Streaming Instability in Turbulent Protoplanetary Disks

Orkan M. Umurhan$^{1,2}$, Paul R. Estrada$^2$, and Jeffrey N. Cuzzi$^2$

$^1$SETI Institute, 189 Bernardo Way, Mountain View, CA 94043, USA; orkan.m.umurhan@nasa.gov
$^2$NASA Ames Research Center, Moffett Field, CA 94035, USA

Received 2019 June 14; revised 2020 April 14; accepted 2020 April 14; published 2020 May 19

Abstract

The streaming instability for solid particles in protoplanetary disks is reexamined assuming the familiar alpha ($\alpha$) model for isotropic turbulence. Turbulence always reduces the growth rates of the streaming instability relative to values calculated for globally laminar disks. While for small values of the turbulence parameter, $\alpha < 10^{-5}$, the wavelengths of the fastest growing disturbances are small fractions of the local gas vertical scale height $H$, we find that for moderate values of the turbulence parameter, i.e., $\alpha \sim 10^{-5}$–$10^{-3}$, the length scales of maximally growing disturbances shift toward larger scales, approaching $H$. At these moderate turbulent intensities and for local particle to gas mass density ratios $\epsilon < 0.5$, the vertical scales of the most unstable modes begin to exceed the corresponding radial scales so that the instability appears in the form of vertically oriented sheets extending well beyond the particle scale height. We find that for hydrodynamical turbulent disk models reported in the literature, with $\alpha = 4 \times 10^{-5}$–$5 \times 10^{-4}$, together with state-of-the-art global evolution models of particle growth, the streaming instability is predicted to be viable within a narrow triangular patch of $\alpha$–$\tau_z$ parameter space centered on Stokes numbers, $\tau_z \sim 0.01$ and $\alpha \sim 4 \times 10^{-5}$, and further, exhibits growth rates on the order of several hundreds to thousands of orbit times for disks with 1% ($Z = 0.01$) cosmic solids abundance or metallicity. Our results are consistent with, and place in context, published numerical studies of streaming instabilities.

Unified Astronomy Thesaurus concepts: Hydrodynamics (1963); Protoplanetary disks (1300)

1 Introduction

The formation of the first 100 km size planetesimals remains a poorly understood chapter in the standard story of planetary formation. A given radial zone in a protoplanetary disk will initially contain micron-size particles of material that is solid under local conditions. These grains grow by sticking until they become submillimeter- to millimeter-sized aggregates (Birnstiel et al. 2012; Estrada et al. 2016). In the inner nebula, some (still mysterious) heating process melts these (ice-poor) aggregates and forms chondrules. These chondrules apparently can form few-centimeter-scale aggregates depending on local disk properties (Simon et al. 2018a). The picture is less clear in the outer nebula, where chondrule formation may never occur and the material remains ice rich.

However, what remains elusive is understanding how these aggregate millimeter- to centimeter-size clusters eventually coalesce into 100 km sized planetesimals, especially if the nebula is even weakly turbulent. This is because further particle evolution by “incremental growth” (by sticking) encounters several barriers—including (but not limited to) the radial drift barrier, the bouncing barrier, and the fragmentation barrier. For a more comprehensive discussion of these barrier mechanisms, see the discussion found in Brauer et al. (2008), Zsom et al. (2010), Birnstiel et al. (2012), and Estrada et al. (2016). Additional barriers to incremental growth in turbulence reappear at 1–10 km size, due to gravitational stirring of such objects by fluctuations in the gas density, much like giant molecular clouds scatter stars in our galaxy (Ida et al. 2008; Yang et al. 2009, 2012; Nelson & Gressel 2010; Gressel et al. 2011). This latter realization has led to a growing suspicion that planetesimals may have been “born big,” close to the typical 100 km sizes we see today (Johansen et al. 2007; Cuzzi et al. 2008; Morbidelli et al. 2009). Moreover, recent meteoritical work suggests that a substantial amount of the accretion that formed the solar nebula’s first planetesimals, some of them leaving behind only their molten Fe–Ni cores, and even the initial 20–50 $M_J$, core of Jupiter, occurred in less than 0.5 Ma after the formation of the first remaining solids (the so-called calcium–aluminum refractory oxide inclusions; Krijger et al. 2017). Thus, formation of sizable planetesimals seemed to have started early, well inside the snow line, and after Jupiter’s initial core formed (which probably required snow-line planetesimals as precursors). It is in this context that current planetesimal formation theories must be assessed. How can planetesimals be born big, starting very early (and continuing for several megayears), directly from millimeter- to centimeter-size objects?

One popular hypothesis is that clumps of small particles are collected into 100 km or larger planetesimals by the streaming instability (SI) and ultimately gravitational collapse. SI is a linear instability (grows without limit from small perturbations, under the right conditions) that can enhance the concentration of particles in protoplanetary disks (Youdin & Goodman 2005; Youdin & Johansen 2005; YG2005 and YJ2007 hereafter, respectively). The dynamics involves the resonant momentum exchange between the disk gas and its embedded particles treated as a pressure-free second fluid (Hopkins & Squire 2018a, 2018b; Squire & Hopkins 2018a, 2018b)—see also Section 2.1. In protoplanetary disks, the SI is strongest for axisymmetric disturbances, and the growth rates are most rapid when the local volume mass densities in the gas ($\rho_g$) and particle components ($\rho_p$) are comparable, i.e., when $\epsilon \equiv \rho_p/\rho_g \gg 1$. Linear stability analyses indicate that for laminar Keplerian flows, the SI grows fastest for small wavelengths and for near-unity Stokes numbers $\tau_z \equiv \Omega t_z$, where $t_z$ is the particle gas drag stopping time and $\Omega$ is the local disk rotation time (also see Section 2.2). Note, in our

---

$^3$ In fact, the inviscid calculation indicates that the growth rates asymptote to finite values as the wavelength of the vertical disturbance approaches zero provided the radial wavelength is larger than some minimum value. This suggests the problem may be ill posed in the inviscid limit. However, this short-wavelength catastrophe is averted when viscosity is included (see our general results below).
discussion throughout this work, we sometimes refer to $\tau_s$ by its more commonly used symbolic designator, “St.”

These features of the SI suggest that this process may play an important role either (a) in the late stage of a protoplanetary disk’s evolution when the disk has lost most of its gas because of photodissociation- or MHD-driven winds from the star or disk (e.g., Carrera et al. 2017; Ercolano et al. 2017), and/or accretion onto the central object, or (b) if the disk is laminar (nonturbulent), in which case the particle component can settle to high $\epsilon$ near the disk midplane. Several detailed numerical simulations of laminar disks have shown that the SI rapidly enhances local particle concentrations (e.g., Bai & Stone 2010; Lyra & Kuchner 2013; Yang & Johansen 2014; Carrera et al. 2015; Yang et al. 2017; Li et al. 2018; Schreiber & Klahr 2018, among many others). Particle concentration is further helped along if particle self-gravity is included in models—e.g., as done in Johansen et al. (2007), Simon et al. (2016, 2017), and most recently Gerbig et al. (2020)—and can drive particle enhancements to the precipice of gravitational collapse and onwards (Johansen et al. 2015; Schäfer et al. 2017; Li et al. 2019).

However, regions of protoplanetary disks in which particle growth is of greatest interest (i.e., 1–100 au) are possibly weakly to moderately turbulent (Turner et al. 2014; Lyra & Umurhan 2019). That is, recent theoretical advances suggest that nonionized regions of protoplanetary disks support several instability processes that can lead to sustained turbulence: the vertical shear instability (VSI), convective overstability (COV), and zombie vortex instability (ZVI).

Turbulence in disks is often thought of in terms of a zero-order closure “alpha-disk” model, wherein gas turbulence is represented by an enhanced viscosity quantified by a turbulent (kinematic) viscosity coefficient, $\nu_t \equiv \alpha c_s H$, where $c_s$ and $H$ are the local isothermal sound speed and the vertical pressure scale heights, respectively (Shakura & Sunyaev 1973; Lynden-Bell & Pringle 1974). The alpha-disk model, notwithstanding its crudeness, does a good job in characterizing disk structure and consequent flow in a turbulent protoplanetary disk medium. Most recently, Stoll et al. (2017) examined the turbulent state of the VSI and found the emergence of large-scale meridional flow to be well predicted by an $\alpha$ model, albeit with effectively nonisotropic diffusion stresses owing to its characteristic large vertical motions. However, MHD turbulence might not be as well represented by an $\alpha$ model, and we keep this in mind throughout. Numerical analyses of the three above-mentioned turbulence generating mechanisms report values of $\alpha$ in the range of $10^{-3} – 10^{-5}$. This turbulence may also, in principle, diffuse particles away from the disk midplane, reducing the values of the density ratio ($\epsilon$) near the midplane, while also radially diffusing and dispersing radial concentrations of particles before they can grow appreciably (Fromang & Papaloizou 2006; Okuzumi & Hirose 2011; Zhu et al. 2015; Riols & Lesur 2018; Yang et al. 2018). On the other hand, some numerical simulations seem to show SI manifesting even in the presence of turbulence (see below).

---

4 A turbulent viscosity assumes that downscale momentum transfer occurs in the inertial range of a fully developed statistically steady turbulent fluid. The “$\alpha$ model” scales this in terms of the typical speed and length scales encountered in locally rotating sections of protoplanetary disks ($c_s$ and $H$). As such, the quantity $\alpha$ is typically interpreted to be the inverse of the local turbulent Reynolds number, $Re_t$, and might be thought of as a measure of the amplitude of the turbulent velocity field compared to the local sound speed. For a more pedagogical review appropriate to astrophysical fluids, see Regev et al. (2016).

So, what really happens to the SI in the presence of turbulence? YG2005 presented a brief and mostly qualitative discussion of the possibilities, but declined to pursue them on the grounds that protoplanetary disks were probably nonturbulent. There are a few numerical simulations examining the fate of the SI in either an axisymmetric or fully 3D model of a protoplanetary disk experiencing some level of turbulent motions, whether self-excited or driven by some other mechanism (including, but not limited to, Johansen et al. 2007; Balsara et al. 2009; Bai & Stone 2010; Tilley et al. 2010; Carrera et al. 2015; Yang et al. 2017, 2018; Li et al. 2018). Some of these studies examined the fate of particle clumping in the presence of magnetorotational turbulence, with turbulent intensity $\alpha \sim 10^{-3} – 10^{-2}$, and for a range of the two particle parameters $\tau_s$ and $\epsilon$ (Fromang & Papaloizou 2006; Johansen et al. 2007; Balsara et al. 2009; Tilley et al. 2010), while others modeled only the self-generated “midplane” turbulence surrounding a layer of particles that had settled toward the midplane of an otherwise laminar disk flow (e.g., Bai & Stone 2010; Carrera et al. 2015; Yang et al. 2017; Li et al. 2018), and most recently in a forced driven model of turbulence (Gole et al. 2020).

While the final state of the SI subject to these different kinds of turbulence indeed varies, a common point of agreement among all of these investigations is that interesting solids clumping and instability emerge for parameter values in which $\epsilon \sim 1$ and, most importantly, when $0.1 < \tau_s < 3$. The results reported in Johansen et al. (2007) are particularly noteworthy as they show that the SI emerges, in turbulence, with a preferred radial length scale of about one pressure scale height ($H$) and has a growth timescale of about 10 local orbital periods. A question facing these and other previous numerical studies of SI in turbulence is whether such combinations of initial conditions—large particles that have somehow grown without being disrupted in such moderate levels of turbulence—are self-consistent (see Section 7.7). Meanwhile, several other studies have shown that quite small particles can undergo SI in disks that are not turbulent at all, globally; the particles experience only a tiny amount of local turbulence, called “midplane turbulence,” generated by the densely settled particle layer (Barranco 2009; Carrera et al. 2015; Yang et al. 2017). Our results explain these different outcomes in a unified and consistent way.

Well-resolved numerical experiments of two-fluid processes are expensive. A theoretical prediction for the fate of the SI under arbitrary turbulent protoplanetary disk conditions would be a useful tool both in developing some quantitative estimate for the expected length and timescales of growth in such a nebula and in planning future detailed numerical experiments. We present such a theory, extending the SI analysis done in YG2005 and YJ2007 with the addition of a simple $\alpha$ model of disk turbulence, that provides an effective turbulent viscosity acting on the gas as well as a prescription for the effective turbulent particle diffusion resulting from the statistically steady stirring of particles by inertial-scale gaseous eddies. The basic assumption is that the fundamental processes driving turbulence are unaffected by the particles and lead to a statistically steady isotropic

---

5 Note that Fromang & Papaloizou’s (2006) two-fluid turbulent set-up is a possible “precursor” turbulent SI analysis, although this is not yet been demonstrated.

6 This study appeared during the revision phase of this work.
turbulence in the gas \(^7\) (see, however, Lin 2019 and Section 7.6).

We present a linear stability analysis of the SI in such an \(\alpha\)-disk model of protoplanetary disk turbulence. We determine the growth rate and wavelength of the fastest growing modes as functions of various properties including gas disk-opening angle, particle Stokes numbers, local particle-to-gas mass density ratio, and the intensity of turbulence. During the revision phase of this manuscript, Chen & Lin (2020) released a study with similar aims, and we find that our results are in mutual agreement. In Section 2, we review the basic properties of the SI, including recent theoretical developments regarding resonant drag instabilities (RDI). In Section 3, we motivate our model equations, describe the steady state, introduce infinitesimal perturbations, and discuss various caveats with respect to our representation of turbulence. Section 4 is concerned with verification. In Section 5, we survey the results of the stability analysis, focusing on the most rapidly growing modes. Section 6 applies our theory to four recently published numerical studies of the SI. In Section 7, we identify regions in the parameter space of particle Stokes numbers and disk turbulent intensity for which the SI remains a feasible path to planetesimal formation at various locations of a model disk with 1\% metallicity. We consider these in light of various known barriers to particle growth. In Section 8, we summarize our findings, discuss various issues spurred by our analysis, and point to future directions.

2. SI Mode Review

2.1. Broad Physical Picture

Because a pressure-supported gaseous disk orbits the central object at sub-Keplerian speeds, momentum exchange between gas and particles via drag forces induces a relative radial drift. In steady state, for example, a radially diminishing steady pressure profile typically causes the gas to spiral out while causing a single-size particle to spiral inward (e.g., see Equations (13)–(14)). If multiple particle size species are considered, one or several (but not all) of their smaller components can spiral outwards with the gas (e.g., Estrada et al. 2016; Benítez-Llambay et al. 2019).

The SI arises from perturbations in this relative drifting steady state and how it modifies oscillatory motions in the disk: momentum exchange is generally modeled as a function of the relative velocities between the two fluids multiplied by the product of the two fluid densities times a drag coefficient representing the type of physically appropriate drag mechanism. The momentum channeled from the mean state and into perturbations through modifications of the drag exchange term due to particle density fluctuations is the root of the linear instability. YG2005 and YJ2007 show that these density fluctuations draw momentum from the mean drift state and destabilize oscillating disk inertial waves (e.g., see YJ2007). The insightful precursor toy model of Goodman & Pindor (2000) argues that this sort of mean-momentum wave-phase-sensitive “tapping” via gas drag can generically lead to instability in otherwise damped oscillating systems.

A recent comprehensive theoretical study (Squire & Hopkins 2018a, 2018b) demonstrates that the SI is a member of a particular class of RDI. A two-fluid system, in which one component is pressure free and streaming with velocity \(w_r\) with respect to the second (nonzero pressure) fluid, is potentially resonantly unstable to any wave phenomenon with wavevector \(k\) supported by the fluid if \(w_r k\) equals the oscillation frequency of the wave phenomenon. In this broad framework, the SI is an RDI arising from the particle stream’s resonance with the inertial waves supported by the gas. We apply this prescription in rationalizing the trends contained in the inviscid models discussed in the verification in Section 4.

Generically speaking, however, the potential for instability holds for any class of waves that the fluid system might support, including sound waves, gravity waves, magnetosonic waves as well as Rossby waves and potentially many others (Hopkins & Squire 2018a, 2018b). For example, Schreiber & Klahr (2018) recently have shown that the SI occurs for non-axisymmetric vertically restricted disturbances in simulations of disks, which means that the waves with which the particle stream becomes resonant are not axisymmetric inertial oscillations but, instead, either non-axisymmetric inertial oscillations or Rossby waves. Similar effects seem to be characterizing the particle-vortex numerical experiments recently reported in Surville & Mayer (2019). Three key ingredients for resonance are that (a) there exist some means of momentum-energy exchange between the two fluid systems, for example, whether it be by means of classical fluid drag (as it is for the SI), or via dynamical drag if the two fluids are self-gravitating (e.g., see Chapter 13 of Chandrasekhar 1961), (b) a relative drift velocity between the particle and gas components manifests, and (c) the fluid component supports some kind of wave phenomenon.

