



Supplement of

Intricate relations among particle collision, relative motion and clustering in turbulent clouds: computational observation and theory

Ewe-Wei Saw and Xiaohui Meng

Correspondence to: Ewe-Wei Saw (ewsaw3@gmail.com) and Xiaohui Meng (mengxh7@mail2.sysu.edu.cn)

The copyright of individual parts of the supplement might differ from the article licence.

Supplementary Material

S1 Further Details of the Direct Numerical Simulation.

The time step in our DNS Δt is 0.001 s. The courant number is $C = 0.073$, (where $C = \Delta t \left[\frac{u'}{\Delta x} + \frac{v'}{\Delta y} + \frac{w'}{\Delta z} \right]_{\max}$, u' etc. are r.m.s. velocities, Δx etc. are grid spacings). The normalized maximum wavenumber simulated is $k_{\max}\eta = 1.2$. The turbulent flow is sustained by randomly forcing the two lowest nonzero shells of wave numbers. The integral length scale of the turbulent flow is estimated to be $L = 0.646$ dm.

We study the statistics of monomers only (i.e. the particle of the same size (d) that we initially introduce into the system and which we later replenish at a constant rate close to the monomer-monomer collision rate). In this sense, the particle (monomers) are naturally lost from our consideration once they collide and become larger particles. Particles that become much larger ($St > 21.6$) are removed from the DNS at each time step.

S2 Estimation of Leading Order Terms in the Drift Flux, e.g $a_{ik}^{(1)}$

Using the DNS data, we estimate e.g. the value of

$$\int_{-\infty}^t a_{ik}^{(1)} dt' \equiv \int_{-\infty}^t \tau_\eta \langle \Gamma_{ik}(t) \Gamma_{lm}(t') \Gamma_{ml}(t') \rangle dt'.$$

Note: the averaging is done over fluid particles (the theory assumed $St \ll 1$ limit, such that all velocity statistics are tied to the fluid's), the integrand is non-vanishing only for t' in the vicinity of $t - \tau_\eta$ to t (where the turbulent velocity gradient Γ_{ij} retains correlation), thus this quantity may be approximated as: $\tau_\eta^2 \langle \Gamma_{ik}(t) \Gamma_{lm}(t) \Gamma_{ml}(t) \rangle$. As shown in Chun et al. (2005), $\langle \Gamma_{ik}(t) \Gamma_{lm}(t) \Gamma_{ml}(t) \rangle$ is by definition zero in fully developed turbulence due to the fact that the small-scale statistics of turbulent flows are almost isotropic Kolmogorov (1941). However, the coagulation constraint dictates that at $r = d$, such averages must be taken with the condition that only fluid-particle pairs with negative radial velocity ($w_r < 0$) are taken into

20 account (that the inertial particles' motion being tied to the fluid's does not imply that inertial pairs sample the fluid particle pairs's motion uniformly). Under this condition, the DNS data gives $\tau_\eta^2 \langle \Gamma_{ik}(t) \Gamma_{lm}(t) \Gamma_{ml}(t) \rangle \approx (-0.171 \times 10^{-3} \text{ dm/s})/d_*$, ($d_* = 9.49 \times 10^{-4} \text{ dm}$); here, it is of value to point out that without such constraint or condition, the result for this quantity from the DNS is two orders of magnitude smaller. Similarly, we found $\int_{-\infty}^t a_{ki}^{(2)} dt' \approx \tau_\eta^3 \langle \Gamma_{ij}(t) \Gamma_{jk}(t) \Gamma_{lm}(t) \Gamma_{ml}(t) \rangle \approx (2.32 \times 10^{-3} \text{ dm/s})/d_*$; for this quantity, the DNS gives roughly the same values with or without the constraint.

