

Ray Casting and Flux Limited Diffusion

Åke Nordlund¹

¹Centre for Star and Planet Formation, and
Niels Bohr Institute, University of Copenhagen
Blegdamsvej 17, 2100 Copenhagen, Denmark
email: aake@nbi.dk

Abstract. Solving radiative transfer problems with ray casting methods is compared with the commonly used ‘Flux Limited Diffusion’ approximation. Whereas ray casting produces solutions that converge to the exact one as the number of rays is increased, flux-limited-diffusion is fundamentally a ‘look-alike’ method, which produces solutions that are reminiscent of the correct solution but which cannot be made to converge to it.

1. Introduction

Solving the ‘radiative transfer problem’ in general means computing specific radiation intensities $I_\nu(\mathbf{r}, \boldsymbol{\Omega})$ at relevant frequencies ν in a given physical situation. With a sufficient set of frequencies and space angle directions $\boldsymbol{\Omega}$, and given also the absorption and scattering cross sections κ_ν and σ_ν , one can then compute the exchange of energy and momentum between the radiation field and the medium the radiation is propagating through.

A number of sub-classes of the ‘radiative transfer (RT) problem’ are of relevance in different astrophysical situations: In some cases the time-of-flight of the radiation is important, in others not. In some cases the momentum and energy transfer are both important, while in other cases only the energy transfer is of significance. In some cases the populations of atomic states, which again determine the absorption cross sections, are strongly influenced by the radiation field, and computing the population numbers becomes in practice an integral part of the ‘radiative transfer problem’. Scattering leads to a similar situation, where the ‘source function’ in the RT problem contains a term directly proportional to the radiation intensity, averaged over space angle with some (isotropic or non-isotropic) angular re-distribution function.

Common to all these cases is the fact that the specific intensity $I_\nu(\mathbf{r}, \boldsymbol{\Omega})$ depends on both frequency ν , direction $\boldsymbol{\Omega}$, and position \mathbf{r} ; i.e., the radiation field has six degrees of freedom. This is often, but as we shall see incorrectly, considered an insurmountable problem with respect to obtaining direct numerical solutions of the ‘radiative transfer problem’. It is argued that it would be too expensive to obtain sufficiently accurate solutions in a 6-dimensional space, and that therefore one needs to use approximate methods that reduce the dimensionality.

Flux limited diffusion (FLD; Minerbo 1978; Levermore & Pomraning 1981) has been, and still is, a popular method for obtaining approximate solutions of radiative transfer problems (Turner & Stone 2001; Krumholz *et al.* 2007). The flux limited diffusion approximation reduces the dimensionality from six to four, by assuming that the radiation field can be reasonably represented by the mean radiation intensity,

$$J_\nu(\mathbf{r}) = \frac{1}{4\pi} \int_{\Omega} I_\nu(\mathbf{r}, \boldsymbol{\Omega}) d\boldsymbol{\Omega}. \quad (1.1)$$

Approximate expressions, which fulfill certain physically motivated asymptotic conditions, are then used to compute estimates of the radiative energy flux,

$$\mathbf{F}_\nu(\mathbf{r}) = \frac{1}{4\pi} \int_{\Omega} I_\nu(\mathbf{r}, \boldsymbol{\Omega}) \boldsymbol{\Omega} d\Omega, \quad (1.2)$$

based upon the mean intensity (as well as gradients thereof) and the local opacity. Note that the flux-limited diffusion approximation is only concerned with approximating the space angle dependence of the radiation field; the frequency dependence of the radiation field is a different affair.

A reduction of the dimensionality from six to four would appear to be a significant advantage and could be seen as such a dramatic gain that it would motivate the use of an approximation, if the results obtained were still reasonable. However, a proper cost comparison must also factor in the relative complexity of the different methods; i.e., what is, in the end, the cost per $(\mathbf{r}, \nu, \boldsymbol{\Omega})$ -point for a direct solution, relative to the cost per (\mathbf{r}, ν) -point for the approximate solution?

