

# A simple phenomenological model for grain clustering in turbulence

Philip F. Hopkins<sup>1,2★</sup>

<sup>1</sup>TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA

<sup>2</sup>Department of Astronomy and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA

Accepted 2015 September 23. Received 2015 September 18; in original form 2013 July 16

## ABSTRACT

We propose a simple model for density fluctuations of aerodynamic grains, embedded in a turbulent, gravitating gas disc. The model combines a calculation for the behaviour of a group of grains encountering a single turbulent eddy, with a hierarchical approximation of the eddy statistics. This makes analytic predictions for a range of quantities including: distributions of grain densities, power spectra and correlation functions of fluctuations, and maximum grain densities reached. We predict how these scale as a function of grain drag time  $t_s$ , spatial scale, grain-to-gas mass ratio  $\tilde{\rho}$ , strength of turbulence  $\alpha$ , and detailed disc properties. We test these against numerical simulations with various turbulence-driving mechanisms. The simulations agree well with the predictions, spanning  $t_s \Omega \sim 10^{-4}$ –10,  $\tilde{\rho} \sim 0$ –3,  $\alpha \sim 10^{-10}$ – $10^{-2}$ . Results from ‘turbulent concentration’ simulations and laboratory experiments are also predicted as a special case. Vortices on a wide range of scales disperse and concentrate grains hierarchically. For small grains this is most efficient in eddies with turnover time comparable to the stopping time, but fluctuations are also damped by local gas-grain drift. For large grains, shear and gravity lead to a much broader range of eddy scales driving fluctuations, with most power on the largest scales. The grain density distribution has a log-Poisson shape, with fluctuations for large grains up to factors  $\gtrsim 1000$ . We provide simple analytic expressions for the predictions, and discuss implications for planetesimal formation, grain growth, and the structure of turbulence.

**Key words:** accretion, accretion discs – hydrodynamics – instabilities – turbulence – planets and satellites: formation – protoplanetary discs.

## 1 INTRODUCTION

Dust grains and aerodynamic particles are fundamental in astrophysics. These determine the attenuation and absorption of light in the interstellar medium (ISM), interaction with radiative forces and regulation of cooling, and form the building blocks of planetesimals. Of particular importance is the question of grain clustering and clumping – fluctuations in the local volume-average number/mass density of grains  $\rho_p$  – in turbulent gas.

Much attention has been paid to the specific question of grain density fluctuations and grain concentration in protoplanetary discs. In general, turbulence sets a ‘lower limit’ to the degree to which grains can settle into a razor-thin sublayer; and this has generally been regarded as a barrier to planetesimal formation (though see Goodman & Pindor 2000; Lyra et al. 2009; Lee et al. 2010; Chiang & Youdin 2010, and references therein). However, it is also well established that the number density of solid grains can fluctuate by multiple orders of magnitude when ‘stirred’ by turbulence,

even in media where the turbulence is highly subsonic and the gas is nearly incompressible (see e.g. Bracco et al. 1999; Cuzzi et al. 2001; Johansen & Youdin 2007; Carballido, Stone & Turner 2008; Bai & Stone 2010a,b,c; Pan et al. 2011). This can occur via self-excitation of turbulent motions in the ‘streaming’ instability (Youdin & Goodman 2005), or in externally driven turbulence, such as that excited by the magnetorotational instability (MRI), global gravitational instabilities, or convection (Dittrich, Klahr & Johansen 2013; Jalali 2013). Direct numerical experiments have shown that the magnitude of these fluctuations depends on the parameter  $\tau_s = t_s \Omega$ , the ratio of the gas ‘stopping’ time (friction/drag time-scale)  $t_s$  to the orbital time  $\Omega^{-1}$ , with the most dramatic fluctuations around  $\tau_s \sim 1$ . These experiments have also demonstrated that the magnitude of clustering depends on the volume-averaged ratio of solids-to-gas ( $\tilde{\rho} \equiv \rho_p/\rho_g$ ), and basic properties of the turbulence (such as the Mach number). These have provided key insights and motivated considerable work studying these instabilities; however, the fraction of the relevant parameter space spanned by direct simulations is limited. Moreover, it is impossible to simulate anything close to the full dynamic range of turbulence in these systems: the ‘top scales’ of the system are  $\lambda_{\max} \sim \text{au}$ , while the viscous/dissipation scales  $\lambda_v$  of

\*E-mail: phopkins@caltech.edu

the turbulence are  $\lambda_v \sim m$  (Reynolds numbers  $\text{Re} \sim 10^6\text{--}10^9$ , under typical circumstances). Reliably modelling  $\text{Re} \gtrsim 10^4$  remains challenging in state-of-the-art simulations (see e.g. Federrath 2013). Clearly, some analytic model (even a very approximate one) for these fluctuations would be tremendously helpful.

The question of ‘preferential concentration’ of aerodynamic particles is actually much more well studied in the terrestrial turbulence literature. There both laboratory experiments (Squires & Eaton 1991; Fessler, Kulick & Eaton 1994; Rouson & Eaton 2001; Gualtieri, Picano & Casciola 2009; Monchaux, Bourgoïn & Cartellier 2010) and numerical simulations (Cuzzi et al. 2001; Hogan & Cuzzi 2007; Yoshimoto & Goto 2007; Bec et al. 2009; Pan et al. 2011; Monchaux, Bourgoïn & Cartellier 2012) have long observed that very small grains, with Stokes numbers  $\text{St} \equiv t_s/t_c(\lambda_v) \sim 1$  (ratio of stopping time to eddy turnover time at the viscous scale) can experience order-of-magnitude density fluctuations at small scales (at/below the viscous scale). Considerable analytic progress has been made understanding this regime: demonstrating, for example, that even incompressible gas turbulence is unstable to the growth of inhomogeneities in grain density (Elperin, Kleeorin & Rogachevskii 1996; Elperin, Kleeorin & Rogachevskii 1998), and predicting the behaviour of the small-scale grain–grain correlation function using simple models of Gaussian random-field turbulence (Sigurgeirsson & Stuart 2002; Bec, Cencini & Hillerbrand 2007). But extrapolation to the astrophysically relevant regime is difficult for several reasons: the Reynolds numbers of interest are much larger, and as a result the Stokes numbers are also generally much larger (in the limit where grains do not cluster below the viscous/dissipation scale because  $t_s \gg t_c(\lambda_{\text{max}})$ ), placing the interesting physics well in the inertial range of turbulence, and rotation/shear, external gravity, and coherent (non-random field) structures appear critical (at least on large scales). This parameter space has not been well studied, and at least some predictions (e.g. those in Sigurgeirsson & Stuart 2002; Bec et al. 2008; Zaichik & Alipchenkov 2009) would naively lead one to estimate much smaller fluctuations than are recorded in the experiments above.

However, these studies still contribute some critical insights. They have repeatedly shown that grain density fluctuations are tightly coupled to the local vorticity field: grains are ‘flung out’ of regions of high vorticity by centrifugal forces, and collect in the ‘interstices’ (regions of high strain ‘between’ vortices). Studies of the correlation functions and scaling behaviour of higher Stokes-number particles suggest that, in the inertial range (ignoring gravity and shear), the same dynamics apply, but with the scale-free replacement of a ‘local Stokes number’  $t_s/t_c$ , i.e. what matters for the dynamics on a given scale are the vortices of that scale, and similar concentration effects can occur whenever the eddy turnover time is comparable to the stopping time (e.g. Yoshimoto & Goto 2007; Bec et al. 2008; Wilkinson, Mehlig & Gustavsson 2010; Gustavsson et al. 2012). Several authors have pointed out that this critically links grain density fluctuations to the phenomenon of intermittency and discrete, time-coherent structures (vortices) on scales larger than the Kolmogorov scale in turbulence (see Bec et al. 2009; Olla 2010, and references therein). In particular, Cuzzi et al. (2001) argue that grain density fluctuations behave in a multifractal manner: multifractal scaling is a key signature of well-tested, simple geometric models for turbulence (e.g. She & Leveque 1994; Boldyrev 2002; Schmidt, Federrath & Klessen 2008). In these models, the statistics of turbulence are approximated by regarding the turbulent field as a hierarchical collection of ‘stretched’ singular, coherent structures (e.g. vortices) on different scales (Dubrulle 1994; She & Waymire 1995; Chainais 2006). Such statistical models have been well tested

as a description of the *gas* turbulence statistics (including gas density fluctuations; see e.g. Burlaga 1992; Sorriso-Valvo et al. 1999; Budaev 2008; She & Zhang 2009; Hopkins 2013b). However, only first steps have been taken to link them to grain density fluctuations: for example, in the phenomenological cascade model fit to simulations in Hogan & Cuzzi (2007).

In this paper, we use these theoretical and experimental insights to build a simple, phenomenological model which attempts to ‘bridge’ between the well-studied regime of small-scale turbulence and that of large, astrophysical particles in shearing, gravitating discs. The key concepts are based on the work above: we first assume that grain density fluctuations are driven by coherent eddies, for which we can calculate the perturbation owing to a single eddy with a given scale. Building on Cuzzi et al. (2001) and others, we then attach this calculation to some simple fractal-like (self-similar) assumptions for the statistics of eddies. This allows us to make predictions for a wide range of quantities, which we compare to simulations and experiments.

## 2 ARBITRARILY SMALL GRAINS: PURE GAS DENSITY FLUCTUATIONS

First consider the case where the grains are perfectly coupled to the gas ( $t_s \rightarrow 0$ ), and their volume-average mass density ( $\rho_p$ , as distinct from the *internal* physical density of a single, typical grain) is small compared to the gas density  $\rho_g$ , so grain density fluctuations simply trace gas density fluctuations.

In both subsonic and supersonic turbulence, the gas experiences density fluctuations directly driven by compressive (longitudinal) velocity fluctuations. This leads to the well-known result, in both subsonic and supersonic turbulence, that the density PDF becomes approximately lognormal, with a variance that scales as  $S_{\ln \rho_g} = \ln[1 + \mathcal{M}_c^2]$  where  $\mathcal{M}_c$  is the rms compressive (longitudinal) component of the turbulent Mach number  $\mathcal{M}$  (component projected along  $\nabla \cdot \mathbf{V}$ ; see Federrath, Klessen & Schmidt 2008; Price, Federrath & Brunt 2011; Konstandin et al. 2012; Molina et al. 2012; Federrath & Banerjee 2015).

However, in subsonic turbulence, the gas density fluctuations quickly become small. Simulations of the (very thin) mid-plane dead-zone dust layers typically record  $\mathcal{M} \lesssim 0.1$ ; they confirm that the scaling above holds, but this produces correspondingly small fluctuations in  $\rho_g$  (see e.g. Johansen & Youdin 2007).<sup>1</sup> Yet these same simulations record orders-of-magnitude fluctuations in  $\rho_p$ .

## 3 PARTIALLY COUPLED GRAINS: THE MODEL

### 3.1 The equations of motion and background flow

Now consider grains with non-zero  $t_s$ , in a gaseous medium and some potential field (for now we take this to be a Keplerian disc, the case of greatest interest, but generalize below). Absent grains and turbulence, the gas equilibrium is in circular orbits, at a cylindrical radius  $R$  from the potential centre, with orbital frequency  $\Omega(R)$ . Because of pressure support, the gas does not orbit at exactly the circular velocity  $V_K$ , but at the reduced speed  $V_{\text{gas}}$ , where

$$\eta V_K \equiv V_K - \langle V_{\text{gas}}(R, \rho_p = 0) \rangle \approx \frac{1}{2 \rho_g V_K} \frac{\partial P}{\partial \ln R}. \quad (1)$$

<sup>1</sup> Note that this does not necessarily mean that Mach numbers in the much larger scaleheight gas disc are small, nor that they are unimportant.

Define a rotating frame with origin in the disc mid-plane at  $R$ , with the  $\hat{x}$ -axis along the radial direction and  $\hat{y}$ -axis in the azimuthal (orbital  $\phi$ ) direction; the frame rotates at the circular velocity  $V_K(R)$ , with the angular momentum vector  $\boldsymbol{\Omega}$  oriented along the  $\hat{z}$ -axis and  $\Omega_R \equiv \Omega(R)$ . The local equation of motion for a grain  $i$  with stopping time  $t_s$  becomes

$$\frac{d\mathbf{v}'_i}{dt} = 2\mathbf{v}'_i \times \boldsymbol{\Omega}_R + 3\Omega_R^2 x_i \hat{x} - \Omega_R^2 z_i \hat{z} - \frac{\mathbf{v}'_i - \mathbf{u}'}{t_s} \quad (2)$$

where  $\mathbf{u}'$  is the gas velocity in the rotating frame. Note that this is a Lagrangian derivative (equation (2) follows the grain path). With no loss of generality, we can conveniently define velocities relative to the linearized Keplerian velocities,  $\mathbf{v} \equiv \mathbf{v}'_i + (3/2)\Omega_R x \hat{y}$  and  $\mathbf{u} \equiv \mathbf{u}' + (3/2)\Omega_R x \hat{y}$ .

Nakagawa, Sekiya & Hayashi (1986) show that for the coupled gas-grain system with dimensionless stopping time  $\tau_s \equiv t_s \Omega_R$  and mid-plane volume-average grain-to-gas mass ratio  $\tilde{\rho} \equiv \rho_p/\rho_g$ , this leads to a quasi-steady-state equilibrium drift solution for the grains and gas, with grain velocity (in the local rotating frame)  $\langle \mathbf{v} \rangle = \mathbf{v}^d = v_x^d \hat{x} + v_y^d \hat{y}$  and gas velocity  $\langle \mathbf{u} \rangle = \mathbf{u}^d = u_x^d \hat{x} + u_y^d \hat{y}$ :

$$v_x^d = -\frac{2\tau_s}{\tau_s^2 + (1 + \tilde{\rho})^2} \eta V_K \quad (3)$$

$$v_y^d = -\frac{1 + \tilde{\rho}}{\tau_s^2 + (1 + \tilde{\rho})^2} \eta V_K \quad (4)$$

$$u_x^d = +\frac{2\tau_s \tilde{\rho}}{\tau_s^2 + (1 + \tilde{\rho})^2} \eta V_K \quad (5)$$

$$u_y^d = -\frac{\tau_s^2 + (1 + \tilde{\rho})}{\tau_s^2 + (1 + \tilde{\rho})^2} \eta V_K \quad (6)$$

$$|\mathbf{v}_{\text{drift}}| = |\mathbf{v}^d - \mathbf{u}^d| = \frac{2\tau_s \sqrt{(1 + \tilde{\rho})^2 + \tau_s^2/4}}{\tau_s^2 + (1 + \tilde{\rho})^2} \eta V_K. \quad (7)$$

So now define the ‘peculiar’ grain/gas velocity relative to the steady-state solution,  $\mathbf{v} \equiv \mathbf{v}^d + \delta\mathbf{v}$  and  $\mathbf{u} \equiv \mathbf{u}^d + \delta\mathbf{u}$ . Insert these definitions into equation (2), and – since the turbulent velocities are much smaller than Keplerian<sup>2</sup> – expand  $\eta = \eta(R)$  and  $V_K(R)$  to leading order in  $x/R$ . We then obtain

$$\delta\dot{v}_x \approx 2\Omega_R \delta v_y - \frac{\delta v_x - \delta u_x}{t_s} \quad (8)$$

$$\delta\dot{v}_y \approx -\frac{1}{2}\Omega_R \delta v_x - \frac{\delta v_y - \delta u_y}{t_s}. \quad (9)$$

The  $\hat{z}$  component of equation (2) forms a completely separable equation which is simply that of a damped harmonic oscillator. Thus retaining it has no effect on our derivation below.

<sup>2</sup> We show below that this is internally consistent, but this amounts to the assumption that the individual eddy sizes *within the dust layer* are small compared to the (full) gas disc gradient scalelength, which is easily satisfied in realistic systems.

## 3.2 Encounters between grains and individual turbulent structures

### 3.2.1 A toy model

Now consider an idealized encounter between a grain group with  $\tau_s$  and single, coherent turbulent eddy. We will first illustrate the key dynamics with a purely heuristic model, then follow with a rigorous derivation (for which the key equations are given in Table 1).

Define the eddy coherence length  $\lambda_e$ , and some characteristic peculiar velocity difference across  $\lambda_e$  of  $\delta u = v_e = \mathcal{M}_e c_s$ , so the eddy turnover time can be defined as  $t_e = \lambda_e/|v_e|$ . In inertial-range turbulence, we expect these to scale as power laws, so define

$$\langle |v_e| \rangle = \langle \mathcal{M}_e \rangle c_s = |v_e(\lambda_{\text{max}})| \left( \frac{\lambda_e}{\lambda_{\text{max}}} \right)^{\zeta_1} \propto \lambda_e^{\zeta_1} \quad (10)$$

$$\langle t_e \rangle \equiv \frac{\lambda_e}{|v_e|} = t_e(\lambda_{\text{max}}) \left( \frac{\lambda_e}{\lambda_{\text{max}}} \right)^{1-\zeta_1}. \quad (11)$$

It is also convenient to define the dimensionless stopping time relative to either the orbital frequency or eddy turnover time:

$$\tau_s \equiv t_s \Omega, \quad \tilde{\tau}_s \equiv t_s/t_e. \quad (12)$$

For now, we will assume  $\rho_p \ll \rho_g$ , so that the back-reaction of the grains on gas can be neglected.

Consider a grain with  $t_s \ll t_e$ , in a sufficiently small eddy that we can ignore the shear/gravity terms across it. Typical eddies are two-dimensional vortices, so the grain is quickly accelerated to the eddy velocity  $v_e$  in an approximately circular orbit. This produces a centrifugal acceleration  $a_{\text{cen}} = \delta v_\theta^2/r \sim v_e^2/\lambda_e = |v_e|/t_e$ , which is balanced by pressure forces for the gas but causes the grain to drift radially out from the eddy centre, at the approximate ‘terminal velocity’ where this is balanced by the drag acceleration  $\sim \delta v_r/t_s$ , so  $\delta v_r \sim t_s v_e^2/\lambda_e = (t_s/t_e) |v_e|$ . If, instead, the eddy is sufficiently large, expansion of the centrifugal force gives  $a_{\text{cen}} \sim 2\Omega v_e$  (the  $2\mathbf{v} \times \boldsymbol{\Omega}$  term in equation 2) – i.e. the global centrifugal force sets a ‘floor’ here, so the terminal velocity is  $\delta v_r \sim 2(t_s/\Omega^{-1}) |v_e|$ .

A real eddy has a velocity gradient across itself. Assume for simplicity that the gradient is strongest across a single dimension, so to first order the local velocity perturbation scales as  $\delta u_y \approx (v_e/\lambda_e)x$  (where  $x = 0$  is the centre of the eddy), while  $\delta u_x \sim \text{constant}$ . If grains have time to come to their terminal velocity while still inside the eddy ( $t_s \ll t_e$ ), then by the arguments above the relative perturbed velocity of two grains on opposite ‘sides’ ( $\pm x$ ) of the eddy will be  $\delta v_x \sim 2x(v_e/\lambda_e)(t_s/t_e)$  or  $\sim 2x(v_e/\lambda_e)\tau_s$  (for small  $t_e \ll \Omega^{-1}$  and large  $t_e \gg \Omega^{-1}$ , respectively). But if the grains do not have time to reach terminal velocity ( $t_s \gg t_e$ ), we can consider them to be sitting ‘in place’ experiencing an approximately constant drag acceleration  $\approx v_e/t_s$  for a time  $\sim t_e$ , so the velocity difference across the eddy at time  $\sim t_e$  is just  $\sim 2x(v_e/\lambda_e)(t_s/t_e)^{-1}$ .

