Modeling planet-induced gaps and rings in ALMA disks: the role of in-plane radiative diffusion

Alexandros Ziampras,[1][*] Richard P. Nelson,[1] Roman R. Rafikov[2,3]

[1]*Astronomy Unit, Dept. of Physics and Astronomy, Queen Mary University of London, London E1 4NS, UK*
[2]*DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*
[3]*Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA*

ABSTRACT

ALMA observations of protoplanetary disks in dust continuum emission reveal a variety of annular structures. Attributing the existence of such features to embedded planets is a popular scenario, supported by studies using hydrodynamical models. Recent work has shown that radiative cooling greatly influences the capability of planet-driven spiral density waves to transport angular momentum, ultimately deciding the number, position, and depth of rings and gaps that a planet can carve in a disk. However, radiation transport has only been treated via local thermal relaxation, not taking into account radiative diffusion along the disk plane. We compare the previous state-of-the-art models of planet–disk interaction with local cooling prescriptions to our new models that include cooling in the vertical direction and radiative diffusion in the plane of the disk, and show that the response of the disk to the induced spiral waves can differ significantly when comparing these two treatments of the disk thermodynamics. We follow up with synthetic emission maps of ALMA systems, and show that our new models reproduce the observations found in the literature better than models with local cooling. We conclude that appropriate treatment of radiation transport is key to constraining the parameter space when interpreting ALMA observations using the planet–disk interaction scenario.

Key words: planet–disc interactions — accretion discs — hydrodynamics — radiative transfer — methods: numerical

1 INTRODUCTION

Annular features are abundant in observations of protoplanetary disks, one of the more famous examples being the HL Tau system (ALMA Partnership et al. 2015). The DSHARP survey (Andrews et al. 2018) has further revealed a multitude of rings and gaps in several systems using dust continuum emission observations (Huang et al. 2018). This provides theorists with high spatial- and spectral-resolution data, useful in constructing realistic physical models that can constrain properties of these systems such as dust opacities (e.g., Birnstiel et al. 2018), the level of turbulence (e.g., Dullemond et al. 2018), as well as the masses of planets that they might harbor (e.g., Zhang et al. 2018, 2021).

The planet–disk interaction scenario is particularly exciting given the detection of a pair of young planets embedded in the disk around PDS 70 (Keppler et al. 2018; Haffert et al. 2019), cementing the concept that planets grow in protoplanetary disks. Planets interact with the gaseous disk they are embedded in by launching spiral density waves (Ogilvie & Lubow 2002; Rafikov 2002a) which then steepen into shocks, depositing angular momentum into the gas non-locally and driving the opening of one or more gaps in the process (Rafikov 2002b). The resulting radial gas structure consists of one or more pressure maxima adjacent to such gaps, which can trap dust grains and form the bright rings that are observed in continuum emission. As a result, planet–disk interaction has the potential to explain the ring and gap structures in such disks (e.g., Zhang et al.

2018; Guzmán et al. 2018), as well as the non-axisymmetric features such as vortices (e.g., Koller et al. 2003; Rafikov & Cimerman 2023) or material trapped around the planet's Lagrange points (e.g., Rodenkirch et al. 2021; Garrido-Deutelmoser et al. 2023).

Planet-driven gap opening is a process highly sensitive to the thermodynamics of the disk environment. Spiral shocks can heat up the surrounding disk (Rafikov 2016; Ziampras et al. 2020a), possibly changing the shape of such spirals (e.g., Zhu et al. 2015) when compared to a locally isothermal approach, where the thermal feedback of spirals on the disk is omitted. At the same time, Miranda & Rafikov (2020a) showed that introducing a finite cooling timescale $t_{\rm cool} = \beta \Omega^{-1}$ as a fraction of the orbital timescale $\Omega_{\rm K}^{-1}$ subjects planetary spirals to linear radiative damping concurrently with their nonlinear dissipation through shocking (Goodman & Rafikov 2001). This effectively reduces the amount of wave angular momentum that is transported further in the disk, thus weakening the capability of a single planet to open multiple gaps. In complementary studies to the above, Zhang & Zhu (2020) confirmed that radiative damping of spiral arms is strongest for $\beta \sim 1$, while Ziampras et al. (2020b) showed that hydrodynamical simulations with radiative cooling could produce fewer rings and gaps in the gas component of realistic ALMA disks compared to locally isothermal models (e.g., Zhang et al. 2018).

The above studies considered a cooling timescale motivated by the thermal emission of dust grains off of the top and bottom surfaces of the disk. While this is typically the dominant mechanism that sets the temperature profile of the disk (Chiang & Goldreich 1997), the diffusion of thermal energy across the disk midplane can be equally

or more important around planet-driven spirals due to the sharp temperature gradients around their crests (Goodman & Rafikov 2001). Miranda & Rafikov (2020b) (hereafter MR20b) explored the role of such in-plane cooling and showed that it may dramatically affect the angular momentum budget of spiral waves, potentially recovering a picture closer to the locally isothermal limit regarding planet-driven rings and gaps.

Their study, while supporting the idea that a single planet could, after all, be responsible for multiple rings and gaps, relied on a simple local model to approximate the effects of in-plane radiation transport. Instead of solving for the diffusion of heat along the disk and across spiral shocks directly, MR20b treated in-plane cooling as an additional effective local cooling channel, similar to the surface cooling described above. Therefore, the details of radiative diffusion across spiral arms, which could affect their angular momentum budget, were not explicitly captured in that work.

In this study, we combine the cooling models of Ziampras et al. (2020b) with the methodology of Miranda & Rafikov (2020a,b), while also including a direct treatment of in-plane radiative diffusion according to Ziampras et al. (2020a) (see Appendix D therein). This allows us to test different prescriptions for handling disk thermodynamics. In doing so, we aim to compare the above findings with those by Zhang & Zhu (2020) and the locally isothermal models of ALMA disks by Zhang et al. (2018), and investigate whether a more realistic treatment of thermodynamics could help recover the "single planet, multiple gaps" scenario.

Our physical framework is described in Sect. 2. Our methods and results are model-dependent, and are presented in Sects. 3 & 4. We discuss our findings in Sect. 5, and conclude in Sect. 6.

2 PHYSICS AND MOTIVATION

In this section we lay out our physical framework. We describe the radiative mechanisms that we consider in our models, and motivate our approach.

2.1 Protoplanetary disk hydrodynamics

We consider a disk of ideal gas with adiabatic index γ and mean molecular weight μ orbiting around a star with mass $M_\star$ and luminosity $L_\star$. In a vertically integrated framework the gas has a surface density Σ, specific internal energy e, and velocity field $\boldsymbol{u} = (u_R, u_\phi)$. The Navier–Stokes equations then read

$$\frac{\partial \Sigma}{\partial t} + \boldsymbol{u} \cdot \nabla \Sigma = -\Sigma \nabla \cdot \boldsymbol{u}, \tag{1a}$$

$$\Sigma \frac{\partial \boldsymbol{u}}{\partial t} + \Sigma(\boldsymbol{u} \cdot \nabla)\boldsymbol{u} = -\nabla P - \Sigma \nabla(\Phi_\star + \Phi_\mathrm{p}) + \nabla \cdot \overline{\boldsymbol{\sigma}}, \tag{1b}$$

$$\frac{\partial \Sigma e}{\partial t} + \boldsymbol{u} \cdot \nabla(\Sigma e) = -\gamma \Sigma e \nabla \cdot \boldsymbol{u} + Q \tag{1c}$$

The stellar potential at distance R is given by $\Phi_\star = -\mathrm{G}M_\star/R$, the viscous stress tensor is denoted by $\overline{\boldsymbol{\sigma}}$, and the vertically integrated pressure $P = (\gamma - 1)\Sigma e$ follows the ideal gas equation of state. The isothermal sound speed is then $c_\mathrm{s} = \sqrt{P/\Sigma} = \sqrt{\mathcal{R}T/\mu}$, with T being the (vertically constant) gas temperature, and relates to the adiabatic sound speed as $c_\mathrm{s}^\mathrm{ad} = c_\mathrm{s}\sqrt{\gamma}$. We can also define the pressure scale height $H = c_\mathrm{s}/\Omega_\mathrm{K}$, where $\Omega_\mathrm{K} = \sqrt{\mathrm{G}M_\star/R^3}$ is the Keplerian angular velocity, and the aspect ratio $h = H/R$. In the above, G and $\mathcal{R}$ denote the gravitational constant and ideal gas constant, respectively.

The term Q encompasses any additional sources or sinks of thermal energy.

Our models include an embedded planet with mass M_p, fixed at $R = R_\mathrm{p}$ in the form of a potential term in Eq. (1b). The planetary potential follows a Plummer prescription

$$\Phi_\mathrm{p} = -\frac{\mathrm{G}M_\mathrm{p}}{\sqrt{d^2 + \epsilon^2}}, \qquad \boldsymbol{d} = \boldsymbol{R} - \boldsymbol{R}_\mathrm{p} \tag{2}$$

with a smoothing length $\epsilon = 0.6\,H_\mathrm{p}$ that aims to account for the vertical stratification of the disk, similar to Müller et al. (2012). The planet is not allowed to migrate or accrete.

Finally, our models consider a reference frame centered on the star. While the indirect term by the star–planet system orbiting about their center of mass is included, we do not account for the disk's feedback on either object.

2.2 Radiative effects

The thermal structure of the disk is determined by a number of different physical processes represented by the terms in Eq. (1c). In this study, we consider the following radiative source and sink terms.

2.2.1 Viscous heating

Following the α-viscosity model of Shakura & Sunyaev (1973), the gas kinematic viscosity is $\nu = \alpha c_\mathrm{s}^\mathrm{ad} H$. We can then write:

$$Q_\mathrm{visc} = \frac{1}{2\nu\Sigma}\mathrm{Tr}(\overline{\boldsymbol{\sigma}}^2) \approx \frac{9}{4}\nu\Sigma\Omega_\mathrm{K}^2. \tag{3}$$

In this study, we explore the nearly-inviscid regime with $\alpha = 10^{-5}$. This allows our disk to essentially not accrete and remain passively heated by stellar irradiation, while minimizing the effects of numerical diffusion. Furthermore, for this low α, viscous damping of the angular momentum flux carried by spiral arms is negligible (MR20b).

2.2.2 Vertical cooling

The gas can transfer its thermal energy to nearby dust grains which then emit as a black body, cooling the disk in the process. Following Menou & Goodman (2004), we have

$$Q_\mathrm{cool} = -2\sigma_\mathrm{SB}\frac{T^4}{\tau_\mathrm{eff}}, \qquad \tau_\mathrm{eff} = \frac{3\tau_\mathrm{R}}{8} + \frac{\sqrt{3}}{4} + \frac{1}{4\tau_\mathrm{P}}, \tag{4}$$

where τ_eff is the effective optical depth according to Hubeny (1990), and $\tau_\mathrm{R,P} = \frac{1}{2}\kappa_\mathrm{R,P}\Sigma$ is the optical depth at the disk midplane for Rosseland and Planck mean opacities κ_R and κ_P, respectively. In our study, and in the analysis below, we assume that $\kappa_\mathrm{P} = \kappa_\mathrm{R}$ and $\tau_\mathrm{R} = \tau_\mathrm{P} = \tau$.

2.2.3 Stellar irradiation

The disk intercepts part of the incoming radiation by the star. Similar to Menou & Goodman (2004) we can approximate this effect as

$$Q_\mathrm{irr} = 2\frac{L_\star}{4\pi R^2}(1 - \epsilon)\frac{\theta}{\tau_\mathrm{eff}}, \tag{5}$$

where we adopt $\epsilon = 0.5$ for the disk albedo, θ is the flaring angle, and τ_eff is the effective optical depth similar to Eq. (4). In principle, θ is radius-dependent (see e.g., Chiang & Goldreich 1997; Menou & Goodman 2004), but constant small values of the order of $\theta = 0.02$–0.03 are often used in the literature (e.g., Zhang et al. 2018). By equating $Q_\mathrm{irr} = Q_\mathrm{cool}$ and assuming $\kappa_\mathrm{R} = \kappa_\mathrm{P}$, one finds the well-known passively-heated disk temperature profiles $T(R) \propto R^{-0.5}$ for $\theta = \mathrm{const.}$, or $T(R) \propto R^{-3/7}$ for $\theta = R\frac{\mathrm{d}h}{\mathrm{d}R}$.

2.2.4 Radiative diffusion

Dust grains emit in all directions, and as a result a local perturbation of temperature drives a radiative flux not only vertically but also through the disk midplane. We model this process using the flux-limited diffusion approximation (Levermore & Pomraning 1981), by writing

$$Q_{\rm rad} = -\sqrt{2\pi}H\,\nabla\cdot\left(\lambda\frac{4\sigma_{\rm SB}}{\kappa_{\rm R}\rho_{\rm mid}}\nabla T^4\right). \tag{6}$$

Here, $c_{\rm v} = \frac{\mathcal{R}}{\mu(\gamma-1)}$ is the heat capacity of the gas at constant volume, $\rho_{\rm mid} = \frac{1}{\sqrt{2\pi}}\frac{\Sigma}{H}$ is the volume density at the disk midplane, and λ is a flux limiter following Kley (1989):

$$\lambda(R') = \begin{cases} \frac{2}{3+\sqrt{9+10R'^2}}, & 0 \leqslant R' \leqslant 2 \\ \frac{10}{10R'+9+\sqrt{180R'+81}}, & 2 \leqslant R' \leqslant \infty. \end{cases}, \quad R' \equiv 4l_{\rm rad}\frac{|\nabla T|}{T}, \tag{7}$$

where $l_{\rm rad} = 1/(\kappa_{\rm R}\rho_{\rm mid})$ is the photon mean free path. This flux limiter handles the transition from the optically thick, diffusion limit ($\lambda \to 1/3$) to the optically thin, free-streaming limit ($\lambda \to 1/R'$).

2.3 Implementing radiative effects

With the above in mind, there are (at least) 4 different recipes of handling gas thermodynamics in our framework. From simplest to most complex:

(i) **locally isothermal**: the energy equation is dominated by radiative processes (Q in Eq. (1c)) such that any excess cooling/heating due to (de)compression is removed instantaneously and the gas is always in thermal equilibrium. The energy equation is then omitted. Instead, we assume that $T(R)$ is time-independent and given by a balance between $Q_{\rm irr}$ (heating) and $Q_{\rm cool}$ (cooling), computed *a priori*.

(ii) **adiabatic**: changes in temperature solely depend on gas (de)compression (first term on RHS of Eq. (1c)), and no radiative effects are considered; entropy is conserved in the absence of shocks.

(iii) **surface cooling**: the gas can heat up via shock heating, $Q_{\rm irr}$, and $Q_{\rm visc}$, and cool through the disk surfaces via $Q_{\rm cool}$. In our vertically-integrated models, this cooling mechanism is local.

