# Orbital alignment and mass segregation in galactic nuclei via vector resonant relaxation

Nathan Magnan, Jean-Baptiste Fouvry, Christophe Pichon, Pierre-Henri Chavanis

MNRAS **000**, 000–000 (0000) Preprint 5 July 2022 Compiled using MNRAS LATEX style file v3.0

# Orbital alignment and mass segregation in galactic nuclei via vector resonant relaxation

Nathan Magnan[1,2⋆], Jean-Baptiste Fouvry[1], Christophe Pichon[1,3] and Pierre-Henri Chavanis[4]

[1] *CNRS and Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, F-75014 Paris, France*
[2] *DAMTP, Centre for Mathematical Sciences, Wilberforce Rd, Cambridge CB3 0WA, United Kingdom*
[3] *Institut de Physique Théorique, DRF-INP, UMR 3680, CEA, Orme des Merisiers Bât 774, 91191 Gif-sur-Yvette, France*
[4] *Laboratoire de Physique Théorique, Université de Toulouse, CNRS, UPS, France*



5 July 2022

**ABSTRACT**

Supermassive black holes dominate the gravitational potential in galactic nuclei. In these dense environments, stars follow nearly Keplerian orbits and see their orbital planes relax through the potential fluctuations generated by the stellar cluster itself. For typical astrophysical galactic nuclei, the most likely outcome of this vector resonant relaxation (VRR) is that the orbital planes of the most massive stars spontaneously self-align within a narrow disc. We present a maximum entropy method to systematically determine this long-term distribution of orientations and use it for a wide range of stellar orbital parameters and initial conditions. The heaviest stellar objects are found to live within a thin equatorial disk. The thickness of this disk depends on the stars' initial mass function, and on the geometry of the initial cluster. This work highlights a possible (indirect) novel method to constrain the distribution of intermediate mass black holes in galactic nuclei.

**Key words:** Diffusion - Gravitation - Galaxies: kinematics and dynamics - Galaxies: nuclei

## 1 INTRODUCTION

Supermassive black holes (BHs) are ubiquitous in external galaxies (Magorrian et al. 1998; Genzel et al. 2010; Kormendy & Ho 2013) and their induced feedback plays a critical role in regulating galaxy formation through cosmic ages (Heckman & Best 2014). The unique proximity of our Galactic centre is an extraordinary opportunity to study and constrain the long-term evolution of galactic nuclei and the stellar clusters orbiting within. Recent developments in that realm include detailed census of stellar populations around SgrA* (Ghez et al. 2008; Gillessen et al. 2017), the observation of a cool accretion disc (Murchikova et al. 2019), as well as the observation of the relativistic precession of S2 (Gravity Collab. et al. 2020). Similarly, the origin of the observed clockwise stellar disc (Bartko et al. 2009; Yelda et al. 2014) has triggered a lot of interest, as its existence may impact the merger rate of the intermediate mass black holes (IMBHs) population (Portegies Zwart & McMillan 2002).

As already pointed in Rauch & Tremaine (1996), the steep potential well generated by the central BH allows for efficient secular orbit-averaged interactions between stars, driving their relaxation through an intricate hierarchy of dynamical processes (Alexander 2017). In this paper, we focus on the process of vector resonant relaxation (VRR) during which stars undergo a random reshuffling of their orbital orientations through long-term coherent torques between the finite number of stellar orbital planes (Kocsis & Tremaine 2015). Given that VRR occurs on a timescale shorter than the stellar ages (see fig. 1 in Kocsis & Tremaine 2011), one may expect that their observed distribution of orientations corresponds to some statistical equilibrium.

Determining the outcome of this long-term rearrangement has been the focus of recent efforts (Roupas et al. 2017; Takács & Kocsis 2018; Szölgyén & Kocsis 2018; Touma et al. 2019; Tremaine 2020a,b; Gruzinov et al. 2020) that jointly offer new clues on the fascinating properties of these long-range interacting systems, such as negative temperatures or phase transitions. In particular, Szölgyén & Kocsis (2018), using an ingenious Monte–Carlo approach, have shown that in systems with a wide range of stellar populations (i.e. various masses and semi-major axes), VRR can lead to the spontaneous formation of a disc through the angular segregation of the most massive stars and IMBHs.

This is the issue that we further investigate in this work. We develop and implement an explicit and efficient optimisation procedure to find maximum entropy solutions compatible with an initial configuration, so as to infer the thermodynamical equilibria of a given galactic nucleus. In the particular context of VRR, this roadmap was already started out by Roupas et al. (2017) in the limit of a single-population system with a quadrupolar interaction, and later improved in Takács & Kocsis (2018) which, while still re-

⋆ nathan.magnan@maths.cam.ac.uk

stricting themselves to a single-population system, considered harmonic expansion beyond the quadrupole. Here, we expand these works to multi-population clusters by emphasising the critical role played by the diversity of stellar orbits and masses to allow for non-trivial anisotropic statistical equilibria, as unveiled in Szölgyén & Kocsis (2018).

The paper is organised as follows. In §2, we briefly review the process of VRR in galactic nuclei. We also present the numerical optimisation method used to determine the thermodynamical equilibria. In §3, we use this approach to carefully investigate the details of equilibrium configurations that appear in systems composed of multiple stellar populations, as well as the impact of the initial mass function and the geometry of the formation scenario. Finally, we conclude in §4.

## 2 VRR MODEL

Let us first set up the framework in which to quantify the long-term effect of VRR on a population of orbital planes.

### 2.1 Interaction of Keplerian annuli via VRR

We consider an isolated cluster of $N \gg 1$ stars orbiting a supermassive BH of mass $M_\bullet$. Following a double orbit-average over all the stars' Keplerian motion and in-plane precession, the VRR Hamiltonian takes the form[1] (Kocsis & Tremaine 2015)

$$H_{\rm tot} = -\sum_{i<j}^{N} \left\langle \frac{G m_i m_j}{|\mathbf{r}_i(t) - \mathbf{r}_j(t')|} \right\rangle_{t,t'}, \tag{1}$$

where the sum over $(i,j)$ runs over all the pairs of particles, and the double orbit-average $\langle \cdot \rangle_{t,t'}$ operates over the fast Keplerian motions and in-plane precessions of both particles, with $\mathbf{r}_i(t)$ and $\mathbf{r}_j(t')$ describing their trajectories.

Using the Legendre expansion of the Newtonian interaction, and the addition theorem for spherical harmonics, one can rewrite equation (1) as (see, e.g., Fouvry et al. 2019b)

$$H_{\rm tot} = -\sum_{i<j}^{N} \sum_{\ell=2}^{\ell_{\rm max}} \sum_{m=-\ell}^{\ell} \mathcal{H}_\ell\big[\mathbf{K}_i, \mathbf{K}_j\big]\, Y_{\ell m}(\widehat{\mathbf{L}}_i)\, Y_{\ell m}(\widehat{\mathbf{L}}_j), \tag{2}$$

where $\widehat{\mathbf{L}}_i$ stands for the unit vector aligned with the instantaneous orientation of the star's orbital plane and the real spherical harmonics, $Y_{\ell m}(\widehat{\mathbf{L}})$, are normalised so that $\int \mathrm{d}\widehat{\mathbf{L}}\, Y_{\ell m} Y_{\ell' m'} = \delta_{\ell\ell'}\delta_{mm'}$. In equation (2), a Keplerian annulus (see Fig. 1) is fully characterised by its conserved quantities

$$\mathbf{K} = (m, a, e), \tag{3}$$

with $m$ the star's individual mass, $a$ its semi-major axis, and $e$ its eccentricity. We stress that, owing to orbit average, each $\mathbf{K}_i$ are taken to be fixed throughout the VRR dynamics. Finally, in equation (2), we introduced the (symmetric) coupling coefficient

$$\mathcal{H}_\ell\big[\mathbf{K}_i, \mathbf{K}_j\big] = G m_i m_j \, \frac{4\pi}{2\ell+1} \big|P_\ell(0)\big|^2 \times \int_0^\pi \frac{\mathrm{d}M_i}{\pi} \int_0^\pi \frac{\mathrm{d}M_j}{\pi} \, \frac{\mathrm{Min}[r_i, r_j]^\ell}{\mathrm{Max}[r_i, r_j]^{\ell+1}}, \tag{4}$$

with $P_\ell$ the Legendre polynomial of order $\ell$ and $M_i$ the mean anomaly of the orbit $\mathbf{K}_i$ (see, e.g., Murray & Dermott 1999).

[1] We also assume that $\sum_{i=1}^N m_i \ll M_\bullet$ (Roupas 2020).

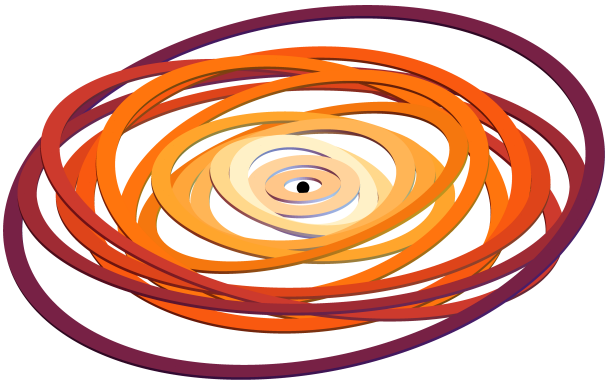

**Figure 1.** Illustration of the process of VRR in a galactic nucleus. Following an orbit-average over the fast Keplerian motion induced by the central BH and the in-plane precessions (due to both the stellar mean potential and the relativistic corrections), stars are replaced by massive annuli (here quasi-circular) which can torque one another (see also fig. 1 in Giral Martínez et al. 2020). This leads ultimately to the relaxation of the stellar orbital orientations with possibly the spontaneous formation of aligned discs.

These coupling coefficients fully encode the impact of the shape of the two orbits on their relative torque and show significant diversity (see, e.g., fig. 1 in Kocsis & Tremaine 2015)[2].

Finally, we point out a few specificities of the multipole expansion from Eq. (2): (i) All odd $\ell$ harmonics have vanishing coupling coefficients, $\mathcal{H}_\ell = 0$, so that they can be dropped; (ii) The Legendre expansion is truncated to the maximum order $\ell \leqslant \ell_{\rm max}$; (iii) We do not account for the harmonics $\ell = 0$ as it does not drive any dynamics. With such a convention, spherically symmetric distributions have a vanishing mean total energy.

Describing the process of VRR amounts to describing the long-term evolution of $\widehat{\mathbf{L}}_i$, as driven by the Hamiltonian from equation (2) (see §2.2 in Kocsis & Tremaine 2015, for the associated equations of motion). We illustrate such a system in Fig. 1. One can equivalently describe the instantaneous state of the stellar cluster with its discrete distribution function (DF)

$$F_{\rm d}(\widehat{\mathbf{L}}, \mathbf{K}, t) = \sum_{i=1}^{N} \delta_{\rm D}(\widehat{\mathbf{L}} - \widehat{\mathbf{L}}_i(t))\, \delta_{\rm D}(\mathbf{K} - \mathbf{K}_i), \tag{5}$$

which follows the normalisation $\int \mathrm{d}\widehat{\mathbf{L}} \mathrm{d}\mathbf{K} F_{\rm d} = N$. Assuming that particles with the same orbital parameters are indistinguishable, this DF entirely describes the system's state. Therefore, characterising the VRR dynamics amounts to characterising $\partial F_{\rm d}/\partial t$, using the Klimontovitch equation (Klimontovich 1967).

### 2.2 Equilibrium configurations

Fortunately, the VRR dynamics ($\sim 1\,$Myr for S2 around SgrA*, see fig. 1 in Kocsis & Tremaine (2011)) is rapid compared to the cluster's age ($\sim 10\,$Myr for the S-cluster, see Habibi et al. (2017)). As such, if one is interested in sufficiently long timescales, rather

[2] Following §A of Fouvry et al. (2019b), the present coupling coefficients simply read $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'] = \mathcal{J}_\ell[\mathbf{K}, \mathbf{K}']\, L(\mathbf{K})$, with $L(\mathbf{K})$ introduced in equation (12). In practice, following the notations from §A of Fouvry et al. (2019b), the coupling coefficients were pre-computed once, up to $\ell_{\rm max} = 50$ and on a $200 \times 100 \times 100$ linear grid in $(\ln(a_{\rm in}/a_{\rm out}), e_{\rm in}, e_{\rm out})$ with $10^{-2} \leqslant a_{\rm in}/a_{\rm out} \leqslant 1$ and $0 \leqslant e_{\rm in}, e_{\rm out} \leqslant 0.99$, to be subsequently interpolated.

than describing the details of $\partial F_{\rm d}/\partial t$, one may solely focus on characterising the expected equilibrium configurations reached at late times. This is given by

$$F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}) = \lim_{t\rightarrow+\infty} \big\langle F_{\rm d}(\widehat{\mathbf{L}},\mathbf{K},t)\big\rangle, \tag{6}$$

with $\langle\,\cdot\,\rangle$ standing for an ensemble average over independent realisations of the system – see §2.3 for the appropriate handling of this ensemble average. Efficiently predicting these long-term equilibrium DFs is the focus of the present work.