2.2. Some Physical Properties

We review some of the basic physical properties of the SI based on the analysis of YJ2007. The analysis here and throughout this paper is based on a (nearly) point analysis performed at some disk cylindrical radial position \(r\). The disk position is assumed to be locally isothermal \(^8\) characterized by a temperature \(T\) and sound speed \(c_s = \sqrt{RT/\mu}\), where the gas constant \(R \approx 8314 J kg^{-1} K^{-1}\) and \(\mu\) is the gas mean molecular weight. The local rotation rate of the disk is \(\Omega\), the Keplerian rotation speed is \(v_K = r \Omega\), and the effective thickness of the disk is measured by the disk-opening angle parameter \(\delta \equiv c_s/v_K = H/r\). We assume there exists a global pressure field \(P\) whose gradient varies on the radial disk scale, i.e., \(\partial P/\partial r = O(P/r)\). The SI operates on length scales \(L_{SI}\) given dimensionally (not quantitatively) by

\[
L_{SI} \sim \frac{\partial P/\partial r}{2 \rho c_s \Omega^2 r} \sim \frac{\delta^2}{r}.
\]

This means that \(L_{SI} \sim \delta H\). In this work, we adopt the definition \(L_{SI} \equiv \delta H\). We refer to \(\delta\) as measuring the relative flow of the particles past the gas at the midplane, typically \(\sim c_s\) (see below). \(^9\) The disk-opening angle is also equal to \(\delta\).

The analysis of YJ2007 (and YG2005) assumes that the vertical variation of gas density plays no role and there is no

---

\(^7\) This tactic was hinted at in Section 3.2.2. of YG2005. Also, an assumption similar to this was employed in studying Tollmien–Schlichting waves of cold (non-MHD) turbulent protoplanetary disks (Umurhan & Shaviv 2009).

\(^8\) “Local” in the sense common in protoplanetary disk literature, namely, that it is only a function of the radial coordinate.

\(^9\) The length scale \(\delta H\) is typical also of the most unstable VSI modes (Umurhan et al. 2016), and the quantity \(\delta^2\) is the same as the pressure gradient parameter \(\eta\) of Nakagawa et al. (1986).
momentum or mass diffusion (i.e., no turbulent viscosity or diffusivity). They show that the SI is the primary instability of axisymmetric inertial modes. The stability of a given mode with radial and vertical wavenumbers, \( k_r \) and \( k_z \), respectively, is a function of the local value of \( \epsilon \) and the stopping time \( t_s \) defined according to whether the particles are in the Epstein or Stokes regimes, respectively:

\[
t_s = \begin{cases} 
\frac{\rho a}{\nu}; & \frac{a}{\lambda_{\text{mfp}}} \leq \left( \frac{4}{9} \right) \lambda_{\text{mfp}}, \\
\frac{\rho a}{\nu \rho C_d \Delta V_{\text{pg}}}; & \frac{a}{\lambda_{\text{mfp}}} > \left( \frac{4}{9} \right) \lambda_{\text{mfp}},
\end{cases}
\]

(2)

where \( \rho \) is the density of a given particle, \( a \) is the particle radius, and \( \lambda_{\text{mfp}} \) is the molecular mean free path. The Epstein regime is appropriate for particles whose radii \( a \leq \left( \frac{4}{9} \right) \lambda_{\text{mfp}} \), while the stopping times for \( a > \left( \frac{4}{9} \right) \lambda_{\text{mfp}} \) are for the Stokes regime, in which \( C_d \) is a particle drag coefficient and \( \Delta V_{\text{pg}} \) is the relative speed between a particle and the gas (Weidenschilling 1977).

Depending on the size of the particle and \( \Delta V_{\text{pg}} \), \( C_d \) may itself be a function of \( \Delta V_{\text{pg}} \) (e.g., see Estrada et al. 2016). As noted previously, the stopping times are scaled by the local orbital time, giving the “Stokes number” \( \tau_s \equiv \frac{\Omega}{t_s} \).

For a given pair of input parameters \((\tau_s, \epsilon)\), instabilities are typically expected for values of \( k_z < k_r \), and growth rates are found to be maximal for values of \( L_{\text{SI}} k_z \approx L_{\text{SI}} k_r \sim 100 \) or more (see Figure 1 of YJ2007 and Figure 2 of YG2005). That is, the wavelengths of fastest growth are usually much smaller than \( L_{\text{SI}} = \delta H \). Instability is most favorable for values of \( \epsilon \sim 1 \). Disturbingly, the problem as set up appears somewhat ill posed, in that instability appears to persist for certain finite values of \( k_z \) as \( k_z \to \infty \). In cases of most physical interest, the instability growth timescales must be much faster than the radial drift rates (see Figure 8 of YG2005). The analysis of YJ2007 assumes the gaseous component is compressible, yet they demonstrate that the contribution of gas compressibility to the instability mechanism is negligible (see also Section 4). As such, they show that the particle fluctuations are the key ingredient of the instability.

3. Model Equations

3.1. Assumptions

We build upon the inviscid model setup of YG2005. We review the assumptions made and indicate what is new to this paper.

1. The gas component is incompressible. As examined in the original studies, the growth rate of the strongest inviscid SI is on the orbit timescale, while its typical length scale \( \sim \delta H \). This means that acoustic disturbances propagate across those relevant length scales on correspondingly shorter timescales \( \sim \delta P_{\text{orb}} \), where \( P_{\text{orb}} \equiv 2 \pi / \Omega \) is the local disk orbit time.

2. Spatial variations in all disk quantities are negligible except for the mean Keplerian shear, which is assumed constant.

3. There are no disk vertical density variations (no vertical density stratification).

4. The Stokes numbers are constant.

5. We consider only particles consisting of a single mass-dominant size and, thus, ours is a “two-fluid” model. Particle growth evolutionary models suggest that the assumption of a mass-dominant size is not bad (Zsom et al. 2010; Estrada et al. 2016). Because Krapp et al. (2019) have recently shown that the streaming instability is diminished for disk models containing most reasonable particle distributions in even weakly turbulent disks, our assumption is conservatively favorable to SI.

6. Scales of interest are small enough so that it is appropriate to use the shearing box assumption, which ignores large-scale effects including curvature.

7. Gas and particle perturbations are axisymmetric.

8. (New to this paper) Turbulence is assumed isotropic and is represented in the gas momentum equation by the standard \( \alpha \)-disk model in which turbulence is modeled as an enhanced kinematic viscosity, \( \nu_t = \alpha_c \nu H \), controlled by the nondimensional parameter \( \alpha \). In other words, internal stresses arising from momentum exchange due to turbulence is modeled with the term, \( \nu_t \rho_s \nabla^2 u_s \), on the right-hand side of the gas momentum equation (e.g., Shakura & Sunyaev 1973; Lynden-Bell & Pringle 1974). We also include the effect of radial accretion of the gas component due to the underlying turbulence-driven viscous evolution.

9. (New to this paper) Turbulence causes stirring of the particle component, represented by a turbulent diffusion term as a source term in the particle mass conservation equation (Cuzzi et al. 1993; Dobrovolskis et al. 1999; Youdin & Lithwick 2007; Carballido et al. 2011; Estrada et al. 2016):

\[
\frac{\alpha_c H}{1 + \tau_s^2} \nabla^2 \rho_p,
\]

which captures the effect of diminished stirring of particles with large inertias (\( \tau_s \gg 1 \)).

10. (New to this paper) Because collisions between particles are not important, the particle phase is typically assumed to be pressure free. However, turbulent stirring introduces an effective pressure gradient upon the particle momentum conservation in the form of an effective particle pressure term,

\[
-c_d^2 \nabla \rho_p, \quad c_d^2 = \frac{\alpha_c^2}{1 + \tau_s^2}
\]

(Cuzzi et al. 1993; Dobrovolskis et al. 1999; Jacquet et al. 2011). The effects of this “particle pressure” term are small, except for the wavelengths of the fastest growing mode.

We note a possible shortcoming in adopting the \( \alpha \)-disk model, namely the assumption that the turbulence is isotropic. It is known that numerical studies of at least one of the proposed hydrodynamical mechanisms that may drive turbulence in protoplanetary disk ohmic (“dead”) zones (i.e., the VSI) have seen turbulent stresses that are clearly nonisotropic (Stoll et al. 2017), as well as in dead zones (DZs) whose activity is driven by sandwiching MHD turbulent layers like in Yang et al. (2018). Incorporating such higher order effects would require formulating a Reynolds averaging type of mixing scale model that takes into account the anisotropy of the

\footnote{Krapp et al. (2019) show that only for particle distributions with a very restricted range in particle sizes and mass loading, i.e., \( 0.0001 < \varsigma < 0.01 \) with \( \epsilon > 0.4 \), are the growth rates of the SI enhanced compared to the two-fluid approach typically taken. See their Figure 5.}
shear stresses following the approaches found in Cuzzi et al. (1993) and Dobrovolskis et al. (1999). This should be revisited in future analyses.

3.2. Equations of Motion and Steady State

We write the fundamental equations of motion in the local frame rotating at \( \Omega \). The radial coordinate is \( r \), the azimuthal coordinate \( \varphi \), and the vertical coordinate \( z \). We represent the mean azimuthal shear as a departure from a mean Keplerian state \( \Omega_0 = -(3/2)\Omega(r - r_0)^2 \) (Umurhan & Regge 2004). The disk gas velocity relative to the mean state is \( v_g = u_g \hat{e}_r + v_g \hat{e}_\varphi + w_g \hat{e}_z \), while the corresponding relative particle velocity is \( v_p = u_p \hat{e}_r + v_p \hat{e}_\varphi + w_p \hat{e}_z \). By writing \( P = \tilde{P} + P \), we split the pressure field into a sum of the background field \( P \)—i.e., which drives the relative mean motion between the gas and particles—and perturbation field \( \tilde{P} \). The axisymmetric equations of motion for the gas are

\[
\rho_g (\partial_t + u_g \partial_r + w_g \partial_z) u_g - 2\Omega \rho_g v_g = -\partial_r P - \partial_r \tilde{P} \partial_r + \rho_g \rho_g \mu (u_g - u_p) + (\alpha_c H \rho_g)^2 \nabla^2 u_g, \tag{4}
\]

\[
\rho_g (\partial_t + u_g \partial_r + w_g \partial_z) v_g + \frac{1}{2} \Omega \rho_g u_g = \rho_g \rho_g \mu (v_g - v_p) + (\alpha_c H \rho_g)^2 \nabla^2 v_g - (3/4)\alpha_c \rho_g \delta_c \Omega, \tag{5}
\]

\[
\rho_g (\partial_t + u_g \partial_r + w_g \partial_z) w_g = -\partial_r P + \rho_g \rho_g \mu (w_g - w_p) + (\alpha_c H \rho_g)^2 \nabla^2 w_g, \tag{6}
\]

\[\partial_t u_g + \partial_r u_g = 0. \tag{7}\]

Momentum exchange between gas and particle phases arises in the terms above of the type \( \rho_g (u_g - u_p) / \rho_g \), where we have also defined the parameter \( \mu \equiv 1 / \rho_g \mu_g \). It is assumed that the conditions in the disk (i.e., gas density and temperature as well as particle abundance) fall into the Epstein regime (though see Section 7.5). The background disk pressure gradient, the primary driver of the SI, is given by

\[-\partial_r \tilde{P} / \partial r = 2\rho_0 \tilde{\rho} \Omega. \tag{8}\]

In order to account for the viscous torque induced by the background alpha-disk model, we include on the left-hand side of Equation (5) the appropriate background viscous forcing in the form of an acceleration, \(- \alpha_c H \rho_g \nabla^2 v_g \). The equations of motion of the the particle phase are

\[
\rho_p (\partial_t + u_p \partial_r + w_p \partial_z) u_p - 2\rho_p \Omega v_p = -c_s^2 \partial_z \rho_p + \rho_p \rho_g \mu (u_g - u_p), \tag{9}
\]

\[
\rho_p (\partial_t + u_p \partial_r + w_p \partial_z) v_p + \frac{1}{2} \rho_p \Omega u_p = \rho_p \rho_g \mu (v_g - v_p), \tag{10}
\]

\[
\rho_p (\partial_t + u_p \partial_r + w_p \partial_z) w_p = -c_s^2 \partial_z \rho_p + \rho_p \rho_g \mu (w_g - w_p), \tag{11}
\]

\[\partial_t \rho_p + \partial_r \rho_p u_p + \partial_z \rho_p w_p = \frac{\alpha_c H}{1 + \tau_s} \nabla^2 \rho_p, \tag{12}\]

where the typical kinetic energy per unit mass of the particles induced by the turbulent stirring of the gas is given by

\[c_s^2 = \alpha_c^2 / (1 + \tau_s^2). \]

Steady uniform solutions of Equations (4)-(12) are sought assuming no vertical velocities and constant steady gas and particle densities, \( \rho_0 \) and \( \rho_p \). Following the procedures of Nakagawa et al. (1986) and YJ2007 (their Equations (7)-(8)), we have that the uniform gas velocities \( U_0 \), \( V_0 \) are

\[U_0 = \frac{2\tau_s - (3\alpha_c^2/2)(1 + \tau_s^2)}{(1 + \tau_s^2)^2 + \tau_s^2 \delta_c}, \]

\[V_0 = \frac{1 + \tau_s^2 + (3\alpha_c^2/4)\delta_c}{(1 + \tau_s^2)^2 + \tau_s^2 \delta_c}, \tag{13}\]

and the uniform particle velocities are

\[U_0 = -\frac{2\tau_s + (3\alpha_c^2/2)(1 + \tau_s^2)\delta_c}{(1 + \tau_s^2)^2 + \tau_s^2 \delta_c}, \]

\[V_0 = -\frac{1 + \tau_s^2 + (3\alpha_c^2/2)\delta_c}{(1 + \tau_s^2)^2 + \tau_s^2 \delta_c}. \tag{14}\]

(see Dipierro et al. 2018). Given the importance of the relative radial velocity between the gas and particle phases in rationalizing the SI in terms of resonant conditions (a la Hopkins & Squire 2018a, 2018b; Squire & Hopkins 2018a, 2018b), we find

\[\Delta U = U_0 - U_p = \tau_s \frac{2 + 2\tau_s - (3\alpha_c^2/2)(1 + \tau_s^2)}{(1 + \tau_s^2)^2 + \tau_s^2 \delta_c}. \tag{15}\]

3.3. Linearized Perturbations

We linearly perturb Equations (4)-(12) around this steady state according to

\[u_g \rightarrow u_g' + U_0, \quad v_g \rightarrow V_0 + v_g', \]

\[w_g \rightarrow w_g', \quad p \rightarrow p'. \tag{16}\]

for the gas quantities and

\[u_p \rightarrow u_p' + U_0, \quad v_p \rightarrow V_0 + v_p', \]

\[w_p \rightarrow w_p', \quad \rho_p \rightarrow \rho_p (1 + \Delta_p). \tag{17}\]

for the particles. Because the gas is incompressible, we further write the quantities \( u_g' \) and \( w_g' \) as derived from a perturbation stream function \( \psi' \):

\[u_g' = \partial_z \psi', \quad w_g' = -\partial_\psi \psi'. \tag{18}\]

One can formally define the azimuthal gas and particle “fluid” vorticity fields as

\[\omega_g' = \partial_z u_g' - \partial_\psi w_g', \quad \omega_p' = \partial_z u_p' - \partial_\psi w_p'. \tag{19}\]

The gas vorticity is related to the stream function via \( \omega_g' = \nabla^2 \psi' \). The perturbed quantities are then Fourier-decomposed. For example, the stream function is written as

\[\psi' \rightarrow \psi \exp(ik_x x + ik_z z - i\omega t) + c.c., \]

where \( k_x \) and \( k_z \) are the radial and vertical wavenumbers (respectively), the frequency is \( \omega \), and \( \psi \) is the normal mode amplitude. \( \omega_i \equiv \text{Im}(\omega) > 0 \) indicates growth with a corresponding e-folding timescale \( \tau_g \equiv 2\pi/\omega_i \). We restrict our
consideration to positive values of \( k_x \) and \( k_z \). As \( k_x > 0 \), values where \( \text{Re}(\omega) > 0 \) indicate outwardly propagating patterns. We are reminded that no such symmetry characteristic exists in the radial direction due to the imposed symmetry breaking provided by both the presence of turbulence and an externally imposed radially dependent pressure field \( P \).