The grain density is determined by the continuity equation  $\partial\rho_p/\partial t + \nabla \cdot (\rho_p \delta\mathbf{v}) = 0$  which we can write as

$$\frac{D \ln \rho_p}{dt} = \left( \frac{\partial}{\partial t} + \delta\mathbf{v} \cdot \nabla \right) \ln \rho_p = -\nabla \cdot \delta\mathbf{v}, \quad (13)$$

where  $D/dt$  is the Lagrangian derivative for a ‘grain population.’ So the  $\delta v_x \propto x$  term means that a population of grains, on encountering this eddy, will expand (be pushed away from the origin of the rotating frame) if  $v_e > 0$ . This is just the well-known result that anticyclonic vortices ( $v_e < 0$ ) on the largest scales tend to collect

**Table 1.** Important variables & key equations derived in this paper.

Variable	Definition	Equation
$\rho_g, c_s$	Mid-plane gas density and sound speed	–
$R, \Omega_R, V_K$	Distance from centre of gravitational potential, Keplerian orbital frequency at $R$ , and Keplerian velocity ( $V_K \equiv \Omega_R R$ )	–
$\lambda_e, v_e, \mathcal{M}_e, t_e$	Characteristic spatial scale, velocity, Mach number ( $\mathcal{M}_e \equiv  v_e /c_s$ ) and turnover time ( $t_e \equiv \lambda_e/ v_e $ ) of a turbulent eddy	–
$\lambda_{\max}, v_e(\lambda_{\max}), \alpha$	Maximum or ‘top’/driving scale of turbulence, with eddy velocity $v_e(\lambda_{\max}) \equiv \alpha^{1/2} c_s$	–
$\lambda_v, \text{Re}, \text{St}$	Viscous/Kolmogorov or ‘bottom’ scale of turbulence; Reynolds number $\text{Re} \equiv (\lambda_{\max}/\lambda_v)^{4/3}$ ; and Stokes $\text{St} \equiv t_s/t_e(\lambda_v)$	–
$\bar{\rho}$	Mean ratio of the volume-average density of solids to gas, in the mid-plane ( $\bar{\rho} \equiv \langle \rho_p \rangle / \langle \rho_g \rangle$ )	–
$\tau_s$	Dimensionless particle stopping time ( $\tau_s \equiv t_s \Omega_R$ )	(8)
$\bar{\tau}_s$	Ratio of particle stopping time to eddy turnover time ( $\bar{\tau}_s \equiv t_s/t_e = \bar{\tau}_s(\lambda_{\max})(\lambda_e/\lambda_{\max})^{1-\zeta_1}$ )	–
$\eta, \Pi$	Difference between the mean gas circular velocity and Keplerian ( $\eta V_K \equiv V_K - \langle V_{\text{gas}} \rangle$ ; $\Pi \equiv \eta V_K/c_s$ )	(1)
$v_{\text{drift}}$	Mean grain-gas relative drift velocity: $v_{\text{drift}} \equiv 2 \eta V_K \tau_s [(1 + \bar{\rho})^2 + \tau_s^2/4]^{1/2} [\tau_s^2 + (1 + \bar{\rho})^2]^{-1}$	(3)
$C_\infty$	Filling factor of eddies: $C_\infty \sim 1-2$ is plausible	(44)
$\zeta_1$	Scaling of one-point gas eddy velocity statistics, $\langle  v_e  \rangle \propto \lambda_e^{\zeta_1}$	(10)
	in the multifractal models used: $\zeta_1 \approx \frac{1}{9} + 2 \left[ 1 - \left( \frac{2}{3} \right)^{1/3} \right]$	(42)
$N_d$	‘Wrapping dimension’ of the singular eddy structures driving density fluctuations	(14)
	( $N_d = 2$ for simple vortices in the disc plane)	
–	Useful variables for Equations below:	Equation
$\beta$	$\beta \equiv \frac{ v_e(\lambda_{\max}) }{ v_{\text{drift}} } = \frac{ v_e(\lambda_{\max})  [(1 + \bar{\rho})^2 + \tau_s^2]}{2 \eta V_K \tau_s [(1 + \bar{\rho})^2 + \tau_s^2/4]^{1/2}} = \frac{(1 + \bar{\rho})^2 + \tau_s^2}{2 \tau_s [(1 + \bar{\rho})^2 + \tau_s^2/4]^{1/2}} \left( \frac{\alpha^{1/2}}{\Pi} \right)$	(37)
$\langle \delta \ln \rho \rangle$	$\langle \delta \ln \rho \rangle \equiv - \frac{N_d \varpi(\tau_s, \bar{\tau}_s)}{1 + h(\lambda_e)^{-1}}$	
	$h(\lambda_e) \equiv -\bar{\tau}_s \ln \left[ 1 - \frac{(\lambda_e/\lambda_{\max})}{\bar{\tau}_s(\lambda_{\max}) g(\lambda_e)^{1/2}} \right], \quad g(\lambda_e) \equiv \frac{1}{\beta^2} + \bar{\tau}_s(\lambda_{\max}) \ln \left[ \frac{1 + \bar{\tau}_s(\lambda_{\max})^{-1}}{1 + \bar{\tau}_s^{-1}} \right],$	(40)
	(Approximation for $\varpi$ : for exact solution see Appendix A):	
	$\varpi = \text{MAX} \left[ \varpi_1, \varpi_0 \equiv 2 \tau_{s,\rho} \left( 1 + \tau_{s,\rho}^2 \right)^{-1} \right] [ \tau_{s,\rho} \equiv \tau_s (1 + \bar{\rho}) \bar{\tau}_{s,\rho} \equiv \bar{\tau}_s (1 + \bar{\rho}) ]$	(A26)
	$0 = 16 \bar{\tau}_{s,\rho}^3 \varpi^4 + 32 \bar{\tau}_{s,\rho}^2 \varpi_1^2 + \bar{\tau}_{s,\rho} (20 + 7 \tau_{s,\rho}^2) \varpi_1^2 + 4 (1 + \tau_{s,\rho}^2 - 3 \tau_{s,\rho} \bar{\tau}_{s,\rho}) \varpi_1 - 4 (\bar{\tau}_{s,\rho} + 2 \tau_{s,\rho})$	(34)
$\rho_{p,\max}$ :	Maximum local density of grains $\rho_p$ :	Equation
	$\ln \left( \frac{\rho_{p,\max}}{\langle \rho_p \rangle} \right) = C_\infty \int_{\lambda=0}^{\lambda_{\max}} [1 - \exp(- \delta \ln \rho )] d \ln \lambda$	(70)
$\Delta^2(k) = \frac{dS_{\ln \rho}}{d \ln \lambda}$ :	(Volume-Weighted) Grain log-density power spectrum (versus scale $\lambda$ ):	Equation
	$\Delta_{\ln \rho}^2 \left( k \equiv \frac{1}{\lambda} \right) = C_\infty  \delta \ln \rho ^2$	(62)
$P_V(\ln \rho_p)$ :	(Volume-weighted) Distribution of Grain Densities $\rho_p$ :	Equation
	$P_V(\ln \rho_p) d \ln \rho_p \approx \frac{(S^{-1} \mu^2)^{m'}}{\Gamma(m' + 1)} \exp(-S^{-1} \mu^2) \frac{\mu}{S} d \ln \rho_p$	(52)–(56)
	$m' \equiv \frac{\mu}{S} \left\{ \frac{\mu^2}{S} \left[ 1 - \exp \left( -\frac{S}{\mu} \right) \right] - \ln \left( \frac{\rho_p}{\langle \rho_p \rangle} \right) \right\}, \quad \mu \equiv C_\infty \int  \delta \ln \rho  d \ln \lambda, \quad S \equiv C_\infty \int  \delta \ln \rho ^2 d \ln \lambda$	
$P_M(\ln \rho_p)$ :	(Mass/particle-weighted) distribution of grain densities $\rho_p$ :	–
	$P_M(\ln \rho_p) d \ln \rho_p = \rho_p P_V(\ln \rho_p) d \ln \rho_p$	–

grains, while cyclonic vortices ( $v_e > 0$ ) disperse them; note that for the small-scale eddies, the sense is *always* dispersal in eddies.<sup>3</sup>

<sup>3</sup> This description of anticyclonic eddies, while common, is actually somewhat misleading. Grains always preferentially avoid regions with high absolute value of vorticity  $|\omega| \sim |v_e \lambda_e^{-1} + \Omega|$ . It is simply that very large eddies ( $t_e \gtrsim \Omega^{-1}$ , with  $t_e = \lambda_e/|v_e|$ ) which are locally anticyclonic and in-plane ( $\hat{v}_e = -\hat{\Omega}$ ) have lower  $|\omega|$  than the mean ( $v_e = 0$ ) Keplerian flow; so grains concentrate there by being dispersed out of higher  $|\omega|$  regions.

Note that above, if the eddy velocity gradient is just one dimensional, the grain population is preferentially dispersed in one dimension; however when the eddies are vortices in two dimensions, the flow is radial (along each dimension). In general, non-zero  $\nabla \cdot \delta \mathbf{v}$  will occur along  $N_d$  dimensions, where  $N_d$  is the number of dimensions along which the eddy flow is locally ‘wrapped.’ For the expected case of simple vortices this is an integer  $N_d = 2$ , but for eddies with complicated structure, or a population of eddies, this can take any non-integer value between zero and the total spatial dimension.



So, on encountering an eddy of scale  $\lambda_e$ , the Lagrangian population of grains with initial extent  $\lambda = \lambda_e$  will shrink or grow in scale according to

$$\frac{D \ln \rho_p}{dt} \sim \begin{cases} \frac{N_d \bar{\tau}_s}{t_e (1 + \mathcal{O}(\bar{\tau}_s^2))} + \mathcal{O}(\bar{\tau}_s \eta^2) + \mathcal{O}(t_e \Omega) & (\text{small } t_e) \\ \frac{2 N_d \bar{\tau}_s}{t_e (1 + \bar{\tau}_s^2)} + \mathcal{O}(\tau_s \eta^2) + \mathcal{O}(t_e \Omega)^{-1} & (\text{large } t_e). \end{cases} \quad (14)$$

We will derive the exact scalings more precisely below, but this simple approximation actually does correctly capture the asymptotic behaviour for small and large eddies.

We now wish to know how long (in an average sense) the perturbation affecting the local grain density distribution in equation (14) is able to act, which we define as the time-scale  $\delta t$ . This obviously cannot be longer than the eddy coherence time, which is about an eddy turnover time  $t_e$ . Since the eddy will not, in equilibrium, accelerate grains to relative speeds *greater* than the flow velocity, it follows that  $t_{\text{sink}} = |\lambda_e / \langle \delta v_{\text{induced}} \rangle|$  (the time-scale for grains to be fully expelled from the eddy region) is always  $> t_e$ , so this does not limit  $\delta t$ . However, if the grains have some non-zero initial relative velocity  $v_0$  with respect to the eddy, and the stopping time is sufficiently long that they are not rapidly decelerated, they can drift through or cross the eddy in finite time  $t_{\text{cross}} \sim \lambda_e / |v_0|$ . If  $t_{\text{cross}} < t_s$  and  $t_{\text{cross}} < t_e$ , then  $\delta t = t_{\text{cross}}$  becomes the limitation. If  $v_0$  is approximately constant (from e.g. global drift or much larger eddies), we have  $t_{\text{cross}} \sim \lambda_e / |v_0| \sim t_e |v_e| / |v_0|$  decreasing on small scales. So we expect  $\delta t / t_e$  is a constant  $\approx 1$  for large-scale eddies, then turns over as an approximate power law  $\propto v_e \propto \lambda_e^{\zeta_1}$ , for  $\lambda_e$  below some  $\lambda_{\text{crit}}$  where  $t_{\text{cross}} < \text{MIN}(t_e, t_s)$ .<sup>4</sup> Again, we will consider this in detail below.

Together, this defines what we call the ‘response function’: the typical density change induced by encounter with an eddy

$$\langle \delta \ln \rho \rangle = \left\langle \int \frac{D \ln \rho_p}{dt} dt \right\rangle \sim \left\langle \frac{D \ln \rho_p}{dt} \right\rangle \delta t. \quad (15)$$

### 3.2.2 Exact solutions in turbulence without external gravity

Now we will derive the previous scalings rigorously.

Consider the behaviour of grain density fluctuations in inertial-range turbulence.<sup>5</sup> First we derive the ‘response function’ above, i.e. the effect of an eddy on a grain distribution. Subtracting the bulk background flow, the grain equations of motion (equation 8) become the Stokes equations

$$\delta \dot{\mathbf{v}} = -\frac{\delta \mathbf{v} - \delta \mathbf{u}}{t_s} \quad (16)$$

with the continuity equation  $\partial \rho_p / \partial t + \nabla \cdot (\rho_p \delta \mathbf{v}) = 0$  which we can write as

$$\frac{D \ln \rho_p}{dt} = \left( \frac{\partial}{\partial t} + \delta \mathbf{v} \cdot \nabla \right) \ln \rho_p = -\nabla \cdot \delta \mathbf{v}, \quad (17)$$

where  $D/dt$  is the Lagrangian derivative for a ‘grain population.’

Many theoretical and experimental studies have suggested that the dynamics of incompressible gas turbulence on various scales can be understood by regarding it as a collection of Burgers vortices (see

Marcu, Meiburg & Newton 1995, and references therein). The Burgers vortex is an exact solution of the Navier–Stokes equations, and provides a model for vortices on all scales (which can be regarded as ‘stretched’ Burgers vortices). In a cylindrical coordinate system centred on the vortex tube, the fluid flow components can be written as  $\delta u_z = 2A z$ ,  $\delta u_r = -A r$ ,  $\delta u_\theta = (B/2\pi r)(1 - \exp(-r^2/2r_0^2))$ , where  $r_0$  is the vortex size,  $B$  the circulation parameter, and  $A = \nu/B$  is the inverse of the ‘vortex Reynolds number.’ Since we consider large  $\text{Re}$ ,  $A \rightarrow 0$  for the large-scale vortices, so  $\delta u_z = \delta u_r = 0$ , and we can specify  $\delta \mathbf{u} = \delta u_\theta \hat{\boldsymbol{\theta}} \equiv u_\theta(r_0/r)(1 - \exp[-r^2/2r_0^2])$ .

On scales  $\lesssim 1.58 r_0$ ,  $\delta u_\theta \propto r + \mathcal{O}(r^2/2.5 r_0^2)$  increases linearly with  $r$ , before turning over beyond the characteristic scale and decaying to zero. So, since we specifically consider the effects of an eddy on scales *within* the eddy size ( $\lesssim r_0$ ), we can take  $\delta u_\theta \propto r$ , in which case the eddy is entirely described by the (approximately constant) turnover time  $t_e$  such that  $\delta u_\theta \equiv r/t_e$ . Note that this is now the general form for *any* eddy with pure circulation and constant turnover time, so while motivated by the Burgers vortex should represent real eddies on a wide range of scales.

We derive an exact, general solution of the problem here in Appendix A, and show this in Figure 1. The qualitative behaviours of the solution derived there is not obvious however, so we illustrate it with a slightly simplified derivation (which captures the correct behaviour in various limits) here.

In the vortex plane, the equations of motion (equation 16) become

$$\begin{aligned} \delta \dot{v}_x &= \delta \dot{v}_{r'} \cos \theta - \delta \dot{v}_\theta \sin \theta - \dot{\theta} (\delta v_{r'} \sin \theta + \delta v_\theta \cos \theta) \\ &= t_s^{-1} (-\delta v_{r'} \cos \theta + \delta v_\theta \sin \theta - \delta u_\theta \sin \theta) \end{aligned} \quad (18)$$

$$\begin{aligned} \delta \dot{v}_y &= \delta \dot{v}_{r'} \sin \theta + \delta \dot{v}_\theta \cos \theta + \dot{\theta} (\delta v_{r'} \cos \theta - \delta v_\theta \sin \theta) \\ &= t_s^{-1} (-\delta v_{r'} \sin \theta - \delta v_\theta \cos \theta + \delta u_\theta \cos \theta) \end{aligned} \quad (19)$$

with  $\dot{\theta} \equiv \delta v_\theta / r'$ . It is straightforward to verify that the peculiar solution is given by  $\delta v_r = \varpi r / t_e$  ( $\delta v_r \propto \delta v_\theta \propto u_\theta \propto r \propto \exp(\varpi t / t_e)$ ) with  $\varpi$  being a root of  $\varpi(1 + \varpi \bar{\tau}_s)(1 + 2\varpi \bar{\tau}_s)^2 - \bar{\tau}_s = 0$ , all of which are decaying solutions except the positive real root:

$$\varpi = \frac{-2 + \sqrt{2(1 + \sqrt{1 + 16\bar{\tau}_s^2})}}{4\bar{\tau}_s} \quad (20)$$

$$\varpi \rightarrow \begin{cases} \bar{\tau}_s & (\bar{\tau}_s \ll 1) \\ (2\bar{\tau}_s)^{-1/2} & (\bar{\tau}_s \gg 1), \end{cases} \quad (21)$$

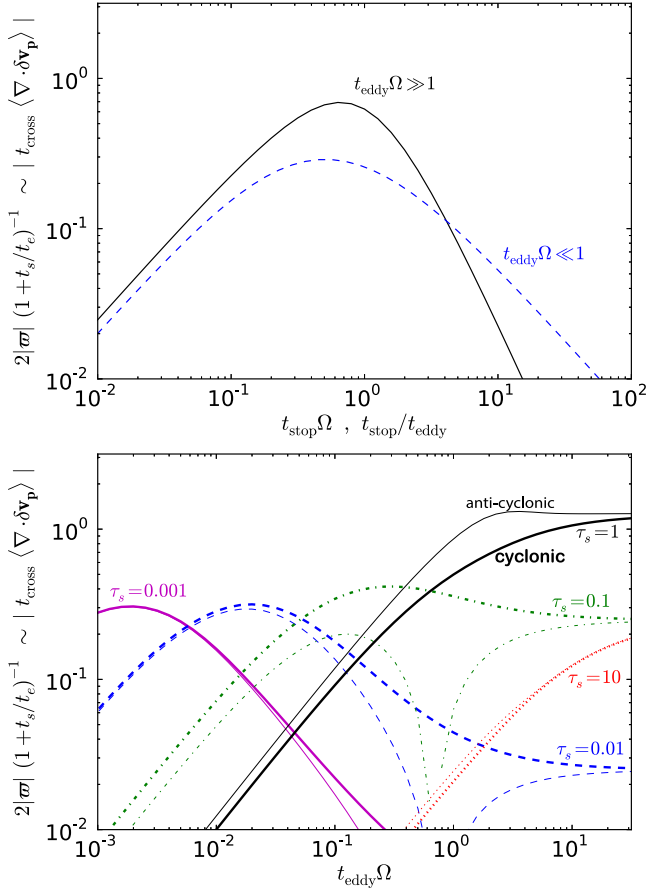
because  $\delta v_r \propto r$  and  $\delta v_\theta$  is independent of  $\theta$ , it follows that along this solution

$$(\nabla \cdot \delta \mathbf{v})_{\text{pec}} = \frac{1}{r} \frac{\partial(r v_r)}{\partial r} + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} = N_d \frac{|v_r|}{r} = \frac{2\varpi}{t_e}. \quad (22)$$

To determine the general solution, we must consider how the eddy evolves in time, since it is able to act on the grains for only finite  $\delta t$ . To approximate this, consider the simplest top-hat model,  $\delta \mathbf{u} = \delta u_\theta \Theta(0 < t < \delta t)$ , where  $\delta \mathbf{u} \propto \Theta = 0$  at  $t < 0$  and  $t > \delta t$  and  $\Theta = 1$  ( $\delta \mathbf{u} = \delta u_\theta(r) \hat{\boldsymbol{\theta}}$ ) for  $0 < t < \delta t$ . We require the net effect of the eddy on the density field (i.e. the late-time result of the perturbation), so we integrate  $d \ln \rho_p / dt = -\nabla \cdot \delta \mathbf{v}$ , from the boundary condition  $\delta \mathbf{v} = \delta \mathbf{v}_0$  at  $t < 0$  until sometime much longer than the eddy lifetime

<sup>4</sup> It is straightforward to show that this time-scale restriction guarantees our earlier assumption (where we dropped higher order terms in the gradient of  $\eta$ ) is valid (for time-scales  $< \delta t$  and subsonic peculiar velocities).

<sup>5</sup> We specifically will assume high Reynolds number  $\text{Re} \gg 1$  and Stokes number  $\text{St} = t_s / t_e(\lambda_v) \gg 1$ , where  $\lambda_v$  is the viscous scale, so will neglect molecular viscosity/diffusion throughout.



**Figure 1.** ‘Response function’  $\overline{|\nabla \cdot \delta \mathbf{v}_p|}$  defined in Section 3.2: this is the mean divergence produced in the peculiar grain velocity distribution by a simple vortex with a turnover time  $t_e$ . Top: limiting cases. First, small eddies ( $t_e \ll \Omega^{-1}$ ), where  $\overline{|\nabla \cdot \delta \mathbf{v}_p|}$  is a function of  $\tilde{\tau}_s \equiv t_s/t_e$  alone ( $\propto \tilde{\tau}_s$  for  $\tilde{\tau}_s \ll 1$ ,  $\propto (2\tilde{\tau}_s)^{-1/2}$  for  $\tilde{\tau}_s \gg 1$ ). Secondly, large eddies ( $t_e \gg \Omega^{-1}$ ), where  $\overline{|\nabla \cdot \delta \mathbf{v}_p|}$  is a function of  $\tau_s \equiv t_s$  alone ( $\propto 2\tau_s$  for  $\tau_s \ll 1$ ,  $\propto 2\tau_s^{-1}$  for  $\tau_s \gg 1$ ). Bottom: full solution (Appendix A), for different  $\tau_s$  and  $t_e$ . For small grains  $\tau_s \ll 0.1$ , there is a broad ‘resonant’ peak around  $t_s \sim t_e$  (spanning  $0.05 t_s \lesssim t_e \lesssim 10 t_s$ ). On the largest scales ( $t_e$ ), the value saturates – this produces a broader and higher amplitude ‘plateau’ for large grains ( $\tau_s \gtrsim 0.1$ ). Both cyclonic (thick) and anticyclonic (thin) eddies are shown: for the anticyclonic cases (eddy angular momentum anti-aligned with  $\Omega$ ), the ‘dip’ near  $t_e \sim \Omega$  comes from a sign change in  $\overline{|\nabla \cdot \delta \mathbf{v}_p|}$ . For  $t_e \ll \Omega$ , the two cases are identical. For  $t_e \gg \Omega$  they have opposite signs.

$t \rightarrow \infty$ . For the simple top-hat form of the eddy lifetime this gives<sup>6</sup>

$$\begin{aligned} \langle \delta \ln \rho \rangle &= \left\langle - \int_t^{\delta t} (\nabla \cdot \delta \mathbf{v}) dt \right\rangle^{\delta u = \delta u_0 \Theta(0 < t < \delta t)} \\ &= -|\nabla \cdot \delta \mathbf{v}|_{\text{pec}} \delta t = -2\overline{|\nabla \cdot \delta \mathbf{v}_p|} \frac{\delta t}{t_e}, \end{aligned} \quad (23)$$

<sup>6</sup> Here, we note that there are two decaying oscillatory solutions to equations (18)–(19) with decay rate  $\omega' = -1/2 \tilde{\tau}_s$ , which correspond to the usual damped modes for grains with drag in a uniform flow. The general solution is derived by matching the modes to the initial velocities  $\delta \mathbf{v}_0$ ; then the solution at  $t = \delta t$  is matched to the solution for the post-eddy field. The exact result is in Appendix A; but if we linearize in  $t_e$ , for example, it is straightforward to show that the integral over time ( $t \rightarrow \infty$ ) of  $\nabla \cdot \delta \mathbf{v}$  is exactly the integral of the positive real peculiar solution  $|\nabla \cdot \mathbf{v}|_{\text{pec}}$  from  $t = 0$  to  $\delta t$ .

where the  $\langle \dots \rangle$  denotes an average over an assumed homogeneous, isotropic initial ensemble in position and velocity space.

Now we need to determine  $\delta t$ . If the grains are ‘trapped’ well within the eddy, this is simply the eddy lifetime  $t_e$ . However, we have not yet accounted for the finite spatial coherence of the eddy. If the grains are moving sufficiently fast and/or if the stopping time is large, they can cross or move ‘through’ the eddy (to  $r \gg r_0 \equiv \lambda_e = |v_e| t_e$ , where the eddy circulation is superexponentially suppressed so becomes negligible) in a time-scale  $t_{\text{cross}} \lesssim t_e$ . Since we are integrating a rate equation, the full  $\delta t$  is simply given by the harmonic mean

$$\delta t^{-1} = t_e^{-1} + t_{\text{cross}}^{-1}, \quad (24)$$

(capturing both limits above; see Voelk et al. 1980; Markiewicz, Mizuno & Voelk 1991). The time-scale for a grain to cross a distance  $\lambda_e = |v_e| t_e$  in a smooth flow (constant  $\delta \mathbf{u}$ ), with initial (relative) grain-gas velocity  $|v_0|$ , is just

$$t_{\text{cross}} = \begin{cases} -t_s \ln \left[ 1 - \frac{\lambda_e}{t_s |v_0|} \right] & (|v_0| > \lambda_e/t_s) \\ \infty & (|v_0| \leq \lambda_e/t_s). \end{cases} \quad (25)$$

Note that for large  $t_s$ , this is just the ballistic crossing time  $\rightarrow \lambda_e/|v_0|$ , but for small  $|v_0| < \lambda_e/t_s$  this diverges because the grain is fully stopped and trapped without reaching  $\lambda_0$ . So now we need to determine  $|v_0|$ , but this is considered in Voelk et al. (1980) and many subsequent calculations (e.g. Markiewicz et al. 1991; Pan & Padoan 2010, 2013). Assuming the turbulence is isotropic and (on long time-scales) velocity ‘kicks’ from independent eddies are uncorrelated, then

$$\langle |v_0|^2 \rangle = |V_L|^2 + \langle |V_{\text{rel}}(\lambda_e)|^2 \rangle, \quad (26)$$

where  $V_L = |V_L|$  is the difference in the laminar bulk flow velocity of grains and gas (due to e.g. settling or gravity) and  $\langle |V_{\text{rel}}(\lambda_e)|^2 \rangle$  represents the rms grain-eddy velocities (averaged on the eddy scale) due to the turbulence itself.

