Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **541,** 1145–1163 (2025)
Advance Access publication 2025 June 24

https://doi.org/10.1093/mnras/staf1018

Infall-driven gravitational instability in accretion discs

Cristiano Longarini [1]★ Daniel J. Price [2] Kaitlin M. Kratter,[3] Giuseppe Lodato [4] and Cathie J. Clarke[1]

[1]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*
[2]*School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia*
[3]*Department of Astronomy and Steward Observatory, University of Arizona, 933 N Cherry Ave, Tucson, AZ, 85721, USA*
[4]*Dipartimento di Fisica, Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy*

Accepted 2025 June 16. Received 2025 June 12; in original form 2025 May 9

ABSTRACT

Gravitational instability (GI) is typically studied in cooling-dominated discs, often modelled using simplified prescriptions such as β-cooling. In this paper, we investigate the onset and evolution of GI in accretion discs subject to continuous mass injection, combining 1D and 3D numerical simulations. We explore an alternative self-regulation mechanism in which mass replenishment drives the system toward marginal stability $Q \sim 1$. In this regime, the disc establishes a steady-state disc-to-star mass ratio, balancing the mass transported to the central object with that added to the disc. Our 3D simulations reveal that the general scaling predicted from the linear theory are respected, however there are important difference compared to the cooling case in terms of morphology and pattern speed. Unlike the flocculent spirals seen in cooling-driven instability, the power is concentrated towards the dominant modes in infall-driven spirals. Additionally, spiral waves generate at the mass injection location, and propagate at constant pattern speed, unlike in the cooling case. This suggests a fundamental difference in how mass-regulated and cooling-regulated discs behave and transport angular momentum.

Key words: accretion, accretion discs – planets and satellites: formation—hydrodynamics – instabilities – gravitation.

1 INTRODUCTION

Young protoplanetary discs are likely to be massive, as most of their mass has not yet been accreted by the central object. Observations support this, as seen in the cumulative mass function of protostellar discs across different star-forming regions and ages (Tobin et al. 2020). Young and massive discs are likely to be gravitationally unstable (Kratter & Lodato 2016). The instability threshold for axisymmetric perturbations is determined by the dimensionless Toomre parameter (Safronov 1960; Lin & Shu 1964; Toomre 1964), according to

$$Q = \frac{c_s \kappa}{\pi G \Sigma}, \tag{1}$$

where $c_{\rm s}$ is the sound speed, Σ the disc surface density and κ the epicyclic frequency, with $Q = 1$ the instability threshold.

In numerical simulations, gravitational instability is usually triggered by cooling the disc. The most widely used cooling prescription is the simple β-cooling model, in which the disc cools on a time-scale proportional to the local dynamical time, with a proportionality factor β (Gammie 2001). For cooling-driven gravitational instability, the system can reach a self-regulated thermal state. An initially hot and stable disc ($Q > 1$) cools according to the β-cooling prescription until it reaches the instability threshold (Q = 1). At this point, gravitational instability sets in, triggering the formation of spiral waves. These waves heat the disc through shocks and increase the Q-parameter. Eventually, the heating from the spiral arms balances the imposed cooling, leading to a self-regulated state. In this scenario, gravitational instability acts as a *thermostat*, maintaining the disc around $Q = 1$.

The $\beta-$cooling model has proved an effective sandbox for studying self-gravitating discs, since one is able to study gravitationally unstable discs in a quasi-steady, equilibrium state. This has enabled detailed study of the angular momentum transport (Lodato 2007; Cossins, Lodato & Clarke 2009), dust accumulation (Rice et al. 2004, 2006; Booth & Clarke 2016; Baehr & Zhu 2021; Longarini et al. 2023; Rowther et al. 2024a), kinematic perturbations (Hall et al. 2020; Longarini et al. 2021; Terry et al. 2022). The main problem is that the $\beta-$cooling prescription does not account for heating by the central star. A number of studies have attempted to account for irradiation in discs (Boss 2002; Mejia 2004; Boley et al. 2007; Stamatellos et al. 2007; Forgan et al. 2009; Young et al. 2024). In particular, Rice et al. (2011) and more recently Leedham, Booth & Clarke (2025,) have shown that thermal self-regulation can be ineffective in strongly irradiated discs, whose behaviour is similar to polytropic models originally studied by Laughlin & Rozyczka (1996). Such studies have been confirmed in the global simulations with the live radiative transfer of Rowther et al. (2024b), who finds negligible shock heating in their irradiated discs.

In the context of gravitational instability, potential candidates include Elias 2–27 (Pérez et al. 2016; Longarini et al. 2024), AB Aurigae (Fukagawa et al. 2004; Tang et al. 2012), and GM Aur (Martire et al. 2024). Interestingly, these sources show signs of infall (Paneque-Carreño et al. 2021; Schwarz et al. 2021; Speedie et al. 2024). Adding mass to a protoplanetary disc can be an alternative way to trigger gravitational instability (Kratter, Matzner & Krumholz

---

★ E-mail: cl2000@cam.ac.uk

2008; Kratter et al. 2010). Considerable theoretical effort has been devoted to understanding the dynamics and consequences of infall on protoplanetary discs (e.g. Cameron 1978; Lin & Pringle 1990; Kenyon, Calvet & Hartmann 1993; Bate, Bonnell & Bromm 2003; Kratter et al. 2008; Zhu, Hartmann & Gammie 2010; Winter, Benisty & Andrews 2024b; Calcino et al. 2025). Observations indicate that this process remains significant even in evolved stages of protoplanetary discs lifetime, suggesting the relevance of late infall (Gupta et al. 2023, 2024; Winter et al. 2024a). Adding mass from the environment offers an alternative way for a disc to self-regulate. In the absence of other instabilities and with a fixed temperature profile, a disc will passively accept mass until the $Q = 1$ threshold is reached. Then, spiral arms will form and angular momentum transport will start, leading to a balance between the mass infalling from the environment, and mass being drained onto the star, decreasing Σ and hence increasing Q.

Our aim in this paper is to explore whether a steady state can be reached in discs that are not isolated from their environment. In particular, we use a combination of 1D and 3D simulations to examine how the long term behaviour of infall-regulated gravitationally unstable discs differs from their temperature-regulated counterparts.

The paper is organized as follows: in Section 2, we present an analytical framework to understand how self-regulation operates in discs undergoing infall; Section 3 describes the numerical framework including 1D and 3D simulations. Section 4 presents the numerical results, and a comparison between the two approaches. In Section 5 we discuss possible observational signatures and possible correlations in the initials condition for planet formation. While our focus is on protostellar discs, the methods, scalings and conclusions are general and can be applied to self-gravitating accretion discs at any scale, such as those around supermassive black holes.

2 GRAVITATIONAL INSTABILITY AND THE ROLE OF INFALL

The linear evolution of gravitational instability is described by the dispersion relation (Safronov 1960; Lin & Shu 1964; Toomre 1964)

$$(\omega - m\Omega)^2 = c_s^2 k^2 - 2\pi G\Sigma|k| + \kappa^2, \tag{2}$$

where ω is the perturbation frequency, m is the azimuthal wavenumber, and κ is the epicyclic frequency, defined as

$$\kappa^2 = \frac{2\Omega}{R}\frac{\mathrm{d}}{\mathrm{d}R}\left(\Omega R^2\right), \tag{3}$$

being equal to the orbital frequency Ω for a Keplerian disc. This relation is valid in the WKB approximation, which requires the radial wavelength of the perturbation to be much smaller than the azimuthal one, i.e. $|kR| \gg m$. The stability of axisymmetric perturbation is controlled by the Toomre parameter, as defined in equation (1), with instability occurring when $Q \leq 1$. In an unstable disc, the most unstable mode corresponds to a radial wavenumber

$$k_{\max} = \frac{\pi G\Sigma}{c_s^2}. \tag{4}$$

For a marginally unstable disc $Q = 1$, a spiral perturbation with a radial wavenumber $k = k_{\max}$ satisfies the dispersion relation in equation (2) with the left-hand side equal to zero. This means that the perturbation satisfies $\omega = m\Omega$, leading to a pattern speed

$$\Omega_{\mathrm{p}} = \frac{\omega}{m} = \Omega. \tag{5}$$

Under these conditions, we expect spiral density waves to generate at co-rotation, and have an instantaneous pattern speed of the order of the Keplerian frequency.

In general, for a given pattern speed Ω_{p}, it is possible to define the Lindblad resonances, that are the location where the gravitational perturbation of the spiral matches the natural epicyclic frequency of the disc material, and the corotation resonance, where the pattern speed matches the angular frequency of the disc. In the case of a constant pattern speed Ω_{p} in a Keplerian disc, the radii of the inner and outer Lindblad resonances are (e.g. Ogilvie & Lubow 2002))

$$R_{\mathrm{ILR,OLR}} = R_{\mathrm{cr}}\left(1 \pm \frac{1}{m}\right)^{2/3}. \tag{6}$$

In non-axisymmetric self-gravitating discs, torques induced by gravitational perturbations transport both energy and angular momentum. Understanding how this happens requires studying the non-linear evolution of gravitational instability. The ability of self-gravitating spirals to transport angular momentum was first identified by Lynden-Bell & Kalnajs (1972) in the context of galactic dynamics. In protostellar discs, Balbus & Papaloizou (1999) computed the stress tensor associated with gravitational potential perturbations, noting that energy transport via gravitational instability cannot be fully described by a local viscous approximation – except at co-rotation. This can be understood considering the spiral wave energy per unit surface $\mathcal{E}_{\mathrm{w}}$, which reads

$$\mathcal{E}_{\mathrm{w}} = \frac{\Sigma}{2}\frac{m^2}{k^2}(\Omega - \Omega_{\mathrm{p}})^2\left(\frac{\delta\Sigma}{\Sigma}\right)^2 + \frac{\Sigma}{2}\frac{m^2}{k^2}\Omega(\Omega - \Omega_{\mathrm{p}})\left(\frac{\delta\Sigma}{\Sigma}\right)^2. \tag{7}$$

The second term – given by the angular momentum per unit area multiplied by the rotation speed – represents a local energy transport process, and can therefore be modelled using the standard α prescription. In contrast, the first term – equal to the same angular momentum term times $\Omega - \Omega_{\mathrm{p}}$ – is a non-local term, and prevents GI discs from behaving as pure α–discs. The degree of non-local angular momentum transport is ultimately determined by the disc-to-star mass ratio (e.g. Laughlin & Bodenheimer 1994; Laughlin & Rozyczka 1996; Laughlin, Korchagin & Adams 1997, 1998), with more massive discs globally redistributing angular momentum within the disc.

2.1 Mass controlled self-regulation

An initially stable disc $Q \gg 1$ does not show any spiral density wave. To trigger gravitational instability, we can either cool the disc, i.e. decrease its sound speed, or add mass, i.e. increase the surface density. In the case of infall-driven GI, the instability is triggered by mass injection, and self-regulation is obtained in terms of mass, rather than temperature. The self-regulated state is characterized by a stationary $Q \approx 1$ profile, meaning that the time derivative of the Toomre parameter is equal to zero. This implies that the injected mass is distributed between the central object and the disc, to ensure that $\dot{Q} = 0$. It is possible to obtain the balance mentioned above as follows: denoting $\dot{M}_{\mathrm{inj}}$ as the mass injection rate, $\dot{M}_\star$ as the mass accretion rate onto the central object, and $\dot{M}_{\mathrm{d}}$ the rate at which the disc mass is increasing, mass conservation implies that

$$\dot{M}_{\mathrm{inj}} = \dot{M}_\star + \dot{M}_d. \tag{8}$$

As mentioned above, the evolution is such that the time derivative of the Toomre parameter is zero. The Toomre parameter can be re-written as

$$Q = A \frac{H}{R} \frac{M_\star}{M_\mathrm{d}}, \tag{9}$$

where A is a constant of order unity and $H \equiv c_\mathrm{s}/\Omega$ is the pressure scale height. As a working hypothesis, we suppose that the disc is locally isothermal, hence the temperature does not change with time, and we also suppose that the disc scale height is constant $H/R \sim$ const. Within these hypotheses, the steady state condition is

$$\dot{Q} = 0 \rightarrow \frac{\dot{M}_\mathrm{d}}{M_\mathrm{d}} = \frac{\dot{M}_\star}{M_\star}. \tag{10}$$

Combining equations (8) and (10), the evolution of the stellar and disc mass is given by

$$\dot{M}_\star = \frac{\dot{M}_\mathrm{inj}}{1+q}, \tag{11}$$

and

$$\dot{M}_\mathrm{d} = \dot{M}_\mathrm{inj} \frac{q}{1+q}, \tag{12}$$

where $q = M_\mathrm{d}/M_\star$. These are the key characteristics of self-regulation: mass is injected into the disc at a specific location and at a given rate, and it is redistributed between the star and the disc to keep their ratio constant over time. To achieve this, equations (11) and 12 must be respected.

Since the problem is completely rescalable in terms of star mass, disc mass and injection rate, it is useful to define a dimensionless injection rate, which does not depend on the physical scale of the system. The dimensionless injection rate $\dot{\mu}$ is defined as

$$\dot{\mu} = \frac{\dot{M}_\mathrm{inj}}{M_\star \Omega_\mathrm{inj}}. \tag{13}$$

where $\Omega_\mathrm{inj} = \Omega(R_\mathrm{inj})$, being R_inj being the injection radius. According to the stationary solution for a self-gravitating accretion disc (Bertin 1997; Bertin & Lodato 1999; Lodato 2007), the mass accretion rate is related to the α, i.e. the efficiency of angular momentum transported throughout the disc, according to

$$\dot{M}_\star = \frac{2\alpha c_s^3}{G} \left| \frac{\mathrm{d}\log\Omega}{\mathrm{d}\log \mathrm{R}} \right| = \frac{3\alpha c_s^3}{GQ}, \tag{14}$$

where in the second equality we assumed that the disc is Keplerian. Hence, using the self-regulation condition of equation (11) the α viscosity can be written as

$$\alpha = \frac{\dot{M}_\mathrm{inj}}{M_\star \Omega} \frac{1}{3(H/R)^3(1+q)} = \frac{\dot{\mu}}{3(H/R)^3(1+q)}. \tag{15}$$

3 NUMERICAL METHODS

In this section, we present the numerical framework of this paper. We describe the 1D and the 3D codes, together with the set of simulations we performed.

3.1 1D grid evolution code

Table 1. Parameters of the 1D simulations: injection rate $\dot{M}_\mathrm{inj}$, dimensionless injection rate, injection location $\hat{R}_\mathrm{inj}$ and aspect ratio H/R.

