Streaming Instability Code Comparison Problem Set

Stanley A. Baronett[1,2,4]★ and Wladimir Lyra[2,4]
[1]University of Nevada, Las Vegas, Department of Physics and Astronomy, Box 454002, 4505 S. Maryland Pkwy., Las Vegas, NV 89154, USA
[2]Nevada Center for Astrophysics, University of Nevada, Las Vegas, 4505 S. Maryland Pkwy., Las Vegas, NV 89154-4002, USA
[3]New Mexico State University, Department of Astronomy, PO Box 30001 MSC 4500, Las Cruces, NM 88001, USA
[4]Planet Formation in the Southwest + (PFITS+)

Created 1 June 2022. Revised 31 October 2024

1 INTRODUCTION

The streaming instability is a promising mechanism to drive planetesimal formation. Since its discovery (Youdin & Goodman 2005), several hydrodynamics codes have explored the parameters, non-linear properties, and implications of this aerodynamic instability that requires feedback between dust and gas momenta. However, the non-trivial differences between numerical techniques (e.g., finite difference or finite volume) and dust modeling (e.g., as a pressureless fluid or as Lagrangian particles) can make it difficult to disentangle unique scientific results from the potential idiosyncrasies of a particular code or implementation. In an effort to address these issues, this collaborative project aims to comprehensively compare various multipurpose codes across some of the key models and problems previously studied in investigations into the streaming instability.

1.1 Repositories

1.1.1 Figure scripts and source codes

Figure scripts and source codes related to this project can be found in the pfitsplus/sicc GitHub repository. Jupyter Notebooks containing Python scripts to generate the manuscript figures can be found in the `/ipynb` directory. Source and input files for some contributing codes can be found in the `/source_files` directory. To be consistent with the structure of this document (Section 1.2), the subdirectories therein are hierarchically organized first by *model*, next by *problem*, next by *variation*, and last by *code*. For more information, please see the repository `README`, and feel free to create an issue for any questions, feedback, or issues encountered.

1.1.2 Output data

The problem data outputted by contributing codes should be uploaded to the designated Google Shared Drive. Anyone with the link can view and comment on the contents, but please contact Stanley A. Baronett (barons2@unlv.nevada.edu) to request access to add files. To be consistent with the structure of this document (Section 1.2), the subdirectories therein are hierarchically organized first by *model*, next by *problem*, next by *variation*, and last by *code*. Regardless of the inherent data format normally generated by a contributing code, all requested output (e.g., arrays) must be stored in or converted to individual compressed NumPy `.npz` files (see the official "Input and output" documentation for details). All quantities, including times and coordinates, should be saved in the units specified by each problem, as detailed in later sections (e.g., Section 2.2.1; see Section 1.2 for document structure). The `.npz` files should be named and structured as follows.

To verify setup consistency between codes and compliance with specified parameters, simulation snapshots for each problem and variation must be accompanied by a `grid.npz` file containing the cell-centered coordinates in separate arrays for each axes using the keyword arguments (`**kwds`) `x`, `y`, and/or `z`. The snapshots themselves should be named as the corresponding simulation time without leading zeros, with the initial snapshot at $t_{\rm sim} = 0$ named `0.npz`. The requested quantities within each snapshot should be stored in individual arrays using the keyword arguments specified by each problem or variation (e.g., `rhop` for the particle density).

Time series data should be saved as `time_series.npz` and contain individual arrays with the keyword arguments `time` (for the corresponding simulation times) and those specified by each problem or variation for the requested quantities (e.g., `maxrhop` for the maximum particle density). The requested cadence (i.e. time increment between outputs $\mathrm{d}t$) for the time series is also specified by each problem or variation.

1.2 Document structure

The subsequent structure of this document is as follows. The sections themselves (e.g., Section 2) correspond to particular *models* with different source terms (e.g., unstratified vs. stratified). Within each section, the first subsection (e.g., Subsection 2.1) explains the setup and relevant quantities for the corresponding model. The second subsection (e.g., Subsection 2.2) identifies the specific *problems* of interest, the relevant *variations* of parameter values, and the corresponding objectives for the code comparison.

2 UNSTRATIFIED

As detailed in Baronett et al. (2024, sec. 2), the unstratified problems are modeled without the vertical component of stellar gravity in the local-shearing-box approximation (Goldreich & Lynden-Bell 1965), where the equations of motion are linearised with Cartesian x, y, and z axes constantly aligned to the radial, azimuthal, and vertical directions, respectively. The Keplerian reference frame is axisymmetric with periodic boundary conditions in all directions. Pseudo-code for this model that implements the gas and dust source terms detailed in Section 2.1 can be found in the GitHub repository in the `/source_files/unstratified` directory (Section 1.1.1).