The combined system reduces (using Mathematica) to a generalized expression for the dispersion relation of the form

\[
F(\omega; k_x, k_z, \tau, \epsilon, \delta, \alpha) = 0,
\]

in which \( F \) is a sixth-order algebraic equation in \( \omega \). In the inviscid limit, the six temporal modes correspond to two inertial waves in the gas phase, two inertial waves in the particle phase, and two “zero”-temperature acoustic modes in the particle phase (e.g., see discussion in Chapter 10 of Chandrasekhar 1961). The algebraic equation for \( \omega \) is solved using standard root-finding methods found in the commercial software MATLAB. The eigenvalue \( \omega \) depends upon five parameter expressions: the two wavenumbers \( k_x, k_z \), the Stokes number \( \tau \), the density ratio \( \epsilon = \rho_p/\rho \), and the ratio of the turbulent intensity parameter to the disk-opening ratio squared: \( \alpha/\delta^2 \). In all of our parameter scans, we assume \( \delta = 0.05 \), a typical value for nominal disk temperatures and orbit velocities. Dynamic length scales are normalized by \( \delta c_s/\Omega = \delta H = L_{\text{SI}} \) and growth rates normalized by \( \Omega \).

3.4. Turbulent Dilution Model

Finally, we restrict our choice of \( \epsilon \) to be physically consistent with the idea that a turbulent disk will loft particles away from the midplane and, consequently, result in dilution of \( \epsilon \) near the midplane. Following previous authors (Dubrulle et al. 1995; Carballido et al. 2006, 2011; Estrada et al. 2016), we estimate an effective particle scale height \( H_p \) from the balance between particle settling toward the midplane and upward lofting by turbulent motions. Thus, given an initial local ratio of the particle surface mass density to gas surface mass density, \( Z \), then in a region of vertical thickness \( H_p \), we broadly assign a local particle to gas mass density ratio via the simple relationship

\[
\epsilon = \epsilon(\alpha, \tau, Z) = Z \sqrt{\frac{\alpha + \tau}{\alpha}},
\]

which, henceforth, is referred to as the turbulent dilution model (TDM). In this form, \( H_p = \sqrt{\alpha/\left(\alpha + \tau\right)} H \). A disk with a cosmic abundance of about 1% would correspond to \( \epsilon = 0.01 \). Even if a protoplanetary disk starts out with the cosmic abundance of solids, any given radial location within that disk may have values of \( Z \) that vary with the disk’s evolution (Birnstiel et al. 2012; Estrada et al. 2016; Sengupta et al. 2019). As such, we allow for values of \( Z \) departing from the fiducial cosmic abundance value.

4. Verification

As a robustness test, we set \( \alpha \to 0 \) and were able to faithfully reproduce all individual eigenvalues and eigenvectors quoted in Table 1 of YJ2007 as well as the growth rate diagrams shown in their Figure 1. We indeed confirm that the SI is an instability of inertial modes of mixed character (i.e., particle gas modes). We note that the calculation in YJ2007 assumed a compressible gas component while our model assumes the gas to be incompressible. The fact that the growth rates are essentially identical in both calculations strongly suggests that the SI (as considered in their study) is practically insensitive to gas compressibility and, further, it would be sound to examine its evolution with the assumption of an incompressible gas.

Figure 1 shows the maximum growth rates as a function of \( k_x \) and \( k_z \), which reproduces the quality reported in YG2005. There exists a combination of \( k_x \) and \( k_z \) for which the growth rates are locally maximal. According to Squire & Hopkins (2018a), there exists a wave–drift resonance relationship identifying this combination as those values of \( k_x, k_z \) for which some collective fluid mode has a projected phase speed that resonates with the relative drift velocity between the gas and the particles. In the case of the classic SI, one such wave mode is an inertial wave. In cases for which both the mass loading is weak (small \( \epsilon \) and the coupling between gas and particles is strong (\( \tau_s < 1 \)), inertial modes in a collective gas–particle medium can be approximated by the wave response in the gas assuming no coupling to the particles. Figure 1 shows the actual growth rates determined for such a strongly coupled weakly mass-loaded model. In this extreme case, it is
elementary to show that

\[ \omega^2 = \frac{\Omega^2 k_z^2}{k_z^2 + k_z^2} \]  

(Lyra & Umurhan 2019). As there are no \( y \) disturbances, identifying the radial component of the inertial wave to the relative drift velocities means equating

\[ \omega = k_z (U_{\rho 0} - U_{\rho 0}). \]  

Inserting Equation (22) and the appropriate steady radial drift expressions from Equation (15) with \( \alpha = 0 \) into the above expression, we find that the desired \( k_z - k_z \) relationship is

\[ k_z^2 = \frac{k_z^4}{k_z^2 - k_z^2}, \quad k_z = \frac{(1 + \epsilon)^2 + 2\delta}{2\delta(1 + \epsilon)\tau_sH}. \]  

The wave–drift resonance relationship expressed in Equation (24) is shown as a dashed line over the growth rates showcased in Figure 1. We see clearly that the resonance relationship follows the maximum growth rates as one scans along \( k_z \). This lends confidence that the resonance condition is a very good predictor for identifying conditions corresponding to maximal growth of this instance of the RDI in the \( \epsilon \ll 1 \) regime.

5. Results

Our closed-form solutions (see Appendix A) permit a finely resolved sweep in parameter space varying both the Stokes number \( \tau_s \) and turbulence parameter \( \alpha \). The particle-to-gas volume mass density ratio \( \epsilon \) is automatically determined as a function of \( \alpha \) and \( \tau_s \) based on the global (unsettled) solids mass fraction \( Z \) and the TDM, Equation (21). For most parameter sweeps, we usually set \( Z \) at nominal cosmic abundance, \( Z = 0.01 \), but we also consider other values of \( Z = 0.02, 0.03, 0.04, \) and 0.08. For these values of \( Z \) but with a fixed disk-opening angle \( \delta = 0.05 \), we study the fastest growing mode as a function of \( \alpha \) and \( \tau_s \).

We caution against cavalierly linking the results of this work to the original theoretical studies (e.g., YG2005 and YJ2007) that surveyed the inviscid linear stability calculation for \( O(1) \) values of \( \epsilon \). We observe that by adopting the TDM, the inviscid limit is singular as it predicts \( \epsilon \to \infty \). Indeed, the TDM assumes that some type of quasi-steady state has emerged between the particles and the surrounding turbulent state. The only permissible pathway linking results of this model to those of the inviscid limit is to take the double limit \( \alpha \to 0, \tau_s \to 0 \), while setting \( \tau_s/\alpha \) to a finite constant. The latter allows for choosing an arbitrary value of \( \epsilon \), but the limit corresponds to dynamics involving particles instantaneously responding to the gas motions with no relative drift. In other words, this limit represents regular inviscid incompressible single-fluid gas dynamics with a slightly enhanced mean density.

Our most basic result is this: isotropic turbulence, as measured by \( \alpha \), causes the growth rates of the SI to diminish, while also increasing the wavelengths corresponding to fastest growth. This is not surprising, because the shortest wavelength modes are eaten away by the turbulent diffusion of momentum and particle concentration.

\[ \text{Figure 2. Growth of the SI in a sequential progression of turbulent intensity.} \]

Growth rates for \( k_z \) and \( k_z \) are shown scaled by \( H^{-1} \). All plots assume \( \delta = 0.05 \), \( Z = 0.01 \), and \( \tau_s = 1 \). Values of \( \epsilon \) follow the TDM, Equation (21). The solid black line denotes zero growth. The top panel shows \( \alpha = 4 \times 10^{-6} \) (weakly turbulent), middle panel shows \( \alpha = 8 \times 10^{-5} \) (moderately turbulent), and bottom panel shows \( \alpha = 10^{-3} \) (strongly turbulent). As \( \alpha \) increases, the wavenumber of peak growth systematically gets smaller, and the spatial orientation becomes more like vertically oriented, radially narrow sheets. The bottom panel represents conditions closely approximating those investigated in Johansen et al. (2007).
5.1. Individual Model Results

In this section, we show the properties of individual models. In particular, Figure 2 displays contour plots of growth rates as a function of $k_z$ and $k_x$ for three values of $\alpha$ together with $\tau_s = 1$ fixed.

For weak turbulence, there exist wavelengths corresponding to maximum growth that are both very short and have growth rates on the order of the disk rotation frequency. For example, the top panel of Figure 2 shows results for the very low value of $\alpha \approx 4 \times 10^{-6}$ with $\tau_s = 1$, perhaps corresponding to “midplane turbulence” around a settled particle layer in a globally laminar nebula. The wavelengths of maximum growth are $\lambda_x(\text{max}) \approx 2\pi/k_x(\max) \approx 0.037H$ and $\lambda_z(\text{max}) \approx 2\pi/k_z(\max) \approx 0.027H$, and thus, only a little smaller than $\delta H$. This fastest growing mode has an e-folding timescale $t_{gm} \approx 0.5P_{\text{orb}}$, where $t_{gm} \equiv t_{\text{e}}(\text{max}) \equiv [\Omega/(2\pi\omega_z(\text{max}))]P_{\text{orb}}$. As the intensity of turbulence increases (middle and lower panels), the wavelengths corresponding to maximal growth get larger and the corresponding growth rates diminish.

The middle panel of Figure 2 shows a wavenumber survey for $\alpha \approx 8 \times 10^{-5}$ and $\tau_s = 1$—a so-called weak-to-moderately turbulent model, even if the particle layer is rather densely settled. The fastest growing wavelengths in both directions are nearly equal, $\lambda_x(\text{max}) \approx 0.25H$ and $\lambda_z(\text{max}) \approx 0.34H$, and start becoming of the same order of magnitude as the disk local scale height. The corresponding growth timescale is now considerably longer with $t_{gm} \approx 6.9P_{\text{orb}}$.

The bottom panel of Figure 2 similarly shows a wavenumber survey for $\alpha \approx 10^{-3}$ and $\tau_s = 1$, a model we term strongly turbulent. Even here, according to the TDM, the particle layer has settled to a thickness of only $H/30$ because of the large $\tau_s$. The wavelengths of fastest growth become even larger, and the relative $(x, z)$ length scales become more disparate with $\lambda_x(\text{max}) \approx 1.55H$ and $\lambda_z(\text{max}) \approx 7.85H$ (implying vertical sheet-like disturbances). The corresponding e-folding growth timescale is $t_{gm} \approx 10.6P_{\text{orb}}$.

The pattern propagation of the turbulent SI also depends upon the degree of turbulence. We measure the radial pattern speed to be given by $c_r \equiv \omega_r/k_x$, where $\omega_r \equiv \text{Re}(\omega)$. The pattern propagation of the fastest growing mode is denoted by $c_r(\text{max})$ and equal to $\omega_z(\text{max})/k_z(\text{max})$. With reference to Figure 2, we see that the pattern propagation is inward for the two largest values of $\alpha$ shown while it is outward for the nominally weakly turbulent model.

The most strongly turbulent model shown in Figure 2 closely corresponds to the conditions modeled in Johansen et al. (2007). Figures 1(b)–(c) of Johansen et al. (2007) depict the growth of the SI in a magnetorotational instability (MRI)-induced turbulent setting in which $\alpha = \mathcal{O}(10^{-3})$ and where the Stokes number $\tau_s \sim 1$. A cursory examination of the growing modes in those simulations during the early linear growth phase ($t < 30P_{\text{orb}}$) shows that the radial wavelength of the most prominently growing structure is $\sim 1.2H$, with a growth timescale of $10–15P_{\text{orb}}$. Furthermore, our model predicts that the pattern speed $c_r(\text{max}) \approx 0.12H/P_{\text{orb}}$ and inward. The apparent propagation of the growing mode in Johansen et al. (2007) is also inward with a pattern speed of $0.1H/P_{\text{orb}}$ (we discuss pattern speeds further in Section 5.3). While this certainly does not prove that our simple model is sufficiently predictive, it is encouraging that it predicts features that are in both qualitative and approximate quantitative agreement with previously published numerical simulations.

5.2. Growth: General Survey

Figure 3 shows the growth timescale of the fastest growing mode, as a function of $\tau_s$ and $\alpha$ for $Z = 0.01$. There are several notable results. There exists a critical branch line, defined nearly by $\epsilon = 1$, in which the growth timescales are infinite. This critical curve extends from $\tau_s = 0$ and terminates at a...
Simulations where the SI was observed are indicated by filled symbols—seven-sided star: Gerbig et al. (2020), and crosses: Gerbig et al. (2020). Simulations where SI was not observed are indicated by open symbols: five-sided star: iMHD simulations of Yang et al. (2018). SI is only observed in numerical simulations of small-St particles in Zone I. The filled circle with red outline for Z = 0.04 corresponds to the early phase of the corresponding simulation of Yang et al. (2017); see also Table 2.

Figure 4. Like Figure 3, except different values of metallicity are shown: (top left) Z = 0.02, (top right) Z = 0.03, (bottom left) Z = 0.04, (bottom right) Z = 0.08. Simulations where the SI was observed are indicated by filled symbols—seven-sided star: Li et al. (2018), circles (squares): 2D (3D) simulations Yang et al. (2017), and crosses: Gerbig et al. (2020). Simulations where SI was not observed are indicated by open symbols: five-sided star: iMHD simulations of Yang et al. (2018). SI is only observed in numerical simulations of small-St particles in Zone I. The filled circle with red outline for Z = 0.04 corresponds to the early phase of the corresponding simulation of Yang et al. (2017); see also Table 2.

| Z    | $\tau_c$ | $\alpha_c$ | $\epsilon_c$ |
|------|----------|------------|-------------|
| 0.001| 0.044    | $4.4 \times 10^{-8}$ | $\approx 1$ |
| 0.010| 0.45     | $3.7 \times 10^{-5}$ | 1.10        |
| 0.020| 1.05     | $2.0 \times 10^{-4}$ | 1.44        |
| 0.030| 1.41     | $4.3 \times 10^{-4}$ | 1.72        |
| 0.040| 1.71     | $7.1 \times 10^{-4}$ | 1.96        |
| 0.080| 3.50     | $2.5 \times 10^{-3}$ | 3.74        |

Note. General growth timescales not shown for this parameter value.

critical value of the Stokes parameter, which we denote by $\tau_c$. For the parameter combination considered here ($Z = 0.01, \delta = 0.05$), $\tau_c \approx 0.45$ at $\alpha = \alpha_c \approx 3.3 \times 10^{-5}$. Although not perceptible for the $Z = 0.01$ case shown in Figure 3, the actual location of $\tau_c$ corresponds to a value of $\epsilon \approx 1.1$ based on the TDM. Similar growth timescale plots, calculated for several different values of Z, are shown in Figure 4. The corresponding value of $\tau_c$ more clearly corresponds to increasingly larger values of $\epsilon$ as Z increases. The critical line appears to hug the $\epsilon = 1$ line until $\tau_c$ begins to approach $\tau_c$ from below, whereupon the curve bends downward in $\alpha$, forming a beak-like shape (this is most starkly apparent for Z = 0.08 in Figure 4). This critical point always corresponds to values of $\epsilon$ larger than 1. We have summarized the observed trends in Table 1. Generally speaking, $\tau_c$ is some function of $Z$ and $\delta$, but a theory clarifying the meaning of this point and its mapping as a function of these and other parameters is not undertaken here.

The character of the turbulent SI is sensitive to whether $\tau_s < \tau_c$ or $\tau_s > \tau_c$. In the case where $\tau_s < \tau_c$, the growth rate dramatically depends on which side of the branch line one is on. For the region below the branch line (i.e., for $\epsilon > 1$)—the so-called laminar zone (Zone I)—the growth timescales are generally fairly short (less than orbit times) in broad accordance with the low-turbulence results depicted in the top panel of Figure 2 as well as in line with expectations based on published numerical and theoretical studies examining the SI in the inviscid limit (also see introductory discussion of Section 6).

