{\text{star}}(E, J_\phi) = \xi F_M(E, J_\phi) T_{\text{inner}}(J_\phi) T_{\text{outer}}(J_\phi). \quad (49)$$

We place ourselves in the same units system as in S12, so that we have  $V_0 = G = R_i = 1$ . The other numerical factors are given by  $q = 11.4$ ,  $\nu_t = 4$ ,  $\mu_t = 5$ ,  $\xi = 0.5$ ,  $R_0 = 11.5$ , and  $R_{\text{max}} = 20$ . The contours of the tapered DF  $F_{\text{star}}$  are illustrated in Fig. 1. At this stage, it is important to note that S12 restricted the perturbations forces to the harmonic sector  $m_\phi = 2$ , so that we may use the same restriction on the considered azimuthal number  $m_\phi$ . As a consequence, in the double resonance sum on  $\mathbf{m}_1$  and  $\mathbf{m}_2$  present in the Balescu-Lenard flux from Eq. (2), we assume that  $\mathbf{m}_1$  and  $\mathbf{m}_2$  belong to the restricted set  $\{m_r, m_\phi\} \in \{(-1, 2), (0, 2), (1, 2)\}$ ,



**Fig. 1.** Contours of the initial distribution function  $F_{\text{star}}$  from Eq. (49), in action space  $(J_\phi, J_r)$ . The contours are spaced linearly between 95% and 5% of the distribution function maximum.

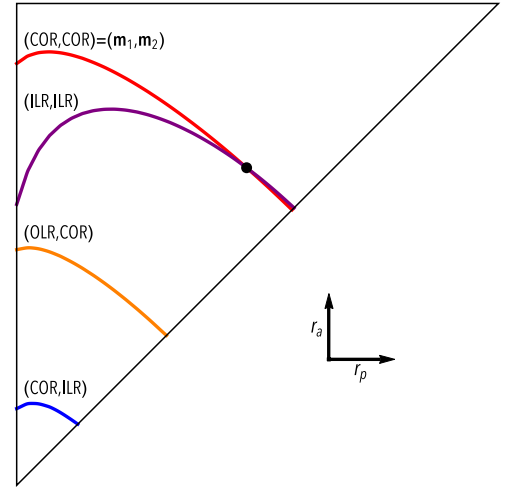
where  $(m_r, m_\phi) = (-1, 2)$  corresponds to the Inner Lindblad resonance (ILR),  $(m_r, m_\phi) = (0, 2)$  to the Corotation resonance (COR), and  $(m_r, m_\phi) = (1, 2)$  to the Outer Lindblad Resonance (OLR). All the upcoming calculations have also been performed while taking into account the contributions from the resonances associated with  $m_r = \pm 2$ , which were checked to be largely subdominant.

#### 4.2. Initial drift and diffusion

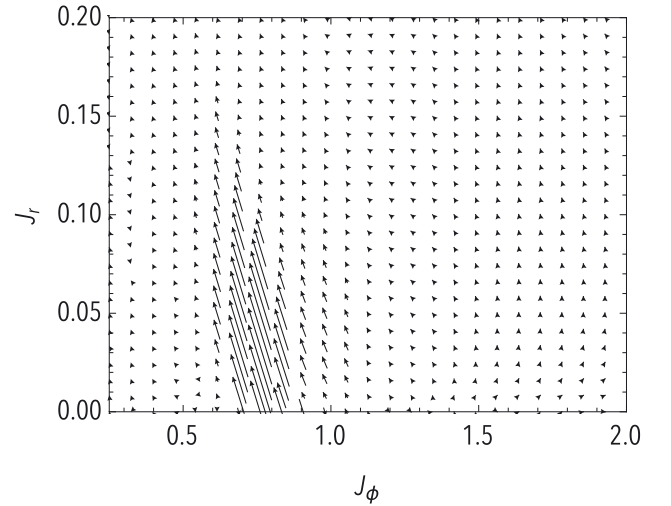
As detailed in Eq. (33), the computation of the response matrix requires to consider a grid in the  $(r_p, r_a)$ -space. We considered a grid such that  $r_p^{\min} = 0.08$ ,  $r_a^{\max} = 4.92$  and  $\Delta r = 0.05$ . The sum on  $m_1$  appearing in Eq. (33) was reduced to  $|m_1| \leq m_1^{\max} = 7$ . The basis considered was Kalnajs 2D basis (Kalnajs 1976) with the parameters  $k_{\text{Ka}} = 7$  and  $R_{\text{Ka}} = 5$ . One should note that despite having a disc which extends up to  $R_{\text{max}} = 20$ , one can still consider a basis truncated at such a small  $R_{\text{Ka}}$ , so as to be able to efficiently capture the diffusion properties of the system in its inner regions, from where the secular diffusion is known to start. The radial basis elements were restricted to  $0 \leq n \leq 8$ . When evaluating the response matrix, as in Eq. (32), one has to add a small imaginary part  $\eta$  to the frequency so as to regularise the resonant denominator. Throughout the calculations presented below, we considered  $\eta = 10^{-4}$  and checked that this choice had no impact on our results.

Since the total potential  $\psi_M$  is known via Eq. (44), the mapping to the angle-action coordinates is completely determined. The two intrinsic frequencies of the system can then be computed on the  $(r_p, r_a)$ -grid via Eqs. (13) and (14). Once these frequencies are known, the critical resonant lines introduced in Eq. (41) can be determined and are illustrated in Fig. 2. It is along these lines that one has to perform the integration present in the definitions of the drift and diffusion coefficients from Eq. (42).

Thanks to this expression, one can then compute the secular diffusion flux  $\mathcal{F}_{\text{tot}}$  defined in Eq. (11). Because the mass of the particles is given by  $\mu = M_{\text{tot}}/N$ , it is natural to consider the quantity  $N\mathcal{F}_{\text{tot}}$  which is independent of  $N$ . The vector field  $-N\mathcal{F}_{\text{tot}} = -(N\mathcal{F}_{J_\phi}, N\mathcal{F}_{J_r})$ , which represents the direction of diffusion of individual particles, is illustrated in Fig. 3. One can already note in Fig. 3 that the diffusion vector field is along a



**Fig. 2.** Illustration of 4 different critical resonant lines in the  $(r_p, r_a)$ -space. As defined in Eq. (41), a critical line is characterised by the resonant vectors  $\mathbf{m}_1, \mathbf{m}_2$  and a location  $\mathbf{J}_1 \leftrightarrow (r_p^1, r_a^1)$  in action space. Each of the 4 plotted critical lines are associated with the same location  $(r_p^1, r_a^1)$ , represented by the black dot. The critical lines correspond to various choices of resonant vectors  $\mathbf{m}_1$  and  $\mathbf{m}_2$  among the three inner, outer and corotation Lindblad resonances, respectively ILR, OLR and COR. One should also note that for  $\mathbf{m}_1 = \mathbf{m}_2$ , the critical lines go through the point  $(r_p^1, r_a^1)$ .

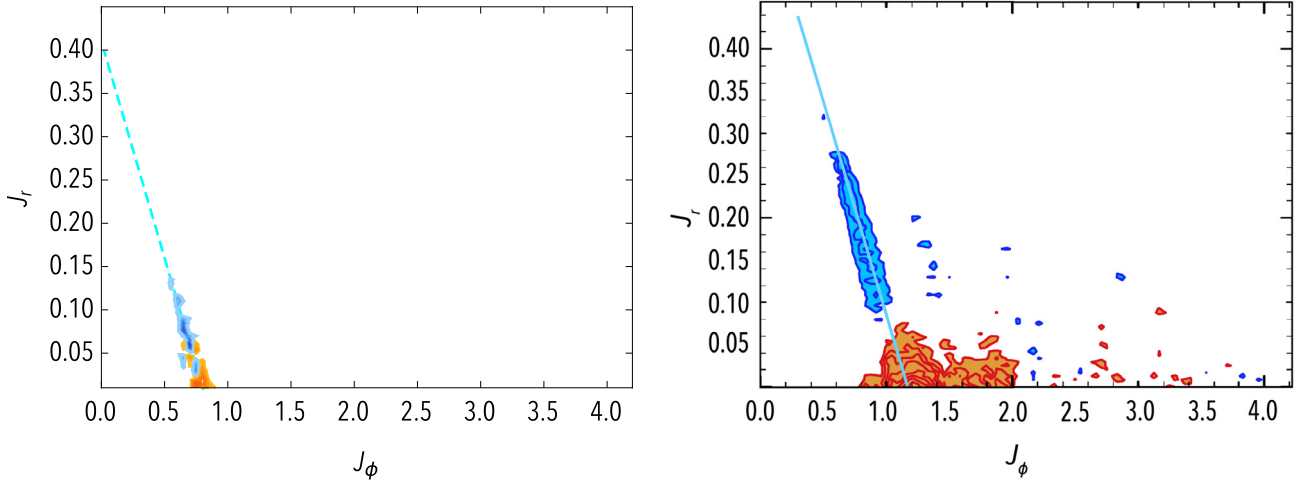


**Fig. 3.** Map of  $-N\mathcal{F}_{\text{tot}}$ , where the flux has been computed with  $\mathbf{m}_1, \mathbf{m}_2 \in \{\mathbf{m}_{\text{ILR}}, \mathbf{m}_{\text{COR}}, \mathbf{m}_{\text{OLR}}\}$ . Following Eq. (11),  $-N\mathcal{F}_{\text{tot}}$  corresponds to the direction of diffusion of individual particles in action-space.

narrow resonant direction. Along this ridge, one typically has  $\mathcal{F}_{J_\phi} \simeq -2\mathcal{F}_{J_r}$ , so the diffusion appears as aligned with the direction of the ILR resonance given by  $\mathbf{m}_{\text{ILR}} = (2, -1)$ .

Once the diffusion flux  $N\mathcal{F}_{\text{tot}}$  has been obtained, one can compute the divergence of this flux, to determine the regions for which the DF is expected to change during the secular diffusion. Figure 4 illustrates the contours of  $N\text{div}(\mathcal{F}_{\text{tot}})$ . In Fig. 4, we obtained that the Balescu-Lenard formalism indubitably predicts the formation of a narrow resonant ridge aligned with the ILR-direction, as was observed in S12's simulation. One also recovers that the stars which will populate the resonant ridge originate from the base of the ridge and diffuse along the direction associated with the ILR resonance. It is most likely that the slight shift in the position of the ridge is due to the fact that the Balescu-Lenard prediction was carried at  $t = 0^+$ , while





**Fig. 4.** *Left panel:* map of  $N\text{div}(\mathcal{F}_{\text{tot}})$ , where the total flux has been computed with  $m_1, m_2 \in \{m_{\text{ILR}}, m_{\text{COR}}, m_{\text{OLR}}\}$ . Red contours, for which  $N\text{div}(\mathcal{F}_{\text{tot}}) < 0$  are associated with regions from which the orbits will be depleted, whereas blue contours, for which  $N\text{div}(\mathcal{F}_{\text{tot}}) > 0$  correspond to regions where the value of the DF will be increased during the secular diffusion. The contours are spaced linearly between the minimum and the maximum of  $N\text{div}(\mathcal{F}_{\text{tot}})$ . The maximum value for the positive blue contours corresponds to  $N\text{div}(\mathcal{F}_{\text{tot}}) \approx 350$ , while the minimum value for the negative red contours is associated with  $N\text{div}(\mathcal{F}_{\text{tot}}) \approx -250$ . *Right panel:* from Sellwood (2012) – Fig. 7, contours of the change in the DF between the time  $t_{\text{S12}} = 1400$  and  $t_{\text{S12}} = 0$ , for a run with  $50M$  particles. Similarly to the left panel, red contours correspond to negative differences, i.e. regions emptied from their orbits, while blue contours correspond to positive differences, i.e. regions where the DF has increased during the diffusion. Both of these contours are aligned with the ILR direction of  $m_{\text{ILR}} = (2, -1)$  in the  $(J_\phi, J_r)$ -plane, corresponding to the cyan line.

S12’s measurements are at  $t = 1400$ , so that we do not expect a perfect match. Other sources of discrepancies might be the use of a softening length in numerical simulations, which modifies the two-body interaction potential, or the difference between an ensemble average (as predicted by the secular formalism) and one specific realisation – our own simulations suggest that there is some variation in the position of the ridge between one run and another. Because we explicitly determined the value of  $N\text{div}(\mathcal{F}_{\text{tot}})$ , we may now study the typical timescale of collisional relaxation predicted by this Balescu-Lenard estimation as detailed in Sect. 4.3. One may also investigate the respective roles of the self-gravitating amplification and the limitation to the tightly wound basis elements as presented in Appendices D.1 and D.2.