2.1 Model setup

2.1.1 Gas

From Baronett et al. (2024, sec. 2.1), the continuity and momentum equations for the inviscid gas ($\nu = 0$) are

$$\frac{\partial \rho_\mathrm{g}}{\partial t} + \nabla \cdot (\rho_\mathrm{g} \mathbf{u}) = 0, \tag{1}$$

$$\frac{\partial \rho_\mathrm{g} \mathbf{u}}{\partial t} + \nabla \cdot (\rho_\mathrm{g} \mathbf{u}\mathbf{u} + P\mathbf{I}) = \rho_\mathrm{g} \left[2\Omega_\mathrm{K} u_y \hat{\mathbf{x}} - \frac{1}{2}\Omega_\mathrm{K} u_x \hat{\mathbf{y}} + 2\Omega_\mathrm{K} \Pi c_\mathrm{s} \hat{\mathbf{x}} - \frac{\rho_\mathrm{p}}{\rho_\mathrm{g}} \left(\frac{\mathbf{u} - \mathbf{v}}{t_\mathrm{stop}} \right) \right], \tag{2}$$

respectively. Solving for the gas density ρ_g, the gas velocity $\mathbf{u}$ is measured relative to the background Keplerian shear flow $\mathbf{u}' = -(3/2)\Omega_\mathrm{K} x \hat{\mathbf{y}}$, where Ω_K is the local Keplerian angular frequency. In equation (2), $P = \rho_\mathrm{g} c_\mathrm{s}^2$ for an isothermal equation of state with speed of sound c_s, and $\mathbf{I}$ is the identity matrix. The first two source terms on the right-hand side of equation (2) are a combination of the radial component of the stellar gravity and the Coriolis and the centrifugal forces. The third term is a constant outward force on the gas due to an external radial pressure gradient, determined by the dimensionless parameter (Bai & Stone 2010, eq. 1)

$$\Pi \equiv \frac{\eta v_\mathrm{K}}{c_\mathrm{s}} = \frac{\eta r}{H_\mathrm{g}} = 0.05, \tag{3}$$

where v_K is the local Keplerian velocity, $H_\mathrm{g} = c_\mathrm{s}/\Omega_\mathrm{K}$ is the vertical gas scale height, and

$$\eta \equiv -\frac{1}{2} \frac{1}{\rho_\mathrm{g} \Omega_\mathrm{K}^2 r} \frac{\partial P}{\partial r} = -\frac{1}{2} \left(\frac{H_\mathrm{g}}{r} \right)^2 \frac{\partial \ln P}{\partial \ln r} \sim \left(\frac{c_\mathrm{s}}{v_\mathrm{K}} \right)^2, \tag{4}$$

is the fractional reduction in orbital speed of the gas from Keplerian (when $\eta > 0$) if the dust were not present (Nakagawa et al. 1986, eq. 1.9). The fourth and final term is the frictional drag force from the solid particles back to the gas, where $\mathbf{v}$ is the ensemble-averaged local velocity of the particles (measured relative to the background shear) and t_stop is their stopping time. The factor of the dust-to-gas density ratio $\rho_\mathrm{p}/\rho_\mathrm{g}$ ensures the conservation of the total linear momentum of the gas and dust particles, where ρ_p is the spatially averaged dust density in the gas cell.

The gas density field is initially uniform with $\rho_\mathrm{g}(x, y, z, t = 0) = \rho_{\mathrm{g},0}$. By assuming a total dust-to-gas mass ratio

$$\epsilon \equiv \frac{\langle \rho_\mathrm{p} \rangle}{\rho_{\mathrm{g},0}}, \tag{5}$$

where

$$\langle f \rangle \equiv \frac{1}{L_x L_y L_z} \iiint f \mathrm{d}x\mathrm{d}y\mathrm{d}z \tag{6}$$

is the instantaneous volume average of quantity f over the computational domain of dimensions $L_x \times L_y \times L_z$, the initial components of the gas velocity take the equilibrium solution by Nakagawa et al. (1986):

$$u_{x,0} = -\epsilon v_{x,0}, \tag{7}$$

$$u_{y,0} = -\left[1 + \frac{\epsilon \tau_\mathrm{s}^2}{(1+\epsilon)^2 + \tau_\mathrm{s}^2} \right] \frac{\eta v_\mathrm{K}}{1+\epsilon}, \tag{8}$$

$$u_{z,0} = 0, \tag{9}$$

where

$$\tau_\mathrm{s} \equiv \Omega_\mathrm{K} t_\mathrm{stop}. \tag{10}$$

is the dimensionless stopping time (a.k.a. Stokes number; Youdin & Goodman 2005).