On the other hand, for $\tau_s < \tau_c$ and above the branch line (i.e., $\epsilon < 1$), in the so-called turbulent regime (Zone II), the growth timescales are extremely long—anywhere from tens to thousands of local orbit times. For optically thick disks in which $\tau_s = 0.01$ and where the VSI is operating $\alpha \sim 2 \times 10^{-4}$ (Estrada et al. 2016; Malygin et al. 2017), the growth timescales are just under $10^4$ orbit times (at Jupiter, this corresponds to about $10^5$ yr). Approaching the $\epsilon = 1$ line from either side of the branch line results in growth timescales approaching infinity. In other words, at $\epsilon = 1$ the mode is marginal, neither growing nor decaying.
For $\tau_s > \tau_c$, the fate of the linear SI is different. The branch line $\epsilon = 1$ ceases to have any consequence for the growth rates. In this region, the growth rates are relatively fast, anywhere between tenths and tens of orbit times. There appears to be a boundary that separates the turbulent zone from the laminar zone (in this range around $\alpha \sim 10^{-4}$) but this is apparent only when looking at the neutral pattern propagation speed line (discussed further below). For $\alpha > 10^{-4}$ (Zone III), the growth timescales are about tens of orbit times while for $\alpha < 10^{-4}$ (Zone IV), the growth timescales are even 10–100 times shorter. Zone III is further distinguished from Zone IV in the character of the pattern speeds and propagation directions (next section). Finally, in a broad sense, we identify Zones II–III as comprising the “turbulent regime” because they embody regions of relatively large values of $\alpha$, while we refer to Zones I, IV as comprising the “laminar regime.”

We surveyed our results to test whether or not the incompressibility assumption remains valid. For a given maximally growing mode characterized by wavelength $\lambda$, we find that both the growth timescale and pattern propagation timescale (across radial distance $\lambda$) are always much longer than the sound propagation time across the same length scale, consistent with the physical basis for ignoring compressibility in the gaseous phase.

5.3. Pattern Speeds: General Survey

Figure 5 depicts the pattern propagation speed of the fastest growing modes for the same parameter sweep discussed above. We restrict our attention to $Z = 0.01$, noting that the qualitative character we report here carries over to the other values of $Z$. We immediately observe that the pattern speed is 0 along the branch line ($\epsilon = 1$) separating Zone I from II. This is consistent with predictions made for the inviscid limit (YG2005), wherein the pattern speed is zero for $\epsilon = 1$. However, beyond the critical point $\tau_c$, the zero pattern speed curve does not lie on the $\epsilon = 1$ line but, instead, follows the curve designating zero pattern speed, which roughly separates Zone III and IV—see dashed line of Figure 5.

In analyzing the inviscid SI, YG2005 predict that its pattern speed depends on whether or not $\epsilon > 1$. We find that those trends carry over to this turbulent model: the pattern speeds are outward ($c_P > 0$) in laminar Zone I ($\epsilon > 1$) while they are inward ($c_P < 0$) for turbulent Zone II. Pattern speeds in this regime are typically less than 0.01$H/P_{\text{orb}}$. A stark qualitative distinction appears in examining $c_P$ in Zones III and IV. Within the relatively turbulent Zone III, $c_P < 0$, that is, inward, just as in Zone II. However, the pattern speeds are very high, on the order 0.1–0.2$H/P_{\text{orb}}$. This high drift rate was already observed in our earlier analysis (Section 5.1) of Johansen et al.’s (2007) simulations. Meanwhile, in Zone IV, the outwardly propagating patterns also drift with relatively high speeds ($<0.05H/P_{\text{orb}}$) that are still, however, slower than those of Zone III. In either case, such high pattern speeds mean that such structures might drift out of the region of interest on relatively short timescales, before they are able to produce planetesimals—see further discussion in Section 7.2.

5.4. Mode Structure: General Survey

Similar to the previous subsection, we restrict our discussion to $Z = 0.01$, noting that the qualitative character we report here carries over to the other values of $Z$. In Figures 6 and 7, we plot the structure of the fastest growing modes, where Figure 6 shows its radial wavelength $\lambda = 2\pi/k_x$ (max), while Figure 7 displays its aspect ratio, defined to be $\lambda_c/\lambda_x = k_x/\lambda_c$. (max).
In accordance with the trends predicted in the inviscid limit, within the weakly turbulent regions (Zones I, IV) the radial wavelengths of the fastest growing modes are on the order of $H$ and get increasingly shorter as $\alpha$ decreases. In Figure 7, we indicate the location where $\lambda_{i}/\lambda_{s} \approx 1$; this lies mostly in Zone IV with some spillover into Zone I. Scanning around this region, we see that $\lambda_{i}/\lambda_{s}$ indeed remains $O(1)$, implying that the mode structure is that of axially symmetric ring structures in this rough patch of parameter space.

On the other hand, in Zone II, above the branch line $\epsilon = 1$, $\lambda_{i}$ steadily increases with $\alpha$. This trend continues into Zone III where the turbulence is still relatively strong but $\tau_{s} > \tau_{c}$. Moreover, the aspect ratio rapidly goes from flattened or tubular rings into vertically oriented sheets as $\lambda_{i}/\lambda_{s}$ steadily grows with increasing turbulent intensity. In both Figures 6 and 7, we have indicated regions for which $\lambda_{i}/\lambda_{s} > 5$; this property represents the entirety of Zones II and III. Naturally, caution should be exercised in interpreting the results in the limit that $\lambda_{i}$ greatly exceeds several scale heights. That is, when the predicted vertical length scales greatly exceed $H$ throughout the majority of Zone II, the growth timescales indicated throughout Zone II of Figures 3–4 may be lower bounds because there may be no expression of the SI for this parameter regime. It is, however, encouraging that the trends predicted here appear to be consistent with the simulation results reported in Balsara et al. (2009). In their Figure 7, for particle sizes in which the SI appears to manifest, the vertical structure in the particle density appears uniform up to the scale height of the particles themselves.

6. Comparison to Some Recent Published Studies

It is revealing to compare the predictions of the theory with published numerical simulations where SI is (filled symbols) or is not (open symbols), seen to operate. The values of $(\tau_{s}, \alpha)$ for the turbulent runs are given explicitly in the papers cited (Johansen et al. 2007; Balsara et al. 2009), and we have indicated their approximate locations on the growth timescale plot Figure 3. In particular, for these simulations (with $\delta = 0.05$, $Z = 0.01$), SI is only manifest squarely inside of Zone III, for relatively large $\tau_{s}$. We note that these simulations considered several particle species with differing Stokes numbers while the growth rates shown in the figure correspond to the linear theory inferred from assuming a single-size species into which all mass is assigned. We consider our theoretical predictions of growth rates to be upper bounds for simulations conducted with several species simultaneously present, based on the recent findings of Krapp et al. (2019). In the following, we examine several recent studies that provide enough information for us to compare their results with our theoretical predictions.

6.1. Yang et al. (2017)

The numerical experiments run by Yang et al. (2017) are more nuanced. They considered a single particle species of a given Stokes number initiated in a globally nonturbulent nebula—effectively laminar flow.$^{14}$ They initialized the particles with a Gaussian vertical distribution symmetric about the midplane with a corresponding effective scale height $H_{p} = 0.2H$ and allowed the particles to settle toward the midplane. In these simulations and others like it (Bai & Stone 2010), the layer collapses until the particle scale height reaches a minimum “bounce” value, out of which the SI begins to become manifest.

We conjecture that this first bounce corresponds to the rapid transition of the fluid into an unsteady (possibly turbulent) state. Whether or not the dynamically active fluid state during this first bounce is driven by the SI or some other mechanism like the Kelvin–Helmholtz instability remains to be established. Bai & Stone (2010) state that the long-time development of the unsteady particle layer under their conditions is not maintained by Kelvin–Helmholtz overturn but, rather, by the SI itself. However, in a footnote, they allow that this would not be the case for smaller $\delta$. Indeed, the recent study of Gerbig et al. (2020) shows that whether SI or the Kelvin–Helmholtz instability maintains the vertical diffusion of particles depends upon the value of St in concert with $\delta$, with the latter process being important for values of $Z$ in the vicinity of 0.01 for $\delta = 0.05$, or toward higher values of $Z$ as $\delta$ increases.$^{15}$ Deeper understanding of this so-called early bounce phase requires further analysis. In either case, we conjecture that it is out of an unsteady, possibly turbulent, state that the subsequent SI develops in these simulations, and it may in turn drive yet another component of turbulence. We do not think these candidate mechanisms can be, nor need to be, distinguished at this stage. To keep our terminology as generic and descriptive as possible, we refer to the fluid state during this early bounce phase as “midplane turbulence.”

In light of the preceding discussion, for all the simulations presented, Yang et al. (2017) record the characteristic value of $H_{p}/H$ at every time step. In practice, the particles settle toward a nominal minimum base state value of $(H_{p}/H)_{b}$ (the aforementioned “bounce”) representing the first instance of balance between turbulent stirring and midplane directed settling, and it is from this base state that the SI begins to emerge. For values of $\tau_{s} \approx 0.01$, this state appears to be reached by $20P_{orb}$, while for $\tau_{s} \approx 0.001$, the base state is reached at more like $80P_{orb}$.

Thus, to place the “detected SI” results of Yang et al. (2017) onto the theoretical growth timescale plots of Figures 3–4, we had to infer an effective value of the $\alpha$ parameter, henceforth $\alpha_{eff}$, characterizing the near-midplane turbulence. Following the relationship between $H_{p}$ and $H$ found after Equation (21) defining the TDM, under the assumption $\tau_{s} \gg \alpha_{eff}$, we approximate $\alpha_{eff} \approx \tau_{s}(H_{p}/H)^{2}$, yielding a corresponding $\epsilon_{eff} \approx Z\sqrt{\tau_{s}/\alpha_{eff}}$. This procedure works well for all of the simulations presented in Yang et al. (2017) that present a time series for $H_{p}/H$. In the two cases where this information is not provided, we estimate an average midplane value of $\epsilon_{est} = \rho_{p}/\rho_{g}$ by matching the color of the quantity against the provided color bar. Equating $\epsilon_{eff}$ to $\epsilon_{est}$ and subsequently inserting it into the rewritten TDM, leads to $\alpha_{eff} = \tau_{s}Z^{2}/(\epsilon_{eff})^{2}$. We have indicated the parameter locations of most of the simulations reported in Yang et al. (2017) throughout Figures 3–4.

With $\alpha_{eff}$, we computed the e-folding timescale of the fastest growing mode, $t_{eff}$, and its associated radial wavelength $\lambda_{eff} = \lambda_{c}(\alpha_{eff})$. The observed e-folding growth timescales, $t_{eff}$, in these simulations are difficult to assess based on the graphs provided. Thus, we estimate $t_{eff}$ to be less than the

---

$^{14}$ The simulations of Bai & Stone (2010) are similarly set up; however, we do not analyze them here because no spacetime plots of vertically integrated density were provided. A follow-up study examining these in more detail is warranted.

$^{15}$ This study was released during the revision phase of this article. Also, the streaming parameter $\delta$ is $2/3$ in their study.
observed saturation timescales (hereafter, \( t_{\text{sat}} \)) found tabulated in Table 2 of Yang et al. (2017). We visually determined an effective length scale of structures observed to emerge during the linear phase (hereafter, \( \lambda_{\text{obs}} \)) according to the following procedure: Figures 3 and 7 of Yang et al. (2017) show spacetime plots for the azimuthally averaged, vertically integrated, scaled particle density which they denote by \((\Sigma_p)/\Sigma_{p,0}\). Approximately after the time particles have settled near the midplane (\( t = 20P_{\text{orb}} \) or \( t = 100P_{\text{orb}} \), depending upon \( \tau_s \); see above) but well before \( t_{\text{sat}} \), we count the number \((N_p)\) of peaks in \((\Sigma_p)/\Sigma_{p,0}\) across the radial size of the simulation domain \( L_z \), and then we estimate \( \lambda_{\text{obs}} \approx L_z/N_p \). In practice, we focused on counting resolvable peaks within the first two predicted e-folding growth timescales because once coherent filamentary structures emerge, they appear to nonlinearly interact, resulting in a series of mergers before the saturated state becomes manifest. In some cases, it was difficult to resolve an unambiguous number of peaks, and this uncertainty was noted. In all cases, we counted peaks as soon as the structures appeared coherent to the eye.

The results of this activity are summarized in Table 2. Despite the crudeness of this approximate analysis, we find that our theory predicts the general properties reported in Yang et al. (2017). In particular, it is clear for those Yang et al. (2017) initial conditions, which did manifest SI, that the effective \( \epsilon_{\text{eff}} \) were always greater than 1, consistent with our predictions that growth is relatively strong for such conditions. In the two cases discussed by Yang et al. (2017) in which the SI was not observed, we predict that the growth times for those input parameter conditions are much longer than the time for which those simulations were conducted. Of interest is the case \( \tau_s = 0.001 \), \( Z = 0.03 \) where no structures were observed to emerge up to simulation time \( t = 5000P_{\text{orb}} \). According to our predictions, if the simulation were run past \( t = 5 \times 10^7 P_{\text{orb}} \), then the SI should become apparent. Similar reasoning also applies to their \( \tau_s = 0.01 \), \( Z = 0.01 \) simulation. We also note that their 2D simulation for the case \( \tau_s = 0.01 \), \( Z = 0.04 \) for the box \( \sim 0.4H \times 0.4H \) experiences a brief dip in \( H_p/H \) to \( \approx 0.004 \) before leveling out to \( H_p/H \approx 0.01 \) at \( t \approx 50P_{\text{orb}} \) (see red dotted time series shown in the middle panel row of their Figure 2 as well as the right panel of their Figure 3(b)). During this initial bounce phase (\( t < 50P_{\text{orb}} \)), the number of filamentary structures is impossible to assess. As such, we examined the behavior and character of this simulation after the initial response phase passed (after \( t = 50P_{\text{orb}} \)), when the number of filaments was first countable. For the corresponding value of \( H_p/H \) (\( \approx 0.01 \)), the predicted \( \lambda_{\text{runs}} \) better matches the corresponding measured \( \lambda_{\text{obs}} \).

### Table 2

| Run Type | \( \delta \) | \( \tau_s \) | \( Z \) | \( (\Delta H/H)_{1/3} \) | \( \epsilon_{\text{eff}} \) | \( \lambda_{\text{obs}} \) | \( \lambda_{\text{run}} \) | Identifiers |
|----------|------|------|------|-----------------|--------|-----------------|-----------------|-----------------|
| 2D       | 0.05 | 0.01 | 0.01 | \( \ldots \) | 13.8\(^d\) | \( \sim 0.90\) & | 11000 & \( \ldots \) | 1.2 & \( \ldots \) | \( (0.2H)^2 \), 2560H\(^{-1}\) |
|          | 0.05 | 0.01 | 0.02 | 0.012 | 14.5 & | 1.66 & | 166 & \(< 200 \) & | 0.13 & | \( \sim 0.10\) & 2560H\(^{-1}\) |
|          | 0.05 | 0.01 | 0.02 | 0.011 | 12.1 & | 1.81 & | 106 & \(< 100 \) & | 0.10 & | 0.08 | \( (0.4H)^2 \), 2560H\(^{-1}\) |
|          | 0.05 | 0.01 | 0.04 | 0.010 | 10.0 & | 4.0 & | 65 & \(< 50 \) & | 0.06 & | 0.05 | \( (0.2H)^2 \), 1280H\(^{-1}\) |
|          | 0.05 | 0.01 | 0.04 | 0.0035 | 1.27 & | 11.2 & | 16 & \(< 50 \) & | 0.01 & | \( n_f/\lambda \) | \( (0.4H)^2 \), 1280H\(^{-1}\) |
|          | 0.05 | 0.001 | 0.03 | \( \ldots \) | 5.4\(^d\) & | \( \sim 1.3\) & | \( 39000 \) & \( \ldots \) | 10 & | \( \ldots \) | \( (0.2H)^2 \), 1280H\(^{-1}\) |
|          | 0.05 | 0.001 | 0.04 | 0.016 | 2.25 & | 2.66 & | 1850 & \(< 1200 \) & | 0.14 & | 0.10 & 0.20 | \( (0.2H)^2 \), 1280H\(^{-1}\) |
|          | 0.05 | 0.001 | 0.04 | 0.015 | 2.56 & | 2.5 & | 2300 & \(< 2200 \) & | 0.17 & | 0.10 & 0.20 | \( (0.4H)^2 \), 1280H\(^{-1}\) |
| 3D       | 0.05 | 0.01 | 0.02 | 0.010 | 10.0 & | 2.0 & | 71 & \(< 250 \) & | 0.08 & | 0.067\(^i\) | \( (0.2H)^2 \), 640H\(^{-1}\) |
|          | 0.05 | 0.001 | 0.04 | 2.0 & | 3.6 & | 800 & \(< 650 \) & | 0.07 & | \( \sim 0.067\) & \( (0.2H)^2 \), 640H\(^{-1}\) |

**Notes.**

\(^a\) Estimated, based on the moment the particle layer has settled but before SI grows (see text).