Simulation	$\dot{\mathrm{M}}_\mathrm{inj}$ ($\mathrm{M}_\odot\,\mathrm{yr}^{-1}$)	$\dot{\mu}$	$\hat{R}_\mathrm{inj}$	H/R
S1D_1 – reference	5×10^{-5}	2.5×10^{-4}	1	0.1
S1D_2	5×10^{-5}	2.5×10^{-4}	1	0.05
S1D_3	5×10^{-5}	2.5×10^{-4}	1	0.15
S1D_4	10^{-6}	5×10^{-6}	1	0.1
S1D_5	10^{-5}	5×10^{-5}	1	0.1
S1D_6	10^{-4}	5×10^{-4}	1	0.1
S1D_7	5×10^{-5}	2.5×10^{-4}	0.5	0.1
S1D_8	5×10^{-5}	2.5×10^{-4}	0.2	0.1

We evolve the surface density of an accretion disc undergoing mass injection with a 1D evolution code[1]. The diffusion equation for the surface density of a Keplerian disc with mass injection is (Lynden-Bell & Kalnajs 1972; Lin & Pringle 1990)

$$\frac{\partial \Sigma}{\partial t} = \frac{3}{R}\frac{\partial}{\partial R}\left[\sqrt{R}\frac{\partial}{\partial R}\left(\sqrt{R}\Sigma\nu\right)\right] + \dot{\Sigma}_\mathrm{inj}, \tag{16}$$

where $\dot{\Sigma}_\mathrm{inj}$ is a source term describing the mass injection and ν is the kinematic viscosity. For simplicity, we suppose that mass is added at a given radial location in the disc R_inj with a constant mass injection rate $\dot{M}_\mathrm{inj}$, according to

$$\dot{\Sigma}_\mathrm{inj} = \frac{\dot{M}_\mathrm{inj}}{2\pi R_\mathrm{inj}} \delta(R - R_\mathrm{inj}). \tag{17}$$

To solve equation (16) numerically, we follow Bath & Pringle (1981). Details can be found in Appendix A. To mimic the effect of gravitational instability, we force the kinematic viscosity to switch on only when the Toomre parameter Q is below a given threshold Q_c, according to (Lin & Pringle 1987, 1990)

$$\nu_\mathrm{GI} = \begin{cases} \eta\left(\frac{Q_c^2}{Q^2} - 1\right)\left(\frac{c_s^2}{\Omega}\right), & Q \leq Q_\mathrm{c} \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

We choose $Q_c = 1$ and $\eta = 0.5$; we tested different values of this parameter $\eta = 0.1$, 1, and we found that the results do not change, as found by Lodato & Natarajan (2006), who used the same picture of infall dominated accretion to explain the formation of supermassive black hole seeds in the early Universe.

Since in the simulation the star mass evolves because of accretion, at each time-step we update its value and as a consequence Ω, Q, and ν.

We tested different viscosity prescription for gravitationally unstable discs (Kratter et al. 2008; Rafikov 2015). The results are presented in Appendix B. The take home message is that the overall evolution is not impacted when using more refined viscosity models.

3.1.1 *Numerical set-up*

We perform simulations assuming a locally isothermal disc with the sound speed c_s constant with time and with radius. We explore the parameters $\dot{M}_\mathrm{inj}$, H/R, and $\hat{R}_\mathrm{inj} = R_\mathrm{inj}/R_\mathrm{out}$. Table 1 lists the values used in each simulation. We initialize a disc with a surface density profile $\Sigma \propto R^{-1.5}$, initially extending from $R_\mathrm{in} = 1$ to $R_\mathrm{out} = 10$ in code units. The linear radial grid extends from 1 to 10^3 in code units, to allow the disc to expand without reaching the outer boundary. The resolution is set by the number of radial zones, that is 2×10^3. The initial disc mass is chosen so that the Toomre parameter at the injection radius is initially 1.1. The inner and outer boundary conditions are set so that $\Sigma(1) = \Sigma(10^3) = 0$. We evolved the simulations for 5 injection times, where we define the injection time

---

[1] https://github.com/crislong/discfusion

as the star mass doubling time

$$t_{\rm inj} \equiv \frac{M_{\star,0}}{\dot{M}_{\rm inj}}. \qquad (19)$$

Table 1 describes the set of simulations we performed, where **S1D_1** is the reference simulation. Starting from its parameters, we vary the aspect ratio, injection rate and injection location.

In the following plots, we will show the evolution up to 2 injection times, since the results afterwards do not change.

3.2 3D SPH simulations

We perform a suite of 3D smoothed particle hydrodynamics (SPH; Gingold & Monaghan 1977; Lucy 1977; Price 2012) simulations of gas discs using the code PHANTOM (Price et al. 2018). The self-gravity in PHANTOM is described in the code paper, with the algorithm for softening the short range interaction described in Price & Monaghan (2007). This code has been widely used in the astrophysics community to study gas dynamics in accretion discs (e.g. Lodato & Price 2010; Cuello et al. 2019; Ragusa et al. 2020; Nealon et al. 2022; Ceppi et al. 2023), and recently has been used for studying gravitationally unstable discs (e.g. Hall et al. 2020; Cadman et al. 2022; Rowther, Nealon & Meru 2022; Longarini et al. 2023). To date, gravitational instability has been numerically triggered by cooling the disc, using the simple $\beta-$prescription (Gammie 2001), or variations thereof (Stamatellos et al. 2007; Young et al. 2024). In this work, we trigger gravitational instability by adding mass with a chosen angular momentum to a disc with a prescribed vertically and locally isothermal temperature profile. While previous numerical simulations have explored infall-driven gravitational instability (Kratter et al. 2010), these studies focused on simulating the collapse of a quasi-spherical core mediated by an accretion disc. In contrast, we consider a more controlled experiment with two free parameters by injecting SPH particles directly into the simulation, without modelling the collapse of a gas sphere.

We inject SPH particles at the injection radius $R_{\rm inj}$ with a constant mass injection rate $\dot{M}_{\rm inj}$, at Keplerian velocity $v_{\rm inj} = \sqrt{GM_\star/R_{\rm inj}}$ and distributed vertically according to a Gaussian distribution with standard deviation equal to the pressure scale height H, which we choose by setting the aspect ratio $(H/R)_{\rm inj}$ to match that of the disc at the injection radius. To minimize asymmetries during the injection process, which could cause a secular displacement of the system's centre of mass, we add particles in pairs with symmetric positions relative to the central object. The injection module is publicly available in PHANTOM v2025.0.0 or higher[2] using the `isosgdisc` set-up.

3.2.1 Numerical set-up

3D hydrodynamical simulations with mass injection are computationally expensive, as the resolution increases over time due to the continuous addition of particles, and the typical time-scale – the injection time-scale – is long. To address this, we first verify that 1D simulations accurately reproduce the global quantities $(\dot{M}_\star, \dot{M}_d, \dot{q})$. Consequently, we rely on the 1D code for a comprehensive exploration of the parameter space. To complement this, we conduct three PHANTOM simulations with varying aspect ratios to perform a Fourier analysis of the spiral modes. Additionally, we run a simulation with cooling rather than infall, for morphological comparison. Table 2 summarizes the set of simulations.

Table 2. Parameters of the 3D simulations: injection rate, dimensionless injection rate, injection location and aspect ratio.

Simulation	$\dot{\rm M}_{\rm inj}$ (M$_\odot$ yr^{-1})	$\dot{\mu}$	$\hat{R}_{\rm inj}$	H/R
S3D_1—reference	5×10^{-5}	2.5×10^{-4}	1	0.1
S3D_2	1.25×10^{-5}	5×10^{-5}	1	0.075
S3D_3	1×10^{-4}	5×10^{-4}	1	0.125
Cool_0325	$\beta = 10$	$q = 0.325$		

We initially set up the disc around a central sink particle with mass $M_\star = 1$ in code units and accretion radius $R_{\rm acc} = 1$. The disc extends between $R_{\rm in} = 1$ and $R_{\rm out} = 10$, with a power law surface density $\Sigma \propto R^p$, with $p = -1.5$. We assume that the disc is locally isothermal, with the sound speed constant with time and radius. We set the initial disc mass by choosing the initial value of Q to be 1.1 at the injection radius. The setup is the same as the 1D simulations, and so it is the reference simulation (see Table 2). In the simulations we are not using the disc–viscosity flag (Lodato & Price 2010), meaning that the viscosity is not described by an $\alpha_{\rm SS}$ coefficient. We did so since in these systems the main driver of angular momentum transport is gravitational instability through spiral arms. We instead apply dissipation only as necessary for shock capturing, with the shock viscosity coefficient $\alpha_{\rm AV} \in [0, 1]$ using the Cullen & Dehnen (2010) switch as described in Price et al. (2018) and $\beta_{\rm AV} = 2$. The initial number of particles is $N_{\rm in} = 5 \times 10^5$, and during the simulations the number significantly increases, reaching a final value of active particles $N_{\rm f} \gtrsim 1 \times 10^6$.

As a comparison with the infall driven GI, we perform an additional simulation where gravitational instability is triggered by cooling **Cool_0325**. We choose the same disc structure of the reference simulation, except for the inner radius that is set $R_{\rm in} = 0.2$. The disc to star mass ratio of the simulation is chosen to match the self-regulated value of the reference simulation, that is $q = 0.325$.

4 RESULTS

4.1 1D simulations

Fig. 1 shows the time evolution of the stellar mass, disc mass, and disc to star mass ratio in the 1D evolution code for the reference simulation **S1D_1**, where time is given in units of injection time. Both the star and disc mass increase with time, eventually reaching a constant disc to star mass ratio. This is particularly evident in the time derivative of these quantities (right panel of Fig. 1). The time needed to achieve self-regulation is of the order of the injection time. In general, we observe that this trend is respected in all the simulations. The right panel of Fig. 1 also shows that $\dot{M}_\star$ and $\dot{M}_{\rm d}$ are in line with the expected values given by equations (11) and (12), where q is chosen as the final one, that are the dashed lines.

Fig. 2 shows the evolution of the surface density, kinematic viscosity, $\alpha-$viscosity, and Toomre Q parameter. As a consequence of the mass injection, the disc mass increases and hence Q decreases. When the $Q = 1$ threshold is hit, the viscosity switches on and the angular momentum is transported throughout the disc. Equation (15) describes how the α-viscosity depends on the injection rate and aspect ratio: α is proportional to the injection rate, and a thinner disc results in more efficient angular momentum transport. Fig. 3 shows radial profiles of the $\alpha-$viscosity for different injection rates (left panel), aspect ratios (central panel) and injection radius (right panel)

---

[2] https://github.com/danieljprice/phantom

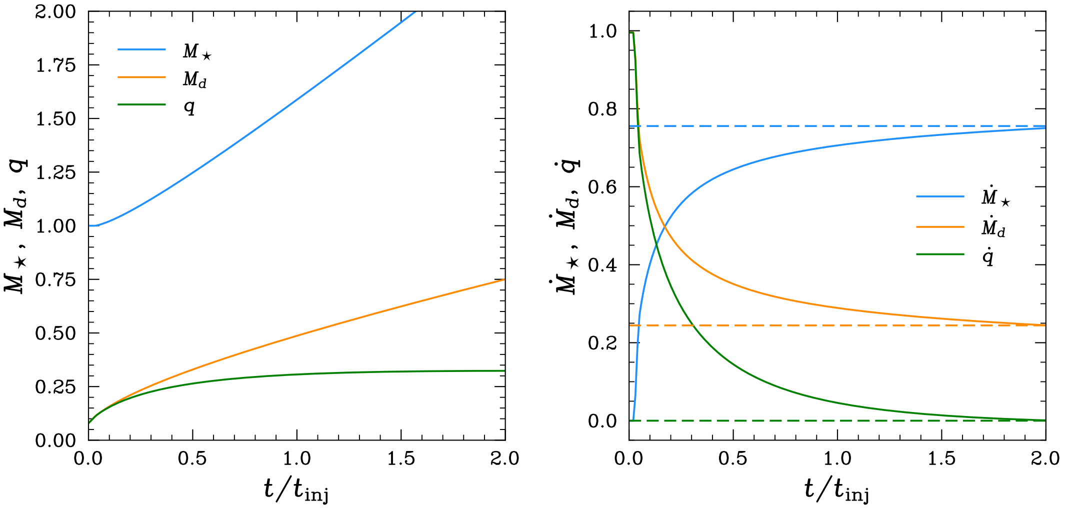

Figure 1. Left panel: time evolution of the stellar mass, disc mass, and disc to star mass ratio as a function of time in units of injection time. Right panel: stellar accretion rate ($\dot{M}_\star$), disc accretion rate ($\dot{M}_{\rm d}$) and time derivative of the disc to star mass ratio ($\dot{q}$) in units of mass injection rate as a function of time in units of injection time. The dashed lines correspond to the steady state prediction: equation (11) for $\dot{M}_\star$, equation (12) for $\dot{M}_{\rm d}$ and zero for $\dot{q}$.