In such a late-time limit, the only information retained by the cluster are its invariants. These are:

- The number density of stars with orbital parameters $\mathbf{K}$

$$N(\mathbf{K}) = \int \mathrm{d}\widehat{\mathbf{L}}\, F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}). \tag{7}$$

- The total energy

$$E_{\rm tot} = \tfrac{1}{2}\int \mathrm{d}\widehat{\mathbf{L}}\mathrm{d}\mathbf{K}\, F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K})\, \varepsilon(\widehat{\mathbf{L}},\mathbf{K}), \tag{8}$$

where $\varepsilon(\widehat{\mathbf{L}},\mathbf{K})$ stands for the energy of a particle of orientation $\widehat{\mathbf{L}}$ and orbital parameter $\mathbf{K}$ as

$$\varepsilon(\widehat{\mathbf{L}},\mathbf{K}) = -\sum_{\ell=2}^{\ell_{\rm max}} \sum_{m=-\ell}^{\ell} \int \mathrm{d}\mathbf{K}'\, \mathcal{H}_\ell\big[\mathbf{K},\mathbf{K}'\big] M_{\ell m}(\mathbf{K}') Y_{\ell m}(\widehat{\mathbf{L}}), \tag{9}$$

with the magnetisations

$$M_{\ell m}(\mathbf{K}) = \int \mathrm{d}\widehat{\mathbf{L}}\, Y_{\ell m}(\widehat{\mathbf{L}})\, F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}) \tag{10}$$

defined as spherical harmonic moments of the DF.

- The total angular momentum

$$\mathbf{L}_{\rm tot} = \int \mathrm{d}\widehat{\mathbf{L}}\, \mathrm{d}\mathbf{K}\, L(\mathbf{K})\, \widehat{\mathbf{L}}\, F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}), \tag{11}$$

with the norm of the angular momentum vector

$$L(\mathbf{K}) = m\sqrt{G M_\bullet\, a\,(1-e^2)}. \tag{12}$$

As usually carried out in the microcanonical ensemble (i.e. for an isolated cluster), for a given set of invariants $\{N(\mathbf{K}), E_{\rm tot}, \mathbf{L}_{\rm tot}\}$, the admissible equilibrium configurations are obtained by maximising the Boltzmann entropy

$$S = -k_{\rm B}\int \mathrm{d}\widehat{\mathbf{L}}\mathrm{d}\mathbf{K}\, F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K})\, \ln\big[F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K})\big], \tag{13}$$

with $k_{\rm B}$ the Boltzmann constant, under the previous conservation constraints. Such a maximisation generically yields (see §A)

$$F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}) = N(\mathbf{K})\, \frac{\mathrm{e}^{-\beta\,\varepsilon(\widehat{\mathbf{L}},\mathbf{K}) + L(\mathbf{K})\,\boldsymbol{\gamma}\cdot\widehat{\mathbf{L}}}}{\displaystyle\int \mathrm{d}\widehat{\mathbf{L}}'\, \mathrm{e}^{-\beta\,\varepsilon(\widehat{\mathbf{L}}',\mathbf{K}) + L(\mathbf{K})\,\boldsymbol{\gamma}\cdot\widehat{\mathbf{L}}'}}, \tag{14}$$

where $\beta$ and $\boldsymbol{\gamma}$ are the Lagrange multipliers respectively associated with the total energy and angular momentum conservation. We note that equation (14) is very similar to its co-planar counterpart found in the HMF model (Antoni & Ruffo 1995; Chavanis et al. 2005) to capture bar formation (Pichon & Lynden-Bell 1993) as a phase transition towards orbit alignment.

Gruzinov et al. (2020) recently used a similar mean-field approximation and maximum entropy method to determine the thermodynamical equilibria of black hole star clusters. It is therefore no surprise that their equation (3) is so similar to the present equation (14). The main difference is that Gruzinov et al. (2020) considered the case of massive Keplerian elliptic wires, while we consider the case of massive Keplerian annuli. Phrased differently, Gruzinov et al. (2020) investigated the simultaneous equilibria of scalar resonant relaxation (SRR) and VRR – i.e. the joint relaxation of eccentricities and orientations – whereas we focus here on the equilibria of VRR – i.e. the sole relaxation of orientations.

### 2.3 Axisymmetric assumption

For a given nuclear cluster, the total angular momentum vector $\mathbf{L}_{\rm tot}$ provides us with only one specific direction, taken to be the $+z$ axis throughout the paper. Unfortunately, this does not imply that the thermodynamical equilibrium of VRR necessarily has an axial symmetry around $\mathbf{L}_{\rm tot}$, as a spontaneous symmetry breaking could occur (see, e.g., Kocsis & Tremaine 2011; Gruzinov et al. 2020).

For simplicity however, we assume that VRR does not exhibit here any such symmetry breaking, and restrict ourselves to axisymmetric DFs, i.e.

$$\forall\,\mathbf{K},\ \forall\,\ell,\ \forall\, m\neq 0,\ M_{\ell m}(\mathbf{K}) = 0. \tag{15}$$

From there, we can make the simplifications $\mathbf{L}_{\rm tot}\rightarrow L_{\rm tot}>0$ and $\boldsymbol{\gamma}\rightarrow\gamma>0$. This assumption greatly reduces the total number of spherical harmonics to consider, therefore it significantly alleviates the numerical complexity. However, let us stress that this assumption is not always physically motivated. It is still legitimate in some cases, as for single-population and single-harmonic clusters (Roupas et al. 2017) or multi-population clusters with null inverse temperatures (§C3), but it does not hold in some other regimes, as emphasised by the finding of a warped VRR disc in the simulations of Kocsis & Tremaine (2011) (see fig. 6 therein). Overall, the axisymmetric assumption is an important limitation of the present work, and needs to be challenged in future studies.

To comply with this approach, the ensemble average of equation (6) is carried out over realisations which are all rotated to have their $\mathbf{L}_{\rm tot}$ aligned along $+z$.

### 2.4 Self-consistency

Of course, one needs to impose self-consistency on equations (8) and (11), as well as on equation (14). Indeed, $F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K})$ involves the one-particle energy $\varepsilon(\widehat{\mathbf{L}},\mathbf{K})$ which, via equations (9) and (10), involves $F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K})$ itself. Within the present microcanonical ensemble, imposing the cluster's total energy, angular momentum and orbital distribution ultimately sets up its temperature (via $\beta$), rate of rotation (via $\gamma$), and shape (via $M_{\ell 0}(\mathbf{K})$)[3].

To effectively solve such a generic problem, we discretise the distribution as a finite set of stellar populations, indexed by $k$ and described by the orbital parameters $\mathbf{K}_k$. We then write $F_{\rm eq}(\widehat{\mathbf{L}},\mathbf{K}) = \sum_k F_k(\widehat{\mathbf{L}})\,\delta_{\rm D}(\mathbf{K}-\mathbf{K}_k)$ with $F_k(\widehat{\mathbf{L}})$ the distribution of orientation of the $k^{\rm th}$ population and $N_k = \int \mathrm{d}\widehat{\mathbf{L}} F_k(\widehat{\mathbf{L}})$ its number of stars (see equation 7). After this discretisation, a cluster's configuration is fully characterised by its set of order parameters

$$\boldsymbol{\theta} = \big(\beta, \gamma, \{M_{\ell,k}\}\big), \tag{16}$$

with $M_{\ell,k} = \int \mathrm{d}\widehat{\mathbf{L}} Y_{\ell 0}(\widehat{\mathbf{L}}) F_k(\widehat{\mathbf{L}})$ the axisymmetric $\ell$-magnetisation of the $k^{\rm th}$ population (see equation 10).

For a given initial condition, i.e. a given $(E_{\rm tot}, L_{\rm tot}, \{N_k\})$,

[3] In principle, there are infinitely many harmonics $\ell$, but in practice we found it acceptable to stop at $\ell_{\rm max}=10$, see §D2.

the cluster's equilibrium configurations are obtained for the parameters $\boldsymbol{\theta}$ that are joint roots of the consistency functions

$$\begin{aligned} C_E &= \big(E_{\rm tot} - E_{\rm tot}\big[F_{\rm eq}(\boldsymbol{\theta})\big]\big)/E_{\rm tot}, \\ C_L &= \big(L_{\rm tot} - L_{\rm tot}\big[F_{\rm eq}(\boldsymbol{\theta})\big]\big)/L_{\rm tot}, \\ \forall\, \ell,\, \forall\, k,\, C_{M_{\ell,k}} &= \big(M_{\ell,k} - M_{\ell,k}\big[F_{\rm eq}(\boldsymbol{\theta})\big]\big)/(N_k\, y_\ell), \end{aligned} \tag{17}$$

with $y_\ell = \sqrt{(2\ell+1)/(4\pi)}$ and $F_{\rm eq}(\boldsymbol{\theta})$ coming from equation (14). Importantly, in order to place all the constraints from equation (17) on equal footings, the consistency functions $\mathbf{C} = \big(C_E, C_L, C_{M_{\ell,k}}\big)$ are all dimensionless, and rescaled to be of order unity.

In order to find roots of the function $\boldsymbol{\theta} \mapsto \mathbf{C}[\boldsymbol{\theta}]$, we use Newton–Raphson's method (e.g., Press et al. 2007), following Takács & Kocsis (2018). More precisely, starting from a configuration $\boldsymbol{\theta}_n$, we compute

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \mathbf{J}^{-1}[\boldsymbol{\theta}_n]\, \mathbf{C}[\boldsymbol{\theta}_n], \tag{18}$$

to obtain the next iteration[4], with $\mathbf{J}[\boldsymbol{\theta}]_{ij} = \partial C_i/\partial \theta_j$ the Jacobian of the consistency function $\mathbf{C}$ at point $\boldsymbol{\theta}$. Given that $\mathbf{C}$ has a simple analytical form in equation (17), so does its Jacobian, as detailed in §B.

Equation (18) is the second important difference between Gruzinov et al. (2020) and the present work. Indeed, Gruzinov et al. (2020) solves the self-consistency requirement on their equation (3) by iteratively computing a sequence $\psi \to F \to \rho \to \psi \to \cdots$, of potentials, DFs and densities, up to convergence. Here, we directly ensure self-consistency in equation (14) by using Newton's method.

### 2.5 Optimisation strategy

To finalise our algorithm, it only remains to specify our choice for the starting point $\boldsymbol{\theta}_0$ of the iteration process.

In §C1 we explore a first method where $\boldsymbol{\theta}_0$ is initialised at random. Such an agnostic approach allows us to recover a cluster's both stable, metastable and unstable equilibria, and to extend the results of Takács & Kocsis (2018) to (axisymmetric) multi-population clusters. Using this method, we recover all the qualitative behaviours of the single-population clusters reported in Roupas et al. (2017) and Takács & Kocsis (2018), and confirm that the branch that goes through $\beta = 0$ (see Fig. C1) always has the highest entropy. Unfortunately, as one increases the number of stellar populations, the efficiency of this protocol drastically drops.

Dealing with systems with many stellar populations therefore requires improvements to the initialisation process. To alleviate most of these difficulties, we restrict ourselves, and predict only the clusters' global thermodynamical equilibria, and none of the other possible equilibria, should they be unstable or metastable. Benefiting from the insight of Fig. C1, we obtain these global equilibria by iteratively moving along the series of equilibria associated with the branch that has a solution for $\beta = 0$ (see §C2). Of course, the main drawback of this approach is that it cannot be used to determine any of the unstable thermodynamical equilibria.

[4] Note that in practice, it is faster and numerically more stable to solve the linear equation $\mathbf{J}\,\mathbf{x} = -\mathbf{C}$ for the unknown $\mathbf{x} = \boldsymbol{\theta}_{n+1} - \boldsymbol{\theta}_n$, rather than to explicitly compute the inverse matrix, $\mathbf{J}^{-1}$.