For the ‘pure turbulence’ case here  $V_L = 0$ , and  $\langle |V_{\text{rel}}(\lambda_e)|^2 \rangle$  is derived in Voelk et al. (1980) as

$$\begin{aligned} \langle |V_{\text{rel}}(\lambda_e)|^2 \rangle &= \int_{k(\lambda_{\text{max}})}^{k(\lambda_e)} dk P(k) \frac{t_s}{t_s + t_e(k)} \\ &= |v_e(\lambda_{\text{max}})|^2 \tilde{\tau}_s(\lambda_{\text{max}}) \ln \left[ \frac{1 + \tilde{\tau}_s(\lambda_{\text{max}})^{-1}}{1 + \tilde{\tau}_s(\lambda_e)^{-1}} \right], \end{aligned} \quad (27)$$

where  $k$  is the wavenumber and  $P(k) = (p-1)k^{-p}$  is the velocity power spectrum ( $\int dk P(k) = \langle \delta u^2 \rangle = v_e(\lambda_{\text{max}})^2$ ; the closed-form expression here follows for any power law  $P(k)$ ; see Ormel & Cuzzi 2007).<sup>7</sup>

After some simple substitution, we now have

$$\frac{\delta t}{t_e} = \left[ 1 + \left( \frac{t_{\text{cross}}}{t_e} \right)^{-1} \right]^{-1} \quad (28)$$

$$\frac{t_{\text{cross}}}{t_e} = -\tilde{\tau}_s \ln \left[ 1 - \frac{(\lambda_e/\lambda_{\text{max}})}{\tilde{\tau}_s(\lambda_{\text{max}}) g_0(\lambda_e)^{1/2}} \right] \quad (29)$$

$$g_0(\lambda_e) \equiv \tilde{\tau}_s(\lambda_{\text{max}}) \ln \left[ \frac{1 + \tilde{\tau}_s(\lambda_{\text{max}})^{-1}}{1 + \tilde{\tau}_s(\lambda_e)^{-1}} \right] \quad (30)$$

<sup>7</sup> Due to intermittency, the eddy intensity will vary at a given scale, which will in turn lead to non-linear variations in the time-scale for particle crossing. We consider a heuristic model for this below, but it is worth further investigation.

giving a complete description of  $|\delta \ln \rho| = 2 \varpi (\delta t / t_e)$ .

Note that we have implicitly assumed a simple and time-constant structure for the eddies in deriving this; another important effect, discussed in Falkovich & Pumir (2004), is that the eddy can stretch and deform the ‘parcel’ of particles such that the crossing time varies across the eddy, or even to deform the parcel until its largest dimension is longer than that of the eddy itself. This could revise the time-scales above significantly. Our simplified model essentially folds this into the ‘effective’ eddy lifetime. If the effect were systematic, it would manifest as a systematic change in the eddy lifetime relative to  $t_e$ , which is ultimately degenerate with the amplitude of the effects we predict (in e.g. the parameter  $N_d$ ). More likely, it is another cause which in more detail would contribute to a distribution of eddy-crossing times.

### 3.2.3 Solution with shear

Here, we discuss the full solution (including shear) for the density fluctuations for  $N_d = 2$  vortices above. Again, the exact result is derived in Appendix A, and shown in Figure 1. Consider the same eddies as in Section 3.2.2, but retain the shear terms from equation (2). Equation (18)–(19) become

$$\begin{aligned} \delta \dot{v}_x &= \delta \dot{v}_{r'} \cos \theta - \delta \dot{v}_{\theta'} \sin \theta - \dot{\theta} (\delta v_{r'} \sin \theta + \delta v_{\theta'} \cos \theta) \\ &= t_s^{-1} (-\delta v_{r'} \cos \theta + \delta v_{\theta'} \sin \theta - \delta u_{\theta'} \sin \theta) \\ &\quad + 2 \Omega_R (\delta v_{r'} \sin \theta + \delta v_{\theta'} \cos \theta) \end{aligned} \quad (31)$$

$$\begin{aligned} \delta \dot{v}_y &= \delta \dot{v}_{r'} \sin \theta + \delta \dot{v}_{\theta'} \cos \theta + \dot{\theta} (\delta v_{r'} \cos \theta - \delta v_{\theta'} \sin \theta) \\ &= t_s^{-1} (-\delta v_{r'} \sin \theta - \delta v_{\theta'} \cos \theta + \delta u_{\theta'} \cos \theta) \\ &\quad - \frac{1}{2} \Omega_R (\delta v_{r'} \cos \theta - \delta v_{\theta'} \sin \theta). \end{aligned} \quad (32)$$

First note that, even when  $t_e \gg t_s$ , there is no  $\theta$ -independent solution if we retain all terms (unlike the case in Section 3.2.2, valid at all  $\theta$ ). Because the shear terms break the symmetry of the problem, in the equilibrium solution a grain drifts on an approximately elliptical orbit, with an epicyclic correction to the circular solution which extends the orbit along the shear direction. The exact result in Appendix A accounts for this by computing the Jacobian for the distorted ellipse; however we can gain considerable intuition by considering the simplest ( $\theta \approx 0$ ) case.

With that caveat, we can follow the identical procedure as in Section 3.2.2. In this regime, there are two solution branches: the first is again an exponentially growing solution with frequency  $= \varpi / t_e$  and divergence

$$\langle \delta \ln \rho \rangle_{\lambda_e} = \left\langle - \int_0^\infty (\nabla \cdot \delta \mathbf{v}) dt \right\rangle \approx -2 \varpi (\lambda_e) \frac{\delta t}{t_e} \quad (33)$$

but now with  $\varpi = \varpi_1$  given by the positive, real root of

$$\begin{aligned} 0 &= 16 \tilde{\tau}_s^3 \varpi_1^4 + 32 \tilde{\tau}_s^2 \varpi_1^3 + \tilde{\tau}_s (20 + 7 \tau_s^2) \varpi_1^2 \\ &\quad + 4(1 + \tau_s^2 - 3 \tau_s \tilde{\tau}_s) \varpi_1 - 4(\tilde{\tau}_s + 2 \tau_s) \end{aligned} \quad (34)$$

$$\varpi_1 \rightarrow \begin{cases} \tilde{\tau}_s & (\tilde{\tau}_s \ll 1) \\ 2(\tau_s + \tau_s^{-1})^{-1} & (\tau_s \gg \tilde{\tau}_s). \end{cases} \quad (35)$$

As expected, on small scales where  $t_e \ll \Omega^{-1}$  ( $\tilde{\tau}_s \gg \tau_s$ ), this reduces to the solution for turbulence without shear (equation 20): we can write  $\varpi_1$  in this limit just in terms of  $\tilde{\tau}_s$ , and it scales as  $\tilde{\tau}_s$  for

$\tilde{\tau}_s \ll 1$ . On sufficiently large scales,  $\tau_s \gg \tilde{\tau}_s$ , we recover the solution we estimated in Section 3.2.1, the approximately constant  $\varpi_1 \rightarrow 2/(\tau_s + \tau_s^{-1})$ .

Nominally for  $\tilde{\tau}_s \gg 1$  this gives  $\varpi_1 \rightarrow (2 \tilde{\tau}_s)^{-1/2}$ ; however, in the full solution there is an additional ‘early time’ solution branch we have dropped. Upon first encountering the eddy, the grains have zero mean (peculiar) vorticity, so the coherent  $\delta v$  grows with time. At sufficiently early times,  $\delta v$  is small and the solution to equations (31)–(32) is obtained by expanding to leading order in  $\delta v$ . After substitution to eliminate  $\delta v_{\theta'}$ , the system simplifies to  $\tau_s^2 \delta \ddot{v}_r + 2 \tau_s \delta \dot{v}_r + (1 + \tau_s^2) \delta v_r = 2 \tau_s \delta u_{\theta'}$ . Just as the solution in Section 3.2.2 above, this has two decaying oscillatory solutions which do not contribute to the integrated  $t \rightarrow \infty$  divergence (since the equations are linearized), and peculiar solution  $\delta v_r = 2 \delta u_{\theta'} / (\tau_s + \tau_s^{-1})$ , i.e. just the ‘large scale’ solution from before. This leads to an identical expression (equation 33) for  $\langle \delta \ln \rho \rangle$ , but with  $\varpi = \varpi_0 = 2/(\tau_s + \tau_s^{-1})$ . This solution track dominates when  $\varpi_1 < \varpi_0$ ; when  $\varpi_1 > \varpi_0$ , the  $\varpi_0$  solution track dominates only for an initial time  $t \ll t_e$ , until (as  $\delta v_r$  grows) the second-order terms in  $\delta v$  become important and  $\varpi \rightarrow \varpi_1$ . Comparing to the exact numerical integration, it is straightforward to verify that general solution for both regimes is qualitatively represented by

$$\varpi (\lambda_e) = \text{MAX} \left[ \varpi_1, \varpi_0 = 2 (\tau_s + \tau_s^{-1})^{-1} \right]. \quad (36)$$

This is directly analogous to our heuristic estimate in Section 3.2.1; global angular momentum sets a ‘floor’ in the (second-order) centrifugal force, here represented by  $\varpi_0$ .

Likewise,  $\delta t$  obeys the same scalings as in Section 3.2.2, with  $\mathbf{V}_{\text{rel}}$  contributed by the turbulence, but now there is a non-zero laminar relative gas-grain flow, given by the equilibrium drift solution  $\mathbf{V}_L = \mathbf{v}^d - \mathbf{u}^d$  in equations (3)–(6):

$$|\mathbf{V}_L|^2 = \frac{4(1 + \tilde{\rho})^2 \tau_s^4 + \tau_s^4}{[(1 + \tilde{\rho})^2 + \tau_s^2]^2} (\eta V_K)^2 = \frac{1}{\beta^2} |v_e(\lambda_{\text{max}})|^2. \quad (37)$$

Together with equations (33)–(34), we can now write

$$\frac{\delta t}{t_e} = \left[ 1 + \left( \frac{t_{\text{cross}}}{t_e} \right)^{-1} \right]^{-1} \quad (38)$$

$$\frac{t_{\text{cross}}}{t_e} = -\tilde{\tau}_s \ln \left[ 1 - \frac{(\lambda_e / \lambda_{\text{max}})}{\tilde{\tau}_s (\lambda_{\text{max}}) g(\lambda_e)^{1/2}} \right], \quad (39)$$

where

$$\begin{aligned} g(\lambda_e) &\equiv \frac{\langle |\mathbf{v}_0|^2 \rangle}{|v_e(\lambda_{\text{max}})|^2} = g_0(\lambda_e) + \frac{1}{\beta^2} \\ &= \frac{1}{\beta^2} + \tilde{\tau}_s (\lambda_{\text{max}}) \ln \left[ \frac{1 + \tilde{\tau}_s (\lambda_{\text{max}})^{-1}}{1 + \tilde{\tau}_s (\lambda_e)^{-1}} \right]. \end{aligned} \quad (40)$$

### 3.3 Hierarchical encounters with many structures

Now, if we assume the gas turbulence follows some simplified scalings, we can embed our estimates for the behaviour in individual eddy encounters into the statistics of the eddies themselves.

Consider the following. We consider a random point  $\mathbf{x}$  in space. Assume that it begins at the mean density. But if there are ‘eddies’ present, which intersect the point, then the density will be modified according to our derivation above. So we need to determine ‘how many’ eddies of different sizes are present.

First, we will make the ad hoc and heuristic assumption that the entire velocity field can be decomposed into a superposition of

‘eddies’ of various ‘sizes’  $\lambda_e$ , which are described by the toy model in the previous section. This is a tremendous simplification, but it allows us to phenomenologically model how a complicated velocity field might non-linearly affect grain clustering. Now, consider the scaling of the gas velocity statistics. Given the assumption of an eddy model, then from standard models of turbulent structure we expect  $|v_e| \propto \lambda_e^{\zeta_1}$  with  $\zeta_1 \approx 1/3$ . In more detail, the cascade models of She & Waymire (1995) and Dubrulle (1994) predict that the structure functions  $\sigma_p(\lambda) = \langle |\Delta \mathbf{u}(\lambda)|^p \rangle \equiv \langle |\delta \mathbf{u}(\mathbf{x}) - \delta \mathbf{u}(\mathbf{x} + \lambda)|^p \rangle$  scale as power laws  $\sigma_p(\lambda) \propto \lambda^{\zeta_p}$  with

$$\zeta_p = (1 - \gamma) \frac{p}{3} + c_g \left[ 1 - \left( 1 - \frac{\gamma}{c_g} \right)^{p/3} \right], \quad (41)$$

where  $\gamma = 2/3$  follows generically from the Navier–Stokes equations (from the Kolmogorov 4/5ths law), and  $c_g = 2$  follows from geometric arguments and fitting to experimental data.<sup>8</sup> What we refer to as  $v_e$  is the one-point function (peculiar eddy velocity differences across the eddy)  $p = 1$ , so we have

$$\zeta_1 = \frac{1}{9} + c_g \left[ 1 - \left( 1 - \frac{2}{3c_g} \right)^{1/3} \right] \approx 0.36 \quad (42)$$

If all eddies/structures had a single ‘size’  $\lambda_e$ , then the probability that our point  $\mathbf{x}$  lies within the domain of an eddy is simply given by the volume filling factor of such structures. Since we have already assumed that eddies are statistically independent and discrete in Section 3.2, the number  $m$  of such eddies (within the context of our very simplified assumptions) intercepting a random point must be Poisson-distributed (because this is just a counting exercise):

$$P(m) = P_{\Delta N}(m) = \frac{\Delta N^m}{m!} \exp(-\Delta N), \quad (43)$$

where the mean  $\langle m \rangle \equiv \Delta N$  is simply related to the eddy filling-factor. Note that if the flow is ergodic, we can equivalently consider this the number of structures (e.g. vortices) encountered by a Lagrangian parcel over a coherence time-scale (Cuzzi et al. 2001; Hopkins 2013a). As argued in Hopkins (2013b) and confirmed by Federrath (2013) and Federrath & Banerjee (2015), that assumption leads to a remarkably accurate description of the first-order intermittency corrections to the *gas* density PDFs in supersonic isothermal turbulence.

Of course, in our eddy decomposition, within the inertial range, there is no single ‘eddy size.’ Rather there is a hierarchy of all sizes from the Kolmogorov scale (vanishingly small, in the limit we consider) to the driving scale. Consider ‘searching’ around our point  $\mathbf{x}$  for eddies which intercept  $\mathbf{x}$ : we begin with a ‘counting sphere’ of radius  $\lambda_1$  centred at  $\mathbf{x}$  and then increase the size of the sphere by a differential interval  $\Delta \ln \lambda = \ln \lambda_2 - \ln \lambda_1$ . Obviously, we do not care if ‘new’ eddies appear within the sphere with sizes  $\lambda_e \ll \lambda_1$ , since they cannot interact with the point  $\mathbf{x}$ . Nor do we care about eddies with sizes  $\lambda_e \gg \lambda_1$ , since we only move a tiny fraction across a single such eddy with this extension. What does matter is whether we find eddies with  $\lambda_e \sim \lambda_1$ . In a differential volume element, we expect to find some differential number  $\Delta N_1(\lambda_e \sim \lambda_1)$  of eddies which intercept the point  $\mathbf{x}$ .

<sup>8</sup> In the original Kolmogorov (1941) model,  $\zeta_p = p/3$ , giving the familiar  $|v_e| \propto \lambda_e^{1/3}$ . The low-order differences between this and the more detailed multifractal models are small. So for our purposes, forcing  $\zeta_1 = 1/3$  instead of equation (41) gives very similar predictions. However, a wide range of experiments favour the scaling in equation (41).

Now, if the turbulent structure is self-similar (or truly fractal), then the only allowed scaling is  $\Delta N \propto |\Delta \ln \lambda|$  on all scales, i.e.

$$\Delta N = C_\infty |\Delta \ln \lambda|. \quad (44)$$

This is equivalent to the statement that the volume-filling factor of eddies with different sizes  $\lambda_e$  is constant – the only possibility if the turbulence is truly self-similar.<sup>9</sup>

Thus far, this is a purely geometric argument, which follows from the idea that the flow structure is self-similar over the inertial range. For the specific geometric assumptions, we have made already about the size and characteristic velocity structure of eddies, then if we do assume the eddies of interest are two-dimensional circular vortices as described above (and that such eddies contain *all* the turbulent power), we can infer  $C_\infty = 2$  by simply normalizing the integrated power. But one could imagine only part of the power is in such structures, or trade off between the assumptions above about the relation between eddy size, turnover time, and  $C_\infty$ , and hence we will consider variations in this parameter below.

It is now straightforward to combine this with the arguments in Section 3.2 to obtain the predicted grain density statistics. Following Section 3.2, assume that each encounter in  $\Delta \ln \lambda$  produces a multiplicative effect on the density statistics with the mean expected magnitude  $\langle \delta \ln \rho [\lambda_e = \lambda] \rangle$  on that scale.<sup>10</sup> In a probabilistic sense, as we sample the density statistics about some random point in space – integrating the effects of all eddies, beginning by counting those with sizes comparable to the largest scales and successively counting smaller and smaller structures – the statistics on successive scales  $\lambda_1$  and  $\lambda_2$  are given by

$$\begin{aligned} \ln [\rho_p(\lambda_2)] &= \ln [\rho_p(\lambda_1)] + m \langle \delta \ln \rho [\lambda_e = \lambda_1] \rangle + \epsilon_0 \\ &= \ln [\rho_p(\lambda_1)] - m |\delta \ln \rho| + \epsilon_0, \end{aligned} \quad (45)$$

where  $m$  is Poisson-distributed as equation (43)–(44), i.e.

$$P \left( \ln \left[ \frac{\rho_p(\lambda_2)}{\rho_p(\lambda_1)} \right] \right) d \ln \left[ \frac{\rho_p(\lambda_2)}{\rho_p(\lambda_1)} \right] = P(m) dm. \quad (46)$$

Mass conservation trivially determines the integration constant

$$\epsilon_0 = \Delta N [1 - \exp(-|\delta \ln \rho|_{\lambda_e=\lambda_1})]. \quad (47)$$

Physically, this should be interpreted as follows. Beginning at the ‘top’ scale  $\lambda_{\max}$  (where  $\rho_p = \langle \rho_p \rangle$  by definition), we can recursively divide the volume about a random point into smaller subvolumes of size  $\lambda$ , each containing a (discrete) number of structures (vortices) with characteristic scale  $\sim \lambda$  (which overlap the ‘point’ we are zooming in on). Each such vortex produces another multiplicative

<sup>9</sup> Note, this is only a statement about the filling factor. The eddies can have any dimensionality, in principle. If eddies are, for example, thin filaments (co-dimension  $C_d = 3 - 1 = 2$ ), then if the filling factor is constant, we expect the *total* number of eddies we encounter in the differential volume element to increase as  $\Delta N_{\text{tot}} \propto \lambda^{C_d} |\Delta \ln \lambda|$ . This is the standard expectation in most geometric models of turbulence (see e.g. She & Leveque 1994; She & Zhang 2009). However, *each* of those eddies only has a probability  $\propto \lambda^{-C_d}$  of intercepting the specific point  $\mathbf{x}$ . Thus, what we care about – the number of eddies we ‘find’ that interact with the point  $\mathbf{x}$  – scales  $\propto |\Delta \ln \lambda|$  as equation (44).

<sup>10</sup> This simplification, that structures produce fluctuations of mean magnitude (given their scale), is substantial; yet for the gas velocity statistics it appears sufficient to capture the power spectrum and PDF shape, and higher-order structure function/correlation statistics to  $\gtrsim 10$ th order in experiments (see She & Zhang 2009). However, we will consider below what happens if there is a variation in the fluctuations produced by structures of the same scale.



effect on the local density field  $\delta \ln \rho$  (additive in log-space); we simplify the statistics by assigning each its mean expected effect ( $\langle \delta \ln \rho(\lambda, t_s, \dots) \rangle$ ). Per Section 3.2, this effect applies *within* the eddy region (dispersing grains from of the eddy centre). But ‘dispersed’ grains must go somewhere; the  $\epsilon_0$  term simply represents the mean effect on the density in the interstices *between* vortices, created by their expulsion of grains.<sup>11</sup> These qualitative effects are well known from simulations and experiments (see references in Section 1); this is their quantitative representation.

### 3.4 Behavior at high grain densities

Thus far, we have neglected the back-reaction of grains on the gas (our predictions are appropriate when  $\rho_p < \rho_g$ ). To extrapolate to  $\rho_p \gtrsim \rho_g$ , we require additional assumptions.