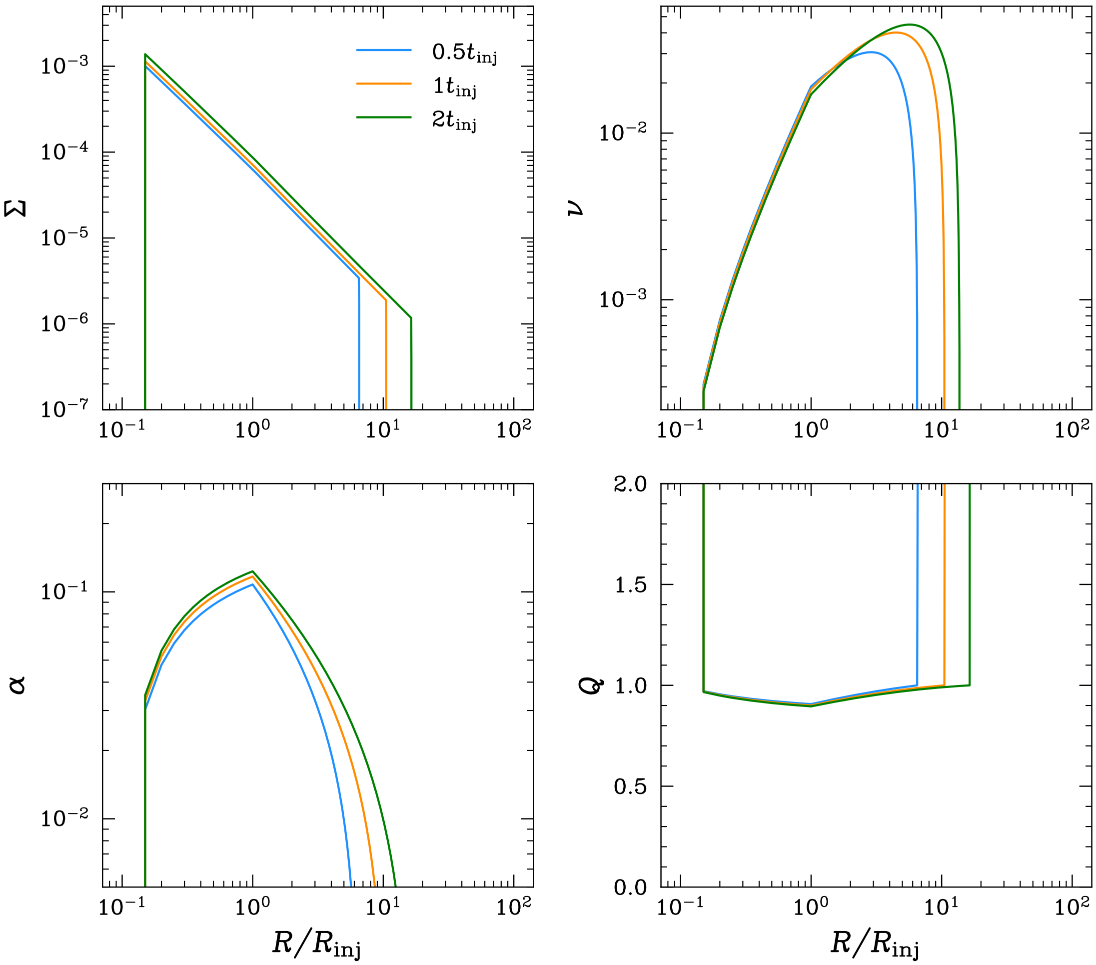

Figure 2. Surface density, kinematic viscosity, alpha viscosity, and Toomre Q parameter at different times for the reference simulation.

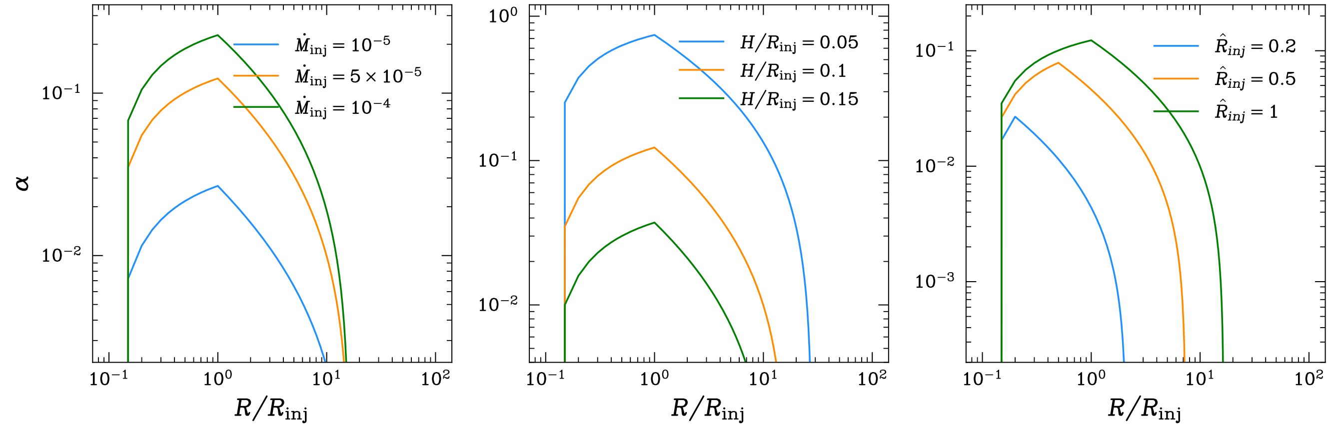

Figure 3. Radial profiles of the α−viscosity for different injection rates (left panel), aspect ratios (central panel), and injection radius (right panel) at $t = 2t_{\rm inj}$.

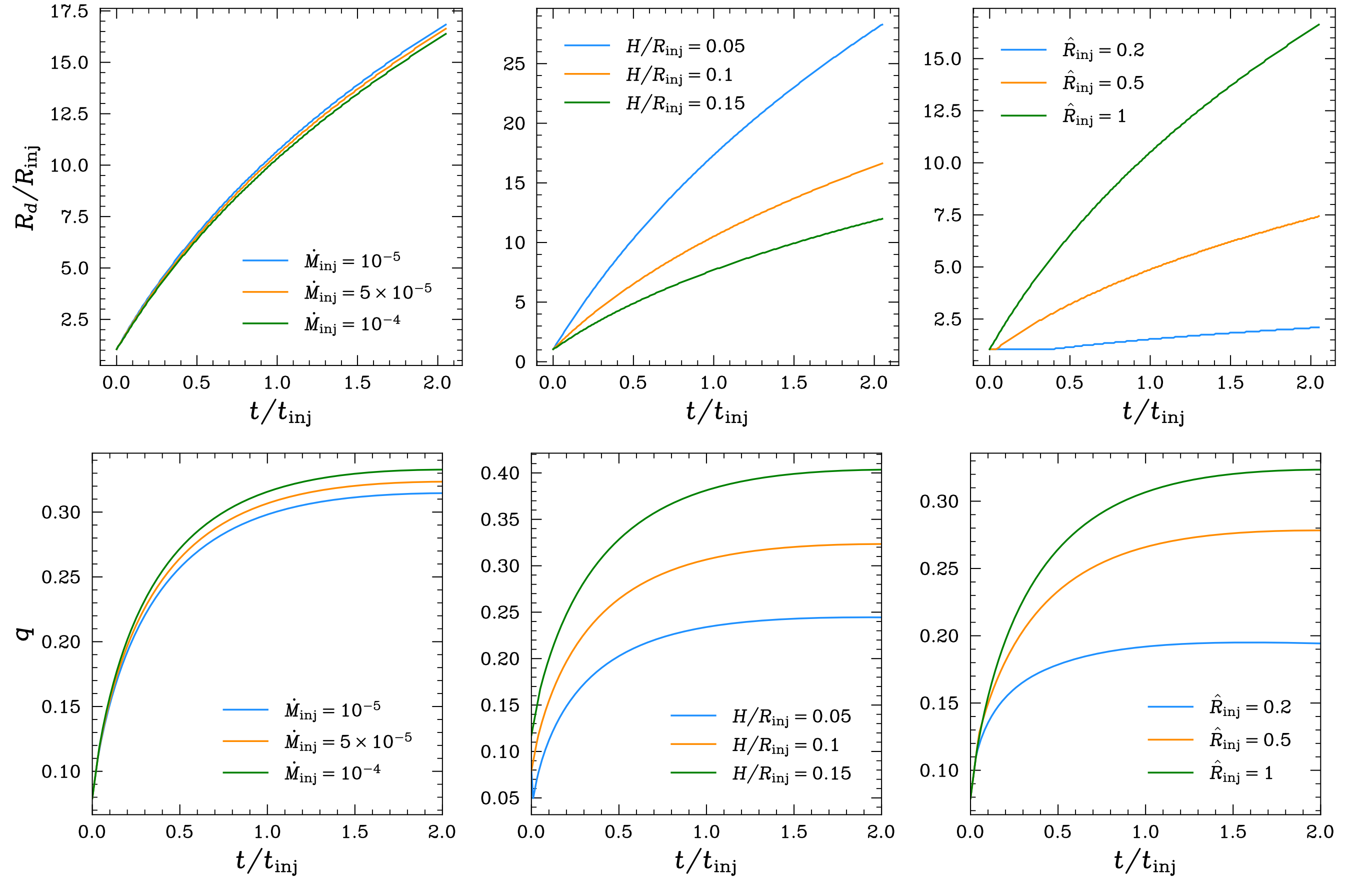

Figure 4. Top row: Time evolution of the outer disc radius for different injection rates (left panel), aspect ratios (central panel), and injection radius (right panel). **Bottom row:** Time evolution of the disc to star mass ratio for different injection rates (left panel), mass aspect ratios (central panel) and injection radius (right panel).

at $t = 2t_{\rm inj}$. The profiles peak at the injection radius, where the mass is injected and hence the Toomre parameter reaches its minimum. The trends with $\dot{M}_{\rm inj}$ and H/R are recovered as expected from equation (15). As for the relationship between α and $\hat{R}_{\rm inj}$, the viscosity is ultimately determined by the value of Ω at the injection radius. Since the Keplerian frequency decreases with radius, we expect a larger α for a greater $\hat{R}_{\rm inj}$, as observed in Fig. 3.

As a result of angular momentum transport, the disc undergoes viscous expansion, as illustrated in the top-left panel of Fig. 2. The extent of this spreading depends on the viscosity, which is influenced by the aspect ratio, mass injection rate, and injection radius. The top row of Fig. 4 shows the time evolution of the disc's outer radius as a function of the injection rate (left panel), aspect ratio (central panel), and injection radius (right panel). These trends reflect the dependence of viscosity on these parameters.

As described in the previous section, the self-regulated state is characterized by having a constant disc to star mass ratio q, that ensures that the Toomre parameter $Q \approx 1$ throughout the disc. The

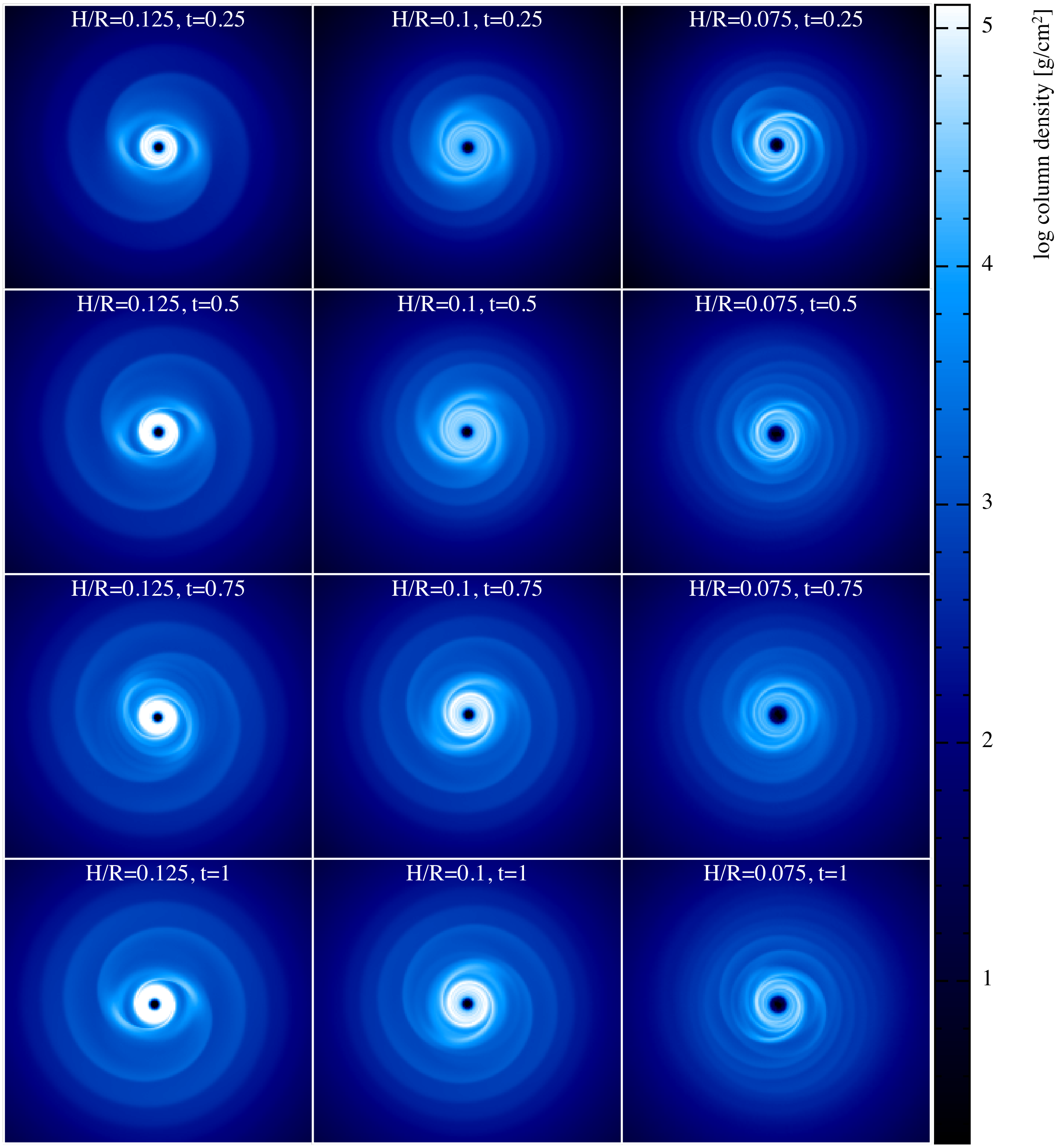

Figure 5. *3D SPH simulations of infall driven gravitational instability.* Snapshots of the logarithmic surface density of SPH simulations with $H/R = 0.05$ (left column), $H/R = 0.1$ (left column) and $H/R = 0.15$ (left column) at $t = 0.25t_{\rm inj}$ (first row), $t = 0.5t_{\rm inj}$ (second row), $t = 0.75t_{\rm inj}$ (third row), $t = t_{\rm inj}$ (fourth row).

bottom row of Fig. 4 shows how the self-regulated q depends on the simulation parameters. Although the injection rate does not have much impact on the equilibrium q, the aspect ratio does, as it determines the amount of mass needed to reach the $Q = 1$ state. Thicker and hotter discs show a higher disc to star mass ratio, as expected. In addition, the location of the injection plays a role in determining the equilibrium disc to star mass ratio, because it determines the region of the disc that must be in a $Q = 1$ state, to transport angular momentum.

The 1D simulations are useful to understand the main trends, since they are computationally cheap and easy to interpret, however they do not capture the complexity of the full three-dimensional simulations. Indeed, with the 1D evolution code it is not possible to simulate disc fragmentation due to GI, that is known to happen in a region of the parameter space (Kratter et al. 2010; Kratter & Lodato 2016).