25 **S3 Full Definition of the Function $f_I(R_0, \mu, t_f)$ in the Model for Non-local Diffusive Flux.**

Derived in Chun et al. (2005), summarized here (with typo corrected), the diffusive action of the turbulence on the particle-pairs is assumed to consist of a random sequence of uniaxial extensional or compressional flows defined, and:

$$f_I(R_0, \mu, t_f) \equiv f_+ I_+(R_0, \mu, t_f) + f_- I_-(R_0, \mu, t_f),$$

where $R_0 \equiv r_0/r$, r_0 is the initial separation distance of a particle pair before a straining event, r is the independent variable of the equation for $g(r)$; f_+ and $f_- \equiv 1 - f_+$ are the fractions of those flows that are extensional and compressional, respectively. Comparing with DNS, Chun et al. (2005) calibrated f_+ and found $f_+ = 0.188$ (a result we use here). I_\pm is an indicator function such that it takes the value $+1$ (-1) when a secondary particle leaves (enters) a sphere of radius r centered on the primary particle, and otherwise zero. μ is the cosine of the angle between the axis of symmetry of the straining flow event and the displacement vector between the two particles, t_f is the lifetime of the event. To obtain a strain rate correlation function that decays exponentially with a characteristic time scale τ_S , Chun et al. (2005) set the probability density function for t_f to be:

$$F(t_f) = \frac{f_s t_f}{\tau_S^2} \exp(-t_f/\tau_S).$$

The indicator function is used to count the net loss of particles from within the sphere over the duration of an (extensional or compressional) event and can be expressed as:

$$I_\pm(R_0, \mu, t_f) = H(1 - R_0)H(R_{f\pm} - 1) - H(R_0 - 1)H(1 - R_{f\pm}),$$

40 where $H(x)$ is the Heaviside function (zero for $x < 0$, unity for $x \geq 0$), $R_{f\pm}$ is the non-dimensional final position of a particle pair with an initial position of R_0 and can be written as:

$$R_{f+} = R_0 \left[\mu^2 \theta_t^2 + \frac{(1 - \mu^2)}{\theta_t} \right]^{1/2},$$

$$R_{f-} = R_0 \left[\frac{\mu^2}{\theta_t^2} + (1 - \mu^2) \theta_t \right]^{1/2},$$

45 for uniaxial extension and compression respectively, where:

$$\theta_t \equiv \exp \left(\frac{t_f}{\tau_\eta \sqrt{3 f_s}} \right).$$

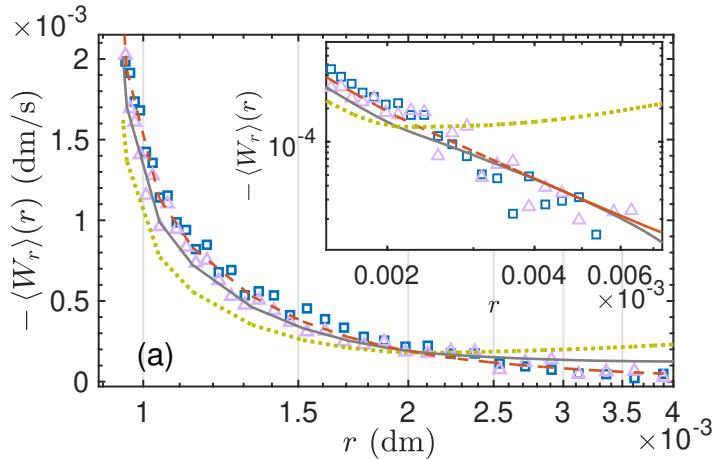


Figure S1. MRVs of particles. Triangles: DNS result for $St = 0.001$. Squares: DNS result for $St = 0.11$. Red-dashed-line: theory's prediction for $St = 0.001$. Gold-dotted-line: theory's prediction for $St = 0.11$. Grey-solid-line: modified-theory's prediction for $St = 0.11$ (details of modification in Sec. S5). Note: the predictions are based on the A_τ values valid in the $r \sim d$ regime ($d = 9.49 \times 10^{-4}$ dm). **Inset)** Similar plots in logarithmic axes highlighting the large- r regime. Note: the predictions are based on the A_τ values valid in the $r \gg d$ regime. It is clear that the modified-theory's predictions agree much better with the DNS.

50 S4 MRV Predictions by the Theory for Other Stokes Numbers.

As mentioned in the main text of this manuscript, even though the theory assumes that MRVs are St -independent for small St 's. Nevertheless, it could produce separate predictions for each St . Here we show the predictions for $St = 0.001$ and 0.11 in Fig. 1. The prediction for $St = 0.001$ (red dash line) agrees well with the DNS results (symbols), but the prediction for $St = 0.11$ (gold dotted line) deviates significantly, suggesting that finite St effects not captured by the theory start to become 55 significant and thus diminish the accuracy of the theory.