#### 4.3. Timescale of diffusion

The most significant disagreement found in Paper I, while applying the WKB approximation of the Balescu-Lenard equation to S12’s simulation was a discrepancy between the time required to observe the resonant ridge in S12’s simulation and the collisional timescale for which the finite- $N$  effects come into play. As already noted in Paper I, since the Balescu-Lenard Eq. (2) only depends on  $N$  through the mass of the individual particles  $\mu = M_{\text{tot}}/N$ , we may rewrite it in the form

$$\frac{\partial F}{\partial t} = \frac{1}{N} C_{\text{BL}}[F], \quad (50)$$

where  $C_{\text{BL}}[F] = N\text{div}(\mathcal{F}_{\text{tot}})$  is the  $N$ -independent Balescu-Lenard collisional operator, i.e. the r.h.s. of Eq. (2) multiplied by  $N = M_{\text{tot}}/\mu$ . As expected, the larger the number of particles, the slower the secular evolution. This also illustrates the fact that the Balescu-Lenard equation comes from a kinetic Taylor expansion in the small parameter  $\varepsilon = 1/N \ll 1$ . We introduce the rescaled time  $\tau = t/N$ , so that Eq. (50) reads

$$\frac{\partial F}{\partial \tau} = C_{\text{BL}}[F], \quad (51)$$

letting us express the Balescu-Lenard equation without any explicit appearance of  $N$ . In Paper I, we estimated the time  $\Delta\tau_{\text{S12}}$  required to observe the ridge as  $\Delta\tau_{\text{S12}} \approx 3 \times 10^{-5}$ . When performing the same measurement thanks to the contours of the diffusion flux  $\text{div}(\mathcal{F}_{\text{tot}})$  computed within the WKB approximation, we obtained  $\Delta\tau_{\text{WKB}} \approx 3 \times 10^{-2}$ , so that Paper I obtained the ratio  $\Delta\tau_{\text{S12}}/\Delta\tau_{\text{WKB}} \approx 10^{-3}$ . This discrepancy was due to the limitation to tightly wound spirals. Because the estimation of the secular diffusion flux  $\mathcal{F}_{\text{tot}}$  presented in Fig. 4 was made using the matrix method (Kalnajs 1976) with a full basis, it captures the additional swing amplification. Indeed, given the map of  $N\text{div}\mathcal{F}_{\text{tot}}$  obtained in Fig. 4, one may estimate the typical time  $\Delta\tau_{\text{BL}}$  required for such a flux to lead to the diffusion features observed in S12. The contours presented in Fig. 7 of S12 are separated by an increment equal to  $0.1 \times F_0^{\text{max}}$ , where  $F_0^{\text{max}} \approx 0.12$  is the maximum of the normalised DF (via Eq. (49)). In order to observe the resonant ridge, the value of the DF should typically change by an amount of the order  $\Delta F_0 \approx 0.1 \times F_0^{\text{max}}$ . From Fig. 4, one can note that the maximum of the divergence of the diffusion flux is given by  $|N\text{div}(\mathcal{F}_{\text{tot}})|_{\text{max}} \approx 350$ . Thanks to Eq. (11), one can immediately write the relation  $\Delta F_0 \approx \Delta\tau_{\text{BL}} |N\text{div}(\mathcal{F}_{\text{tot}})|_{\text{max}}$ , where  $\Delta\tau_{\text{BL}}$  is the time during which the Balescu-Lenard equation has to be evolved in order to develop a ridge. With the previous numerical values, one obtains  $\Delta\tau_{\text{BL}} \approx 3 \times 10^{-5}$ . Comparing the numerically measured time  $\Delta\tau_{\text{S12}}$  and the time  $\Delta\tau_{\text{BL}}$  predicted from the Balescu-Lenard equation, one obtains

$$\frac{\Delta\tau_{\text{S12}}}{\Delta\tau_{\text{BL}}} \sim 1. \quad (52)$$

As expected, the projection of the response over an unbiased basis leads to over a hundredfold increase of the susceptibility of the disc and therefore to a very significant acceleration of secular diffusion. Thanks to this mechanism, we now find a very good agreement between the diffusion timescales observed in numerical simulations and the predictions from the Balescu-Lenard formalism. This quantitative match is rewarding, both from the point of view of the accuracy of the integrator (symplecticity, timestep size, softening...), and from the relevance of



the successive approximations underpinning the Balescu-Lenard formalism (timescale decoupling, truncation of the BBGKY hierarchy, neglect of the close encounter term ...).

In Appendix D, we show that when considering either Fig. D.1, for which the self-gravity of the system has been turned off, or Fig. D.3 for which the loosely wound basis elements were not taken into account, one does not recover a narrow resonant ridge appearing on timescales compatible with S12’s simulations. Therefore, the main source of secular collisional diffusion observed in S12 and recovered in Fig. 4 has to be the strong self-gravitating amplification of loosely wound perturbations, i.e. a sequence of uncorrelated swing amplified spirals sourced by finite- $N$  effects is indeed the main driver of secular diffusion. The WKB formalism from Paper I identified correctly the family of orbits involved, but fell short in predicting how narrow the resonant ridge is and how strongly amplified the response is.

## 5. Comparison to $N$ -body simulations

In Paper I, we relied on simulations presented in Sellwood (2012) to compare the divergence of the diffusion flux to the WKB prediction. In order to probe the expected scalings with the number of particles or with the active fraction of the disc, we now resort to our own  $N$ -body simulations.

### 5.1. $N$ -body integration

The initial sampling of particles is critical when investigating the origin of secular evolution, as one must ensure that the disc is initially in a state of equilibrium. The sampling strategy we implemented is described in some detail in Appendix E.

Once sampled, we evolve the initial conditions using a straightforward particle-mesh  $N$ -body code with a single-timestep leapfrog integrator (e.g. Binney & Tremaine 2008, Sect. 3.4.1). We follow S12 and split the potential in which the particles move into two parts: (i) an axisymmetric contribution  $\psi_M$  from the unperturbed Mestel disc, as in Eq. (44), and (ii) a non-axisymmetric contribution  $\psi_1(R, \phi)$  that develops as perturbations grow in the disc. This splitting avoids difficulties in the treatment of the rigid component of the potential that is not included in the DF, due to the tapering functions and active fraction introduced in Eq. (49). We calculate  $\psi_1$  using cloud-in-cell interpolation (e.g. Binney & Tremaine 2008, Sect. 2.9.3) of the particles’ masses onto an  $N_{\text{mesh}} \times N_{\text{mesh}}$  mesh of square cells spaced  $\Delta x$  apart, then filtering the resulting density field  $\rho(x, y)$  to isolate the disc response (see below), before applying the usual Fourier-space “doubling up” procedure to obtain the potential  $\psi_1$  at the cell vertices. The contribution of  $\psi_1$  to each particle’s acceleration is then obtained using the same cloud-in-cell interpolation scheme.

When computing the density mesh, we added a filtering scheme to include only the  $m_\phi = 2$  disc response, similarly to what was considered in S12. We isolate this  $m_\phi = 2$  mode by calculating

$$\rho_2(r) = \frac{1}{2\pi} \int d\phi \rho(r \cos(\phi), r \sin(\phi)) e^{-2i\phi} \quad (53)$$

immediately after the cloud-in-cell assignment of mass to the mesh at each timestep, then imposing the new mesh mass distribution

$$\rho(x_k, y_k) = \rho_2(r_k) e^{2i\phi_k} \quad (54)$$

with  $(r_k, \phi_k)$  chosen according to  $(x_k, y_k) = (r_k \cos(\phi_k), r_k \sin(\phi_k))$ . To obtain  $\rho_2(r)$  we use brute-force computation of

Eq. (53) on a series of  $N_{\text{ring}}$  radial rings with spacing  $\Delta r \ll \Delta x$ , using the trapezium rule with  $N_\phi = 720$  points in  $\phi$  for the angular integrals. These models are designed to reproduce as closely as possible the essential details of S12’s simulations. There are a couple of deliberate technical differences: S12 uses a polar mesh to obtain  $\psi_1$ , whereas we use a cartesian mesh with a  $m_\phi = 2$  prefiltering of the density field; S12 has a block timestep scheme instead of our simpler single-timestep one.

For the results presented here we used a timestep  $\Delta t = 10^{-3} R_i / V_0$  on a mesh that extends to  $\pm R_{\text{max}} = 20 R_i$  with  $N_{\text{mesh}} = 120$  cells, so that  $\Delta x = R_i / 3$ . The filtering of the potential perturbations to the harmonic sector  $m_\phi = 2$  was performed with  $N_{\text{ring}} = 1000$  radial rings, so that  $\Delta r = 2 R_i / 100$ , and  $N_\phi = 720$  points in the azimuthal direction. Finally, the computation of the potential from the density via Fourier transform was performed with a softening length  $\varepsilon = R_i / 6$ , comparable to the value used in Sellwood (2012), which considered a Plummer softening with  $\varepsilon = R_i / 8$ . The results are not significantly changed when we halve the timestep or the mesh size. In Appendix C, we detail the validation of our  $N$ -body code by recovering known unstable modes of truncated Mestel discs (Zang 1976; Evans & Read 1998b; Sellwood & Evans 2001).

### 5.2. Scaling with $N$

In order to rid our measurements of individual fluctuations, we run multiple simulations for the same number of particles and perform an ensemble average of different evolution realisations for the same number of particles. This procedure allows us to estimate only the mean evolution, which is effectively what is described by the Balescu-Lenard formalism.

In order to study the scaling with  $N$  of these numerical simulations, one has to extract from the simulations a quantity on which to test this scaling and compare it with the predictions from the Balescu-Lenard formalism. The statistical nature of the initial sampling presents an additional difficulty. Indeed, because one only samples  $N$  stars as described in Appendix E, the initial effective DF fluctuates around the smooth background DF from Eq. (49) as a Poisson shot noise. These statistical fluctuations originate from the initial sampling and are not as such specific to the physical process captured by the Balescu-Lenard formalism, so that one should carefully disentangle these two contributions. Hence we introduce the function  $\tilde{h}(t, N)$  defined as

$$\tilde{h}(t, N) = \langle h_i(t, N) \rangle, \quad (55)$$

where the operator  $\langle \cdot \rangle$  corresponds to the ensemble average, approximated here with the arithmetic average over the  $p = 32$  different realisations of simulations for the same number of particles, indexed by  $i$ :  $\langle \cdot \rangle = 1/p \sum_i (\cdot)$ . In Eq. (55) the function  $h_i(t, N)$  is a lag function which reads

$$h_i(t, N) = \int d\mathbf{J} [F_i(t, \mathbf{J}, N) - \langle F(t=0, \mathbf{J}, N) \rangle]^2, \quad (56)$$

where we defined as  $F_i(t, \mathbf{J}, N)$  the normalised DF of the  $i$ th realisation for a number  $N$  of particles. Such a quantity is designed to quantify the “distance” between the initial mean DF  $\langle F(t=0) \rangle$  and the evolved DF  $F_i$ . We are interested in the early time behaviour of the lag function  $h$  from Eq. (55), so that we may perform its Taylor expansion

$$\tilde{h}(t, N) \simeq \tilde{h}_0(N) + \tilde{h}_1(N) t + \tilde{h}_2(N) \frac{t^2}{2}, \quad (57)$$

where it is important to note that the coefficients  $\tilde{h}_0$ ,  $\tilde{h}_1$  and  $\tilde{h}_2$  depend only on  $N$ . Let us now estimate each of these coefficients in turn. Thanks to Eq. (56), one can compute  $\tilde{h}_0(N)$  which reads

$$\tilde{h}_0(N) = \int d\mathbf{J} \langle [F - \langle F_0 \rangle]^2 \rangle, \quad (58)$$

where we used the shortened notations  $\langle F_0 \rangle = \langle F(t=0, \mathbf{J}, N) \rangle$  and  $F = F(t=0, \mathbf{J}, N)$ . We note that this coefficient only depends on the properties of the initial sampling, and not on its dynamics. Because discrete sampling obeys Poisson statistics, one can write

$$\tilde{h}_0(N) = \frac{\alpha_0}{N}, \quad (59)$$

where  $\alpha_0$  is a constant independent of  $N$ . One may then compute  $\tilde{h}_1(N)$ , which takes the form

$$\tilde{h}_1(N) = 2 \int d\mathbf{J} \langle [F - \langle F_0 \rangle] F' \rangle, \quad (60)$$

where we used the shortened notation  $F' = [\partial F / \partial t](t=0)$ . One should note that the terms appearing in Eq. (60) have two different physical contents. Indeed, the term  $[F - \langle F_0 \rangle]$  involves the initial sampling, whereas  $F'$  is driven by the dynamics of the system. If we assume that the sampling and the system's dynamics are uncorrelated, one writes

$$\langle [F - \langle F_0 \rangle] F' \rangle = \langle F - \langle F_0 \rangle \rangle \langle F' \rangle = 0. \quad (61)$$

As a consequence, one immediately obtains from Eq. (60) that  $\tilde{h}_1(N) = 0$ . One can finally compute the coefficient  $\tilde{h}_2(N)$  which reads

$$\tilde{h}_2(N) = 2 \int d\mathbf{J} \langle [F']^2 + [F - \langle F_0 \rangle] F'' \rangle, \quad (62)$$

where we used the shortened notation  $F'' = [\partial^2 F / \partial t^2](t=0)$ . Using the same argument as in Eq. (61), we may get rid of the second term in the r.h.s. of Eq. (62). If we now also assume that the variance of  $[F']^2$  is small compared to its expectation, one can write  $\langle [F']^2 \rangle = [\langle F' \rangle]^2$ , so that Eq. (62) becomes

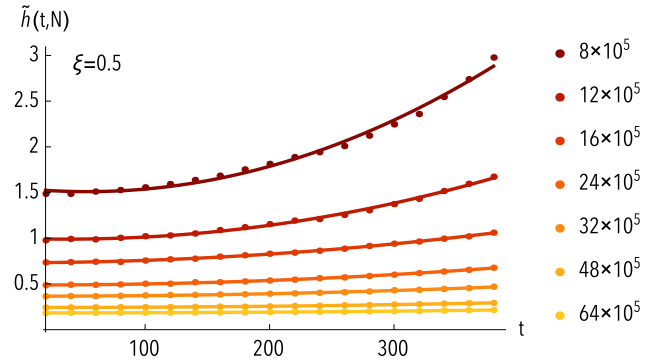
$$\tilde{h}_2(N) = 2 \int d\mathbf{J} \langle \langle F' \rangle \rangle^2. \quad (63)$$

Now the dependence of the term  $\langle F' \rangle$  with  $N$  follows from Eq. (50), so that one can write

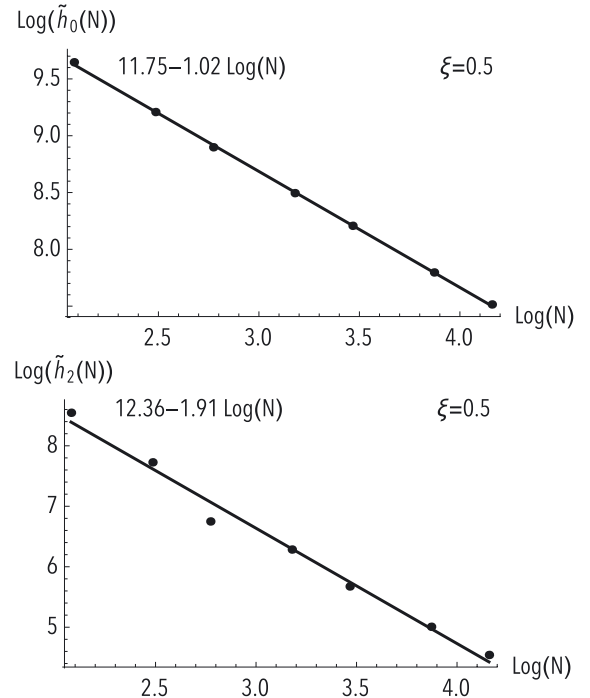
$$\tilde{h}_2(N) = \frac{\alpha_2}{N^2}, \quad (64)$$

where  $\alpha_2$  is an amplitude independent of  $N$ . This scaling is a prediction from the Balescu-Lenard formalism. If the secular evolution observed in S12's simulation was a Vlasov-only evolution, i.e. a collisionless evolution, one would expect a scaling of  $\tilde{h}_2$  such that  $\partial \tilde{h}_2 / \partial N = 0$ .