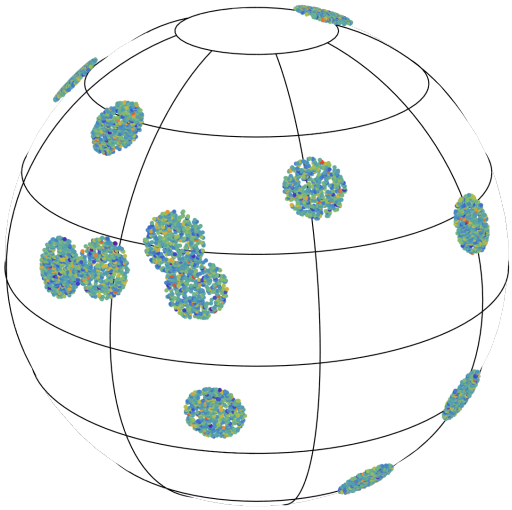

**Figure 2.** Illustration of a cluster's typical initial condition. It consists of $N_{\rm disc} = 16$ dense patches of stars scattered uniformly on the unit sphere, and associated with distinct episodes of star formation or infall events. Each patch is made of $N_{\rm part} = 512$ stars distributed uniformly within a small angular section around the patch's centre. Each star is coloured according to the norm of its angular momentum vector, with the smallest norm in blue and the largest one in red. Clearly, stars with small angular momenta are the most numerous.

## 3 THERMODYNAMICAL NUCLEAR EQUILIBRIA

We can now make use of our entropy optimisation algorithm to investigate the typical equilibrium distribution of orientations in galactic nuclei. Benefiting from the efficiency and versatility of this method, we will also use it to explore the (large) parameter space describing possible initial stellar clusters.

### 3.1 Parametrisation of the stellar population

Let us first parametrise the distribution of the orbital parameters $\mathbf{K}_k$, i.e., the distribution of masses, semi-major axes, and eccentricities. Our fiducial model is the same as in Szölgyén & Kocsis (2018). We assume that stars are formed through a series of distinct episodes of star formation or infall events. More precisely, for a given realisation, we generate $N_{\rm disc} = 16$ discs, each of them composed of $N_{\rm part} = 512$ stars, so that the cluster's total number of particles is $N = N_{\rm disc} \times N_{\rm part} = 8\,192$. For each disc, its average orientation, $\widehat{\mathbf{L}}_{\rm disc}$, is drawn uniformly on the the unit sphere, while the orientations of the stars within that disc are drawn uniformly within the small region $\widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}_{\rm disc} \geqslant 0.994$. Finally, for each star, the orbital parameters $\mathbf{K} = (m, a, e)$ are drawn independently from one another according to power law probability distribution functions (PDFs) proportional to $(m^{-2}, a^0, e)$ respectively within the ranges $m_{\rm max}/m_{\rm min} = 100$, $a_{\rm max}/a_{\rm min} = 100$, and $(e_{\rm min}, e_{\rm max}) = (0, 0.3)$. In Fig. 2, we illustrate one typical realisation of such a protocol. Following §2.3, we recall that the realisations are always rotated to have their $\mathbf{L}_{\rm tot}$ aligned along $+z$.

Once an initial distribution has been drawn, one may compute its two key invariants, $E_{\rm tot}$ and $|\mathbf{L}_{\rm tot}|$. In practice, we keep track of these two invariants through the two dimensionless quantities

$$E = -\frac{E_{\rm tot}}{N^2\, G m_{\rm min}^2/a_{\rm min}}; \quad s = \frac{|\mathbf{L}_{\rm tot}|}{\sum_i L(\mathbf{K}_i)}, \tag{19}$$

which we respectively call the binding energy and spin of the stellar cluster. In Fig. 3, we represent the typical distribution of $(E, s)$

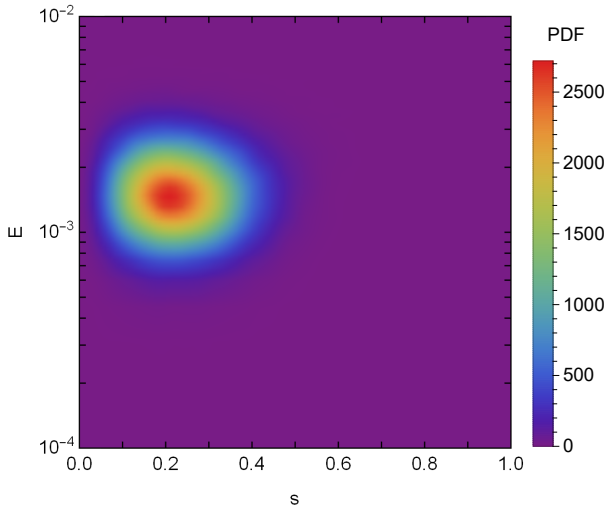


**Figure 3.** Distribution of the spin and binding energy of clusters drawn following the fiducial protocol from §3.1, for $2^{18}$ realisations of initial conditions and $\ell_{\max} = 10$.

for a large number of stellar clusters drawn from our fiducial procedure. We find that clusters exhibit a wide range of binding energy and spin, and in turn this somewhat impacts the diversity of thermodynamical equilibria.

### 3.2 A typical equilibrium distribution

In Fig. 3, we determined the distribution of the clusters' invariants $(E, s)$. In order to characterise the properties of the associated systems, we now investigate in detail the particular case $(E, s) = (2\times10^{-3}, 0.2)$, i.e. a typical cluster realisation. We detail in §D1 our precise choices for the discretisation of the stellar populations. By maximising the entropy, we determine this cluster's equilibrium configuration, $F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K})$. This is first illustrated in Fig. 4 in the $(m, \widehat{L}_z)$-plane. The same cluster is alternatively represented in Fig. 5, where we present the stellar density distribution on the sphere for two different mass bins.

As is clearly visible, heavy particles tend to have their unit angular momentum vector $\widehat{\mathbf{L}}$ oriented towards the North pole, meaning that the associated stars tend to orbit near the equatorial plane, a conclusion already reached by Szölgyén & Kocsis (2018) using an alternative Monte–Carlo approach[5], and by Gruzinov et al. (2020) in the more general case of Keplerian elliptic wires. Having the heavy particles, e.g., the IMBHs, orbit close to the same orbital plane should drastically impact the rate of their pairwise mergers in galactic nuclei. In practice, we also repeated the experiment from Fig. 4 a hundred times by letting $(E, s)$ explore the distribution from Fig. 3. The associated ensemble-averaged DF was found to be similar to the one from Fig. 4.

For the heaviest stars, we also note the presence of an additional over-density near the South pole indicating the presence of a counter-rotating equatorial disc. However, owing to the conservation of angular momentum, this component is less populated than the main prograde disc.

[5] Note that even when accounting for the different normalisation, the details of the DF in $(m, \widehat{L}_z)$-space obtained here in Fig. 4 somewhat differ from the equivalent fig. 2 of Szölgyén & Kocsis (2018).

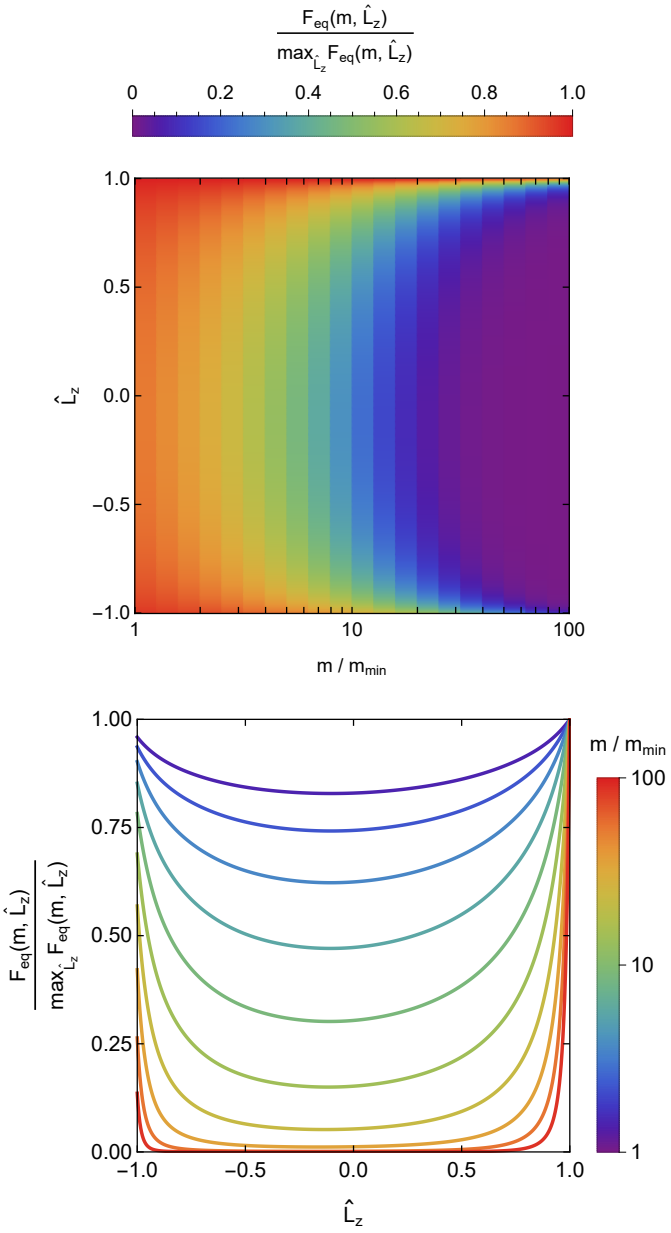


**Figure 4.** *Top panel*: Illustration of the relaxed stellar density distribution, $F_{\rm eq}(m, \widehat{L}_z)$ (integrated over all semi-major axes and eccentricities), as a function of mass and orientation for a relaxed cluster with $(E, s) = (2\times10^{-3}, 0.2)$. See §D for the numerical details. Our normalisation allows for each mass bin to have its maximum equal to 1. Heavier populations are much more segregated towards the poles. *Bottom panel*: Same as the top panel, where each line corresponds to a different mass bin.

As one considers lighter particles, the anisotropy fades away and the lightest particles do not show any strong sign of spontaneous orientation alignment. This concurs with Gruzinov et al. (2020)'s results (see §6.3. therein), which found that light objects follow a spherically symmetric distribution.

Having the full equilibrium distribution $F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K})$ at our disposal, one may also study its dependence w.r.t. the semi-major axis and eccentricity. Limiting ourselves to particles of intermediate mass, we find that the disc is slightly thinner for intermediate semi-major axes, as illustrated in Fig. 6. This is in agreement with

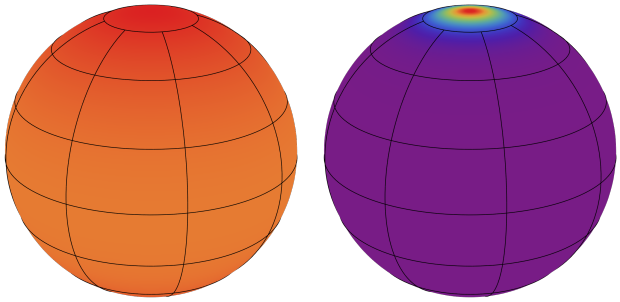

**Figure 5.** Stellar density distribution on the unit sphere, after relaxation, for the lightest particles (left, $1 \leqslant m/m_{\rm min} \leqslant 1.25$) and the heaviest ones (right, $79 \leqslant m/m_{\rm min} \leqslant 100$). We use the same normalisation and colouring as in Fig. 4. Importantly, we note that the heavy particles are significantly more segregated towards the poles, i.e. the corresponding heavy stars live within a thin equatorial disc.

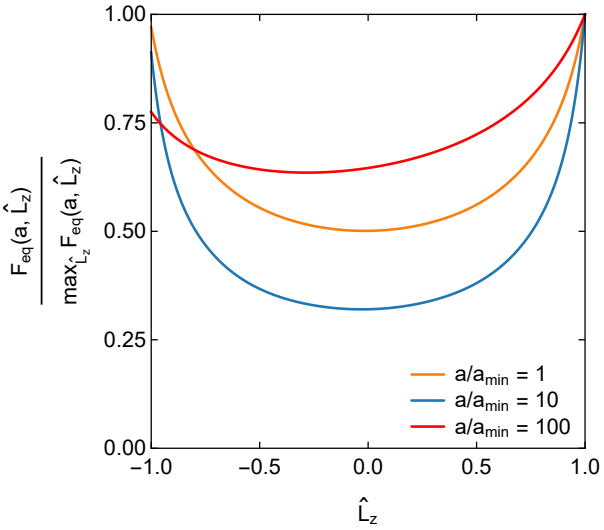


**Figure 6.** Illustration of the relaxed stellar density distribution, $F_{\rm eq}(a, \widehat{L}_z)$ (integrated over all eccentricities, but only intermediate masses in $2 \leqslant m/m_{\rm min} \leqslant 16$), as a function of semi-major axis and orientation, for the same cluster as in Fig. 4. The normalisation allows for each mass bin to have its maximum equal to 1. Populations with an intermediate semi-major axis tend to segregate within a thinner disk.

the recent result from Máthé et al. (2022) (see fig. 7 therein).