\(^b\) Based on the lesser of either visual identification procedure or the quoted saturation time \( t_{\text{sat}} \) from their Table 2.

\(^c\) Simulation box size and resolution used in terms of grid points per scale height.

\(^d\) Based on \( \epsilon_{\text{eff}} \) as no \( H_p/H \) time series was provided for this case.

\(^e\) Based on extracting a midplane-averaged value of \( \rho_p/\rho_g \) estimated from the color bar for this particular run (see text).

\(^f\) No discernible wavelength structure observed in the time range \((t < 50P)\), perhaps because of insufficient simulation resolution.

\(^g\) Based on \( t = 100P \), the time after the observed initial adjustment bounce. Input parameters are the same as row above. See text.

\(^h\) Assessed at \( t = 100P \); see also discussion in Section 4.2 of Yang et al. (2017).

\(^i\) Assessed at \( t = 100P \) before the first filament merger occurs.

6.2. Li et al. (2018)

Li et al. (2018) examined the fate of the SI under similar circumstances and initial setups considered in Yang & Johansen (2014) but using three different boundary conditions—periodic, reflecting, and outflow—and they show that the nonlinear SI state is largely insensitive to the boundary conditions employed. Using a fixed grid resolution \((H/320 \approx 0.003H)\), Li et al. (2018) checked for the emergence of filaments for three different box sizes: \( (0.2H)^3 \), \( (0.4H)^3 \), \( (0.8H)^3 \). Their simulations were run with \( \tau_s = 0.314 \), and they quote time series quantities similar to \( H_p/H \). During the early linear “first-bounce” phase, all simulations collapsed into a layer with \( H_p/H \approx 0.001 \) within \( t = 1P_{\text{orb}} \). The predicted maximum growth timescales for these input parameters are short, i.e., \( t_{\text{run}} \approx 0.5P_{\text{orb}} \). This layer then begins to develop the SI just as it in the simulations reported in both Yang & Johansen (2014) and Yang et al. (2017). Coherent filaments undergoing epicyclic oscillations are clearly discernible by \( t = 3P_{\text{orb}} \), and it is also at this point that they undergo nonlinear merging.
We perform the same crude analysis as above based on the periodic boundary condition runs (see Figure 10 of Li et al., 2018), and the results are summarized in Table 3. We estimated $\lambda_{\text{obs}}$ by counting the number of filaments around $t = 2P_{\text{orb}}$. We choose this time because filaments are not easily identifiable at any time before this point, while for times after this point, the filaments begin their process of merging. For all simulation runs, we find $\lambda_{\text{obs}} \approx 0.015H$, while our theory predicts a value approximately half of that, $\lambda_\text{sim} \approx 0.007H$. This means that the observed fastest growing mode encompasses about five grid points while our predicted value would not be resolvable (at between two and three of their grid points). Indeed, we note that their measured value for $H_p/H$ at $t \approx 1P_{\text{orb}}$ is much smaller than their grid resolution. We conjecture that their simulations are not sufficiently resolved and should be run for at least two to three times the resolution originally used. Convergence would perhaps be indicated if during this early organization phase ($t < 1P_{\text{orb}}$) the minimum value of $H_p/H$ remains fixed with increasing resolution. What is not yet clear is how the response of the particles, and the turbulence they churn up as they approach the midplane, depends upon resolution as such a systematic study remains to be done.

6.3. Gerbig et al. (2020)

Gerbig et al. (2020) considered the fate of SI under setups and conditions similar to those of the previous two studies. One of their main aims was to disentangle to what degree the midplane turbulence churned up by the settling dust particles is driven either by the Kelvin–Helmholtz instability or the SI, and they seek to shed light on this outcome as a function of the dust streaming parameter, $\delta$ (which is equal to $\beta/2$ in their analysis), $Z$, and for a fixed value of $\tau_x = 0.2$. The results they uncovered are subtle, but it is evident from their reported simulations that when SI is not present, the dust layer thickens consistently with what one might predict based on the Richardson criterion for Kelvin–Helmholtz overturn under the influence of gravity. They provided the results for a suite of runs in which they quote a measured dust layer thickness plus its standard deviation, together with snapshots of the particle accumulation as viewed from several perspectives. For one run, $Z = 0.02$, $\tau_x = 0.2$, $\delta = 0.05$, they also provided a spacetime (radius–time) plot of the vertically integrated, azimuthally averaged particle density. Focusing on only a subset of simulations reported in their study—all before self-gravity is turned on—we compare their measured and tabulated quantities against predictions made using our theory. As per the procedure defined earlier, we estimate an $\alpha_{\text{eff}}$ based on their reported value of $H_p$. We also apply our theory to corresponding “high” and “low” values of $\alpha_{\text{eff}}$ based on their quoted one standard deviation values of $H_p$. The results of this exercise are summarized in Table 4.

We find reasonable agreement between our theoretical predictions and their reported study. The predicted versus observed wavelengths are mutually consistent with one another within one standard deviation of their reported values of $H_p$. Most of the filament wavelength structure that emerges in their simulations when SI is present is shown at late times ($t = 40P$), likely well after a significant amount of filament merging has taken place. For this reason, we suppose that the observed wavelength/average separation ($\lambda_{\text{obs}}$) should be larger than the predicted wavelength of our theory. Given the error bounds on their reported $H_p$, this appears to be a plausible interpretation for the simulation with $Z = 0.03$, $\tau_x = 0.2$, and $\delta = 0.05$, in which our predicted wavelength ($0.05H < \lambda_{\text{sim}} < 0.08H$) is smaller than the observed value ($\lambda_{\text{obs}} = 0.10H$). We also observe that for the simulation with $Z = 0.01$, $\tau_x = 0.2$, and $\delta = 0.05$, in which SI is not observed, our theory predicts that the fastest growing wavelength should be $\lambda_{\text{sim}} = 1.1H$ with a growth timescale of $\sim 70P$. Given that their simulation was run in a box with radial length equal to 0.4$H$, we predict that if the same simulation was rerun with a radial extent of at least 1.1$H$ or longer, then SI should appear. A similar prediction can be made for their $Z = 0.02$, $\tau_x = 0.2$, and $\delta = 0.10$ simulation, although the radial scale of their box should be at least twice as much, if not more (see Table 4).

6.4. Yang et al. (2018)

Yang et al. (2018) examined the response of the SI in a setup that supports two scenarios: one in which the full layer is turbulent due to the MRI (i.e., ideal MHD, and hereafter “iMHD”) and the other as a DZ, in which the midplane is ohmic and effectively MHD inactive but where the disk gas about one gas scale height away from the midplane is MRI active. They considered a particle component with $\tau_x = 0.1$ and metallicities of $Z = 0.01, 0.02, 0.04$, and 0.08. In the following, we restrict our attention to the iMHD models. We do not consider their DZ models because the SI physics contained in them are probably strongly influenced by coherent structures, a feature that is outside the purview of our theory—see Appendix B for further discussion.

For the iMHD models, we repeat the calculations done in the previous subsections, and the results are summarized in Table 5. We estimate $H_p$ for each simulation based on what we were best able to surmise to be the particle collective’s first bounce based on the time series found in the first column of their Figure 9. The value of $\alpha_{\text{eff}}$ is different from the value quoted in their Table 1, which is based on the state of the flow

| Run type | $\delta$ | $\tau_x$ | $Z$ | $H_p$ | $\lambda_{\text{sim}}$ | $\tau_x$ | $H_p$ | $\lambda_{\text{obs}}$ | Identifiers |
|---------|---------|---------|-----|------|------------------------|---------|------|------------------------|------------|
| 3D      | 0.05    | 0.314   | 0.02 | 0.001| 3.14                   | 0.5     | 0.05 | ~5                     | 0.007      | ~0.015 (0.2H)², 320H⁻¹ |
|         | 0.05    | 0.314   | 0.02 | 0.001| 3.14                   | 0.5     | 0.05 | ~5                     | 0.007      | ~0.016 (0.4H)², 320H⁻¹ |
|         | 0.05    | 0.314   | 0.02 | 0.001| 3.14                   | 0.5     | 0.05 | ~5                     | 0.007      | ~0.016 (0.8H)², 320H⁻¹ |

Notes. Periodic boundary condition runs shown. Table headings are defined as in Table 2.

* Based on the first clear appearance of coherent structures.

* Peaks were counted at approximately $t = 3P$, the point at which coherence was discernible to the human eye.

Table 3
Selected Simulations of Li et al. (2018) Compared to Theoretical Predictions
at very late stages of their simulation. For all of the metallicities considered, we find that the fastest growing modes have wavelengths (\(\sim 30\)) much larger than the radial computational domain size (\(\sim 4H\)). The predicted growth timescales are several hundred orbit times (at the very least), which are far longer than the 100\(P\) timescales on which their simulations were conducted.

However, inspection of their results shows that the iMHD model achieves very modest increases in particle enhancements over and above the nominal base value of either the vertically integrated particle densities, i.e., \(\Sigma_{p}/\Sigma_{p,0}\) where enhancements range from factors of 2 (on average) to no more than factors of 5 in extremes (see Table 3 of their study). The particles are prevented from settling to the midplane \(H_{p}/H \sim 0.3\) owing to the large vertical fluid stresses arising from the MRI turbulence with effective values of \(\alpha \sim 0.008\). The top row of Figure 10 of Yang et al. (2018) depicts a spacetime diagram showing \(\Sigma_{p}\) as a function of radial position and time for the four iMHD models of differing \(Z\) compared against a sample run in which there is no backreaction of the particles back onto the fluid. Except for the \(Z = 0.08\) model,
inspection of the remaining figures show that there is no obvious distinction between those simulations with and without backreaction, as there is neither clear nor sustained emergent radial structure beyond the periodic box scale. In the $Z = 0.08$ model (far right, top-row panel of their Figure 10), there is a brief period (between $t/P_{\text{orb}} = 20$ and 40) in which strong accumulation appears to take place but then it eventually dissipates.

We interpret the possibility of ephemeral bursts of accumulation to be the result of concentrations driven into place by large-scale coherent structures of the flow field of the turbulent state, or perhaps even by turbulent concentration (Hartlep & Cuzzi 2020). For example, the driving motions of the MRI (channel modes) and its secondary turbulent transition (parasitic/Kelvin–Helmholtz overturn) are primarily axisymmetric, which means that the corresponding fluid eddies and waves of the driving scales are azimuthally aligned—meaning that there are coherent structures in the flow and pressure field being imposed upon the flow by the unstable MRI activity. Such motions may be responsible for any temporary accumulations seen in these iMHD models, instead of SI per se. Moreover, our theory is neither able to treat nor predict what happens to the SI in the presence of unsteady yet coherent structures: the simple turbulence model like we have employed here assumes the motions are uncorrelated (implicitly placing it within the inertial range of a turbulent flow). See more discussion in Appendix B.

7. Turbulent SI Validity Regime Constrained by Realistic Protoplanetary Disk Models

We now put these results into the context of realistic global disk evolution and particle growth models, and other observed constraints. Recall that our theory assumes a single dominant, mass-carrying particle size, Benítez-Llambay et al. (2019) and Krapp et al. (2019) as well as others note that when multiple particle species are present in the mix, the overall growth rate of the SI may be significantly modified (probably diminished) compared to single-sized particles (Krapp et al. 2019).

Thus, we ask the questions: for what turbulent disk conditions and particle properties does the SI provide a direct path to planetesimal formation, and are these initial conditions realistic and self-consistent? We consider these questions for three general locations in the solar nebula, each of which nominally representing (i) the inner disk at $R \approx 3$ au, $T = 265$ K with $c_s \approx 1050$ m s$^{-1}$ (assuming a H$_2$ gas), (ii) the snow line around $R \approx 5$ au, $T = 140$ K with $c_s \approx 760$ m s$^{-1}$, and (iii) the outer disk at $R = 30$ au, $T = 73$ K with $c_s \approx 550$ m s$^{-1}$.

7.1. Disk Lifetime Constraints

For the inner disk, we rule out SI models that predict growth timescales that are significantly longer than $\sim 1$ Ma, based on the evidence that planetesimals were abundantly forming before that time (Kruijer et al. 2017; Scott et al. 2018; see also Section 1). This translates to some number of equivalent $P_{\text{orb}}$ depending upon where the turbulent SI theory is being modeled. For example, we nominally place the snow line at 5 au, which means that model parameters that predict growth timescales in excess of $(0.4-2.0) \times 10^5 P_{\text{orb}}$ are ruled out.

7.2. Particle and Mode Radial Drift

Growing particles will drift radially at different rates due to variable coupling with the nebula gas. The drift rate increases with Stokes number $\tau_s$, reaching a peak when $\tau_s = t_{\text{dr}} \Omega \sim 1$, and then decreases for larger sizes. A Stokes number of unity corresponds to different size particles at different places in the protoplanetary disk, depending on the local gas surface density. For the standard minimum mass solar nebula, meter-sized particles drift the fastest in the terrestrial planet region, whereas farther out in the disk, where gas densities are low and the pressure gradients can be strong (especially if there is a strong gradient in the gas density at the disk outer edge), much smaller (millimeter-size) particles have $\tau_s = 1$. In fact, inward-drifting particles may drift faster than they can grow in size by sticking; this is the so-called “radial drift barrier” (e.g., Brauer et al. 2008; Birnstiel et al. 2012; Estrada et al. 2016; Sengupta et al. 2019).

A similar criterion can be applied to the SI. Using the mean radial velocity of the particle component (Equation (14)) to estimate the time $t_{\text{dr}}$ a particle takes to traverse the scale of the disk, we find (for $\tau_s \gg \alpha$)

$$t_{\text{dr}}^{-1} = \frac{|U_{\text{pol}}|}{r} = \frac{2\pi \delta^2 \Omega}{(1 + \epsilon)^2 + \tau_s^2}. \tag{25}$$

For a given set of parameters, the SI is considered “viable” if the derived growth rate is faster than the drain rate, i.e., when $\omega_i > t_{\text{dr}}^{-1}$. Growth rates depend very much on $\epsilon$, but sufficiently high solids-to-gas ratios that allow the solids to drive the gas motions are hard to achieve in turbulence (e.g., see Estrada et al. 2016), unless one imposes arbitrary trapping mechanisms such as pressure bumps to thwart radial drift (e.g., Drazkowska et al. 2013, 2014). Fractal particle growth leads to highly porous particles, which drift radially much more slowly (Ormel et al. 2007; Estrada & Cuzzi 2008; Okuzumi et al. 2012; P. R. Estrada et al. 2020, in preparation); this may provide a means to weaken the radial drift barrier and generate the necessary solids enhancements, but decreasing $\tau_s$ also shifts the case to the left in Figure 3, generally weakening SI for any $\alpha$.

As in the simulations reported in Johansen et al. (2007), when $\tau_s \sim \mathcal{O}(1)$ the pattern drift of growing modes is relatively fast, and there emerges the possibility that the growing mode drifts in toward the star faster than the overdensity can grow sufficiently for gravitational instability to take root. This concerns the notion of “convective instability” familiar in hydrodynamics (Regev et al. 2016; Drazin et al. 2020) and plasma physics (see Chapter 18 of Bers 2016). As such, we assess the conditions in which the pattern drift timescale (to reach the star) is much shorter than the unstable growth timescale. The time it takes for inwardly propagating structures to drift into the star is $t_w = r/|c_r|$ (provided $c_r < 0$); so, if $\omega_i > t_w^{-1}$, we consider the SI viable.