2.1.2 Lagrangian dust particles

From Baronett et al. (2024, sec. 2.2), the dust is modeled as Lagrangian super-particles, each of which represents an ensemble of numerous identical solid particles described by their total mass and average velocity. The equations of motion for the i-th super-particle is then

$$\frac{\mathrm{d}\mathbf{x}_{\mathrm{p},i}}{\mathrm{d}t} = \mathbf{v}_i - \frac{3}{2}\Omega_\mathrm{K} x_{\mathrm{p},i} \hat{\mathbf{y}}, \tag{11}$$

$$\frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = 2\Omega_\mathrm{K} v_{i,y} \hat{\mathbf{x}} - \frac{1}{2}\Omega_\mathrm{K} v_{i,x} \hat{\mathbf{y}} - \frac{\mathbf{v}_i - \mathbf{u}}{t_\mathrm{stop}}, \tag{12}$$

where the velocity $\mathbf{v}_i$ is measured relative to the background Keplerian shear $\mathbf{v}_i' = -(3/2)\Omega_\mathrm{K} x_{\mathrm{p},i} \hat{\mathbf{y}}$. The right-hand side of equation (12) parallels equation (2) in Lagrangian form without the radial gas pressure gradient. The gas velocity $\mathbf{u}$ is interpolated at the particle position $\mathbf{x}_{\mathrm{p},i}$ using the Triangular-Shaped-Cloud scheme under the standard particle–mesh method (Hockney & Eastwood 1981). For a monodisperse population of dust, the stopping times t_stop and τ_s (equation 10) are the same for all particles. As with the gas (equations 7– 9), the initial components of the particle velocity take the equilibrium solution by Nakagawa et al. (1986):

$$v_{i,x,0} = -\left[\frac{2\tau_\mathrm{s}}{(1+\epsilon)^2 + \tau_\mathrm{s}^2} \right] \eta v_\mathrm{K}, \tag{13}$$

$$v_{i,y,0} = -\left[1 - \frac{\tau_\mathrm{s}^2}{(1+\epsilon)^2 + \tau_\mathrm{s}^2} \right] \frac{\eta v_\mathrm{K}}{1+\epsilon}, \tag{14}$$

$$v_{i,z,0} = 0. \tag{15}$$

2.1.3 Pressureless dust fluid

From Youdin & Johansen (2007, sec. 2.1.1), the continuity and momentum equations for the inviscid ($\nu = 0$) and pressureless dust fluid are

$$\frac{\partial \rho_\mathrm{d}}{\partial t} + \nabla \cdot (\rho_\mathrm{d} \mathbf{v}) = 0, \tag{16}$$

$$\frac{\partial \rho_\mathrm{d} \mathbf{v}}{\partial t} + \nabla \cdot (\rho_\mathrm{d} \mathbf{v}\mathbf{v} + P\mathbf{I}) = \rho_\mathrm{d} \left[2\Omega_\mathrm{K} v_y \hat{\mathbf{x}} - \frac{1}{2}\Omega_\mathrm{K} v_x \hat{\mathbf{y}} - \frac{1}{\rho_\mathrm{d}} \left(\frac{\mathbf{v} - \mathbf{u}}{t_\mathrm{stop}} \right) \right], \tag{17}$$

respectively. Solving for the dust density ρ_d, the dust velocity $\mathbf{v}$ is measured relative to the background Keplerian shear flow $\mathbf{v}' = -(3/2)\Omega_\mathrm{K} x \hat{\mathbf{y}}$. The right-hand side of equation (17) parallels equation (2) without the radial gas pressure gradient. As with the gas (equations 7– 9), the initial components of the dust velocity take the equilibrium solution by Nakagawa et al. (1986):

$$v_{x,0} = -\left[\frac{2\tau_\mathrm{s}}{(1+\epsilon)^2 + \tau_\mathrm{s}^2} \right] \eta v_\mathrm{K}, \tag{18}$$

$$v_{y,0} = -\left[1 - \frac{\tau_\mathrm{s}^2}{(1+\epsilon)^2 + \tau_\mathrm{s}^2} \right] \frac{\eta v_\mathrm{K}}{1+\epsilon}, \tag{19}$$

$$v_{z,0} = 0. \tag{20}$$

2.2 Problems

As in Johansen & Youdin (2007), the problems below are intended to study the non-linear saturation of the streaming instability. The