As for eccentricities, we find that, in essence, $F_{\rm eq}$ is independent of $e$. This was expected given that the coupling coefficients, $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}']$ (equation 4) and the norm of the angular momentum $L(\mathbf{K})$ (equation 12) only weakly depend on $e$, especially for the chosen quasi-circular orbits $0 \leqslant e \leqslant 0.3$.

### 3.3 Impact of the cluster's binding energy and spin

As highlighted in Fig. 3, the clusters exhibit a significant diversity in their binding energies and spins, given our generation protocol. Let us therefore investigate the dependence of the equilibria w.r.t. these invariants. In Fig. 7, we present series of equilibria (caloric curves) giving the inverse temperature, $\beta$, as a function of the normalised total energy, $-E$, for various total angular momentum (i.e. various $s$). We observe that the temperature generically increases with energy, except in a small region of negative specific heat. Also,

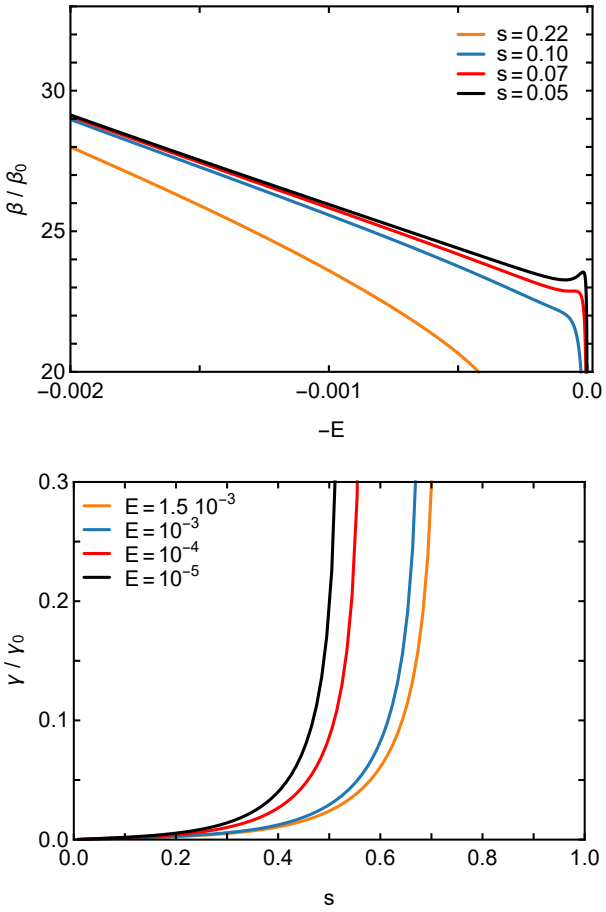


**Figure 7.** *Top panel*: Illustration of the equilibrium inverse temperature $\beta$ as a function of the cluster's normalised total energy, $-E$ (equation 19) for different values of spin. Each line is composed of 500 systems, all with $(m, a, e)$ drawn following the protocol of §3.1, and with binding energies distributed on a logarithmic grid ranging from $10^{-6}$ to $10^{-1}$. *Bottom panel*: Illustration of the equilibrium angular frequency $\gamma$ as a function of the cluster's spin for different binding energies. For each line, 100 systems were simulated, with spins distributed on a linear grid ranging from 0 to 1. See equation (C1) for the definitions of $\beta_0$ and $\gamma_0$. As one lowers the total angular momentum, via $s$, thermodynamical equilibria can exhibit negative specific heats, i.e. $C = \partial E_{\rm tot}/\partial T < 0$ (Roupas et al. 2017). For a given value of $s$, our approach to determine the series of equilibria is to start from the (much) simpler problem at $\beta = 0$ and move up to the target energy step by step iteratively (see §C2).

for any value of the spin there exists a region of negative temperature. Both effects had already been reported in Roupas et al. (2017) and Takács & Kocsis (2018), and are further discussed in §C1. Note however that these behaviours arise for $E \lesssim 10^{-4}$, i.e. well outside of the astrophysical regime, $E \gtrsim 10^{-3}$, found in Fig. 3. In the same figure, we also present series of equilibria giving the angular frequency $\gamma$ as a function of the total angular momentum $s$, for fixed values of the energy $E$. The angular frequency is found to always increase with the cluster's spin.

Within the domain of invariants spanned by Fig. 3, all equilibria are found to remain qualitatively similar but differ in the strength of their anisotropy. To characterise the level of anisotropy in the orbital distribution, we introduce a cluster's segregation rate,

SR, as

$$\mathrm{SR} = \frac{\displaystyle\int_{m\geqslant m_\mathrm{c}} \mathrm{d}\mathbf{K} \int_{\widehat{L}_z \geqslant \widehat{L}_\mathrm{c}} \mathrm{d}\widehat{\mathbf{L}}\, F_\mathrm{eq}(\widehat{\mathbf{L}}, \mathbf{K})}{\displaystyle\int_{m\geqslant m_\mathrm{c}} \mathrm{d}\mathbf{K} \int \mathrm{d}\widehat{\mathbf{L}}\, F_\mathrm{eq}(\widehat{\mathbf{L}}, \mathbf{K})}, \tag{20}$$

This quantity describes which fraction of heavy particles (i.e. with $m \geqslant m_\mathrm{c}$) have an angular momentum vector that is well-aligned with the cluster's total angular momentum (i.e. have $\widehat{L}_z \geqslant \widehat{L}_\mathrm{c}$). With such a definition, we note that anti-aligned particles are not accounted for. The larger SR, the stronger the anisotropic segregation of the heavy populations. In practice, guided by Fig. 4, we consider the values $m_\mathrm{c} = 10\, m_\mathrm{min}$ and $\widehat{L}_\mathrm{c} = \cos(20^\circ)$. We expect that changing the values of $m_\mathrm{c}$ does not qualitatively impact the present conclusions.

In Fig. 8, we illustrate the clusters' segregation rate as a function of the binding energy and spin $(E, s)$. We observe the crucial role played by these two invariants in driving the spontaneous alignment of the orbital orientations of the heaviest particles. The larger $E$ and the larger $s$, the larger the proportion of heavy particles that are aligned near the equatorial plane.

To strengthen this conclusion, let us finally determine the angular size, $\theta_\mathrm{c}$, of the northern polar cap that contains a fraction $f_\mathrm{c}$ of the heavy particles' unit angular momentum vectors $\widehat{\mathbf{L}}$. More precisely, we define $\theta_\mathrm{c}$ through the implicit constraint

$$\frac{\displaystyle\int_{m\geqslant m_\mathrm{c}} \mathrm{d}\mathbf{K} \int_{\widehat{L}_z \geqslant \cos(\theta_\mathrm{c})} \mathrm{d}\widehat{\mathbf{L}}\, F_\mathrm{eq}(\widehat{\mathbf{L}}, \mathbf{K})}{\displaystyle\int_{m\geqslant m_\mathrm{c}} \mathrm{d}\mathbf{K} \int \mathrm{d}\widehat{\mathbf{L}}\, F_\mathrm{eq}(\widehat{\mathbf{L}}, \mathbf{K})} = f_\mathrm{c}, \tag{21}$$

noting once again that anti-aligned stars are not accounted for. Equivalently, $\theta_\mathrm{c}$ is also the angular size of the equatorial disk that contains a fraction $f_\mathrm{c}$ of the heavy particles. Using a fraction $f_\mathrm{c} = 50\%$, the dependence of $\theta_\mathrm{c}$ w.r.t. the clusters' invariants is illustrated in Fig. 8. Once again, the more bound the cluster, and the larger its total angular momentum, the thinner the disc of heavy particles in the relaxed equilibrium.

Interestingly, if equations (20) and (21) were to consider both aligned and anti-aligned stars, through the replacement $\widehat{L}_z \geqslant \cos(\theta_\mathrm{c}) \rightarrow |\widehat{L}_z| \geqslant \cos(\theta_\mathrm{c})$, the dependence w.r.t. $s$ in Fig. 8 would be significantly reduced. This result is expected since in equation (8), the total energy is left invariant by the changes $\widehat{\mathbf{L}}_i \rightarrow -\widehat{\mathbf{L}}_i$, while in equation (11) the total angular momentum changes sign.

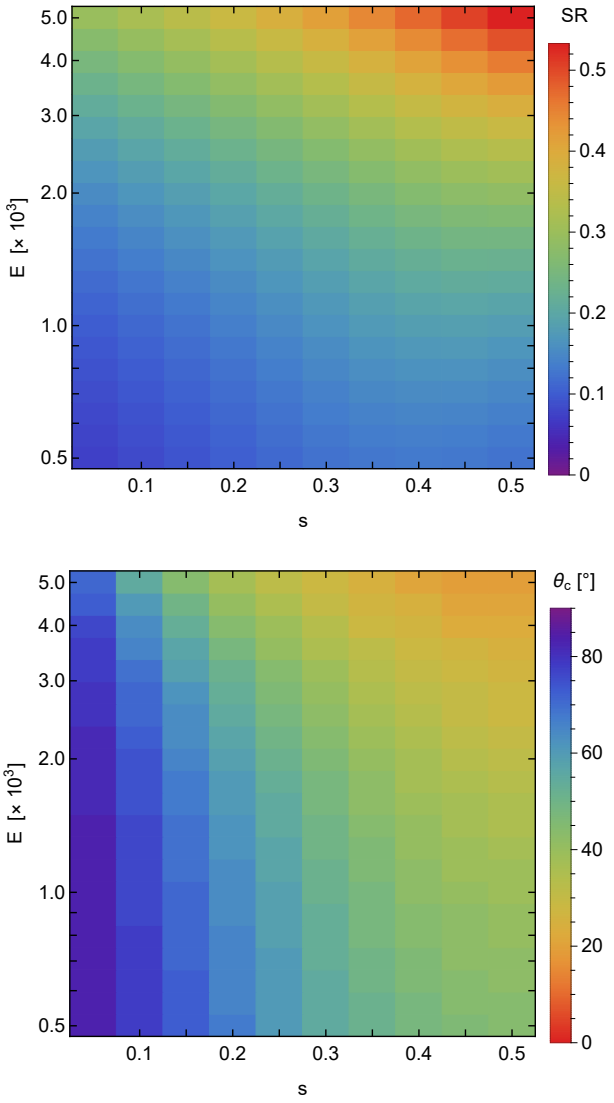


**Figure 8.** *Top panel*: Segregation rate of the relaxed clusters (equation 20), within the range of spins and binding energies expected for typical astrophysical clusters (see Fig. 3). *Bottom panel*: Same domain as the top panel but for the segregation angle $\theta_\mathrm{c}$ (equation 21). Clusters with large binding energies and spins exhibit the strongest segregation of their heavy particles within the equatorial disc.

### 3.4 Impact of the mass and semi-major axes distributions

Let us now step out and explore the impact of the distribution of the orbital stellar parameters themselves on the clusters' thermal equilibria. Following §3.1, we now vary the PDFs w.r.t. which the stars' masses and semi-major axes are drawn, keeping all other parameters the same. More precisely, we still draw $(m, a)$ pairs according to power law distributions, $(m^{\gamma_m}, a^{\gamma_a})$, but this time vary the power law indices $(\gamma_m, \gamma_a)$ between different realisations.

The impact of changing these two parameters on the average segregation rate of the heavy particles, $\langle\theta_\mathrm{c}\rangle$, is illustrated in Fig. 9. We find that the slope $\gamma_m$ has the strongest effect with clusters containing fewer heavy stars segregating in a disk twice thinner than the fiducial cluster. The semi-major axis power law index also has an impact on the segregation angle: clusters with larger semi-major axes tend to display a stronger mass segregation.