4.2 SPH simulations

Fig. 5 shows shapshots of column density evolution in the 3D PHANTOM simulations, for three different disc aspect ratios (left to right) at four different times (top to bottom), where time is given in units of the injection time-scale (equation 19). Thinner discs (right column) show more spiral arms because the saturation value of the disc-to-star mass ratio is lower, and according to the cubic dispersion

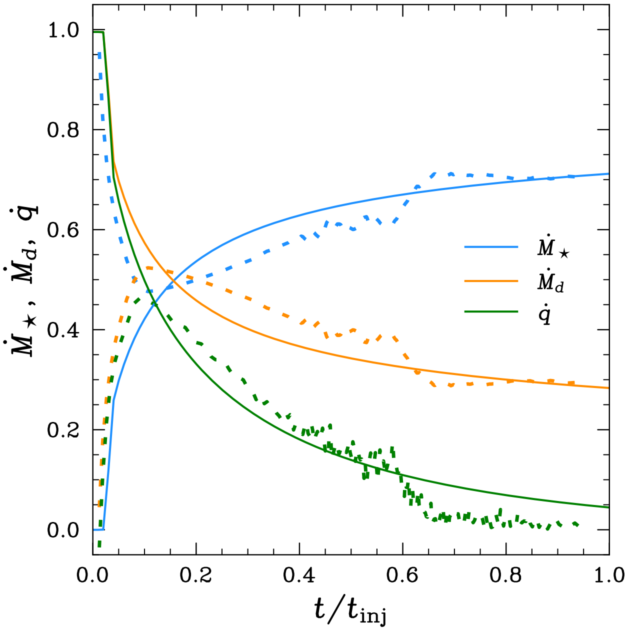

Figure 6. Comparison between the 1D evolution code and the 3D SPH simulation. The solid lines correspond to the evolution of $\dot{M}_\star$ (blue), $\dot{M}_d$ (orange) and $\dot{q}$ (green) according to the 1D evolution code, for the reference simulation. The dotted lines show the behaviour of the reference 3D SPH simulation, with the same initial parameters.

Table 3. Global quantities of the 3D simulations at self-regulation compared with the analytical expectations. The star and disc accretion rates are in units of the injection rate.

$\dot{M}_{\rm inj}$	H/R	q	$\dot{M}_\star$	$\dot{M}_d$	Exp. $\dot{M}_\star$	Exp. $\dot{M}_d$
2.5×10^{-5}	0.075	0.25	0.75	0.25	0.8	0.2
5×10^{-5}	0.1	0.32	0.7	0.3	0.75	0.25
1×10^{-4}	0.125	0.39	0.7	0.3	0.71	0.29

relation of Lau & Bertin (1978), the most unstable mode moves to higher m in discs with lower q. This also explains the trend with time, as q increases with time towards its saturation value (top to bottom), also producing more spiral arms.

4.2.1 *Comparison with 1D code*

Fig. 6 compares the evolution of disc mass, star mass, and disc-to-star mass ratio between the 1D and 3D reference simulations (**S1D_1** and **S3D_1**). Apart from the dimensionality, the key physical difference between the 1D and 3D simulations is the viscosity. Indeed in the 1D case the value of α is a prescribed function of Q which self-regulates so as to achieve the required steady state α; in the 3D case conversely the development of the instability is modelled hydrodynamically and the value of α quantifies angular momentum transport associated with spiral structures.

The results of the PHANTOM simulation appear to be well described by the simple 1D code. In 3D the system also relaxes into a self-regulated state with a constant disc to star mass ratio. In addition, the final values of the accretion rates are in line with the analytical expectations. Table 3 compares the results of the 3D simulations to the theoretical expectations, showing agreement to $\lesssim 10$ per cent between the measured and expected mass growth rates.

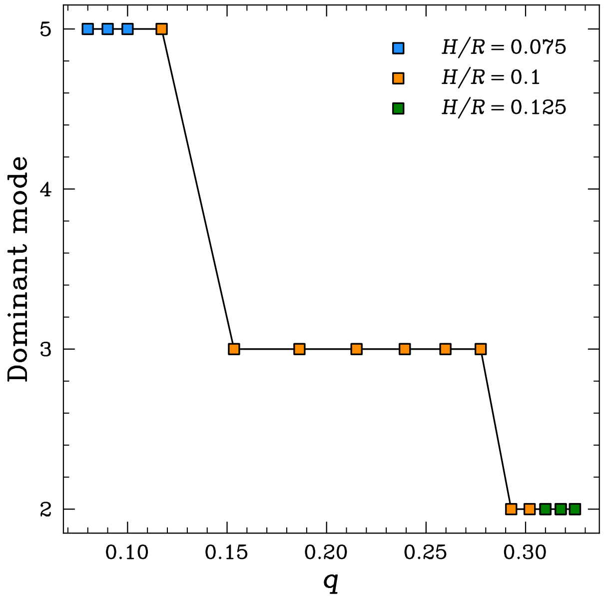

Figure 7. Dominant azimuthal mode as a function of the disc to star mass ratio, combining the three PHANTOM simulations.

4.2.2 *Azimuthal modes analysis*

We start the morphological characterization by studying the azimuthal modes of the perturbation. We extract the azimuthal wavenumber m, i.e. the number of spiral arms, using the procedure described in Cossins et al. (2009). We first divide the disc into radial annuli with fixed width $\Delta R = 0.2$. Then, within each annulus, the azimuthal wavenumber amplitude $\mathcal{A}_m$ is computed directly from the SPH particles' position according to

$$\mathcal{A}_m = \frac{1}{N_{\rm ann}} \left| \sum_{i=1}^{N_{\rm ann}} \exp[-im\phi_i] \right| , \tag{20}$$

where ϕ_i are the azimuthal angles of the individual particles and $N_{\rm ann}$ is the number of particles within an annulus.

The number of spiral arms is determined by the disc to star mass ratio: in particular, high disc mass discs show fewer spiral arms (Lau & Bertin 1978; Laughlin & Bodenheimer 1994). Fig. 7 shows the dominant azimuthal wavenumber as a function of the disc to star mass ratio. To produce the plot, we combined the results of the three PHANTOM simulations, using the fact that q increases with time. As expected, a thicker disc (higher q) has a lower dominant m mode.

This behaviour is clearly seen in Fig. 8, which shows the morphology (top row) and power spectrum (bottom row) of the reference simulation at two distinct time- steps, $t = 0.25$ (left) and $t = 0.75$ (right) at a radius $R = 7.5$. At $t = 0.25$, the disc has not yet reached self-regulation, and its disc-to-star mass ratio is still increasing. Consequently, $q(t = 0.25) < q(t = 0.75)$, leading to a different spiral morphology. At this stage, the dominant mode is $m = 3$, as evidenced by the power spectrum peaks occurring at multiples of three. In contrast, by $t = 0.75$, the disc has achieved self-regulation and the mass ratio has stabilized to a level where the dominant mode shifts to $m = 2$, with peaks of the power spectrum appearing in multiples of two. As mentioned above, the disc to star mass ratio controls the azimuthal wavenumber.

Fig. 9 illustrates the transition between $m = 2$ and $m = 3$ modes more clearly, by plotting the amplitude of these modes at $R = 7.5$

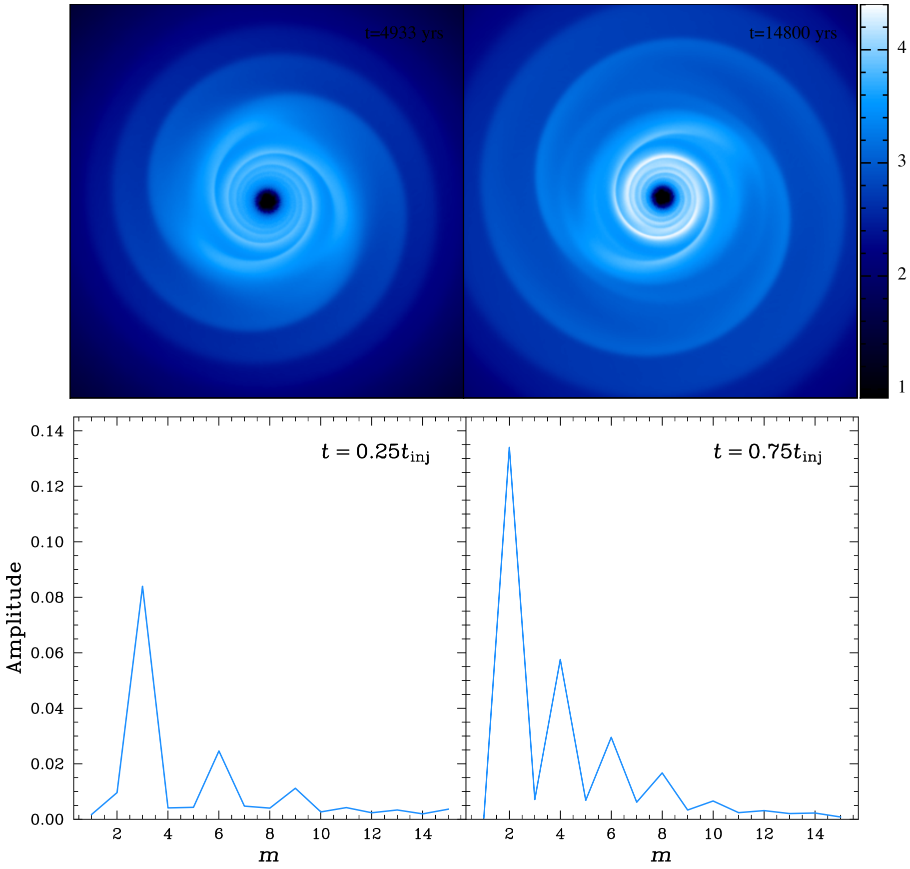

Figure 8. Density snapshots and the corresponding power spectrum for the reference simulation at $t = 0.25$ and $t = 0.75$ at $R = 7.5$.

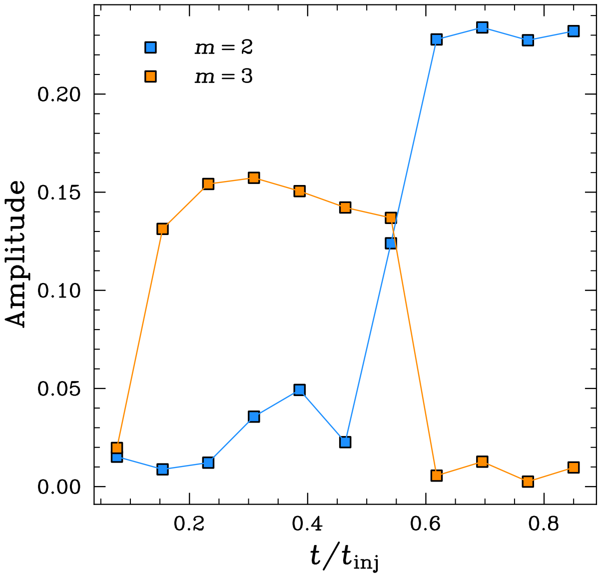

Figure 9. Evolution of the $m = 2$ and $m = 3$ modes at $R = 7.5$ in the reference simulation as function of time.

in the disc as a function of time. A sharp transition from $m = 3$ to $m = 2$ modes can be seen to occur, controlled by the disc to star mass ratio. Comparison with Fig. 6 shows that this transition happens when self-regulation kicks in, at approximately $t = 0.5t_{\rm inj}$.

4.2.3 Morphological differences with cooling driven GI

Fig. 10 compares the reference simulation with its counterpart where the GI is triggered by cooling, both having the same final disc-to-star mass ratio. From the density snapshots, we observe that the dominant mode remains the same, $m = 2$, since it is just determined by the disc to star mass ratio. Yet important morphological differences emerge. In the cooling simulation, the pitch angle remains nearly constant with radius, whereas in the infall case, as we will show below, this is not the case. More importantly, the cooling-driven spirals are more flocculent, exhibiting transient secondary spiral arms that continuously form and dissipate. This is particularly evident in the azimuthal power spectra shown in the bottom panel of Fig. 10. While both simulations display a peak at $m = 2$, indicating the dominant mode, the cooling simulation exhibits power across higher m modes, emphazising its flocculent nature.

We believe that there are two main drivers of the differences between the cooling and the infall-driven cases. The first is the difference in thermodynamic treatment: the cooling simulations solve the full energy equation and allow the temperature to evolve, while the infall-driven simulations assume a locally isothermal equation of state. This certainly affects the dynamics and the development of instabilities. However, it is worth noting that previous studies using a variety of thermodynamic and infall prescriptions – including Kratter et al. (2008), Zhu et al. (2012), and Harsono, Alexander & Levin (2011) – consistently report the emergence of coherent spiral structures, often resembling global modes excited near the infall radius. For instance, Kratter et al. (2010) and Zhu et al. (2012)

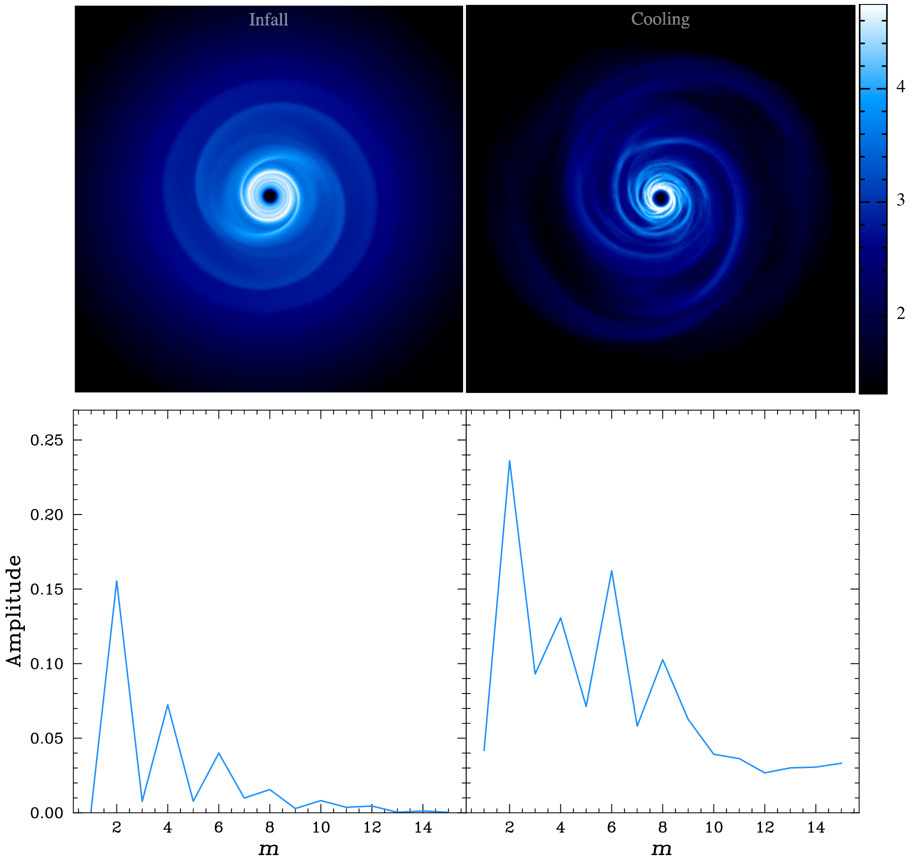

Figure 10. Surface density snapshots (top row) for the reference simulation (left) and the cooling driven (right), and the corresponding azimuthal power spectrum at $R = 5$ (lower panels).

adopt different treatments of the thermal physics (with the former using a locally isothermal approximation and the latter implementing radiative cooling), yet both identify similar morphological features to those found in our simulations. Harsono et al. (2011), on the other hand, include infall onto a disc governed by beta cooling. The resulting morphology exhibits a more coherent global mode structure, albeit with residual flocculent spirals characteristic of beta-cooled discs. These comparisons suggest that the presence of coherent spiral patterns may be a robust outcome of infall, somewhat independent of the exact thermodynamic treatment. Secondly, in the infall-driven case, the angular momentum of the infalling material plays a crucial role in determining the modes that are excited and the onset of the non-linear evolution. In this paper, we have considered an idealized case where the infalling material has a fixed angular momentum, exactly the Keplerian angular momentum at the injection radius. This excites a coherent global mode with a pattern speed matching the Keplerian value at that radius. In more realistic scenarios, the angular momentum of the infalling material would span a broader range, potentially exciting a richer spectrum of modes.