(iv) **fully radiative**: in-plane radiative diffusion is added to the above by including $Q_{\rm rad}$.

It is quite common to account for the mechanisms outlined above using a parametrized approach. Assuming the gas cools over a cooling timescale $t_{\rm cool}$, we define the dimensionless cooling timescale $\beta = t_{\rm cool}\Omega_{\rm K}$. This β can be a constant, allowing for scale-free models, or a function of one or more physical quantities, allowing for more realistic models that might explore certain mechanisms or target specific observations. In that case, one can assume that there exists an equilibrium temperature profile $T_0(R)$ that remains largely unchanged in time (e.g., as the result of a balance between $Q_{\rm irr}$ and $Q_{\rm cool}$), and then remove excess heat from the disk over a timescale β such that

$$Q_{\rm relax} = -\Sigma c_{\rm v}\frac{T - T_0}{\beta}\Omega_{\rm K}. \tag{8}$$

This approach allows the disk temperature profile to remain reasonably close to T_0 while also accounting for the interplay between compression heating and radiative terms. This is useful in studying the radiative damping of processes such as the spiral arms of a planet (e.g., Zhang & Zhu 2020; Miranda & Rafikov 2020a,b), the opening of a circumbinary cavity (e.g., Sudarshan et al. 2022; Pierens &

Nelson 2022), the development of various instabilities (e.g., Béthune et al. 2021; Pfeil & Klahr 2021), etc.

Equations (4) & (6) allow us to define cooling timescales associated with surface (vertical) and in-plane cooling respectively. Following Ziampras et al. (2020b), the surface cooling timescale $\beta_{\rm surf}$ can be written as

$$\beta_{\rm surf} = \frac{\Sigma e}{|Q_{\rm cool}|}\Omega_{\rm K}. \tag{9}$$

Treatment of the in-plane cooling is not as simple, because heat diffuses through the midplane on a timescale that depends on the diffusion lengthscale $l_{\rm d}$. Taking this into account, MR20b argued that the in-plane cooling of the m-th azimuthal Fourier component of the thermal perturbation can be approximated via a β-cooling term with the cooling timescale $\beta_{\rm mid}^m$ given by

$$\beta_{\rm mid}^m \approx \frac{\Omega_{\rm K}}{\eta}\left[(l_{\rm d}^m)^2 + \frac{l_{\rm rad}^2}{3}\right], \qquad \eta = \frac{16\sigma_{\rm SB}T^3}{3\kappa_{\rm R}\rho_{\rm mid}^2 c_{\rm v}}, \tag{10}$$

where the two terms in the square brackets represent the characteristic photon diffusion lengthscales in the optically thick and thin limits, respectively. Here, η is a radiative diffusion coefficient associated with Eq. (6). MR20b further suggested that the cooling effect of many Fourier perturbation harmonics can be reduced approximately to a local cooling term with a single $\beta_\star$ given by Eq. (10) with $l_{\rm d}^m$ replaced with $l_{\rm d}^\star = 1/k_\star$, such that

$$k_\star \approx \frac{1}{l_{\rm rad}}\left(\frac{\pi}{2}\right)^{1/2}\left|\frac{\Omega_{\rm p}}{\Omega} - 1\right|\frac{m_\star}{\tau}, \quad m_\star \approx \frac{1}{2h_{\rm p}}. \tag{11}$$

In our present study, we simplify this local prescription even further by assuming that the dominant diffusion lengthscale is $l_{\rm d}^\star \approx H$ (e.g., Flock et al. 2017), and write

$$\beta_{\rm mid} \approx \frac{\Omega_{\rm K}}{\eta}\left(H^2 + \frac{l_{\rm rad}^2}{3}\right). \tag{12}$$

We note that the dominant diffusion lengthscale will be smaller than H near shocks. As a result, this approach is only an approximation and cannot allow a direct comparison with the results of MR20b. We nevertheless carry out a comparison in Appendix F, finding that the recipe of MR20b does not always capture the the behavior of the spiral AMF appropriately.

Finally, we can combine the two timescales $\beta_{\rm surf}$ and $\beta_{\rm mid}$ into an effective cooling timescale $\beta_{\rm tot}$ that accounts for both vertical and in-plane cooling. Following MR20b, we write

$$\beta_{\rm tot}^{-1} = \beta_{\rm surf}^{-1} + \beta_{\rm mid}^{-1}. \tag{13}$$

With some algebra, we can show using Eqs. (4), (9), (10), (12) that in our approach with $l_{\rm d} = H$, $\beta_{\rm tot}$ can then be written as

$$\beta_{\rm tot} = \frac{\beta_{\rm surf}}{1+f}, \qquad f = \frac{\beta_{\rm surf}}{\beta_{\rm mid}} = 16\pi\frac{\tau\tau_{\rm eff}}{6\tau^2 + \pi}, \tag{14}$$

implying that models that do not account for in-plane cooling will considerably overestimate the cooling timescale, as also shown by MR20b.

An example of β profiles for a passively-heated disk mimicking the setup of MR20b is shown in Fig. 1. As they showed, $\beta_{\rm mid}$ is approximately 3–4 times lower than $\beta_{\rm surf}$, pushing $\beta_{\rm tot}$ to ≈ 0.2–$0.25\beta_{\rm surf}$. This highlights the necessity of treating in-plane cooling, which MR20b demonstrated in their study.

However, a caveat of implementing in-plane cooling with this β approach is that it is inherently localized: a local increase in temperature should result in both surface cooling, which cools the hot spot over time depending on $\beta_{\rm surf}$, as well as radiative diffusion. The latter spreads the heat to neighboring areas, cooling the hot spot over
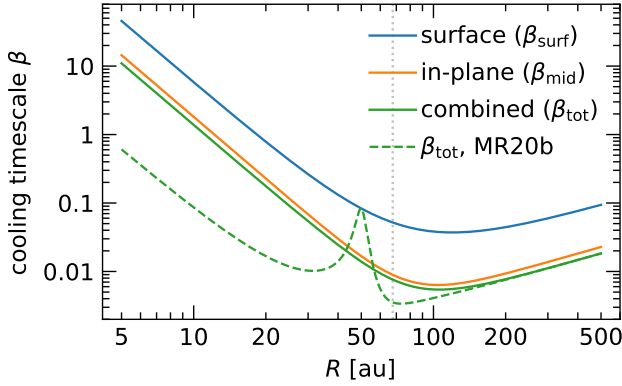
Figure 1. Estimates of the cooling timescale using Eqs. (9), (10) & (13) for an example system similar to the setup of MR20b. The in-plane cooling timescale β_{mid} is about 3–4 times lower than the surface cooling time β_{surf}, highlighting the importance of in-plane cooling. Here, we chose $M_\star = M_\odot$, $R_0 = 50$ au, $\mu = 2$, $h_0 = 0.1$, $\Sigma_0 = 10$ g/cm^2, $\Sigma \propto R^{-1}$, $T \propto R^{-1/2}$, and $\kappa = 2\times10^{-4}T^2$. The solid green line assumes $l_d = H$ (Eq. (12)), while dashed green uses the recipe by MR20b with $l_d \approx 1/k_\star$ (Eq. (11)) and $R_p = 50$ au.

Figure 2. Cartoon example of the temperature profile around a one-dimensional shock front (at $x = 0$). The higher post-shock temperature T_{post} can relax to the pre-shock temperature T_{pre} at a rate that depends on the cooling timescale β (faster for smaller β). Accounting for radiative diffusion further smears the shock front, resulting in a heating of the pre-shock region as well. Of course, spiral shocks are more complex than this example, with Σ and T reverting to their pre-shock values after a few H (see Fig. 14).

β_{mid} but also heating up its surroundings. This last heating effect is missed when using β_{tot} as a proxy for in-plane cooling, as this approach always relaxes the gas towards the equilibrium state T_0. This can result in significant differences around regions of sharp temperature gradients (such as shocks) between a β_{tot}-cooled model and one that explicitly accounts for Q_{rad}.

In the context of planet–disk interaction, MR20b have already shown that in-plane cooling through β_{tot} can affect gap opening by a planet embedded in a disk by essentially allowing the gas to cool through the midplane. This approach, however, cannot capture the heating due to thermal diffusion around a spiral shock, which could further affect the gap opening process. For that, directly accounting for Q_{rad} is required.

To illustrate our point, we consider a one-dimensional shock front in Fig. 2. Accounting for surface cooling relaxes the hot, post-shock region to its initial, pre-shock state over a lengthscale that depends on the cooling timescale (blue curves). By treating in-plane cooling as a local effect (i.e., using β_{mid}) we essentially enable an additional cooling pathway for the gas, reducing the effective β from β_{surf} to β_{tot}. However, when handling in-plane cooling via radiative diffusion, the temperature spike at the shock front spreads into both the pre- and post-shock regions (green curve). Considering how the deposition of angular momentum by planet-driven spiral arms and the subsequent gap opening are very sensitive functions of the radiative damping of such shocks, it becomes clear that a more realistic treatment of the thermodynamics could have an impact on planet–disk interaction in this context.

With fully radiative models (i.e., accounting for Q_{rad}), we aim to explore the interplay between surface and in-plane cooling and combine the findings of Miranda & Rafikov (2020a,b), Zhang & Zhu (2020), and Ziampras et al. (2020b) into one complete, coherent picture.

We investigate two different scenarios: we start with a wide parameter study that aims to quantify the planet's gap opening capacity with a suite of short-term, high resolution numerical models, presented in Sect. 3. We then use our models to identify different regimes of cooling, and demonstrate the effect of in-plane radiative diffusion with long-term simulations and simplistic synthetic obser-
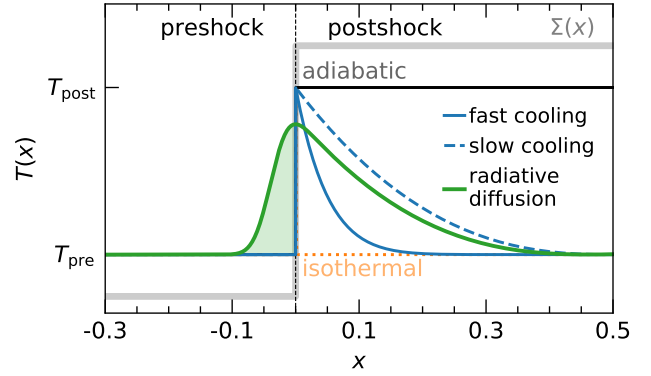
vations of continuum emission for two ALMA systems, AS 209 and Elias 20, in Sect. 4. The specifics of the numerical setup for each set of models are covered in their respective sections.

3 THE ROLE OF COOLING ON SPIRAL AMF

In this section we present a high-resolution, wide parameter study of radiative planet–disk interaction. Our aim is to measure the planet's gap opening capabilities using the angular momentum flux (AMF) carried by the planet-driven spiral arms as a metric of the momentum transport and deposition in the disk. Following MR20b, in the linear approximation we define the spiral AMF as

$$F_J(R) = R^2\bar{\Sigma}\oint_\phi \delta u_R\, \delta u_\phi \mathrm{d}\phi, \qquad \delta u_x = u_x - \bar{u}_x, \qquad x \in \{R, \phi\} \qquad (15)$$

where a bar denotes an azimuthally averaged quantity. The AMF can then be normalized to a reference value (Goldreich & Tremaine 1980)

$$F_{J0} = \left(\frac{M_p}{M_\star}\right)^2 h_p^{-3}\Sigma_p R^4 \Omega_p^2. \qquad (16)$$

In order to quantify the effects of cooling, both surface and in-plane, we explore the β_{surf} parameter space for two different values of the optical depth τ. Varying the cooling timescale β_{surf} allows us to study the spiral AMF as the disk transitions from adiabatic ($\beta_{surf} \to \infty$) to isothermal ($\beta_{surf} \to 0$), while varying the optical depth τ controls the efficiency of both surface cooling and in-plane radiative diffusion. We then carry out models for the equations of state iso, adb, surf, and rad, listed in Table 1.

In the following sections we describe our initial and boundary conditions for each model, and describe our findings.

3.1 Model setup and methods

We utilize the numerical hydrodynamics code PLUTO v4.4 (Mignone et al. 2007) in a 2D cylindrical polar $\{R, \phi\}$ geometry. PLUTO uses a finite-volume Godunov scheme to integrate Eqs. (1)

in time using a Riemann solver. Our computational domain extends radially in the range $R \in [0.4\text{--}2.5]\,R_{\mathrm{p}}$, with $R_{\mathrm{p}} = 1$ au, and covers the full azimuthal range $\phi \in [0, 2\pi]$.

In our models we also use the FARGO method by Masset (2000), which has been implemented into PLUTO by Mignone et al. (2012) and yields a speedup factor of ~ 10 alongside improving the accuracy of the numerical scheme. This is achieved by subtracting the background Keplerian flow before solving the Riemann problem across all cell interfaces, substantially relaxing timestep limitations in the rapidly rotating inner disk.

Regarding our numerical setup, we use second-order accurate reconstruction and time marching schemes (LINEAR and RK2 respectively) with the hllc Riemann solver (Toro et al. 1994) and the VAN_LEER slope limiter Van Leer (1974). To minimize dissipation we make use of a frame corotating with the planet, and activate the option CHAR_LIMITING. Using a higher-order reconstruction did not affect our results.

Throughout this parameter scan we keep the (initial) aspect ratio at the planet's position fixed at $h_{\mathrm{p}} = 0.05$ and the planet's mass at $M_{\mathrm{p}} = 0.3\,M_{\mathrm{th}} = 3.75 \times 10^{-5}\,M_{\star}$. Here, $M_{\mathrm{th}} \approx h_{\mathrm{p}}^3 M_{\star}$ is the disk thermal mass at the planet's location (Rafikov 2002a).

To create an environment as simple as possible, we set Σ, β_{surf} and τ to be constant throughout the domain. Given the nonlinear dependence of β_{surf} and τ on Σ, T, and κ, this requires unphysical choices for Σ and κ for some of our models. Nevertheless, we will later show that our approach can be justified by looking at several ALMA systems observed with the DSHARP survey (Andrews et al. 2018) in the $\beta_{\mathrm{surf}}\text{--}\tau$ space. As a result, our choices for Σ and κ (to control β_{surf} and τ) as well as $L_{\star}$ (to control h_{p}) can be completely arbitrary.

By keeping Σ and κ constant, we can have radially constant profiles for τ. By further adopting a temperature profile $T(R) \propto R^{-0.5}$ we can also maintain a constant β_{surf}. This allows for constant parameters throughout the disk while being consistent with a realistic, passively irradiated disk model. A table with all values of Σ and κ that we explored is provided in Appendix G.