These dependencies may possibly be used in the future to estimate indirectly the stellar and compact objects' initial mass function in galactic centres. Indeed, the rate of pairwise IMBH mergers should be measurable using gravitational waves, and could serve as an indirect probe for the width of the segregated massive disc, which we just have shown is linked to the IMBHs' initial mass function.

### 3.5 Impact of the dispersion of disc orientations

A cluster's long-term equilibrium distribution is fully characterised by its three invariants: the binding energy, $E$, the spin, $s$, and the distribution of orbital parameters, $N(\mathbf{K})$. Unfortunately, while dynamically relevant, such parameters do not translate easily in terms of astrophysical observables. As such, let us finally switch to

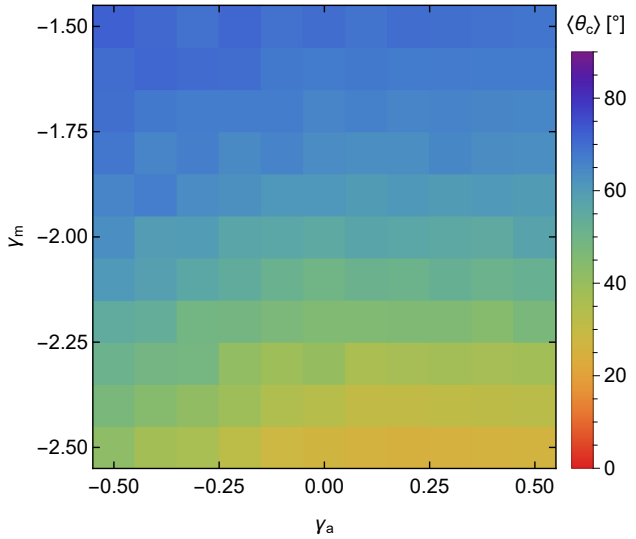


**Figure 9.** Average segregation angle, $\langle\theta_c\rangle$, as a function of the power law indices of the initial distributions in semi-major axes and masses. Here, for each value of $(\gamma_a, \gamma_m)$, the average has been computed over 128 realisations of a cluster. The fiducial case from Fig. 4 corresponds to $(\gamma_a, \gamma_m) = (0, -2)$. Interestingly, both power law indices are found to impact the relaxed segregation angle.

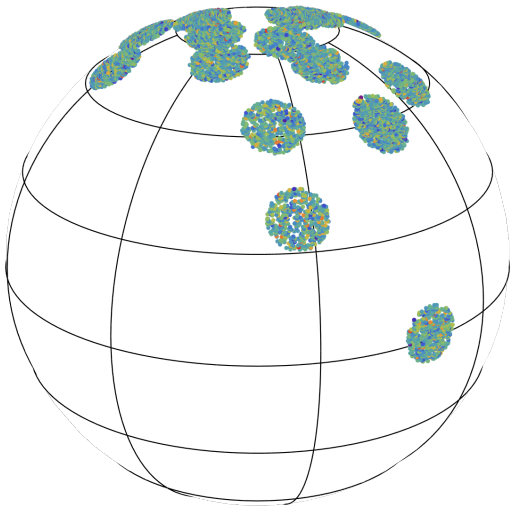

**Figure 10.** Same as Fig. 2, but assuming that the orientations of the discs are drawn from a von Mises–Fisher PDF with concentration $\kappa = 5$ (equation 22).

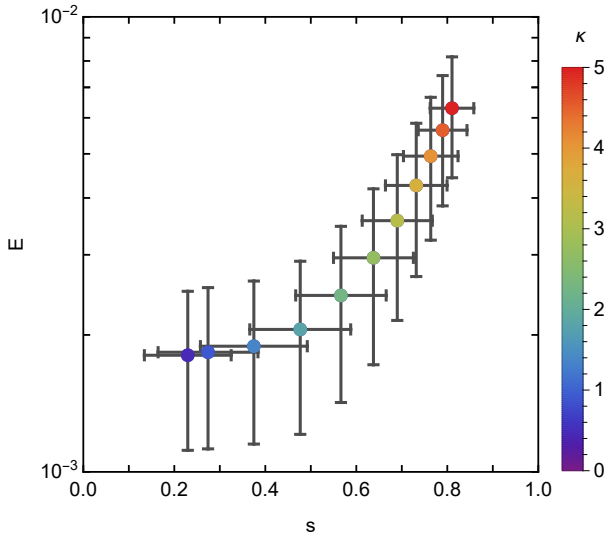


**Figure 11.** Typical distribution of the spin and binding energies as one varies the concentration $\kappa$ of the initial discs. Coloured points correspond to the mean values of $(s, E)$, while black error bars are standard deviations in both parameters. Statistics are computed from $10^4$ realisations. As expected, the more concentrated the discs' orientations, the more packed the initial distributions, and therefore the larger the spins and binding energies.

parametrisations more closely related to the underlying formation process of the galactic centre's stellar cluster, such as the geometry of the initial distribution.

Building upon Fig. 2, rather than drawing the initial orientations of the discs uniformly on the unit sphere, we may draw them along some biased direction, for example to reflect the preferential infall of new star-forming gas along specific directions, e.g., imposed by the geometry of past gaseous accretion events. More precisely, we assume that the average orientation of each disc is drawn according to a Von Mises–Fisher PDF (Wood 1994) of the form

$$P(\widehat{\mathbf{L}}) = \frac{\kappa}{4\pi \sinh(\kappa)} \, \mathrm{e}^{\kappa \widehat{L}_z}, \tag{22}$$

with $\kappa$ the PDF's concentration. Here, $\kappa = 0$ corresponds to the isotropic case considered in Fig. 2. The larger $\kappa$, the smaller the spread of the PDF on the unit sphere, as in Fig. 10 for $\kappa = 5$.

Keeping all other stellar parameters as in §3.1, we may now study the impact of the concentration parameter $\kappa$ on the shape of the relaxed distributions. This is first illustrated in Fig. 11, where we show the dependence of the invariants $(E, s)$ with $\kappa$. As expected, we recover that the more concentrated the distribution of the discs' initial orientations, the larger the value of $s$ and $E$. In Fig. 12, we subsequently illustrate the dependence of the average segregation angle, $\langle\theta_c\rangle$ (equation 21), as a function of $\kappa$. Narrower initial distributions of orientations lead to stronger segregation of the heavy stars at equilibrium. This is expected, since Fig. 8 showed that the segregation strength correlates positively with both binding energy and spin. To emphasise this conclusion, Fig. 12 also illustrates the thickness of the asymptotic massive disc as a function of the initial mass function slope. Similarly to Fig. 9, at fixed initial anisotropy, the steeper the slope, the stronger the mass segregation. Overall, investigations as in Fig. 12 should ultimately prove useful to place some constraints on the origin of SgrA*'s surrounding stellar cluster, in particular in the light of the "paradox of youth" (Ghez et al. 2003; Genzel et al. 2010).

## 4 CONCLUSION

In the spirit of Roupas et al. (2017); Takács & Kocsis (2018), we relied on maximum entropy methods to assess the internal stellar structure of galactic nuclei and their underlying distribution of orbital orientations. We expanded their approach to multi-populations clusters and showed how to jointly account for the constraints of energy, angular momentum, and orbital parameters conservations to efficiently characterise a cluster's expected thermodynamical equi-

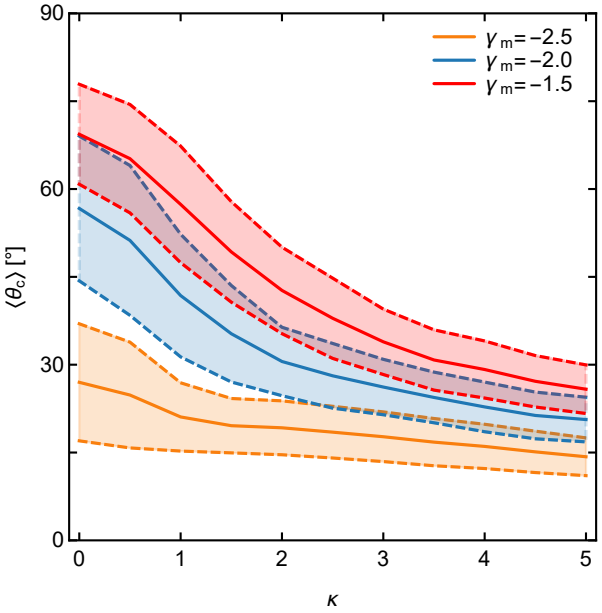


**Figure 12.** Average segregation angle, $\langle\theta_c\rangle$, as a function of the concentration parameter $\kappa$, for three different mass distributions. The solid lines represent the average segregation angle, and the dashed lines represent the standard deviation. The statistics are estimated from $2^{10}$ realisations – though $\sim 10\%$ of the realisations for $(\kappa, \gamma_m) = (5, -2.5)$ failed because of numerical overflows encountered with very strongly anisotropic DFs. The more anisotropic a cluster initially, the higher its binding energy and spin, and therefore the more segregated at equilibrium.

libria. In particular, as already pointed out in Szölgyén & Kocsis (2018), we recovered the spontaneous alignment of the heavier stellar components, such as IMBHs (Fig. 4). The more bound the cluster and the larger its spin, the thinner the asymptotic disc (Fig. 8).

Benefiting from the versatility of the present method, we used it to explore the dependence of these structures w.r.t. the cluster's initial conditions of formation, such as the intrinsic power law spectra in mass and semi-major axis (Fig. 9), or the initial spread in orientations (Fig. 12). Small modifications to the present algorithm would also make it possible to confirm that the final equilibrium does not depend on the specific sequence of star formation events (Szölgyén & Kocsis 2018).

Of course, this work is but one step towards the systematic exploration of the remaining imprints of a galactic nucleus' initial formation on the cluster's long-term distribution of orientations. Let us conclude by listing a few possible venues for future works.

The present approach could first benefit from various extensions on the theoretical and analytical front. (i) We have assumed from the start that stellar eccentricities were conserved throughout the VRR evolution. This amounts to neglecting the process of SRR (Rauch & Tremaine 1996), whose signature is clearly visible in the S-cluster distribution (see, e.g., Tep et al. 2021). As already recently hinted in Szölgyén et al. (2021), allowing for the eccentricities to also vary during the orientations' relaxation might impact the system's VRR equilibria. This deserves careful study. Similarly, stars can also change their semi-major axis through non-resonant relaxation (NR), whose impact on the outcome of VRR should also be investigated. (ii) We recovered here that VRR may lead to highly non-spherical distribution of orientations. Such an end-state distribution is expected to impact in turn the rate of orientation diffusion of a given test star and the efficiency with which newly formed stellar discs can dilute (see, e.g., Giral Martínez et al. 2020).

The stochastic process of spontaneous alignment is not limited to the axisymmetric relaxation of Keplerian annuli around supermassive BHs. As such, at a significantly larger numerical cost, one could extend the present approach to non-axisymmetric distributions, as was already done in Roupas et al. (2017) in the limit of a single-population system with a quadrupolar interaction. Similarly, it would be of interest to investigate the distinction between the present global equilibria and the other possible metastable ones[6]: do they also exhibit a disc-like structure? Finally, one should investigate whether there exist astrophysically relevant regimes with negative temperatures or heat capacities (Figs. 7 and C1).

In equation (1), we assumed that we could average the cluster's Hamiltonian over the in-plane precessions. We could lift this assumption and determine the equilibrium distributions of Keplerian elliptic wires (see, e.g., Gruzinov et al. 2020; Tremaine 2020b). Finally, since the mean potential remains on average spherically symmetric, all these investigations could also be performed in globular clusters, where stellar metallicities or ages could be used as additional tracers of the mixing of stellar orientations.

In the present work, we restricted ourselves to predicting the end-state of VRR, i.e. its thermodynamical equilibrium. By design such an approach cannot provide any estimate of the expected relaxation time required for this asymptotic distribution to be reached. There are, at least, two possible venues to quantify such relaxation time: (i) One could use direct time integrations of the equations of motion to get the equilibrium distributions, a goal already pursued in Kocsis & Tremaine (2015). Building upon Fouvry et al. (2020), one can expect that efficient multipole methods may be designed to perform such direct numerical simulations with a computational complexity scaling linearly with the total number of particles; (ii) In the limit of sufficiently symmetric orbital distributions, e.g., axisymmetric (Fouvry et al. 2019a), one could alternatively derive an explicit kinetic theory for the cluster, from which the relaxation time would naturally follow. Interestingly, we report that the axisymmetric equilibrium recovered in Fig. 4 generically exhibits a monotonic profile of latitudinal precession frequency for all stellar populations. This may induce a situation of "kinetic blocking" and play a critical role in defining the efficiency with which these systems may relax (Fouvry et al. 2019a).