One may now compare these predictions to the scalings obtained from  $N$ -body runs. We considered number of particles given by  $N \in \{8, 12, 16, 24, 32, 48, 64\} \times 10^5$ , and for each of these values of  $N$ , we ran 32 different simulations with different initial conditions while using the  $N$ -body code described in Sect. 5.1. For each value of  $N$ , one may study the function  $t \mapsto \tilde{h}(t, N)$ , as illustrated in Fig. 5. Once the behaviour of the function  $t \mapsto \tilde{h}(t, N)$  is known, one can fit these to parabolas as in Eq. (57), so as to determine the behaviour of the functions  $N \mapsto \tilde{h}_0(N)$  and  $\tilde{h}_2(N)$ . The dependence with  $N$  of these coefficients is illustrated in Fig. 6. From the top panel of Fig. 6, we recover the



**Fig. 5.** Illustration of the behaviour of the function  $t \mapsto \tilde{h}(t, N)$  defined in Eq. (55), for an active fraction  $\xi = 0.5$ , when averaged on 32 different realisations for particles numbers  $N \in \{8, 12, 16, 24, 32, 48, 64\} \times 10^5$ . To compute  $\tilde{h}(t, N)$ , we binned the action-space domain  $(J_\phi, J_r) = [0; 2.5] \times [0; 0.2]$  in  $100 \times 50$  regions. The values of  $\tilde{h}(t, N)$  have also been uniformly renormalised so as to clarify this representation. The dots corresponds to the snapshots of the simulations for which  $\tilde{h}(t, N)$  was computed, whereas the lines correspond to second-order fits. As expected, the smaller the number of particles, the noisier the simulation and the larger  $\tilde{h}(t, N)$ .



**Fig. 6.** *Top panel:* illustration of the behaviour of the function  $\log(N) \mapsto \log(\tilde{h}_0(N))$ , where  $N$  has been rescaled by a factor  $10^{-5}$  so as to simplify the representation. The dots correspond to computed values thanks to Fig. 5, while the line corresponds to a linear fit, which takes the form  $\log(\tilde{h}_0(N)) \simeq 11.75 - 1.02 \log(N)$ . The coefficients  $\tilde{h}_0(N)$  have been uniformly renormalised so as to clarify this representation. *Bottom panel:* similar representation for the behaviour of the function  $\log(N) \mapsto \log(\tilde{h}_2(N))$ , whose linear fit takes the form  $\log(\tilde{h}_2(N)) \simeq 12.36 - 1.91 \log(N)$ . Similarly, the coefficients  $\tilde{h}_2(N)$  have been uniformly renormalised so as to clarify this representation.

scaling of  $\tilde{h}_0(N)$  derived in Eq. (59) due to Poisson shot noise. The bottom panel of Fig. 6 displays the scaling  $\tilde{h}_2(N) \propto N^{-1.91}$ . Given the finite number of simulations considered and the uncertainties in the fits, this is in good agreement with the result presented in Eq. (64). This scaling of  $\tilde{h}_2(N)$  with  $N$  therefore confirms the relevance of the Balescu-Lenard formalism to describe

the secular evolution of S12's stable Mestel disc. Specifically, as explained below Eq. (64), if the features observed in S12's simulation had only been the result of a collisionless mechanism, one would not have observed such a scaling of  $\tilde{h}_2(N)$  with  $N$ . This scaling confirms that the secular evolution of S12's stable Mestel disc is the result of a collisional evolution seeded by the discrete nature of the system, and the effect of amplified distant resonant encounters. Another probe of the collisional scaling, which allows us to get rid of Poisson shot noise as present in Eq. (59), is described in Appendix F.

### 5.3. Scaling with $\xi$

Since the novelty of the Balescu-Lenard formalism is to capture the effect of gravitational polarisation, we now further compare qualitatively the prediction from Sect. 4 with the results obtained from numerical simulations, by studying the impact of the active fraction  $\xi$  of the disc on the observed properties of the secular diffusion. Indeed, as detailed in Sect. 4.1, the disc considered in S12 had an active fraction of  $\xi = 0.5$ , so only one half of the potential was due to the active component. If one increases the active fraction of the disc, one increases the strength of the self-gravitating amplification, and therefore accelerates the secular evolution of the disc, while still remaining in a regime of collisional evolution. Therefore the scaling of  $\tilde{h}_2$  with  $N$  given by Eq. (64) remains the same, but the prefactor  $\alpha_2(\xi)$  increases because the secular evolution is amplified via a more efficient polarisation. The dependence of  $\alpha_2$  with  $\xi$  can be both measured from  $N$ -body simulations but also predicted using the Balescu-Lenard formalism via the calculations presented in Sect. 4.2. Let us consider the same sets of simulations as in Sect. 5.2, so that the number of particles are given by  $N \in \{8, 12, 16, 24, 32, 48, 64\} \times 10^5$ , and for each of these values of  $N$ , 32 different simulations with  $\xi = 0.6$  are performed, in order to carry out ensemble averages.

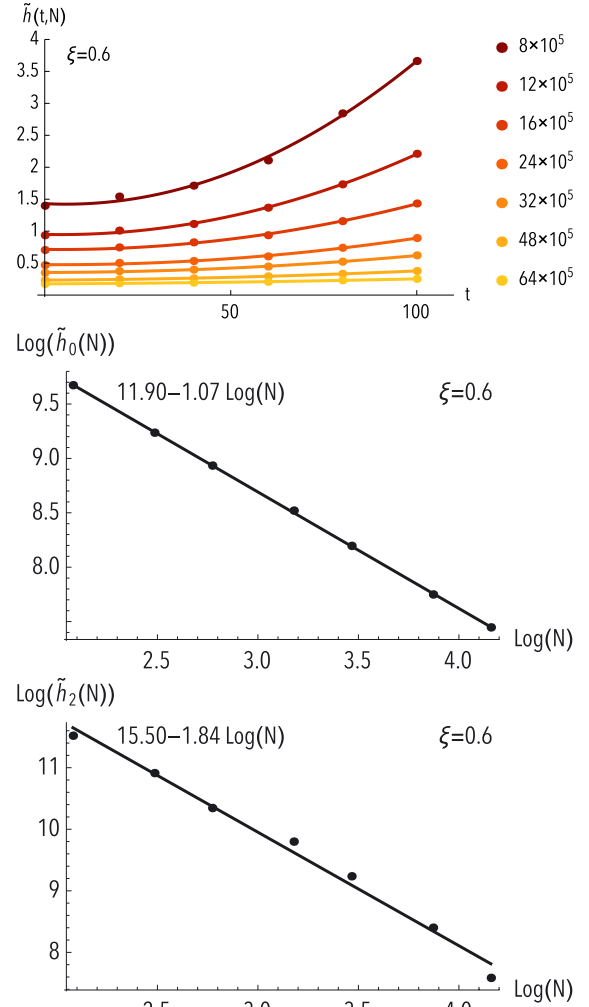
The equivalent of Figs. 5 and 6 for  $\xi = 0.6$  is illustrated in Fig. 7. Even for  $\xi = 0.6$ , one finds that the function  $t \mapsto \tilde{h}(t, N)$  follows a parabola given by Eq. (57). One also recovers the predicted scalings with  $N$  of the functions  $N \mapsto \tilde{h}_0(N)$  and  $N \mapsto \tilde{h}_2(N)$ , respectively representing the initial Poisson shot noise of the sampling and the collisional scaling of the Balescu-Lenard secular evolution. As expected, when the active fraction of the disc is increased the secular evolution is accelerated. Thanks to these fits, one can study the dependence of the ratio  $\alpha_2(\xi = 0.6)/\alpha_2(\xi = 0.5)$ , as defined in Eq. (64), both from numerical simulations as described in Figs. 6 and 7 and from the Balescu-Lenard equation using the matrix method described in Sect. 4.

From the fits of  $N \mapsto \tilde{h}_2(N)$  from Figs. 6 and 7, one can write  $\log(\tilde{h}_2(N)) \simeq 6.40 - 1.91(\log(N) - 3.12)$  for  $\xi = 0.5$  and  $\log(\tilde{h}_2(N)) \simeq 9.76 - 1.84(\log(N) - 3.12)$  for  $\xi = 0.6$ , where we shifted the intercept of the fits to correspond to the centre of the considered region  $\log(N) \in [\log(8); \log(64)]$ . One therefore obtains the ratio

$$\left. \frac{\alpha_2(0.6)}{\alpha_2(0.5)} \right|_{\text{NB}} \simeq \exp[9.76 - 6.40] \simeq 29. \quad (65)$$

One may now compare this  $N$ -body measurement to the same measurement performed via the Balescu-Lenard formalism. Following Eq. (63), one obtains that this ratio is given by

$$\frac{\alpha_2(\xi_1)}{\alpha_2(\xi_2)} = \frac{\int d\mathbf{J} [\text{div}(\mathcal{F}_{\text{tot}}^{\xi_1})]^2}{\int d\mathbf{J} [\text{div}(\mathcal{F}_{\text{tot}}^{\xi_2})]^2}, \quad (66)$$



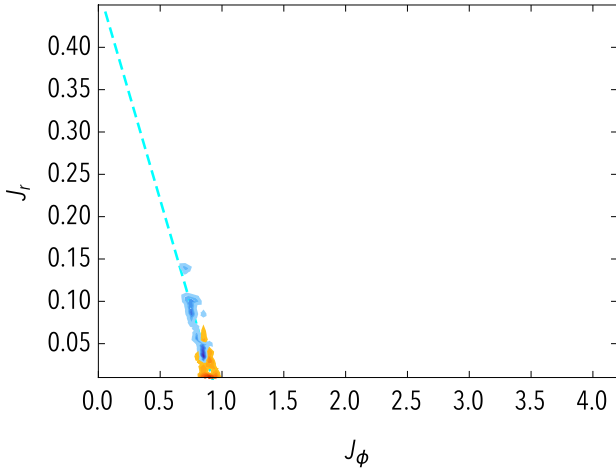
**Fig. 7.** Top panel: illustration of the behaviour of the function  $t \mapsto \tilde{h}(t, N)$  for an active fraction  $\xi = 0.6$ , using the same conventions as in Fig. 5. As expected, for an increased active fraction, the secular evolution of the system is fastened leading to a faster increase of the distance  $\tilde{h}(t, N)$ . Middle panel: behaviour of the function  $\log(N) \mapsto \log(\tilde{h}_0(N))$ , for an active fraction  $\xi = 0.6$ , using the same conventions as in Fig. 6. Its linear fit takes the form  $\log(\tilde{h}_0(N)) \simeq 11.90 - 1.07 \log(N)$ . One recovers the expected scaling of the Poisson shot noise sampling derived in Eq. (59). Bottom panel: behaviour of the function  $\log(N) \mapsto \log(\tilde{h}_2(N))$ , for an active fraction  $\xi = 0.6$ , using the same conventions as in Fig. 6. Its linear fit takes the form  $\log(\tilde{h}_2(N)) \simeq 15.48 - 1.84 \log(N)$ . One recovers the expected collisional scaling with  $N$  obtained in Eq. (64).

where  $\mathcal{F}_{\text{tot}}^{\xi}$  stands for the secular diffusion flux at  $t = 0^+$  with an active fraction  $\xi$ . The value of  $\alpha_2(\xi = 0.5)$  can be determined via Fig. 4, while the secular diffusion flux determined for  $\xi = 0.6$  is illustrated in Fig. 8. Thanks to the contours presented in Figs. 4 and 8, one can perform the same estimation as in Eq. (65) starting from the Balescu-Lenard predictions. In order to focus on the contributions associated with the resonant ridge, the integrals on  $\mathbf{J}$  in Eq. (66) were performed for  $J_{\phi} \in [0.5; 1.2]$  and  $J_r \in [0.06; 0.15]$ . We measured

$$\left. \frac{\alpha_2(0.6)}{\alpha_2(0.5)} \right|_{\text{BL}} \simeq 42. \quad (67)$$

Despite the difficulty of this measurement which required to consider a much more sensitive disc with  $\xi = 0.6$ , the ratios of  $\alpha_2$  measured either via direct  $N$ -body simulations as in Eq. (65)





**Fig. 8.** Map of  $N\text{div}(\mathcal{F}_{\text{tot}})$  with an active fraction  $\xi = 0.6$  and using the same conventions as in Fig. 4. The contours are spaced linearly between the minimum and the maximum of  $N\text{div}(\mathcal{F}_{\text{tot}})$ . The maximum value for the positive blue contours corresponds to  $N\text{div}(\mathcal{F}_{\text{tot}}) \simeq 4200$ , while the minimum value for the negative red contours corresponds to  $N\text{div}(\mathcal{F}_{\text{tot}}) \simeq -3200$ . As expected, one recovers that when the active fraction of the disc is increased, the susceptibility of the disc is increased, so that the norm of  $N\text{div}(\mathcal{F}_{\text{tot}})$  gets larger and the secular diffusion is fastened.

or via application of the Balescu-Lenard formalism in Eq. (67) are within the same order of magnitude. As a consequence, one indeed checks that the Balescu-Lenard equation is able to correctly capture the relative effect of the disc's susceptibility on the characteristics of the collisional secular diffusion. The strong consequence of modifying the active fraction, observed both in Eqs. (65) and (67), illustrates the relevance of the self-gravitating amplification in determining the typical timescale of secular diffusion of the system.

#### 5.4. Late-time evolution

The predictions of the Balescu-Lenard secular diffusion flux presented in Sect. 4 were only applied for the initial time of evolution, i.e. for the estimation of  $\mathcal{F}_{\text{tot}}(t = 0^+)$ . The  $N$ -body simulations presented in Sect. 5 allowed us to verify the appropriate scaling of the response of the system with the number of particles for the initial time of evolution as illustrated in Fig. 6. Using the Balescu-Lenard formalism to probe the late secular evolution of the system would require to evolve iteratively Eq. (2) over secular times. Such iterations are clearly beyond the scope of this first paper, however the use of  $N$ -body simulations allows us to start probing now such late times of evolution.

As discussed in Sect. 2, the Balescu-Lenard equation describes the long-term evolution of a discrete self-gravitating inhomogeneous system. Such a collisional evolution is only relevant for dynamically stable systems, i.e. systems assumed to be stable in the Vlasov sense, and remains valid as the system moves through a series of equilibria, while the drift and diffusion coefficients change with  $F$ . Because it has been obtained via a Taylor expansion of the dynamics at the order  $1/N$  in the number of particles, it remains valid only on secular timescales of the order  $Nt_D$ , with  $t_D$  the dynamical time.