S5 Derivation of c_{st} , its Role and Possibility of Further Corrections to The CK Theory.

In this work, we deviate crucially¹ from the CK theory Chun et al. (2005) by introducing an extra factor c_{st} (positive, of order unity or less) in the model of non-local diffusion:

$$q_r^D = c_{st} r \int d\Omega \int_0^\infty dt_f F(t_f) \int_{d/r}^\infty dR_0 R_0^2 \langle P \rangle(r_{R_0}) f_I(R_0, \mu, t_f). \quad (\text{S1})$$

60 To determine what c_{st} is (or should be), we begin from an important finding in Chun et al. (2005) that if $\langle P \rangle$ is power-law of r , i.e. $\langle P \rangle = Cr^{-c_1}$, then the non-local diffusion q_r^D can be cast into a differential form (which is usually only true for local

¹'Crucial' refers to the fact that without c_{st} the theory would be inconsistent with previous experimental results (as this section will show) and it would also produce results far from our DNS results.

diffusion):

$$q_r^D = -B_{nl} \tau_\eta^{-1} r^2 \frac{\partial \langle P \rangle}{\partial r}, \quad (\text{S2})$$

where:

$$65 \quad B_{nl} = \tau_\eta \int d\Omega \int_0^\infty dt_f F(t_f) \int_{d/r}^\infty dR_0 R_0^{2-c_1} f_I(R_0, \mu, t_f). \quad (\text{S3})$$

This, together with: $q_i^d = -A_{ck} \tau_\eta^{-1} r \langle P \rangle$, eventually leads to the first order equation differential equation for the RDF ($g(r) \equiv V \langle P \rangle$), that has (only) power-law solutions: $g(r) = V C r^{-c_1}$. This result (i.e. $g(r)$ or equivalently $\langle P \rangle(r)$ are power-laws) has seen compelling validations from both experiments (e.g. Saw et al. (2012b); Lu et al. (2010); Yavuz et al. (2018)) and DNS (e.g. Chun et al. (2005); Bec et al. (2007); Saw et al. (2012a)). We now begin from this experimentally validated result
70 and work backward to derive an expression for c_{st} . We plug the power-law form for $\langle P \rangle$ into (S2):

$$\begin{aligned} q_r^D &= -B_{nl} \tau_\eta^{-1} r^2 \frac{\partial (C r^{-c_1})}{\partial r} \\ &= -B_{nl} \tau_\eta^{-1} r^2 C (-c_1) r^{-c_1 - 1} \\ &= B_{nl} \tau_\eta^{-1} r c_1 C r^{-c_1} \\ &= \tau_\eta^{-1} r c_1 C r^{-c_1} \tau_\eta \int d\Omega \int_0^\infty dt_f F(t_f) \int_{d/r}^\infty dR_0 R_0^{2-c_1} f_I(R_0, \mu, t_f) \\ 75 &= r c_1 \int d\Omega \int_0^\infty dt_f F(t_f) \int_{d/r}^\infty dR_0 R_0^2 C (r R_0)^{-c_1} f_I(R_0, \mu, t_f) \\ &= c_1 r \int d\Omega \int_0^\infty dt_f F(t_f) \int_{d/r}^\infty dR_0 R_0^2 \langle P \rangle(r R_0) f_I(R_0, \mu, t_f). \end{aligned}$$

Comparing with (S1), we have:

$$c_{st} = | -c_1 | \equiv |c_1|,$$

80 which is found in experiments (and theories) to be of order 0 to 1 and a function of particle Stokes number St ; in words, this means c_{st} is given by the modulus of the power-law exponent of the RDF that would arise in the collision-less case; in the case with collision and sufficiently small particle ($d/\eta \lesssim 1$), such as in this study, c_{st} equals the modulus of the power-law exponent of the RDF the range of $d \ll r \ll 20\eta$ (note: power-laws RDF are empirically observed for $r \ll 20\eta$ Saw et al. (2008, 2012a)). Note: we have chosen to define c_{st} using the ‘modulus’ (instead of the ‘negative’ of the power-law exponent) since it guarantees
85 that q_r^D is negative (positive) when $g(r)$ is an increasing (decreasing) function of r , so that we are consistent with the fact that q_r^D is a diffusion flux. We note that both the CK theory and the current modified version assume $St \ll 1$.