7.3. Fragmentation Barrier

Several studies (Brauer et al. 2008; Birnstiel et al. 2012; Estrada et al. 2016; Sengupta et al. 2019) show that the fragmentation barrier in a turbulent medium may be estimated by assessing the inequality

$$\alpha c_s U_{\text{pol}}^2 > 2\tau_s \alpha c_s^2 \tag{26}.$$
mass, which is quantified by \( U_f \), the fragmentation speed for the particle (in reality a loose aggregate) in question. For silicate aggregates, \( U_f \approx 1.4 \text{ m s}^{-1} \) (Güttler et al. 2010; Zsom et al. 2010).\(^{16}\)

The value of \( U_f \) for H\(_2\)O ice aggregates is more nuanced. Laboratory studies of material properties suggest that sticking is significantly more effective for H\(_2\)O ice than for silicates (Bridges et al. 1996; Wada et al. 2009), allowing ice particles to grow larger, faster, and more porous, with effective fragmentation velocity thresholds an order of magnitude or more higher than for silicates (Wada et al. 2009; Okuzumi et al. 2012; Wada et al. 2013). On the other hand, Musiolik & Wurm (2019) reported on experimental results testing the surface energy of \( \sim 1 \text{ mm H}_2\text{O} \) ice spheres in the 175 K < \( T < 240 \text{ K} \) temperature range and find that this increased strength only applies in a narrow temperature range plateauing near 200 K, with effective sticking speeds almost 30 times lower when the grains are cold (i.e., \(< 175 \text{ K} \)). They find that at these colder temperatures, the surface energy of H\(_2\)O ice grains is about the same as the surface energy of silicate particles—suggesting that under very cold protoplanetary disk conditions, collisional growth may not favor ice over silicates.

Based on the above findings, we consider two different values of \( U_f \) for H\(_2\)O ice: when grains are near the ice line (\( \sim 5 \text{ au} \)), we adopt a fragmentation velocity of \( U_f \approx 8.9 \text{ m s}^{-1} \), which we refer to as the sticky H\(_2\)O ice fragmentation speed consistent with both previous studies (e.g., Wada et al. 2009) and those of Musiolik & Wurm (2019). Well outside the snow line where temperatures are low (\( \sim 30 \text{ au} \)), we consider both these stronger ice particles and also a fragmentation velocity for ice having the same value as silicate particles (\( U_f \approx 1.4 \text{ m s}^{-1} \); Güttler et al. 2010; Zsom et al. 2010), which we refer to as the cold H\(_2\)O ice fragmentation speed.

### 7.4. Bouncing and Drift Barriers

Numerous laboratory and theoretical studies have found that particle growth can be influenced by bouncing at much smaller sizes than the size that collides so energetically as to fragment the particles (Zsom et al. 2010). Though Estrada et al. (2016) derive an expression for bouncing between similar-sized particles (their Equation (59)), this size is harder to specify rigorously because it likely also depends on material properties (Güttler et al. 2010; Zsom et al. 2010). Estrada et al. (2016) emphasize that the bouncing barrier is not impermeable but merely slows growth to a great degree. For this reason, it is not as instructive to include the bouncing barrier as a constraint as we do for fragmentation. A similar situation holds for the radial drift barrier alluded to earlier; it is hard to write quantitative values for this size limit (e.g., see Equations (60)–(61), Estrada et al. 2016). Our fragmentation barrier upper limits on particle size (Equation (26)) are thus approximate. Under some conditions, particles may never get that large, due to a combination of the bouncing and radial drift barriers (Birnstiel et al. 2012; Takeuchi et al. 2012; Estrada et al. 2016; Sengupta et al. 2019). Under other conditions, growth can proceed somewhat beyond the fragmentation barrier if a distribution of collision velocities and mass transfer in collisions is allowed for (e.g., Windmark et al. 2012; Drazkowska et al. 2013, 2014; Estrada et al. 2016). Bouncing, drift, mass transfer, and a probability distribution for collisional velocities are included in the more realistic growth model constraints discussed in Section 7.7.

### 7.5. Combined Limits on Particle Size: \( \tau_s \)

Global numerical simulations suggest that the series of barriers to growth discussed previously are quite effective at limiting the \( \tau_s \) of the particle size that dominates the mass (e.g., Birnstiel et al. 2012; Estrada et al. 2016; Sengupta et al. 2019). Recent simulations conducted for solid particle growth over a three-order-of-magnitude range of turbulent intensities (\( 10^{-5} < \alpha < 10^{-2} \)) relevant to the first 0.5 Ma where planetesimal formation is thought to begin (see Section 1) further indicate that the maximum achieved Stokes numbers of the mass-dominant particles\(^{17}\) range from \( \sim 0.001 \) to \( 0.01 \) over a wide range of disk models with initial global metallicities of \( Z \approx 0.01 \) (Estrada et al. 2016; P. R. Estrada et al. 2020, in preparation). Table 6 summarizes these Stokes numbers, the corresponding particle radii, and ambient disk conditions at our three representative radial locations within the protoplanetary disk. We note that in these simulations, the local solids abundance and the nebula gas density and temperature (and thus \( \alpha \), as well as the pressure gradient) are evolving with time, which means that \( Z \) can have a range of values over the disk’s radial extent. In particular, the snow line (and various other “evaporation fronts”) evolves with time, leading to both local enhancements and depletions in the amounts of solids, especially in the planet-forming region. The values given in Table 6 are selected at locations where the solids-to-gas ratio \( Z(\tau) \approx 0.01 \) and which correspond to the “inner disk” and “snow line”; they may thus lie at semimajor axes that are slightly different from our nominal definitions of \( R = 3 \) and 5 au, respectively. The “outer disk” location reliably corresponds to 30 au (see caption, Table 6).

Despite the variable metallicity, the particle size distributions have already reached a quasi-equilibrium state at the selected times and radial locations, changing only slowly with time and depending only weakly on the instantaneous value of \( Z \) (see Figure 19 of Estrada et al. 2016 and associated discussion). The physics that limits the mass-dominant particle \( \tau_s \) in different radial regions depends on the ambient nebula conditions. In the inner disk regions where bulk composition is ice free, \( c_i \) and \( \Omega \) are large so turbulent relative velocities are fairly large. These regions tend to be in the fragmentation regime based on previous discussion, using Equation (26) for the fragmentation equilibrium Stokes number. However, for \( \alpha \gtrsim 10^{-3} \), the values for the fragmentation \( \tau_s \) obtained from Equation (26) using the parameters cited in Section 7 are consistently smaller than the values listed in Table 6.

\(^{16}\)We note, however, that it has been suggested that the fragmentation speeds of relatively high-temperature (\( T \sim 200–400 \text{ K} \)) aggregates composed of organic covered silicate submicron grains may approach 100s of m s\(^{-1} \) (Homma et al. 2019). For the above quoted temperature range, a 0.03 \( \mu \text{m} \) monomer Si aggregate with about equal amount (by radius) of organic covering has a fragmentation velocity of \( \sim 100 \text{ m s}^{-1} \), while a 1 \( \mu \text{m} \) sized Si aggregate with about 3% organic coating (by radius) has a smaller fragmentation velocity diminishing to about 10 m s\(^{-1} \) (see their Figure 4), with the assertion that this reduction in critical speeds continues for larger masses. While it is unlikely that such submicron-sized aggregates collide with one another with such high speeds in the turbulent nebula, this structural strengthening feature of organic covered Si grains should nonetheless be incorporated into future analyses.

\(^{17}\)As in some of the models discussed in Section 6, the particle growth models of Estrada et al. (2016) and P. R. Estrada et al. (2020, in preparation) exhibit a broad size distribution and do not employ particles of a single size. However, in general, most of the particle mass is near the fragmentation size when drift is not important, or near the largest particle size in the distribution when in the drift-dominated regime—either way, defining a single particle size.
because the actual growth models include growth beyond the fragmentation barrier (see Section 7.4). On the other hand, for \( \alpha = 10^{-4} - 10^{-5} \), one finds Equation (26) gives \( \tau_s \approx 0.01-0.1 \), which is consistently larger than the corresponding values in Table 6, despite the temperatures being closer to those cited in Section 7, and in fact, these particles have already reached the fragmentation barrier. As noted in Table 6, these mass-dominant particles are in the Stokes regime, which for these relatively small \( \tau_s \) means the relative velocities between them tend to be higher for a given Stokes number compared to the Epstein regime. This discrepancy does not depend on which flow regime the particles are in, but rather because eddy-crossing effects start to become important in weak turbulence, particularly when \( \delta / \alpha \gg \tau_s^{-1} \). Under these circumstances, the relative velocities between the particles are higher, even for similar-sized particles, which means the fragmentation \( \tau_s \) will be smaller than what Equation (26) would predict (Ormel & Cuzzi 2007).

In the colder, ice-rich outer disk, a naive application of Equation (26) would give \( \tau_s \gtrsim 1 \) for \( \alpha \lesssim 10^{-4} \) for the sticky water-ice fragmentation case. However, the simple constraint of Equation (26) assumes that the collision velocity is dominated by turbulent relative velocities, but large values of the pressure gradient (manifested by \( \delta \)) in the outer nebula can drive systematic drift- and headwind-related velocities that can significantly exceed those due to turbulence for these weak values of \( \alpha \). Under these circumstances, bouncing plays an even more influential role by slowing growth, enhancing the importance of the drift barrier in precluding the fragmentation barrier from being reached. Simulations of the outer disk where particles drift faster than they can grow (Figures 12 and 13) are characterized by Stokes numbers \( \tau_s \lesssim 0.01 \) that are nearly constant or modestly varying with radius, which suggest that the decrease in particle radii with distance simply mirrors the decrease in gas density (Birnstiel et al. 2012; Estrada et al. 2016; P. R. Estrada et al. 2020, in preparation). Even when the bouncing barrier was not considered, \( \tau_s \lesssim 0.1 \) in these regions. On the other hand, for large values of \( \alpha \gtrsim 10^{-2} \), one finds that even in the outer disk particles can be in the fragmentation-dominated regime even if their Stokes numbers remain comparable to the lower \( \alpha \) cases, because the fragmentation barrier occurs at much smaller sizes there (see Table 6).

Overall, Equation (26) is a handy but crude approximation in general. Estrada et al. (2016) give a more in-depth discussion and derive estimates of the limits of \( \tau_s \) as it pertains to their models. The variation in \( \tau_s \) and particle size over different radial locations and with time listed in Table 6 is subtle but secondary to our focus here. After looking at many models, we believe the values presented are representative for the purpose of constraining regions of parameter space that may permit SI to form planetesimals over a wide range of conditions (see Section 7.7). An additional description of the

---

Table 6

| \( \alpha \times 10^{-4} \) | Inner Disk (\( R = 3 \) au) | Snow Line (\( R = 5 \) au) | Outer Disk (\( R = 30 \) au) |
|---|---|---|---|
| | \( \tau_s \) | \( \rho_s (g \ cm^{-3}) \) | \( \delta \) | \( \tau_s \) | \( \rho_s (g \ cm^{-3}) \) | \( \delta \) | \( \tau_s \) | \( \rho_s (g \ cm^{-3}) \) | \( \delta \) |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 100^a | 283 | 2.4 \times 10^{-11} | 0.07 | 156 | 1.2 \times 10^{-11} | 0.07 | 73 | 3.2 \times 10^{-13} | 0.11 |
| 40^d | 290 | 3.7 \times 10^{-11} | 0.09 | 148 | 2.3 \times 10^{-11} | 0.08 | 71 | 3.6 \times 10^{-13} | 0.12 |
| 10^g | 275 | 1.2 \times 10^{-10} | 0.08 | 130 | 4.4 \times 10^{-11} | 0.07 | 75 | 8.6 \times 10^{-13} | 0.11 |
| 4^d | 265 | 3.0 \times 10^{-10} | 0.08 | 140 | 7.0 \times 10^{-11} | 0.08 | 75 | 4.8 \times 10^{-13} | 0.11 |
| 1^g | 225 | 8.6 \times 10^{-10} | 0.06 | 152 | 2.2 \times 10^{-10} | 0.06 | 70 | 9.1 \times 10^{-13} | 0.11 |
| 0.1^d | 204 | 6.7 \times 10^{-10} | 0.06 | 148 | 2.7 \times 10^{-10} | 0.06 | 73 | 6.5 \times 10^{-13} | 0.11 |

Notes.

^a Corresponds to the size that dominates the mass in the particle size distribution.

^b Approximate radial locations. For this exercise, the models were evaluated for local values about \( Z = 0.01 \pm 0.001 \) at a mean time of \( \sim 2 \times 10^7 \) yr, where \( 3 \) au corresponds to radial distances where \( 200 \lesssim R \lesssim 290 \) K, the snow line to \( 130 \lesssim R \lesssim 160 \) K, and \( 30 \) au to \( 70 \lesssim R \lesssim 75 \) K, respectively.

^c Data shown with yellow symbols in Figures 11–13 are from P. R. Estrada et al. (2020, in preparation).

^d Values from MTBF models of Estrada et al. (2016) and shown with black symbols in Figures 11–13.

^e These particles are in the Stokes regime with particle Reynolds numbers near unity.

^f Using Musiolik & Wurm (2019) cold ice fragmentation findings for \( U_s \) and shown with orange symbols in Figures 11–13.
colored-symbol models will be given in P. R. Estrada et al. (2020, in preparation). Even with their uncertainties, the message of the realistic models is that plausible, self-consistent combinations of nebula turbulence and particle size typically lie in “Zone II,” the “moderately turbulent” regime above the \( \epsilon = 1 \) line, where SI is only “incipient.” Moreover, any degree of particle porosity leads to even smaller \( \tau_s \) (P. R. Estrada et al. 2020, in preparation).

### 7.6. Turbulence Constraints

It is widely believed that the region where the first planetesimals were assembled was a “DZ” extending from 1 au to \( \lesssim 80 \) au, so-called because the temperature and transparency of protoplanetary disk material to ionizing photons are too low to allow magnetically driven (MHD) turbulence to arise (Turner et al. 2014). These areas of protoplanetary disks are now more commonly referred to as ohmic zones (e.g., Lyra & Umurhan 2019) owing to the dominance of ohmic dissipation that suppresses self-generated activity like the linear MRI process. There remains substantial debate over the true ionization state and consequent effective magnetic resistivity of disk material, depending on the abundance of grains of sufficiently small size (Okuzumi et al. 2012; Ormel & Okuzumi 2013; Simon et al. 2018b). Even if insufficiently ionized for MRI, the disk remains susceptible to nonideal MHD processes, including wind-launching...
mechanisms (Bai 2016; Bai et al. 2016). Nonetheless, this debate leaves some uncertainty as to whether any regions may be susceptible to the MRI, which numerical experiments show to induce very strong turbulent intensities of $\alpha \sim 10^{-3} - 10^{-2}$ (Balbus & Hawley 1998; Armitage 2011).

Therefore, for the purposes of this study, we shall assume that any turbulence that arises in the near-midplane regions of protoplanetary disks, which are of greatest interest to planetesimal formation, stems from any one of the three purely hydrodynamical mechanisms noted in Section 1 that have recently been discussed in the literature—for a review, see Turner et al. (2014) or Lyra & Umurhan (2019). The operation of the three mechanisms depends upon the thermal relaxation or cooling timescale $\tau_{\text{th}}$ of the particle–gas mixture, which in turn depends on the disturbance length scales if these length scales are in the optically thick regime. Specifically, the [VSI/COV/ZVI] (Section 1) is expected to operate in a disk when $[\Omega_{\text{th}} \ll 1, \Omega_{\text{th}} \approx 1, \Omega_{\text{th}} \gg 1]$ (respectively). Exactly which regions of protoplanetary disks are most susceptible to which mechanism remain under discussion (Malygin et al. 2017; Umurhan et al. 2017; Barranco et al. 2018). However, because of their complementary instability criteria, it is plausible that the full extent of protoplanetary disks that is of interest to planetesimal formation will be turbulent, due to at least one of the three mechanisms listed. Recently published numerical studies of the three processes show that the turbulent intensities arising from these mechanisms lie somewhere in the range $10^{-5} < \alpha < 5 \times 10^{-4}$ (Lyra & Umurhan 2019). Of the various published studies of the VSI, Flock et al. (2017) predict the lowest level of $\alpha \sim 4 \times 10^{-5}$, and we use this as our minimum adopted value. Our understanding of the nature of cold protoplanetary disk turbulence will surely continue to evolve.

Caveat: we have assumed that the mechanism driving turbulence in disks is unaffected by the degree of particle loading. This is valid a posteriori in regions covered by all of the detailed particle growth models (Figures 11–13), which imply $\epsilon < 1$, and limited or negligible particle feedback. This assumption should be regarded with caution in parameter combinations where $\epsilon > 1$, but these combinations (corresponding to Zone I) are already known to lead to SI in numerical experiments.