Table 1. Parameters for the unstratified problems, the different dust implementations, and their variations (Section 2.2). The columns are (1) problem name, (2) dimensionless stopping time[a], (3) total dust-to-gas mass ratio[b], (4) domain size, (5) snapshot times, (6) snapshot keywords, (7) time series cadence, (8) time series keywords, (9) average number of Lagrangian particles per cell n_p or Gaussian velocity noise $\mathrm{d}\mathbf{v}$, and (10) grid resolution. Length, time, and density are in units of gas scale height H_g, orbital period T[c], and initially uniform gas density $\rho_{\mathrm{g},0}$, respectively.

Problem	τ_s	ϵ	$L_x = L_y = L_z$ (H_g)	Snapshots t_sim/T	Snapshots keywords	Time series $\mathrm{d}t/T$	Time series keywords	Dust	$N_x \times N_y \times N_z$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
BA	1.0	0.2	2.0	0, 5, 10, 20, 50, 100	`rhop`	0.1	`maxrhop`	$n_\mathrm{p} = 1$	$512 \times 1 \times 512$
								$n_\mathrm{p} = 1$	$1024 \times 1 \times 1024$
								$n_\mathrm{p} = 9$	$512 \times 1 \times 512$
								$\mathrm{d}\mathbf{v} = 0.01c_\mathrm{s}$	$512 \times 1 \times 512$
								$\mathrm{d}\mathbf{v} = 0.01c_\mathrm{s}$	$1024 \times 1 \times 1024$

[a] Defined by equation (10)
[b] Defined by equation (5)
[c] Defined by equation (21)

parameter values for the following problems, the different dust implementations, and their associated variations are summarized in Table 1.

2.2.1 BA

This problem and its associated variations are based on run "BA" from Johansen & Youdin (2007). As key parameters, $\tau_\mathrm{s} = 1.0$, and $\epsilon = 0.2$. The domain size is $L_x \times L_y \times L_z = 2\ H_\mathrm{g} \times 2\ H_\mathrm{g} \times 2\ H_\mathrm{g}$ (Section 2.1.1).

Output data should include simulation snapshots and a time series (Section 1.1.2). The grid coordinates should be in units of the vertical gas scale height H_g. The snapshots should be taken at $t_\mathrm{sim}/T = 0$, 5, 10, 20, 50, and 100, where

$$T \equiv 2\pi/\Omega_\mathrm{K}, \tag{21}$$

is the local orbital period. These should contain dust density field in units of the initially uniform gas density, i.e. $\rho_\mathrm{p}(x/H_\mathrm{g}, z/H_\mathrm{g})/\rho_{\mathrm{g},0}$ (Section 2.1.1), stored with the keyword argument `rhop`. The time series should include the maximum particle density in the domain in units of the initially uniform gas density, i.e. $\max(\rho_\mathrm{p})/\rho_{\mathrm{g},0}$, stored with the keyword argument `maxrhop` at a cadence of $\mathrm{d}t = 0.1T$.

For simulations that implement Lagrangian dust particles, the requested dust density quantities should be mapped onto the gas grid via the particle–mesh assignment (Section 2.1.2). The first two variations should use a total number of particles, which are randomly distributed throughout the domain, such that there are $n_\mathrm{p} = 1$ particles per cell on average. For $n_\mathrm{p} = 1$, the first and second variations should have grid resolutions of $N_x \times N_y \times N_z = 512 \times 1 \times 512$ and $1024 \times 1 \times 1024$, respectively. The third variation should have $n_\mathrm{p} = 9$ with $N_x \times N_y \times N_z = 512 \times 1 \times 512$.

Simulations that implement a pressureless dust fluid (Section 2.1.3) should initially perturb the fluid with Gaussian noise at 1% of the sound speed for all velocities, i.e. $\mathrm{d}\mathbf{v} = 0.01c_\mathrm{s}$. The first and second variations for this implementation should have resolutions of $N_x \times N_y \times N_z = 512 \times 1 \times 512$ and $1024 \times 1 \times 1024$, respectively

The parameter values for the different dust implementations and their associated variations are included in Table 1. Code comparison objectives for this problem and its variations include comparing morphologies, maximum density evolution, and cumulative distribution functions.