4.2.4 *Spiral tracking and pattern speed*

In order to perform the radial mode analysis and to calculate the pattern speed of the perturbation in the infall-driven scenario, we track the spiral arms. To identify the spiral arms in the simulations, we find the peaks in the radial derivative of the surface density[3] using `find_peaks` implemented in the PYTHON library SCIPY (Virtanen et al. 2020). The left panel of Fig. 11 shows the tracking procedure for the **S3D_3** simulation after 1 injection time. The tracking procedure allows us to compute the pattern speed of the spiral, comparing the position of the maxima at different times. The right panel of Fig. 11 shows the pattern speed of the spirals as a function of radius, compared with $\Omega(R)$ extracted from the simulation. The pattern speed is constant throughout the radial extent of the disc, marking a fundamental difference from the cooling-driven case. Indeed, in the standard GI scenario (Cossins et al. 2009), the instantaneous pattern speed of the spiral is close to the local Keplerian frequency. Spiral density waves are typically generated near corotation, and shock quickly as they propagate away from it, dissipating.

Interestingly, in the infall scenario Ω_p matches the angular frequency at the injection radius $\Omega_{\rm inj}$, and this recovered in every simulation, as shown in Appendix C. This behaviour may be interpreted as the infall triggering a spiral mode with a pattern speed that matches the velocity of the injected material.

Since the pattern speed of the perturbation is constant with radius, it is possible to compute the Lindblad resonances according to equation (6). In this case $m = 2$, and the resonances are at $R_{\rm ILR} = 0.63 R_{\rm inj}$

---

[3] We tried to track the spiral arms both from the maxima in the surface density and in the radial derivative of Σ, and we found that the latter quantity better trace the perturbations.

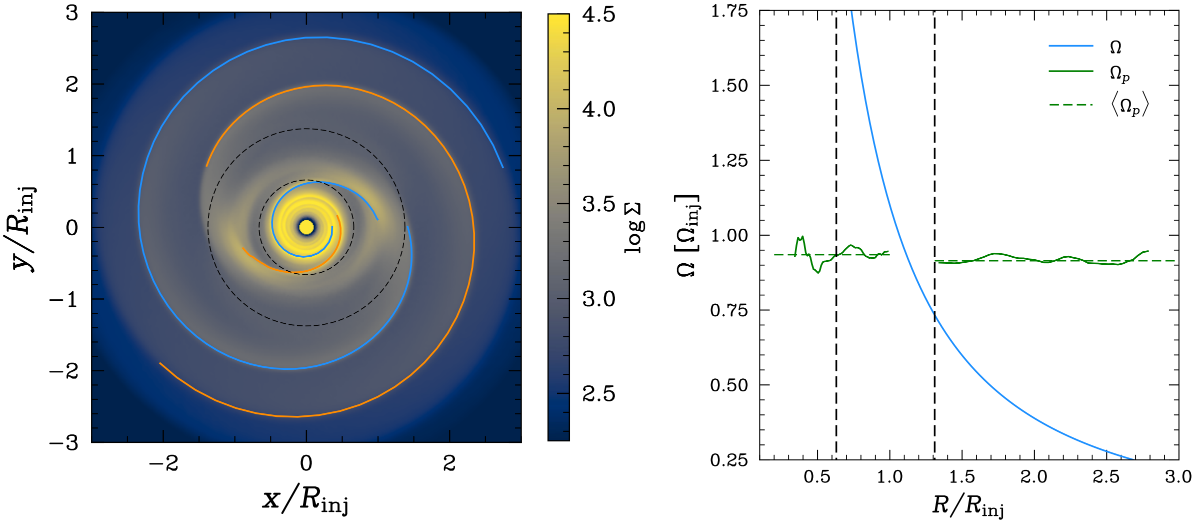

Figure 11. Left panel: spiral arms identification procedure for the **S3D_3** simulation after 1 injection time. The dominant mode is $m = 2$, and the two spiral arms are underlined with blue and orange lines. Right panel: pattern speed of the identified spirals (green solid line), radially averaged value (green dashed line) and Ω extracted from the simulation (blue line). The vertical dashed lines correspond to the inner and outer Lindblad resonances.

and $R_{\rm OLR} = 1.31 R_{\rm inj}$. Clearly, the corotation resonance occurs at the injection radius. In the following, the Lindblad resonances are marked on the figures with black dashed lines.

By visually inspecting the spiral morphology, as shown in Fig. 12, something interesting happens between the injection radius and the outer Lindblad resonance. The bottom panel of Fig. 12 shows the divergence of the velocity field ∇v, a quantity that tracks the spirals. Indeed, when $\nabla v < 0$ the flow is converging, and that happens at the wave front. The spiral structure clearly breaks down between the injection radius and the outer Lindblad resonance. To understand the cause of this behaviour, we need to analyse wave propagation in that region using the dispersion relation.

The dispersion relation can be written in dimensionless units according to

$$s^2 = Q^2\hat{k}^2 - 2|\hat{k}| + 1, \tag{21}$$

where $s = m(\Omega_{\rm p} - \Omega)/\kappa$ and $\hat{k} = k/(\kappa^2/\pi G\Sigma)$. This variable is particularly useful because $s = 0$ corresponds to corotation, and $s = \pm 1$ to the Lindblad resonances. Under the assumption of Keplerian potential, which is justified in our simulations, the variable s can be re-written as

$$s = m\left(\hat{R}^{3/2} - 1\right), \tag{22}$$

where $\hat{R}$ is the radius in units of the corotation one, that is the injection radius in this case. Combining this with the dispersion relation, we can obtain how kH changes with radius, depending on the Toomre parameter Q

$$\hat{R} = \left[\pm\frac{1}{m}\sqrt{Q^2\hat{k}^2 - 2|\hat{k}| + 1} + 1\right]^{2/3}. \tag{23}$$

As shown in Binney & Tremaine (1987), Chapter 6, there are four families of possible solutions:

(i) Short trailing waves ($k > 0$ and kH increasing with $|s|$),
(ii) Long trailing waves ($k > 0$ and kH decreasing with $|s|$),
(iii) Short leading waves ($k < 0$ and kH increasing with $|s|$),
(iv) Long leading waves ($k < 0$ and kH decreasing with $|s|$).

In this analysis, we are interested in trailing waves ($k > 0$), so we can discard the leading solutions. The left panel of Fig. 13 shows the dispersion relation in the $(\hat{R}, kH)$ plane, where we fixed $m = 2$, using the Toomre parameter extracted from the simulation, shown in the right panel of Fig. 13, i.e. $Q \sim 1$ for $\hat{R} < 1$ and $Q \sim 1.5$ for $\hat{R} > 1$. Since the Toomre parameter is equal to 1 inside the corotation, a wave that is created at the injection radius can propagate inward following the short trailing branch (decreasing the pitch angle far from the injection), k being a real number. Conversely, since $Q > 1$ outside the corotation, the radial wavenumber k is imaginary, hence the wave is evanescent and its amplitude decreases. At some point, far out enough from the injection radius, the radial wavenumber becomes real again, and the wave can propagate, following again the short trailing branch. It happens that the location where k becomes real is close to the outer Lindblad resonance for $Q \sim 1.5$, which matches what we see in the simulation (see Fig. 12).

This argument explains why the spirals break between the injection radius and the outer Lindblad resonance[4] Hence, in that region we are not able to track them, and in the following plot that region will be excluded from our analysis.

We stress that the analysis presented here on wave propagation is necessarily simplified. A more rigorous treatment would require the use of higher-order dispersion relations (Lau & Bertin 1978), or ideally a global stability analysis that properly accounts for boundary conditions and normal modes (Shu 1970a, b) and allows one to compute growth rates of the perturbations. Such an approach would provide a more complete understanding of the nature and evolution of the modes triggered by the infall. We leave this detailed investigation to a forthcoming paper.

4.2.5 *Radial mode analysis*

Performing the radial mode analysis is more challenging compared to the azimuthal modes. As pointed out by Cossins et al. (2009),

---

[4] Formally, the propagation of the wave across the evanescent region is analogous to the quantum mechanical tunneling effect. In both cases, the wavenumber becomes imaginary, leading to an exponential decay of the wave amplitude. However, it is important to stress that this is only a mathematical analogy: no actual quantum effects are involved.

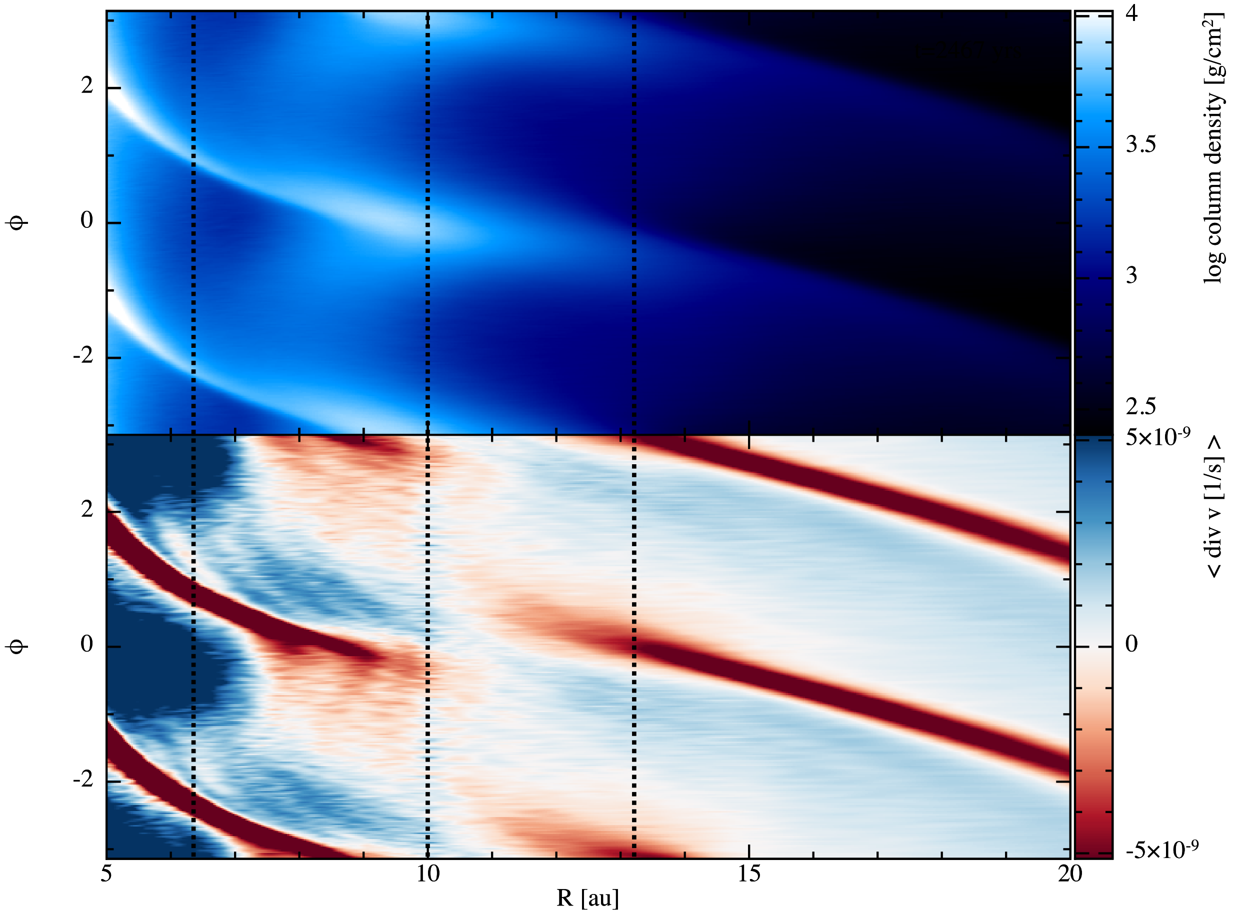

Figure 12. Polar plot of the surface density Σ (top panel) and of the divergence of the velocity field $\nabla\cdot v$ (bottom panel) of the **S3D_3** simulation after 1 injection time.