Chun et al. Chun et al. (2005) went further to provide a solution for c_1 (for collision-less particles, in the $St \ll 1$ limit):

$$c_1 = \frac{A_{ck}}{B_{nl}} \equiv \frac{A_{\tau, r \gg d} \tau_\eta}{B_{nl}}, \quad (\text{S4})$$

where we have clarified that A_τ in our work is defined differently from "A" in (Chun et al., 2005) (we denote the latter as A_{ck} to avoid confusion), and $A_{\tau, r \gg d}$ is our A_τ evaluated at the large- r limit. In the current context, c_1 maybe obtained via (S4) or alternatively directly from the power-law exponent of $g(r)$ in the range $d \ll r \ll 20\eta$ as discussed above. Using values of the relevant parameters in our DNS, we found $\frac{A_\tau \tau_\eta}{B_{nl}} \approx \frac{2.4St^2 \times 0.0925}{0.0397} = 5.6St^2$, which is 15% smaller than the one found in Chun et al. (2005), i.e. $\frac{A_{ck}}{B_{nl}} \approx \frac{0.61St^2}{0.0926} = 6.6St^2$. However, we have observed in our DNS that the direct method (by fitting power-laws to the RDFs in the suitable r -range) gives c_1 which is 3.2 (1.9) times larger than the one obtained using (S4) for the case of $St = 0.054$ (0.11).

A plausible interpretation of the discrepancy described just above is that there may be another missing dimensionless factor (of order unity, possibly weakly dependent on Reynolds-number) in the correct definition of q_r^D . This is beyond the scope of this present study (to avoid confusion, we currently restrict ourselves to the least speculative correction only) and is a good subject for future works. However it may be informative to note that, by inspection, we find that if we further include a factor of $\sim 1/3$ to $1/2$ in the definition of q_r^D , then the agreement between the theoretical (the integral version) and DNS produced $\langle W_r \rangle$ is strikingly better in the $r \gg d$ limit, while in the $r \sim d$ regime, it is slightly better (the former should not come as a surprise as this is the regime of power-law RDFs and the factor of $\sim 1/3$ is exactly designed to reproduce the correct c_1).

To demonstrate the point just discussed, we show in Fig. 1 the predictions by the theory for the case of $St = 0.11$. We see that, in the $r \sim d$ regime, the prediction by the original theory (dotted line in the main figure) is somewhat below the DNS result, while the prediction by the modified theory (with a factor of $1/2$ appended to the definition of q_r^D), shown as the solid line, is much closer to DNS. In the $r \gg d$ regime, the modified theory's superiority in terms of accuracy is even more pronounced (see inset of Fig. 1).

S6 Relation Between $g(r)$ and $\langle P \rangle$.

In the main text, we state that $g(r) \equiv V \langle P \rangle$, where V is the spatial volume of the full domain of the problem i.e. $(2\pi)^3$ in the DNS. Justification: let $g(\mathbf{r})$ be the ratio of probability of finding a second particle at \mathbf{r} from a particle, to the probability of such finding in a perfectly random distributed particle population, thus: $g(\mathbf{r}) \equiv \frac{\langle P \rangle \delta x \delta y \delta z}{(\delta x \delta y \delta z)/V} \equiv \langle P \rangle V$. Further, since system is isotropic, $g(\mathbf{r}) \equiv g(r)$.

S7 Modeling of MRV based on Distribution of Particle Approach Angles $P(\theta)$.