Our fiducial model consists of a disk with $\tau = 10$, $\beta_{\mathrm{surf}} = 1$, $h_{\mathrm{p}} = 0.05$, and $M_{\mathrm{p}} = 0.3\,M_{\mathrm{th}}$. After a resolution analysis, presented in detail in Appendix A, we opt for a resolution of 32 cells per scale height in both directions at the planet's location, for a fiducial grid of 1200×4096 cells.

We use a locally isothermal and an adiabatic model (labeled "iso" and "adb") with $h_{\mathrm{p}} = 0.05$ and $M_{\mathrm{p}} = 0.3\,M_{\mathrm{th}}$ as references. Then, we also run models with surface cooling only ("surf"), which keep only Q_{irr} and Q_{cool} nonzero, and models with both surface and in-plane cooling ("rad"), which additionally include the in-plane radiative diffusion term Q_{rad}. These four sets of models, which represent the different recipes listed in Sect. 2.3, are described in more detail in Table 1.

We note here that we use β_{surf} as a parameter to control the efficiency of cooling in models surf and rad. These models utilize the radiative terms introduced in Sect. 2.2 and implemented in Eq. (1c), with disk parameters tailored such that the surface cooling timescale β_{surf} (defined via Eq. (9)) has a desired value. It is very important to keep in mind that β_{surf} is used here only for parametrization of these models and does not assume a β cooling prescription, which would implement β_{surf} via Eq. (8). For a comparison with models using β cooling, see Sect. 3.3.

In all models the planet grows over one orbit using the formula by de Val-Borro et al. (2006), and each model is run for ten planetary orbits so that the global structure of the spiral AMF and the thermal structure of the disk around spirals are established. In addi-

tion, wave damping zones are employed in the radial direction between $R \in [0.4, 0.5] \cup [2.1, 2.5]\,R_{\mathrm{p}}$ following de Val-Borro et al. (2006), with a damping timescale of 0.1 orbits at the corresponding radial boundary. This ensures that spiral arms are not reflected at the boundary edge, which is otherwise closed.

As we will show in the next section, it is difficult to quantify the differences in radial profiles of F_J between our models due to the size of our parameter study (over 200 models). For this reason, we define a proxy for the planet's "capacity to open multiple gaps" for a given thermodynamic prescription as

$$G = \frac{J}{J^{\mathrm{adb}}}, \qquad J = \int_{0.5\,R_{\mathrm{p}}}^{1.5\,R_{\mathrm{p}}} F_J \, \mathrm{d}R, \qquad (17)$$

where J^{adb} represents the adiabatic case. Effectively, a high G translates to weak radiative wave damping and a multiple-gap configuration ($\beta \to 0$ or $\beta \to \infty$), whereas a low G translates to strong wave damping and a single gap around the planet and is expected for $\beta_{\mathrm{surf}} \sim 1$ (see Fig. 4 in Miranda & Rafikov 2020a). We choose these radial limits to cover the entire inner disk, but exclude the outskirts of the disk to avoid some oscillations in the radial profiles of the AMF (see e.g., Fig. 5) which slightly polluted our results without changing their general behavior.

3.2 Results

In this section we present our results for the models described in Sect. 3.1 in the context of spiral AMF.

3.2.1 Fiducial model

We start by looking at the global structure of the disk for our models with $\tau = 10$, $\beta_{\mathrm{surf}} = 1$, $h_{\mathrm{p}} = 0.05$, $M_{\mathrm{p}} = 0.3\,M_{\mathrm{th}}$ at $t = 10$ orbits to provide some context. Figures 3 & 4 show the global deviations in surface density Σ and temperature T respectively, after subtracting the azimuthally averaged disk profile.

As expected, these deviations are localized around the planet-driven spiral arms, with minor features inside the planet's corotating region. In particular, model surf shows weaker spiral arms compared to other models (Fig. 3), while spirals in model rad are significantly more diffuse in terms of temperature contrast compared to those in models surf or adb (Fig. 4).

The latter is to be expected: the model with in-plane radiation transport allows diffusion of heat along the midplane, which smears temperature gradients about the crests of spiral arms. In addition, it has an effectively shorter cooling timescale since the gas can cool via both the vertical and in-plane channel. As a result, model rad behaves more akin to a locally isothermal model—where no temperature deviations are observed—as far as temperature is concerned.

We then plot radial profiles of the normalized spiral AMF F_J/F_{J0} for these four different thermodynamic assumptions in Fig. 5, also showing for illustration the values of G defined by Eq. (17) for each model. Since $\beta_{\mathrm{surf}} = 1$, the AMF in model surf is weaker than either the isothermal or adiabatic limits. At the same time, we find a considerably higher F_J in model rad compared to surf for $\beta_{\mathrm{surf}} = 1$, consistent with the findings of MR20b.

Table 1. Models referenced throughout this study. Model `surf` aims to reproduce Ziampras et al. (2020b), while `bsurf` and `btot` aim to reproduce MR20b. Model pairs `surf`–`bsurf` and `rad`–`btot` consider the same physics with different implementations (source terms and β cooling, respectively). Models `iso`, `adb`, `surf`, and `rad` are introduced in Sect. 3.1 and discussed throughout Sect. 3, while models `bsurf` and `btot` are discussed in Sect. 3.3.

tag	physics	relevant terms and equations
`iso`	locally isothermal (instantaneous cooling)	no energy equation
`adb`	adiabatic (no cooling)	adiabatic compression: (1c)
`surf`	surface cooling only	(1c); Q_{irr}: (5), Q_{cool}: (4)
`rad`	surface cooling and in-plane radiative diffusion	(1c); Q_{irr}: (5), Q_{cool}: (4), Q_{rad}: (6)
`bsurf`	β relaxation with surface term	(1c); Q_{relax}: (8), β_{surf}: (9)
`btot`	β relaxation with surface and in-plane terms	(1c); Q_{relax}: (8), β_{tot}: (13)
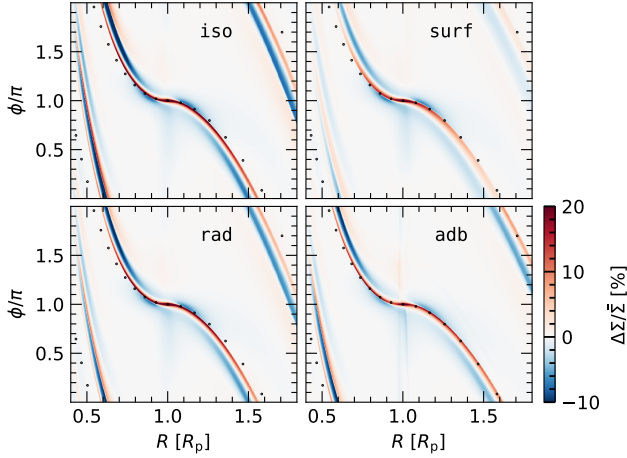
Figure 3. Two-dimensional heatmaps of surface density perturbations in our fiducial model for various thermodynamic treatments. Planet-driven spiral arms permeate the disk, with minor deviations around the planet's corotating region as well. Model `surf` shows fainter spirals due to the effects of surface cooling, while accounting for in-plane radiation transport (model `rad`) helps recover a more isothermal-like picture. Dots help guide the eye, showing the different propagation speed of spirals in different panels.
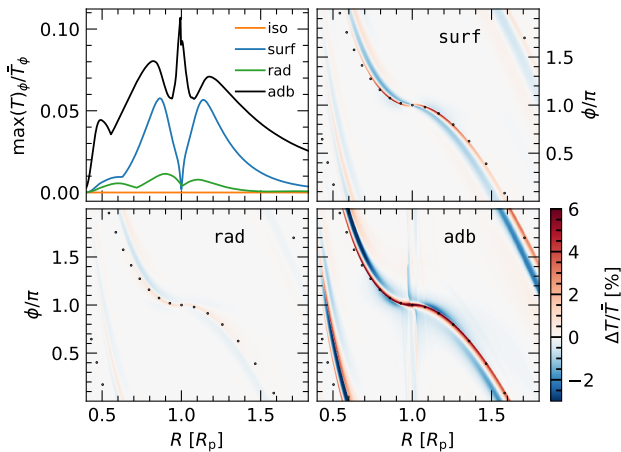
Figure 4. Temperature deviations plotted similar to Fig. 3. Depending on the rate of cooling, the temperature contrast around spiral arms can be very weak (`rad`) to very strong (`adb`). The top left panel shows the maximal contrast at a given radius to quantify these differences. Temperature spikes near $R = 0.5$–0.6 in the inner disk exist due to the excitation of secondary spiral shocks, while a spike at R_p in the adiabatic model corresponds to hot material in the planet's Hill sphere.
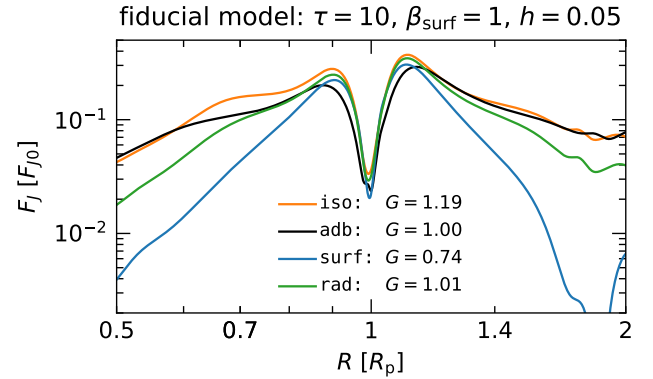

Figure 5. Normalized angular momentum flux (AMF) F_J due to spiral arms for our fiducial model as a function of equation of state. Values of the metric G defined by Eq. (17) are indicated for each model. We see a distinctly lower AMF for model `surf`, consistent with previous findings given that $\beta_{surf} = 1$.

3.2.2 Dependence of spiral AMF on β_{surf}

We now build on our fiducial model by repeating our sets of runs for various values of β_{surf}. Figure 6 shows an example of the dependence of F_J on β_{surf} for model `surf`, with F_J exhibiting fastest damping for $\beta_{surf} \approx 1$ in agreement with results by Miranda & Rafikov (2020a) and Zhang & Zhu (2020).

We then compute the factors G^{surf} and G^{rad} via Eq. (17) and plot them as a function of β_{surf} in Fig. 7. Focusing on the solid blue and green curves, three main takeaways can be extracted from this figure:

(i) We confirm that radiative damping of planet-driven density waves is strongest and therefore the planet's capacity to open multiple gaps is weakest for $\beta_{surf} \approx 1$ in a model where surface emission is the only cooling channel, in agreement with Zhang & Zhu (2020).

(ii) We recover both the locally isothermal and adiabatic limits at very small and very large β_{surf} respectively, adding to the robustness of our method.

(iii) Accounting for both surface cooling and in-plane radiative diffusion (`rad` models) does not always imply weaker radiative damping of spirals. Instead, the corresponding curve of $G(\beta_{surf})$ in Fig. 7 is simply shifted to larger β_{surf} compared to model `surf` but behaves similarly otherwise.

The last point can be explained with the same argument as in the previous section: in-plane radiative diffusion enables an additional cooling channel for spiral shocks and allows them to behave more akin to locally isothermal shocks.

The above also implies that there is a critical cooling timescale β_{crit} for which the G parameter in `rad` models shows a minimum,
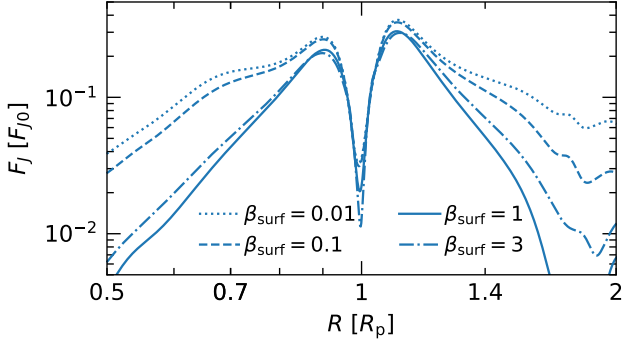
Figure 6. Dependence of the spiral AMF on the surface cooling timescale β_{surf} for model `surf`, similar to Fig. 5. We find a minimum in F_J for $\beta_{\text{surf}} \approx 1$, consistent with previous reports (Miranda & Rafikov 2020a).

similarly to how $\beta_{\text{crit}} = \beta_{\text{surf}}$ for `surf` models. Assuming that in this case $\beta_{\text{crit}} = \beta_{\text{tot}}$, we define the factor f through Eq. (14) and then shift the blue curve in Fig. 7 towards the right by $\beta_{\text{surf}}/\beta_{\text{tot}} = f + 1 \approx 4.51$. We find that the `rad` (in green) and shifted `surf` curves (faint blue), do not quite match, highlighting that the non-local radiative diffusion in our `rad` models cannot be modeled accurately with a local cooling approach. Instead, the `surf` models would need to be shifted by a further factor of ≈ 1.5 in order to overlap with the `rad` curve.

We note that f is a sensitive function of τ, peaking at $\tau \approx 0.6$ ($f \approx 6.1$, see Fig. 8). Essentially, this means that the green curve in Fig. 7 would shift further to the right (towards longer β, as in-plane cooling would be more efficient), while the blue curve would not be displaced as in-plane cooling is not considered. To showcase this effect, and to add to the robustness of our results, we repeat our models for $\tau = 1$ and present our findings in Fig. 9. Once again the `rad` and shifted `surf` models do not quite match, with the two curves being offset by a factor of ≈ 2, but at the qualitative level the agreement between these models is reasonably good.

3.3 Comparison with various β-cooling prescriptions

Until this point we implemented our heating and cooling prescriptions similar to Ziampras et al. (2020b) by directly adding radiative source terms (Q_{irr}, Q_{cool}, Q_{rad}) to the energy equation. Previous studies on this topic, however, used the simpler and more popular β-cooling approach by defining the cooling timescale similarly to Eqs. (9) or (13) and then implementing a relaxation term similar to Q_{relax} in Eq. (8) in the energy equation (e.g., Miranda & Rafikov 2020a,b; Zhang & Zhu 2020). While our findings agree qualitatively with these studies in that gap opening is weakest in models with surface cooling and $\beta_{\text{surf}} \sim 1$ (Miranda & Rafikov 2020a; Zhang & Zhu 2020) and that in-plane cooling dramatically changes spiral AMF (Miranda & Rafikov 2020b), we would like to quantify the differences between the two approaches.

To investigate this, we carry out two more sets of models to complement our `surf` and `rad` runs. By substituting the balance between Q_{irr} and Q_{cool} with a cooling term Q_{relax} where $\beta = \beta_{\text{surf}}$ (models tagged "`bsurf`"), and by further implementing in-plane cooling via $\beta = \beta_{\text{tot}}$ through Eqs. (12) & (13) in a similar manner (models "`btot`"), we aim to compare models in the pairs `surf`–`bsurf` and `rad`–`btot` with each other. Table 1 provides a description of all physical options and implementations.