Because the FLD leads to an elliptic equation for the mean intensity, while the direct solution only involves a small number of floating point operations per degree of freedom, the cost balance actually comes out in favor of the direct method, even for rather large numbers of rays. To put this differently: The key circumstance that renders direct solutions of the 6-dimensional RT problem not only tractable but even advantageous is that, even though the computation effort scales with six degrees of freedom,

$$t_{\text{update}} = c_{\text{update}} N_{\mathbf{r}} N_\nu N_\Omega, \quad (1.3)$$

the scaling constant c_{update} is very small, of the order of just a few nanoseconds per update per point with current CPU-cores. Since the time required to update the physical state of an MHD-model is typically several microseconds per point on the same CPU-cores, one can afford of the order of a thousand frequency points times space-angle directions, without increasing the computing time per time step with more than a factor of two.

Therefore, in many common astrophysical situations it is indeed possible, and actually advantageous, to solve radiative transfer directly, using ray tracing.

The main advantage with ray-tracing methods is that solutions converge to the correct ones as the number of rays are increased, and that they therefore lend themselves to traditional types of convergence studies and validation. However, as discussed below, ray tracing methods in general also have several other advantages relative to flux-limit-diffusion: They are simpler to implement, often use less computing time, and they can be made to have near-perfect parallelization properties.

2. Radiative Transfer: Formal Solutions

The central task that needs to be performed in all ray-tracing based methods is computing the radiation intensity $I_\nu(\mathbf{r}, \boldsymbol{\Omega})$, given a source function $S_\nu(\mathbf{r}, \boldsymbol{\Omega})$, from the radiative transfer equation

$$\frac{dI_\nu(\mathbf{r}, \boldsymbol{\Omega})}{d\tau} = S_\nu(\mathbf{r}, \boldsymbol{\Omega}) - I_\nu(\mathbf{r}, \boldsymbol{\Omega}), \quad (2.1)$$

where $d\tau$ is the optical depth increment in the direction $\boldsymbol{\Omega}$, defined as

$$d\tau = (\kappa_\nu + \sigma_\nu) ds. \quad (2.2)$$

This is often referred to as the ‘formal problem’, and the task is thus to compute the solution to the formal problem as rapidly and accurately as possible.

Note that solving the 'formal problem' is part of (while still totally independent of) any requirement of consistency between $I_\nu(\mathbf{r}, \boldsymbol{\Omega})$ and $S_\nu(\mathbf{r}, \boldsymbol{\Omega})$, e.g. due to scattering or atomic level population balance. A number of well-known and thoroughly tested iterative techniques exist for solving for source-intensity consistency (e.g. Hubeny 2003), but these need not be discussed here, since they all have similar requirements with respect to solving the formal problem—several solutions of the formal problem are typically needed to evaluate residual imbalance between the radiation field and the source function. Thus, for the purpose of this discussion we can focus on considering the cost of obtaining the radiation field from a given source function.

As a test case where the results of using direct ray casting on the one hand and the flux-limited-diffusion approximation on the other can be easily compared, I consider below the computation of the rate of energy transfer between the radiation and the gas, in a case where velocities are small with respect to the speed of light, so the radiation field can be considered to be determined instantaneously from the source function.

2.1. Ray Tracing / Ray Casting

Direct solutions of the radiative transfer problem generally rely on solving the radiative transfer equation (2.1) along a number of rays through the volume under consideration. In general one distinguishes between *short characteristics* methods and *long characteristics* methods.

Short characteristics solutions are obtained incrementally, going typically from one grid point to the next cell boundary, where the solution then needs to be interpolated to the nearest grid point in order to continue.

Long characteristics methods rely on solutions obtained along rays traversing the entire volume. For each direction $\boldsymbol{\Omega}$ parallel inclined rays are passed through all grid points in a central plane. One thus obtains solutions near all grid point in one sweep. Rays generally then do not pass through grid points, but with a density of rays similar to the density of grid points across rays, values at grid points can be found by interpolation between rays. With this method all information that is obtained along each single ray is being utilized, and the total cost thus scales as $N_{\text{update}} \sim N_r N_\Omega N_\nu$.

The denomination 'long characteristics' is sometimes used to refer to a theoretical case where each point in three dimensions is connected by a specific ray to every other point. Such a method can of course never be used in practice, since it would scale as the square of the number of points; $N_{\text{update}} \sim N_r^2 N_\nu$. The method corresponds to setting $N_\Omega = N_r$, and would thus correspond to a huge over-resolution in space angle. Not even discrete sources needs to be handled in this way; these can instead be handled by adaptive schemes (Wang *et al.* 2004; Razoumov & Cardall 2005).