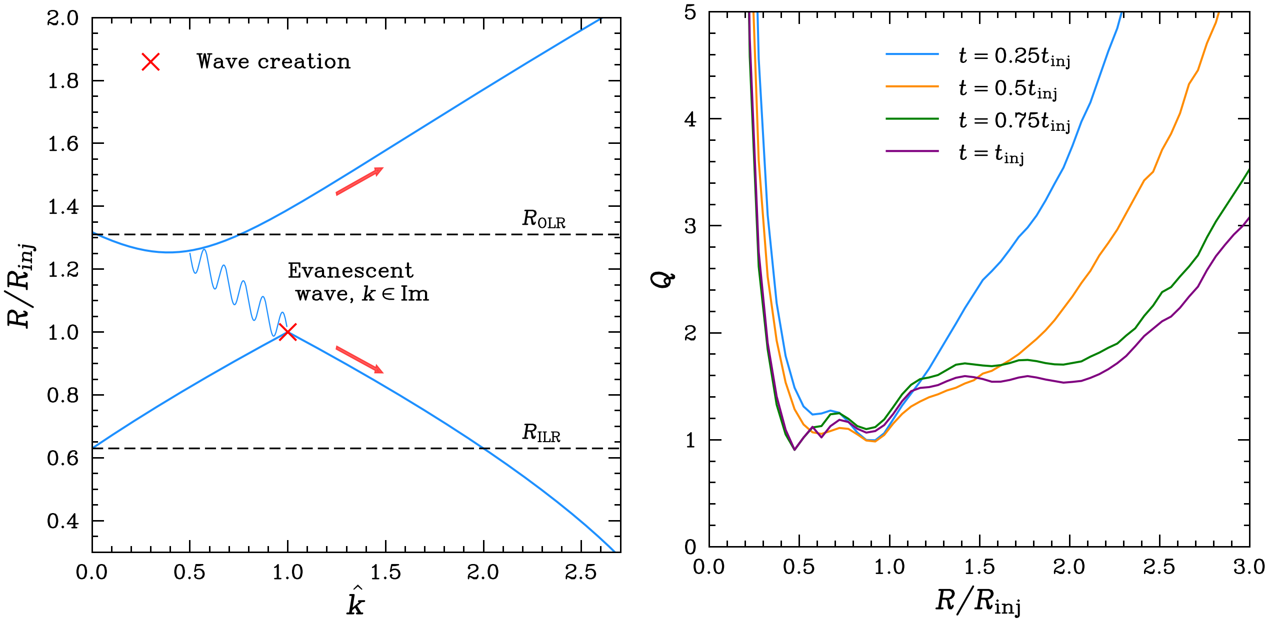

Figure 13. Left panel: propagation diagram of trailing waves in a Keplerian disc, according to equation (23). The waves are generated at the corotation-injection radius (red cross) and they propagate inwards-outwards following the short trailing branches (arrows). The outward wave'tunnels' through the evanescent zone, where $k \in \mathrm{Im}$. Right panel: Toomre parameter of the **S3D_3** simulation as a function of radius for different times.

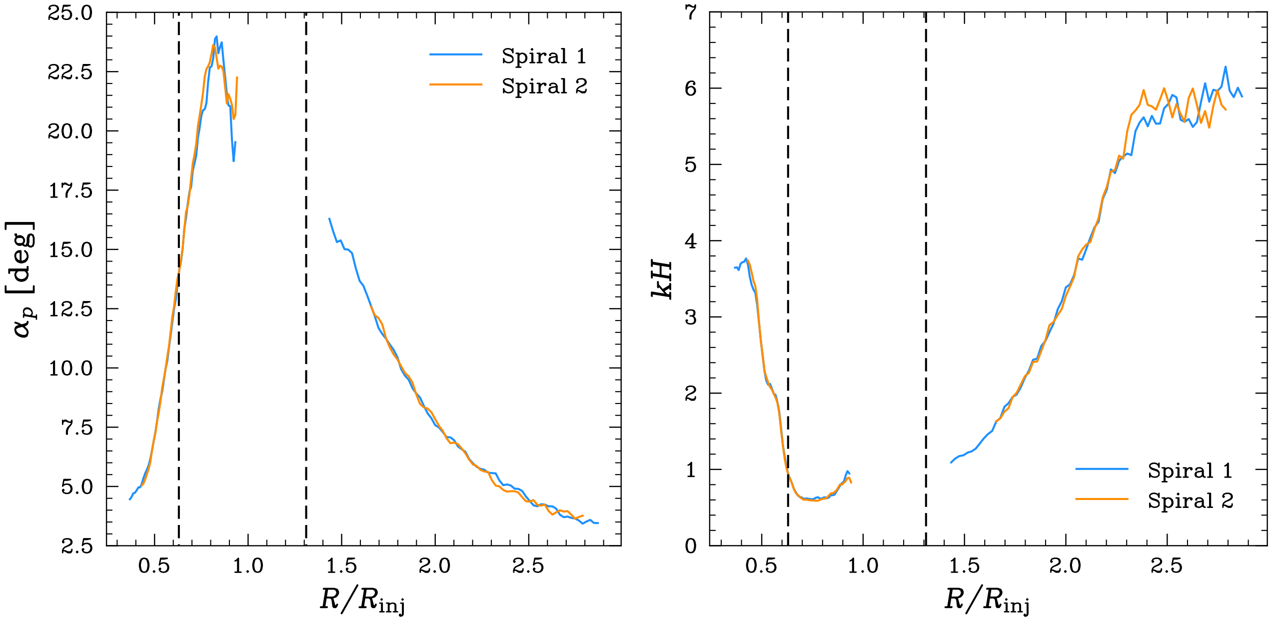

Figure 14. Left panel: Pitch angle of the **S3D_3** simulation after 1 injection time computed using equation (25). Right panel: kH profile of the **S3D_3** simulation after 1 injection time. This quantity increases with distance from corotation, pointing to the short-trailing nature of the waves.

radial binning of the data is crucial in order to obtain the correct result. We tried to use the same procedure of Cossins et al. (2009), but we found that our results strongly depended on the radial binning, and did not converge. For this reason, we instead derive the radial wavenumber k by measuring the pitch angle of the spiral, since they can be related according to

$$k = \frac{m}{R \tan \alpha_p}. \tag{24}$$

To compute the opening angle $\alpha_p(R)$, we leverage on the tracking procedure outlined before. We compute the opening angle of the identified spirals through geometrical arguments, where α_p represents the angle between the tangent to the spiral arm and the tangent to a circular arc at a given radius, according to

$$\tan \alpha_p = \left| \frac{1}{R} \frac{\mathrm{d}R}{\mathrm{d}\phi} \right|. \tag{25}$$

In principle, each identified spiral arm has a distinct k; however, we verify that the individual radial wavenumbers are consistent and take the average of k across all identified spirals for the purposes of our analysis.

The right panel of Fig. 14 shows the pitch angle of the **S3D_3** simulation after 1 injection time. Again, the Lindblad resonances are shown as black dashed lines. The pitch angle profile shows a maximum at $R \sim R_{\rm inj}$ with smaller opening angles both inside and outside this radius. This behaviour resembles the planetary case, where the wave propagation is almost radial at the planetary location, and then the perturbations become tightly wound (e.g. Ogilvie & Lubow 2002; Rafikov 2002). From the pitch angle we then compute the dimensionless radial wavenumber kH using equation (24), where we define H as[5]

$$H = \frac{c_s^2}{\pi G \Sigma}. \tag{26}$$

The right panel of Fig. 14 shows kH as a function of radius. This quantity increases with distance from the injection, confirming the short-trailing nature of these spirals.

5 DISCUSSION

5.1 Cooling or infall?

Gammie (2001) showed that the non-linear outcome of gravitational instability in cooling accretion discs is either fragmentation or a state of self-regulated turbulence, depending on the cooling rate. This picture assumes that the only source of heat in the disc is due to spiral shocks caused by the gravitational instability itself. Hence, if the cooling rate is long enough, the disc self regulates as a temperature-controlled thermostat around $Q = 1$.

However, in a realistic scenario, heating from the central object plays a crucial role in determining the disc's thermal equilibrium, potentially influencing its ability to reach a self-regulated state. Gravitational instability in irradiated discs is a delicate topic. Simulations of protostellar discs incorporating realistic heating and cooling mechanisms suggest that stellar irradiation primarily governs the disc temperature (Rice et al. 2011; Haworth et al. 2020; Young et al. 2024; Rowther et al. 2024b), with internal heating from shocks playing only a minor role on a global scale.

Recent results by Rowther et al. (2024b), using on-the-fly Monte Carlo radiative transfer, reinforce the idea that protostellar discs are *passive*, their thermal structure dictated by the central star. Similar conclusions were drawn by previous studies (Matzner & Levin 2005; Krumholz, Klein & McKee 2007; Kratter et al. 2010; Meru & Bate 2010). These are all global simulations, where the large-scale temperature profile is set by stellar irradiation. However, local deviations can occur. For instance, Leedham et al. (2025,), using 2D shearing sheet simulations with external irradiation, found that when a spiral arm crosses, the temperature varies due to local shock heating. This suggests that while the disc's overall thermal structure

---

[5] We underline that for a gravitationally unstable disc, this definition of H is equivalent to $\hat{k}/k$, as used in equation(21).

is dictated by the star, local heating from spiral shocks may still play a significant role.

Our work adopts the extreme assumption, treating the disc as locally isothermal. The reality likely lies between these two extremes, necessitating live radiative transfer simulations with infall to assess the impact of material injection.

Interestingly, all of the discs we believe are gravitationally unstable, Elias 2–27, (Paneque-Carreño et al. 2021) AB Aurigae, (Tang et al. 2012), and GM Aur, (Schwarz et al. 2021), show evidence of infall from the environment. Our numerical experiments in this paper suggest a picture closer to a bathtub than a thermostat: the disc fills by adding mass from the environment until the $Q = 1$ threshold is reached, upon which the disc starts to 'release water' onto the central star. The steady state is that the disc and star mass grow together.

Fig. 6 confirms that this simple picture holds in both 1D and 3D simulations. While we have injected mass in an axisymmetric way to minimize the number of free parameters, we expect the same steady state behaviour independent of the way mass is added to the disc. That a steady state is reached means that past work using β-cooling as a numerical sandbox to understand turbulence, spiral arms, dust aggregation and angular momentum transport driven by gravitational instability can be ported to this new and more realistic framework of infall driven mass-regulated gravitational instability.

5.2 Observational signatures

A complication is that infall itself can drive spiral arms in the disc, if the mass is added in a non-axisymmetric way (Calcino et al. 2025). Hence, it is important to be able to distinguish spiral arms driven by infall from those driven by gravitational instability. Infall is a necessary but not sufficient condition for gravitational instability to operate. The most robust way to distinguish the two is to measure the disc mass (e.g. with dynamical disc mass measurements; Veronesi et al. 2021; Lodato et al. 2023; Martire et al. 2024; Longarini et al. 2025) and hence constrain the Toomre Q parameter directly.

Hall et al. (2020), using $\beta-$cooling simulations, showed that gravitational instability can leave a unique signature in channel maps of molecular line emission from discs – the 'GI wiggle'. Essentially spirals produce global scale perturbations in the velocity field which shift emission into neighbouring frequency channels. This characteristic wiggle has been observed by ALMA in Elias 2–27 (Paneque-Carreño et al. 2021; Longarini et al. 2024) and AB Aur (Speedie et al. 2024) and claimed as evidence for gravitational instability operating in these discs. We expect GI-induced wiggles to persist in the infall-driven gravitational instability scenario. However, the amount of angular momentum transported by the spirals – ultimately setting the wiggle amplitude (Longarini et al. 2021, 2024) – would no longer be governed by $\beta-$cooling. Instead, it would depend on a combination of the mass injection rate and the disc's aspect ratio. As noted by Longarini et al. (2024), if the wiggle amplitude is interpreted as a measure of the angular momentum transport driven by GI – specifically, $\alpha_{\rm GI}$ as defined in Longarini et al. (2024) – then it remains independent of the mechanism driving self-regulation.

5.3 Setting the initial stages for planet formation

In this paper, we showed that if a protostellar disc undergoes infall driven gravitational instability, it reaches a self-regulated state where star mass, disc mass and accretion rate are related. Large surveys of star-forming regions have revealed robust power-law correlations between stellar mass and disc properties (Ansdell et al. 2016; Barenfeld et al. 2016; Alcalá et al. 2017; Ansdell et al. 2017; Manara et al. 2017, 2020; Testi et al. 2022). These correlation seem to evolve, steepening with time. Somigliana et al. (2022) showed that these correlations cannot be solely explained by viscous evolution unless an initial correlation is already imprinted in the system's initial conditions.

Infall-driven gravitational instability is a promising mechanism capable of imprinting correlations in the initial conditions for the viscous evolution of protostellar discs. In particular, we have shown that this mechanism naturally leads to a disc-to-star mass ratio proportional to the disc aspect ratio, while the mass accretion rate scales with α, which itself depends on the stellar mass, aspect ratio, and mass injection rate. To robustly predict the scaling relations of M_d and $\dot{M}_\star$ with stellar mass, one must consider several physical factors that regulate disc structure and evolution during the early phases. For instance, the disc temperature plays a critical role, as it sets the aspect ratio.

Although in later stages the temperature of a passively irradiated disc is known to depend on the stellar luminosity – and therefore the stellar mass – during the earliest phases, accretion luminosity is the dominant heating source (Krumholz et al. 2007; Offner et al. 2009; Hennebelle et al. 2022), effectively setting the thermal structure of the disc. In an infall-driven scenario, this luminosity is tightly linked to the mass injection rate, which in our model is assumed to be constant in time and radially localized. In reality, this is not the case (Cassen & Moosman 1981; Terebey, Shu & Cassen 1984; Hueso & Guillot 2005; Bae et al. 2013; Hartmann & Bae 2018). Exploring these effects in more detail could provide further insight into the phenomena setting the initial stage of planet formation.

5.4 Rescaling the injection rate

Our study is performed with protostellar discs in mind, but our analysis is applicable across a wide mass range, from protostellar discs to discs around black holes where there is irradiation from a central source (which could be the inner disc itself). Since all our simulations are scale-free, they can be used to interpret any accretion disc. The same principle applies to the injection rate, which must be considered in relation to the disc's physical size and the injection radius. The key variable that governs angular momentum transport is not simply $\dot{M}_{\rm inj}$, but rather the dimensionless quantity defined $\dot{\mu}$, as defined in equation (13).