7.7. Regions of Parameter Space that Permit the SI to Operate for Z = 0.01

Given the physical constraints outlined in the previous six subsections, we can now assess in what parts of the $\alpha$–$\tau_s$ parameter space the SI is likely to operate. We examine this for three locations: $R = 3$ au (“inner disk”), $R = 5$ au (“snow line”), and $R = 30$ au (“outer disk”) in a $Z = 0.01$ disk with constant $\delta = 0.05$. By implication, we imagine the following survey to be predictive for an early phase of the solar nebula with uniform $Z = 0.01$ and well before wind loss substantially evaporates the disk’s gas (say, < 2.5 Ma; Carrera et al. 2017). It is increasingly thought that planetesimal formation had to proceed in the inner nebula before the 0.5 Ma time marker (relative to calcium–aluminum-rich inclusions), and that even Jupiter’s core may have formed in less than 1 Ma (e.g., Kruijer et al. 2017; Simon et al. 2018a). For higher, perhaps local, values of $Z$, to which particle size is insensitive, the reader is referred to Figure 4 for comparison of how Zones I and II shift.

We approach this task by first comparing the SI’s predicted growth timescales with constraining timescales based on the various disk processes and particle growth barriers discussed in

---

**Figure 10.** Like Figure 8, except at $R = 30$ au, nominally the outer disk. Note, only the 2.5 Ma line (white) is shown. The SI incipient regime is shown with the red-hatched triangle and corresponds to water ice being stickier beyond the snow line (and thus bounded on the right by the sticky H$_2$O ice fragmentation line). The incipient region would be significantly decreased if one adopts the ice cold H$_2$O ice fragmentation line, which may be the case away from the snow line in the colder regions of the disk (Musiolik & Wurm 2019).
Sections 7.1–7.6 for our nominal disk model with $\delta = 0.05$. We then delineate regimes of $(\alpha, \tau_s)$ space deemed implausible by the best current models of growth-by-sticking (Estrada et al. 2016, P. R. Estrada et al. 2020, in preparation, Sections 7.4 and 7.5) over the first 0.5 Ma, which are typically characterized by larger $\delta \gtrsim 0.05$–0.1. We declare the SI to be “incipient,” or capable of leading to some degree of enhancement, in those regions of $(\alpha, \tau_s)$ space which are both reasonable from the standpoint of particle growth and where the SI’s growth timescales are shorter than the aforementioned constraining timescales.

Nominal fragmentation constraints. The results for our nominal models are shown in Figures 8–10, and those with the added constraints of particle growth models are shown in Figures 11–13. All of these plots overlay various excluded regions on the growth timescale plot similar to Figure 3 for the three disk zones of interest. The timescales in Figures 11–13, however, were recalculated using the higher $\delta$ values associated

Figure 11. Regions of $\alpha$–$\tau_s$ parameter space accessible to the SI based on the same input disk model values as shown in Figure 3 (except with $\delta = 0.07$), $Z = 0.01$, and at the nominal location of the asteroid belt, $R = 3$ au (inside the snow line for these models). A consequence of higher $\delta$ values is that the radial and mode pattern drift lines shift downwards relative to our nominal $\delta = 0.05$ case, further limiting the SI incipient region. The indicated (upper) limits on particle size (yellow and black stars) are based on the predicted mass-bearing Stokes numbers for evolutionary models of Estrada et al. (2016) and P. R. Estrada et al. (2020, in preparation) as listed in Table 6. The yellow patch demarcated by the dashed line identifies the parameter region ruled out based on the interpolation of the yellow stars. The orange star is for a model in which water ice is only sticky near the snow line (last row of Table 6). The latter model’s different $\tau_s$ value is attributable to less material drifting in to the inner disk relative to the sticky-ice model where particles can grow larger. Note also that these evolution models do not examine values of $\alpha$ beyond the vertical extent of the yellow shading. There is no SI incipient region accessible in this case.

Figure 12. Like Figure 11, except at $R = 5$ au, nominally the snow line. The orange star is for a model in which water ice has the low stickiness appropriate for cold temperatures (last row of Table 6) and thus leads to smaller particles. We note that a tiny SI incipient regime survives, indicated by the red-hatched triangle.
with the particle growth models. If a zone of incipient SI is allowed, it is indicated by a red-hatched triangle in these figures.

Figure 8 shows the situation for the inner disk where the temperature is too large to support water-ice particles. We find that the SI is incipient in a relatively small triangular patch of parameter space centered at about a value of \( \alpha \approx 5 \times 10^{-5} \) and \( \tau_s \approx 0.01-0.02 \). The region of accessible parameter space is bounded on the small side of \( \tau_s \) lines showing conditions where the growing SI pattern drifts into the star faster than it can achieve an e-folding level of growth (see discussion in Section 7.2). The region is also bounded on the higher side of \( \tau_s \) by the constraint imposed by the silicate fragmentation line. Within this SI incipient region, the growth timescales are relatively long—thousands of orbit times.

The situation at the snow line can vary significantly depending on whether one adopts the sticky or cold H\(_2\)O fragmentation condition, as can be seen in Figure 9. Taking the sticky-ice constraint opens up a much larger region for incipient growth ranging from \( \tau_s \approx 0.01-1 \), and for turbulent intensities as high as \( \alpha \approx 5 \times 10^{-5} \) that extends into Zone III, and even Zone I for the largest \( \tau_s \) (see Figure 3), and growth timescales are as little as 10s of orbits. However, the sticky-ice condition may only apply to regions that are close to the water-ice evaporation temperature (Musiolik & Wurm 2019). Adopting the cold water-ice fragmentation condition returns the situation to the one seen in Figure 8 with only a tiny incipient zone. Likewise at 30 au, the region of SI incipient growth depends strongly on how sticky water ice is. Sticky water ice, as has been adopted in most growth models to date, leads to an even larger incipient region where turbulence can be as high as \( \alpha \approx 10^{-3} \) and encompasses a range from \( \tau_s \approx 0.008-4 \), again extending as far as Zones III and I. On the other hand, the cold water-ice fragmentation condition (Musiolik & Wurm 2019) restricts the incipient region to a tiny region in Zone II, centered about \( \alpha \approx 7 \times 10^{-5} \) and bounded by \( \tau_s \approx 0.008-0.1 \).

Modeled particle growth constraints. We now examine how the constraints imposed by the evolutionary growth model results of Estrada et al. (2016) and P. R. Estrada et al. (2020, in preparation) in the epoch of interest affect the regions of SI incipient growth. In Figure 11, we return to the inner disk, but now plot the maximum achieved Stokes numbers from Table 6 (colored stars) which correspond to the particle sizes that carry most of the mass (see Section 7.5). Note that Figure 11 contains new SI mode growth rates, recalculated for \( \delta = 0.07 \) which represents a characteristic value for the particle growth and drift models. The main effect of this is that the radial and pattern drift constraint lines have shifted downwards, with the latter further restricting the SI incipient region to a vanishingly small range even before particle constraints are folded in. However, when one considers the growth models—even for the minimum \( \alpha \) case—this region becomes inaccessible. As was discussed in Section 7.5, not all the plotted points fall on the representative silicate fragmentation line which assumes that turbulence dominates their relative velocities and that eddy-crossing effects are unimportant. The points defined by \( \alpha \gtrsim 10^{-5} \) broadly lie above their respective fragmentation lines (as can be determined from the temperatures listed in Table 6) because there has been significant growth beyond the fragmentation barrier (Section 7.5) due to both “mass transfer” (e.g., Wurm et al. 2005; Windmark et al. 2012) and “lucky particle” growth in those models (e.g., Garaud et al. 2013; Drążkowska et al. 2014), whereas for \( \alpha \lesssim 10^{-4} \), the mass-dominant particles are entering a low-turbulence regime in which they are subject to eddy-crossing effects so that their fragmentation Stokes numbers are smaller than what Equation (26) would predict. Thus, all the particle evolution models discussed here are in fact in the fragmentation regime. Overall then, we find that in the silicate-dominated region illustrated by Figure 11, all plausible, self-consistent combinations of turbulent intensity and particle size lie well within Zone II, and do not overlap any of the region in \( \alpha-\tau_s \) parameter space in which SI incipient growth is permissible.
For both the nominal snow line and the outer disk, Figures 12 and 13 show that the situation is also strongly limited by realistic growth models, with respect to regions of incipient growth. As was the case in Figure 11, Figure 12 corresponds to a value of $\delta = 0.07$, which represents a characteristic value across all models listed in Table 6 for the snow line. Like before, the higher $\delta$ further restricts the incipient growth region, though not as dramatically as for the inner disk. In this case, we see that the SI appears to be incipient only within a tiny triangular patch of parameter space, again around $\alpha \sim 5 \times 10^{-5}$ and $\tau_s \approx 0.01$, indicated in the figure by the red-hatched triangular boundary. The constraining processes of the minimum level of turbulence, and the mode pattern drift timescales, bound the region from the bottom and left. However, the evolutionary growth models between $\alpha \sim 10^{-5} - 10^{-4}$ cut off access to the regions of $\alpha - \tau_s$ parameter space that would be permissible from either water-ice fragmentation constraint. Figure 13 shows the case for the outer nebula at 30 au, in which $\delta = 0.1$. Here, the mode pattern drift line moves even farther downward, increasing the lower bound of $\tau_s \approx 0.04$ by a factor of 5 more than for the $\delta = 0.05$ case (Figure 10) and shrinking the SI incipient zone. Meanwhile, particle Stokes numbers remain at about $\tau_s \approx 0.01$, farther to the left and eliminating the small incipient zone.

In Figure 12, almost all models with $\alpha \leq 4 \times 10^{-3}$ are in the drift-dominated regime, whereas in Figure 13 the fragmentation limit has only been reached for $\alpha = 10^{-2}$ (recall that the fragmentation limit for the black and yellow symbols are more closely associated with the sticky H$_2$O ice fragmentation line). In Figure 12, the models for $\alpha = 10^{-4} - 10^{-3}$ are again in a regime where eddy crossing is starting to become important, so their fragmentation $\tau_s$ will be smaller than what the sticky-ice line predicts, but still larger than their current values. Here, only the model for $\alpha = 10^{-5}$ is in the Stokes flow regime. The remaining models in Figures 12 and 13 (orange stars), which are associated with the cold H$_2$O ice model, have reached the fragmentation size (though significant growth beyond it has occurred in Figure 12 at the snow line). The curious inflection that leads to peak values for $\tau_s$ for $\alpha \sim 10^{-3}$ are real and due in part to enhanced growth about the snow line, whereas the trend toward larger $\tau_s$ seen for $\alpha \sim 10^{-2}$ at 30 au is due to both reaching the fragmentation barrier and the lower gas surface density as a result of the more rapid viscous evolution compared to the other models. Overall though, as was the case for the silicate-dominated inner disk region, all self-consistent combinations of turbulent intensity and particle size lie within Zone II, from the standpoint of SI.

The interplay between particle drift, bouncing, and fragmentation, and the manner in which particle growth proceeds, especially around evaporation fronts like the snow line, highlights the complexity of modeling particle growth and gas evolution with time and call for more in-depth analyses of this type in the context of the theory presented herein.

8. Summary and Conclusions

Conducting high-resolution numerical simulations of the interaction of gas and particles to shed light upon the streaming instability is a computationally expensive undertaking. It should be of value to have some kind of theoretical guide—however approximate—to constrain the parameter range of its validity from the standpoint of planetesimal formation. The purpose of this study is to provide a theoretical framework to address the question of how and to what extent the streaming instability might be effective for planetesimal formation under globally turbulent disk conditions.

Our model extends and generalizes behavior initially studied by YG2005 by representing the effects of local turbulence by an $\alpha$ model and makes additional predictions that are consistent with previously reported numerical studies of the presence or absence of the streaming instability in turbulent disk simulations which include the following: Johansen et al. (2007), Balsara et al. (2009), Yang et al. (2017, 2018), Li et al. (2018), and Gerbig et al. (2020). The $\alpha$-model representation of turbulence—which characterizes its effect in the form of an enhanced isotropic turbulent viscosity and diffusion—acts locally both to stir particles and to exchange momentum. Underpinning its use here is the assumption that the processes leading to turbulence, especially in protoplanetary disk ohmic zones, do so independently of the presence of particles with the realistic sizes and abundances treated here.

We have examined the normal mode response of the streaming instability as a function of the disk turbulence parameter $\alpha$ and particle Stokes numbers $\tau_s$, by identifying the wavelengths, growth times, and pattern speeds of the fastest growing modes. For given values of $\alpha$ and $\tau_s$, the particle to gas mass density ratio ($\epsilon = \rho_p / \rho_g$) is calculated using a TDM (Equation (21)) to represent the balance between the gravitational settling of particles toward the disk midplane and the vertical diffusion of the same particles due to turbulence (Dubrulle et al. 1995; Youdin & Lithwick 2007; Estrada et al. 2016). We hope the theoretical framework proposed and examined in this study is useful in similar future studies.

Simple turbulence models like ours involve taking higher order moments of the equations of motion which are then truncated with some closure relationship. In general, such resulting model equations do not conserve momentum (see extended discussion in Davidson 2004). Tominaga et al. (2019) highlighted the lack of angular momentum conservation in a 2D non-axisymmetric disk setting with a turbulence model similar to ours and proposed a solution to the problem. Together with the authors of Chen & Lin (2020), we have conducted a proper reanalysis of the matter in the framework of these equations (not shown here) and find that this nonconservation has a negligible effect on the growth rates determined for all values of $\alpha$ pertinent for realistic protoplanetary disks. A future follow-up study explicitly detailing this result is warranted.

The study conducted here has revealed several interesting trends for a nominal (minimum mass) solar nebular model with global solids-to-gas mass ratio of $Z = 0.01$ and disk-opening angle $\delta = H/r = 0.05$ (results are also given for values of $Z$ as large as 0.08). While the specific values summarized below primarily pertain to disk models with $Z = 0.01$, the conclusions might be extended to higher $Z$ by comparison with Figure 4 (although further analysis should be done to verify this assertion as well):

1. As turbulent intensity increases, the wavelengths of the maximally growing SI modes increase, while the growth rates of the maximally growing SI modes diminish.

For the one set of simulations we examined where the correspondence was weakest (i.e., Li et al. 2018), we conjecture that those simulations were not run with sufficient resolution to see the short wavelengths predicted by our model.
2. The combination of \((\tau_s, \alpha)\) that leads to initial \(\epsilon = 1\) according to the TDM traces a critical curve with important implications for SI behavior. This curve terminates at a critical point \((\tau_s, \alpha_c)\) corresponding to values of \(\epsilon\) that monotonically increase with \(Z\) for a given \(\delta\). For cosmic abundance \(Z = 0.01\), this critical point occurs at \(\tau_s = 0.45\) and \(\alpha_c \approx 3.7 \times 10^{-3}\), corresponding to \(\epsilon = \epsilon_c = 1.1\). Selected critical point values for other \(Z\) are summarized in Table 1.

3. For values of \(\tau_s < \tau_c\), and for parameter combinations of \(\alpha\) and \(\tau_s\) that lead to \(\epsilon = 1\) according to the TDM (Equation (21))—that is, along the critical \(\epsilon = 1\) curve—the least stable SI mode neither grows nor decays, and the growth timescale is effectively infinite.

4. Provided \(\tau_s < \tau_c\), we identify two regimes that straddle the above-mentioned critical line as being either “laminar/unstable” (Zone I) or “turbulent/saturated” (Zone II; Figure 3).

5. The spatial structure of the fastest growing mode in Zone II (the turbulent/saturated regime) typically corresponds to vertically oriented sheets with radial scale of about a pressure scale height \(H\). In practice, the sheet’s vertical extent should follow the particle scale height, and this appears to be consistent with the simulation results reported in Balsara et al. (2009); see their Figure 7. The mode structure in Zone I (the laminar/unstable regime) exhibits narrow, azimuthally oriented tubes with length scales much less than \(H\), consistent with previous predictions and simulations (e.g., see the recent results of Carrera et al. 2015; Yang et al. 2017; Li et al. 2018).

6. The theory developed here appears to reasonably predict the onset or absence of the SI in several recently published simulations. We have plotted these correspondences in our \(\alpha\)—St growth timescale plots in Figures 3 and 4. For laminar disk models in which particles settling to the midplane generate their own midplane turbulence from either the Kelvin–Helmholtz overturn and/or the SI itself, we find that our theory (i) reasonably predicts the radial spacing of emerging filaments when SI is active, especially during its early onset phases well before filament–filament merging occurs and (ii) predicts when the SI does not appear. We have come to this conclusion after careful analysis of the three recent studies by Yang et al. (2017), Li et al. (2018), and Gerbig et al. (2020).