A common misconception is that direct, ray tracing solutions are only suitable for discrete sources. In fact, it is easy to use a combination of a set of rays emanating from point sources with sets of parallel rays that take care of diffuse sources. For diffuse sources, the number of space-angles needed is generally quite small, as illustrated by the test below. However, since the requirements depend on the detailed properties (e.g. smoothness) of the source field, no general rule can be given.

In analogy with the concept of 'multi-casting' on the internet (where many clients can receive information transmitted by a single source), one could argue for referring to the efficient long characteristics method as 'ray multi-casting', or just 'ray casting'.

2.2. Flux Limited Diffusion

In the flux-limited-diffusion approximation one finds the mean radiation intensity as follows: One assumes that the radiative flux can be estimated by an expression of the

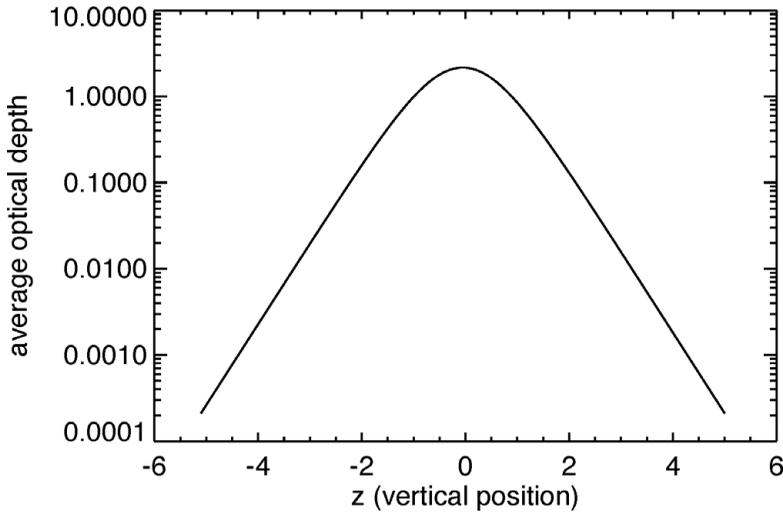


Figure 1. Horizontally averaged optical depth in the test case.

form

$$\mathbf{F}_{\text{rad},\nu} = -\frac{\lambda}{\kappa_\nu} \nabla J, \quad (2.3)$$

where $\lambda = \lambda(|\nabla J|, J, \kappa_\nu)$ is a ‘flux limiter’ function (e.g. Levermore & Pomraning 1981). One then requires that the divergence of the radiative flux equals the rate of energy transfer between the radiation field and the gas, so

$$\nabla \cdot \mathbf{F}_{\text{rad},\nu} = \rho \kappa_\nu (S - J), \quad (2.4)$$

where S and J are, respectively, the source function and the mean radiation intensity (Turner & Stone 2001).

The result of combining Eqs. 2.3 and 2.4 is a non-linear elliptic equation in J . In the test cases reported on below this problem was solved using successive over-relaxation, starting from the ‘exact’ solution for J provided by the ray casting method. To monitor convergence the residual between the left and right hand sides of Eq. 2.4 was used.

Tests with three different flux limiters showed that the Levermore & Pomraning (1981) limiter produced slightly better results than the Minerbo (1978) and Krumholz *et al.* (2007) limiters in the current case.

It should be noted that, even though the radiative flux is fully specified by Eq. 2.4, the actual boundary value of the mean intensity needs to be determined independently. To ensure that this choice did not disfavor the FLD solution the mean intensity of the most accurate (reference) ray casting case was adopted as a boundary condition on J .

3. Results

The test case represents a ‘fragmented disk’ — a proxy of a turbulent, low temperature stellar accretion disk. The opacity is assumed to be due to pure absorption, with a constant absorption κ_ν per unit mass. The opacity $\rho \kappa_\nu$ thus varies in proportion to the mass density. The source function is chosen to vary in unison with the density, mimicking temperature fluctuations at near-constant entropy. The horizontal variation of the mass density is chosen to be a pattern generated by smoothing an image of white noise with a