Recall, the background drift solution in equations (3)–(6) already accounts for  $\tilde{\rho}$ . So, after subtracting this flow, the equations (8) for peculiar grain motion remain identical; but the gas equation of motion is (dropping the shear terms for simplicity)

$$\delta \dot{\mathbf{u}} = -\tilde{\rho} \left( \frac{\delta \mathbf{u} - \delta \mathbf{v}}{t_s} \right) - \frac{1}{\rho_g} \nabla \delta P_g, \quad (48)$$

where  $\nabla \delta P_g$  represents the peculiar hydrodynamic forces. In subsonic turbulence, it seems reasonable to make the *ansatz* that the back-reaction, while it may distort the flow  $\delta \mathbf{u}$ , does not alter the driving force ( $\nabla \delta P_g$ ) that forms the eddy.<sup>12</sup> But we know that the ‘zero back-reaction’ eddy structure  $\delta \mathbf{u}_0 \equiv \delta \mathbf{u}(\tilde{\rho} = 0)$  is, by definition, a solution to the equation  $\delta \dot{\mathbf{u}}_0 = -\rho_g^{-1} \nabla \delta P_g$ . So decompose  $\mathbf{u}$  into the sum  $\delta \mathbf{u} \equiv \delta \mathbf{u}_0 + \delta \mathbf{u}'$ , substitute, and obtain

$$\delta \dot{\mathbf{u}}' = -\tilde{\rho} \left( \frac{\delta \mathbf{u}_0 + \delta \mathbf{u}' - \delta \mathbf{v}}{t_s} \right) = \tilde{\rho} \left( \frac{\delta \mathbf{v} - \delta \mathbf{u}_0}{t_s} + \frac{\delta \mathbf{u}'}{t_s} \right). \quad (49)$$

In the limit  $\tilde{\rho} \rightarrow 0$ , we know  $\delta \mathbf{u}' \rightarrow 0$ . For  $0 < \tilde{\rho} \ll 1$ , we expect  $|\delta \mathbf{u}'| \ll \delta \mathbf{u}$ , so we can linearize the equations of motion and obtain  $\delta \dot{\mathbf{u}}' \approx (\tilde{\rho}/t_s)(\delta \mathbf{v} - \delta \mathbf{u}_0) \sim \tilde{\rho} d(\delta \mathbf{v} - \delta \mathbf{u}_0)/dt + \mathcal{O}(\tilde{\rho}^2)$  (where the latter follows from the grain and gas momentum equations assuming  $\langle |\delta \mathbf{u}_0| \rangle > \langle |\delta \mathbf{v}| \rangle$ ), so  $\delta \mathbf{u} \approx \tilde{\rho}(\delta \mathbf{v} - \delta \mathbf{u}_0)$ . In the limit  $\tilde{\rho} \rightarrow \infty$  ( $\rho_g/\rho_p = \tilde{\rho}^{-1} \rightarrow 0$ ), the gas is perfectly dragged by the grains, so  $\delta \mathbf{u} \rightarrow \delta \mathbf{v}$  and  $\delta \mathbf{u}' \rightarrow \delta \mathbf{v} - \delta \mathbf{u}_0$ . Linearizing in this limit in  $\rho_g/\rho_p = \tilde{\rho}^{-1}$  similarly gives  $\delta \mathbf{u}' \sim (1 - \tilde{\rho}^{-1})(\delta \mathbf{v} - \delta \mathbf{u}_0) + \mathcal{O}(\tilde{\rho}^{-2})$ . We can

<sup>11</sup> More rigorously, we could calculate the effects of the strain part of the turbulent field on actively ‘trapping’ grains, rather than assuming purely ‘passive’ trapping via this mass conservation argument. However, calculating the strain field in detail requires a more specific model for the structure of eddies: specifically, for their velocity decay at large radii, *and* the characteristic separations between eddies. If, however, we assume a configuration of Burgers vortices separated by lengths comparable to their core radii, then our simple model produces a PDF quite similar to what one would obtain by direct numerical integration, because the characteristic decay length for the eddy vorticity is itself similar to the eddy size and coherence size of the strain regions,  $\sim \lambda_e$ . If we accounted for the exact spatial distribution of eddies, it would also not much change our result so long as the rms separation in space between eddies was  $\leq \lambda_e$ , which is our ‘overlapping’ limit. That said, it is a particularly interesting subject for future work to consider the role of special strain regions which, although they may be rare, can act as ‘caustics’ (for example, a ridgeline between two overlapping and counter-rotating eddies) and may be important local sites of extreme concentration. We will discuss the accuracy of this approximation further below.

<sup>12</sup> Although we caution that this cannot be strictly true in the regime of high  $\tilde{\rho}$  where the streaming instability operates (Goodman & Pindor 2000; Youdin & Goodman 2005).

simply interpolate between these limits by assuming that, in equilibrium

$$\delta \mathbf{u}' \sim \frac{\tilde{\rho}}{1 + \tilde{\rho}} (\delta \mathbf{v} - \delta \mathbf{u}_0). \quad (50)$$

We stress that this is *not* exact, but it at least gives the correct asymptotic behaviour. Inserting this into the equation for  $\delta \mathbf{v}$ , we have

$$\delta \dot{\mathbf{v}} = -\frac{(\delta \mathbf{v} - [\delta \mathbf{u}_0 + \delta \mathbf{u}'])}{t_s} \rightarrow -\frac{(\delta \mathbf{v} - \delta \mathbf{u}_0)}{t_s(1 + \tilde{\rho})}. \quad (51)$$

But this is our original equation for  $\delta \mathbf{v}$ , modulo the substitution  $t_s \rightarrow t_s(1 + \tilde{\rho})$ . So – given the extremely simple ansatz here – our derivation of  $\varpi$  and previous quantities is identical, but we should replace  $t_s$  with an ‘effective’  $t_{s,\rho} \equiv t_s(1 + \tilde{\rho})$ . In this lowest order approximation, back-reaction lessens the relative velocities (hence friction strength) by dragging gas with grains, and thus lengthens the ‘effective’ stopping time.

The time-scale  $\delta t$  is of course still limited by the eddy lifetime  $t_e$ , and the crossing time solution we previously derived already accounted for  $\tilde{\rho} > 0$  (in the drift time), so we do not need to re-derive it. Finally, we will further assume that the back-reaction, while it may distort individual eddies, does not alter their fractal dimensions (hence other statistics like the gas power spectrum). Of course, this cannot be true in detail, and all of these assumptions are questionable in the limit of  $\tilde{\rho} \sim 1$ . None the less, it provides us with a plausible ‘guess,’ and allows us to phenomenologically extend our model to simulations with large  $\tilde{\rho}$ . Below, we discuss the accuracy of these assumptions, and how well this simplistic extension actually performs.

## 4 PREDICTIONS

### 4.1 The shape of the grain density distribution

The full grain density PDF, averaged on any spatial scale, can now be calculated.

To do, we start on the initial scale  $\lambda = \lambda_{\max}$ . By definition, since this is the top scale of the turbulence and/or box, the density distribution is a delta function with  $\rho_p(\lambda_{\max}) = \langle \rho_p \rangle$ . Now take a differential step in scale  $\ln \lambda \rightarrow \ln \lambda_{\max} - d \ln \lambda$  and convolve this density with the PDF of density changes  $P(\ln[\rho_p(\lambda_2)/\rho_p(\lambda_1)])$  from Section 3.3. Repeat until the desired scale is reached; the one-point density PDF is just this iterated to  $\lambda \rightarrow 0$ .<sup>13</sup>

It is easiest to do this with a Monte Carlo procedure; each point in a large ensemble represents a random point in space (thus they equally sample volume) with its own independent  $\ln \rho_i$ . For each step in scale  $\Delta \ln \lambda$ , draw  $m = m_i$  for each point from the appropriate Poisson distribution for the step (equation 43), and use  $|\delta \ln \rho(\lambda)|$  to calculate the change in  $\ln \rho_i$  (equation 45), and repeat until the desired scale is reached.

Recall that for the largest eddies  $|\delta \ln \rho| \approx 2 N_d/(\tau_s + \tau_s^{-1})$  is approximately constant. In that case, the log-Poisson distribution in Section 3.3 is scale-invariant and infinitely divisible, meaning that the integrated PDF of  $\ln \rho$  is *also* exactly a log-Poisson distribution on all scales, with the same  $|\delta \ln \rho| \sim \text{constant}$ , and  $m$  drawn from a Poisson distribution with the integrated  $\Delta N = C_\infty \ln(\lambda_{\max}/\lambda)$ .

<sup>13</sup> Really, we should truncate (or modify) this at the viscous scale  $\lambda_v$ ; here we assume large Reynolds number  $\text{Re} \rightarrow \infty$ . However, the distinction is important for modest Stokes numbers  $\text{St} = t_s/t_e(\lambda_v)$  or simulations with limited resolution (small ‘effective’  $\text{St}$ ).

However, if  $|\delta \ln \rho|$  depends on scale (as it does on small scales), then the convolved distribution is not exactly log-Poisson. But it is quite accurately approximated by a log-Poisson with the same mean and variance as the exact convolved distribution (see Stewart et al. 2006). These quantities add linearly with scale. Over the differential interval in scale  $d \ln \lambda$  ( $\Delta N = C_\infty d \ln \lambda$ ), the added variance ( $\Delta S$ ) in  $\Delta \ln \rho = \ln(\rho_p[\lambda_2]/\rho_p[\lambda_1])$  is  $\Delta S = \Delta N |\delta \ln \rho|^2$ . So the exact integrated variance in the final volume-weighted  $\ln \rho$  distribution is

$$\begin{aligned} S_{\ln \rho, V}(\lambda) &= \int \frac{dS_{\ln \rho, V}}{d \ln \lambda} d \ln \lambda = \int_{\lambda}^{\lambda_{\max}} \Delta N |\delta \ln \rho|^2 \\ &= \int_{\lambda}^{\lambda_{\max}} C_\infty |\delta \ln \rho(\lambda)|^2 d \ln \lambda. \end{aligned} \quad (52)$$

And the integrated first moment (subtracting the  $\epsilon_0$  term) is

$$\mu = \int \Delta N |\delta \ln \rho| = \int_{\lambda}^{\lambda_{\max}} C_\infty |\delta \ln \rho(\lambda)| d \ln \lambda. \quad (53)$$

The approximate integrated PDF on a scale  $\lambda$  is then given by

$$\begin{aligned} P_V(\ln \rho_p) d \ln \rho_p &\approx \frac{\Delta N_{\text{int}}^m \exp(-\Delta N_{\text{int}})}{\Gamma(m+1)} \frac{d \ln \rho_p}{|\delta \ln \rho|_{\text{int}}} \\ m &= |\delta \ln \rho|_{\text{int}}^{-1} \left\{ \Delta N_{\text{int}} [1 - \exp(-|\delta \ln \rho|_{\text{int}})] - \ln \left( \frac{\rho_p}{\langle \rho_p \rangle} \right) \right\} \end{aligned} \quad (54)$$

which is just the log-Poisson distribution (equation 45) with

$$\Delta N \rightarrow \Delta N_{\text{int}} \equiv \frac{\mu^2}{S_{\ln \rho, V}} \quad (55)$$

$$|\delta \ln \rho| \rightarrow |\delta \ln \rho|_{\text{int}} \equiv \frac{S_{\ln \rho, V}}{\mu}. \quad (56)$$

Note  $\Delta N_{\text{int}}$  is now  $\sim C_\infty \langle \ln(\lambda_{\max}/\lambda) \rangle$ , where  $\langle \dots \rangle$  denotes an average over integration weighted by  $|\delta \ln \rho|$  (i.e. the ‘effective’ dynamic range of the cascade which contributes to fluctuations). And  $|\delta \ln \rho|_{\text{int}}$  similarly reflects a variance-weighted mean.

This determines the volumetric grain density distribution, i.e. the probability, per unit volume, of a given mean grain density  $\rho_p = M_p(V)/V$  within that volume  $V$

$$P_V(\ln \rho_p) = \frac{dP_{\text{Vol}}(\ln \rho_p)}{d \ln \rho_p}. \quad (57)$$

This is trivially related to the mass-weighted grain density distribution  $P_M$ , or equivalently the Lagrangian grain density distribution (distribution of grain densities at the location of each grain, rather than at random locations in the volume):

$$P_M(\ln \rho_p) = \frac{dP_{\text{Mass}}(\ln \rho_p)}{d \ln \rho_p} = \rho_p P_V(\ln \rho_p). \quad (58)$$

Note that as  $\Delta N_{\text{int}} \rightarrow \infty$ , this distribution becomes lognormal. This is generically a consequence of the central limit theorem, for sufficiently large number of independent multiplicative events in the density field.

## 4.2 The grain density power spectrum

The power spectrum of a given quantity is closely related to the real-space variance as a function of scale. Specifically, the variance in some field smoothed with an isotropic real-space window function of size  $\lambda$  is related to the power spectrum by

$$S(\lambda) = \int d^3 \mathbf{k} P(\mathbf{k}) |W(\mathbf{k}, \lambda)|^2, \quad (59)$$

where  $W$  is the window function. If we isotropically average, and adopt for convenience a window function which is a Fourier-space top-hat<sup>14</sup> we obtain

$$S(\lambda) = \int_{k=1/\lambda}^{\infty} \Delta^2(k) d \ln k, \quad (60)$$

where  $\Delta^2(k)$  is now defined as the isotropic, dimensionless power spectrum, and is related to  $S(\lambda)$  by

$$\frac{dS}{d \ln \lambda} = \Delta^2(k[\lambda]). \quad (61)$$

But we know how the variance ‘runs’ as a function of scale, for the logarithmic density distribution. Specifically, for the distribution in equation (45), over some differential interval in scale  $d \ln \lambda$  corresponding to  $\Delta N = C_\infty d \ln \lambda$ , the variance in  $\ln \rho$  is just  $\Delta S = \Delta N |\delta \ln \rho|^2$  (and this adds linearly in scale). So

$$\Delta_{\ln \rho}^2(k) = \frac{dS_{\ln \rho}}{d \ln \lambda} = C_\infty |\delta \ln \rho|^2 = C_\infty \left[ N_d \varpi(\lambda) \frac{\delta t}{t_e} \right]^2. \quad (62)$$

Recall the turnover in  $(\delta t/t_e)$  below  $\lambda_{\text{crit}}$  (Section 3.2), which leads to a two-power-law behaviour in  $\Delta^2$

$$\Delta_{\ln \rho}^2(k) \propto \begin{cases} \text{constant} & (\lambda \gg \lambda_{\text{crit}}) \\ v_e^2 \propto k^{-2\zeta_1} & (\lambda \ll \lambda_{\text{crit}}), \end{cases} \quad (63)$$

i.e. we predict a turnover/break in the power spectrum at a characteristic scale  $\lambda_{\text{crit}}$  (defined in Section 3.2 as the scale where the time-scale for grains to cross an eddy is shorter than the stopping time). On large scales where eddy turnover times are long, the logarithmic statistics are nearly scale-free, but on small scales, where eddy turnover times are short compared to the stopping and drift times, the variance is suppressed.

The power spectrum for the *linear* density field  $\rho$  is similarly trivially determined as

$$\Delta_\rho^2 = \frac{dS_\rho}{d \ln \lambda}. \quad (64)$$

However  $S_\rho$  is not so trivially analytically tractable, since the total variance does not sum simply in *linear*<sup>15</sup>  $\rho$ . But it is straightforward to construct  $S_\rho$ , by simply using equation (45) to build the density PDF at each scale, directly calculating the variance in the linear  $\rho$ , and then differentiating. If the density PDF is approximately log-Poisson, then we can have

$$S_\rho \approx \exp\{\Delta N_{\text{int}} (1 - e^{-|\delta \ln \rho|_{\text{int}}})^2\} - 1. \quad (65)$$

This leads to the somewhat cumbersome expression for  $\Delta_\rho^2$ :

$$\begin{aligned} \Delta_\rho^2 &\approx C_\infty S_\rho \frac{|\delta \ln \rho|_{\text{int}}}{|\delta \ln \rho(\lambda)|^2} e^{-2|\delta \ln \rho|_{\text{int}}} (e^{|\delta \ln \rho|_{\text{int}}} - 1) \\ &\times [2 |\delta \ln \rho|_{\text{int}} (e^{|\delta \ln \rho|_{\text{int}}} - 1 + |\delta \ln \rho(\lambda)|) \\ &- 2 |\delta \ln \rho|_{\text{int}}^2 - |\delta \ln \rho(\lambda)| (e^{|\delta \ln \rho|_{\text{int}}} - 1)]. \end{aligned} \quad (66)$$

<sup>14</sup> We treat the isotropically averaged, Fourier-space top-hat case purely for convenience, because it is usually measured and more relevant on small scales. This is not the same as assuming the power spectrum is intrinsically isotropic or that Fourier modes are uncoupled.

<sup>15</sup> This is a general point discussed at length in Hopkins (2013a), Appendices F–G; it is not, in general, possible to construct a non-trivial field distribution that is simultaneously scale-invariant under linear-space and logarithmic-space convolutions. However, as shown therein, the compound log-Poisson cascade is *approximately* so, to leading order in the expansion of the logarithm.

But the limits are easily understood: if  $|\delta \ln \rho|_{\text{int}} \approx |\delta \ln \rho(\lambda)|$ ,  $\Delta_\rho^2 \rightarrow C_\infty [1 - \exp(-|\delta \ln \rho|)]^2 S_\rho$ : this just reflects the scaling from the ‘number of structures’. If  $|\delta \ln \rho| \ll 1$ , this further becomes  $\Delta_\rho^2 \sim C_\infty |\delta \ln \rho|^2 = \Delta_{\ln \rho}^2$ , since for small fluctuations the linear and logarithmic descriptions are identical. If  $|\delta \ln \rho| \gtrsim 1$  is large,  $\Delta_\rho^2 \sim C_\infty \exp(\Delta N_{\text{int}}) \sim C_\infty (\lambda/\lambda_{\text{max}})^{-C_\infty}$  is a power law, whose scaling (slope  $C_\infty \sim 2$ ) only depends on geometric scaling of the ‘number of structures’ (fractal dimension occupied by vortices).

### 4.3 Correlation functions

The (isotropically averaged) autocorrelation function  $\xi(r)$  is

$$\xi(r) \equiv \frac{1}{\langle \rho_p \rangle^2} (\langle \rho_p(\mathbf{x}) - \langle \rho_p \rangle \rangle (\rho_p(\mathbf{x}) - \langle \rho_p \rangle)) \quad (67)$$

equivalently, this is the excess probability of finding a number of grains in a volume element  $dV$  at a distance  $r$  from a given particle (not a random point in space)<sup>16</sup>

$$\langle dN_p(r, r + dr) \rangle = \langle n_p \rangle dV [1 + \xi(r)]. \quad (68)$$

$\xi(r)$  is directly related to the variance  $\langle (\rho_p[r] - \langle \rho_p \rangle)^2 \rangle$  of the linear density field  $\rho_p[r]$  averaged on the scale  $r$ , by

$$\frac{1}{V(r)} \int_{V(r)} \xi(\mathbf{r}') d^3 \mathbf{r}' = \frac{\langle (\rho_p(r) - \langle \rho_p \rangle)^2 \rangle}{\langle \rho_p \rangle^2} \equiv S_\rho \quad (69)$$

(Peebles 1993).<sup>17</sup> So the correlation function contains the same statistical information as the density power spectrum; and if we calculate  $S_\rho$  above it is straightforward to determine  $\xi(r)$  by equation (69).

Note that if  $\xi(r)$  is a power law,  $S_\rho(r) \sim \xi(r)$ . And if  $|\delta \ln \rho|_{\text{int}} \ll 1$  ( $\tau_s \ll 1$ ), equation (65) simply becomes  $S_\rho \sim \Delta N_{\text{int}} |\delta \ln \rho|_{\text{int}}^2$ . On the largest scales  $t_e \gtrsim \Omega^{-1}$ ,  $|\delta \ln \rho| \sim \text{constant}$  so  $\xi(r)$  rises weakly (with  $\Delta N_{\text{int}}$ ) with decreasing  $r$  as  $1 + \xi(r) \propto \ln(1/r)$ ; approaching scales  $t_e \sim t_s$ ,  $|\delta \ln \rho| \propto \bar{\tau}_s$  rises so  $\xi(r) \propto \bar{\tau}_s^2 \propto \lambda^{-2(1-\zeta_1)}$  rises as a power law with a slope near unity; finally on small scales  $t_e \lesssim t_s$ ,  $|\delta \ln \rho|$  falls rapidly, so  $|\delta \ln \rho|_{\text{int}}$  and  $\Delta N_{\text{int}}$  converge and  $\xi(r) \rightarrow \text{constant}$ .

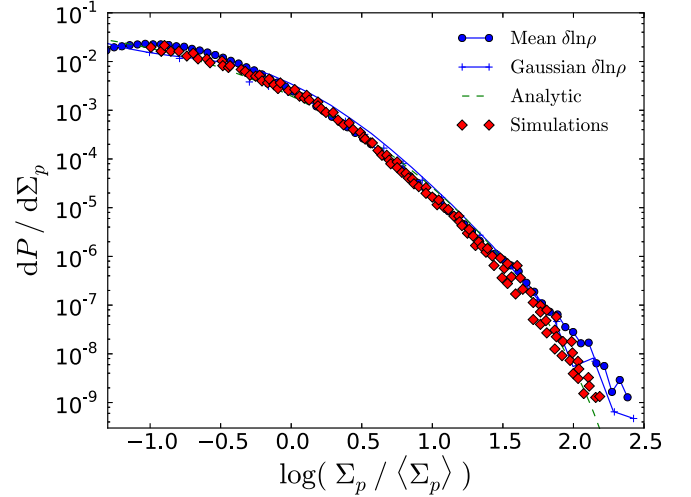
### 4.4 Maximum grain densities

Using the predicted grain density PDFs, we can predict the maximum grain densities that will arise under various conditions.

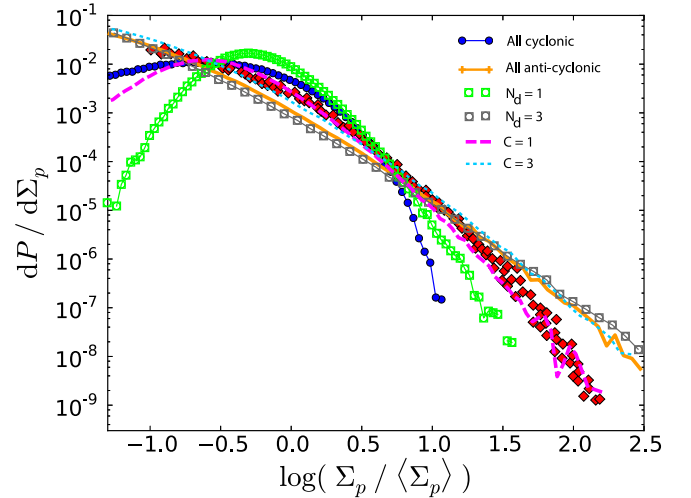
In equation (45), note that there is, in fact, a maximum density, given by  $m = 0$  (and  $\epsilon_0 = 0$ ) on all scales. This is approximately the density where the distributions ‘cut off’ in Figs 2–4, steeper than a Gaussian. It is straightforward to estimate this using

<sup>16</sup> Since we assume a uniform grain population, we treat grain mass and number densities as equivalent.

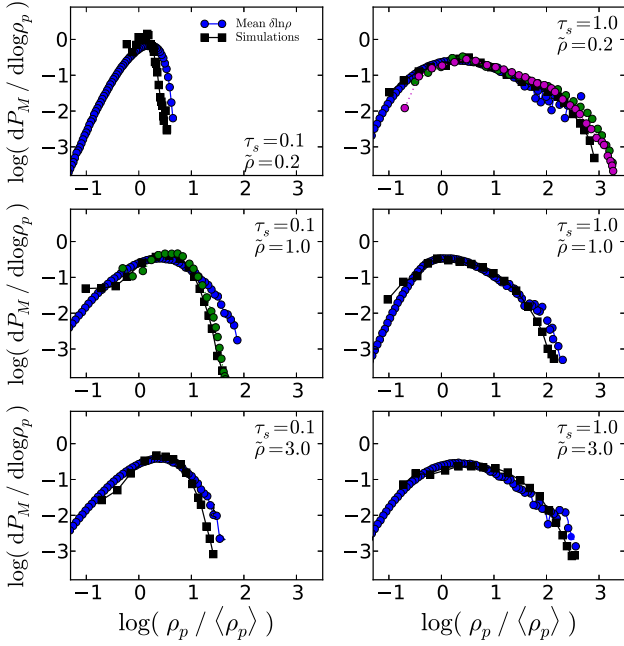
<sup>17</sup> We assume the absolute number of grains is large so we can neglect Poisson fluctuations.