We imagine the particles are small i.e. $d \ll \eta$ and $St \ll 1$. The latter implies their trajectories are almost like fluid particles', while the former implies that, viewed at the scale of interest $r \sim d$, their trajectories are almost rectilinear (since the radii of curvature are proportional to η). Thus in the reference frame of a primary particles, no secondary particle could have a trajectory, being straight-line, that has a history of collision with the volume of the primary (otherwise coagulation would have occurred and the secondary particle in question would cease to exist). In trigonometric terms, let θ be the angle between the secondary

particle's velocity and its vector position in the rest frame of the primary particle, then we must have: $\sin^{-1}(d/r) \leq \theta \leq \pi$,
120 with the convention that $\sin^{-1}(x) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

From the above, we could then compute the MRV, $\langle w_r \rangle_*$ based on fluid particles' statistics. Since collision-coagulation affects positive and negative relative particle velocities differently, we begin by writing $\langle w_r \rangle_*$ as a sum of the positive (i.e. $w_r > 0$) and negative branches (with proper statistical weights p_{\pm} to account for possible skewness of the probability distribution of velocity):

$$125 \quad \langle W_r \rangle \equiv \langle w_r \rangle_* = p_- \langle w_r | w_r < 0 \rangle_* + p_+ \langle w_r | w_r \geq 0 \rangle_* .$$

The negative branch $p_- \langle w_r | w_r < 0 \rangle_*$ is unaffected by collision-coagulation and we thus express it as a simple linear function of r that follows from the K41-phenomenology (Kolmogorov, 1941), i.e. $-p_- \xi_- r$, where $\xi_{\pm} \sim \sqrt{\varepsilon/(15\nu)}$, ε is the (kinetic) energy dissipation rate of the flow. For the positive branch, we further assume that the (fluid particles') joint probability density function (PDF) of magnitude of relative velocity (secondary particle relative to primary particle) $|\mathbf{w}|$ and approach-angle θ ,
130 $P(|\mathbf{w}|, \theta)$, is separable (note: $w_r \equiv |\mathbf{w}| \cos(\theta)$), hence:

$$\begin{aligned} p_+ \langle w_r | w_r \geq 0 \rangle_* &= \int_0^{\infty} d|\mathbf{w}| \int_{\theta_m}^{\frac{\pi}{2}} d\theta P(|\mathbf{w}|, \theta) |\mathbf{w}| \cos(\theta) \\ &= \int_0^{\infty} d|\mathbf{w}| P_w(|\mathbf{w}|) |\mathbf{w}| \int_{\theta_m}^{\frac{\pi}{2}} d\theta P_\theta(\theta) \cos(\theta) \\ &= p_+ \int_0^{\infty} d|\mathbf{w}| P_w(|\mathbf{w}|) |\mathbf{w}| \int_{\theta_m}^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta), \end{aligned}$$

135 where all the P 's are PDFs, note that $p_+ \equiv \int_0^{\frac{\pi}{2}} P_\theta d\theta$, $\int_0^{\frac{\pi}{2}} P_\theta^+ d\theta \equiv \int_0^{\frac{\pi}{2}} (P_\theta/p_+) d\theta = 1$ and $\int_0^\pi P_\theta d\theta = 1$, also note that $P_\theta^+ \equiv P_\theta(\theta \mid w_r \geq 0)$; more importantly $\theta_m = \sin^{-1}(d/r)$ as previously explained. Further:

$$\begin{aligned}
& p_+ \langle w_r \mid w_r \geq 0 \rangle_* \\
&= p_+ \int_0^\infty d|\mathbf{w}| P_w(|\mathbf{w}|) |\mathbf{w}| \int_{\theta_m}^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta) \\
&= p_+ \int_0^\infty d|\mathbf{w}| P_w(|\mathbf{w}|) |\mathbf{w}| \left[\int_0^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta) + \int_{\theta_m}^0 d\theta P_\theta^+(\theta) \cos(\theta) \right] \\
140 &= p_+ \int_0^\infty d|\mathbf{w}| P_w(|\mathbf{w}|) |\mathbf{w}| \int_0^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta) \left[1 + \frac{\int_{\theta_m}^0 d\theta P_\theta^+(\theta) \cos(\theta)}{\int_0^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta)} \right] \\
&= p_+ \xi_+ r \left[1 + \frac{\int_{\theta_m}^0 d\theta P_\theta^+(\theta) \cos(\theta)}{\int_0^{\frac{\pi}{2}} d\theta P_\theta^+(\theta) \cos(\theta)} \right],
\end{aligned}$$

where in the last line, we have replaced the first two integrals, combined, with its K41 estimate, where $\xi_\pm \sim \sqrt{\varepsilon/(15\nu)}$.