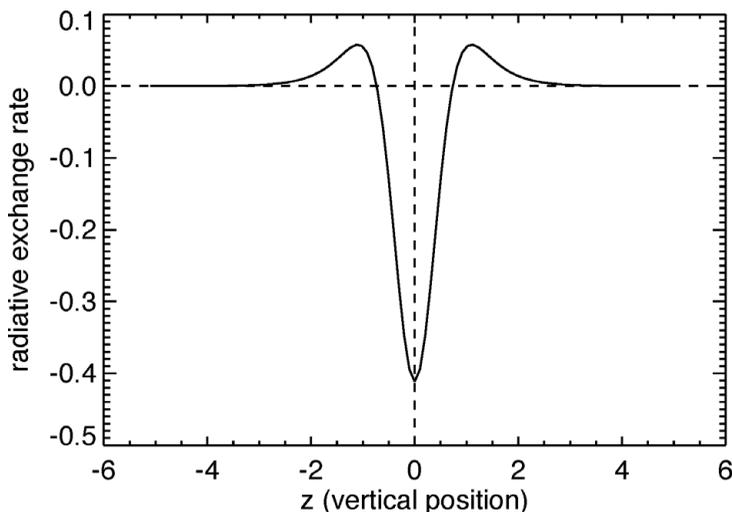


Figure 2. Average heating / cooling rate per unit volume for the reference solution, as a function of the vertical coordinate.

Gaussian filter, leaving approximately 16 degrees of freedom in each horizontal direction (cf. Fig 5a).

The opacity scaling has been chosen so the maximum optical depth across the disk (from bottom to top) is about 5, while in the least dense spots on the disk the total optical depth is about 0.5. Above and below the mid plane the mass density (and hence the optical depth) decreases exponentially. The optical depth is of the order of 10^{-4} at the upper / lower boundaries of the domain. Figure 1 shows the horizontally averaged optical depth in the disk, as a function of the vertical position. The horizontal variation is illustrated in Fig. 5a.

Figure 2 shows the average heating / cooling rate for the reference solution as a function of the vertical coordinate. Generally, there is cooling (energy loss) in a region near the disk mid plane, because the finite optical depth allows energy to leak out. However, sideways heating can occur in spots, as radiation ‘leaks’ from hotter regions into cooler ones. There is a general re-heating above the optical surface, as the mean intensity is larger than the local source function. Because of the exponential drop in density away from the mid plane the heating / cooling rate drops to very small values there.

Figure 3 shows the horizontal root-mean-square variation of the heating / cooling rate as a function of the vertical coordinate. The root mean square variation is largest in the mid plane, then drops, but is again fairly large in the re-heating layer just outside of optical depth unity.

The FLD solutions are initialized with the ‘exact’ mean intensity from the reference solution, and the elliptic problem is then iterated to near consistency (note that, since the iterations drive the solution *away* from the correct solution any residual iteration error generally tends to *reduce* the difference with respect to the ray tracing reference solution).

Figure 4 shows minimum and maximum errors of the FLD heating / cooling rate, relative to the global root-mean-square variation of the reference heating / cooling rate. Note that the largest error of the FLD solution is more than twice as large as the overall root-mean-square variation of the heating / cooling rate.

The horizontal variation of the error is illustrated in Fig. 5, where the right hand side panel shows an image of the variation of the error in the horizontal plane.

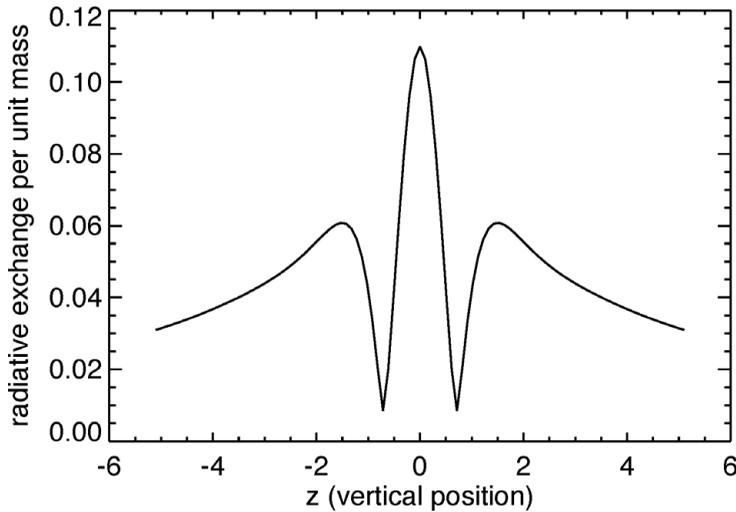


Figure 3. Horizontal root-mean-square variation of the heating / cooling rate, as a function of vertical position