**Figure 2.** Predicted grain density distribution in numerical simulations of MRI-driven turbulence with  $\tau_s = 1$  and  $\bar{\rho} = 0$  (no grain-gas back-reaction). The exact prediction from our Monte Carlo method, given the simulation parameters, is shown either assuming vortices with fixed  $t_e$  each produce the same, mean multiplicative effect (‘mean  $\delta \ln \rho$ ’) or draw from a Gaussian distribution (‘Gaussian  $\delta \ln \rho$ ’). We also show the simple closed-form fitting function (‘Analytic’) derived for fluctuations on large scales (Table 2). This all assumes our ‘default’ model ( $N_d = 2$  dimensional vortices,  $C_\infty = 2$ , and a random cyclonic/anticyclonic distribution). We compare the simulation results from Dittrich et al. (2013). The agreement is very good; the simulations are not able to distinguish the (very similar) ‘mean  $\delta \ln \rho$ ’ and ‘Gaussian  $\delta \ln \rho$ ’ models.



**Figure 3.** Effect of model choices on the predicted distribution in Fig. 2. If we assume all vortices are cyclonic, the distribution is too broad at low densities and cuts off too sharply at high densities. But assuming all vortices are anticyclonic overpredicts the high-density tail (these would also predict net angular momentum in the vortices, which is not allowed in our assumptions). Decreasing/increasing the assumed wrapping dimension of eddies decreases/increases the predicted scatter correspondingly. Interestingly, the effective ‘filling factor’  $C_\infty$  of eddies is only weakly constrained: a range  $C_\infty \sim 0.5\text{--}2$  is broadly consistent with the distribution, though much larger values  $C_\infty \gtrsim 3$  are ruled out.



**Figure 4.** Grain density distribution, as Fig. 2, for simulations of streaming-instability turbulence with  $\tau_s = 0.1$ –1 and  $\tilde{\rho} = 0.2$ –3 (labelled). The simulations are from Johansen & Youdin (2007) and Bai & Stone (2010a); where multiple simulations with different numerical methods are available we plot them, to represent differences owing purely to numerics. As expected from the stronger response functions on large scales in Fig. 1, the PDF width is larger for  $\tau_s = 1$ . Increasing  $\tilde{\rho}$  broadens the PDF for  $\tau_s \ll 1$ , but narrows it when  $\gg 1$ , consistent with our lowest order estimate that it changes the ‘effective stopping time’ as  $\tau_s \rightarrow \tau_s (1 + \tilde{\rho})$ . All the predictions have some growing discrepancies at the highest densities, probably because our assumption that the grains represent a perturbation on the gas turbulence structure is no longer valid.

$\rho_p(\lambda_{\max}) = \langle \rho_p \rangle$ , and taking  $m = 0$  on all scales

$$\begin{aligned} \ln \left( \frac{\rho_{p,\max}[\lambda]}{\langle \rho_p \rangle} \right) &= \int_{\lambda}^{\lambda_{\max}} \epsilon_0 \\ &= C_{\infty} \int_{\lambda}^{\lambda_{\max}} [1 - \exp(-|\delta \ln \rho|)] d \ln \lambda. \end{aligned} \quad (70)$$

Trivially, we see

$$\frac{d \ln \rho_{p,\max}}{d \ln \lambda} = -C_{\infty} [1 - \exp(-|\delta \ln \rho(\lambda)|)], \quad (71)$$

i.e.  $\rho_{p,\max}$  behaves *locally* over some scale range in  $\lambda$  as a power-law  $\rho_{p,\max} \propto \lambda^{-\gamma}$  with slope  $\gamma \equiv C_{\infty} [1 - \exp(-|\delta \ln \rho|)]$ . When  $|\delta \ln \rho| \ll 1$  is small,  $\gamma \sim C_{\infty} |\delta \ln \rho|$  is also small, so  $\rho_{p,\max}$  grows slowly. For sufficiently large  $|\delta \ln \rho| \gtrsim 1$ ,  $\gamma \sim C_{\infty} \sim 2$  saturates at a value determined by the fractal filling factor of vortices –  $\rho_{p,\max}$  grows rapidly with scale, in a power-law fashion with slope  $\sim 2$  determined by the density of structures in turbulence.

## 5 COMPARISON WITH SIMULATIONS AND EXPERIMENTS

Going forward, unless otherwise specified we will assume the ‘default’  $C_{\infty} = 2$  and  $N_d = 2$ . The values  $\tilde{\rho}$  and  $\tau_s$  are necessarily specified for each experiment. With these values, we only need two or three additional parameters to completely determine our model predictions. One is the ratio of eddy turnover time to stopping time on the largest scales  $\tilde{\tau}_s(\lambda_{\max})$  (or equivalently, ratio

$t_e(\lambda_{\max})/\Omega^{-1}$ ), the other is the ratio of mean drift to turbulent velocity  $\beta \equiv |v_e(\lambda_{\max})|/|v_{\text{drift}}|$  (or equivalently the disc parameters  $\alpha^{1/2}/\Pi$ ). These are properties of the gas turbulence and mean flow, so in some cases are pre-specified but in other cases are determined in a more complicated manner by other forces. To compare to simulations, we also (in some cases) need to account for their finite resolution, i.e. minimum  $\lambda/\lambda_{\max}$  or effective Reynolds number. Recall, we assumed a full inertial-range scaling for the turbulence; but many simulations resolve only a very limited (or no) inertial range. As a result, grain clustering (especially on small scales) may be underestimated, and it would be inappropriate to compare our model assuming a fully resolved cascade to the limited dynamic range of the simulations. Finally, the one additional parameter which is sometimes important is the fraction of eddies which are anticyclonic (anti-aligned with the shear flow in the disc) as opposed to cyclonic, when we compare to simulations with an externally imposed shear flow (e.g. a Keplerian gravitational field). For reasons we discuss below, our ‘default’ model assumes these are equally likely.

### 5.1 Density PDFs

#### 5.1.1 Externally driven MRI turbulence

First consider the simulations in Dittrich et al. (2013). These solve the equations of motion for the coupled gas-grain system, in full magnetohydrodynamics (MHD), for a grain population with a single stopping time  $t_s$ . The simulations are performed in a three-dimensional, vertically stratified shearing box in a Keplerian potential, and there is a well-defined local  $\Omega$ ,  $\eta$ ,  $\Pi = 0.05$ . The simulations develop the MRI, which produces a nearly constant  $\alpha \approx 0.004$  (in our units defined here) in the dust layer, and back reaction on gas from grains is ignored so we can take  $\tilde{\rho} \rightarrow 0$ . The authors record a grain density PDF for  $\tau_s = 1$  with large  $\rho_p$  fluctuations arising as the MRI develops (their fig. 11), to which we compare in Fig. 2. Since the disc is vertically stratified,  $\langle \rho_p \rangle$  depends on vertical scale-height, so we would (ideally) compare our predictions separately in each vertical layer (though they are most appropriate for the disc mid-plane). Lacking this information, we should compare instead to the local surface density of grains relative to the mean grain surface density  $\Sigma_p/\langle \Sigma_p \rangle$ , which is independent of stratification and (as shown therein) closely reflects the distribution of mid-plane grain densities.

First, we compare our exact prediction (computed via the Monte Carlo method in Section 4.1). Our natural expectation  $C_{\infty} = N_d = 2$  gives a remarkably accurate prediction of the simulation results! In fact, freeing  $N_d$  and fitting to the data does not significantly improve the agreement (best-fitting  $N_d \approx 1.9 \pm 0.1$ ). We also compare with our closed-form analytic approximation to the integrated density distribution, from Table 2. This gives a very similar result, indicating that for this case (modest resolution simulations, so densities are not averaged on extremely small scales, and large  $\tau_s = 1$ ), the large-scale approximation is good.

In Section 3.3, we adopt the simplest assumption for the effects of an eddy (multiplication by  $(\delta \ln \rho)$ ). As noted there, one might extend this model by instead adopting a distribution of multipliers, with characteristic magnitude  $\delta \ln \rho$ . Here, we consider one such example. For each ‘event’ in  $m$  in the log-Poisson hierarchy, instead of taking  $\ln \rho_p \rightarrow \ln \rho_p + (\delta \ln \rho)$ , assume the ‘multiplier’ is drawn from a Gaussian distribution, so  $\ln \rho_p \rightarrow \ln(\rho_p [1 + \mathcal{R}])$  where  $\mathcal{R}$  is a Gaussian random variable with dispersion  $\langle \mathcal{R}^2 \rangle^{1/2} = (\delta \ln \rho)$ . This is a somewhat arbitrary choice, but illustrative and motivated by Gaussian-like distributions in eddy velocities and lifetimes (and



**Table 2.** Approximations for Large Scales and/or Grains (Appendices A–B).

	Useful variables:
$\varpi,  \delta_0 , \lambda_{\text{crit}}$	$\varpi \sim 2\phi \frac{\tau_{s,\rho}}{1 + \tau_{s,\rho}^2} (\phi \sim 0.8), \quad  \delta_0  \equiv N_d \varpi \sim 2N_d \phi \frac{\tau_{s,\rho}}{1 + \tau_{s,\rho}^2}, \quad \lambda_{\text{crit}} \equiv \beta^{-1/\xi_1} \lambda_{\text{max}}, \quad \xi_1 = 0.36$
$\langle \delta \ln \rho \rangle$	$\langle \delta \ln \rho \rangle = \frac{N_d \varpi}{1 + h(\lambda_e)^{-1}} \sim \frac{2N_d \phi}{\tau_{s,\rho} + \tau_{s,\rho}^{-1}} \left[ 1 + \beta^{-1} \left( \frac{\lambda_e}{\lambda_{\text{max}}} \right)^{-\xi_1} \right]^{-1} =  \delta_0  [1 + (\lambda_e/\lambda_{\text{crit}})^{-\xi_1}]^{-1}$
$\rho_{p,\text{max}}$ :	Maximum local density of grains $\rho_p$ : $\ln \left( \frac{\rho_{p,\text{max}}(\lambda \rightarrow 0)}{\langle \rho_p \rangle} \right) \sim \frac{C_\infty}{\xi_1} \frac{ \delta_0 }{1 +  \delta_0 } \ln[1 + \beta(1 +  \delta_0 ) + \beta^{3/2}\{(1 +  \delta_0 ^2)^{1/2} - 1\}]$ $\rho_{p,\text{max}}(\lambda) \propto \lambda^{-\gamma}$ , with $\gamma \sim \begin{cases} C_\infty [1 - \exp(- \delta_0 )] & (\lambda \gg \lambda_{\text{crit}}) \\ C_\infty  \delta_0  (\lambda/\lambda_{\text{crit}})^{\xi_1} & (\lambda \ll \lambda_{\text{crit}}) \end{cases}$
$\Delta^2(k)$ :	(Volume-weighted) grain linear-density and log-density power spectrum (versus scale $\lambda$ ): $\Delta_{\ln \rho}^2 \left( k \equiv \frac{1}{\lambda} \right) \sim C_\infty  \delta_0 ^2 [1 + (\lambda/\lambda_{\text{crit}})^{-\xi_1}]^{-2}$ $\Delta_\rho^2 \sim \begin{cases} C_\infty  \delta_0 ^2 & (\lambda \gg \lambda_{\text{crit}}, \Delta_{\ln \rho}^2 \ll 1) \\ C_\infty  \delta_0 ^2 (\lambda/\lambda_{\text{crit}})^{2\xi_1} & (\lambda \ll \lambda_{\text{crit}}, \Delta_{\ln \rho}^2 \ll 1) \end{cases} \sim \begin{cases} C_\infty (\lambda/\lambda_{\text{max}})^{-C_\infty} & (\lambda \gg \lambda_{\text{crit}},  \delta_0  \gg 1) \\ 2C_\infty e^{\Delta N_{\text{int}}} \frac{ \delta_0 }{ \delta \ln \rho _{\text{int}}} (\lambda/\lambda_{\text{crit}})^{\xi_1} & (\lambda \ll \lambda_{\text{crit}},  \delta_0  \gg 1) \end{cases}$
$P_V(\ln \rho_p)$ :	(Volume-weighted) Distribution of Grain Densities $\rho_p$ : $P_V(\ln \rho_p) d \ln \rho_p \approx \frac{(\Delta N_{\text{int}})^{m'}}{\Gamma(m' + 1)} \frac{\exp(-\Delta N_{\text{int}})}{ \delta \ln \rho _{\text{int}}} \frac{d \ln \rho_p}{ \delta \ln \rho _{\text{int}}}, \quad m' =  \delta \ln \rho _{\text{int}}^{-1} \left\{ \Delta N_{\text{int}} [1 - \exp(- \delta \ln \rho _{\text{int}})] - \ln \left( \frac{\rho_p}{\langle \rho_p \rangle} \right) \right\}$ $\Delta N_{\text{int}} = \frac{\mu^2}{S}, \quad  \delta \ln \rho _{\text{int}} = \frac{S}{\mu}, \quad \mu \sim \frac{C_\infty}{\xi_1}  \delta_0  \ln(1 + \beta), \quad S \sim \frac{C_\infty}{\xi_1}  \delta_0 ^2 \left( \ln(1 + \beta) - \frac{\beta}{1 + \beta} \right)$

it has the advantage of continuously extending the predictions to all finite  $\rho_p$ , while producing the same change in variance as our fiducial model over small steps in  $\delta \ln \rho$ ;<sup>18</sup> we could instead adopt a  $\beta$ -model as in Hogan & Cuzzi (2007), but it would require additional parameters. In either case, this could, for example, represent the known effects of intermittency leading to variations in eddy intensity from region to region within the flow. Here, however, we see that this makes little difference to the predicted PDF. The reason is that the variance predicted in  $\rho_p$  is dominated by the variance in the local turbulent field (the ‘number of structures’ on different scales), and by the scales on which those structures appear – not by the variance inherent to an individual structure on a specific scale.

However, a number of model assumptions are important here, and we illustrate this in Fig. 3. Recall, this simulation includes an external Keplerian field, therefore eddies lead to different outcomes depending on whether they are cyclonic or anticyclonic. If we assume all eddies are cyclonic, we obtain too broad a distribution at low densities, and too rapid a cutoff at high densities: this is because cyclonic eddies do not actively concentrate grains, but disperse them, and the concentration is a secondary effect resulting from their trapping in strain regions. But sufficiently large anticyclonic eddies lead to large positive concentration effects, hence a stronger tail towards high concentrations. But, assuming all eddies are anticyclonic is similarly problematic, predicting too much concentration. Moreover, because we define all eddies relative to comoving coordinates with the Keplerian disc gas orbits, the total angular momentum in eddies should vanish; if all eddies were cyclonic or anticyclonic,

it would change the global angular momentum of the system, invalidating our original assumptions. To physically maintain this configuration, the gas would have to develop an additional drift and different orbits as angular momentum is transferred into eddies.

Remarkably, if we simply assume random eddy orientations (half cyclonic and half anticyclonic, with the relative number drawn randomly from a binomial distribution for each number of eddies in the log-Poisson distribution we model), we obtain a good fit to the simulations. This is also the simplest physical configuration which gives negligible net angular momentum in eddies, consistent with our original solution for the gas orbits. For now, we will adopt it as our ‘default,’ but wish to stress the caveat that we do not have a full model for what drives the distribution of eddy orientations.

Interestingly, the role of anticyclonic vortices at the high-density tail of the PDF suggests that one shortcoming of our model – the fact that we do not explicitly treat the strain field, but only vorticity, in modelling this regime – may be partially ameliorated by the fact that, in cases where large positive-density fluctuations occur, they are dominated by such eddies, rather than by trapping in the strain field around cyclonic eddies.

We can also explore the effects of varying  $N_d$  and  $C_\infty$  in our simple model. At otherwise fixed properties, lower  $N_d$  leads to lower variance, as expected, since the effects of a single eddy scale  $\propto N_d$ . As noted above, our simplest choice  $N_d = 2$  appears to work very well, while  $N_d = 1$  and 3 do not. The effects of  $C_\infty$  are more subtle. Naively, this also simply multiplies in the variance, and if we keep the product  $N_d^2 C_\infty$  fixed (the combination which enters our estimate of the variance), we do not see large effects from changing  $C_\infty$ . In fact, a range of  $C_\infty \sim 1$ –2 is permitted, and agrees plausibly with a number of the simulation and experiment metrics to which we compare, if we allow for (relatively) small changes in  $N_d$  or the cyclonic/anticyclonic ratio.

<sup>18</sup> Note that we do have to enforce a truncation where  $\mathcal{R} > -1$  to prevent an unphysical negative density. However because the  $|\delta \ln \rho|$  along individual ‘steps’ is small, this has only a small effect on the predictions.

### 5.1.2 Self-driven (Streaming & Kelvin–Helmholtz) turbulence

Next, Fig. 4 repeats this comparison, with a different set of simulations from Johansen & Youdin (2007) and Bai & Stone (2010a,b). These are two- and three-dimensional simulations, of non-MHD hydrodynamic shearing boxes but ignoring vertical gravity/stratification (so we can directly take the statistics in  $\rho_p/\langle\rho_p\rangle$  as representative of our predictions). The simulations fix  $\eta = 0.005$ ,  $\Pi = 0.05$ ,  $\Omega$ , and  $\tau_s$  for monolithic grain populations, and have no external driving of turbulence. However, they do include the grain-gas back-reaction for  $\tilde{\rho} = 0.2, 1.0, 3.0$ , and so develop some turbulence naturally via a combination of streaming and Kelvin–Helmholtz-like shear instabilities, with  $\alpha \sim 10^{-8}$ – $10^{-2}$  depending on the simulation properties (but recorded for each simulation therein). The two studies adopt entirely distinct numerical methods, so where possible we show the differences owing to numerics.

In nearly every case, we see good agreement with our simple mathematical model. This is especially true at smaller  $\tau_s$  and  $\rho_p$ ; at the highest absolute grain densities, our model is less applicable but still performs reasonably well. In particular, the non-linear behaviour seen therein, where for example there is a large difference between  $\rho_p = 0.2$ – $1$  for  $\tau_s = 0.1$  (but little change between  $\rho_p = 1$ – $3$ ), and the much larger fluctuations seen for  $\rho_p = 0.2$  compared to  $\rho_p = 1$ – $3$  for  $\tau_s = 1$ , are all predicted. The large increase with  $\rho_p = 0.2$ – $1$  for  $\tau_s = 0.1$  follows from  $\rho_p$  increasing the ‘effective’ stopping time as discussed in Section 3.4, and then the effect saturates with increasing  $\rho_p \gtrsim 1$ . However much of the difference also owes to the different values of  $\alpha$  in each simulation (the case  $\tau_s = 1$ ,  $\rho_p = 0.2$  produces a very large  $\alpha$ , driving much of the very large variance).

If we repeat the experiments from Section 5.1.1, we come to the same conclusions. For example, adopting the pure log-Poisson model (mean  $\langle\delta \ln \rho\rangle$ ) makes little difference compared to the log-Poisson–Gaussian model discussed in Section 5.1.1 above. However, we see more clearly here that allowing for additional variance in the effects of an individual eddy does, as one might expect, increase the variance at the high- $\rho_p$  tail of the distribution, whereas a ‘strict’ log-Poisson model (our default model) has an absolute cutoff at some  $\rho_{p,\max}$ . Interestingly, this gives a slightly better fit at high  $\tau_s$ , and poorer at low  $\tau_s$ , perhaps indicating the relative importance of in-eddy variance in these two cases.

We should also note that the finite simulation resolution limit is important here when  $\tau_s = 0.1$  and  $\tilde{\rho} \ll 1$ , we will quantify this below.

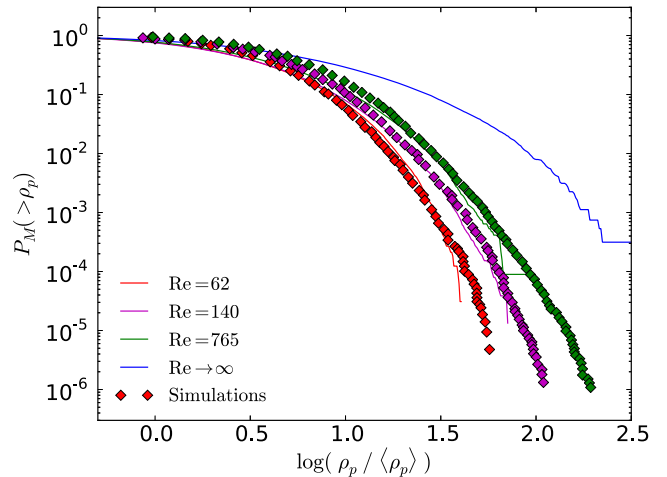
Interestingly, the streaming instability relies on the cooperation of the rotation (shear terms), radial drift (non-zero  $\eta$ ), and the back-reaction of grains on gas, in order to develop a growing mode. Although we do attempt to include a simplistic treatment of these effects, the qualitative behaviour of our model does not directly rely on a drift term – i.e. one could imagine applying it with zero  $\eta$  and obtaining a similar result. This is because the model we apply here is really a model for the grain density fluctuations in the non-linear phase of evolution with fully developed turbulence. We are ‘just’ modelling what the turbulence, once present in some saturated amplitude, does to the grains, rather than the cooperation of the grains and gas in driving the turbulence to grow in the first place. As a result, if we took the *initial* conditions of the streaming instability experiments above (where there is no turbulence) our instantaneous prediction would be that there are no grain density fluctuations. And indeed, this would be instantaneously correct – but the point is that the model here cannot predict whether such fluctuations *should* develop from a smooth initial condition, because

we are not attempting to predict or model the time-evolution of the turbulence. Rather, we rely on some other model to tell us the growth history and state of the turbulence, then simply apply this to obtain an estimate of the ensuing grain clustering statistics. As a result, what the agreement here suggests – although it clearly merits further study in future numerical simulations – is that the physics unique to the streaming instability may not qualitatively change the key statistical properties of grain density fluctuations, once it powers turbulence at a given level.

### 5.1.3 Turbulent concentration

In Fig. 5, we now compare the density PDFs measured in ‘turbulent concentration’ simulations in Hogan et al. (1999); these simulations follow a driven turbulent box (no shear or self-gravity), so we should apply the version of the model from Section 3.2.2. We expect the same  $N_d$  and  $C_\infty$ , and back-reaction is not included so  $\tilde{\rho} \rightarrow 0$ . We then need to know over what range to integrate the cascade: this is straightforward since each simulation has a well-defined Reynolds number  $\text{Re} = (\lambda_{\max}/\lambda_v)^{4/3}$ . Lacking a model for the dissipation range, we simply truncate the power exponentially when  $\lambda < \lambda_v$ . The simulation follows particles with Stokes numbers  $\text{St} \equiv t_s/t_e(\lambda_v) = 1$ .