S8 Prediction of the Peak Location of the RDF Using the Differential Form of the Drift-Diffusion Equation.

$$-\tau_\eta^{-1} B_{nl} r^4 \frac{\partial g}{\partial r} + g(r) [r^2 \langle W_r \rangle - A_\tau r^3] = -R_c^*, \quad (S5)$$

145 A finite R_c^* inhibit us from locating the peak of the RDF using (S5) à la Lu et al. (2010) i.e. without knowing $g(r)$, since $g(r)$ could no longer be factored out when $\frac{\partial g}{\partial r} = 0$. However, we argue that (S5) could still give a reasonably accurate account of the peak location. For the case of $St = 0.05$, at $r = 3d$ (the approximate peak location), we found the DNS data gives $-\tau_\eta B_{nl} r^4 \frac{\partial g}{\partial r} \Big|_{\approx 0} + g(r) [r^2 \langle W_r \rangle - A_\tau r^3] \approx -1.05 \times 10^{-9}$ and $-R_c^* \approx -1.01 \times 10^{-9}$

S9 General Analytical Solution for the Differential Form of the Drift-Diffusion Equation.

150 The general solution for the first-order non-homogenous ordinary differential equation (see e.g. Arfken and Weber (1999)), with $\langle w_r \rangle_*$ given by the model in the main text, is:

$$g(r) = \frac{1}{\beta(r)} \left[\int \beta(r) q(r) dr + C \right], \quad (S6)$$

with $q(r) = R_c^*/(\tau_\eta B_{nl} r^4)$; $\beta(r) = \exp \left[\int p(r) dr \right]$ and $p(r) = [A_\tau r - \langle w_r \rangle_*]/(\tau_\eta B_{nl} r^2)$. For the current model described in the main text, the integral in (S6) could not be expressed in terms of simpler canonical functions. Hence, for specific applications, we currently anticipate that some sort of power-law expansion or asymptotic reduction (if not numerical integration) would be needed to produce problem specific analytical approximations.

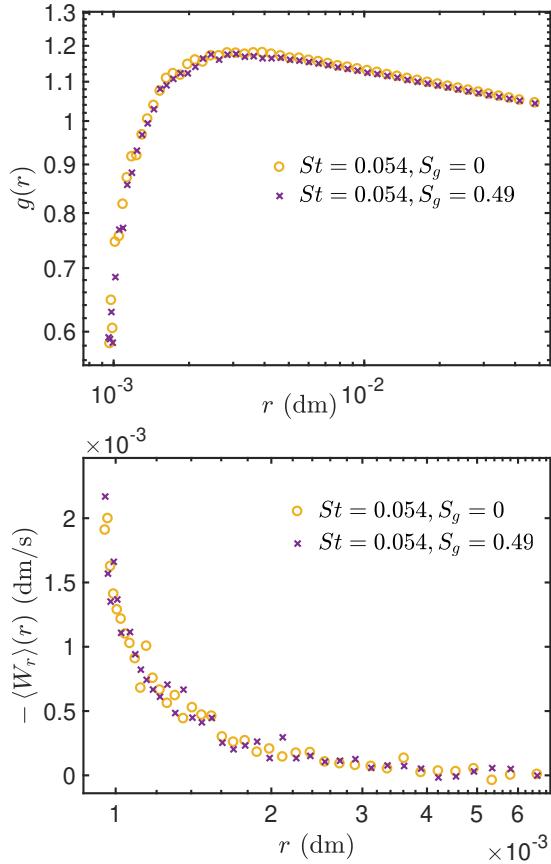


Figure S2. Top) RDFs of particles ($St = 0.054$) subject to action of turbulence, collision-coagulation with and without gravity. Circles: $S_g = 0$ (zero gravity); triangles: $S_g = 0.49$ (nonzero gravity). No discernible effect of gravity. **Bottom**) MRVs of the same cases. No discernible effect of gravity.