Finally, benefiting from the planned upgrade on VLTI (Eisenhauer 2019; Gravity Collab. et al. 2021), as well as the future thirty-meter class telescopes such as ELT (Pott et al. 2018; Davies et al. 2018) and TMT (Do et al. 2019), we will soon have a wealth of orbitally-resolved stars around SgrA* along with their stellar ages. Building upon the present work, a detailed characterisation of their orbital distribution should prove paramount to place constraints on the properties of the (likely present) un-observed IMBHs, and the impact of their effective distribution of orientations on their overall in-situ merger rates.

[6] For systems with long-range interactions, metastable states have a very long lifetime scaling as $e^N$ (see, e.g., Chavanis 2006), because a system trapped in a metastable state (local entropy maximum) has to cross a huge barrier of potential to reach the fully stable equilibrium (global entropy maximum).

**ACKNOWLEDGEMENTS**

This work is partially supported by the grant Segal ANR-19-CE31-0017 of the French Agence Nationale de la Recherche, and by the Idex Sorbonne Université. NM acknowledges support from SUPAERO, the Université Paul Sabatier and the École polytechnique. We thank Stéphane Rouberol for running smoothly the Infinity cluster, where the simulations were performed.

**DATA AVAILABILITY**

The data and numerical codes underlying this article were produced by the authors. They will be shared on reasonable request to the corresponding author. The code is distributed on Github at the following URL: https://github.com/NathanMagnan/VRROOMe.

**REFERENCES**

Alexander T., 2017, ARA&A, 55, 17
Antoni M., Ruffo S., 1995, Phys. Rev. E, 52, 2361
Bartko H., et al., 2009, ApJ, 697, 1741
Chavanis P.-H., 2006, Int. J. Mod. Phys. B, 20, 3113
Chavanis P.-H., Vatteville J., Bouchet F., 2005, Eur. Phys. J. B, 46, 61
Davies R., et al., 2018, in Ground-based and Airborne Instrumentation for Astronomy VII. p. 107021S
Do T., et al., 2019, BAAS, 51, 530
Eisenhauer F., 2019, in The Very Large Telescope in 2030. p. 30
Fouvry J.-B., Bar-Or B., Chavanis P.-H., 2019a, Phys. Rev. E, 99, 032101
Fouvry J.-B., Bar-Or B., Chavanis P.-H., 2019b, ApJ, 883, 161
Fouvry J.-B., Dehnen W., Tremaine S., Bar-Or B., 2020, arXiv, 2011.01673
Genzel R., Eisenhauer F., Gillessen S., 2010, Rev. Mod. Physics, 82, 3121
Ghez A. M., et al., 2003, ApJ, 586, L127
Ghez A. M., et al., 2008, ApJ, 689, 1044
Gillessen S., et al., 2017, ApJ, 837, 30
Giral Martínez J., Fouvry J.-B., Pichon C., 2020, MNRAS, 499, 2714
Gravity Collab. et al., 2020, A&A, 636, L5
Gravity Collab. et al., 2021, A&A, 645, A127
Gruzinov A., Levin Y., Zhu J., 2020, ApJ, 905, 11
Habibi M., et al., 2017, ApJ, 847, 120
Heckman T. M., Best P. N., 2014, ARA&A, 52, 589
Katz J., 1978, MNRAS, 183, 765
Klimontovich I., 1967, The statistical theory of non-equilibrium processes in a plasma. M.I.T. Press
Kocsis B., Tremaine S., 2011, MNRAS, 412, 187
Kocsis B., Tremaine S., 2015, MNRAS, 448, 3265
Kormendy J., Ho L. C., 2013, ARA&A, 51, 511
Lynden-Bell D., Wood R., 1968, MNRAS, 138, 495
Magorrian J., et al., 1998, AJ, 115, 2285
Máthé G., Szölgyén Á., Kocsis B., 2022, arXiv, 2202.07665
Murchikova E. M., Phinney E. S., Pancoast A., Blandford R. D., 2019, Nature, 570, 83
Murray C., Dermott S., 1999, Solar System Dynamics. Cambridge Univ. Press
Onsager L., 1949, Il Nuovo Cimento, 6, 279
Pichon C., Lynden-Bell D., 1993, in Statistical Description of Transport in Plasma, Astro- and Nuclear Physics. p. 261
Portegies Zwart S. F., McMillan S. L. W., 2002, ApJ, 576, 899
Pott J. U., et al., 2018, in Ground-based and Airborne Instrumentation for Astronomy VII. p. 1070290
Press W., et al., 2007, Numerical Recipes 3rd Ed.. Cambridge Univ. Press
Rauch K. P., Tremaine S., 1996, New Astron., 1, 149
Roupas Z., 2020, J. Phys. A, 53, 045002
Roupas Z., Kocsis B., Tremaine S., 2017, ApJ, 842, 90
Szölgyén Á., Kocsis B., 2018, Phys. Rev. Lett., 121, 101101
Szölgyén Á., Máthé G., Kocsis B., 2021, ApJ, 919, 140
Takács Á., Kocsis B., 2018, ApJ, 856, 113
Tep K., Fouvry J.-B., Pichon C., Heißel G., Paumard T., Perrin G., Vincent F., 2021, MNRAS, 506, 4289
Touma J., Tremaine S., Kazandjian M., 2019, Phys. Rev. Lett., 123, 021103
Tremaine S., 2020a, MNRAS, 491, 1941
Tremaine S., 2020b, MNRAS, 493, 2632
Wood A. T., 1994, Commun. Stat.-Simul. Comput., 23, 157
Yelda S., et al., 2014, ApJ, 783, 131

**APPENDIX A: THERMODYNAMICAL EQUILIBRIUM**

Following Roupas et al. (2017), we derive the generic expression of the DF of thermodynamical equilibria, $F_{\rm eq}$, from equation (14). This DF maximises the Boltzmann entropy (equation 13), under three joint constraints: the conservation of orbital parameters (equation 7), the conservation of energy (equation 8) and the conservation of angular momentum (equation 11).

Following the double orbit-average from equation (1), we stress that the Hamiltonian of VRR (equation 2) is such that (i) the phase space domain for $\widehat{\mathbf{L}}$ is of finite volume and (ii) for a finite $\ell_{\rm max}$, there is no divergence of the pairwise VRR interaction for $\widehat{\mathbf{L}}_i = \widehat{\mathbf{L}}_j$. As such, the VRR dynamics is simplified and does not present the typical peculiarities of the statistical mechanics of self-gravitating systems such as the evaporation of stars, the formation of binaries, the gravothermal instability, and the absence of a true statistical equilibrium (see, e.g., Chavanis 2006). In the present case, we essentially deal with the dynamics of self-interacting spins on a sphere, a system which possesses well-defined statistical equilibria in all cases.

Let us consider $F_{\rm eq}$ a local extremum of entropy. Lagrange multipliers then guarantee the existence of a function $\alpha(\mathbf{K})$, a scalar $\beta$ and a vector $\boldsymbol{\gamma}$ that ensure the differential equality

$$\mathcal{D}_{F_{\rm eq}} S + \int \mathrm{d}\mathbf{K}\, \alpha(\mathbf{K}) \mathcal{D}_{F_{\rm eq}} N(\mathbf{K}) - \beta \mathcal{D}_{F_{\rm eq}} E_{\rm tot} + \boldsymbol{\gamma} \cdot \mathcal{D}_{F_{\rm eq}} \mathbf{L}_{\rm tot} = 0. \tag{A1}$$

In that expression, the differentials are given by the linear forms

$$\begin{aligned}
&\mathcal{D}_{F_{\rm eq}} S : \delta F \mapsto -k_{\rm B} \int \mathrm{d}\widehat{\mathbf{L}} \mathrm{d}\mathbf{K} \left[1 + \ln\big(F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K})\big)\right] \delta F(\widehat{\mathbf{L}}, \mathbf{K}),\\
&\mathcal{D}_{F_{\rm eq}} N(\mathbf{K}) : \delta F \mapsto \int \mathrm{d}\widehat{\mathbf{L}}\, \delta F(\widehat{\mathbf{L}}, \mathbf{K}),\\
&\mathcal{D}_{F_{\rm eq}} E_{\rm tot} : \delta F \mapsto \int \mathrm{d}\widehat{\mathbf{L}} \mathrm{d}\mathbf{K}\, \varepsilon(\widehat{\mathbf{L}}, \mathbf{K})\, \delta F(\widehat{\mathbf{L}}, \mathbf{K}),\\
&\mathcal{D}_{F_{\rm eq}} \mathbf{L}_{\rm tot} : \delta F \mapsto \int \mathrm{d}\widehat{\mathbf{L}} \mathrm{d}\mathbf{K}\, L(\mathbf{K})\, \widehat{\mathbf{L}}\, \delta F(\widehat{\mathbf{L}}, \mathbf{K}),
\end{aligned} \tag{A2}$$

where we used the symmetry $\mathcal{H}_\ell[\mathbf{K}, \mathbf{K}'] = \mathcal{H}_\ell[\mathbf{K}', \mathbf{K}']$ (see equation 4) to compute $\mathcal{D}_{F_{\rm eq}} E_{\rm tot}$. Injecting these relations into equation (A1), we find

$$\begin{aligned}
\forall\, \delta F, \int \mathrm{d}\widehat{\mathbf{L}} \mathrm{d}\mathbf{K} \big\{ &- k_{\rm B} \big[1 + \ln\big(F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K})\big)\big] + \alpha(\mathbf{K}) \\
&- \beta\, \varepsilon(\widehat{\mathbf{L}}, \mathbf{K}) + L(\mathbf{K})\, \boldsymbol{\gamma} \cdot \widehat{\mathbf{L}} \big\}\, \delta F(\widehat{\mathbf{L}}, \mathbf{K}) = 0,
\end{aligned} \tag{A3}$$

Since this integral must vanish whatever the small displacement $\delta F$ considered, we have

$$\begin{aligned}
\forall\, \widehat{\mathbf{L}}, \forall\, \mathbf{K},\ &- k_{\rm B} \big[1 + \ln\big(F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K})\big)\big] + \alpha(\mathbf{K}) \\
&+ \beta\, \varepsilon(\widehat{\mathbf{L}}, \mathbf{K}) + L(\mathbf{K})\, \boldsymbol{\gamma} \cdot \widehat{\mathbf{L}} = 0.
\end{aligned} \tag{A4}$$

Inverting this relation, we find that the equilibrium DF is necessarily of the form

$$F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K}) = \exp\big[\big(\alpha(\mathbf{K}) - 1\big) - \beta\, \varepsilon(\widehat{\mathbf{L}}, \mathbf{K}) + L(\mathbf{K})\, \boldsymbol{\gamma} \cdot \widehat{\mathbf{L}}\big], \tag{A5}$$

where, for convenience, the factor $k_\mathrm{B}$ has been absorbed in the definition of $\alpha$, $\beta$ and $\boldsymbol{\gamma}$. Following equation (7), we impose the number density of stars with orbital parameters $\mathbf{K}$ to be $N(\mathbf{K})$, and we finally recover equation (14).