Perhaps surprisingly, the model agrees fairly well with the simulations. At larger  $\text{Re}$ , the density PDF becomes more broad, because of contributions to fluctuations over a wider range of scales (the response function in Fig. 1 is broad). However, this does not grow indefinitely – as  $\text{Re} \rightarrow \infty$  we predict convergence to a finite PDF width (with  $\rho_{p,\max} \sim 300$ – $1000$ ). This is both because the response function declines, and, as the ‘top’ of the cascade becomes larger in



**Figure 5.** Grain density distribution, as Fig. 2, for simulations of ‘turbulent concentration’ from Hogan, Cuzzi & Dobrovolskis (1999). Here there is no shear/gravity (see Section 3.2.2), and the flow is simulated from a fixed viscous scale to various Reynolds numbers ( $\text{Re} = 62, 140, 765$ ). In each case, the Stokes number is unity ( $t_s \approx t_e(\lambda_v)$ , where  $\lambda_v$  is the viscous scale); this gives  $\tilde{\tau}_s(\lambda = \lambda_{\max})$  at the top of the cascade of  $\tilde{\tau}_s(\lambda_{\max}) \approx 0.13, 0.03, 0$  ( $\text{Re} = 62, 765, \infty$ ). The PDF width grows logarithmically with  $\text{Re}$  as we integrate over more of the broad response function in Fig. 1. However, the decay in this function at  $\tilde{\tau}_s \ll 1$ , and the increase in rms turbulent velocities of grains lowering their eddy crossing times, means that the PDF does not grow indefinitely as  $\text{Re} \rightarrow \infty$ . For comparison, pure uncorrelated (Markovian) fluctuations predict a PDF with dispersion in  $\rho_p$  of  $\approx 1.7$ , giving a PDF that falls below the minimum plotted range here at  $\log(\rho_p/\langle\rho_p\rangle) \approx 0.8$  (Zaichik & Alipchenkov 2009).

velocity scale, the residual (logarithmically growing) offset between grain and eddy velocities becomes larger (equation 40), suppressing the added power.

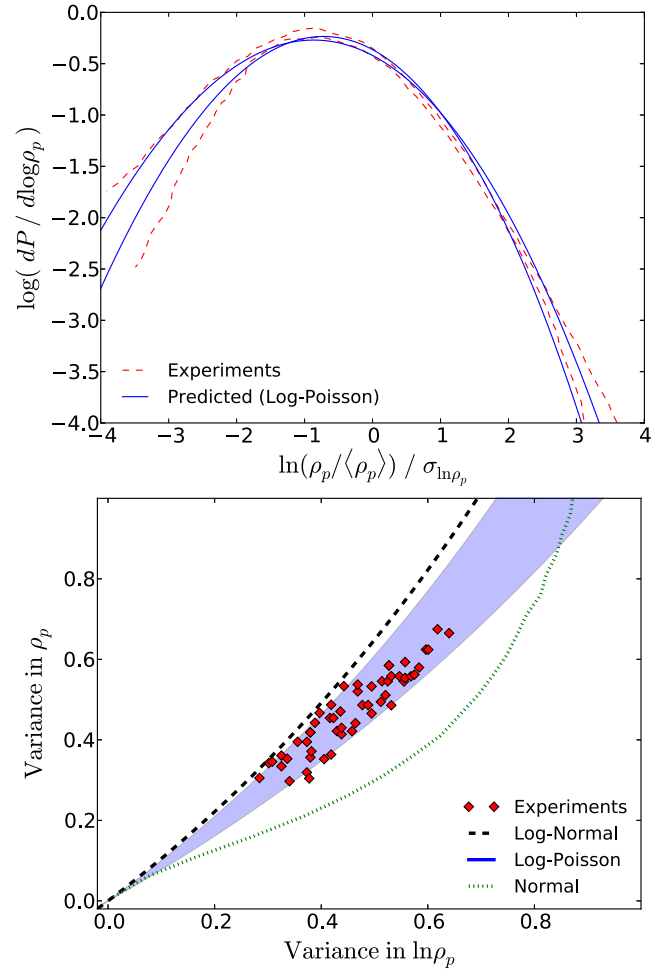
Note, though, that our model is not designed for small  $St$ , and we see the effect here. The highest density tail of the PDF is not fully reproduced (the model predictions, especially at  $Re = 765$ , cut off more steeply). We show below that this is because grains with small  $St \sim 1$  can continue to cluster and experience strong density fluctuations on very small scales  $\lambda \lesssim \lambda_v$ , which are not accounted for in our calculation. We also do not explicitly account for the role of the strain field in enhancing grain clustering outside of regions of high vorticity; this may be fine in the previous cases where high-density fluctuations are dominated by anticyclonic vortices, but needs further investigation in the cases here. In any case, we caution care for particles where the key fluctuations lie outside the inertial range.

#### 5.1.4 Experiments and non-Gaussianity

In Fig. 6, we extend our comparisons to experimental data. There is a considerable experimental literature for  $St \lesssim 1$  particles in terrestrial turbulence (see Section 1); unfortunately many of the measurements are either in regimes where our model does not apply or of quantities we cannot predict. However, Monchaux et al. (2010) measure the density PDF in laboratory experiments of water droplets in wind tunnel turbulence,<sup>19</sup> with  $St \sim 0.2-6$  and  $Re \sim 300-1000$  (Taylor  $Re_\lambda = 70-120$ ). They measure the PDF shape (normalized by its standard deviation) for a large number of experiments with different properties. The range of results, including time variation and variation across experiments, is shown in Fig. 6. We compare this with the predicted log-Poisson distribution: from Section 4.1, the variance in the log-Poisson is  $S = \Delta N_{\text{int}} |\delta \ln \rho|_{\text{int}}^2$ , so normalizing to fixed  $\sigma = \sqrt{S}$ , the PDF shape varies with the ratio  $|\delta \ln \rho|_{\text{int}} / \Delta N_{\text{int}}$ . Taking the predicted (modest) range in this parameter for the same range in simulation properties, we show the predicted PDF shapes. Within this range, the experimental PDF is consistent with the prediction.

The lower panel makes this more quantitative. For each PDF in Monchaux et al. (2010), we record the variance in linear  $\rho_p$  ( $S_\rho$ ) and logarithmic  $\ln \rho_p$  ( $S_{\ln \rho_p}$ ). This scaling for different distributions is discussed in detail in Hopkins (2013b, see Fig. 4 in particular). If the distribution of  $\rho_p$  were exactly lognormal, then there is a one-to-one relation between the two:  $S_{\ln \rho_p} = \ln(1 + S_\rho)$ . This appears to form an ‘upper envelope’ to the experiments. If  $\rho_p$  is distributed as a Gaussian in linear- $\rho_p$ , there is also a one-to-one relation (straightforward to compute numerically); this predicts relatively small  $\ln \rho_p$  variation, in conflict with the simulations. For the log-Poisson distribution, the relation depends on the second parameter  $|\delta \ln \rho|_{\text{int}} / \Delta N_{\text{int}}$ . As this  $\rightarrow 0$ , the distribution becomes lognormal; for finite values,  $S_\rho$  is smaller than would be predicted for a lognormal with the same  $S_{\ln \rho_p}$ ; we compare the range predicted for plausible values of  $|\delta \ln \rho|_{\text{int}}$  in these experiments (similar parameters to the simulations in Fig. 8).

We could extend this comparison by including the numerical simulations at higher  $t_s$  (and including gravity and shear). But the agreement with our predictions is already discussed, and it is evident by-eye that the distributions in Fig. 4 are not exactly lognormal



**Figure 6.** Density PDF in laboratory experiments of water droplets in wind tunnel turbulence (Monchaux et al. 2010). Top: range of particle density PDFs obtained (dashed), normalized by their variance. We compare the predicted log-Poisson distribution, with the same variance and a range of  $|\delta \ln \rho|_{\text{int}} \sim 0.01-0.2$  corresponding to model predictions. Bottom: test of lognormality: We compare the variance in  $\rho_p$  to that in  $\ln \rho_p$  from the same experiments, to the range predicted for log-Poisson PDFs with the predicted range in  $|\delta \ln \rho|_{\text{int}}$ , the prediction from a lognormal distribution, and from a normal (Gaussian) distribution. The experiments favour a log-Poisson distribution as opposed to a linear Gaussian distribution or lognormal.

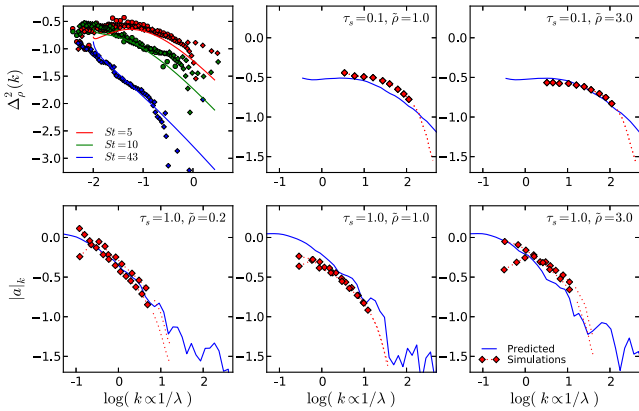
(they are asymmetric in log-space about the median), nor can they be strictly Gaussian in linear  $\rho$  (which for such large positive fluctuations would require negative densities). While it is less obvious by eye, the non-trivial fractal spectrum in the turbulent concentration experiments, discussed at length in Cuzzi et al. (2001), also requires non-Gaussian PDFs.

#### 5.2 Density power spectra

Direct measurements of the power spectrum of  $\ln \rho$  are not available for the simulations we examine here. However, Johansen & Youdin (2007) do measure the average one-dimensional power spectra of the linear grain volume density  $\rho_p$ .<sup>20</sup> We can compute this as

<sup>19</sup> Actually they measure the local Voronoi area around each particle: as noted therein, this is strictly equivalent to a local density PDF. We convert between the two as they do, taking the density to be the inverse area.

<sup>20</sup> We exclude their simulation ‘AA’ which uses the two-fluid approximation, that the authors note cannot capture the full gas-grain cascade and so predicts an artificially steep power spectrum.



**Figure 7.** Power spectra of linear grain density fluctuations. We compare the streaming-instability simulations in Fig. 4 for  $\tau_s = 0.1$ –1, and the turbulent concentration simulations (top left) in Yoshimoto & Goto (2007, squares) and Pan et al. (2011, diamonds) for Stokes numbers  $\sim 5, 10, 50$ . Simulation results within five cells of the resolution limit are shown as dashed lines; the power suppression here is artificial. Agreement is good. For large particles, the power is preferentially concentrated on large scales. For small particles, the power spectrum is quite flat, until a turnover at scales below which  $t_e \ll t_s$ .

described in Section 4.2, and show the results in Fig. 7.<sup>21</sup> Down to the simulation resolution limit (where the simulated power is artificially suppressed) we see good agreement. Because the power spectrum here is in linear  $\rho_p$ , saturation effects dilute the clarity of the predicted transition near  $t_e = t_s$ , but it is still apparent. Moreover, we confirm the qualitative prediction that for large  $\tau_s \gtrsim 0.1$ , most of the power is on relatively large scales where  $t_e \gtrsim t_{\text{cross}}$  and  $t_s$ . With smaller  $t_s$ , there is a larger dynamic range on large scales where  $t_e \gg t_s$ , over which the power spectrum is flatter. For very small grains, the power would become more concentrated near  $t_e \sim t_s$ , as in Fig. 1.

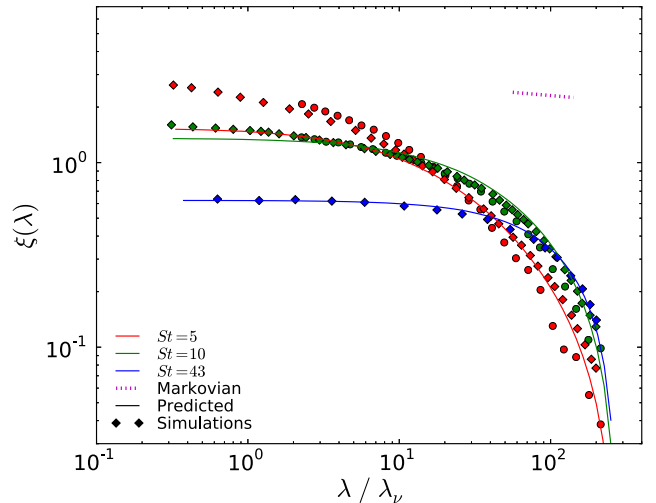
Freeing  $C_\infty$  and  $N_d$ , the simulations with  $\tau_s = 1$  and  $\tilde{\rho} \gtrsim 1$  try to fit a slightly steeper slope compared to that predicted in the default model, but this is not very significant at present (and no significant change in  $N_d$  is favoured). If confirmed, though, this might imply that the gas turbulence in this regime behaves more like compressible, supersonic turbulence (see Boldyrev 2002; Schmidt et al. 2008); this might be expected when  $\tau_s$  and  $\rho_p$  are large, since the dominant grains can efficiently compress the gas.

We also compare the power spectra from turbulent concentration simulations. This is the same information as contained in the correlation function, converted via equation (69), so we discuss it below.

### 5.3 Grain correlation functions

In Fig. 8, we compare our predictions to published grain correlation functions  $\xi$  in turbulent concentration experiments. The same information, represented as the linear density power spectrum, is in Fig. 7. Here, we compare the simulations from Pan et al. (2011) and

<sup>21</sup> To match what was done in that paper precisely, we calculate the volumetric density PDF and corresponding variance in linear  $\rho$  ( $S_\rho$ ) explicitly for each scale, use this to obtain the isotropic power spectrum, then use this to realize the density distribution repeatedly on a grid matching the simulations and compute the discrete Fourier transform, and finally plot the mean absolute magnitude of the coefficients as a function of  $k$ .



**Figure 8.** Radial grain correlation functions, for the same simulations in Fig. 7. The predictions agree well at  $St \gg 1$  and/or scales  $\lambda \gg \lambda_v$ , but the clustering is underpredicted at  $\lambda \lesssim \lambda_v$  for  $St \sim 1$ , owing to non-inertial range effects we do not include. The shallow dotted line shows the slope predicted for Markovian (pure random-field) fluctuations (the amplitude is below the range plotted) in the inertial range ( $\lambda \gg \lambda_v$ ), following Bec et al. (2007): uncorrelated/incoherent fluctuations lead to negligible clustering when  $St \gg 1$ .

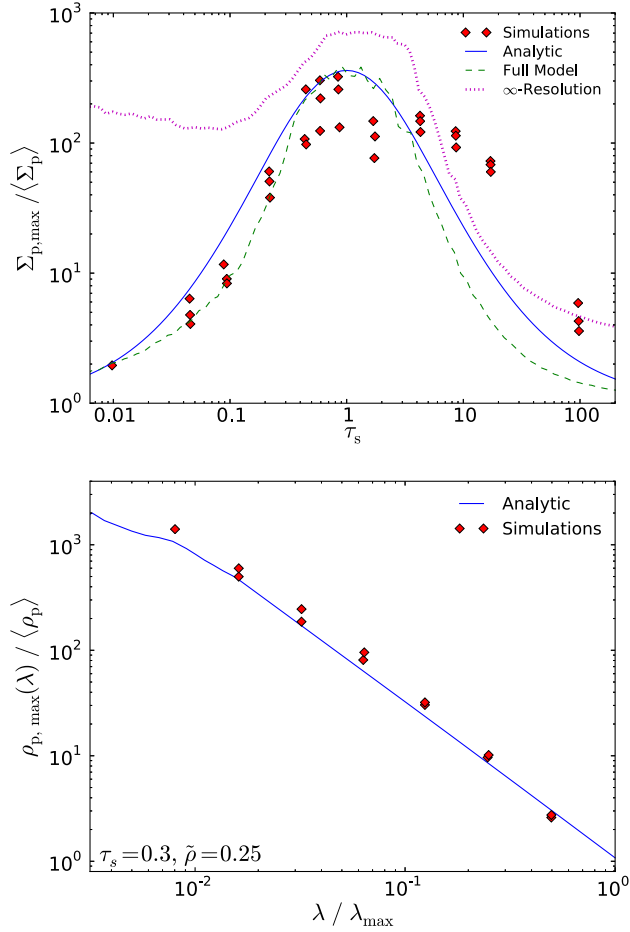
Yoshimoto & Goto (2007), with the model appropriate for ‘pure turbulence’ (see Section 5.1.3 above).<sup>22</sup> The authors each simulate a range of Stokes numbers; here we only compare with  $St > 1$  simulations since our model is largely inapplicable to  $St \lesssim 1$ .

For large  $St = 43 \gg 1$ ,  $\xi(\lambda) \sim \text{constant}$  on small scales (there is no power here since  $t_e \ll t_s$ ). But there is significant power on larger scales, where  $t_e \sim t_s$  (for  $St = 43$ , this is when  $\lambda/\lambda_v \gtrsim 100$ ). And  $\xi(\lambda)$  in all cases truncates at the very largest scales because of the finite box size/driving scale  $\lambda_{\text{max}}$ . For smaller  $St = 10$ , the rising portion of  $\xi(\lambda)$  continues to smaller scales, since  $t_e \sim t_s$  at  $\lambda/\lambda_v \sim 40$ . For still smaller  $St = 5$  this extends to  $\lambda/\lambda_v \sim 10$ . These are all confirmed in the simulations. However, for the smallest  $St \sim 1$ ,  $\xi(\lambda)$  (and hence the power in density fluctuations) continues to rise even at  $\lambda \lesssim \lambda_v$ , where  $t_e \ll t_s$ . This is well-known, and in fact for  $St \sim 1$  a power-law rise in  $\xi(\lambda)$  appears to continue to  $\lambda \rightarrow 0$ , which does not occur when  $St \gg 1$  (see Squires & Eaton 1991; Bec et al. 2007; Yoshimoto & Goto 2007; Pan et al. 2011). This effect is fundamentally related to the dissipation range and viscous effects not included in our model, so we do not expect to capture it (and, for example, trapping of grains in strain regions outside of vortices). And this is why we do not reproduce the full small-scale power in the density PDFs for  $St = 1$  in Fig. 5.

For large-scale eddies and grains with  $\tau_s \gtrsim 0.1$ , we can compare with the simulations in Carballido, Stone & Turner (2008). The results are consistent, but due to limited resolution the simulations only measure significant clustering in the couple smallest bins/cells quoted (see their fig. 9); so the constraint is not particularly useful (significantly more information is available in Fig. 7).

<sup>22</sup> Perhaps because of different definitions, the normalizations for the correlation functions  $\xi$ , at identical Stokes and Reynolds number, disagree between the authors at the factor  $\sim$  few level. However, the shape of  $\xi$  in all cases agrees extremely well in both studies. So we treat the normalization as arbitrary at this level and focus on the shape comparison.





**Figure 9.** Top: maximum grain density measured in the MRI simulations from Fig. 2, as a function of population stopping time  $\tau_s$ . Different simulations with various box sizes and resolution are plotted, the predictions should form an ‘upper envelope.’ We compare our simple approximation for large eddies and full prediction, given the (finite) simulation resolution and turbulence properties. These agree very well up to  $\tau_s \sim$  a few, though they underpredict fluctuations when  $\tau_s \gg 1$ . On large scales, maximum densities increase rapidly with  $\tau_s$  up to  $\tau_s \sim 1$ . We also show the prediction if the simulation were infinitely high resolution (densities measured on arbitrarily small scales). In this regime,  $\tau_s \ll 1$  grains also show large fluctuations; however, these high densities are manifest on very small scales. Roughly, convergence to this solution requires resolving eddies with  $t_e \gtrsim 0.05 t_s$ ; for the smallest  $t_s$  and given simulation turbulence properties and box sizes here, this would require a minimum  $\sim (10^6)^3$ -cell simulation. Bottom: maximum grain density in simulations from Johansen, Youdin & Lithwick (2012), averaged on different smoothing scales  $\lambda$ . For large grains, the predicted maximum increases on smaller smoothing scales  $\propto \lambda^{-(1-2)}$ .

## 5.4 Maximum grain densities

### 5.4.1 Dependence on stopping time

Fig. 9 compares these predictions for the maximum grain concentration to the maximum measured in the MRI-unstable simulations from Section 5.1.1 (Dittrich et al. 2013). Recall, here  $\alpha$  and  $\Pi$  are approximately constant in all cases, and grain-gas back-reaction is ignored ( $\tilde{\rho} \rightarrow 0$ ), so the only varied parameter is  $\tau_s$ . In the range  $\tau_s \sim 0.01$ –1, our predictions are in remarkably good agreement with the simulations, with the maximum grain concentration increasing from tens of per cent to factors  $\gtrsim 300$ , for the simple assumption  $N_d = 2$ . Only a small range  $N_d \sim 1.8$ –2.2 is allowed if we free this

parameter. For  $1 < \tau_s \lesssim 5$ , the predictions are also reasonably accurate. At very large  $\tau_s \gtrsim 10$ ; however, we appear to underestimate the magnitude of fluctuations; though it also appears that there is some change in the vertical structure in these simulations relative to what is expected (discussed in Dittrich et al. 2013), so the large  $\Sigma_{p,\max}$  may not entirely reflect mid-plane density fluctuations.

As noted above, it is important that we account for finite resolution here. We compare the predictions using our best estimate of  $\lambda_{\max}$  (the driving scale) relative to the finite resolution limit (factor of  $\sim 100$  in scale), to the prediction assuming infinite resolution (and  $\text{Re} \rightarrow \infty$ ), with density measured on infinitely small scales. For large grains, this makes little difference (most power is on large scales). For small grains  $\tau_s \ll 1$ , however, the difference is dramatic. Very small grains with  $\tau_s \sim 0.01$  may still experience factor  $\gtrsim 100$  fluctuations on small scales. This should not be surprising, however – this is already evident in the turbulent concentration simulations, which exhibit such large fluctuations (even over limited Reynolds number, but covering the range where  $t_s \sim t_e$ ) despite  $\tau_s \rightarrow 0$ , effectively. These  $\rho_p$  fluctuations, for small grains, occur on scales where  $t_s \sim t_e$ , and resolving their full dynamic range (getting convergence here) requires resolution of the broad peak in the response function (Fig. 1), crudely we estimate  $0.05 t_s \lesssim t_e \lesssim 20 t_s$  should be spanned. This translates, even in idealized simulations, to large Reynolds numbers  $\text{Re} \gtrsim 10^4$ – $10^5$  (not surprising, since the simulations in Fig. 5 are not converged yet at  $\text{Re} \sim 1000$ ). And for the simulations here, which have a fixed box size at of the order of the dust layer scaleheight, this would require resolution of a factor  $\sim 10^6$  below the largest eddies scales  $\lambda_{\max}$  (far beyond present capabilities).