On such secular timescales, a Balescu-Lenard evolution can lead to two different outcomes. On the one hand, if the system remains stable during its entire evolution, the Balescu-Lenard

equation will tend towards a  $1/N$ -stationary state<sup>2</sup>. Once such a stationary state of evolution has been reached, the dynamics is then governed by the next order kinetic effects in  $1/N^2$ , which are not captured by the Balescu-Lenard equation. On the other hand, the Balescu-Lenard collisional evolution may also lead to a destabilisation of the system. Indeed, the long-term effects of the collisional diffusion, because they lead to an irreversible diffusion of the DF, may change its current state w.r.t. the collisionless (Vlasov) dynamics. After a slow and stable evolution sourced by collisional  $1/N$  effects, the system may then become dynamically unstable with respect to collisionless dynamics, which becomes the main driver of its later-time evolution, as was suggested by Sellwood (2012). S12 observes an out-of-equilibrium transition between the  $1/N$  Balescu-Lenard collisional evolution and the collisionless Vlasov evolution.

One can illustrate such a transition using the  $N$ -body simulations presented previously. In order to capture the change of regime of evolution within the disc (collisional vs. collisionless), for a given value of the number of particles, we define the quantity  $\Sigma_2(t, N)$  as

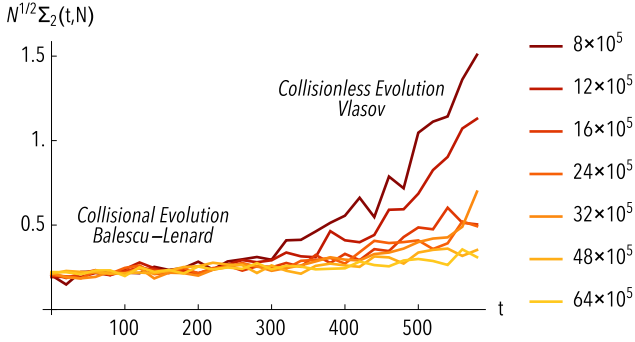
$$\Sigma_2(t, N) = \left\langle \int_{R_{\text{inf}}}^{R_{\text{sup}}} dR R d\phi \Sigma_{\text{star}}(t, N, R, \phi) e^{-i2\phi} \right\rangle = \left\langle \mu \sum_n e^{-i2\phi_n} \right\rangle, \quad (68)$$

where as in Eq. (55), the operator  $\langle \cdot \rangle$  corresponds to the ensemble average, approximated here with the arithmetic average over the  $p = 32$  different realisations of simulations for the same number of particles  $N$ . The radii considered are restricted to the range  $R \in [R_{\text{inf}}; R_{\text{sup}}] = [1.2; 5]$ , where the active surface density of the disc is little affected by the inner and outer tapers. Finally, to obtain the second equality in Eq. (68), as in Eq. (C.2), we replaced the active surface density of the disc by a discrete sum over all the particles of the system, where the sum on  $n$  is restricted to all the particle whose radius lies between  $R_{\text{inf}}$  and  $R_{\text{sup}}$ , while their azimuthal phase was written as  $\phi_n$ . Such a quantity allows us to probe easily the presence of strong non-axisymmetric features within the disc. During the initial Balescu-Lenard collisional evolution of the system, one expects low values of  $\Sigma_2$ . Indeed, during this evolution, one relies on the phase averaging approximation, which assumes that  $F = F(\mathbf{J}, t)$ , so that the DF of the system does not depend on the angles  $\theta$ . During this collisional phase,  $\Sigma_2$  still remains non-zero because the system develops transient spiral waves, which sustain the secular evolution. On the long-term, this collisional evolution, through an irreversible diffusion of the DF, leads to a destabilisation of the system. Eventually, the dynamical drivers of evolution are no longer discrete resonant collisional effects but exponentially growing dynamical instabilities. In this regime of collisionless unstable evolution, one expects much larger values of  $\Sigma_2$ , because of the appearance of strong non-axisymmetric bars within the disc. This bifurcation between these two regimes of diffusion is illustrated in Fig. 9, through the behaviour of the function  $t \mapsto \sqrt{N}\Sigma_2(t, N)$ <sup>3</sup>. One can similarly observe this transition

<sup>2</sup> Boltzmann DF of the form  $\exp[-\beta H(\mathbf{J})]$ , when physically reachable, are obvious stationary states of the Balescu-Lenard equation.

<sup>3</sup> A similar dynamical phase transition has been observed (Campa et al. 2008) in a toy model of systems with long-range interactions called the Hamiltonian Mean Field (HMF) model. During the slow collisional evolution, because of finite- $N$  effects, the distribution function of the system changes with time. In certain cases, the system may become dynamically (Vlasov) unstable and undergo a rapid phase transition from a homogeneous phase to an inhomogeneous phase. This phase transition can be monitored by the magnetisation (see Fig. 1 of Campa et al. 2008) which is an order parameter playing a role similar to  $\Sigma_2(t, N)$ .





**Fig. 9.** Behaviour of the function  $t \mapsto \sqrt{N} \Sigma_2(t)$  as defined in Eq. (68), for various values of the number of particles. The prefactor  $\sqrt{N}$  has been added so as to mask Poisson shot noise, allowing the initial values of  $\sqrt{N} \Sigma_2$  to be independent of  $N$ . It illustrates the bifurcation between the initial Balescu-Lenard collisional evolution, for which low values of  $\Sigma_2$  are expected and the collisionless Vlasov evolution for which the system is no more axisymmetric leading to larger values of  $\Sigma_2$ . As expected, the larger the number of particles, the later the transition.

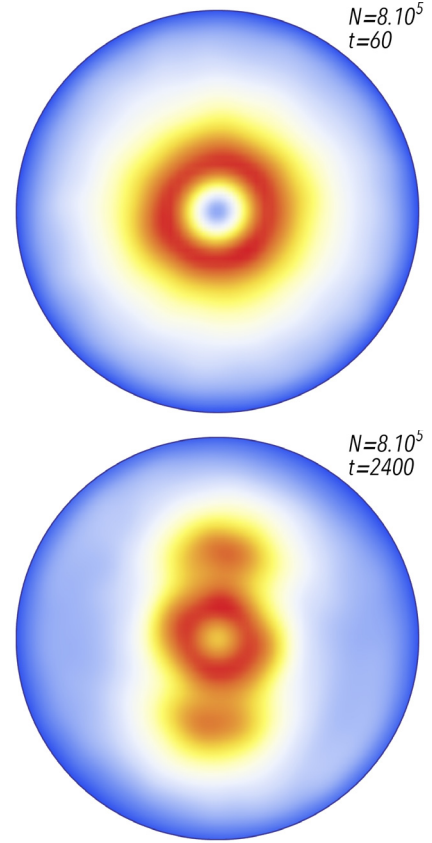
directly by looking at the active surface density  $\Sigma_t(R, \phi, t)$  for these two different regimes. This is illustrated in Fig. 10, where one recovers that in the late time collisionless regime of evolution, the galaxy becomes strongly non-axisymmetric. S12 found that just after the disc becomes unstable, the pattern of the spiral response is consistent with the ILR frequency corresponding to the ridge<sup>4</sup>. Hence the phase transition observed in Fig. 10 is driven by all the free energy available in a cold disc, which via spiral transients secularly heats the disc, but only along a very tight resonant direction. This in turn leads the disc towards an orbital instability, transverse to the resonance (via the direct azimuthal analogue to the two stream instability in plasma physics, Lynden-Bell 1979; Pichon 1994). Qualitatively, one expects that the more massive and the narrower the ridge, the larger the number of orbits trapped in ILR resonance with little relative azimuthal dispersion, and the earlier the instability (Penrose 1960; Pichon & Lynden-Bell 1993).

In closing, it is quite striking that an isolated galactic disc, fully stable in the mean field sense, will, given time, drive itself through two-point resonant correlations towards instability, demonstrating the extent to which such cold systems are truly secularly metastable.

## 6. Conclusion

Most astrophysical discs formed through dissipative processes and have typically evolved over many dynamical times. Even in isolation, the long-range force of gravity allows their components to interact effectively through resonances, which given time may drive them secularly towards more likely equilibria. Such processes are captured by recent extensions of kinetic theories rewritten in angle-action variables (Heyvaerts 2010; Chavanis 2012). Solving these equations provide astronomers with a unique opportunity to quantify the induced secular angular momentum redistribution within these discs (Lynden-Bell & Kalnajs 1972) on cosmic timescales. While challenging, the numerical computation of the corresponding diffusion and drift coefficients is, as demonstrated, within reach of a relatively straightforward extension of the so-called matrix method

<sup>4</sup> One could also check that the disc's distribution function corresponds at that stage to an unstable configuration, using the matrix method described in Appendix C.



**Fig. 10.** Illustration of the active surface density  $\Sigma_t$  for a  $N$ -body run with  $N = 8 \times 10^5$ , restricted to the range  $R \leq 6$ . *Top panel:* active surface density  $\Sigma_t$  at an early time  $t = 60$ , for which the galaxy remains globally axisymmetric. In this regime, the dynamics of the system is collisional and governed by the Balescu-Lenard Eq. (2). *Bottom panel:* active surface density  $\Sigma_t$  at a much later time  $t = 2400$ . The galaxy is then strongly non-axisymmetric. In this regime, the dynamics of the system is collisionless and governed by the Vlasov equation.

(Kalnajs 1976), which computes the orbital response of self-gravitating discs using quadratures and linear algebra.

Paper I presented asymptotic expressions in the tightly wound limit and provided a qualitative insight into the physical processes at work during the secular diffusion of a self-gravitating discrete disc. Conversely, in this paper, we computed numerically the drift and diffusion coefficients of the inhomogeneous Balescu-Lenard equation for such infinitely thin stellar discs. The self-gravity of the disc was taken into account via the matrix method, validated on unstable Mestel discs. We computed the divergence of the flux density in action space,  $\text{div}(\mathcal{F}_{\text{tot}})$ . Swing amplification was shown to provide a significant boost for the diffusion timescale, which now matches the numerically measured one. These computations are the first exact calculation of the Balescu-Lenard diffusion and drift coefficients in the context of inhomogeneous multi-periodic systems. They capture the essence of self-induced evolution (nature), which should compete with environmentally induced evolution (nurture). We then compared these predictions to idealised numerical simulations of stable stationary and truncated Mestel discs sampled by pointwise particles, which were evolved for hundreds of dynamical times. Using ensemble averages of our  $N$ -body runs, we also identified a clear signature of the Balescu-Lenard process in the scaling of the diffusion features with  $N$  and  $\xi$ , the fraction of the mass within the disc. As originally

identified by Goldreich & Lynden-Bell (1965), Julian & Toomre (1966) in the context of their linear response, the susceptibility of cold self-gravitating discs plays a critical role for their secular evolution as it is squared in the Balescu-Lenard equation, which boosts considerably the effect of discreteness. Indeed, both the numerical experiments and our computation of the fluxes show that  $N_{\text{eff}} \sim N/10^4$ , which is consistent with the predicted rescaling in  $1/\mathcal{D}^2$  (it was shown forty years ago that for a Mestel disc  $1/\mathcal{D} \sim 10^2$ , depending on the exact temperature of the disc, Toomre 1981).

Jointly with Paper I we now have a qualitative and quantitative understanding of the initial secular orbital diffusion process induced by the discreteness of galactic discs. Our qualitative understanding allows us to identify the role played by the square susceptibility in boosting the diffusion. Our quantitative agreement in both amplitude, position, width and scaling of the induced orbital signatures strongly suggests that secular evolution is indeed driven by resonances as captured by the Balescu-Lenard formalism, and that it does not depend on the initial phases of the disc (since the matching Balescu-Lenard fluxes are phase averaged). It demonstrates that this equation initially reproduces the observed evolution of self-gravitating discs driven by resonant two-point correlations beyond the mean field approximation.

The next step will be to evolve iteratively Eq. (2) over a Hubble time, and compare with the result of  $N$ -body simulations. One should also model it jointly with an externally induced orbital diffusion (Fouvry et al. 2015b) arising from for example a (possibly anisotropic) cosmic environment (Codis et al. 2012, 2015) so as to assess which process dominates. At the technical level, the Balescu-Lenard formalism should be used to (in)validate  $N$ -body integrators accuracy on secular timescales. There are indeed very few analytical predictions on which to calibrate  $N$ -body experiments in this regime. Such an exploration would also allow us to get a better grasp of the impact of the numerical parameters used in the  $N$ -body integration (such as timestep, mesh size or softening length) on the long-term dynamics of the system.

Beyond the application described in this paper, the Balescu-Lenard formalism may in the future also be numerically implemented to describe for instance the secular diffusion of giant molecular clouds in galactic discs (which in turn could play a role in migration-driven metallicity gradients and disc thickening), the secular migration of planetesimals in partially self-gravitating proto-planetary debris discs, or even the long-term evolution of population of stars and gas blobs near the Galactic centre. In 3D, assuming spherical symmetry, its implementation could be useful to describe spherical systems dominated by radial orbits, or the secular evolution of tidal debris in our possibly flattened galactic halo using Stäckel potentials.

*Acknowledgements.* We thank the referee, James Binney, for numerous valuable comments. J.B.F. thanks the Institute of Astronomy, Cambridge, for hospitality while this investigation was initiated. J.B.F. and C.P. also thank the theoretical physics sub-department, Oxford, for hospitality and the CNRS-Oxford exchange program for funding. J.B.F., C.P. and P.H.C. also thank the CNRS Inphyniti program for funding. C.P. thanks Clare and Churchill college, Cambridge, the French embassy and the community of <http://mathematica.stackexchange.com> for their help. J.M. thanks the Institut d’Astrophysique de

Paris for their hospitality and the Ville de Paris for a “Research in Paris” award. We thank Donald Lynden-Bell, John Papaloizou, Walter Dehnen, Rebekka Bieri, Laura Monk, Gordon Ogilvie, Dmitry Pogosyan and Simon Prunet for stimulating discussions, and Eric Pharabod for his help with Fig. 4. Special thanks to Stephane Rouberol for customising the Horizon cluster for our purposes. This work is partially supported by the Spin(e) grants ANR-13-BS05-0005 of the French Agence Nationale de la Recherche (<http://cosmicorigin.org>) and by the LABEX Institut Lagrange de Paris (under reference ANR-10-LABX-63) which is funded by ANR-11-IDEX-0004-02.