## APPENDIX B: CONSISTENCY FUNCTIONS AND JACOBIAN

We detail here the expressions of the consistency functions from equation (17). In order to shorten the notations, let us first write the exponential factor from equation (14) as

$$\mathrm{e}[\widehat{\mathbf{L}},\mathbf{K}] = \exp\big[-\beta\,\varepsilon(\widehat{\mathbf{L}},\mathbf{K}) + \gamma\,L(\mathbf{K})\,\widehat{L}_z\big]. \tag{B1}$$

We may then introduce the four angular integrals

$$\begin{aligned}
I(\mathbf{K}) &= \int \mathrm{d}\widehat{\mathbf{L}}\,\mathrm{e}[\widehat{\mathbf{L}},\mathbf{K}],\\
I_\varepsilon(\mathbf{K}) &= \int \mathrm{d}\widehat{\mathbf{L}}\,\varepsilon(\widehat{\mathbf{L}},\mathbf{K})\,\mathrm{e}[\widehat{\mathbf{L}},\mathbf{K}],\\
I_L(\mathbf{K}) &= \int \mathrm{d}\widehat{\mathbf{L}}\,L(\mathbf{K})\,\widehat{L}_z\,\mathrm{e}[\widehat{\mathbf{L}},\mathbf{K}],\\
I_\ell(\mathbf{K}) &= \int \mathrm{d}\widehat{\mathbf{L}}\,Y_{\ell 0}(\widehat{\mathbf{L}})\,\mathrm{e}[\widehat{\mathbf{L}},\mathbf{K}].
\end{aligned} \tag{B2}$$

The consistency functions from equation (17) then simply read

$$\begin{aligned}
C_E &= \frac{1}{E_\mathrm{tot}}\left\{E_\mathrm{tot} - \tfrac{1}{2}\int \mathrm{d}\mathbf{K}\,N(\mathbf{K})\,\frac{I_\varepsilon(\mathbf{K})}{I(\mathbf{K})}\right\},\\
C_L &= \frac{1}{L_\mathrm{tot}}\left\{L_\mathrm{tot} - \int \mathrm{d}\mathbf{K}\,N(\mathbf{K})\,\frac{I_L(\mathbf{K})}{I(\mathbf{K})}\right\},\\
C_{M_{\ell,k}} &= \frac{1}{N_k y_\ell}\left\{M_{\ell,k} - N_k\,\frac{I_\ell(\mathbf{K})}{I(\mathbf{K})}\right\}.
\end{aligned} \tag{B3}$$

As required by equation (18), it is straightforward to compute the gradients of $(C_E, C_L, \{C_{M_{\ell,k}}\})$ w.r.t. the order parameters $(\beta, \gamma, \{M_{\ell,k}\})$. All these expressions are analytical and involve angular integrals similar to the ones present in equation (B2). In practice, having discretised the stellar populations (see §D1), the integrals w.r.t. $\mathrm{d}\mathbf{K}$ simply become $\sum_k$. In addition, benefiting from the assumption of axisymmetry (see §2.3), the integrals w.r.t. $\mathrm{d}\widehat{\mathbf{L}}$ become integrals w.r.t. $\mathrm{d}\widehat{L}_z$, which are computed using a Gauss–Legendre quadrature (see, e.g., Press et al. 2007) with 100 nodes.

## APPENDIX C: OPTIMISATION STRATEGIES

In this Appendix, we present our two main approaches to solving the iteration problem from equation (18).

### C1 Random initialisations

A first approach to initialise equation (18) is to consider starting points $\boldsymbol{\theta}_0$ taken at random. More precisely, we draw the initial order parameters independently from one another and uniformly within the domains

$$\begin{aligned}
&0 \leqslant \beta \leqslant 20\,\beta_0 \quad \text{with} \quad \beta_0 = 1/\big(N\,G m_\mathrm{min}^2/a_\mathrm{min}\big),\\
&0 \leqslant \gamma \leqslant 2\,\gamma_0 \quad \text{with} \quad \gamma_0 = 1/\big(m_\mathrm{min}\sqrt{G M_\bullet a_\mathrm{min}}\big),\\
&|M_{\ell,k}| \leqslant y_\ell\,N_k,
\end{aligned} \tag{C1}$$

with $y_\ell = \sqrt{(2\ell+1)/(4\pi)}$. The main advantage of such an agnostic approach is that it allows us to recover a cluster's both stable and unstable equilibria.

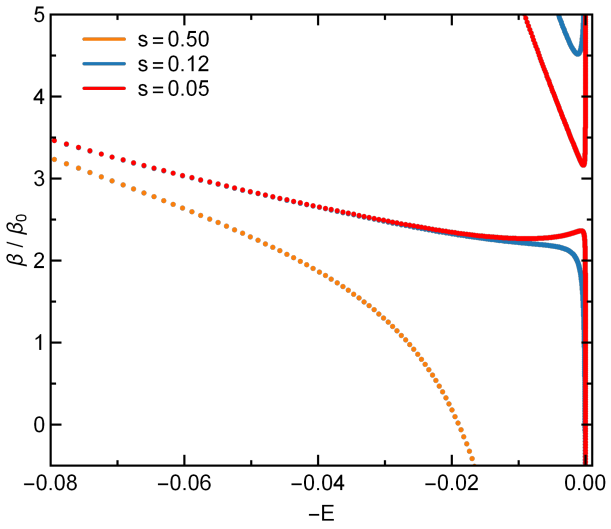


**Figure C1.** Inverse temperature $\beta$ as a function of the normalised total energy, $-E$ (see equation 19), for various spins $s$, when considering a cluster composed of 5 populations of different individual masses interacting with $\ell_\mathrm{max} = 10$. This figure was obtained without restricting ourselves to the global thermodynamical equilibria (see C1), as emphasised by the second set of (red and blue) solutions at high $\beta$.

In Fig. C1, we use this protocol for a multi-population and a multi-harmonic cluster with $(N_m, N_a, N_e) = (5, 1, 1)$, and $\ell_\mathrm{max} = 10$ – see §D for details on the cluster's properties. In a sense, this figure is similar to the top panel of Fig. 7, except that in Fig. C1, we do not restrict ourselves solely to the global thermodynamical equilibria. Interestingly, it also appears that all the qualitative behaviours reported in Roupas et al. (2017) and Takács & Kocsis (2018) are still present in the multi-population case, as can be seen by comparing with Figs. D3 and D4.

We first note that for any spin $s$, there exists a region of the caloric curve where equilibria have negative temperatures. Negative temperatures occur when the total volume of phase space is finite (which is the case for VRR), as first shown by Onsager (1949) in the context of $2D$ vortex dynamics. However, in the present context, the negative temperature states do not seem to have any particular impact and, furthermore, the region $\beta < 0$ is found to be outside of the astrophysically relevant regime (see Fig. 3).

For some binding energies and spins, there can exist several solutions to equation (14). Indeed, the Lagrange multipliers method finds all local extrema of the entropy. These could be local minima, saddle points, metastable local maxima, or the one global maximum. In the present context, the turning point method of Poincaré (see, e.g., Lynden-Bell & Wood 1968; Katz 1978) states that along a continuous branch in the $(E_\mathrm{tot}, \beta)$-plane, there can be a loss of stability around a given solution in the microcanonical ensemble (i.e. fixed $E_\mathrm{tot}$ and $L_\mathrm{tot}$) only if $E_\mathrm{tot} \mapsto \beta(E_\mathrm{tot})$ exhibits an infinite derivative there. In Fig. C1, no such turning points are observed. Given that the solution for $\beta = 0$ is always found to be stable (see §C3), we can conclude that the branch of solutions that goes through $\beta = 0$ are always, at least, metastable. In practice, we systematically computed the entropy of the various branches, and the one going through $\beta = 0$ was always found to correspond to the global (axisymmetric) maxima. This justifies our choice in §C2 to limit ourselves to only determining the equilibria along $\beta = 0$, therefore significantly alleviating the numerical burden.

In Fig. C1, we also recover that low-spin multi-population clusters can exhibit negative specific heats at small energies. However, when considering invariants $(E, s)$ provided by the protocol from §3.1, none of the realisations considered were found to fall within that domain. Finally, we note that the region of negative specific heat is delimited by a pair of equilibrium points where the function $E_{\rm tot} \mapsto \beta(E_{\rm tot})$ has a vanishing derivative. The Poincaré turning point method can be used to show that the series of equilibria in between these two points of vanishing derivative is unstable in the canonical ensemble (fixed $\beta$ and $\gamma$). This leads to a situation of ensemble inequivalence – since equilibria with negative specific heats are allowed in the microcanonical ensemble but forbidden in the canonical ensemble – characteristic of systems driven by long-range interactions (see, e.g., Chavanis 2006). Given the similarities of Fig. C1 with the results already presented in Roupas et al. (2017); Takács & Kocsis (2018), we refer to these previous works for detailed discussions of these canonical phase transitions. We stress, however, that the relevant ensemble considered in the main text is the microcanonical one since our system is assumed to be isolated.

### C2 Iterative resolutions

When dealing with systems with a much larger number of stellar populations, the random initialisation from equation (C1) is not efficient anymore. Indeed, because the total number of order parameters scale like $\mathcal{O}(N_{\rm pop}\, \ell_{\rm max})$ (see equation 16), the likelihood of starting sufficiently close to one equilibrium for equation (18) to converge drastically drops. As such, benefiting from the insight of Fig. C1, we limit ourselves to predicting the cluster's global thermodynamical equilibrium, i.e. predicting the equilibrium that lies on the branch that goes through $\beta = 0$.

More precisely, assuming that the cluster at hand is characterised by the two invariants $(E_{\rm tot}, L_{\rm tot})$, we first determine the energy $E^{(0)}_{\rm tot}$ and order parameters $\boldsymbol{\theta}^{(0)}$ associated with the thermodynamical equilibrium at temperature $\beta = 0$ and total angular momentum $L_{\rm tot}$. We highlight in §C3 how such a problem is significantly easier to solve, as it does not involve any self-gravitating contribution, therefore making equation (14) explicit.

From this initial configuration, we use $\boldsymbol{\theta}^{(0)}$ as an initial condition to solve a new self-consistency problem for the invariants $(E^{(1)}_{\rm tot}, L_{\rm tot})$, with $E^{(1)}_{\rm tot}$ lying between $E^{(0)}_{\rm tot}$ and the final target $E_{\rm tot}$. Provided that $E^{(1)}_{\rm tot}$ is close enough to $E^{(0)}_{\rm tot}$ the Newton method from §2.4 is expected to converge rapidly and to provide us with a new configuration $\boldsymbol{\theta}^{(1)}$, solution of the problem with invariants $(E^{(1)}_{\rm tot}, L_{\rm tot})$. This new solution may then be used as an appropriate initial condition to solve the problem $(E^{(2)}_{\rm tot}, L_{\rm tot})$, getting us closer to our target energy $E_{\rm tot}$. Repeating this iterative process, we can ultimately solve the problem truly at hand, i.e. the one associated with the invariants $(E_{\rm tot}, L_{\rm tot})$. In practice, we generically used a total of 50 different energies $E^{(i)}_{\rm tot}$ spread logarithmically between $E^{(0)}_{\rm tot}$ and the target total energy $E_{\rm tot}$.

### C3 Equilibria for $\beta = 0$

We derive the DF that solves the entropy maximisation problem, without any constraint on energy. This DF is the starting point of our optimisation strategy with small steps in total energy (see §C2).

First, for a given non-zero value of $\mathbf{L}_{\rm tot}$ (taken to be along $+z$), the solution of the optimisation problem without energy constraint must be axisymmetric. Indeed, if $F(\phi, \widehat{L}_z, \mathbf{K})$ satisfies the constraints from equations (7) and (11), then so does $F_\Delta : (\phi, \widehat{L}_z, \mathbf{K}) \to F(\phi + \Delta, \widehat{L}_z, \mathbf{K})$. Given that equation (11) is linear w.r.t. $F$,

$$\overline{F} : (\phi, \widehat{L}_z, \mathbf{K}) \to \int_0^{2\pi} \frac{{\rm d}\Delta}{2\pi}\, F_\Delta(\phi, \widehat{L}_z, \mathbf{K}) = \overline{F}(\widehat{L}_z, \mathbf{K}) \quad \text{(C2)}$$

also meets the constraints on $\mathbf{L}_{\rm tot}$. We note that $\overline{F}$ is a barycentre of the phase-shifted DFs, which all have the same entropy (equation 13). Given that $s \colon x \mapsto x \ln(x)$ is a convex function, Jensen's inequality gives $S[\overline{F}] \geqslant S(F)$. As such, in the absence of any constraint on energy, the entropy maximum must be axisymmetric.

Given this axisymmetry and remembering that $\beta = 0$ since there is no constraint on the energy, equation (14) becomes

$$F_{\rm eq}(\widehat{\mathbf{L}}, \mathbf{K}) = N(\mathbf{K})\, \frac{\exp\big[L(\mathbf{K})\, \gamma\, \widehat{L}_z\big]}{\displaystyle\int {\rm d}\widehat{\mathbf{L}}'\, \exp\big[L(\mathbf{K})\, \gamma\, \widehat{L}'_z\big]}. \quad \text{(C3)}$$

Equation (C3) is much simpler than equation (14) because its r.h.s. does not contain $F_{\rm eq}$ anymore, i.e. the equation becomes explicit.