Interestingly, we predict a ‘partial’ convergence: because  $|\delta \ln \rho|$  is non-monotonic in  $\lambda$ , the fluctuations on large scales can converge at reasonable resolution (factor  $\sim 2$  changes relative to the simulations here make little difference to the predicted curve). Only when the resolution is increased by the much larger factor described above does the additional power manifest. So, if the ‘interesting’ fluctuations are those on large scales, such simulations, or the approximations in Table 2, are reasonable.

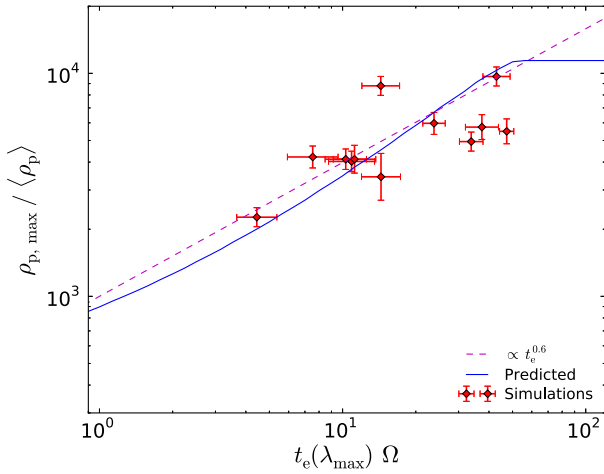
### 5.4.2 Dependence on scale

Johansen et al. (2012) present the maximum density as a function of scale in streaming-instability simulations similar to those in Johansen & Youdin (2007).<sup>23</sup> Given  $\alpha$ ,  $\Pi$ ,  $\tilde{\rho}$  and  $\tau_s = 0.3$  specified in the simulation, it is straightforward to predict  $\rho_{p,\max}(\lambda)$  and compare to their result (using  $\rho_p(\lambda_{\max}) = \langle \rho_p \rangle$  at the dust scale-height). As discussed in Section 4.4 and in Table 2, on large scales  $\rho_{p,\max} \propto \lambda^{-\gamma}$  with  $\gamma \approx C_\infty [1 - \exp(-|\delta_0|)]$ ; for  $\tau_s = 0.3$  and  $\tilde{\rho} = 0.25$ , this gives  $\gamma \approx 1.5$ , a power-law like scaling in excellent agreement with the simulations.

### 5.4.3 Dependence on eddy turnover time

In the simulations of Dittrich et al. (2013) from Fig. 9, the authors also note that in a separate series of simulations with fixed  $\tau_s = 1$ , they see a significant dependence of the maximum  $\rho_p$  on the lifetime/coherence time of the largest eddies. They quantify this by comparing  $\rho_{p,\max}$  to (twice) the correlation time of the longest lived Fourier modes, which should be similar to our  $t_e(\lambda_{\max})$ . As discussed

<sup>23</sup> We specifically compare their simulations with no collisions and no grain self-gravity.



**Figure 10.** Dependence of the maximum grain density in the  $\tau_s = 1$  MRI simulations from Figs 2 and 9 on the correlation time of the largest eddies in the simulation box ( $t_e(\lambda_{\max})$ ). Because of differences in the definition of correlation time, and the unknown vertical sedimentation, we treat the normalization of each axis as arbitrary: what matters here is the predicted trend. At fixed numerical resolution,  $\rho_{p, \max}$  increases with  $t_e(\lambda_{\max})^{0.4-0.7}$  when  $t_e(\lambda_{\max}) \lesssim 1$ , until saturating (with the  $t_e(\lambda_{\max})$ -independent scalings in Table 2) when  $t_e(\lambda_{\max}) \gg 1$ .

in Appendix B, on sufficiently large scales ( $t_e \gg \Omega^{-1}$ ),  $|\delta \ln \rho|$  and fluctuation properties asymptote to values independent of  $t_e$  (the scalings in Table 2). However for smaller  $t_e(\lambda_{\max}) \lesssim 1$ , since the power for large grains  $\tau_s \sim 1$  is concentrated on scales with  $t_e \sim t_s \sim \Omega^{-1}$ , the integrated power will decline if the top scales only include smaller eddies. Given the asymptotic scaling of  $\overline{\varpi} \propto \tilde{\tau}_s^{-1/2}$  for  $\tilde{\tau}_s \gg \tau_s$  ( $t_e \ll \Omega^{-1}$ ) and  $\tilde{\tau}_s \gg 1$  ( $t_e \ll t_s$ ), we expect that the power at the largest scales will scale  $\propto \tilde{\tau}_s(\lambda_{\max})^{-1/2} \propto t_e(\lambda_{\max})^{1/2}$ . Performing the full calculation for comparable Reynolds number to the simulations, we indeed predict a scaling  $\rho_{p, \max} \propto (t_e(\lambda_{\max}) \Omega)^{0.4-0.7}$  (for a range  $\tau_s \sim 0.5-2$ ), until saturation. Fig. 10 explicitly compares this to the scaling in those simulations; the agreement is good.

#### 5.4.4 Effects of simulation dimension

Note that, in Johansen & Youdin (2007) and Bai & Stone (2010a,b), some significant differences are found between two-dimensional and three-dimensional simulations. To first order, the differences are accounted for in our model, not because of a fundamental change in the behaviour of grains in response to eddies, but rather because of the different amplitudes of turbulence and/or vertical stratification in the simulations. In three dimensions, it appears more difficult for the streaming instability to generate large- $\alpha$  turbulence. That result itself is not part of our model here, however, for the given  $\alpha$  in the different simulations, our predictions appear to agree with the different simulation PDFs.

#### 5.4.5 Effects of pressure gradients and metallicity

Pressure gradients ( $\Pi$ ) and metallicity ( $Z$ ) enter our model only indirectly, by altering the value of the parameters  $\tilde{\rho}$  and  $\beta$ , and – under some circumstances such as streaming-instability turbulence – by altering the bulk turbulent properties (velocities/eddy turnover times). We have already derived the dependence of  $\beta$  on  $\alpha$ ,  $\tilde{\rho}$ , and  $\Pi$ . And we expect  $\tilde{\rho} = (\Sigma_p / \Sigma_{\text{gas}})(h_{\text{gas}} / h_p) \approx Z(\tau_s / \alpha)^{1/2}$  (see Carballido, Fromang & Papaloizou 2006; Youdin & Lithwick 2007).

When other parameters are fixed (for example, if the turbulence is externally driven), it is straightforward to accommodate these parameter variations. However, the dependence of bulk turbulent properties  $\alpha$  and  $t_e(\lambda_{\max})$  on  $Z$  and  $\Pi$  (or other disc properties) requires some additional model for the driving and generation of turbulence. Bai & Stone (2010b) consider a survey of these parameters, in the regime where the turbulence is driven by the streaming instability, and find that the dust layer scaleheight, turbulent dispersion  $\alpha$ , and largest eddy scales depend in highly non-linear and non-monotonic fashion on  $Z$  and  $\Pi$ . If we adopt some of the simple dimensional scalings they propose therein for these quantities, we qualitatively reproduce the same trends they see:  $\rho_{p, \max}$  increases with both increasing  $Z$  (increasing  $\tilde{\rho}$ ) and decreasing  $\Pi$  (weaker drift, larger  $\beta$ ), but it is difficult to construct a quantitative comparison.

## 6 DISCUSSION

### 6.1 Summary

We propose a simple, analytic, phenomenological model for the clustering of aerodynamic grains in turbulent media (with or without external shear and gravity). We show that this leads to unique, definite predictions for quantities such as the grain density distribution, density fluctuation power spectrum, maximum grain densities, and correlation functions, as a function of grain stopping/friction time, grain-to-gas volume density ratio, and properties of the turbulence. Our predictions are specifically appropriate for inertial-range turbulence, with large Reynolds numbers and Stokes numbers ( $t_s$  large compared to the eddy turnover time at the viscous scale), the regime of most astrophysical relevance. Within this range, we compare these predictions to numerical simulations and laboratory experiments, with a wide range in stopping times and turbulence properties, and show that they agree well.

The model assumes that grain density fluctuations are dominated by coherent turbulent eddies, presumably in the form of simple vortices. Such eddies act to accelerate grains and preferentially disperse them away from the eddy centre (concentrating grains in the interstices between eddies). Qualitatively, such behaviour has been observed in a wide range of simulations and experiments (see Section 1). Quantitatively, our model first adopts a simple calculation of the effects of a vortex with a given eddy turnover time (and lifetime of the same duration) acting on an initially homogeneous, isotropic Lagrangian grain population. We then attach this calculation to a simple assumption for eddy structures on different scales: namely that eddies on different scales are self-similar, statistically independent, and reproduce a Kolmogorov-type scaling.

### 6.2 Key conclusions and predictions

(i) *Large grain density fluctuations are expected even in incompressible turbulence.* We predict that even a small aerodynamic de-coupling between gas and grains allows for large (order-of-magnitude) fluctuations in  $\rho_p$ , even while gas density fluctuations are negligible.

(ii) *Grain density fluctuations do not explicitly depend on the driving mechanisms of turbulence.* Given the simplistic level of detail in our model, it applies equally to simulations with turbulence arising via MRI, Kelvin–Helmholtz, and streaming instabilities, or artificially (numerically) driven. Still, even with limited accounting for the detailed structure of turbulence, we are able to predict some dependence of fluctuations on the stopping time  $t_s$ , the ratio of volume-averaged grain-to-gas densities  $\tilde{\rho}$ , and

some basic properties of the turbulence (the Reynolds number and velocity/length/time-scale at the driving scale). These can, of course, change depending on the driving.

(iii) *The grain density distribution  $\rho_p$  is log-Poisson*

$$P_V(\ln \rho_p) d \ln \rho_p \approx \frac{\Delta N_{\text{int}}^m \exp(-\Delta N_{\text{int}})}{\Gamma(m+1)} \frac{d \ln \rho_p}{|\delta \ln \rho|_{\text{int}}}$$

$$m = |\delta \ln \rho|_{\text{int}}^{-1} \left\{ \Delta N_{\text{int}} [1 - \exp(-|\delta \ln \rho|_{\text{int}})] - \ln \left( \frac{\rho_p}{\langle \rho_p \rangle} \right) \right\}. \quad (72)$$

We predict  $\Delta N_{\text{int}}$  and  $|\delta \ln \rho|_{\text{int}}$  as a function of turbulent properties. This arises (in the model here) because the number of eddies is quantized (Poisson), and each produces a multiplicative (logarithmic) effect on the grain density field.

Generically, we suggest that this can be used as a fitting function, where the best-fitting value of  $\Delta N_{\text{int}} \sim C_\infty \ln(\lambda_{\text{max}}/\lambda_{\text{min}})$  crudely measures the dynamic range of the cascade over which density fluctuations occur, and the value  $|\delta \ln \rho|_{\text{int}}$  reflects the rms fluctuation amplitude ‘per event’ in the turbulence.

(iv) *On large scales ( $t_e \gtrsim \Omega^{-1}$ ) shear/gravity dramatically enhances density fluctuations.* In this model, the fluctuation ‘response’ in large ( $t_e \gtrsim \Omega^{-1}$ ) eddies is approximately scale-free, with amplitude  $|\delta \ln \rho| \sim 2 N_d (\tau_s + \tau_s^{-1})^{-1}$ . The variance in  $\ln \rho_p$ , and maximum values of  $\rho_p$ , increase with  $\tau_s$  up to a maximum near  $\tau_s \sim 1$  (where maximum  $\rho_p$  values can reach thousands of times the mean); much larger grains are too weakly coupled to experience fluctuations and behave in an approximately ‘free-steaming’ manner.

The maximum  $\rho_p(r)$  on a smoothing scale  $r$  scales  $\propto r^{-\gamma}$  with  $\gamma \sim C_\infty [1 - \exp(-|\delta \ln \rho|)]$ ; for small grains with  $|\delta \ln \rho| \ll 1$   $\gamma \sim 2 |\delta \ln \rho|$  is small, so the scale-dependence is shallow. For large grains with  $|\delta \ln \rho| \gtrsim 1$ ,  $\gamma$  saturates at  $\sim 2$  (isothermal-like). Most of the power in  $\rho_p$  fluctuations is on large scales for large grains, while for small grains the power spectrum is approximately flat over a range of scales down to a scale  $\lambda_{\text{crit}}$  where the rms eddy-crossing time becomes shorter than the grain stopping time, below which power is suppressed.

(v) *On small scales ( $t_e \ll \Omega^{-1}$ ), grain clustering depends only on the ratio  $t_s/t_e$ .* Within the context of our model, the fluctuation amplitude in small eddies is maximized around  $t_s \sim t_e$ , declining  $\propto t_s/t_e$  for  $t_s \ll t_e$  (where eddies are ‘flung out’ of vortices at speeds limited by the eddy terminal velocity  $\propto t_s$ ) and  $\propto (t_s/t_e)^{-1/2}$  for  $t_s \gg t_e$  (where eddies cannot fully trap grains, so their effects add incoherently in a Brownian random walk).

Integrated over a sufficiently broad cascade (Reynolds number  $\rightarrow \infty$ ), this means that some eddies will always have  $t_s \sim t_e$ , so the integrated density variance and maximum  $\rho_p$  always converge to values only weakly dependent on the absolute value  $t_s$ . The maximum  $\rho_p$  can reach several hundred times the mean grain density, even in the limit  $\rho_p \ll \rho_g$  and  $\tau_s \ll 1$ . The variance is concentrated on small scales, however, and the ‘resonance region’ of eddy turnover time is broad – so resolving this in simulations or experiments requires resolved eddies at least over the range  $0.05 t_s \lesssim t_e \lesssim 20 t_s$  (Reynolds numbers at least  $\gtrsim 10^4$ – $10^5$ ).

The grain–grain correlation function  $\xi(r)$  in this limit scales weakly on the largest scales ( $\propto \ln(1/r)$ ), until approaching  $t_e \sim t_s$  where it rises as  $\xi(r) \propto (t_s/t_e)^2 \propto r^{-2(1-\xi)}$  (a slope near unity), then converges (flattens to  $\xi(r) \rightarrow \text{constant}$ ) below  $t_e \lesssim t_s$ .

(vi) *Stronger turbulence enhances clustering.* At otherwise identical properties, larger values of the Mach number, Reynolds number, or driving-scale eddy turnover time  $t_e(\lambda_{\text{max}})\Omega$  give rise to a larger dynamic range of the cascade driving  $\rho_p$  fluctuations.

Stronger turbulence may decrease  $\langle \rho_p \rangle$ , so it is not necessarily the case that these lead to larger absolute maximum values of  $\rho_p$ , but only stronger grain clumping. Conversely, larger *drift* (laminar relative grain-gas velocities) weakens the clustering, by suppressing the time grains interact with single eddies.

(vii) *Higher grain-to-gas density ratios enhance clustering of small grains.* We attempt to consider some very simple approximations to account for the regime where the back-reaction of grains on gas is important. Remarkably, the predictions of the model here appear to reasonably describe simulations in this limit (despite many of our assumptions formally breaking down). Bearing these caveats in mind, this predicts that, up to a saturation level where  $t_s(1+\bar{\rho}) \gtrsim 1$ , increasing the volume density of grains increases their effective stopping time by dragging gas in a local wake, leading to larger terminal velocities and eddy effects.

(viii) *Coherent eddy structure is critical.* Our predictions rely fundamentally on locally coherent (albeit short-lived) structures in turbulence. We show that a purely Markovian (Gaussian random field) approximation does not produce fluctuations nearly as large (nor with the correct scaling). So in some sense inertial-range  $\rho_p$  fluctuations depend intrinsically on structure in the gas turbulence. This may not necessarily be captured in models which treat density perturbations purely as a ‘turbulent diffusion’ term or Brownian motion.

### 6.3 Limitations of this model & areas for future work

This paper is intended as a first step to a model of grain clustering in inertial-range turbulence, and many aspects could be improved. We have *intentionally* excluded many important details of turbulent structure, higher order grain-gas coupling terms, and other effects, in order to construct a simplest possible phenomenological model which is able to reproduce certain basic statistics of grains in turbulence. As such, many areas ripe for further investigation.

For example, we fundamentally assume that grains have no effect on the character of gas turbulence statistics (though it may drive that turbulence), which is probably not true when  $\rho_p \gtrsim 1$ . And indeed, we see our predictions do not agree well with the simulations when  $\bar{\rho} \gg 1$ . Similarly, we appear to predict too rapid a turnover in grain clustering when  $\tau_s \gg 1$ . In these limits, it might be more accurate to begin from the statistics of a purely collisionless grain system, and treat the gas perturbatively (essentially the opposite of our approach). Further investigation of this regime is warranted.

We also simplify tremendously by only considering the mean effects of eddies with a given scale; but even at fixed eddy scale there should be a distribution of eddy structure, meaning that  $\delta \ln \rho(\lambda)$  is not simply a number but itself a distribution. More detailed models could generalize our model here to allow this. Such generalizations have been developed for the pure gas statistics (see Castaing 1996; Chainais 2006); however, experimental data has been largely unable to distinguish that case from the simplified model. We show one example of such a model, which suggests that the character of the grain density PDF at high- $\rho_p$  may be able to distinguish such higher order models. Other such models, for example the  $\beta$ -models proposed in Hogan et al. (1999), may be more accurate still.

We have also, of course, assumed a very simplistic model for the eddy statistics, which assumes they are self-similar and statistically independent. In detail, neither of these assumptions is true in real turbulence. A more detailed model for the hierarchy of eddies could be coupled to the results we have derived for the behaviour of grains encountering individual eddies; this is straightforward. However, it is more complicated to properly account for the non-independent

nature of eddies, since if one wishes to treat this case in detail, it is no longer a good approximation to apply the results of our calculation for single eddies to ‘each’, because of non-linear terms in the grain response function. It remains to be seen if the non-linear behaviour can be captured outside of direct numerical simulations.

We have also ignored the role of the strain field in determining the clustering of grains; we have essentially considered vorticity as a force ‘expelling’ grains (except in anticyclonic vortices in a rotating disc, which lead to concentration effects we explicitly model), and modelled their concentration with a simple mass-conservation argument (based on what is expelled being ‘spread about’). This is obviously an oversimplification, and will be most significant in the high-grain density regime in small-scale turbulence. Future work should attempt to explicitly model the form and dynamics of grain density fluctuations in a non-linear strain field.

Additionally, our calculation ignores grain–grain collisions, which may significantly alter the statistics on the smallest scales in high-density regions (see Johansen et al. 2012). And our scalings are derived for inertial-range turbulence; the case appropriate for small Stokes numbers where concentration occurs at/below the viscous scale is much more well studied in the terrestrial turbulence literature (see Section 1), and may be more relevant for the smallest grains. Thus on the smallest scales where collisions and/or viscous effects dominate, our predictions are expected to break down.

#### 6.4 Implications

Despite these significant limitations, the model here has a range of implications for many important astrophysical questions. An analytic model for grain clustering is particularly important in order to extrapolate to regimes which cannot easily be simulated (large Reynolds numbers and/or small scales). With an analytic description of the grain density power spectrum, it becomes straightforward to apply methods such as those in Hopkins & Christiansen (2013) to estimate the mass and/or size spectra of grain aggregations meeting various ‘interesting’ criteria (such as those aggregations which are self-gravitating).

Large grain clustering is of central importance to planetesimal formation. In grain overdensities reaching  $\sim 100$ – $1000$  times the mean, local grain densities in protoplanetary discs can easily exceed gas densities, triggering additional processes such as the streaming instability. It is even possible that such large fluctuations can directly exceed the Roche density and promote gravitational collapse (see Cuzzi, Hogan & Shariff 2008). In future work, we will use the model here to investigate the conditions under which such collapse may be possible.

Grain–grain collisions in protoplanetary discs and the ISM depend sensitively on the small-scale clustering of grains, i.e.  $\langle n^2(\lambda \rightarrow 0) \rangle$ , which we show can differ dramatically from a homogeneous medium, even for very small grains. Even simple clumping factors  $\langle n^2 \rangle / \langle n \rangle^2$  can be large ( $\gg 1$ ). Thus grain clustering can make substantial differences to quantities such as grain collision rates and approach velocities.

Radiative transfer through the dusty ISM (and consequences such as emission, absorption, and cooling via dust) also depend on dust clumping. Depending on the geometry and details of the problem, this can even extend to extremely small-scale clustering properties within the dust, where inhomogeneities cannot be resolved in current simulations and may depend critically on dust clustering even independent of gas density fluctuations.

This model should be equally applicable to terrestrial turbulence, in the case of large Reynolds and Stokes numbers. We predict that

even relatively large or heavy aerosols may undergo large number density fluctuations in inertial-range turbulence. We specifically provide a theoretical framework for the observations of preferential concentration of large-St grains with amplitudes larger than those corresponding to pure random-phase models (Bec et al. 2007), with scale-dependent Stokes number  $\tilde{\tau}_s = \tau_s/t_c(\lambda)$ . Measurements of the clustering scales of these particles and their amplitudes can strongly constrain the role of coherent structures in preferential concentration and their geometry/fractal structure.

The intention here is to provide a simple framework in which to interpret simulations and experiments of grain clumping. We provide general fitting functions, which can be used in simulations to quantify important properties of turbulent fluctuations, such as the dynamic range of the cascade contributing to fluctuations and the magnitude of coherent ‘events.’ They also provide a guideline for understanding on which scales simulations can resolve clumping, and to understand the regimes to which these results can be generalized.

#### ACKNOWLEDGEMENTS

We thank Jessie Christiansen and Eugene Chiang for many helpful discussions during the development of this work. We also thank Andrew Youdin, Anders Johansen, and Jeff Cuzzi for several suggestions. We also thank our referee for a number of constructive and useful suggestions, especially pointing out the means to replace some earlier simplified approximations with exact solutions. Support for PFH was provided by NASA through Einstein Postdoctoral Fellowship Award Number PF1-120083 issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the NASA under contract NAS8-03060.