## References

- Aumer, M., & Binney, J. J. 2009, *MNRAS*, **397**, 1286  
 Balescu, R. 1960, *Phys. Fluids*, **3**, 52  
 Binney, J., & Tremaine, S. 2008, *Galactic Dynamics*, 2nd edn., Princeton Series in Astrophysics (Princeton University Press)  
 Born, M. 1960, *The Mechanics of the Atom* (F. Ungar Pub. Co.)  
 Campa, A., Chavanis, P.-H., Giansanti, A., & Morelli, G. 2008, *Phys. Rev. E*, **78**, 040102  
 Chandrasekhar, S. 1942, *Principles of Stellar Dynamics* (University of Chicago Press)  
 Chavanis, P.-H. 2012, *Physica A*, **391**, 3680  
 Chavanis, P.-H. 2013, *A&A*, **556**, A93  
 Codis, S., Pichon, C., Devriendt, J., et al. 2012, *MNRAS*, **427**, 3320  
 Codis, S., Pichon, C., & Pogosyan, D. 2015, *MNRAS*, **452**, 3369  
 Dehnen, W. 1998, *AJ*, **115**, 2384  
 Earn, D. J. D., & Sellwood, J. A. 1995, *ApJ*, **451**, 533  
 Evans, N. W., & Read, J. C. A. 1998a, *MNRAS*, **300**, 83  
 Evans, N. W., & Read, J. C. A. 1998b, *MNRAS*, **300**, 106  
 Fouvry, J.-B., Pichon, C., & Chavanis, P.-H. 2015a, *A&A*, **581**, A139  
 Fouvry, J.-B., Pichon, C., & Prunet, S. 2015b, *MNRAS*, **449**, 1967  
 Goldreich, P., & Lynden-Bell, D. 1965, *MNRAS*, **130**, 125  
 Goldstein, H. 1950, *Classical mechanics* (Addison-Wesley)  
 Heyvaerts, J. 2010, *MNRAS*, **407**, 355  
 Hörmander, L. 2003, *The analysis of linear partial differential operators. I. Classics in Mathematics* (Springer-Verlag)  
 Jeans, J. 1929, *Astronomy and Cosmogony* (Cambridge Univ. Press)  
 Julian, W. H., & Toomre, A. 1966, *ApJ*, **146**, 810  
 Kalnajs, A. J. 1972, in *IAU Colloq. 10: Gravitational N-Body Problem*, ed. M. Lecar, *Astrophys. Space Sci. Lib.*, **31**, 13  
 Kalnajs, A. J. 1976, *ApJ*, **205**, 745  
 Landau, L. 1936, *Phys. Z. Sowj. Union*, **10**, 154  
 Lenard, A. 1960, *Annals of Physics*, **10**, 390  
 Lynden-Bell, D. 1979, *MNRAS*, **187**, 101  
 Lynden-Bell, D., & Kalnajs, A. J. 1972, *MNRAS*, **157**, 1  
 McMillan, P. J. 2013, *MNRAS*, **430**, 3276  
 Mestel, L. 1963, *MNRAS*, **126**, 553  
 Penrose, O. 1960, *Phys. Fluids*, **3**, 258  
 Pichon, C. 1994, Ph.D. Thesis, University of Cambridge  
 Pichon, C., & Cannon, R. C. 1997, *MNRAS*, **291**, 616  
 Pichon, C., & Lynden-Bell, D. 1993, in *Statistical Description of Transport in Plasma, Astro- and Nuclear Physics*, eds. J. Misquich, G. Pelletier, & P. Schuck, 261  
 Quinlan, G. D., & Tremaine, S. 1992, *MNRAS*, **259**, 505  
 Rosenbluth, M., MacDonald, W., & Judd, D. 1957, *Phys. Rev.*, **107**, 1  
 Sellwood, J. A. 1983, *J. Computat. Phys.*, **50**, 337  
 Sellwood, J. A. 2010, *MNRAS*, **409**, 145  
 Sellwood, J. A. 2012, *ApJ*, **751**, 44  
 Sellwood, J. A., & Evans, N. W. 2001, *ApJ*, **546**, 176  
 Toomre, A. 1964, *ApJ*, **139**, 1217  
 Toomre, A. 1977, *ARA&A*, **15**, 437  
 Toomre, A. 1981, in *Structure and Evolution of Normal Galaxies*, eds. S. M. Fall, & D. Lynden-Bell, 111  
 Tremaine, S., & Weinberg, M. D. 1984, *MNRAS*, **209**, 729  
 Vlasov, A. 1938, *Zh. Eksp. i Teor. Fiz.*, **8**, 291  
 Weinberg, M. D. 1993, *ApJ*, **410**, 543  
 Wielen, R. 1977, *A&A*, **60**, 263  
 Zang, T. A. 1976, Ph.D. Thesis, Massachusetts Institute of Technology

## Appendix A: Kalnajs basis

We now detail the properties of the basis introduced in Kalnajs (1976) to describe 2D discs<sup>5</sup>. The basis depends on two parameters: an index  $k_{\text{Ka}} \in \mathbb{N}$  and a scale radius  $r_{\text{Ka}} \in \mathbb{R}^+$ . In all the upcoming formulae of this section, in order to shorten the notations, we write  $r$  for the dimensionless quantity  $r/r_{\text{Ka}}$ . As introduced previously in Eq. (16), the basis elements depend on two indices: the azimuthal number  $\ell$  and the radial index  $n$ . One should note that we have  $\ell \geq 0$  and  $n \geq 0$ . The radial component of the potential elements are then of the form

$$\mathcal{U}_n^\ell(r) = -\frac{\sqrt{G}}{r_{\text{Ka}}^{1/2}} \mathcal{P}(k_{\text{Ka}}, \ell, n) r^\ell \sum_{i=0}^k \sum_{j=0}^n \alpha_{\text{Ka}}(k_{\text{Ka}}, \ell, n, i, j) r^{2i+2j}. \quad (\text{A.1})$$

The radial component of the density elements is given by

$$\begin{aligned} \mathcal{D}_n^\ell(r) &= \frac{(-1)^n}{\sqrt{G} r_{\text{Ka}}^{3/2}} \mathcal{S}(k_{\text{Ka}}, \ell, n) (1-r^2)^{k_{\text{Ka}}-1/2} r^\ell \\ &\times \sum_{j=0}^n \beta_{\text{Ka}}(k_{\text{Ka}}, \ell, n, j) (1-r^2)^j. \end{aligned} \quad (\text{A.2})$$

In Eqs. (A.1) and (A.2), the coefficients  $\mathcal{P}(k, \ell, n)$  and  $\mathcal{S}(k, \ell, n)$  are defined by

$$\begin{aligned} \mathcal{P}(k, \ell, n) &= \left\{ \frac{[2k+\ell+2n+(1/2)] \Gamma[2k+\ell+n+(1/2)]}{\Gamma[2k+n+1] \Gamma^2[\ell+1] \Gamma[n+1]} \right. \\ &\quad \left. \times \Gamma[\ell+n+(1/2)] \right\}^{1/2}, \end{aligned} \quad (\text{A.3})$$

and

$$\begin{aligned} \mathcal{S}(k, \ell, n) &= \frac{\Gamma[k+1]}{\pi \Gamma[2k+1] \Gamma[k+(1/2)]} \\ &\times \left\{ \frac{[2k+\ell+2n+(1/2)] \Gamma[2k+n+1] \Gamma[2k+\ell+n+(1/2)]}{\Gamma[\ell+n+(1/2)] \Gamma[n+1]} \right\}^{1/2}. \end{aligned} \quad (\text{A.4})$$

Finally, in Eqs. (A.1) and (A.2), we have also introduced

$$\begin{aligned} \alpha_{\text{Ka}}(k, \ell, n, i, j) &= \frac{[-k]_i [\ell+(1/2)]_i [2k+\ell+n+(1/2)]_j}{[\ell+1]_i [1]_i [\ell+i+1]_j [\ell+(1/2)]_j [1]_j} \\ &\times [i+\ell+(1/2)]_j [-n]_j, \end{aligned} \quad (\text{A.5})$$

and

$$\beta_{\text{Ka}}(k, \ell, n, j) = \frac{[2k+\ell+n+(1/2)]_j [k+1]_j [-n]_j}{[2k+1]_j [k+(1/2)]_j [1]_j}. \quad (\text{A.6})$$

In the two previous expressions, we introduced the rising Pochhammer symbol  $[a]_i$  defined as

$$[a]_i = \begin{cases} 1 & \text{if } i = 0, \\ a(a+1) \dots (a+n-1) & \text{if } i > 0. \end{cases} \quad (\text{A.7})$$

<sup>5</sup> See also Earn & Sellwood (1995) for a similar rewriting of the basis normalisations.

## Appendix B: Calculation of $\mathfrak{N}$

We now detail how the analytical function  $\mathfrak{N}$  from Eq. (32) may be computed. In order to ease the implementation of its computation, we rewrite  $\mathfrak{N}$  in an undimensionalised way as follows

$$\begin{aligned} \mathfrak{N}(a_g, b_g, c_g, a_h, b_h, c_h, \eta, \Delta r) &= \\ &\int_{-\frac{\Delta r}{2}}^{\frac{\Delta r}{2}} \int_{-\frac{\Delta r}{2}}^{\frac{\Delta r}{2}} dx_p dx_a \frac{a_g + b_g x_p + c_g x_a}{a_h + b_h x_p + c_h x_a + i\eta} \\ &= \frac{a_g}{a_h} (\Delta r)^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx dy \frac{1 + \frac{b_g \Delta r}{a_g} x + \frac{c_g \Delta r}{a_g} y}{1 + \frac{b_h \Delta r}{a_h} x + \frac{c_h \Delta r}{a_h} y + i \frac{\eta}{a_h}} \\ &= \frac{a_g}{a_h} (\Delta r)^2 \mathfrak{N}_D \left[ \frac{b_g \Delta r}{a_g}, \frac{c_g \Delta r}{a_g}, \frac{b_h \Delta r}{a_h}, \frac{c_h \Delta r}{a_h}, \frac{\eta}{a_h} \right], \end{aligned} \quad (\text{B.1})$$

where we assumed that  $a_g, a_h \neq 0$  and used the change of variables  $x = x_p/\Delta r$  and  $y = x_a/\Delta r$ . We also defined the dimensionless function  $\mathfrak{N}_D$  as

$$\mathfrak{N}_D(b, c, e, f, \eta) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} dx dy \frac{1 + bx + cy}{1 + ex + fy + i\eta}. \quad (\text{B.2})$$

To compute this integral, we may now exhibit a function  $G(x, y)$  such that

$$\frac{\partial^2 G}{\partial x \partial y} = \frac{1 + bx + cy}{1 + ex + fy + i\eta}. \quad (\text{B.3})$$

One possible choice for  $G$  is given by

$$\begin{aligned} G(x, y) &= \frac{1}{4e^2 f^2} \log \left[ e^2 x^2 + 2e(fxy + x) + f^2 y^2 + 2fy + \eta^2 + 1 \right] \\ &\times \left\{ bf(e^2 x^2 - (fy + i\eta + 1)^2) + 2ef(ex + i\eta + 1) - ce(ex + i\eta + 1)^2 \right\} \\ &+ \frac{i}{2e^2 f^2} \left\{ \frac{\pi}{2} - \tan^{-1} \left[ \frac{ex + fy + 1}{\eta} \right] \right\} \\ &\times \left\{ bf(e^2 x^2 - (fy + i\eta + 1)^2) + 2ef(ex + i\eta + 1) - ce(ex + i\eta + 1)^2 \right\} \\ &+ \frac{y}{4e^2 f} \left\{ f(-4e + b(2ex + fy + 2i\eta + 2)) \right. \\ &\quad \left. + ce(2ex - fy + 2i\eta + 2) + 2ef(cy + 2) \log[ex + fy + i\eta + 1] \right\}. \end{aligned} \quad (\text{B.4})$$

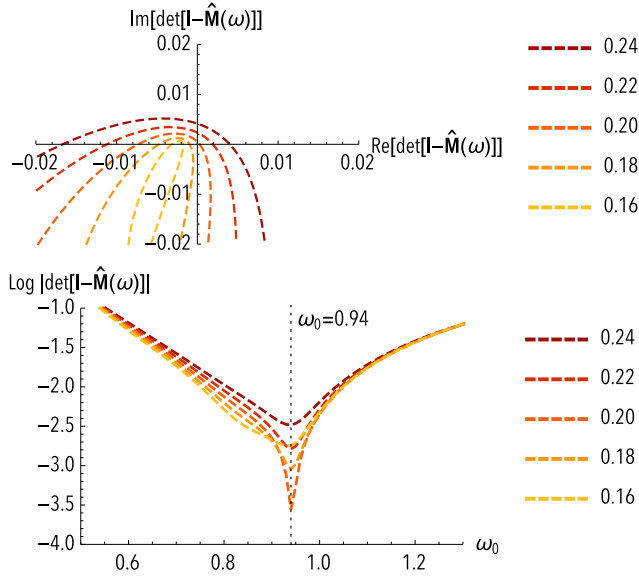
One should note in the previous expression the presence of a complex logarithm and a  $\tan^{-1}$ . However, because  $e, f, \eta \in \mathbb{R}$  and  $\eta \neq 0$ , one can easily show that the arguments of both of these functions never cross the usual branch-cut of these functions  $\{\text{Im}(z) = 0; \text{Re}(z) \leq 0\}$ . As a consequence, the expression (B.2) can immediately be computed as

$$\mathfrak{N}_D = G\left[\frac{1}{2}, \frac{1}{2}\right] - G\left[\frac{1}{2}, -\frac{1}{2}\right] - G\left[-\frac{1}{2}, \frac{1}{2}\right] + G\left[-\frac{1}{2}, -\frac{1}{2}\right]. \quad (\text{B.5})$$

## Appendix C: Response matrix and $N$ -body validations

The computation of the response matrix as described in Sect. 3 was validated by recovering the results of the pioneer work of Zang (1976), extended in Evans & Read (1998a,b), and recovered numerically in Sellwood & Evans (2001). These papers predicted the precession rate  $\omega_0 = m_\phi \Omega_p$  and growth rate  $\eta = s$  of the unstable modes of a truncated Mestel disc similar to the stable one described in Sect. 4.1. To build up an unstable disc similar to the ones considered in these previous works,