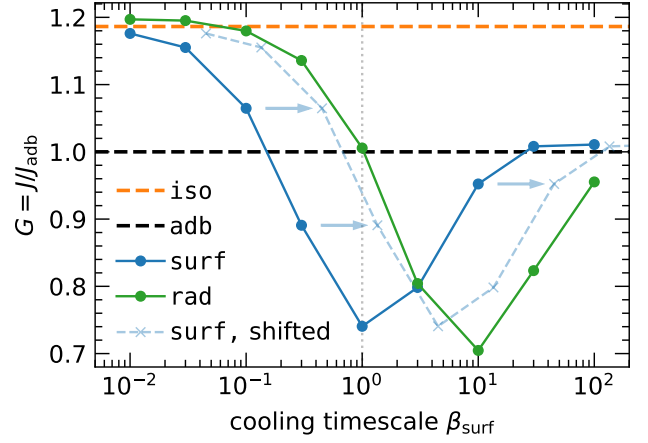Figure 10 shows $G(\beta_{\text{surf}})$ for the four approaches described above

Figure 7. Dependence of the normalized, radially integrated AMF G (Eq. (17)) on the surface cooling timescale β_{surf} for our models with $\tau = 10$ and various treatments of cooling. We note that model `surf` (blue curve) shows a minimum for $\beta_{\text{surf}} \approx 1$, which can also be inferred from Fig. 6. Including radiative diffusion (model `rad`, green) should then simply shift the curve towards larger β_{surf} by a factor of $\beta_{\text{surf}}/\beta_{\text{tot}}$ (see Eq. (14)), which is not exactly the case. Shifting the blue curve by $\beta_{\text{surf}}/\beta_{\text{tot}}$ (dashed blue) instead shows that the two curves are still offset horizontally by a factor of ≈ 1.5.
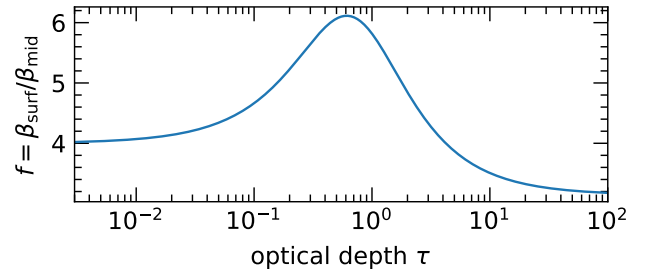

Figure 8. The factor f in Eq. (14) as a function of the optical depth τ.

and for $\tau = 10$. The purple curve, which uses Eqs. (9), (12), and (13), is offset towards longer β when compared to the red curve (prescription of Miranda & Rafikov 2020a; Zhang & Zhu 2020) by the factor $1 + f$ described in Eq. (14), showing that their findings using a β-cooling approach are qualitatively consistent with our explicit treatment of heating and cooling terms in terms of spiral AMF. The same behavior can be seen in Fig. 11, where $\tau = 1$.

However, these figures also show that model pairs that treat the same cooling mechanisms with different implementations (`surf`–`bsurf` and `rad`–`btot`) do not match each other. We identify that the reason behind this mismatch is inherent to the treatment of cooling using a β approach where $Q_{\text{relax}} \propto \beta^{-1} \propto Q_{\text{cool}} \propto T^4$, which overestimates the true cooling timescale by a factor of 4 when small deviations around an equilibrium state are considered. Our setup reflects such a case, with temperature deviations in the order of $\delta T/T_0 \lesssim 0.1$ around spiral arms for $M_{\text{p}} = 0.3\,M_{\text{th}}$ (see top left panel in Fig. 4). This factor of 4 has also been reported by Dullemond et al. (2022). We provide a detailed derivation of this factor in Appendix C.

Shifting β models to the right by this factor of 4 and further adjusting models `surf` and `bsurf` by the factor $1 + f$ as in Sect. 3.2.2 (faint curves in Figs. 10 & 11) shows that, once again, no prescrip-
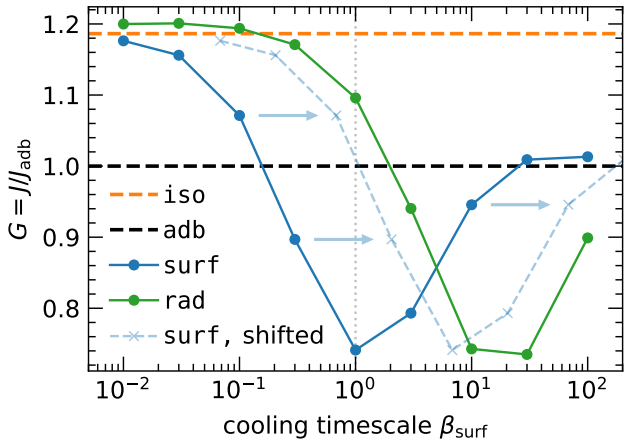
Figure 9. Similar to Fig. 7 but for models with $\tau = 1$, such that the ratio of surface to in-plane cooling is larger. Here, we find that the offset between G^{surf} and G^{rad} requires a further correction of β_{surf} by a factor of ≈ 2.
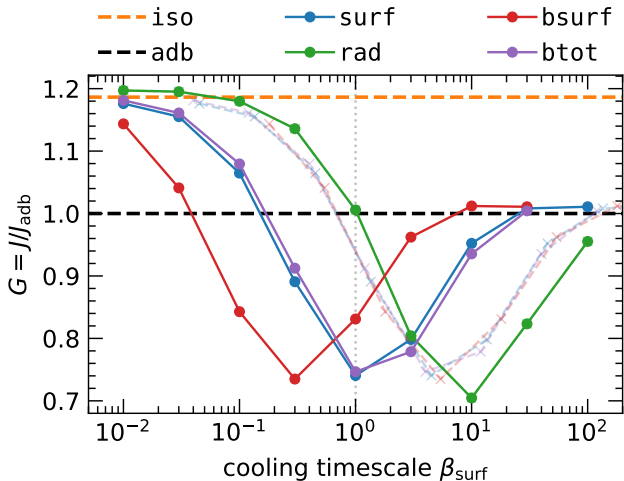

Figure 10. The parameter $G(\beta_{\mathrm{surf}})$ similar to Fig. 7 ($\tau = 10$), this time including models where the β cooling approach was used (bsurf, btot). They are shifted towards shorter β by a factor of 4 compared to models with explicit source terms (surf and rad, respectively). Accounting for this factor and the ratio $\beta_{\mathrm{surf}}/\beta_{\mathrm{tot}}$ (when applicable) shows that all local cooling prescriptions behave similarly (faint curves), matching rad models only qualitatively.

tion matches the results with FLD (rad) at the quantitative level. This highlights the importance of modeling FLD self-consistently instead of using a local approximation.

We note that the overlap between models surf (blue) and btot (purple) in Fig. 10 before shifting is simply a coincidence. Model surf overestimates the true cooling timescale β_{tot} by a factor of $1+f$ due to the lack of in-plane cooling (see Eq. (14)), while model btot overestimates β_{tot} by a factor of 4 due to using β cooling. Since $1 + f \approx 4.5$ for $\tau = 10$, the two curves seemingly overlap. For $\tau = 1$ (Fig. 11) we have $1 + f \approx 6.8$, and the two curves are distinctly different.

To test the limits of our approach in the nonlinear regime, we repeated a series of models with $\tau = 10$ and $M_{\mathrm{p}} = 3\,M_{\mathrm{th}}$. Our results
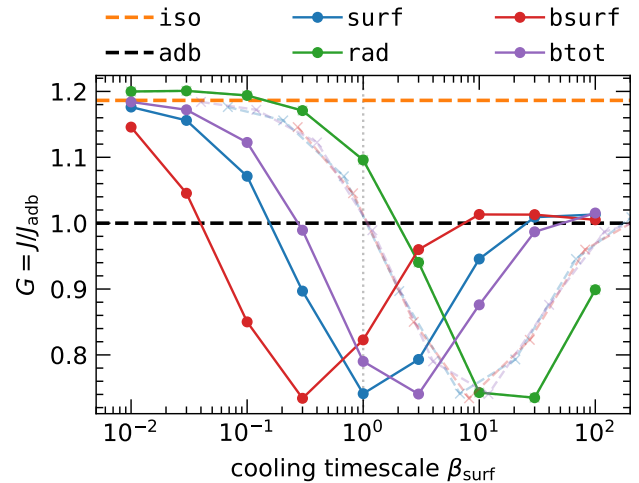

Figure 11. Same as Fig. 10, for $\tau = 1$, showing that our models are robust regardless of τ. Here it becomes clear that the overlap of models surf and btot for $\tau = 10$ (blue and purple curves in Fig. 10) was simply a coincidence.

can be found in Appendix D, where we show that the quoted factor of 4 reduces to approximately 3.5 for this more massive planet. To further check whether our choice of h_{p} could affect our results, we carried out a set of models with $h_{\mathrm{p}} = 0.1$ and a resolution of 32 cells per scale height ($N_R \times N_\phi = 600 \times 2048$) and found very good agreement with our models with $h_{\mathrm{p}} = 0.05$. These results are shown in Appendix E.

3.4 An approximate recipe for FLD

Based on the above, one could in principle account for all necessary factors that differentiate the β models using Eqs. (12) & (13) and our fully radiative FLD models and arrive at a model that behaves similar to model rad in terms of AMF. To test this, we design an adjusted cooling timescale that aims to account for the factor 4 discussed above (see Appendix C) and write:

$$\beta_{\mathrm{adj}} = \frac{1}{4a}\beta_{\mathrm{tot}} = \frac{1+f}{4a}\beta_{\mathrm{surf}}, \qquad a \approx 1.5\text{--}2, \qquad (18)$$

where β_{surf} and f are defined through Eqs. (9) & (14), respectively, and a is a correction factor motivated by the offset between surf and rad models in Figs. 10 & 11. A comparison between a set of models using this recipe—tagged "adj"—and our rad models in terms of $G(\beta_{\mathrm{surf}})$ is shown in Fig. 12 for $\tau = 1$, and we find excellent agreement between the two sets of models as far as this metric is concerned for $a = 2$.

However, while G is the same between the two sets of models, that does not mean that the radial AMF profiles $F_J(R)$ will also match, as seen in Fig. 13 for $\tau = 1$, $\beta_{\mathrm{surf}} = 1$ and $a = 2$. We find that F_J is higher (lower) in the inner (outer) disk for model adj compared to model rad, even though integrating F_J radially yields the same result.

Furthermore, the two sets of models are quite different in terms of the structure of the spiral arms themselves. As we discussed in Sect. 2 and highlighted in Fig. 1, our FLD models allow the diffusion of heat into the pre-shock region, something that β models fail to capture.

This effect is illustrated in Fig. 14, where we compare the sur-
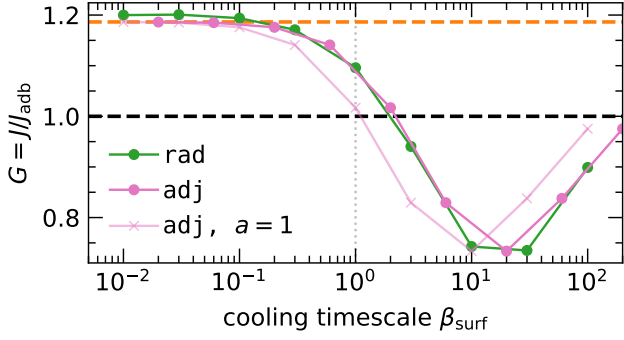
Figure 12. Similar to Fig. 11 ($\tau = 1$), but comparing models with explicit treatment of FLD (green) to the approximate recipe given by Eq. (18) that accounts for in-plane cooling through β (pink). We find an excellent agreement between the two curves.
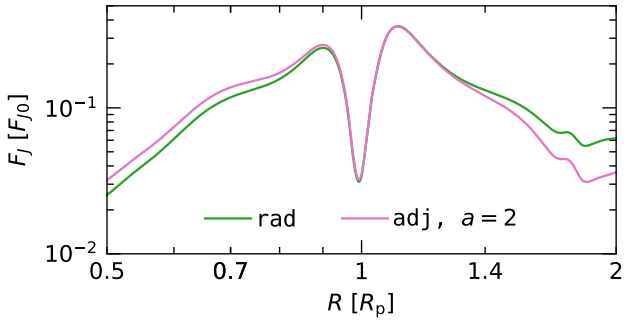

Figure 13. Radial profiles of F_J for models rad (green) and adj (pink) for $\tau = 1$, $\beta_{\rm surf} = 1$ and $a = 2$. While the two models have the same G (see Fig. 12), their F_J profiles are somewhat different, mainly far from the planet.

face density and temperature structure of the shock front in a model with $\beta_{\rm surf} = 3$, $\tau = 1$ for these two implementations of in-plane cooling using an azimuthal slice at $R = 1.2\,R_{\rm p}$. Visible is a stark contrast in temperature between models rad (green) and adj (pink), with the former showing a heat excess in the pre-shock region (right half) while the latter shows a clear, sharp temperature spike along the shock front.

3.5 Section summary

In summary, we confirm the findings of MR20b in that in-plane cooling strongly affects the spiral wave AMF, and by extension the propensity for planet-induced gap opening. We further show that fully radiative models (rad) are distinctly different from surf models, as simply accounting for the ratio $\beta_{\rm surf}/\beta_{\rm tot}$ locally does not yield a good match between the two models, with an offset of ≈ 1.5–2 along the β axis depending on τ. Nevertheless, the existence of a minimum in G naturally splits the β space into a fast- and slow-cooling regime, where omitting the effects of in-plane cooling under- and overestimates the planet's capacity to open multiple gaps, respectively. We investigate this further in Sect. 4.

We also find that a local, β prescription with $\beta = \beta_{\rm surf}$, similar to that by Miranda & Rafikov (2020a) and Zhang & Zhu (2020), results in an effective cooling timescale that is larger by a factor of 4 than in our surf models due to their use of β cooling with a term
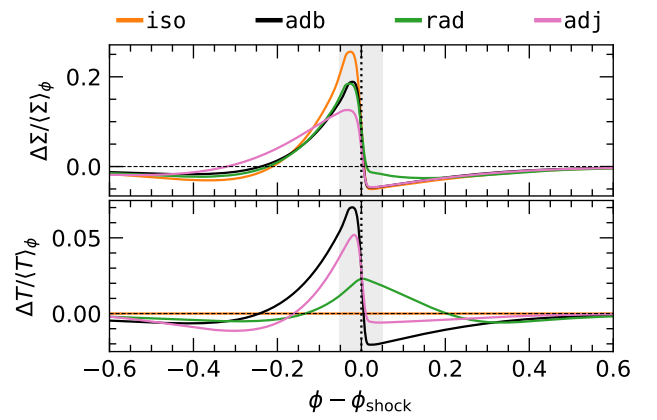

Figure 14. A slice along the spiral shock front at $R = 1.2\,R_{\rm p}$, showing its azimuthal structure. While models rad and its approximate equivalent adj (green and pink curves, respectively) are functionally identical in terms of our AMF metric G (see Fig. 12), the shock structure is completely different. Thanks to in-plane radiative diffusion, model rad shows a weaker temperature spike and thus behaves more similar to the locally isothermal model (iso, orange curves). See also Fig. 2.