S10 Further Details on the Effects of Gravity.

We repeat the DNS case of $St = 0.054$ and $St = 0.54$ with the particles subjected to gravity (body force), and compares results with the zero-gravity case. Fig. 2 shows the results for case $St = 0.054$. There is no discernible difference between the cases
160 with and without gravity.

For case $St = 0.54$, the main RDF and MRV results is shown the main text. Here we show only the compensated-RDFs ($g_c(r)$), where each $g_c(r)$ is calculated via $g(r)$ divide by a power law ($c_0 r^{-c_1}$) that resulted from curve-fitting to the original $g(r)$ in the range $0.6\eta \leq r \leq 3\eta$. Fig. 3 compares $g_c(r)$ for cases with and without gravity. The fact that there is no discernible difference implies that the uncompensated $g(r)$ could be model as $g_c \times g_g$ where g_c is function that depends only on the particle
165 collision process while g_g depends on other factors e.g. gravity and is independent of particle collision.

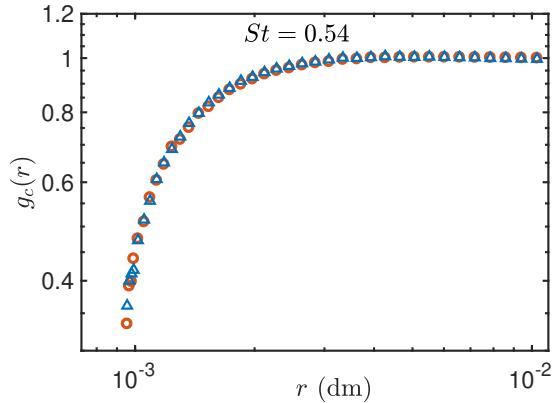


Figure S3. Compensated RDFs of particles ($St = 0.54$) subject to action of turbulence, collision-coagulation with and without gravity.

Circles: $S_g = 0$ (zero gravity); triangles: $S_g = 4.9$ (nonzero gravity). Interpretation in the text.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant 11872382) and by the Thousand Young Talent Program of China. We thank Jialei Song for helps. We thank Wai Chi Cheng, Jianhua Lv, Liubin Pan, Raymond A. Shaw for discussion and suggestions.

References

170 Arfken, G. B. and Weber, H. J.: Mathematical methods for physicists, 1999.

Bec, J., Biferale, L., Cencini, M., Lanotte, A., Musacchio, S., and Toschi, F.: Heavy Particle Concentration in Turbulence at Dissipative and Inertial Scales, *Phy. Rev. Lett.*, 98, 084 502, 2007.

Chun, J., Koch, D. L., Rani, S. L., Ahluwalia, A., and Collins, L. R.: Clustering of aerosol particles in isotropic turbulence, *J. Fluid Mech.*, 536, 219–251, 2005.

175 Kolmogorov, A. N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers, *Dokl. Akad. Nauk SSSR*, 30, 299–303, 1941.

Lu, J., Nordsiek, H., Saw, E. W., and Shaw, R. A.: Clustering of Charged Inertial Particles in Turbulence, *Phys. Rev. Lett.*, 104, 184 505, 2010.

Saw, E. W., Shaw, R. A., Ayyalasomayajula, S., Chuang, P. Y., and Gylfason, A.: Inertial Clustering of Particles in High-Reynolds-Number 180 Turbulence, *Phys. Rev. Lett.*, 100, 214 501, 2008.

Saw, E.-W., Salazar, J. P., Collins, L. R., and Shaw, R. A.: Spatial clustering of polydisperse inertial particles in turbulence: I. Comparing simulation with theory, *New Journal of Physics*, 14, 105 030, 2012a.

Saw, E.-W., Shaw, R. A., Salazar, J. P., and Collins, L. R.: Spatial clustering of polydisperse inertial particles in turbulence: II. Comparing simulation with experiment, *New Journal of Physics*, 14, 105 031, 2012b.

185 Yavuz, M., Kunnen, R., Van Heijst, G., and Clercx, H.: Extreme small-scale clustering of droplets in turbulence driven by hydrodynamic interactions, *Phys. Rev. Lett.*, 120, 244 504, 2018.