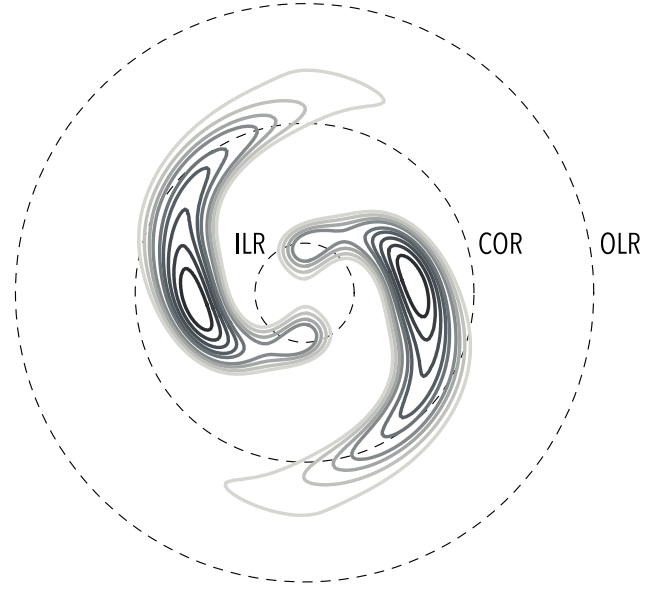




**Fig. C.1.** *Top panel:* zoomed Nyquist contours in the complex plane of  $\omega_0 \mapsto \det[\mathbf{I} - \hat{\mathbf{M}}(\omega_0, \eta)]$  obtained via the matrix method for a truncated Mestel disc with  $\nu_t = 6$  and  $q = 6$ , looking for  $m_\phi = 2$  modes. Each contour corresponds to a fixed value of  $\eta$ . For a growth rate of  $\eta \approx 0.20$ , one can note that the contour crosses the origin, which corresponds to the presence of an unstable mode. *Bottom panel:* illustration of the function  $\omega_0 \mapsto \log |\det[\mathbf{I} - \hat{\mathbf{M}}(\omega_0, \eta)]|$  for the same truncated Mestel disc. Each line corresponds to a fixed value of  $\eta$ . Such a representation allows us to determine the pattern speed  $\omega_0 = m_\phi \Omega_p \approx 0.94$  of the unstable mode.

one has to consider a fully active disc, so that  $\xi = 1$ . So as to have  $Q = 1$  (Toomre 1964), the velocity dispersion within the disc is given by  $q = 6$ , where the parameter  $q$  has been introduced in Eq. (46). Finally, a last parameter one can tune in order to modify the properties of the disc is the truncation index of the inner tapering  $\nu_t$  defined in Eq. (48). While looking only for  $m_\phi = 2$  modes, we considered three different truncations indices given by  $\nu_t = 4, 6, 8$ . To compute the response matrix, we used the same numerical parameters as described in Sect. 4.2. Looking for unstable modes amounts to looking for complex frequencies  $\omega = \omega_0 + i\eta$  such that the response matrix  $\hat{\mathbf{M}}(\omega_0, \eta)$  from Eq. (24) possesses an eigenvalue equal to 1. Such a complex frequency is then associated with an unstable mode of pattern speed  $\omega_0$  and growth rate  $\eta$ . To determine the growth rate and pattern speed of the unstable modes, we relied on Nyquist contours similarly to the technique presented in Pichon & Cannon (1997). For a fixed value of  $\eta$ , one can study the continuous complex curve  $\omega_0 \mapsto \det[\mathbf{I} - \hat{\mathbf{M}}(\omega_0, \eta)]$ . Because for  $\eta \rightarrow +\infty$ , one has  $|\hat{\mathbf{M}}(\omega_0, \eta)| \rightarrow 0$ , the number of windings of this curve around the origin gives a lower bound on the number of unstable modes with a growth rate superior to  $\eta$ . By decreasing the value of  $\eta$ , one can then determine the largest value of  $\eta$  admitting an unstable mode, and therefore the most unstable mode of the disc. The Nyquist contours obtained for the truncation index  $\nu_t = 6$  are illustrated in Fig. C.1, while the measurements are gathered in Table C.1.

Once the characteristics  $(\omega_0, \eta)$  of the unstable modes have been determined, one can study in the physical space the shape of the mode. Indeed, for  $\omega = \omega_0 + i\eta$ , one can compute  $\hat{\mathbf{M}}(\omega_0, \eta)$ , and numerically diagonalise this matrix. One then considers its eigenvector  $\mathbf{X}_{\text{mode}}$  (of size  $n_{\text{max}}$ , where  $n_{\text{max}}$  is the number of basis elements considered) associated with the



**Fig. C.2.** Unstable mode for the truncated  $\nu_t = 4$  Mestel disc recovered via the matrix method presented in Sect. 3. Only positive contour levels are shown and are spaced linearly between 10% and 90% of the maximum norm. The radii associated with the resonances ILR, COR and OLR have been represented, as given by  $\omega_0 = \mathbf{m} \cdot \boldsymbol{\Omega}(R_m)$ , where the intrinsic frequencies  $\boldsymbol{\Omega}(R) = (\Omega_\phi(R), \kappa(R))$  have to be computed within the epicyclic approximation. For a Mestel disc, they are given by  $\Omega_\phi(R) = V_0/R$  and  $\kappa(R) = \sqrt{2} \Omega_\phi(R)$ .

eigenvalue almost equal to 1. The shape of the mode is then immediately given by

$$\Sigma_{\text{mode}}(R, \phi) = \text{Re} \left[ \sum_p \mathbf{X}_{\text{mode}}^p \Sigma^{(p)}(R, \phi) \right], \quad (\text{C.1})$$

where  $\Sigma^{(p)}$  are the considered surface density basis elements. The shape of the recovered unstable mode for the truncated  $\nu_t = 4$  Mestel disc is illustrated in Fig. C.2.

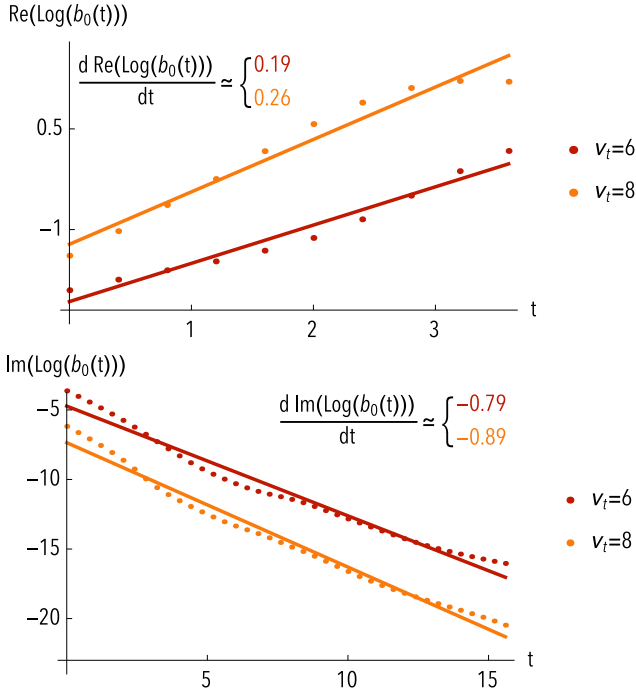
The same unstable modes were also used to validate the  $N$ -body code presented in Sect. 5. To run these simulations, we used the same sampling technique as described in Appendix E. In order not to be significantly impacted by the absence of a quiet start sampling (Sellwood 1983), for each value of  $\nu_t$ , the measurements were performed with simulations of  $20M$  particles. As observed in Sellwood & Evans (2001), the appropriate setting of the parameters of the  $N$ -body code are crucial to recover correctly the unstable modes of a disc. We considered a grid made with  $N_{\text{mesh}} = 120$  grid cells, while using a softening length equal to  $\varepsilon = R_i/60$ . As described in Sect. 5, we similarly restricted the perturbing forces only to the harmonic sector  $m_\phi = 2$ , using  $N_{\text{ring}} = 2400$  radial rings, with  $N_\phi = 720$  azimuthal points. In order to extract the properties of the mode present within the disc, one may proceed as follows. For each simulation snapshot, one can estimate the active surface density within the disc via

$$\Sigma_{\text{star}}(\mathbf{x}, t) = \mu \sum_i \delta_{\text{D}}(\mathbf{x} - \mathbf{x}_i(t)), \quad (\text{C.2})$$

where the sum on  $i$  is made on all the particles of the simulation and  $\mathbf{x}_i(t)$  is the position of the  $i$ th particle at time  $t$ . Such a surface density can be decomposed on the basis elements from Eq. (3) in the form

$$\Sigma_{\text{star}}(\mathbf{x}, t) = \sum_p b_p(t) \Sigma^{(p)}(\mathbf{x}), \quad (\text{C.3})$$





**Fig. C.3.** Measurement of the growth rate  $\eta$  and pattern speed  $\omega_0$  of the  $m_\phi = 2$  unstable modes of truncated Mestel discs with a random velocity given by  $q = 6$  for various values of the truncation index  $v_t = 6, 8$ . The basis coefficient plotted is associated with the basis element  $(\ell, n) = (2, 0)$ , using the same basis elements as for the matrix method in Sect. 4.2.

where the sum on  $p$  is made on all the basis elements  $\Sigma^{(p)}$  considered. The effective basis elements used during our measurements are the same as the ones used in the matrix method from Sect. 4.2. Thanks to the biorthogonality property from Eq. (3), the coefficients  $b_p(t)$  can be immediately determined as

$$b_p(t) = - \int d\mathbf{x} \Sigma_{\text{star}}(\mathbf{x}, t) \psi^{(p)*}(\mathbf{x}) = -\mu \sum_i \psi^{(p)*}(\mathbf{x}_i(t)). \quad (\text{C.4})$$

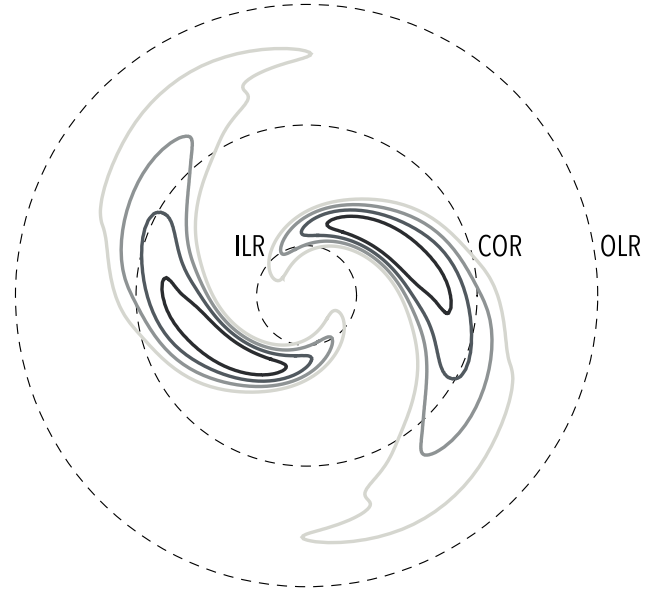
As we are looking for unstable modes within the disc, we expect to have  $b_p(t) \propto \exp[-i(\omega_0 + i\eta)t]$ , where  $\omega_0 = m_\phi \Omega_p$  is the pattern speed of the mode and  $\eta = s$  its growth rate. As a consequence, one immediately obtains that

$$\frac{d \text{Re}(\log(b_p(t)))}{dt} = \eta; \quad \frac{d \text{Im}(\log(b_p(t)))}{dt} = -\omega_0, \quad (\text{C.5})$$

if one is sufficiently careful with the branch-cut of the complex logarithm. Such a linear scaling with  $t$  of  $\text{Re}(\log(b_p(t)))$  and  $\text{Im}(\log(b_p(t)))$  is therefore the appropriate measurement procedure to use in order to estimate the growth rate and pattern speed of the unstable modes of these truncated Mestel discs. The measurements for the various values of the truncation index  $v_t$  are illustrated in Fig. C.3. Once the basis coefficients  $b_p(t)$  have been determined, one can study the shape of the recovered unstable modes in the physical space. Indeed, similarly to Eq. (C.1), the shape of the modes is given by

$$\Sigma_{\text{mode}}(R, \phi, t) = \text{Re} \left[ \sum_p b_p(t) \Sigma^{(p)}(R, \phi) \right]. \quad (\text{C.6})$$

In analogy with Fig. C.2, for which the unstable modes have been obtained via the matrix method, Fig. C.4 illustrates the unstable mode of the same truncated  $v_t = 4$  Mestel disc.



**Fig. C.4.** Unstable mode for the truncated  $v_t = 4$  Mestel disc recovered via direct  $N$ -body simulations as presented in Sect. 5. Only positive contour levels are shown and are spaced linearly between 20% and 80% of the maximum norm. Similarly to Fig. C.2, the radii associated with the resonances ILR, COR and OLR have been represented.

**Table C.1.** Measurements of the pattern speed  $\omega_0 = m_\phi \Omega_p$  and growth rate  $\eta = s$  for unstable  $m_\phi = 2$  modes of tapered Mestel discs.

Unstable $m_\phi = 2$ modes of truncated Mestel discs, $q = 6$ .						
Method	$v_t = 4$		$v_t = 6$		$v_t = 8$	
	$\omega_0$	$\eta$	$\omega_0$	$\eta$	$\omega_0$	$\eta$
Linear theory	0.88	0.13	0.90	0.22	0.92	0.27
Matrix method	0.93	0.11	0.94	0.20	0.95	0.24
$N$ -body	0.99	0.13	0.79	0.19	0.89	0.26

**Notes.** The velocity dispersion within these discs is characterised by  $q = 6$ , and the inner truncation power indices are given by  $v_t = 4, 6, 8$ . The theoretical values were obtained from linear theory in Evans & Read (1998b). Our measurements were either performed via the response matrix method as in Eq. (24), or via direct  $N$ -body simulations, using the  $N$ -body integrator described in Sect. 5.

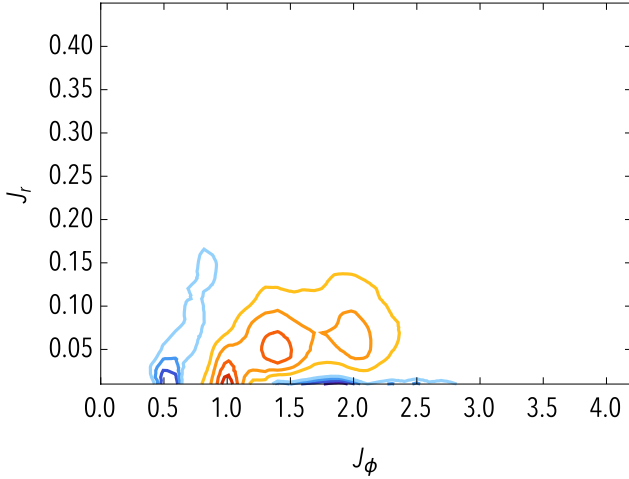
As a conclusion, the growth rates and pattern speeds obtained either via the matrix method or direct  $N$ -body simulations are gathered in Table C.1. As observed in Sellwood & Evans (2001), the recovery of the unstable modes characteristics from direct  $N$ -body simulations when performed for truncated Mestel discs is a difficult task, for which convergence to the values predicted through linear theory may be difficult.