Fig. 15 illustrates how to rescale the *physical* injection radius. The left panel shows the α-viscosity as a function of $\dot{\mu}$ and H/R, as described in equation (15), where, for simplicity, the term $(1+q)$ has been omitted. A specific value of α is determined by the combination of H/R and $\dot{\mu}$. To convert $\dot{\mu}$ into physical units – specifically, an injection rate in solar masses per year – the right panel of Fig. 15 provides the corresponding scaling relation based on the injection radius, according to

$$\log \dot{M}_{\rm inj}[{\rm M}_\odot/{\rm yr}] \simeq \log \dot{\mu} + \frac{3}{2}\log M_\star[{\rm M}_\odot] - \frac{3}{2}\log R_{\rm inj}[{\rm au}] + 1.6. \qquad (27)$$

For instance, in a protoplanetary disc with an aspect ratio of $H/R = 0.1$ and a viscosity parameter $\alpha = 10^{-2}$, the dimensionless mass injection rate is $\log_{10}\dot{\mu} = -4.5$. Injecting material at 50 au results in a physical injection rate of $6\times10^{-7}\,{\rm M}_\odot\,{\rm yr}^{-1}$. A similar calculation applies to AGN discs around black holes. For a disc with $H/R = 0.05$ and $\alpha = 10^{-2}$, the dimensionless injection rate is $\log_{10}\dot{\mu} = -5.42$. If mass is injected at $R_{\rm inj} = 0.1$ pc around a $10^6\,{\rm M}_\odot$ black hole, the corresponding physical injection rate is $5.5\times10^{-2}\,{\rm M}_\odot\,{\rm yr}^{-1}$.

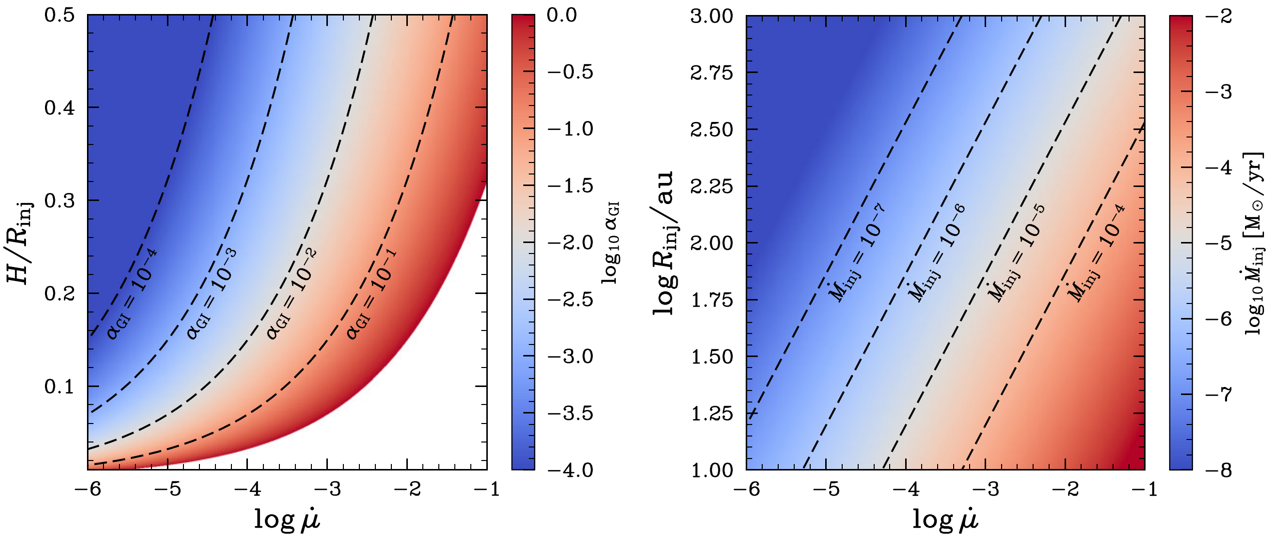

Figure 15. Left panel: $\alpha-$viscosity as a function of Γ and the aspect ratio. Right panel: physical injection rate as a function of the injection radius and Γ.

6 CONCLUSIONS

In this work, we characterized gravitational instability triggered by mass injection through 1D and 3D numerical simulations. Our findings can be summarized as:

(i) Self-regulation can be achieved in terms of mass rather than temperature, leading to a constant disc-to-star mass ratio over time. We verify that self-regulation occurs in both 1D and 3D simulations and find that the characteristic timescale is the injection time, corresponding to the stellar mass doubling time-scale. Additionally, self-regulation is reached more rapidly in 3D than in 1D.

(ii) The 1D evolution code enables us to study the evolution of key global parameters, including $M_\star$, $M_{\rm d}$, their ratio, and the disc radius $R_{\rm d}$, as functions of the aspect ratio, injection rate, and injection location. We find that the aspect ratio is the critical parameter, as it sets the maximum mass the disc can sustain.

(iii) 3D SPH simulations show excellent agreement with the 1D code and provide insights into the morphology of the spirals. Unlike the cooling-driven case, the spirals are less flocculent, with an azimuthal power spectrum peaking at the dominant mode and exhibiting little power in higher m modes. Additionally, we observe a time correlation between different modes, particularly between $m = 2$ and $m = 3$.

(iv) Unlike the cooling scenario, the pattern speed of the spiral perturbations in the infall driven GI have a constant pattern speed, matching the angular velocity at the injection location. In this picture, the injection radius is the corotation resonance, and it is possible to identify the inner and outer Lindblad resonances. Spiral waves generate at the injection, and freely propagate inwards since the Toomre parameter is $Q \sim 1$, while between the injection and the outer Lindblad resonance the wave is evanescent, since $Q > 1$ and $k \in \mathrm{Im}$.

ACKNOWLEDGEMENTS

The authors thank the referee for the constructive report which improved the clarity of the paper. In addition, the authors thank Rebecca Nealon for the de-refinement routine in PHANTOM, Duncan Forgan for contributing the disc stresses and mode analysis module and Alice Somigliana for helping in the discussion about the initial conditions. The authors finally thank Juan Garrido-Deutelmoser, Andrew Winter, Francesco Zagaria, Jess Speedie and Cat Leedham for useful discussions.

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement # 823823 (RISE DUSTBUSTERS project). CL and CJC have been supported by the UK Science and Technology Research Council (STFC) via the consolidated grant ST/W000997/1. GL acknowledges support from PRIN-MUR 20228JPA3A and from the European Union Next Generation EU, CUP:G53D23000870006. KMK acknowledges support NASA under agreement no. 80NSSC21K0593 for the programme 'Alien Earths'.

The simulations presented in this work were performed using the DiRAC Data Intensive service at Leicester (DiAL3), operated by the University of Leicester IT Services, which is part of the STFC DiRAC HPC Facility (www.dirac.ac.uk) within the RAC large project DISCSIM IV and Cambridge Service for Data Driven Discovery (CSD3).

In this work, we used SPLASH to create figures of the hydrodynamical simulations (Price 2007) and SARRACEN (Harris & Tricco 2023) for part of the analysis.

DATA AVAILABILITY

The simulations performed in this work will be made available on Zenodo 10.5281/zenodo.14906500.

REFERENCES

Alcalá J. M. et al., 2017, A&A, 600, A20
Ansdell M. et al., 2016, ApJ, 828, 46
Ansdell M., Williams J. P., Manara C. F., Miotello A., Facchini S., van der Marel N., Testi L., van Dishoeck E. F., 2017, AJ, 153, 240
Bae J., Hartmann L., Zhu Z., Gammie C., 2013, ApJ, 764, 141
Baehr H., Zhu Z., 2021, ApJ, 909, 135
Balbus S. A., Papaloizou J. C. B., 1999, ApJ, 521, 650
Barenfeld S. A., Carpenter J. M., Ricci L., Isella A., 2016, ApJ, 827, 142
Bate M. R., Bonnell I. A., Bromm V., 2003, MNRAS, 339, 577
Bath G. T., Pringle J. E., 1981, MNRAS, 194, 967
Bertin G., 1997, ApJ, 478, L71

Bertin G., Lodato G., 1999, A&A, 350, 694
Binney J., Tremaine S., 1987, Galactic dynamics, Princeton University Press
Boley A. C., Hartquist T. W., Durisen R. H., Michael S., 2007, ApJ, 656, L89
Booth R. A., Clarke C. J., 2016, MNRAS, 458, 2676
Boss A. P., 2002, Earth Planet. Sci. Lett., 202, 513
Cadman J., Hall C., Fontanive C., Rice K., 2022, MNRAS, 511, 457
Calcino J., Price D. J., Hilder T., Christiaens V., Speedie J., Ormel C. W., 2025, MNRAS, 537, 2695
Cameron A. G. W., 1978, Moon and Planets, 18, 5
Cassen P., Moosman A., 1981, Icarus, 48, 353
Ceppi S., Longarini C., Lodato G., Cuello N., Lubow S. H., 2023, MNRAS, 520, 5817
Cossins P., Lodato G., Clarke C. J., 2009, MNRAS, 393, 1157
Cuello N. et al., 2019, MNRAS, 483, 4114
Cullen L., Dehnen W., 2010, MNRAS, 408, 669
Forgan D., Rice K., Stamatellos D., Whitworth A., 2009, MNRAS, 394, 882
Fukagawa M. et al., 2004, ApJ, 605, L53
Gammie C. F., 2001, ApJ, 553, 174
Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375
Gupta A. et al., 2023, A&A, 670, L8
Gupta A., Miotello A., Williams J. P., Birnstiel T., Kuffmeier M., Yen H.-W., 2024, A&A, 683, A133
Hall C. et al., 2020, ApJ, 904, 148
Harris A., Tricco T., 2023, J. Open Source Softw., 8, 5263
Harsono D., Alexander R. D., Levin Y., 2011, MNRAS, 413, 423
Hartmann L., Bae J., 2018, MNRAS, 474, 88
Haworth T. J., Cadman J., Meru F., Hall C., Albertini E., Forgan D., Rice K., Owen J. E., 2020, MNRAS, 494, 4130
Hennebelle P. et al., 2022, A&A, 668, A147
Hueso R., Guillot T., 2005, A&A, 442, 703
Kenyon S. J., Calvet N., Hartmann L., 1993, ApJ, 414, 676
Kratter K., Lodato G., 2016, ARA&A, 54, 271
Kratter K. M., Matzner C. D., Krumholz M. R., 2008, ApJ, 681, 375
Kratter K. M., Matzner C. D., Krumholz M. R., Klein R. I., 2010, ApJ, 708, 1585
Krumholz M. R., Klein R. I., McKee C. F., 2007, ApJ, 656, 959
Lau Y. Y., Bertin G., 1978, ApJ, 226, 508
Laughlin G., Bodenheimer P., 1994, ApJ, 436, 335
Laughlin G., Rozyczka M., 1996, ApJ, 456, 279
Laughlin G., Korchagin V., Adams F. C., 1997, ApJ, 477, 410
Laughlin G., Korchagin V., Adams F. C., 1998, ApJ, 504, 945
Leedham C. S., Booth R. A., Clarke C. J., 2025, MNRAS, 539, 2780
Lin D. N. C., Pringle J. E., 1987, MNRAS, 225, 607
Lin D. N. C., Pringle J. E., 1990, ApJ, 358, 515
Lin C. C., Shu F. H., 1964, ApJ, 140, 646
Lodato G., 2007, Nuovo Cimento Rivista Serie, 30, 293
Lodato G., Natarajan P., 2006, MNRAS, 371, 1813
Lodato G., Price D. J., 2010, MNRAS, 405, 1212
Lodato G. et al., 2023, MNRAS, 518, 4481
Longarini C., Lodato G., Toci C., Veronesi B., Hall C., Dong R., Patrick Terry J., 2021, ApJ, 920, L41
Longarini C., Armitage P. J., Lodato G., Price D. J., Ceppi S., 2023, MNRAS, 522, 6217
Longarini C. et al., 2024, A&A, 686, L6
Longarini C. et al., 2025, ApJ, 984, L17
Lucy L. B., 1977, AJ, 82, 1013
Lynden-Bell D., Kalnajs A. J., 1972, MNRAS, 157, 1
Manara C. F. et al., 2017, A&A, 604, A127
Manara C. F. et al., 2020, A&A, 639, A58
Martire P. et al., 2024, A&A, 686, A9
Matzner C. D., Levin Y., 2005, ApJ, 628, 817
Mejia A. C., 2004, PhD thesis, Indiana University, Bloomington
Meru F., Bate M. R., 2010, MNRAS, 406, 2279
Nealon R., Ragusa E., Gerosa D., Rosotti G., Barbieri R., 2022, MNRAS, 509, 5608
Offner S. S. R., Klein R. I., McKee C. F., Krumholz M. R., 2009, ApJ, 703, 131
Ogilvie G. I., Lubow S. H., 2002, MNRAS, 330, 950
Paneque-Carreño T. et al., 2021, ApJ, 914, 88
Pérez L. M. et al., 2016, Science, 353, 1519
Price D. J., 2007, PASA, 24, 159
Price D. J., 2012, J. Comp. Phys., 231, 759
Price D. J., Monaghan J. J., 2007, MNRAS, 374, 1347
Price D. J. et al., 2018, PASA, 35, e031
Rafikov R. R., 2002, ApJ, 569, 997
Rafikov R. R., 2015, ApJ, 804, 62
Ragusa E., Alexander R., Calcino J., Hirsh K., Price D. J., 2020, MNRAS, 499, 3362
Rice W. K. M., Lodato G., Pringle J. E., Armitage P. J., Bonnell I. A., 2004, MNRAS, 355, 543
Rice W. K. M., Lodato G., Pringle J. E., Armitage P. J., Bonnell I. A., 2006, MNRAS, 372, L9
Rice W. K. M., Armitage P. J., Mamatsashvili G. R., Lodato G., Clarke C. J., 2011, MNRAS, 418, 1356
Rowther S., Nealon R., Meru F., 2022, ApJ, 925, 163
Rowther S., Nealon R., Meru F., Wurster J., Aly H., Alexander R., Rice K., Booth R. A., 2024a, MNRAS, 528, 2490
Rowther S., Price D. J., Pinte C., Nealon R., Meru F., Alexander R., 2024b, MNRAS, 534, 2277
Safronov V. S., 1960, Annales d'Astrophysique, 23, 979
Schwarz K. R. et al., 2021, ApJS, 257, 20
Shu F. H., 1970a, ApJ, 160, 89
Shu F. H., 1970b, ApJ, 160, 99
Somigliana A., Toci C., Rosotti G., Lodato G., Tazzari M., Manara C. F., Testi L., Lepri F., 2022, MNRAS, 514, 5927
Speedie J. et al., 2024, Nature, 633, 58
Stamatellos D., Whitworth A. P., Bisbas T., Goodwin S., 2007, A&A, 475, 37
Tang Y. W., Guilloteau S., Piétu V., Dutrey A., Ohashi N., Ho P. T. P., 2012, A&A, 547, A84
Terebey S., Shu F. H., Cassen P., 1984, ApJ, 286, 529
Terry J. P., Hall C., Longarini C., Lodato G., Toci C., Veronesi B., Paneque-Carreño T., Pinte C., 2022, MNRAS, 510, 1671
Testi L. et al., 2022, A&A, 663, A98
Tobin J. J. et al., 2020, ApJ, 890, 130
Toomre A., 1964, ApJ, 139, 1217
Veronesi B., Paneque-Carreño T., Lodato G., Testi L., Pérez L. M., Bertin G., Hall C., 2021, ApJ, 914, L27
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Winter A. J., Benisty M., Manara C. F., Gupta A., 2024a, A&A, 691, A169
Winter A. J., Benisty M., Andrews S. M., 2024b, ApJ, 972, L9
Young A. K., Celeste M., Booth R. A., Rice K., Koval A., Carter E., Stamatellos D., 2024, MNRAS, 531, 1746
Zhu Z., Hartmann L., Gammie C., 2010, ApJ, 713, 1143
Zhu Z., Hartmann L., Nelson R. P., Gammie C. F., 2012, ApJ, 746, 110