$Q_{\rm cool} \propto T^4$. While this does not take away from their findings, as they simulate a generic system to showcase the effect of in-plane cooling, this difference becomes important for ALMA disks where the cooling timescale in the observable range is commonly $\beta_{\rm surf} \sim 0.1$–10.

To reconcile this, we design a recipe for β cooling (Sect. 3.4) that accounts for all theoretically motivated differences between β models and our model rad which treats radiation transport directly. While reproducing the AMF behavior of the density waves reasonably well, however, this adjusted β approach misses the diffusion of heat around spiral shocks, something that affects their density and temperature structure. In the next section, we will demonstrate how this adjusted β recipe performs at predicting gap profiles carved by planets in realistic disks with radially-varying quantities.

4 APPLICATION TO REAL SYSTEMS

In the previous section we explored the differences in AMF between models that omit or treat in-plane radiation transport, and used the parameter G as a proxy for estimating the planet's gap-opening capabilities. We found two regimes of $\beta_{\rm surf} \gtrsim 1$ and $\beta_{\rm surf} \lesssim 1$, where the omission of in-plane cooling under- and overestimates spiral radiative damping respectively. Since stronger radiative damping translates to a single, deep gap rather than several, shallow gaps (MR20b), accounting for in-plane cooling will result in fewer gaps for $\beta_{\rm surf} \gtrsim 1$ and vice versa for $\beta_{\rm surf} \lesssim 1$.

Since the latter scenario is more common at radii where planets are predicted to exist in ALMA-observed systems ($R \sim 50\,{\rm au}$), accounting for in-plane radiative diffusion should typically result in more rings and gaps than when neglecting it. This is consistent with the findings of MR20b, supports the hypothesis that a planet can be responsible for multiple gaps (e.g., Zhang et al. 2018), and explains the weaker substructure in Ziampras et al. (2020b), where the effect of in-plane cooling was missed. The opposite is true, however, in the slower-cooling, more optically thick inner disk ($\beta_{\rm surf} \gtrsim 10$, $\tau \gtrsim 10$). Here, in-plane radiative diffusion serves to reduce the AMF
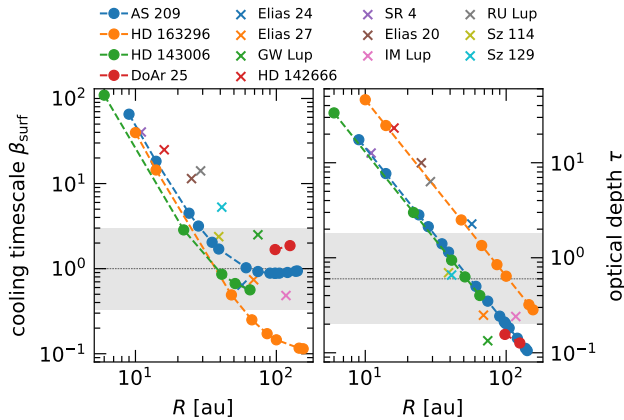
Figure 15. Surface cooling timescale β_{surf} and optical depth τ for a variety of DSHARP systems (Andrews et al. 2018), using estimates of gas and dust properties discussed in Sect. 4.1. Shaded bands mark areas where the inclusion of in-plane cooling can strongly affect gas thermodynamics and therefore planet-driven gap opening.

of planetary spirals close to the planet, reducing their ability to open multiple gaps.

In this section, we test these assumptions by running long-term models of systems with parameters inspired by ALMA observations.

4.1 Estimates of β_{surf} and τ in ALMA systems

As a proof of concept, we compute τ and β_{surf} at the position of annular features (rings, gaps) for several systems present in the DSHARP survey (Andrews et al. 2018). We perform this by combining the list of features in Table 1 of Huang et al. (2018) with estimates of the surface density Σ at each position using Table 3 in Zhang et al. (2018) for all applicable systems. Stellar properties $(M_\star, L_\star)$ are given by Andrews et al. (2018). We then assume for a background state a passively heated disk model where $Q_{irr} = -Q_{cool}$, or

$$T = T_0 R^{-1/2}, \qquad T_0 = \left(\frac{\theta L_\star}{8\pi\sigma_{SB}}\right)^{1/4}, \qquad (19)$$

and compute the Rosseland and Planck mean opacities using the joint power-law formula of Lin & Papaloizou (1985) (for details see Table 1 in Müller & Kley 2012), which for our models practically reduces to $\kappa_R = \kappa_P = 2 \times 10^{-4} \, (T/\text{K})^2 \, \text{cm}^2$ per gram of gas. We finally calculate the optical depth τ and the cooling timescale β_{surf} using Eqs. (4) & (9).

Our results are shown in Fig. 15, where we mark interesting regions—where treating in-plane cooling is expected to produce significant differences to surface-cooling-only models—with shaded bands. Notably, we find a large cluster of features for $\beta_{surf} \sim 1$ and $\tau \sim 0.6$, the latter being consistent with estimates of the optical depth by Dullemond et al. (2018). We note, however, that given our liberal estimates of Σ and our relatively simple opacity model, our estimated values of β_{surf} and τ might not be very accurate.

Based on these estimates, we choose the systems AS 209 and Elias 20 to represent the marginally optically thin ($\beta_{surf} \sim 1$, $\tau \lesssim 1$) and optically thick regimes ($\beta_{surf} \sim 10$, $\tau \sim 10$). For these two systems the inclusion of in-plane cooling should result in an enhancement/suppression of gap opening compared to models with surface cooling only, respectively. We describe our setup for these simulations in the next section.

Table 2. Parameters used in our models of AS 209 and Elias 20. Values with a subscript 'p' refer to the planet's position.

Quantity	AS 209	Elias 20
$M_\star$ [$M_\odot$]	0.83	0.48
$L_\star$ [$L_\odot$]	1.41	2.24
R_p [au]	99	29
M_p [M_J]	0.144	0.091
Σ_p [g/cm^2]	10	80
T_p [K]	14.2	29.7
h_p	0.082	0.084

4.2 Model setup

Our setup is very similar to that described in Sect. 3.1 with minor changes. Namely, we base our initial conditions on our estimates in Sect. 4.1 and choose initial surface density and temperature profiles $\Sigma_0(R) \propto R^{-1}$ and $T(R) \propto R^{-1/2}$, as well as reference aspect ratios $h_p \approx 0.08$ for the two systems, the latter given by a balance between Q_{irr} and Q_{cool}. The Rosseland and Planck mean opacities now follow the recipe of Lin & Papaloizou (1985) as well. Our initial conditions are consolidated in Table 2 and plotted on the left panels of Fig. 16.

Similar to Sect. 3.1, we use grids logarithmically spaced in $R \in [0.1, 4] R_p$ and with a resolution of 16 cells per scale height at R_p, for a grid size of $N_R \times N_\phi = 721 \times 1225$ for AS 209 and 701×1189 for Elias 20. This resolution strikes a good balance between appropriately resolving gas dynamics, maintaining feasible computation runtimes, and comparing with the literature.

The planets have a mass of $M_p = 0.3 \, M_{th}$ to be consistent with our analysis in the previous section. This translates to $M_p = 0.144 \, M_J$ and $0.09 \, M_J$ for AS 209 and Elias 20, respectively, which are reasonable values compared to the fits of Zhang et al. (2018) if we extrapolate to the lower viscosity of $\alpha = 10^{-5}$ that we use. The planets are then embedded at $R_p = 99$ au and 29 au for AS 209 and Elias 20, respectively. We note that our goal here is not to reproduce these two systems optimally, but instead to showcase the importance of in-plane cooling in gap opening in the optically thin and thick regimes, and therefore the exact values of M_p and α are not important.

Both systems are then evolved for 500 planetary orbits using the following four cooling models: locally isothermal ("`iso`"), surface cooling ("`surf`"), surface and in-plane cooling ("`rad`"), and adjusted β cooling ("`adj`"). The first three are described in Table 1, and the latter is introduced in Eq. (18) and discussed in Sect. 3.4.

4.3 Results

In this section we present our findings for the models of AS 209 and Elias 20 that were described above. We first discuss differences in the gas profiles, and continue with simplistic synthetic observations of continuum emission.

4.3.1 Gas profiles

The gas surface density structure after 500 orbits is shown in the right panels of Fig. 16. We note the lack of a secondary gap at $R \approx 0.5 R_p$ as well as a deeper primary gap for model `surf` for AS 209 (panel b̲), compared to other models, fully consistent with the findings of Ziampras et al. (2020b). At the same time, the two models that treat in-plane cooling (`rad` and `adj`), while slightly different further into the inner disk, both capture the secondary gap as well as the bump inside the planet's corotating region, leading to the possibility of two additional rings compared to model `surf`.
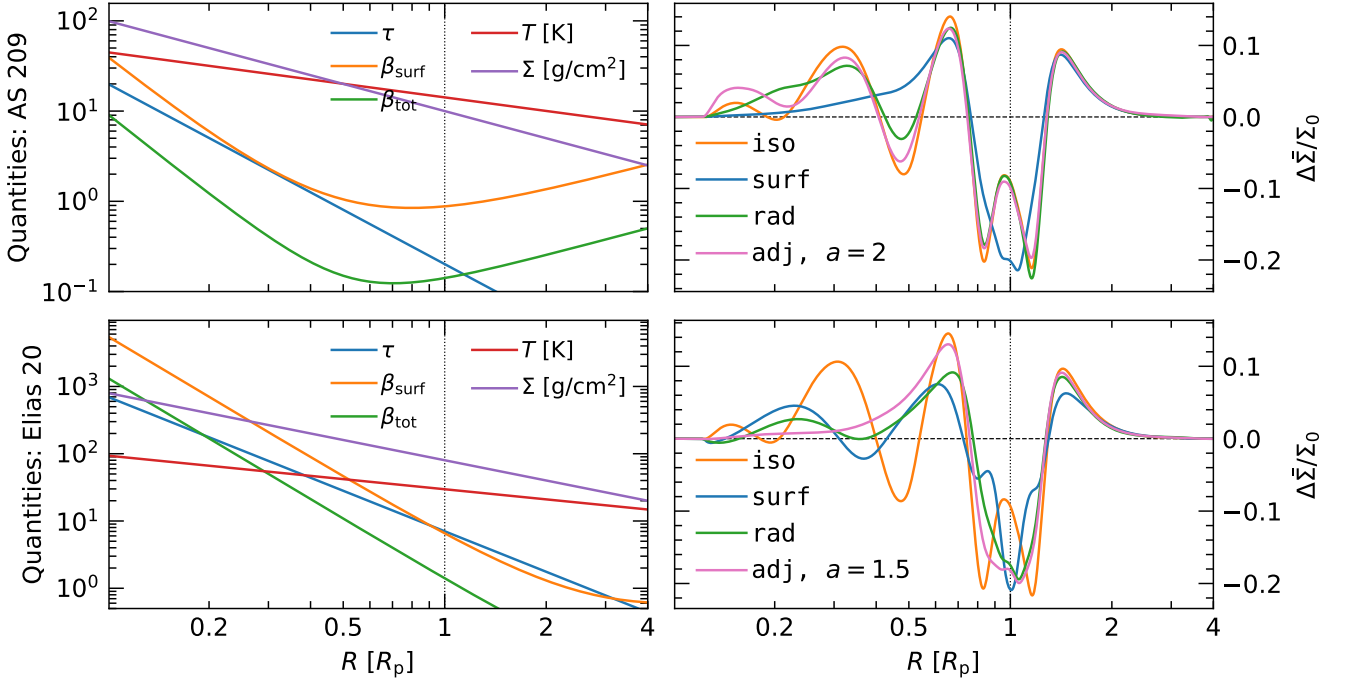
Figure 16. Gas profiles for AS 209 (top) and Elias 20 (bottom). Left panels show various relevant quantities at $t = 0$, while right panels show azimuthally averaged surface density deviations after 500 planetary orbits. Undulations suggest possible rings or gaps, which are explored further in Sect. 4.3.2. We note a moderate damping of such undulations for models with cooling when compared to the locally isothermal one (iso, orange curves on the right), and the different behavior of models with and without in-plane cooling (rad, green and surf, blue, respectively). Treating in-plane cooling with our adjusted β recipe via Eq. (18) (model adj, pink) matches the fully radiative model rad reasonably well for AS 209, but not so much for Elias 20.

In other words, we confirm that in the (marginally) optically thin regime ($\tau \lesssim 1$, $\beta_{surf} \sim 1$) the omission of in-plane cooling incorrectly weakens the planet's capacity to open multiple gaps.

Regarding Elias 20 (panel d̲ of Fig. 16), comparing models surf (blue) and rad (green) indicates that in-plane cooling slightly suppresses the opening of a secondary gap at $R \approx 0.4\,R_p$ and results in weaker radial features inside the planet's corotating region. This is once again in agreement with our predictions in Sect. 3, validating our findings.

However, we find that model adj, which is designed to approximately mimic the effects of in-plane radiative diffusion, results in a rather different radial structure for Elias 20. This is particularly obvious from the lack of a secondary undulation in the inner disk. This is not entirely unexpected: while models adj and rad are functionally identical in terms of the parameter G (a proxy for the AMF), that does not mean their AMF profiles $F_J(R)$ will also be identical (see Fig. 13). In addition, we showed in Sect. 3.4 that there are significant differences in the density and temperature structure around shock fronts between the two models (see Fig. 14). These effects, in conjunction with our temperature-dependent opacity model (see also Appendix C and Dullemond et al. 2022) and the existence of radial gradients in Σ, κ, τ, and β—which did not exist in our simple models in Sect. 3—shows that this adjusted β approach might not always be applicable for realistic disks.

4.3.2 Continuum emission

In the interest of comparing our findings with observations, we follow up with synthetic observations of millimeter emission simi-

lar to MR20b. This is achieved by postprocessing our azimuthally-averaged disk profiles with a simple model of radial dust drift and computing the resulting emission flux at a wavelength of 1 mm. Our method is described in detail in Appendix B.