## Appendix D: Why swing matters?

Let us investigate here briefly the importance of self-gravitation and the completeness of the projection basis in capturing the role of swing amplification.

### D.1. Turning off the self-gravitating amplification

In order to investigate the role of the self-gravitating amplification, one may perform the same estimation as presented in Fig. 4, while neglecting collective effects. When neglecting collective



**Fig. D.1.** Map of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}})$ , corresponding to the bare secular diffusion flux, using the same conventions as in Fig. 4. The contours are spaced linearly between the minimum and the maximum of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}})$ . The maximum value for the positive blue contours corresponds to  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}}) \approx 0.30$ , while the minimum value for the negative red contours is associated with  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}}) \approx -0.50$ . This figure is qualitatively similar to the one obtained in Fig. 9 of Paper I.

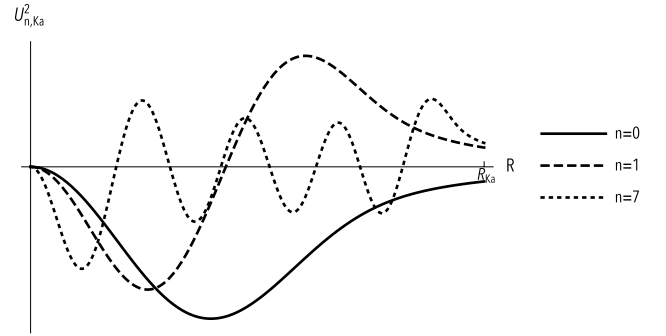
effects, i.e. when assuming that  $\widehat{\mathbf{M}} \equiv 0$ , one recovers the inhomogeneous Landau equation (Chavanis 2013) which reads

$$\frac{\partial F}{\partial t} = \pi(2\pi)^d \mu \frac{\partial}{\partial J_1} \cdot \left[ \sum_{m_1, m_2} m_1 \int dJ_2 |A_{m_1, m_2}(J_1, J_2)|^2 \times \delta_D(m_1 \cdot \Omega_1 - m_2 \cdot \Omega_2) \left( m_1 \cdot \frac{\partial}{\partial J_1} - m_2 \cdot \frac{\partial}{\partial J_2} \right) F(J_1, t) F(J_2, t) \right]. \quad (\text{D.1})$$

Equation (D.1) involves the bare susceptibility coefficients  $|A_{m_1, m_2}(J_1, J_2)|^2$ , which can be equivalently defined (see Appendix B of Paper I) by

$$\begin{aligned} A_{m_1, m_2}(J_1, J_2) &= - \sum_p \psi_{m_1}^{(p)}(J_1) \psi_{m_2}^{(q)*}(J_2) \\ &= \frac{1}{(2\pi)^4} \int d\theta_1 d\theta_2 u(|\mathbf{x}(\theta_1, J_1) - \mathbf{x}(\theta_2, J_2)|) e^{i(m_1 \cdot \theta_1 - m_2 \cdot \theta_2)}, \quad (\text{D.2}) \end{aligned}$$

where  $u(\mathbf{x})$  is the binary potential of interaction given by  $u(\mathbf{x}) = -G/|\mathbf{x}|$  for gravity. This estimation therefore does not require to estimate the response matrix from Eq. (24), but one still has to perform integrations along the resonant lines as in Eq. (42). Let us provide a first numerical implementation of this equation in the context of galactic dynamics. The contours of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}})$  are illustrated in Fig. D.1. Comparing the maps of the dressed diffusion flux  $N\text{div}(\mathcal{F}_{\text{tot}})$  from Fig. 4 and the bare diffusion flux  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{bare}})$ , allows us to assess the strength of the self-gravitating amplification. As expected, when turning off the self-gravity of the system, one reduces significantly the susceptibility of the system and therefore slows down its secular evolution by a factor of about 1000. One may also remark that while the secular appearance of a resonant ridge in the dressed diffusion from Fig. 4 was obvious, the shape of the contours obtained in the bare Fig. D.1 do not emphasise as clearly the appearance of such a narrow resonant ridge. One can still remark that the structure of the bare contours obtained in Fig. D.1 is similar to what was obtained in Fig. 9 of Paper I, through the WKB limit of the Balescu-Lenard equation. One can finally note that the amplitudes of the bare divergence contours obtained previously

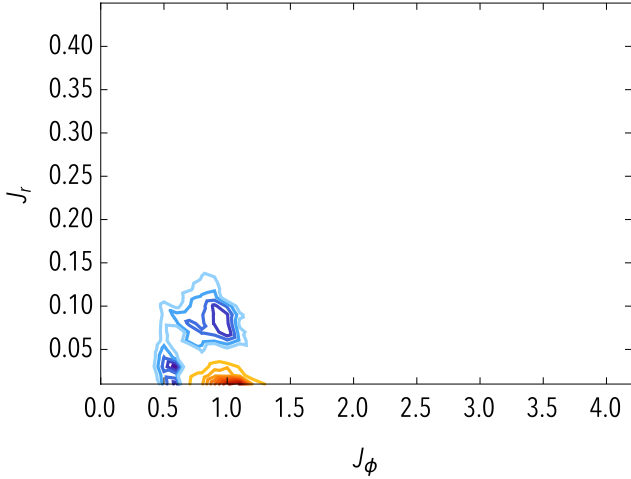


**Fig. D.2.** Illustration of the radial basis elements of the  $\ell = 2$  Kalnajs basis elements for  $k_{\text{Ka}} = 7$ , defined in Appendix A, which were used in the estimation of the Balescu-Lenard diffusion flux in Sect. 4.2. As the radial index  $n$  increases, the basis elements get more and more wound.

are similar to the WKB values obtained in Paper I. As a consequence, the comparison of Figs. 4 and D.1 emphasises that the strong self-gravitating amplification of loosely wound perturbations is indeed responsible for the appearance of a narrow ridge, while also ensuring that this appearance is sufficiently rapid, as observed in the diffusion timescales comparison from Eq. (52).

## D.2. Turning off loosely-wound contributions

As emphasised in the Introduction, the WKB limit of the Balescu-Lenard equation presented in Fouvry et al. (2015a) was not able to capture the mechanism of swing amplification, which involves unwinding perturbations. By considering a complete and global basis as in Eq. (16), we have shown in Fig. 4 how the missing amplification from Fouvry et al. (2015a) could be recovered. Using the numerical method of estimation of the secular diffusion flux as presented in Sect. 3, one can try to recover the results obtained within the WKB formalism by carefully choosing the considered basis elements generically introduced in Eq. (16) and chosen to be Kalnajs basis elements as detailed in Appendix A. We recall that each basis element depends on two indices: an azimuthal index  $\ell$  and a radial one  $n$ . Because in S12's simulation perturbations were restricted to the harmonic sector  $m_\phi = 2$ , one only has to consider basis elements associated with  $\ell = 2$ . Moreover, as illustrated in Fig. D.2, the larger  $n$  the radial index, the faster the radial variation of the basis elements and therefore the more tightly wound the basis elements. So as to get rid of the loosely-wound basis elements which are the ones which can get swing amplified, we perform a truncation of the radial indices considered. Therefore, we define the secular diffusion flux  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$  computed in the same way than  $N\text{div}(\mathcal{F}_{\text{tot}})$  as presented in Sect. 4.2, except that the basis elements are such that  $n_{\text{cut}} \leq n \leq n_{\text{max}}$ , with  $n_{\text{cut}} = 2$  and  $n_{\text{max}} = 8$ . By keeping only the tightly wound basis elements, one can therefore consider the same contribution as the one considered in the WKB limit presented in Paper I. The contours of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$  are illustrated in Fig. D.3. One can note that the values of the contours obtained in the map of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$  illustrated in Fig. D.3 are in the same order of magnitude as the ones which were presented in Fig. 9 of Paper I in the WKB limit. The presence of positive blue contours of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$  is also in agreement with a secular heating of the disc (i.e. an increase of  $J_r$ ). However, these contours do not display a narrow resonant ridge as was observed in S12's simulation or in Fig. 4.



**Fig. D.3.** Map of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$ , corresponding to the dressed secular diffusion flux, using the same conventions as in Fig. 4. In order to limit ourselves only to tightly wound contributions, the basis elements associated with the radial basis index  $n \in \{0, 1\}$  have not been taken into account. The contours are spaced linearly between the minimum and the maximum of  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}})$ . The maximum value for the positive blue contours corresponds to  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}}) \simeq 0.7$ , while the minimum for the negative red contours is associated with  $N\text{div}(\mathcal{F}_{\text{tot}}^{\text{WKB}}) \simeq -4.5$ . This figure is to be compared to Fig. 9 of Paper I.

## Appendix E: Sampling of the DF

In order to use the  $N$ -body integrator described in Sect. 5, one has to sample the particles according to the DF given by Eq. (49). We introduce the probability distribution function  $F_{\text{sp}}$ , normalised to 1 and thanks to which the sampling is performed. This probability DF  $F_{\text{sp}}$  is directly proportional to the active distribution function  $F_{\text{star}}$  from Eq. (49), so that we may write

$$F_{\text{sp}}(E, L) = C_{\text{sp}} L^q \exp[-E/\sigma_r^2] T_{\text{inner}}(L) T_{\text{outer}}(L), \quad (\text{E.1})$$

where  $C_{\text{sp}}$  is a normalisation constant which is determined in the upcoming calculations. Because the mapping  $(E, L) \mapsto (J_r, J_\phi)$  from Eq. (12) is not a trivial one, we do not perform the sampling of the stars in the action space  $(J_r, J_\phi)$ , but rather in the  $(E, L)$ -space. Moreover, one should pay attention to the fact that the DF  $F_{\text{sp}}$  from Eq. (E.1) is a probability distribution function in the  $(\mathbf{x}, \mathbf{v})$ -space, so that  $d^2\mathbf{x}d^2\mathbf{v}F_{\text{sp}}(\mathbf{x}, \mathbf{v})$  is proportional to the number of particles in the infinitesimal volume  $d^2\mathbf{x}d^2\mathbf{v}$  around the position  $(\mathbf{x}, \mathbf{v})$ . As we want to sample the particles in the  $(E, L)$ -space, we introduce the function  $h_{\text{sp}}(E, L)$  such that  $dEdLh_{\text{sp}}(E, L)$  is proportional to the number of particles in the volume  $dEdL$  around the location  $(E, L)$ . One can now determine  $h_{\text{sp}}(E, L)$  as a function of  $F(E, L)$ . Indeed, we have

$$\begin{aligned} h_{\text{sp}}(E', L') &= \int d\mathbf{x}d\mathbf{v} \delta_D(E' - E) \delta_D(L' - L) F_{\text{sp}}(E, L) \\ &= 2\pi \int dr r \int dv_r dv_t \delta_D(E' - E) \delta_D(L' - L) F_{\text{sp}}(E, L) \\ &= 2\pi \int dr \int dv_r \delta_D(E' - E) F_{\text{sp}}(E, L'), \end{aligned} \quad (\text{E.2})$$

using the fact that the tangential velocity satisfies  $v_t = L/r$ . The last step is then to perform the change of variable  $v_r \rightarrow E$ . One has  $v_r^2 = 2(E - \psi_M(r)) - L^2/r^2$ , so that

$dv_r = dE / \sqrt{2(E - \psi_M(r)) - L^2/r^2}$ . Because the radial velocity can be both positive and negative, Eq. (E.2) takes the form

$$\begin{aligned} h_{\text{sp}}(E', L') &= 4\pi \int dr \int dE \frac{\delta_D(E' - E) F_{\text{sp}}(E, L')}{\sqrt{2(E - \psi_M(r)) - L'^2/r^2}} \\ &= \frac{4\pi^2}{\Omega_1(E', L')} F_{\text{sp}}(E', L'), \end{aligned} \quad (\text{E.3})$$

where we used the definition (13) of the radial intrinsic frequency  $\Omega_1$ . One can then correctly normalise the probability distribution  $h_{\text{sp}}$  and determine the value of the constant  $C_{\text{sp}}$  from Eq. (E.1). One should pay attention to the fact that in addition to the tapering functions  $T_{\text{inner}}$  and  $T_{\text{outer}}$  from Eqs. (48), we also assume that no stars have orbits that extend beyond  $R_{\text{max}}$ . As a consequence, the allowed region in the  $(E, L)$ -space has to satisfy two constraints. First of all, the angular momentum  $L_{\text{star}}$  has to satisfy

$$L_{\text{min}} = 0 \leq L_{\text{star}} \leq R_{\text{max}} V_0 = L_{\text{max}}. \quad (\text{E.4})$$

Then, for a given value of  $L_{\text{star}}$ , one can show that the energy of the star  $E_{\text{star}}$  must satisfy the constraint

$$E_{\text{min}}(L_{\text{star}}) = \psi_M\left[\frac{L_{\text{star}}}{V_0}\right] + \frac{V_0^2}{2} \leq E_{\text{star}} \leq \frac{L_{\text{star}}^2}{2R_{\text{max}}^2} = E_{\text{max}}(L_{\text{star}}). \quad (\text{E.5})$$

These two constraints allow us to completely characterise the  $(E, L)$ -space on which the sampling  $(E, L)$  has to be performed. One finally has to satisfy the constraint

$$\int_{L_{\text{min}}}^{L_{\text{star}}} dL \int_{E_{\text{min}}(L)}^{E_{\text{max}}(L)} dE h_{\text{sp}}(E, L) = 1. \quad (\text{E.6})$$

Given the parameters presented after Eq. (49), one can numerically determine the value of the constant  $C_{\text{sp}}$  which reads

$$C_{\text{sp}} \simeq 1.4723 \times 10^{-15}. \quad (\text{E.7})$$

We may now proceed to the sampling of the coordinates of the particles. Up to the sign of its radial velocity, one star is characterised by the set  $\{E_{\text{star}}, L_{\text{star}}, R_{\text{star}}, \phi_{\text{star}}\}$ . Given that the initial state is axisymmetric, the azimuthal angle of the star can be uniformly sampled between 0 and  $2\pi$ . The next step is then to successively sample  $(L_{\text{star}}, E_{\text{star}})$  and finally  $R_{\text{star}}$ , using successive rejection samplings as we now detail.