REFERENCES

Bai X.-N., Stone J. M., 2010, ApJ, 722, 1437
Baronett S. A., Yang C.-C., Zhu Z., 2024, MNRAS, 529, 275
Goldreich P., Lynden-Bell D., 1965, MNRAS, 130, 125
Hockney R. W., Eastwood J. W., 1981, Computer Simulation Using Particles. McGraw-Hill
Johansen A., Youdin A., 2007, ApJ, 662, 627
Li R., Youdin A. N., 2021, ApJ, 919, 107
Nakagawa Y., Sekiya M., Hayashi C., 1986, Icarus, 67, 375
Youdin A. N., Goodman J., 2005, ApJ, 620, 459
Youdin A., Johansen A., 2007, ApJ, 662, 613

APPENDIX A: PENDING

The following sections are works in progress, some of which may be included in the main text above in future revisions of this document.

APPENDIX B: UNSTRATIFIED

B1 Problems

B1.1 AB

Unstratified monodisperse streaming instability. This is run "AB" of Johansen & Youdin (2007).

The box should have dimensions $L_x = L_z = 0.1H \times 0.1H$, resolution is 1024×1024, Stokes number St = 0.1, dust-to-gas ratio $\varepsilon = 1$, number of particles: 4 particle per cell (4,194,304 particles).

20 snapshots taken between 0 and 2 (on intervals of 0.1) orbits and single snapshots at 3 and 4 orbits (units of $2\pi/\Omega$). Snapshots should contain densities and velocities for the gas and particles, as well as the particle positions (can be separate snapshots).

Submit the results as numpy savez files (.npz), containing, respectively

(i) a file with the grid arrays, x and z
(ii) files with the particle density for each snapshot;
(iii) the particle positions for each snapshot;
(iv) time series (with $dt = 0.01$ orbits), containing the time and the particle and gas density and velocity dispersions, as defined by equations 10 and 11 of Baronett et al. (2024, sec. 3.1).

Objective: compare dispersions, cumulative distribution function, and morphological evolution.

B1.2 lin A

Linear, unstratified monodisperse streaming instability. This is run "lin A" of Youdin & Johansen (2007).

TO DO

Objective: reduce non-linearity of initial conditions, identify close to pure code comparisons.

APPENDIX C: STRATIFIED

C1 2D

C1.1 Lagrangian dust particles

Clumping threshold for streaming instability. This is run Z0.4t30 of Li & Youdin (2021).

The box should have dimensions $L_x = L_z = 0.8H \times 0.4H$, resolution is 1024×512, Stokes number $\mathrm{St} = 0.3$, dust-to-gas ratio $Z = 0.01$.

Number of particles should be 4 particles per cell, but considering the effective particle scale height ($H_p \approx 0.1H \approx 2\eta r$). For $\Pi = 0.05$, that's 262,144 particles.

Vertical boundary condition: reflective (zero normal velocity u_z, zero gradient for u_x and u_y).

Initial condition: Gaussian for gas density, particles settled with particle scale height $H_p = 0.025$.

Snapshots taken at 5, 10, 20, 50, 100 orbits (units of $2\pi/\Omega$). Snapshots should contain particle density and particle positions. Time series of maximum particle density.

Objective: Do codes agree on clumping?

C1.2 Pressureless dust fluid

Same as Problem 3A but for fluid. Start fluid with Gaussian noise at 1% of sound speed for all velocities.

C2 3D

3D Streaming Instability. This is a 3D extension of Problem 4A (with higher Z also, for shorter computation time).

The box should have dimensions $L_x = L_z = 0.8H \times 0.4H \times 0.4H$, resolution is $1024 \times 512 \times 512$, Stokes number $\mathrm{St} = 0.3$, dust-to-gas ratio $Z = 0.01$, number of particles: $N_w \times \Pi = 13,421,772$.

Vertical boundary condition: reflective (zero normal velocity u_z, zero gradient for u_x and u_y).

Initial condition: Gaussian for gas density, particles settled with particle scale height $H_p = 0.025$.

Midplane and vertical slice of particle density at 2, 5, 10, and 20 orbits. Full datacube at 20 orbits (units of $2\pi/\Omega$). Snapshots should contain particle density and particle positions. Time series of maximum particle density.

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