To find a value of $\gamma$ for which $F_{\rm eq}$ meets the constraint on $L_{\rm tot}$, we compute the integral on the r.h.s. of equation (11). We get

$$\begin{aligned} L(\mathbf{K}) \int {\rm d}\widehat{\mathbf{L}}\, \widehat{L}_z\, \exp\big[L(\mathbf{K})\, \gamma\, \widehat{L}_z\big] &= 2\pi\, L(\mathbf{K}) \int_{-1}^{+1} {\rm d}z\, z\, {\rm e}^{L(\mathbf{K})\gamma\, z} \\ &= \frac{4\pi}{\gamma} \big\{ \cosh[L(\mathbf{K})\, \gamma] - {\rm sinch}[L(\mathbf{K})\, \gamma] \big\}, \end{aligned} \quad \text{(C4)}$$

and

$$\begin{aligned} \int {\rm d}\widehat{\mathbf{L}} \exp\big[L(\mathbf{K})\, \gamma\, \widehat{L}_z\big] &= 2\pi \int_{-1}^{+1} {\rm d}z\, {\rm e}^{L(\mathbf{K})\gamma\, z} \\ &= 4\pi\, {\rm sinch}[L(\mathbf{K})\, \gamma], \end{aligned} \quad \text{(C5)}$$

with ${\rm sinch}(x) = \sinh(x)/x$. Therefore, equation (11) finally gives

$$L_{\rm tot}(\gamma) = \int {\rm d}\mathbf{K}\, N(\mathbf{K}) \big\{ L(\mathbf{K})\, {\rm cotanh}[L(\mathbf{K})\, \gamma] - (1/\gamma) \big\}. \quad \text{(C6)}$$

This is a strictly increasing function of $\gamma$, so that there exists at most one solution $\gamma$ for a given value of $L_{\rm tot}$. Although not analytical, such a solution is straightforward to obtain by dichotomy.

## APPENDIX D: NUMERICAL APPLICATIONS

We briefly detail some of our choices in the effective numerical implementation of the method of entropy optimisation from §2.4 We also validate it in single-population systems by reproducing previous published results.

### D1 Discretisation of the stellar populations

As explained in §2.4, in order to be effectively implemented the entropy optimisation requires a discretisation of the stellar populations. Each population is associated with an index $k$ and is characterised by some orbital parameters $\mathbf{K}_k$ and a particle number $N_k$.

In §3.1, we assumed that $(m, a, e)$ are drawn independently from one another, so that they may be discretised independently as well. For the stellar mass, the considered range $m_{\rm min} \leqslant m \leqslant m_{\rm max}$ is discretised in $N_m$ logarithmic bins. Similarly, for $a$ we use $N_a$ logarithmic bins, and $N_e$ linear bins for $e$. As a consequence, the effective number of populations is set by $N_{\rm pop} = N_m N_a N_e$. Then, for a given population, the number of particles $N_k$ is simply set by the value of the sampling power-law PDFs in the centre of the bins, multiplied by the volumes of the bins. Similarly, the value of the

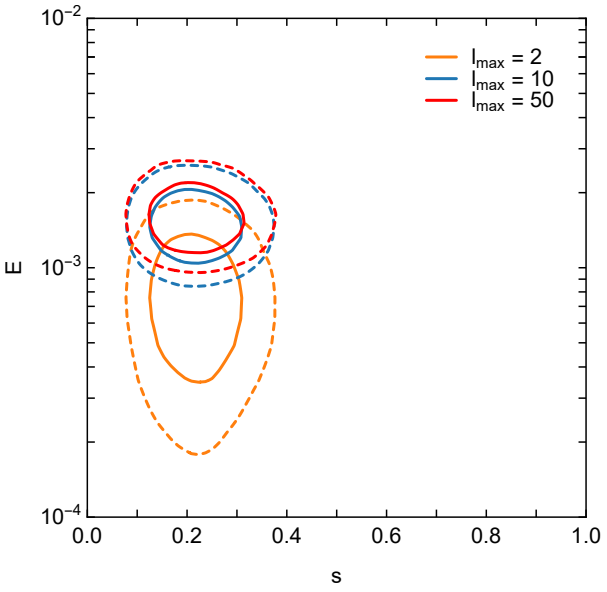


**Figure D1.** Illustration of the dependence of $(E, s)$ as a function of $\ell_{\max}$ for the protocol from §3.1. For a given value of $\ell_{\max}$, the two plotted contours correspond to 33% (dashed) and 66% (solid) of the PDF's maximum. As already pointed out in Takács & Kocsis (2018), for $\ell_{\max} \gtrsim 10$, the values of the two invariants $(E, s)$ can be considered as converged.

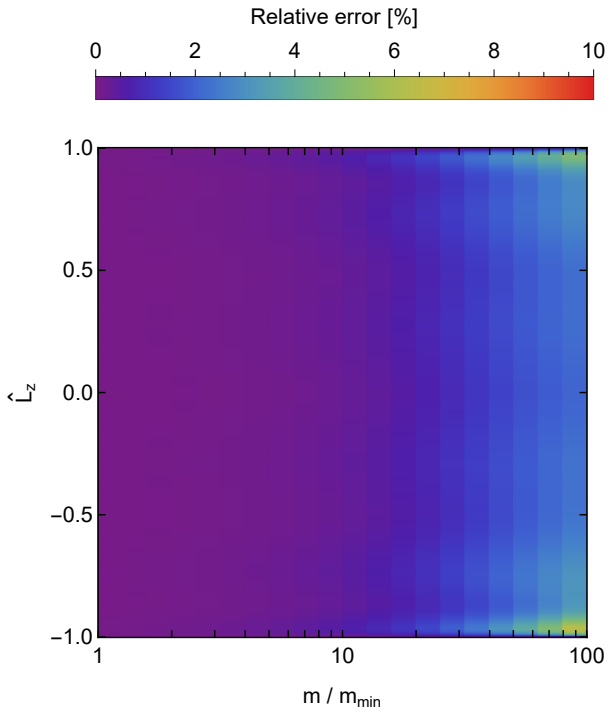


**Figure D2.** Relative error in Fig. 4 as one increases the harmonic truncation from $\ell_{\max} = 10$ to $\ell_{\max} = 50$. The maximum relative error is $\sim 7\%$.

population's orbital parameter, $\mathbf{K}_k$, is set by the value of $\mathbf{K}$ in the centre of the bin. As such, all our integrals over $\mathrm{d}\mathbf{K}$ are formally approximated by Riemann sums using the midpoint rule.

To obtain Figs. 4 and 5, we used $(N_m, N_a, N_e) = (20, 10, 5)$. For the subsequent sections, we considered the simpler regime $(N_m, N_a, N_e) = (10, 10, 3)$. We checked that such choices of population numbers did not affect the results. Given that the numerical complexity scales like $\mathcal{O}(N_{\mathrm{pop}}^3)$, this greatly eases the overall parameter exploration.

### D2 Convergence w.r.t. $\ell_{\max}$

In practice, the Hamiltonian from equation (2) has to be truncated to some finite harmonic order $\ell_{\max}$. While Roupas et al. (2017) limited themselves to the quadrupolar case $\ell_{\max} = 2$, Takács & Kocsis (2018) lifted this restriction in single-population systems and showed that an effective truncation at $\ell_{\max} = 10$ is sufficient to get converged results. This is what we briefly explore for multi-population clusters in this section.

First, in Fig. D1, considering the same initialisation protocol as in §3.1, we illustrate the dependence of the invariants $(E, s)$ as a function of $\ell_{\max}$. In that figure, we recover that the typical value of $E$ increases with $\ell_{\max}$, while, of course, the average value of $s$ is independent of it. In addition, we note that restricting oneself to $\ell_{\max} = 2$ seems insufficient, while $\ell_{\max} = 10$ offers reasonably well converged values of the binding energy.

In addition to affecting the values of the cluster's invariants, increasing the value of $\ell_{\max}$ might also impact the overall shape of the reconstructed anisotropic equilibria DF. This is what we explore in Fig. D2, where we compute the relative error in the reconstructed DF between $\ell_{\max} = 10$ and $\ell_{\max} = 50$. Given that the maximum relative error is $\sim 7\%$ for the present clusters, we systematically truncated the pairwise interaction at $\ell_{\max} = 10$ in all the figures presented in the main text. As the complexity of the entropy optimisation scales like $\mathcal{O}(\ell_{\max}^3)$, this significantly alleviates the numerical difficulty of the computations.

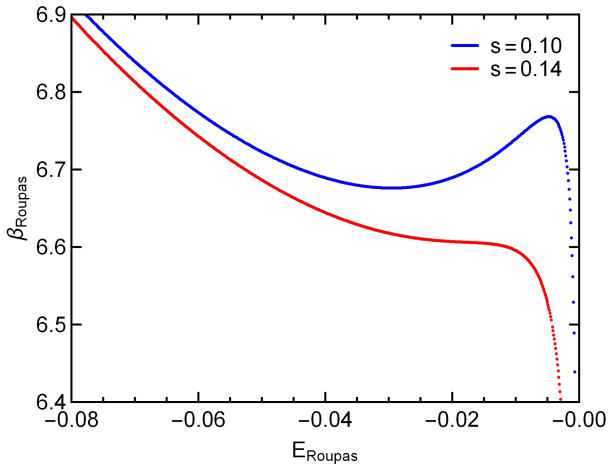


**Figure D3.** Inverse temperature $\beta$ as a function of the energy $E_{\mathrm{Roupas}}$, for various fixed total angular momentum, $s$, for a single-population cluster with $\ell_{\max} = 2$. This figure reproduces fig. 11 of Roupas et al. (2017). For low $s$, the cluster can exhibit a negative specific heat, i.e. $\partial\beta_{\mathrm{Roupas}}/\partial E_{\mathrm{Roupas}} > 0$.

### D3 Validation in single-population systems

We validate our implementation of entropy maximisation by recovering previous results from the literature.

First, in Fig. D3, we recover the caloric curve presented in fig. 11 of Roupas et al. (2017) for a single-population system interacting only through $\ell = 2$. In order to exactly match our present

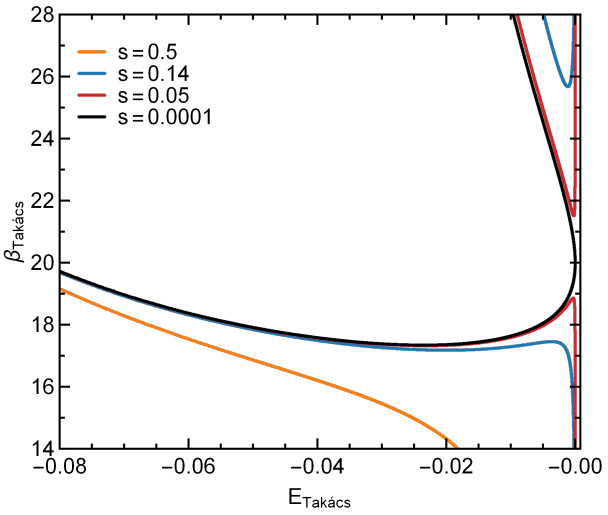


**Figure D4.** Same as Fig. D3 but considering $\ell_{\rm max} = 58$. This figure reproduces fig. 4 of Takács & Kocsis (2018). In addition to the presence of negative specific heats for low $s$, we recover for weakly bound clusters the presence of more than one equilibrium solutions. Configurations with the smallest $\beta$ were always found to have the highest entropy, hence correspond to the cluster's (axisymmetric) global statistical equilibria.

normalisation convention with the ones from Roupas et al. (2017), one has to consider

$$\beta_{\rm Roupas} = \beta\, \tfrac{3}{8} N G m^2/a; \;\; E_{\rm Roupas} = E_{\rm tot}/\big(\tfrac{3}{8} N^2 G m^2/a\big). \quad \text{(D1)}$$

In this figure, it is interesting to recall that clusters with sufficiently small total angular momentum can present a negative specific heat.

In Fig. D4, we recover fig. 4 from Takács & Kocsis (2018), for a single-population system coupled beyond the quadrupolar interaction. To exactly match the normalisation conventions from Takács & Kocsis (2018), one has to consider

$$\beta_{\rm Takács} = \beta\, N G m^2/a; \quad E_{\rm Takács} = E_{\rm tot}/\big(N^2 G m^2/a\big). \quad \text{(D2)}$$

In addition, following equation (2) of Takács & Kocsis (2018), we also had to replace the coupling coefficients from equation (4) with the simpler asymptotic scaling

$$\mathcal{H}_\ell^{\rm Takács} = \frac{G m^2}{a} \frac{4\pi}{\ell^2\,(2\ell+1)}. \quad \text{(D3)}$$

For sufficiently small values of the total spin $s$, one recovers more than one equilibrium solutions. The solutions with the smallest $\beta$, i.e. the branch that goes through $\beta = 0$, were always found to have the largest entropy, hence corresponding to the (axisymmetric) thermodynamical equilibria.