#### REFERENCES

- Bai X.-N., Stone J. M., 2010a, *ApJS*, 190, 297  
 Bai X.-N., Stone J. M., 2010b, *ApJ*, 722, L220  
 Bai X.-N., Stone J. M., 2010c, *ApJ*, 722, 1437  
 Bec J., Cencini M., Hillerbrand R., 2007, *Phys. Rev. E*, 75, 025301  
 Bec J., Cencini M., Hillerbrand R., Turitsyn K., 2008, *Physica D Nonlinear Phenomena*, 237, 2037  
 Bec J., Biferale L., Cencini M., Lanotte A. S., Toschi F., 2009, preprint (arXiv:0905.1192)  
 Boldyrev S., 2002, *ApJ*, 569, 841  
 Bracco A., Chavanis P. H., Provenzale A., Spiegel E. A., 1999, *Phys. Fluids*, 11, 2280  
 Budaev V., 2008, *Plasma Phys. Rep.*, 34, 799  
 Burlaga L. F., 1992, *J. Geophys. Res.*, 97, 4283  
 Carballido A., Fromang S., Papaloizou J., 2006, *MNRAS*, 373, 1633  
 Carballido A., Stone J. M., Turner N. J., 2008, *MNRAS*, 386, 145  
 Castaing B., 1996, *J. Phys. II*, 6, 105  
 Chainais P., 2006, *Eur. Phys. J. B*, 51, 229  
 Chiang E., Youdin A. N., 2010, *Annu. Rev. Earth Planet. Sci.*, 38, 493  
 Cuzzi J. N., Hogan R. C., Paque J. M., Dobrovolskis A. R., 2001, *ApJ*, 546, 496  
 Cuzzi J. N., Hogan R. C., Shariff K., 2008, *ApJ*, 687, 1432  
 Dittrich K., Klahr H., Johansen A., 2013, *ApJ*, 763, 117  
 Dubrulle B., 1994, *Phys. Rev. Lett.*, 73, 959  
 Elperin T., Kleorin N., Rogachevskii I., 1996, *Phys. Rev. Lett.*, 77, 5373  
 Elperin T., Kleorin N., Rogachevskii I., 1998, *Phys. Rev. Lett.*, 81, 2898  
 Falkovich G., Pumir A., 2004, *Phys. Fluids*, 16, L47  
 Federrath C., 2013, *MNRAS*, 436, 1245  
 Federrath C., Banerjee S., 2015, *MNRAS*, 448, 3297  
 Federrath C., Klessen R. S., Schmidt W., 2008, *ApJ*, 688, L79



Fessler J. R., Kulick J. D., Eaton J. K., 1994, *Phys. Fluids*, 6, 3742  
 Goodman J., Pindor B., 2000, *Icarus*, 148, 537  
 Gualtieri P., Picano F., Casciola C. M., 2009, *J. Fluid Mech.*, 629, 25  
 Gustavsson K., Meneguz E., Reeks M., Mehlig B., 2012, *New J. Phys.*, 14, 115017  
 Hogan R. C., Cuzzi J. N., 2007, *Phys. Rev. E*, 75, 056305  
 Hogan R. C., Cuzzi J. N., Dobrovolskis A. R., 1999, *Phys. Rev. E*, 60, 1674  
 Hopkins P. F., 2013a, *MNRAS*, 430, 1653  
 Hopkins P. F., 2013b, *MNRAS*, 430, 1880  
 Hopkins P. F., Christiansen J. L., 2013, *ApJ*, 776, 48  
 Jalali M. A., 2013, *ApJ*, 772, 75  
 Johansen A., Youdin A., 2007, *ApJ*, 662, 627  
 Johansen A., Youdin A. N., Lithwick Y., 2012, *A&A*, 537, A125  
 Kolmogorov A., 1941, *Dokl. Akad. Nauk SSSR*, 30, 301  
 Konstandin L., Girichidis P., Federrath C., Klessen R. S., 2012, *ApJ*, 761, 149  
 Lee A. T., Chiang E., Asay-Davis X., Barranco J., 2010, *ApJ*, 725, 1938  
 Lyra W., Johansen A., Zsom A., Klahr H., Piskunov N., 2009, *A&A*, 497, 869  
 Marcu B., Meiburg E., Newton P. K., 1995, *Phys. Fluids*, 7, 400  
 Markiewicz W. J., Mizuno H., Voelk H. J., 1991, *A&A*, 242, 286  
 Molina F. Z., Glover S. C. O., Federrath C., Klessen R. S., 2012, *MNRAS*, 423, 2680  
 Monchaux R., Bourgoin M., Cartellier A., 2010, *Phys. Fluids*, 22, 103304  
 Monchaux R., Bourgoin M., Cartellier A., 2012, *Int. J. Multiph. Flow*, 40, 1  
 Nakagawa Y., Sekiya M., Hayashi C., 1986, *Icarus*, 67, 375  
 Olla P., 2010, *Phys. Rev. E*, 81, 016305  
 Ormel C. W., Cuzzi J. N., 2007, *A&A*, 466, 413  
 Pan L., Padoan P., 2010, *J. Fluid Mech.*, 661, 73  
 Pan L., Padoan P., 2013, *ApJ*, 776, 12  
 Pan L., Padoan P., Scalo J., Kritsuk A. G., Norman M. L., 2011, *ApJ*, 740, 6  
 Peebles P. J. E., 1993, Princeton Univ. Press, Princeton  
 Price D. J., Federrath C., Brunt C. M., 2011, *ApJ*, 727, L21  
 Rouson D. W. I., Eaton J. K., 2001, *J. Fluid Mech.*, 428, 149  
 Schmidt W., Federrath C., Klessen R., 2008, *Phys. Rev. Lett.*, 101, 194505  
 She Z.-S., Leveque E., 1994, *Phys. Rev. Lett.*, 72, 336  
 She Z.-S., Waymire E. C., 1995, *Phys. Rev. Lett.*, 74, 262  
 She Z.-S., Zhang Z.-X., 2009, *Acta Mech. Sin.*, 25, 279  
 Sigurgeirsson H., Stuart A. M., 2002, *Phys. Fluids*, 14, 4352  
 Sorriso-Valvo L., Carbone V., Veltri P., Consolini G., Bruno R., 1999, *Geophys. Res. Lett.*, 26, 1801  
 Squires K. D., Eaton J. K., 1991, *Phys. Fluids A*, 3, 1169  
 Stewart T., Strijbosch L., Moors J., Batenburg P. v. 2006, A simple approximation to the convolution of gamma distributions. Discussion Paper, Tilburg University, Center for Economic Research, Available at: <http://ideas.repec.org/p/dgr/kubcen/200627.html>, p. 27  
 Voelk H. J., Jones F. C., Morfill G. E., Roeser S., 1980, *A&A*, 85, 316  
 Wilkinson M., Mehlig B., Gustavsson K., 2010, *Europhys. Lett.*, 89, 50002  
 Yoshimoto H., Goto S., 2007, *J. Fluid Mech.*, 577, 275  
 Youdin A. N., Goodman J., 2005, *ApJ*, 620, 459  
 Youdin A. N., Lithwick Y., 2007, *Icarus*, 192, 588  
 Zaichik L. I., Alipchenkov V. M., 2009, *New J. Phys.*, 11, 103018

## APPENDIX A: AN EXACT SOLUTION FOR THE RESPONSE FUNCTION

### A1 General case

In the text, we discuss the response of grains to a vortex with pure vorticity  $1/t_e$  ( $\delta u_\theta = r/t_e$ ,  $\delta u_r = 0$ ). Here, we derive this more exactly.

Consider a vortex which appears for a time  $\delta t = t_e$ , so  $\delta \mathbf{u}(t) = \delta \mathbf{u} \Theta(0 < t < t_e)$ . First note that, for this particular choice of  $\delta \mathbf{u}$ ,

we have  $\delta u_x = \delta \mathbf{u} \cdot \hat{x} = -y/t_e$ ,  $\delta u_y = +x/t_e$ .<sup>24</sup> In Cartesian coordinates, the equations of motion (equation 2) in the vortex plane then become

$$\delta \dot{v}_x = 2 \Omega_R \delta v_y - t_s^{-1} (\delta v_x + t_e^{-1} y) \quad (\text{A1})$$

$$\delta \dot{v}_y = -\frac{1}{2} \Omega_R \delta v_x - t_s^{-1} (\delta v_y - t_e^{-1} x). \quad (\text{A2})$$

Note that  $\delta v_x = \dot{x}$  and  $\delta v_y = \dot{y}$ . For convenience, define the time in units of  $t_s$ , so  $x' \equiv t_s \dot{x}$  and so on. Now the equations become

$$x'' = 2 \tau_s y' - x' - \tilde{\tau}_s y \quad (\text{A3})$$

$$y'' = -\frac{1}{2} \tau_s x' - y' + \tilde{\tau}_s x. \quad (\text{A4})$$

If we define the vector  $\mathbf{x} \equiv (x, y, x', y')$ , this is a linear system:

$$\mathbf{x}' = \mathbf{M} \cdot \mathbf{x} \quad (\text{A5})$$

$$\mathbf{M} \equiv \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\tilde{\tau}_s & -1 & 2 \tau_s \\ \tilde{\tau}_s & 0 & -\frac{1}{2} \tau_s & -1 \end{pmatrix}. \quad (\text{A6})$$

The eigenvalues  $\lambda_i$  of  $\mathbf{M}$  solve the characteristic polynomial:

$$\lambda_i^4 + 2 \lambda_i^3 + (1 + \tau_s^2) \lambda_i^2 - \frac{5}{2} \tau_s \tilde{\tau}_s \lambda_i + \tilde{\tau}_s^2 = 0. \quad (\text{A7})$$

Since this is a simple system of linear ordinary differential equations (ODEs), the solution  $\mathbf{x}(t/t_s)$  is given by

$$\mathbf{x}(t/t_s) = \mathbf{V} \cdot [\mathbf{\Lambda} \cdot (\mathbf{V}^{-1} \cdot \mathbf{x}_0)], \quad (\text{A8})$$

where  $\mathbf{x}_0 \equiv \mathbf{x}(t=0)$ ,  $\mathbf{\Lambda}$  is the diagonal matrix of eigenfunctions

$$\Lambda_{ij} = \delta_{ij} \exp(-\lambda_i t/t_s) \quad (\text{A9})$$

and  $\mathbf{V}$  is the column matrix of eigenvectors  $\mathbf{v}_i$  corresponding to each eigenvalue  $\lambda_i$

$$\mathbf{V} \equiv \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \quad (\text{A10})$$

At  $t = t_1 = |t_e|$ , the vortex is removed, so we have

$$\mathbf{x}_1 \equiv \mathbf{x}(t = |t_e|) = \mathbf{V} \cdot [\mathbf{\Lambda}_1 \cdot (\mathbf{V}^{-1} \cdot \mathbf{x}_0)] \quad (\text{A11})$$

$$\Lambda_{1,ij} = \delta_{ij} \exp(-\lambda_i |t_e|/t_s) \quad (\text{A12})$$

and the equations of motion become

$$\delta \dot{v}_x = 2 \Omega_R \delta v_y - t_s^{-1} \delta v_x \quad (\text{A13})$$

$$\delta \dot{v}_y = -\frac{1}{2} \Omega_R \delta v_x - t_s^{-1} \delta v_y. \quad (\text{A14})$$

It is straightforward to see that this has the solution  $\delta v_x = a \exp(-[1 + \tau_s](t - t_1)/t_s) + b \exp(-[1 - \tau_s](t - t_1)/t_s)$ , and a corresponding form for  $\delta v_y$ , where the constants are determined by matching to  $\delta v_x(t = t_1) = t_s^{-1} x'(t = t_1)$  and  $\delta v_y(t = t_1) = t_s^{-1} y'(t = t_2)$ . The velocities are exponentially

<sup>24</sup> Here, the sign of  $t_e$  can be positive or negative, reflecting cyclonic or anticyclonic vortices, respectively.

damped (since  $\delta \mathbf{u}(t > t_e) = 0$ , and the final positions  $x_f, y_f$  are given by

$$x_f = x(t_1) + \int_{t_1}^{\infty} \delta v_x(t) dt = x(t_1) + \frac{x'(t_1) + 2\tau_s y'(t_1)}{1 + \tau_s^2} \quad (\text{A15})$$

$$y_f = y(t_1) + \int_{t_1}^{\infty} \delta v_y(t) dt = y(t_1) + \frac{y'(t_1) - \frac{1}{2}\tau_s x'(t_1)}{1 + \tau_s^2} \quad (\text{A16})$$

we can write this as a linear transform

$$\mathbf{x}_f = \mathbf{P} \cdot \mathbf{x}_1 = \mathbf{P} \cdot \{\mathbf{V} \cdot [\mathbf{\Lambda}_1 \cdot (\mathbf{V}^{-1} \cdot \mathbf{x}_0)]\} \quad (\text{A17})$$

$$\mathbf{P} \equiv \begin{pmatrix} 1 & 0 & (1 + \tau_s^2)^{-1} & 2\tau_s(1 + \tau_s^2)^{-1} \\ 0 & 1 & -\frac{1}{2}\tau_s(1 + \tau_s^2)^{-1} & (1 + \tau_s^2)^{-1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{A18})$$

So the final coordinates are simply a linear transformation of the initial coordinates:

$$\mathbf{x}_f = \mathbf{T} \cdot \mathbf{x}_0 \quad (\text{A19})$$

$$\mathbf{T} \equiv \mathbf{P} \cdot \mathbf{V} \cdot \mathbf{\Lambda}_1 \cdot \mathbf{V}^{-1}. \quad (\text{A20})$$

Now, consider particles with negligible initial velocities  $\delta \mathbf{v}(t = 0) \approx 0$  (since we wish to focus on the mean perturbation to their distribution); we will include the effects of non-zero initial velocity below. We can then write

$$\begin{pmatrix} x_f \\ y_f \end{pmatrix} = \mathbf{J} \cdot \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \quad (\text{A21})$$

$$\mathbf{J} \equiv \begin{pmatrix} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{pmatrix} \quad (\text{A22})$$

i.e. the Jacobian matrix  $\mathbf{J}$  is just the upper-left block of the matrix  $\mathbf{T}$ . The mapping from  $(x_0, y_0) \rightarrow (x_f, y_f)$  is just a linear coordinate transformation, so the density distribution  $\rho$  of particles must obey

$$\rho_f dx_f dy_f = \rho_0 dx_0 dy_0 \quad (\text{A23})$$

$$\rho_f = |\text{Det}[\mathbf{J}]|^{-1} \rho_0 \quad (\text{A24})$$

so  $f_f$  is simply related to  $f_0$  by the determinant of  $\mathbf{J}$ . Since this is independent of position, we have

$$\langle \delta \ln \rho \rangle = \delta \ln \rho = -\ln |\text{Det}[\mathbf{J}]| = -2\varpi \quad (\text{A25})$$

$$\varpi \equiv \frac{1}{2} \ln |\text{Det}[\mathbf{J}]|. \quad (\text{A26})$$

In general, this must be evaluated numerically, since  $\lambda_i$  has no closed-form general solution. However, we can exactly evaluate several limiting expressions.

## A2 Large eddies

First, consider the limit of large eddies,  $|t_e \Omega| \gg 1$ . While we can derive the limiting expression directly from the above, it is easier to start with our equations of motion, and linearize in terms of  $\mathcal{O}(|t_e \Omega|)^{-1} \ll 1$ . It is easy to verify this admits a solution of the form  $x \propto \exp(\lambda_{|t_e \Omega| \gg 1} [\tau_s] t / |t_e|)$ , with eigenvalues

$$\lambda_{|t_e \Omega| \gg 1} = \frac{5\tau_s/4 \pm \sqrt{(3\tau_s/4)^2 - 1}}{1 + \tau_s^2} \text{SIGN}(t_e), \quad (\text{A27})$$

where  $\text{SIGN}(t_e) \equiv +1$  if  $t_e > 0$ , which in these units means the eddy is cyclonic (the sign of the vorticity is the same as that of the Keplerian flow), and  $\text{SIGN}(t_e) \equiv -1$  if  $t_e < 0$  refers to an anticyclonic eddy. For  $t_e \Omega > 0$  this is a growing mode, so the eddy sees exponential dispersal of the particles (expansion of their Lagrangian radius), while for  $t_e \Omega < 0$  the eddy concentrates particles. Note that the second-derivative terms  $\mathcal{O}(|t_e \Omega|)^{-2}$  drop here, so this becomes a first-order linear system of ODEs, and the governing matrices are now  $2 \times 2$ , with the form (written in the same notation as our general derivation above)

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{A28})$$

$$\mathbf{M} \rightarrow \frac{t_s}{t_e} \frac{1}{1 + \tau_s^2} \begin{pmatrix} 2\tau_s & -1 \\ -1 & 2\tau_s \end{pmatrix}. \quad (\text{A29})$$

After some tedious but straightforward linear algebra, we obtain the Jacobian determinant:

$$\varpi(|t_e \Omega \gg 1|) \rightarrow \frac{5}{4} \text{SIGN}(t_e) \frac{\tau_s}{1 + \tau_s^2}. \quad (\text{A30})$$

We see that this is just the linear growth rate ( $\lambda/t_e$ ) of the radius (the  $\pm \sqrt{(3\tau_s/4)^2 - 1}$  term in the eigenvalue represents motion of the particles on elliptical trajectories), times the eddy lifetime ( $t_e$ ) as we should anticipate for a weak perturbation. Interestingly, this means when  $|t_e \Omega| \gg 1$ , the perturbation becomes independent of  $t_e$  and depends only on  $\tau_s$ . The  $5/4$  coefficient here comes from the fact that shear enhances the dispersal more strongly in the  $x$ -direction (the  $+2\Omega$  term) than in the  $y$ -direction (the  $-1/2\Omega$  term). Averaged over initial positions, the mean ‘stretching’ coefficient for the ellipse along which particles in some initial ring are sheared goes as the mean of the absolute value of the two coefficients:  $(1/2)(|2| + |-1/2|) = 5/4$ .

## A3 Small eddies

Now consider small eddies,  $|t_e \Omega| \ll 1$ . To leading order, we can now drop the shear terms in  $\mathbf{M}$ , and obtain

$$\lambda_i = -\frac{1}{2} (1 \pm \sqrt{1 \pm i 4\tau_s}). \quad (\text{A31})$$

The eigenvalues come in two conjugate pairs, a growing pair and a decaying pair. The conjugation simply represents rotation, since the equations without shear are symmetric under rotation. Taking  $r^2 \propto x^2 + y^2$  with  $x \sim \sum a_i \exp(\lambda_i t/t_s)$ , it is straightforward to see that this translates to four radial modes: three are damped, one with decay rate (real part of the radial eigenvalue)  $= 1/t_s$  and a conjugate pair with decay rate  $= 1/(2t_s)$ . And one is a growing mode, with growth rate  $= [-2 + \sqrt{2(1 + \sqrt{1 + 16\tau_s^2})}]/(4t_s)$ , the mode noted in the main text.<sup>25</sup>

With these eigenvalues, we obtain

$$\varpi \rightarrow -\frac{2 + a_+ + a_-}{4|\tau_s|} - \ln 2 - \frac{1}{4} \ln(1 + 16\tau_s^2) + \quad (\text{A32})$$

<sup>25</sup>To show this, apply Euler’s formula to the  $\lambda_i$  values above, and use the identities  $\cos(\tan^{-1}[x]/2) = 2^{-1/2} [1 + (1 + x^2)^{-1/2}]^{1/2}$  and  $\sin(\tan^{-1}[x]/2) = 2^{-1/2} x [1 + x^2]^{-1/2} [1 + (1 + x^2)^{-1/2}]^{-1/2}$ .

$$\begin{aligned} & \frac{1}{2} \ln \left[ -(1 + 2\iota \tilde{\tau}_s - a_+) + (1 + 2\iota \tilde{\tau}_s + a_+) \exp(a_+ |\tilde{\tau}_s|^{-1}) \right] \\ & + \frac{1}{2} \ln \left[ -(1 - 2\iota \tilde{\tau}_s - a_-) + (1 - 2\iota \tilde{\tau}_s + a_-) \exp(a_- |\tilde{\tau}_s|^{-1}) \right] \\ & a_{\pm} \equiv \sqrt{1 \pm \iota 4 \tilde{\tau}_s}. \end{aligned} \quad (\text{A33})$$

Here,  $\varpi$  is identical under the transformation  $\tilde{\tau}_s \rightarrow -\tilde{\tau}_s$  ( $t_e \rightarrow -t_e$ ), so cyclonic and anticyclonic vortices behave identically. This should be obvious from the symmetry of the problem absent shear. Since  $\varpi > 0$  for all  $\tilde{\tau}_s$ , this means grains are always dispersed by small eddies, independent of the eddy orientation.

The expression for  $\varpi$  above is rather opaque; however, it can be approximated to  $\sim 1$  per cent accuracy at all  $\tilde{\tau}_s$  by

$$\varpi \approx |\tilde{\tau}_s| \left[ 1 + \frac{|\tilde{\tau}_s|}{\ln \sqrt{2}} (1 + |\tilde{\tau}_s|^{-1})^{-1/5} \right]^{-1}. \quad (\text{A34})$$

And the limits are easily evaluated:

$$\varpi \rightarrow |\tilde{\tau}_s| \quad (|\tilde{\tau}_s| \ll 1) \quad (\text{A35})$$

$$\varpi \rightarrow \ln \sqrt{2} \quad (|\tilde{\tau}_s| \gg 1). \quad (\text{A36})$$

## APPENDIX B: APPROXIMATE EXPRESSIONS FOR LARGE GRAINS

For large grains ( $\tau_s \gtrsim 0.1$ ), and/or large scales ( $t_e \gtrsim 0.1 \Omega^{-1}$ ) with appreciable levels of turbulence ( $\alpha \gtrsim 10^{-8}$ ), we can

simplify our derivations considerably. Over the dynamic range where nearly all the power in density fluctuations is concentrated in Fig. 1, we see that the response function is approximately constant, with its asymptotic value  $\varpi \approx (5/4)/(\tau_s + \tau_s^{-1})$ . Over the same range, the  $\beta$  term in  $g(\lambda_e)$  (Table 1) is almost always dominant in the function  $g$  ( $g(\lambda_e) \approx \beta^{-2}$ ), so  $h(\lambda_e) \approx (\tilde{\tau}_s/\tilde{\tau}_s(\lambda_{\max}))(\lambda_e/\lambda_{\max})g(\lambda_e)^{-1/2} \approx \beta |v_e|/|v_e(\lambda_{\max})| = |v_e|/|v_{\text{drift}}|$ . Using  $|v_e|/|v_e(\lambda_{\max})| = (\lambda_e/\lambda_{\max})^{\zeta_1}$ , we obtain

$$\langle \delta \ln \rho \rangle \approx -\frac{5 N_d/4}{\tau_s + \tau_s^{-1}} \left[ 1 + \beta^{-1} \left( \frac{\lambda_e}{\lambda_{\max}} \right)^{-\zeta_1} \right]^{-1} \text{SIGN}(t_e). \quad (\text{B1})$$

With this approximation,  $\Delta_{\ln \rho} = C_{\infty} |\delta \ln \rho|^2$  follows trivially. The integral quantities  $\mu$  and  $S_{\ln \rho}$  used to estimate the density distribution and  $S_{\rho}$  can be evaluated exactly in closed form; these are presented in Table 2. The integral  $\rho_{p, \max}$  is straightforward to evaluate numerically, but the closed-form expression is rather unwieldy; using the fact that neither  $|\delta_0|$  (Table 2) or  $\beta$  are extremely large ( $\gtrsim 100$ ), the exact integral can be approximated to very good accuracy by the function (from the logarithmic series expansion) given in Table 2.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.