Our calculated brightness temperature T_b for AS 209 is shown in the top panel of Fig. 17. We find that, as expected, the model with surface cooling only ("surf", blue curve) shows distinctly different results. All other approaches reproduce the ring inside the planet's gap region at $R \approx 0.9\,R_p \approx 90$ au as well as a darker gap at $R \approx 0.4\,R_p \approx 40$ au, consistent with the continuum observation (Huang et al. 2018) and previous locally isothermal models of this system (e.g., Zhang et al. 2018; Ziampras et al. 2020b). To further demonstrate the differences between models, we provide heatmaps of T_b in Fig. 18 (left panel) as well as a comparison to the ALMA observation of this system in Fig. 19.

Regarding Elias 20, our results are shown in the bottom panel of Fig. 17, the right panel of Fig. 18, and in Fig. 20. Here, the surf model shows a clear, deep gap about the planet's orbit (29 au). Models with a treatment of in-plane cooling (rad, adj), however, show much weaker radial structure. To an extent, this is in line with the observation of Elias 20, which only shows two very shallow gaps about 29 au (Huang et al. 2018), but a more accurate model is necessary to correctly fit this system. Such a fit is not within the scope of this work, but will be addressed in followup work.

For both systems, we note the very good agreement between the fully radiative model rad and its approximate approach (adj) in the figures described above.
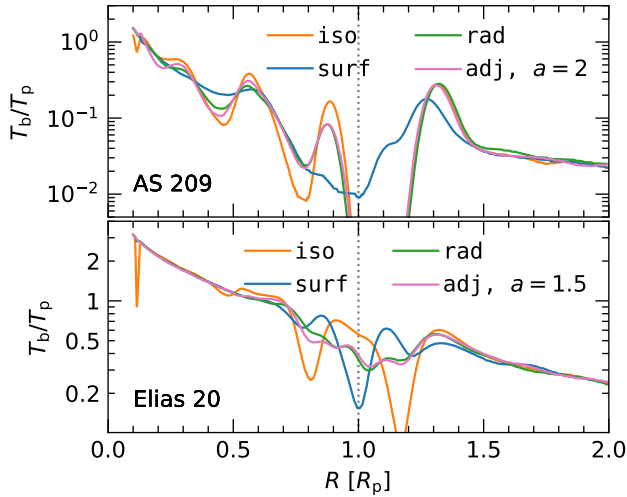
Figure 17. Radial profiles of brightness temperature T_b for mm-sized dust grains, normalized to the temperature at the planet's location T_p. Model `surf` shows distinctly different structure, with too few and too many rings for AS 209 (top) and Elias 20 (bottom), respectively. On the other hand, models `rad` and `adj` agree quite well.
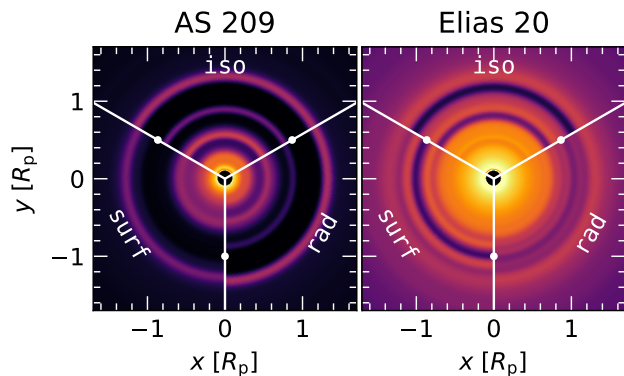

Figure 18. Two-dimensional heatmaps of T_b similar to Fig. 17 to help visualize the differences between models. The planet's radial position is labeled with white dots. The colormap limits match Figs. 19 and 20 for AS 209 and Elias 20, respectively.

4.4 Section summary

In Sect. 3 we identified two regimes of $\beta_{surf} \gtrsim 1$ and $\beta_{surf} \lesssim 1$, and showed that in-plane cooling corrects the AMF of planet-driven spirals by damping or enhancing it compared to models with surface cooling only, depending on the regime. We tested this assumption by selecting two representative systems from the DSHARP survey (Andrews et al. 2018), one for each of the two regimes.

Our findings match our theoretical expectations and previous work very well, in that our `rad` models consistently recover a more realistic image of both systems regardless of cooling regime when compared to `surf` models. In addition, we found that model `adj` recovers a dust structure very similar to that of model `rad` for both systems, but with differences in the inner disk that could become significant if a more realistic treatment of dust is considered.
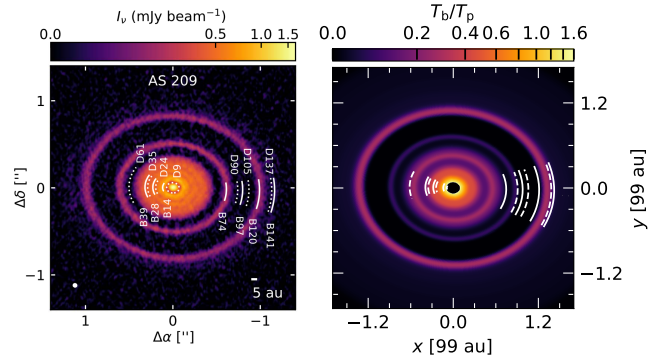
Figure 19. Left: ALMA observation of AS 209 (Huang et al. 2018). Right: Brightness temperature heatmap for AS 209 for our fully radiative model ("`rad`") similar to the left panel of Fig. 18 but projected in a way similar to the observation shown in Huang et al. (2018). Overlaid in white curves are the positions of rings and gaps as reported by that study.
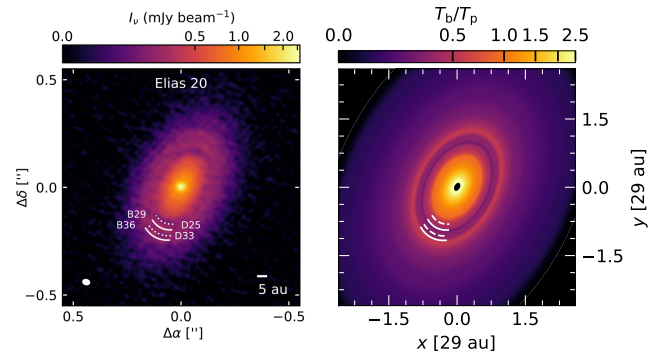

Figure 20. Left: ALMA observation of Elias 20 (Huang et al. 2018). Right: Brightness temperature heatmap for our `rad` model of Elias 20, similar to Fig. 19 and based on the right panel of Fig. 18.

5 DISCUSSION

In this section we touch on our assumptions and their limitations. We also discuss possible applications of our findings.

5.1 On the quality of our models of real sources

In Sect. 4.3 we focused on two ALMA systems, AS 209 and Elias 20, which were part of the DSHARP survey (Andrews et al. 2018). There, we made several assumptions about the properties of the gas and dust in both systems (as well as their ages through our model runtime) and computed millimeter emission maps based on the results of our hydrodynamics models in Sect. 4.3.2.

It is clear that, looking at Figs. 19 & 20, our models do not reproduce the two systems very well: radial features are sometimes observed at different radii, while a number of rings and gaps are missing altogether. We stress that an accurate model of either system is not the goal of this work, but instead the idea that cooling—and in particular in-plane radiation transport—can strongly influence gas hydrodynamics and affect a planet's gap opening capacity. After all, we treat dust transport with a post-processed, radial drift approach, without accounting for the real-time effects of dust concentration and grain size distribution on gas opacity and dynamics.

Nevertheless, our results show that the inclusion of in-plane cool-

ing is a step in the right direction. Accounting for in-plane cooling recovers an additional ring for AS 209 compared to a model that does not, and forms several, shallow gaps about the planet's orbit instead of a single deep gap at $R_{\rm p}$ for Elias 20.

5.2 Using our findings to constrain observables

Zhang et al. (2018) showed that at least five of the observed rings in AS 209 could be explained by a single planet at 99 au. While their approach certainly involved some fine-tuning in their parameters in order to match the observed emission ($h_{\rm p} = 0.05$ and $\alpha = 3 \times 10^{-4}$ at 99 au), their findings clearly showed that the "single planet, multiple rings" scenario is certainly feasible and a realistic fit to observations is possible with the right choice of physical parameters.

This parameter space, however, is not only difficult to explore with high-resolution simulations, but also degenerate to an extent: a disk with low enough $h_{\rm p}$–α or a massive enough planet can overcome the gap opening threshold (Crida et al. 2006) and lead to one or more gaps, depending on the combination of these three parameters. Unfortunately, however, even with constraints on $h_{\rm p}$ via the assumption of a passively irradiated disk, the turbulent parameter α and the planet's mass $M_{\rm p}$ are very difficult to constrain from observations (though some limits on α can be derived, e.g., Dullemond et al. 2018).

Given that in-plane radiation transport can act as a means of diffusion in that it smears temperature peaks around spiral shocks, its inclusion in a hydrodynamical model could help relax the need to explore very high values of α. This helps constrain the parameter space and shift the focus to surveying different planet masses.

Of course, a cooling prescription raises the question of how one should treat other quantities that become relevant with it, namely the dust opacity and the related properties of dust grains. Progress is constantly being made on this front, with access to up-to-date opacity models for ALMA systems (Birnstiel et al. 2018) and tools for opacity-dependent radiative postprocessing of dust and synthetic imaging such as `RADMC-3D` (Dullemond et al. 2012).

Nevertheless, regardless of the accuracy of the physical model when accounting for cooling, the reality is that the locally isothermal approach will consistently overestimate the planet's ability to open gaps in the inner disk when compared with a model with any form of cooling, as noted by Miranda & Rafikov (2019). This can be translated to the planetary mass estimates of Zhang et al. (2018) simply being a lower limit, and can be leveraged to further constrain the parameter space.

5.3 On our approximate recipe for FLD

In Sect. 3.4 we defined the parameter G as a proxy for the planet's capacity to open multiple gaps. This was done since we were not interested at the absolute number of rings or gaps that a planet can produce, but rather the relative AMF by planet-driven spiral arms when comparing different equations of state and methods of cooling.

When comparing our models that account for in-plane cooling through FLD (`rad`) to those that do so with a β approach (`adj`), we found excellent agreement in terms of this metric G. However, we also showed that the two methods show somewhat different radial profiles of the angular momentum flux F_J (Fig. 13), especially far from the planet, as well as very different spiral shock structures due to the diffusion of thermal energy around the shock front in `rad` models. This difference became visible in our models of ALMA systems in Sect. 3.2, with the gas structure in the inner disk being dis-

tinctly different between the two approaches (see Fig. 16). Nevertheless, the differences in gas profiles were ultimately not as noticeable in our dust emission profiles (Fig. 17).

All in all, while this approximate β recipe in Eq. (18) is considerably cheaper compared to a model with full radiation transport from a computational point of view (1.4× for Elias 20, 2.1× for AS 209), we highlight that caution should be exercised when using it.

5.4 3D effects

In this work we approached planet–disk interaction in a 2D, vertically integrated framework. In principle, however, the processes relevant to our study are 3D. Compressed gas can expand in the vertical direction, resulting in an additional degree of freedom around shock fronts and therefore weaker spiral shocks (Lyra et al. 2016). Meridional flows can also slow down or even stall planet-induced gap opening by refilling the gap region (Morbidelli et al. 2014; Lega et al. 2021). Finally, the gas can cool and diffuse in the vertical direction, something we incorporate in a simplistic way with $\tau_{\rm eff}$ in Eq. (4) but that is missed in Eq. (6). The dependence of β on height will result in spiral arms cooling at different rates at different altitudes, possibly affecting their AMF. A full 3D investigation with self-consistent global cooling (e.g., an FLD approach) is necessary to compare and contrast our findings with more realistic models. This will be the focus of future work.

6 CONCLUSIONS

Our aim in this work was twofold: to demonstrate the effects of in-plane radiation transport in planet-driven gap opening, and to compare our findings to those of MR20b, who treated this process using a specially designed local β cooling prescription. We approached this topic with nearly-inviscid numerical hydrodynamics simulations of planet–disk interaction.

We first conducted a high-resolution parameter study where we measured the planet-driven spiral angular momentum flux (AMF) as a function of the surface cooling timescale $\beta_{\rm surf}$ and for various treatments of cooling. Namely, we considered the locally isothermal ("`iso`") and adiabatic ("`adb`") equations of state, as well as models with surface cooling ("`surf`") and both surface cooling and in-plane radiation transport via flux-limited diffusion (FLD, "`rad`").

This comparison showed that spiral wave AMF, which can be used as a proxy for the capability of a planet to open one or more gaps, damps fastest when the cooling timescale approaches unity regardless of cooling method. For `surf` models this translates to $\beta_{\rm surf} \approx 1$, while for `rad` models the condition is instead $\beta_{\rm tot} \approx 1.5$–2 (see Eq. (13)). The former result is consistent with the literature (Miranda & Rafikov 2020a; Zhang & Zhu 2020), while the effect of in-plane radiation transport by means of radiative diffusion is original to this work and shows that a local cooling prescription cannot accurately reproduce models with a full treatment of FLD. We also identified the existence of two regimes, for $\beta_{\rm surf} \lesssim 1$ and $\beta_{\rm surf} \gtrsim 1$, where the omission of in-plane cooling results in an under- or overestimation of spiral AMF, respectively.

We then repeated the radiative models described above (`surf` and `rad`) using a local β relaxation approach instead of allowing the gas to freely heat up and cool down through stellar irradiation and a cooling prescription, respectively. After making necessary adjustments, we showed in Sect. 3.4 that our approximation (Eqs. (12)–(14) for in-plane cooling reproduces our fully radiative models (`rad`) very accurately, making it a suitable approach when studying a planet's

spiral AMF. However, we showed that this β approach cannot capture the diffusion of heat around shock fronts, which significantly affects the temperature and density structure of spiral arms and therefore limits the applications of this method.

We followed up with long-term, semi-realistic models of two sources from the DSHARP study (Andrews et al. 2018; Huang et al. 2018), namely AS 209 and Elias 20. The two were chosen to probe the two ends of the β spectrum (short and long cooling timescales, respectively), in order to verify our findings from the previous numerical experiment (Sect. 3). For these models we estimated gas properties using stellar and disk properties from the literature (Andrews et al. 2018; Zhang et al. 2018) and a temperature-dependent opacity model (Lin & Papaloizou 1985).

We found that our models agree very well with our predictions from our parameter study in Sect. 3. Namely, the inclusion of in-plane radiation transport results in brighter ring features compared to `surf` models for AS 209, and weaker gap opening for Elias 20. We then showcased our findings with simplistic synthetic observations of millimeter dust emission for the two systems, showing the presence of an additional (albeit faint) ring at the planet's orbit for AS 209 and much shallower gaps for Elias 20 in `rad` models.