APPENDIX A: 1D EVOLUTION CODE

We solve the 1D evolution equation for the surface density of a Keplerian disc using a first-order explicit finite-volume update following (Bath & Pringle 1981). We work in dimensionless units, where $\tau = t/t_{\rm inj}$, with $t_{\rm inj} = {\rm M}_{\odot}/\dot{M}_{\rm inj}$, $\sigma = \Sigma/\Sigma(R_{\rm inj})$, $X = R/R_{\rm inj}$ and $D = \nu \times (t_{\rm inj}/R_{\rm inj}^2)$. Equation 16 becomes

$$\frac{\partial \sigma}{\partial \tau} = \frac{3}{X}\frac{\partial}{\partial X}\left[\sqrt{X}\frac{\partial}{\partial X}\left(\sqrt{X}\sigma D\right)\right] + \gamma_{\rm inj}, \tag{A1}$$

where $\gamma_{\rm inj} = \dot{\Sigma}_{\rm inj} \times (t_{\rm inj}/\Sigma(R_{\rm inj}))$. We solve the previous equation using the following discretisation

$$\sigma_j^{n+1} = \sigma_j^n + \frac{3\Delta t}{X_j}\frac{\sqrt{X_{j+1/2}}\left(A_{j+1}^n - A_j^n\right) - \sqrt{X_{j-1/2}}\left(A_j^n - A_{j-1}^n\right)}{(\Delta X)^2} + \gamma_j \Delta t, \tag{A2}$$

where

$$A_j^n = \sqrt{X_j}\sigma_j^n D_j. \tag{A3}$$

The code used to evolve the surface density of the disc is publicly available on https://github.com/crislong/discfusion.

APPENDIX B: VISCOSITY PRESCRIPTION FOR GRAVITATIONALLY UNSTABLE DISCS

In this appendix, we assess the robustness of our 1D results against different viscosity prescriptions. We initially adopted the GI viscosity model from Lin & Pringle (1987), where angular momentum transport occurs only when $Q < 1$. This model is very simple and features an abrupt transition between active and inactive viscosity. To test the sensitivity of our results, we also consider an alternative viscosity prescription from Kratter et al. (2008), given by

$$\alpha_{\rm Kratter} = \max\left\{0.14\left[\left(\frac{1.3}{\max(Q,1)}\right)^2 - 1\right], 0\right\}. \tag{B1}$$

According to this prescription, the transition between active and inactive is shallower. Fig. B1 shows the evolution of star mass, disc mass, and disc to star mass ratio, comparing the two viscosity prescriptions. We find an excellent agreement, meaning that our results are robust.

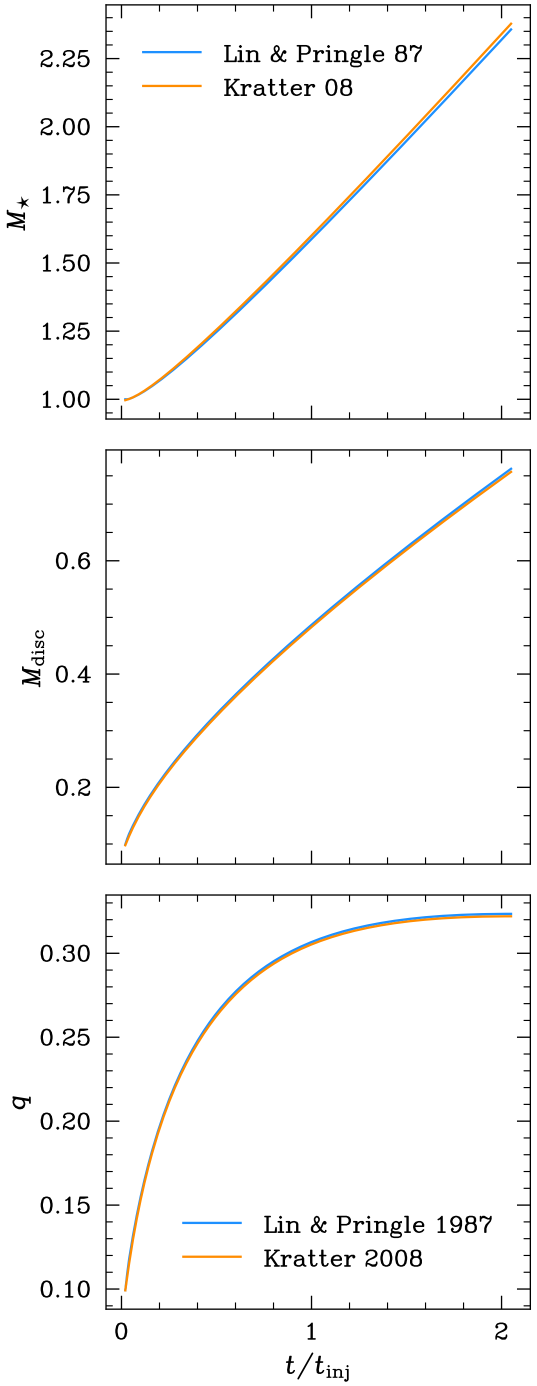

Figure B1. Evolution of star mass, disc mass, and disc to star mass ratio in the 1D evolution code with two different viscosity prescriptions, namely Lin & Pringle (1987) and Kratter et al. (2008).

APPENDIX C: SPIRAL TRACKING

In this appendix, (Fig. C1 and C2) we show the results of the tracking procedure for the simulation **S3D_1** and **S3D_2**.

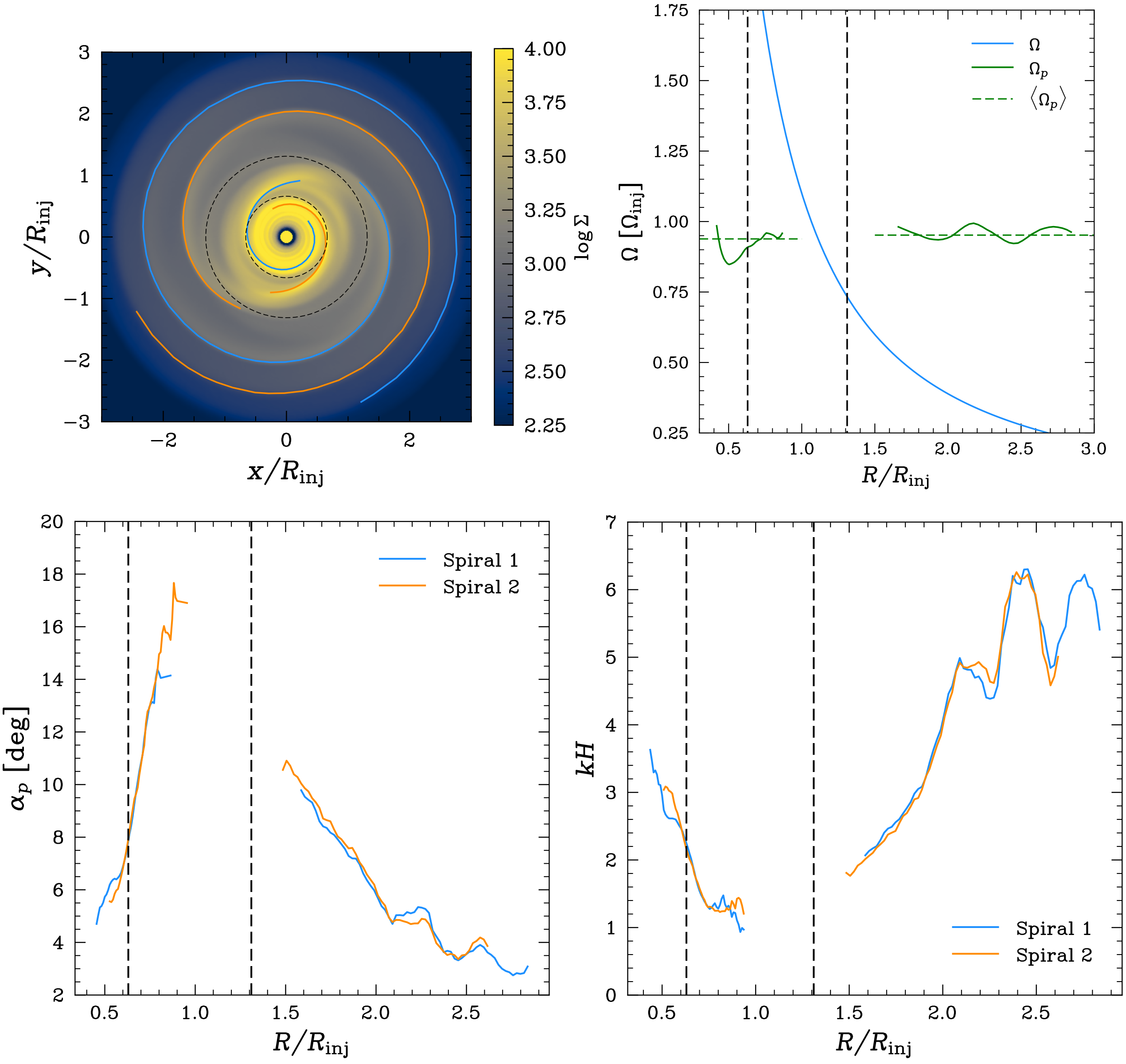

Figure C1. Results of the spiral tracking procedure for the **S3D_1** simulation.

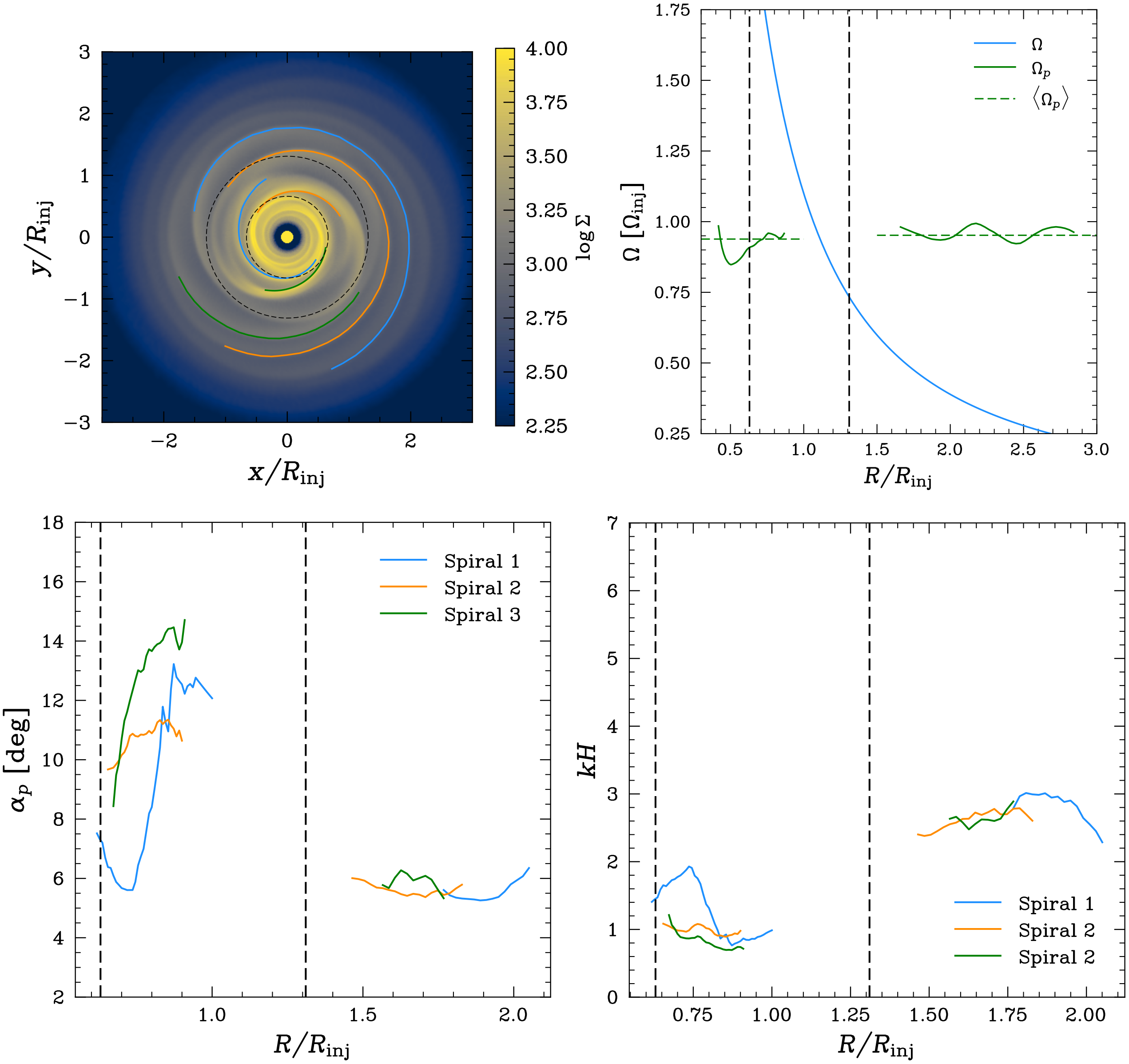

Figure C2. Results of the spiral tracking procedure for the **S3D_2** simulation.

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