In simulations where turbulence is externally driven throughout the disk model by the MRI like Johansen et al. (2007), Balsara et al. (2009), and Yang et al. (2018), the theory does a reasonable job at predicting whether or not SI should or should not be present and, in some cases, predicts the radial wavelength structure and growth timescale of filaments like in Johansen et al. (2007). In particular, the theory predicts the absence of the SI in the simulations of Yang et al. (2018) for their ideal MHD model in which the MRI is rampant throughout the computational domain. On the other hand, Yang et al. (2018) also present results of SZ disk models driven by waves coming from sandwiching MRI-active layers. In the absence of backreaction, particles accumulate along filaments likely because the wave forcing engenders azimuthally oriented pressure fluctuations with coherent pressure maxima toward which particles naturally drift. When backreaction is turned on, the particle densities within these filaments become enhanced. Our theory does not explicitly handle SI physics in the presence of coherent structures: it implicitly assumes the scales of the system take place well inside the inertial range of the turbulent forcing where coherence is lost. This matter must be addressed in future work.

7. We believe that realistic protoplanetary disk conditions relevant to the early solar system \((\leq 0.5 \text{ Ma})\) in which \(Z = 0.01\) and where typically \(\delta \geq 0.05–0.1\)—taking into account various barriers to particle growth, age constraints, and the disk’s likely degree of turbulence as quantified by \(\alpha\) restrict the SI to Zone II where it is “incipient,” or allowed, for a narrow range of self-consistent disk and particle properties. For example, we find at a disk location nominally representative of the ice line of Jupiter’s early core \((r = 5 \text{ au})\) that the incipient range of parameter space as indicated by the red-hatched triangle in Figure 12 is restricted about a turbulent intensity \(\alpha \sim 4 \times 10^{-5}\) and Stokes numbers in a narrow range between \(\sim 0.01 < \tau_s < 0.02\). We treat this vanishingly small allowable region as only suggestive—given the uncertainties in all the constraints, it may be larger. On the other hand, no such permissible regions are found in the inner and outer disk cases for the given models. Thus, achieving growth sufficient to breach the incipient regions in the inner and outer disks appears even much more challenging.

8. In Zone I, SI is robust and plausibly proceeds to planetesimal formation, as routinely observed in numerical simulations. In Zone II, we identify a new kind of behavior in which SI is only “incipient,” the growth timescales are very long \((10^2–10^4 \text{ local orbit times})\), and the growth of particle density is limited.

9. It should be kept in mind that while the analysis in Section 7.5 suggests that it may be difficult for the SI to trigger sufficient particle enhancements in turbulent disks with spatially uniform values of \(Z = 0.01\) and that this prediction is also likely to be applicable only for the earliest phases of a model solar disk in which there is negligible particle enrichment. As a disk evolves, the disk’s gas content will evolve both globally due to wholesale loss via winds and local variations induced by transport. Over time, then, regions will emerge with substantially enhanced values of \(Z\). The regions that are “SI incipient” can, in principle, slip from being in Zone II to Zone I, wherein the growth is expected to be more rapid even in the face of turbulence. A counterargument here might be that decreasing the gas density through winds might not necessarily lead to higher values of St nor enhanced local values of \(Z\) because the efficiency of radial drift can efficiently deplete a local region of its particles while particles might encounter their fragmentation barriers at smaller sizes (Birnstiel et al. 2012; Estrada et al. 2016; Carrera et al. 2017). We therefore expect the situation to be very different for conditions corresponding to the latter stages of the disk’s evolution beyond its thick gas phase, and we caution against applying our prediction of limited to no growth to the whole of the disk’s lifetime. Further follow-up analyses are therefore necessary and warranted.
Other future work. There is much to understand about the implications of this new physics. Included in any list of future work should be (a) a better physical understanding of the critical line and critical point corresponding to the special condition $\epsilon = 1$ and, especially, why this special combination of parameters leads to exactly marginal modes, (b) the role of a particle size distribution, (c) the possible role of particle loading on damping turbulence, (d) a self-consistent analytical model of particle layers and self-generated turbulence in the limit of vanishingly small global turbulence, (e) the behavior when the predicted vertical wavelength exceeds the thickness of the particle layer, and of course, (f) the regime where purely hydrodynamical turbulence is operative, and its intensity. It is worthwhile to redo this analysis within the framework of (i) a turbulent MHD model and/or (ii) a model with multiple particle sizes as in Krapp et al. (2019), and (iii) to expand the theory to account for the presence of coherent flow structures, e.g., like those present in Yang et al. (2018).

Finally, understanding the physical mechanism of the turbulent SI remains a priority. While the mathematical description of the onset of the inviscid SI in terms of resonant drag energy exchange between gas waves and particles is satisfying (i.e., the RDI, Squire & Hopkins 2018a), it might need modification to explain the onset of instability in this kind of turbulent model setting. The mechanistic explanation of the inviscid SI presented by Jacquet et al. (2011) offers a framework upon which an explanation for the viscous case might be built.

We thank our two reviewers who provided very thorough readings and suggestions that improved both the rigor of our analysis and the quality of this manuscript. We are indebted to Allan Treiman for organizing “Accretion: Building New Worlds Conference” held at the Lunar and Planetary Institute (2017 August 15–18), where the motivation for this study first germinated via in-depth conversations with A. Carballido, D. Sengupta and others at NASA Ames Research Center’s Origins Group as well as input and perspectives shared with us especially by M-K. Lin and K. Chen who developed very similar results independently of—but in parallel with—us, as well as X. Bai, H. Klahr, W. Lyra, J. Drazkowska. O.M.U. acknowledges support from NASA Astrophysics Theory Program grant NNX17AK59G and The New Horizons Kuiper Belt Extended Mission, contract NASA NASW02008. P.R.E. and J.N.C. acknowledge NASA Emerging Worlds grant NNX17AL60A for providing resources and support for this project. Finally, we also acknowledge NASA’s “Internal Scientist Funding Program” grant to Ames Research Center on Origins of Planetary Systems as well as NASA’s Senior NPP program.

Appendix A

Linearized Matrix Equation

The linearized equations of motion are nondimensionalized on $\Omega$ and $H$. Furthermore, in order to ensure incompressibility, the radial and vertical components of the perturbation gas velocity (respectively, $u'_r, w'_z$) is expressed in terms of the azimuthal perturbation stream function $\psi'$, which is related to the gas vorticity via $\omega'_z = -K^2 \psi'$ to insure incompressibility. The result may be cast into the following matrix form:

$$
-i\omega \begin{pmatrix} u'_{r} \\ v'_{r} \\ v'_{\theta} \\ w'_{r} \\ w'_{\theta} \\ \Delta'_p \end{pmatrix} + M \begin{pmatrix} u'_{r} \\ v'_{r} \\ v'_{\theta} \\ w'_{r} \\ w'_{\theta} \\ \Delta'_p \end{pmatrix} = 0,
$$

(A1)

where $M$ is given by

\[
M \equiv \begin{pmatrix}
 a & \frac{2ik_c}{K^2} & \frac{ik_c \epsilon}{K^2\tau_s} & 0 \\
 \frac{ik_c}{2} & a & 0 & -\frac{\epsilon}{\tau_s} \\
 -\frac{ik_c}{\tau_s} & 0 & \frac{1}{\tau_s} & \frac{2ik_c \tau_s}{(\epsilon + 1)^2 + \tau_s^2} \\
 0 & -\frac{1}{\tau_s} & \frac{1}{2} & \frac{1}{\tau_s} - \frac{2ik_c \tau_s}{(\epsilon + 1)^2 + \tau_s^2} \\
 ik_c & 0 & 0 & 0 \\
 0 & 0 & ik_c & 0 \\
 \end{pmatrix}
\]

\[
-\frac{ik_c \epsilon}{K^2\tau_s} & -\frac{2ik_c \epsilon (\epsilon + 1)}{K^2((\epsilon + 1)^2 + \tau_s^2)} & 0 & \frac{\epsilon \tau_s}{(\epsilon + 1)^2 + \tau_s^2} & 0 \\
 0 & 0 & ik_c \epsilon & 0 & \frac{ik_c \epsilon}{(\epsilon + 1)^2 + \tau_s^2} \\
 0 & 0 & 0 & 0 & \frac{K^2 \alpha / \delta^2}{\tau_s + 1} - \frac{2ik_c \tau_s}{(\epsilon + 1)^2 + \tau_s^2} \\
\]

(A2)

where $K^2 \equiv k_r^2 + k_z^2$, and for notational convenience we define

$$
a \equiv \alpha K^2 / \delta^2 + \frac{\epsilon}{\tau_s} + \frac{2ik_c \epsilon \tau_s}{(\epsilon + 1)^2 + \tau_s^2}.
$$

J. Simon and D. Carrera. We thank Phil Armitage for organizing the “Theoretical and Computational Challenges in Planet Formation Workshop” at the Flatiron Institute (2019 May 20–22) where mature results of this study were presented. We are grateful for extended conversations with K. Shariff and M. K. Lin.
We could approach solutions to this by solving for $\omega$ from the sixth-order dispersion relation arising from

$$\det (-i \omega \mathbb{I} + M) = 0,$$  \hspace{1cm} (A3)

where $\mathbb{I}$ is the $6 \times 6$ identity matrix. However, the expression resulting from the above operation is extremely unwieldy and offers no insight toward the mechanisms operating in the instability. Instead, we choose to solve this directly by determining the eigenvalues of $M$ using standard numerical techniques found in MATLAB. We arrange the solutions in descending order of $\omega$. The solutions plotted throughout the text are the least stable mode. In some instances, there are two unstable modes but the second mode is usually dwarfed in magnitude by the first mode. A detailed examination of the second unstable mode and its interpretation remains to be done, but may be of interest. We are reminded that in the inviscid theory that the characteristic length scales of the SI are $H \delta$ while length scales in the $\alpha$-disk model are measured on $H$. Hence, a disparity of length scales between the large-scale global turbulence and the SI appears with the ratio $\alpha^{1/2} / \delta$.

Appendix B
On the Dead Zone Models of Yang et al. (2018)

By contrast, the DZ model results (especially $Z = 0.04, 0.08$) exhibit particle accumulation along azimuthally elongated filaments that grow tighter, become increasingly coherent, and achieve high densities, plausibly due to SI. However, we think that the physics in this case is more complicated than our turbulence model can address. Yang et al. (2018) depicted the particle behavior in DZ models with and without backreaction (see bottom two rows of Figure 10 of Yang et al. 2018). Even in the models without backreaction (where SI is inactive), particles can be seen to accumulate into azimuthally elongated filamentary structures almost from the very beginning of the model runs. In these models, the value of $\epsilon$ in the densest regions appears to exceed unity in many places. In comparison, all DZ models with backreaction show the same early accumulation as in the non-backreacting case. However, with backreaction, these density enhancements continue to evolve because the SI drives further density increases.

Yang et al. (2018) characterize the vertical distribution of particles in terms of an effective particle diffusion parameterized by an analogous $\alpha$ parameter called $\alpha_{\text{g,z}}$. This diffusion parameter achieves high values ($\alpha_{\text{g,z}} \sim 4 \times 10^{-3}$) similar to the values of $\alpha$ in the iMHD case, even though the value of $\alpha$ in the midplane regions of the DZ model is quite low ($\sim 2 \times 10^{-3}$). Nonetheless, particles appear to concentrate into filaments, which become reinforced and further enhanced as the SI takes root; this is especially apparent in the simulations involving $Z = 0.02, 0.04, 0.08$. Such filaments are likely pressure extrema with concomitant zonal flows that are known to be characteristic of MRI-driven turbulence at the large scales of a simulation (Johansen et al. 2009; Simon et al. 2012).

These scales are, at best, at the top of the inertial range of the turbulent flow and are not contained inside the turbulence’s inertial range. To what extent the midplane layers in their DZ models may or may not be characterized as a uniform turbulent flow (see Section 4.1 of Yang et al. 2018) and what effect external wave forcing has on inducing a priori high-density azimuthally aligned filaments remain to be fully examined. By examining the development of filaments in models with backreaction against those without backreaction, it can be seen in the latter that filamentary particle enhancements emerge as a matter of course in response to the large-scale wave forcing coming from the magnetically active layers (see also Figure 8 of Yang et al. 2018).21 With backreaction turned on, these particle enhancements readily continue their condensation into narrow coherent structures. It is difficult to disentangle the causal sequence of events: are the observed structures growing solely attributable and intrinsic to the SI, or does the preexistence of overdense filaments aid in triggering the SI in these cases? It would seem that the large-scale azimuthally elongated structural forcing imprinted by the waves emanating from the overlying turbulent layers predisposes the midplane particles into undergoing the SI because of the a priori axisymmetric clumping caused by the radial pressure maxima induced by the waves. The coherent large-scale structure of the nonuniform wave forcing is a physical complexity not reflected in the simple turbulence model of our theory, nor of that of Chen & Lin (2020). We therefore view the physical structures observed in the midplane regions of their DZ model as outside the scope of the theoretical construct discussed here and could be addressed in future work. Finally, the relevance of models with such deep MRI-active layers is itself open to debate (e.g., Bai & Stone 2010).

References
Armitage, P. J. 2011, ARA&A, 49, 195
Bai, X.-N. 2016, ApJ, 821, 80
Bai, X.-N., & Stone, J. M. 2010, ApJ, 722, 1437
Bai, X.-N., Ye, J., Goodman, J., & Yuan, F. 2016, ApJ, 818, 152
Balbus, S. A., & Hawley, J. F. 1998, RMP, 70, 1
Balsara, D. S., Tilley, D. A., Reettig, T., & Brittan, S. D. 2009, MNRAS, 397, 24
Barranco, J.-A. 2009, ApJ, 691, 907
Barranco, J. A., Pei, S., & Marcus, P. S. 2018, ApJ, 869, 127
Benítez-Llambay, P., Krapp, L., & Pessah, M. E. 2019, ApJS, 241, 25
Bers, A. 2016, Plasma Physics and Fusion Plasma Electrodynamics (Oxford: Oxford Univ. Press), 2528
Birnstiel, T., Andrews, S. M., & Ercolano, B. 2012, A&A, 544, A79
Brauer, F., Henning, T., & Dullemond, C. P. 2008, A&A, 487, LI
Bridges, F. G., Sepulver, K. D., Lin, D. N. C., et al. 1996, Icar, 123, 422
Carballido, A., Bai, X.-N., & Cuzzi, J. N. 2011, MNRAS, 415, 93
Carballido, A., Fromang, S., & Papaloizou, J. 2006, MNRAS, 373, 1633
Carrera, D., Gorti, U., Johansen, A., et al. 2017, ApJ, 839, 16
Carrera, D., Johansen, A., & Davies, M. B. 2015, A&A, 579, A43
Chandrasekhar, S. 1961, International Series of Monographs on Physics (Oxford: Clarendon) 1961
Chen, K., & Lin, M.-K. 2020, ApJ, 891, 132
Cuzzi, J. N., Dobrovolskis, A. R., & Champney, I. M. 1993, Icar, 106, 102
Cuzzi, J. N., Hogan, R. C., & Shariff, K. 2008, ApJ, 687, 1432
Davidson, P. A. 2004, Turbulence: An Introduction for Scientists and Engineers (Oxford: Oxford Univ Press)
Dipierro, G., Laibe, G., Alexander, R., et al. 2018, MNRAS, 479, 4187
Dobrovolskis, A. R., Dacles-Mariani, J. S., & Cuzzi, J. N. 1999, IGR, 104, 30805
Drazin, P. G., & Reid, W. H. 2020, in Plasma Physics and Fusion Plasma Electrodynamics, ed. P. G. Drazin & W. H. Reid (Cambridge: Cambridge Univ. Press), 626
Drazkowska, J., Windmark, F., & Dullemond, C. P. 2013, A&A, 556, A37
Drazkowska, J., Windmark, F., & Dullemond, C. P. 2014, A&A, 567, A38
Dubrulle, B., Morfill, G., & Sterzik, M. 1995, Icar, 114, 237
Ercolano, B., Jennings, J., Rosotti, G., et al. 2017, MNRAS, 472, 4117

21 Similar structuring can be seen in the simulations of Johansen et al. (2007).