The heart of the rejection sampling is as follows. Let us assume that we want to generate sampling values from a function  $f(x)$ , from which it is difficult to sample. However, we assume that we have at our disposal another distribution function  $g(x)$  from which the sampling is simple, and such that there exists a bound  $M > 1$  satisfying  $f(x) < M g(x)$ . The smaller  $M$ , the more efficient the sampling. One then has to proceed as follows: sample both a proposal  $x$  from  $g$  and  $u$  uniformly between  $[0; 1]$ . One then applies the selection

$$\alpha = \frac{f(x)}{Mg(x)}: \begin{cases} u < \alpha \implies x \text{ is kept.} \\ u \geq \alpha \implies x \text{ is rejected.} \end{cases} \quad (\text{E.8})$$

In order to have an efficient sampling, one should try to consider a function  $g$  close to  $f$ .

We may now directly sample  $(E, L)$  thanks to this algorithm. The true sampling function is  $f_{(E,L)} = h_{\text{sp}}$  from Eq. (E.2). The simple sampling function is  $g_{(E,L)} \propto 1$ , defined on the domain characterised by the constraints from Eqs. (E.4) and (E.5). When

performing a rejection sampling with such a uniform  $g_{(E,L)}$ , in order to determine the bound  $M_{(E,L)}$ , one only has to determine a uniform bound for  $f_{(E,L)}$ . With the numerical values introduced after Eq. (49), one can check that  $f_{(E,L)}$  is such that

$$f_{(E,L)}(E, L) \leq 1.4. \quad (\text{E.9})$$

The final element required to be able to perform the rejection sampling with  $f_{(E,L)}$  is to be able to draw uniformly candidate  $(E, L)$  in the domains defined by the constraints from Eqs. (E.4) and (E.5), which is equivalent as sampling candidates  $(E, L)$  from the uniform probability distribution function  $g_{(E,L)}$ . To perform this uniform sampling, since the constraints from Eq. (E.5) are expressed for a given value of  $L_{\text{star}}$ , it is more natural to first draw  $L_{\text{star}}$  and then  $E_{\text{star}}$ . The probability distribution according to which  $L_{\text{star}}$  has to be drawn is of the form  $f_L(L) \propto (E_{\text{max}}(L) - E_{\text{min}}(L))$ . When correctly normalised, it reads

$$f_L(L) = \frac{3}{2} \frac{1}{R_{\text{max}} V_0} \left[ \frac{L^2}{2R_{\text{max}}^2 V_0^2} - \frac{1}{2} - \log \left[ \frac{L}{R_{\text{max}} V_0} \right] \right]. \quad (\text{E.10})$$

To sample  $L$  from  $f_L$ , we use another rejection sampling by introducing the additional simple probability distribution function  $g_L$  defined as

$$g_L(L) = -\frac{1}{R_{\text{max}} V_0} \log \left[ \frac{L}{R_{\text{max}} V_0} \right]. \quad (\text{E.11})$$

It is straightforward to check that  $f_L < (3/2) g_L$ , so that we may use the bound  $M_L = 3/2$  to perform the rejection sampling of  $L_{\text{star}}$ . The final remark is to note that sampling  $L$  from  $g_L$  is simple since its cumulative distribution function  $G_L = \int_{L_{\text{min}}}^L dL' g_L(L')$  can be inverted so as to read

$$G_L^{-1}(u) = R_{\text{max}} V_0 \exp \left[ 1 + W_{-1} \left( -\frac{u}{e} \right) \right], \quad (\text{E.12})$$

where  $W_{-1}$  is the lower branch of the Lambert function  $W(x)$  for  $x \in [-1/e; 0]$ . With all these elements, the rejection sampling of  $L_{\text{star}}$  following  $f_L$  from Eq. (E.10) can be performed.

Once  $L_{\text{star}}$  has been drawn, it only remains to sample uniformly  $E_{\text{star}}$  on the interval  $E_{\text{star}} \in [E_{\text{min}}(L_{\text{star}}); E_{\text{max}}(L_{\text{star}})]$ , as given by Eq. (E.4). Thanks to these uniformly drawn candidates  $(E_{\text{star}}, L_{\text{star}})$  and the uniform bound from Eq. (E.9), one can perform the rejection sampling from the probability distribution  $f_{(E,L)}$ .

For  $E_{\text{star}}$  and  $L_{\text{star}}$  successfully sampled, one may then sample the radius  $R_{\text{star}}$  using a similar rejection sampling. The radius has to be sampled according to the probability distribution  $f_R$  given by

$$f_R(r) = \frac{\Omega_1/\pi}{\sqrt{2(E - \psi_M(r)) - L^2/r^2}}, \quad (\text{E.13})$$

so that one has  $f_R \propto v_r$ . However, one should note that for  $r \rightarrow r_{p/a}$ , one has  $f_R(r) \rightarrow +\infty$ , so that the rejection sampling cannot be used without considering a probability DF  $g$  which also diverges for  $r \rightarrow r_{p/a}$ . In order to get rid of these divergences, instead of sampling the variable  $r$ , we sample the angle  $u \in [-\pi/2; \pi/2]$ , where we defined the mapping  $r \mapsto u(r)$  as

$$r(u) = \frac{r_p + r_a}{2} + \frac{r_a - r_p}{2} \sin(u), \quad (\text{E.14})$$

so that one naturally has  $r(-\pi/2) = r_p$  and  $r(\pi/2) = r_a$ . The probability distribution function from which  $u$  has to be sampled is immediately given by

$$f_u(u) = \frac{r_a - r_p}{2} \cos(u) p_r(r(u)). \quad (\text{E.15})$$

Using the fact that the maximum of  $f_u$  is reached for  $u = \pi/2$ , one can then sample  $u$  from  $f_u$  using a rejection sampling with a uniform control probability distribution function  $g_u(u) = 1/\pi$ . Once  $u$  is known, it only remains to compute  $R_{\text{star}} = r(u_{\text{star}})$ , so that the sampling of all the required quantities for one star has been performed.

The final step of the sampling of the particles is to determine the physical coordinates of the particles  $(\mathbf{x}, \mathbf{v})$  associated with the set  $\{E_{\text{star}}, L_{\text{star}}, R_{\text{star}}, \phi_{\text{star}}\}$ . These physical coordinates are the ones which are given to the  $N$ -body integrator. We draw uniformly the sign of the radial velocity  $\varepsilon_r \in \{-1, 1\}$ . Because we are considering a disc made only of prograde stars, one immediately obtains that the radial and tangential velocities  $v_r$  and  $v_t$  are given by

$$\begin{cases} v_r = \varepsilon_r \sqrt{2(E - \psi_M(R_{\text{star}})) - L_{\text{star}}^2/R_{\text{star}}^2}, \\ v_t = L_{\text{star}}/R_{\text{star}}. \end{cases} \quad (\text{E.16})$$

The final step of the transformation to the  $(\mathbf{x}, \mathbf{v})$ -coordinates is then straightforward, since one naturally has

$$\begin{cases} x = R_{\text{star}} \cos(\phi_{\text{star}}), \\ y = R_{\text{star}} \sin(\phi_{\text{star}}), \\ v_x = v_r \cos(\phi_{\text{star}}) - v_t \sin(\phi_{\text{star}}), \\ v_y = v_r \sin(\phi_{\text{star}}) + v_t \cos(\phi_{\text{star}}). \end{cases} \quad (\text{E.17})$$

One should note that the sampling procedure described previously does not correspond to a quiet start procedure (Sellwood 1983), which would lead to a reduction of the initial shot noise within the disc, as briefly discussed in Appendix C, with regard to the validation of the  $N$ -body code.

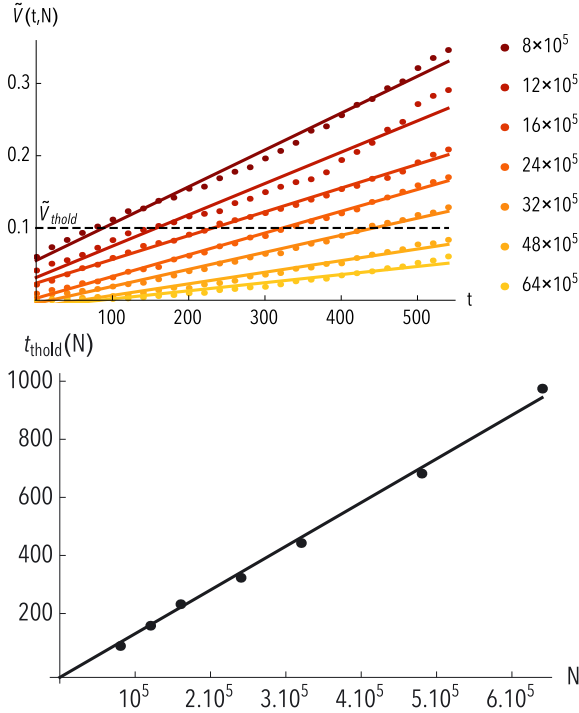
## Appendix F: Another test of the scaling with $N$

One difficulty with the measurement of the scaling with  $N$  presented in Sect. 5.2 is that one has to disentangle the contributions from the initial sampling Poisson shot noise present through  $h_0$  from Eq. (59) and the effects due to the collisional Balescu-Lenard diffusion scaling through  $h_2$  from Eq. (64). Indeed, Poisson shot noise leads to fluctuations of the system DF about its mean value. In order not to be sensitive to such fluctuations, one could only consider fluctuations sufficiently large, i.e. fluctuations caused by an effective secular diffusion rather than caused by inevitable Poisson fluctuations. As a consequence, by restricting ourselves only to large fluctuations, we can get rid of Poisson's effects. We therefore define the function  $\hat{V}(t, N)$  as

$$\hat{V}(t, N) = \int d\mathbf{J} \chi [ \langle F(t, \mathbf{J}, N) \rangle - \langle F(t=0, \mathbf{J}, N) \rangle < C_{\hat{V}} ], \quad (\text{F.1})$$

where we introduced a threshold  $C_{\hat{V}} < 0$ . Here  $\chi[\cdot]$  is a characteristic function equal to 1 if  $(\langle F(t, \mathbf{J}, N) \rangle - \langle F(t=0, \mathbf{J}, N) \rangle < C_{\hat{V}})$ , and 0 otherwise. As a consequence,  $\hat{V}(t, N)$  measures the volume in action space of the regions (depleted from particles, since  $C_{\hat{V}} < 0$ ) for which the mean DF has changed by more than  $C_{\hat{V}}$ . For a sufficiently large value of the threshold  $C_{\hat{V}}$ , such a construction allows us not to be polluted by Poisson sampling shot noise. For the initial





**Fig. F.1.** *Top panel:* illustration of the behaviour of the function  $t \mapsto \tilde{V}(t, N)$  from Eq. (F.1), when averaged on 32 different realisations for particles numbers  $N \in \{8, 12, 16, 24, 32, 48, 64\} \times 10^5$ , along with the associated linear fits. To compute  $\tilde{V}(t, N)$ , we used the same binning of the action-space  $(J_\phi, J_r)$  as in Fig. 5. As obtained in Eq. (F.3), one recovers that for a fixed value of  $N$ , the function  $t \mapsto \tilde{V}(t, N)$  is linear. The horizontal dashed line illustrates the threshold value  $\tilde{V}_{\text{thold}}$  for which the threshold time  $t_{\text{thold}}$  is determined. *Bottom panel:* illustration of the behaviour of the function  $N \mapsto t_{\text{thold}}(N)$ . As derived in Eq. (F.5), one recovers a linear dependence of  $t_{\text{thold}}$  with  $N$ .

times, as in Eq. (50), it is straightforward to study the scaling of  $\tilde{V}(t, N)$  with  $t$  and  $N$ . Indeed, one can write

$$\langle F(t, \mathbf{J}, N) \rangle - \langle F(t=0, \mathbf{J}, N) \rangle \simeq \Delta\tau \operatorname{div}(\mathcal{F}_{\text{tot}}) \simeq \frac{t}{N} \operatorname{div}(\mathcal{F}_{\text{tot}}). \quad (\text{F.2})$$

Introducing  $\tilde{V}_0 = \int d\mathbf{J} \chi [\operatorname{div}(\mathcal{F}_{\text{tot}}) < C_{\tilde{V}}]$ , one can rewrite Eq. (F.2) as

$$\tilde{V}(t, N) \simeq \frac{t}{N} \tilde{V}_0. \quad (\text{F.3})$$

Therefore, for a fixed value of  $N$ , one expects to observe a linear time dependence of the function  $\tilde{V}(t, N)$ , as illustrated in Fig. F.1. In order to test the scaling of Eq. (F.3) with  $N$ , one may proceed as follows. Introducing a threshold value  $\tilde{V}_{\text{thold}}$ , for each value of  $N$ , we define the associated threshold time  $t_{\text{thold}}(N)$  as

$$\tilde{V}(t_{\text{thold}}(N), N) = \tilde{V}_{\text{thold}}. \quad (\text{F.4})$$

Thanks to the scalings from Eq. (F.3), one immediately obtains that

$$t_{\text{thold}}(N) \simeq N \frac{\tilde{V}_{\text{thold}}}{\tilde{V}_0}. \quad (\text{F.5})$$

Such a linear scaling of  $t_{\text{thold}}(N)$  with  $N$  is a prediction from the Balescu-Lenard formalism and is nicely recovered in Fig. F.1.

## Appendix G: Distributed code description

For the sake of reproducibility, which has been lacking in the context of the linear response of stellar systems, we distribute the linear matrix response code we wrote for this paper both as a MATHEMATICA package<sup>6</sup>, and a notebook<sup>7</sup>. The functions therein allow for:

- The determination as a function of  $(r_p, r_a)$  of the orbits quantities:  $E, L, J_r, \Omega_1$  and  $\Omega_2$ .
- The construction of the 2D basis from Kalnajs (1976).
- The computation of the Fourier transform w.r.t. the angles, i.e. the computation of  $\mathcal{W}_{\ell p m_2 m^p}^{m_1}(\mathbf{J})$  from Eq. (21).
- The calculation of the 2D response matrix via Eq. (33).

It has been tested for the isochrone and the Mestel discs. A copy of the code is also available at the CDS.

<sup>6</sup> <http://www.iap.fr/users/pichon/MatrixMethod/code/MatrixMethod2D.m>

<sup>7</sup> <http://www.iap.fr/users/pichon/MatrixMethod/code/MatrixMethod2D.nb>