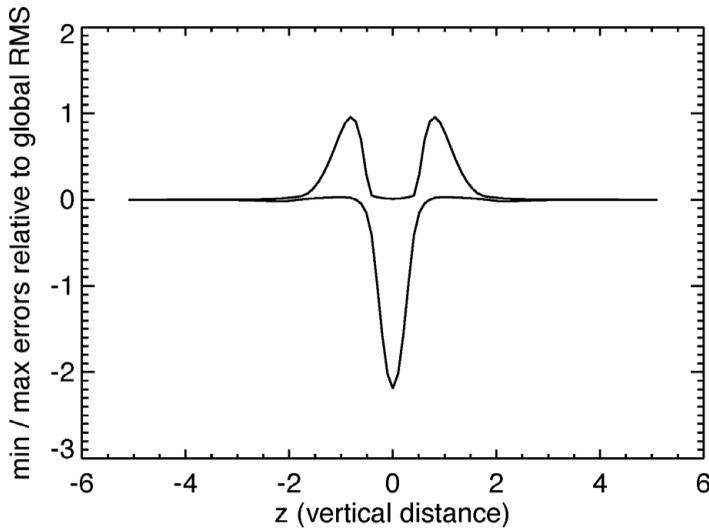


Figure 4. Minimum and maximum errors of the FLD heating / cooling rate, relative to the global root-mean-square variation of the reference heating / cooling rate

Ray tracing solutions were obtained by using Radau integration over inclination angle (Radau integration differs from Gaussian integration in that the vertical direction is always included). For inclined rays integration in the azimuth direction is performed with evenly distributed ray directions.

Figure 6 shows a comparison of the horizontal root-mean-square errors of the heating / cooling rates for the FLD solution and ray tracing solutions with varying number of ray directions (note that there are two rays for each ray direction). As a reference case we use a ray-casting solution with 96 ray directions; this is sufficient for a very accurate solution, given the source function and opacity profiles of this test case. As illustrated by Fig. 6, even the ray tracing solution with only four ray direction is superior to the

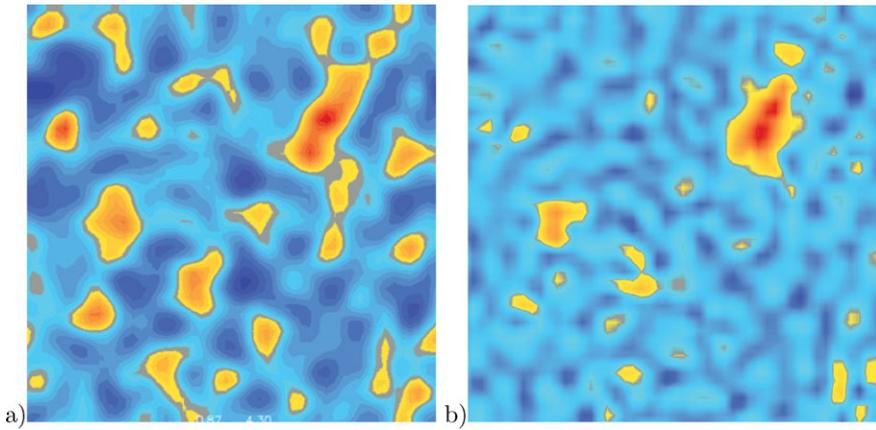


Figure 5. Images of the horizontal variation of the optical depth in the mid plane (left) and of the error in the heating / cooling rate (right). Positive and negative values are shown in distinct colors / shades of gray.