Our findings highlight the importance of proper treatment of radiation transport in protoplanetary disk modeling, while also showing that the "single planet, multiple gaps" scenario demonstrated by Zhang et al. (2018) is certainly possible with realistic gas thermodynamics. Such models can be used to constrain the masses of planets possibly present in the gaps seen in millimeter emission, as well as the turbulence and dust properties in protoplanetary disks (Ziampras et al., in prep.).

ACKNOWLEDGEMENTS

DATA AVAILABILITY

Data from our numerical models are available upon reasonable request to the corresponding author.

REFERENCES

ALMA Partnership et al., 2015, ApJ, 808, L3
Andrews S. M., et al., 2018, ApJ, 869, L41
Béthune W., Latter H., Kley W., 2021, A&A, 650, A49
Birnstiel T., et al., 2018, The Astrophysical Journal, 869, L45
Chiang E. I., Goldreich P., 1997, ApJ, 490, 368

Crida A., Morbidelli A., Masset F., 2006, Icarus, 181, 587
Dullemond C. P., Juhasz A., Pohl A., Sereshti F., Shetty R., Peters T., Commercon B., Flock M., 2012, RADMC-3D: A multi-purpose radiative transfer tool, Astrophysics Source Code Library, record ascl:1202.015 (ascl:1202.015)
Dullemond C. P., et al., 2018, ApJ, 869, L46
Dullemond C. P., Ziampras A., Ostertag D., Dominik C., 2022, A&A
Flock M., Nelson R. P., Turner N. J., Bertrang G. H. M., Carrasco-González C., Henning T., Lyra W., Teague R., 2017, ApJ, 850, 131
Garrido-Deutelmoser J., Petrovich C., Charalambous C., Guzmán V. V., Zhang K., 2023, ApJ, 945, L37
Goldreich P., Tremaine S., 1980, ApJ, 241, 425
Goodman J., Rafikov R. R., 2001, ApJ, 552, 793
Guzmán V. V., et al., 2018, ApJ, 869, L48
Haffert S. Y., Bohn A. J., de Boer J., Snellen I. A. G., Brinchmann J., Girard J. H., Keller C. U., Bacon R., 2019, Nature Astronomy, 3, 749
Huang J., et al., 2018, ApJ, 869, L42
Hubeny I., 1990, ApJ, 351, 632
Hunter J. D., 2007, Computing In Science & Engineering, 9, 90
Keppler M., et al., 2018, A&A, 617, A44
Kley W., 1989, A&A, 208, 98
Koller J., Li H., Lin D. N. C., 2003, ApJ, 596, L91
Lega E., et al., 2021, A&A, 646, A166
Levermore C. D., Pomraning G. C., 1981, ApJ, 248, 321
Lin D. N. C., Papaloizou J., 1985, in Black D. C., Matthews M. S., eds, Protostars and Planets II. pp 981–1072
Lyra W., Richert A. J. W., Boley A., Turner N., Mac Low M.-M., Okuzumi S., Flock M., 2016, ApJ, 817, 102
Masset F., 2000, A&AS, 141, 165
Menou K., Goodman J., 2004, ApJ, 606, 520
Mignone A., Bodo G., Massaglia S., Matsakos T., Tesileanu O., Zanni C., Ferrari A., 2007, The Astrophysical Journal Supplement Series, 170, 228
Mignone A., Flock M., Stute M., Kolb S. M., Muscianisi G., 2012, A&A, 545, A152
Miranda R., Rafikov R. R., 2019, ApJ, 878, L9
Miranda R., Rafikov R. R., 2020a, ApJ, 892, 65
Miranda R., Rafikov R. R., 2020b, ApJ, 904, 121
Morbidelli A., Szulágyi J., Crida A., Lega E., Bitsch B., Tanigawa T., Kanagawa K., 2014, Icarus, 232, 266
Müller T. W. A., Kley W., 2012, A&A, 539, A18
Müller T. W. A., Kley W., Meru F., 2012, A&A, 541, A123
Ogilvie G. I., Lubow S. H., 2002, MNRAS, 330, 950
Pfeil T., Klahr H., 2021, ApJ, 915, 130
Pierens A., Nelson R., 2022, arXiv e-prints, p. arXiv:2211.03816
Rafikov R. R., 2002a, ApJ, 569, 997
Rafikov R. R., 2002b, ApJ, 572, 566
Rafikov R. R., 2016, ApJ, 831, 122
Rafikov R. R., Cimerman N. P., 2023, MNRAS, 519, 5800
Rodenkirch P. J., Rometsch T., Dullemond C. P., Weber P., Kley W., 2021, A&A, 647, A174
Shakura N. I., Sunyaev R. A., 1973, A&A, 500, 33
Sudarshan P., Penzlin A. B. T., Ziampras A., Kley W., Nelson R. P., 2022, A&A, 664, A157
Takeuchi T., Lin D. N. C., 2002, ApJ, 581, 1344
Toro E. F., Spruce M., Speares W., 1994, Shock waves, 4, 25
Van Leer B., 1974, Journal of computational physics, 14, 361
Zhang S., Zhu Z., 2020, MNRAS, 493, 2287
Zhang S., et al., 2018, ApJ, 869, L47
Zhang K., et al., 2021, ApJS, 257, 5
Zhu Z., Dong R., Stone J. M., Rafikov R. R., 2015, ApJ, 813, 88
Ziampras A., Ataiee S., Kley W., Dullemond C. P., Baruteau C., 2020a, A&A, 633, A29
Ziampras A., Kley W., Dullemond C. P., 2020b, A&A, 637, A50
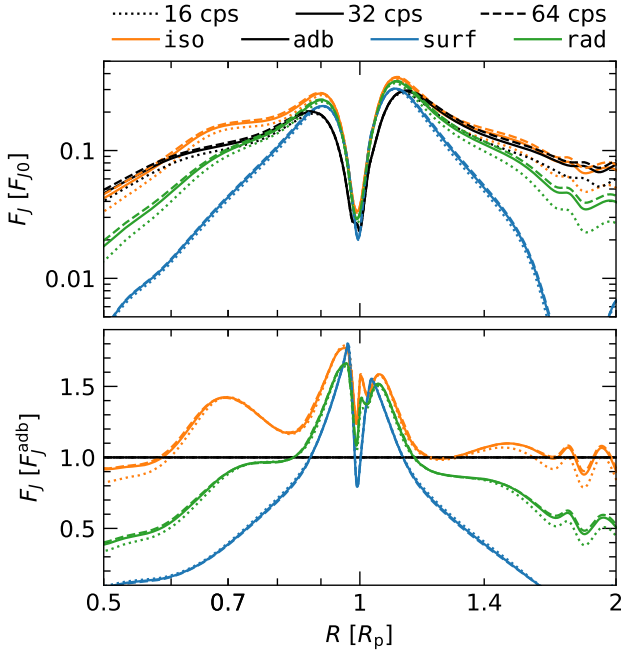de Val-Borro M., et al., 2006, MNRAS, 370, 529

Figure A1. Spiral AMF as a function of resolution in units of cells per scale height at R_p (cps, various line styles) and equation of state (various colors). Top: F_J normalized to the reference F_{J0} in Eq. (16). Bottom: F_J normalized to the curves for adiabatic models ("adb") at a given resolution.

APPENDIX A: SPIRAL AMF — RESOLUTION STUDY

To make sure we resolve the planet-driven spirals accurately enough to calculate their AMF in Sect. 3, we carry out a resolution study where we vary the number of cells per scale height (cps) at R_p from 8 to 64 for the fiducial model with $\tau = 10$, $\beta_{\text{surf}} = 1$ and for various cooling prescriptions. We then compute F_J for each model, and present our results in Fig. A1.

We find that a resolution of 32 cps is sufficient for most models, with model `surf` converging even at 16 cps (top panel of Fig. A1). There are minor differences between 32 and 64 cps, but the overall behavior is well-resolved. Furthermore, normalizing all curves to F_J^{adb} instead makes any differences between 32 and 64 cps indistinguishable (bottom panel), suggesting that 32 cps is enough to carry out our study in Sect. 3 while claiming numerical convergence in terms of the parameter G. Finally, we note that results at 8 cps were quite poor, and advise against using such low resolutions for similar studies.

APPENDIX B: INTEGRATION OF DUST PARTICLES

To produce the continuum emission maps in Sect. 4.3.2 we closely followed the approach of MR20b. We output azimuthally averaged radial disk profiles (denoted as $\langle x \rangle$) at a frequency of 10 snapshots per planetary orbit, and then interpolated linearly between these snapshots to generate time-dependent profiles for all quantities. The dust is initialized as 2×10^5 particles of size $a^d = 1$ mm and bulk density of $\rho^d = 1$ g/cm^3, distributed radially as $\Sigma_0^d(R) = 0.01 \Sigma_0(R)$. The Stokes number is then

$$\text{St}(R) = \frac{\pi}{2} \frac{\rho_d a_d}{\Sigma(R)}. \tag{B1}$$

We then compute the radial velocity of each dust particle using the dust drift formula by Takeuchi & Lin (2002)

$$\bar{u}_R^d = \frac{1}{1 + \text{St}^2} \left(\bar{u}_R + \frac{\text{St}}{\langle \Sigma \rangle \Omega_K} \frac{d\langle P \rangle}{dR} \right), \tag{B2}$$

where $\bar{u}_R(R) = \langle \Sigma u_R \rangle / \langle \Sigma \rangle$ is the density-weighted, azimuthally averaged gas radial velocity profile. The dust particles are then integrated using an RK2 scheme with a timestep of 0.1 orbits, up to a time of 500 planetary orbits. Particles that drift past the inner radial boundary of our domain ($R \leq 0.1R_p$) are dropped.

After that, we compute the dust surface density $\Sigma^d(R)$ by binning the dust particles in as many bins as there are radial cells in the domain and convolving with a Gaussian curve with a FWHM of $0.075 R$, similar to MR20b. Finally, we assume a dust opacity of $\kappa^d = 2.2$ cm^2 per gram of dust and compute the brightness temperature

$$T_b(R) = \left(1 - e^{-\tau^d(R)} \right) T(R), \tag{B3}$$

where $\tau^d = \frac{1}{2} \kappa^d \Sigma^d$ and $T(R)$ is the gas temperature. In the optically thin limit, T_b reflects the surface density distribution of dust grains.

APPENDIX C: ON THE DIFFERENCES BETWEEN EXPLICIT SOURCE TERMS AND β COOLING

In Sect. 3.3 we argued that a β cooling approach with β_{surf} in the form given by Eq. (9) consistently overestimates the true cooling timescale (obtained by explicitly considering Q_{irr} and Q_{cool}) by a factor of 4 for small deviations about the equilibrium temperature. Here, we derive this factor by linearizing the energy equation (Eq. (1c)) and using different source terms, similar to Dullemond et al. (2022).

We start by computing an equilibrium state with temperature T_0 given by $Q_{\text{irr}}^0 = Q_{\text{cool}}^0$. We then assume a perturbed temperature $T = T_0 + \delta T$, with a small deviation $\delta T \ll T_0$. When β cooling using Eq. (9) we have:

$$\frac{d\Sigma e}{dt} = Q_{\text{relax}} = -\frac{T - T_0}{T} Q_{\text{cool}}, \qquad Q_{\text{cool}} = 2\sigma_{\text{SB}} \frac{T^4}{\tau_R^{\text{eff}}}, \tag{C1}$$

see Eqs. (4)) & (8). Assuming, for simplicity, that τ does not depend on temperature, we can now Taylor-expand Q_{relax} to first order with respect to T and obtain:

$$Q_{\text{relax}} = -2\sigma_{\text{SB}} \frac{(T_0 + \delta T)^3}{\tau_R^{\text{eff}}} \delta T \approx -Q_{\text{cool}}^0 \frac{\delta T}{T_0}. \tag{C2}$$

We can repeat this exercise for the case where cooling is determined by

$$\frac{d\Sigma e}{dt} = Q_{\text{irr}} - Q_{\text{cool}} = \Delta Q. \tag{C3}$$

Using Eq. (5), and for a given radial position R and flaring angle θ, we find that Q_{irr} is independent of temperature. As a result, we can write

$$\Delta Q = Q_{\text{irr}}^0 - \frac{2\sigma_{\text{SB}}}{\tau_R^{\text{eff}}} (T_0 + \delta T)^4$$
$$= Q_{\text{irr}}^0 - Q_{\text{cool}}^0 - Q_{\text{cool}}^0 \frac{4\delta T}{T_0} = 4Q_{\text{relax}}. \tag{C4}$$

In other words, cooling with a β approach with β_{surf} in the form of Eq. (9) will result in small perturbations damping with an effective cooling timescale that is roughly 4 times longer compared to treating cooling with the radiative source terms explicitly. This is verified in Fig. 10, where curves corresponding to β models are offset along the β axis, with the offset translating to an effective cooling timescale of a factor 4 times longer for β models.
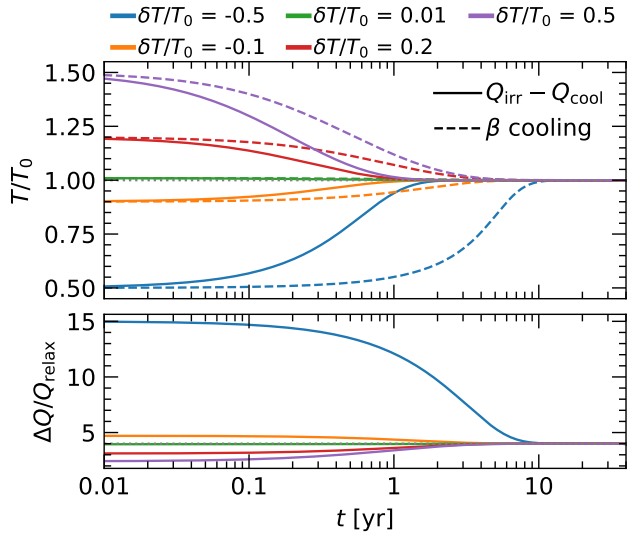
Figure C1. Time evolution of the temperature T for a fixed point in the disk where T evolves via a balance of heating versus cooling (solid curves, ΔQ, Eq. (C3)) or via β relaxation (dashed lines, Q_{relax}, Eq. (C1)) to the equilibrium solution T_0. The top panel shows that the β approach consistently overestimates the true cooling timescale. The bottom panel shows that the ratio $\Delta Q/Q_{\text{relax}}$ converges to a value of 4 for small deviations from T_0 (see Eq. (C4)) but can be substantially different for larger deviations.

We also note that this factor of 4 can vary significantly as perturbations grow in strength. To showcase this, we integrate Eqs. (C1) & (C3) in time with a simple python script for various values of $\delta T/T_0$ and consolidate our findings in Fig. C1. As shown in this figure, this approach is no longer valid for $|\delta T|/T_0 \gtrsim 0.1$, making the β approach unsuitable for modeling massive planets in disks subject to shock heating (Rafikov 2016; Ziampras et al. 2020a). In particular, a β cooling approach results in a completely different relaxation rate for $\delta T/T_0 \lesssim -0.5$.

Finally, it should be noted that for a temperature-dependent opacity this approximation becomes significantly more complicated. Dullemond et al. (2022) showed that this factor of 4 is more generally $4+b$ in the optically thin limit (and therefore $4-b$ in the optically thick limit), where $b = \frac{d\log\kappa}{d\log T}$. In the interest of only highlighting the shortcomings of β cooling, we chose not to repeat models that explore different values of b.

APPENDIX D: NONLINEAR REGIME

In the majority of our runs we use a planet mass $M_{\text{p}} = 0.3\,M_{\text{th}}$ such that our models probe the (quasi-)linear regime of planet–disk interaction. For more massive planets this interaction becomes nonlinear, with the planet launching multiple strong spiral shocks that result in a more complex disk structure due to the opening of multiple deep gaps and stronger shock heating (Rafikov 2016; Ziampras et al. 2020a). The latter in particular can lead to a temperature increase of up to a factor of 2 from the equilibrium solution for $M_{\text{p}} \gtrsim M_{\text{th}}$ in optically thick disks (Ziampras et al. 2020a).

This situation creates two problems in the scope of our study. For one, the excitation of multiple spiral shocks causes the AMF $F_J(R)$ to show multiple peaks in the inner disk, making the parameter G less effective at characterizing a planet's capacity to open multiple
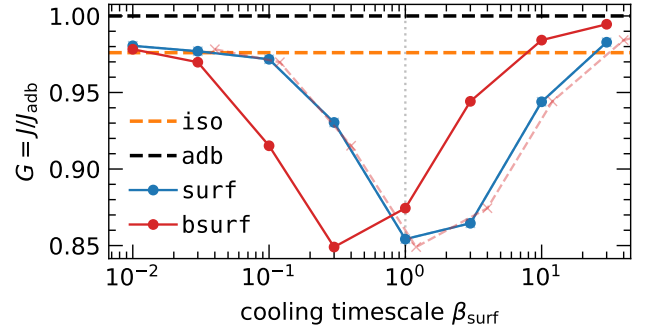
gaps. At the same time, large deviations from the equilibrium temperature T_0 will change the correction factor of 4 derived for the linear regime in Appendix C in a non-trivial way, which could render β models less suitable for the treatment of cooling. To investigate the above, we carry out a set of models with $\tau = 10$, $h_{\text{p}} = 0.05$ and a higher mass planet with $M_{\text{p}} = 3\,M_{\text{th}}$ and compute once again $G(\beta_{\text{surf}})$.

Our results are shown in Fig. D1. We find that shifting bsurf models to the right by a correction factor of 4 yields a very good match with the surf models, albeit the optimal factor would be $\approx$ 3.5. This is in line with our estimate that this correction factor should decrease for larger $\delta T/T_0$ (stronger shock heating) in Fig. C1. Given that a planet mass of $3\,M_{\text{th}}$ is already quite high, we conclude that a correction factor of 4 will be sufficient in most cases.

At the same time, however, the data becomes quite noisy for this planet mass. The planet strongly influences its surrounding disk, with the surface density changing locally by a factor of up to 2–3 depending on equation of state. This is in stark contrast to our models with $M_{\text{p}} = 0.3\,M_{\text{th}}$, where Σ barely changes by 20% in most cases (see Fig. 3), and complicates the interpretation of G as a measure of the planet's capacity to open multiple gaps.

We also find that, here, $G^{\text{adb}} \approx G^{\text{iso}}$. This is not unexpected, as for such high M_{p} the density waves are highly nonlinear and dissipate through shocks immediately around the planet, before the linear propagation of the spiral can change their AMF.

Figure D1. The parameter G similar to Fig. 7, but for models with $M_{\text{p}} = 3\,M_{\text{th}}$. We find a very good match between models bsurf and surf after shifting by the correction factor of 4 discussed in Appendix C, even in this nonlinear regime.

APPENDIX E: EFFECT OF THE ASPECT RATIO ON SPIRAL AMF

In all models shown in the main text we used a reference aspect ratio $h_{\text{p}} = 0.05$. It is nevertheless worth investigating whether the critical cooling timescale β_{crit} for which G is smallest might depend on h. To answer this question, we carry out a set of models with $\tau = 10$ and $h_{\text{p}} = 0.1$ (with $M_{\text{p}} = 0.3\,M_{\text{th}} = 10^{-4}\,M_{\star}$) and plot the computed $G(\beta_{\text{surf}})$ in Fig. E1. As shown there, while the absolute value of G is typically higher compared to models with $h_{\text{p}} = 0.05$, G once again shows a minimum at $\beta_{\text{surf}} = 1$ and 10 for models surf and rad, respectively, consistent with our findings for $h_{\text{p}} = 0.05$. The values of Σ and κ used in these models are listed in Table G.
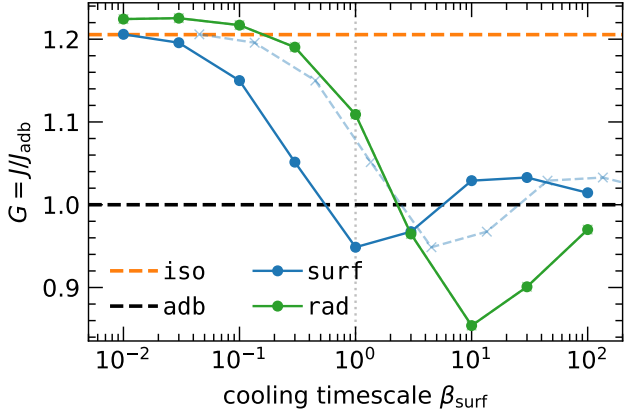
Figure E1. The parameter G similar to Fig. 7, but for models with $h_p = 0.1$. The dashed blue line results from shifting the `surf` model by $1 + f$ (see Eq. (14)).

APPENDIX F: COMPARISON WITH RECIPE OF MR20B

We also discuss the performance of the effective β-cooling prescription proposed in MR20b (see Eqs. (10) & (11)). In Figs. F1–F3 we compare the performance of just the β_{mid} part of their prescription (marked "MR20b mid") against models with only Q_{rad} (i.e., no surface cooling or irradiation heating are included), marked "FLD only". This enables a direct comparison of the two ways to treat the in-plane cooling. Different plots make different assumptions about τ and h_p, as indicated in captions. One can see that the simple prescription of MR20b does not do very well at reproducing the direct radiation diffusion calculations: it is often offset horizontally and features a deeper minimum than the latter.

In Fig. F4 we also show the surface density profiles obtained for AS 209 and Elias 20 using different approximations of surface and in-plane cooling. Here, models marked "MR20b+" use β cooling as in Eq. (13) where β_{surf} is corrected by the factor 4 discussed in Sec. 3.3 and β_{mid} is given by the recipe in MR20b through Eqs. (10) & (11). One can see that the latter approach gives results different from both the adjusted (`adj`) and most realistic models (`rad`) when it comes to details around rings/gaps.

The likely reason for the poor match between the MR20b and the `rad` models is that the former effectively assumed the characteristic radial scale of the temperature perturbation in the planet-driven density wave to be the same as the radial scale of its Σ perturbation, which becomes considerably smaller than H as the wave propagates. However, as Fig. 14 clearly illustrates, thermal diffusion ahead of the shock makes the thermal length scales larger than the length scale of Σ variation, which acts to slow down the loss of the wave AMF compared to MR20b. In fact, one can argue that in real disks the temperature length scale in the disk plane should always be around H: once the photons diffuse by $\sim H$ horizontally, it becomes easier (more likely) for them to then diffuse vertically by $\sim H$ than to continue horizontally. This may justify the better performance of the simple prescription (Eq. (12)) in reproducing `rad` models upon the simple adjustment described in Sect. 3.4.
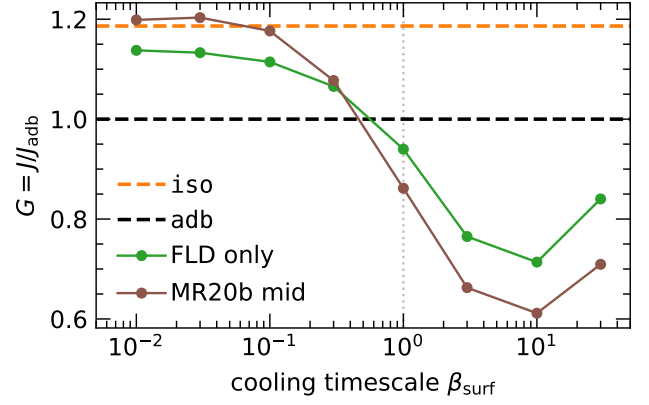

Figure F1. The parameter $G(\beta_{surf})$ for $\tau = 10$ similar to Fig. 7, comparing models where only in-plane cooling is considered. Brown: β models with the in-plane cooling recipe given by Eqs. (10) & (11), following MR20b. Green: models with full in-plane radiative diffusion (Q_{rad}, Eq. (6)).
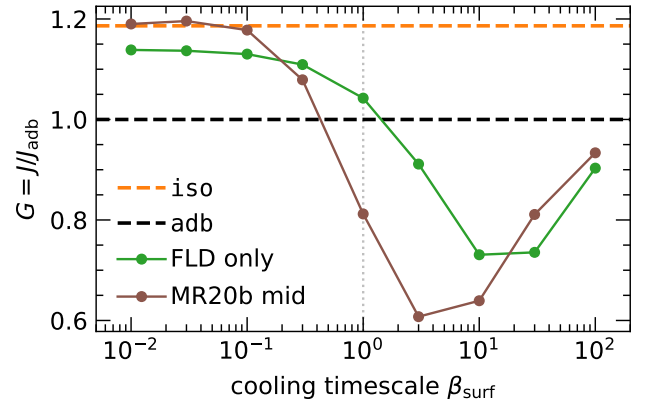

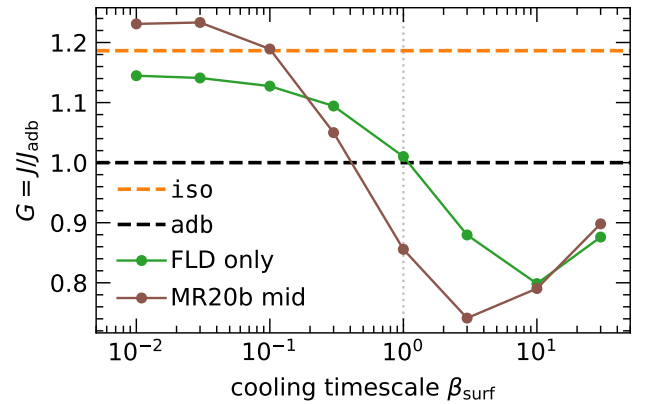**Figure F2.** Similar to Fig. F1 for $\tau = 1$.
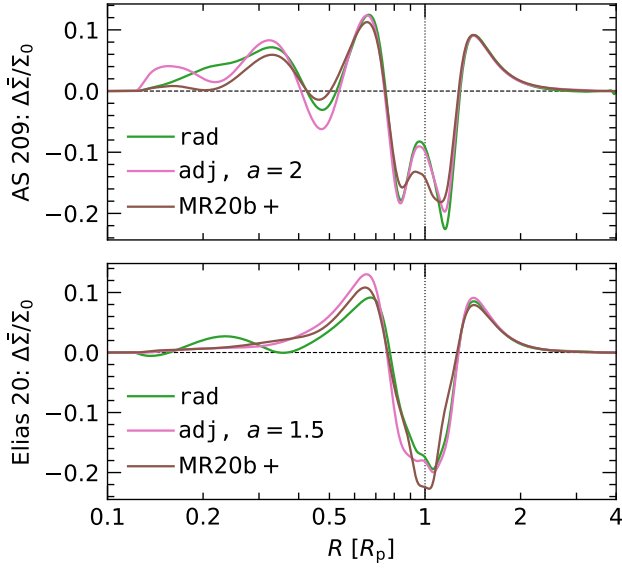

Figure F3. Similar to Fig. F1 for $h_p = 0.1$.

Figure F4. Azimuthally averaged gas density after 500 planetary orbits for AS 209 and Elias 20, similar to Fig. 16.

APPENDIX G: VALUES OF Σ AND κ IN OUR AMF STUDY

Here we provide a summary of the values of Σ and κ used to achieve different combinations of β_{surf} and τ in Sect. 3. We note that some values or combinations of Σ and κ are unphysical, but this does not matter in the scope of our study as the intent is simply to achieve a given combination of β_{surf} and τ. Our parameters are listed in Table G.

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

Table G1. Values for surface density Σ and opacity κ used in our models in Sect. 3, and the corresponding surface cooling timescale β_{surf} and optical depth τ. Alongside an aspect ratio profile $h = 0.05\,(R/R_{\text{p}})^{1/4}$, or equivalently a temperature profile $T \propto R^{-1/2}$, our choices ensure constant β_{surf} and τ throughout the disk.

quantity	value								
τ	10								
β_{surf}	0.01	0.03	0.1	0.3	1	3	10	30	100
Σ [g cm^{-2}]	3.8146	11.552	37.49	113.54	377.09	1141.9	3792.7	11367	37892
κ [cm^2 g^{-1}]	5.2699	1.7103	0.5362	0.17503	0.05362	0.017402	0.0053311	0.0017594	5.278×10^{-4}
quantity	value								
τ	1								
β_{surf}	0.01	0.03	0.1	0.3	1	3	10	30	100
Σ [g cm^{-2}]	15.151	45.883	148.91	450.96	1506.4	4588.3	14891.0	45211	150710
κ [cm^2 g^{-1}]	0.13192	0.044067	0.0135	0.0043813	0.0013345	4.356×10^{-4}	1.335×10^{-4}	4.4236×10^{-5}	1.33×10^{-5}

Table G2. Values for surface density Σ and opacity κ used in our models in Appendix E, where $h_{\text{p}} = 0.1$, similar to Table G.

quantity	value								
τ	10								
β_{surf}	0.01	0.03	0.1	0.3	1	3	10	30	100
Σ [g cm^{-2}]	242.5	727.51	2425.1	7275.1	24250	72751	2.425×10^5	7.275×10^5	2.43×10^6
κ [cm^2 g^{-1}]	0.082473	0.027491	0.0082472	0.0027491	8.247×10^{-4}	2.749×10^{-4}	8.247×10^{-5}	2.749×10^{-5}	8.247×10^{-6}