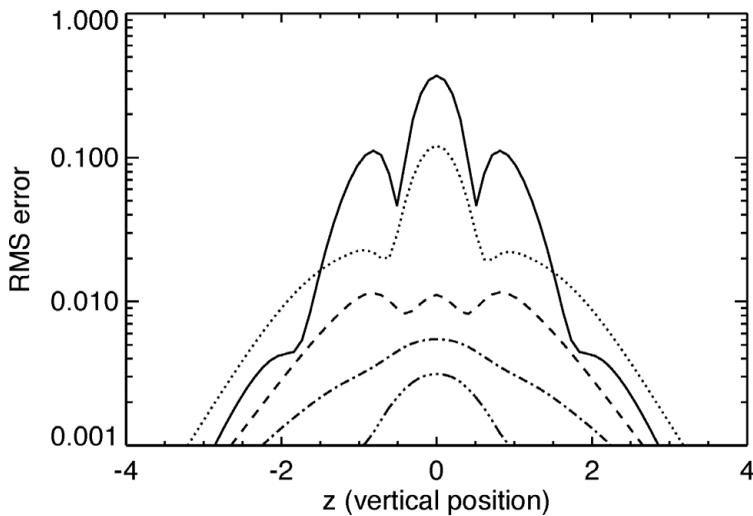


Figure 6. Horizontal root-mean-square errors of the heating / cooling rate, relative to the global root-mean-square variation of the reference solution. The full drawn curve is for the FLD solution. The other curves are for ray tracing with 4 (dotted), 7 (dashed), 16 (dashed-dotted), and 26 (dash-dot-dot-dot) ray directions.

FLD solution at the levels where the heat exchange is strongest, and a solution with 7 ray directions is good to within 1% of the reference solution.

4. Conclusions and concluding remarks

Ray-tracing radiative transfer methods have a challenging ('6-dimensional') scaling of the required work. However, because the scaling constant is very small (about 1-5 nano-seconds per mesh point) a fair number of angle-frequency pairs is still affordable. In practice, the solution of the radiative transfer equation is so fast that similar amounts of time is needed for table lookup of opacities and source functions, and for interpolations of variables to and from the inclined rays.

Ray-tracing (or 'ray-casting') solutions have the big advantage over approximate methods such as FLD that they actually converge towards the exact solution. FLD methods, on the other hand, are 'look-alike methods'; the solutions behave more or less as the correct solutions, but there is no way to force the solution to converge to the correct one—there is no 'cost parameter' with which one can buy higher accuracy at the expense of longer computing time. FLD solutions converge to the diffusion approximation at large optical depth, and behave more or less reasonably at small optical depths, but as illustrated by the test here the errors made can be quite large in a neighborhood of optical depth unity. These layers are typically of central importance when calculating the thermal structure of stars and accretion disks.

Whether one uses ray-tracing or flux-limited-diffusion it is often important to include several frequency points (or frequency bins)—representing for example heating in one frequency domain and cooling in another (Nordlund 1982). With ray-tracing, there is enough capacity to employ several frequencies (or frequency bins), while still being able to afford a reasonable space-angle resolution.

An important requirement for ray tracing to be affordable is that the solution along all rays should be fully utilized; the solution along each ray should contribute to the knowledge of the radiation field in all cells that it passes through. This is trivial in Cartesian geometries, where rays can be chosen to be parallel. However, in more complicated geometries, such as for example disks with near-cylindrical symmetry and huge ratios of outer to inner radius (Pascucci *et al.* 2004), choosing an optimal set of rays and organizing interpolations of results back and forth can be non-trivial.

Ray tracing methods with long characteristics have the great advantage to be easy to parallelize. One can parallelize over ray positions, ray directions, frequencies and frequency bins, but also along each ray (Heinemann *et al.* 2006). The latter is particularly useful when using MPI-domain decomposition, since it minimizes the amount of information that must be passed from cell to cell, and parallelizes without problems to thousands of cores. In this method local solutions of the radiative transfer equation, which can be obtained in 'embarrassingly parallel mode', can be patched together by sending boundary data up-stream and down-stream in the ray direction (Heinemann *et al.* 2006).

The work of ÅN was supported by the Danish Natural Research Council and by the Danish Research Foundation, through its establishment of the Centre for Star and Planet Formation. Computing resources were provided by the Danish Center for Scientific Computing.

References

- Heinemann T., Dobler W., Nordlund Å., & Brandenburg A. 2006, *AA* 448, 731
Hubeny, I. 2003, ASP Conference Proceedings, Vol. 288, 17
Krumholz M. R., Klein R. I., McKee C. F., & Bolstad J. 2007, *ApJ* 667, 626
Levermore C. D. & Pomraning G. C. 1981, *ApJ* 248, 321
Minerbo G. N. 1978, *JQSRT* 31, 149
Nordlund, A. 1982, *AA* 107, 1
Pascucci I., Wolf S., Steinacker J., Dullemond C. P., Henning T., Niccolini G., Woitke P., & Lopez B. 2004, *AA* 417, 793
Razoumov A. O. & Cardall C. Y. 2005, *MNRAS* 362, 1413
Turner N. J. & Stone J. M. 2001, *ApJS* 135, 95
Wang P., Abel T., & Zhang W. 2008, *ApJS* 176, 467