Measuring concentration with Voronoï diagrams: study of possible biases

Romain Monchaux

Measuring concentration with Voronoï diagrams: study of possible biases

Monchaux Romain

ENSTA-ParisTech Unité de Mécanique
chemin de la Hunière \\91761 Palaiseau cedex\\

E-mail: romain.monchaux@ensta-paristech.fr

Abstract. In the context of turbulent flows laden with inertial particles, accurate estimation of preferential concentration is particularly relevant. We have recently proposed to use Voronoï diagrams to estimate concentration fields from 2D imaging techniques implemented around wind tunnel experiments. Due to various experimental biases, the relevance of such an analysis gets questionable. In this article we show the robustness of the Voronoï analysis with respect to the three more important identified biases possibly present in such experiments.

PACS numbers: 47.55.Kf,47.80.Jk

Contents

1. Context and motivations

In the increasing number of studies on particle laden turbulent flows, three aspects are usually pointed out in the case of inertial particles: the preferential concentration or clustering, the possible increase in collision rate and the enhancement of the particle settling velocity since all of them are of prime importance regarding the various practical applications (pollutant dispersion, rain formation, optimisation of chemical reactors...). Preferential concentration plays a crucial role since it is clearly involved in the two other aspects; the collision kernel is expressed as the product of the particles radial relative velocity and of the particle radial distribution function, the latter containing the information related to the preferential concentration [17, 9]. Regarding the settling velocity enhancement, that has been reported both numerically and experimentally [13, 15, 1], the commonly invoked mechanism is based on the fact that particles tend to preferentially explore the downward side of turbulent eddies. Nevertheless, Aliseda and coworkers have shown that collective effects should be involved to explain the further increase in the settling velocity measured in their experiment when the seeding density is increased [1]. They found that the fall velocity conditional on the local concentration linearly increases with the latter suggesting the key role played by preferential concentration in this issue.

Properly simulating particle laden flows requires an accurate model for the particle dynamics equation that is lacking so far: only the point particle limit is usually considered [5, 6] and most of the simulations make further simplifications and often do not consider the back-reaction of the particles on the fluid phase. Thus, experimental investigations are required to assess the preferential concentration problem and its consequences on the settling velocity and collision rate enhancement. We study in the

present the possible biases arising from classical experiments regarding the quantification of preferential concentration.
The relevant parameters in this context are the particles diameter or their Stokes number defined as the ratio of the particle viscous relaxation time to the dissipation time scale of the carrier flow ($St = \tau_p/\tau_\eta$), the average seeding density and the Reynolds number. A first bias arises from the nature of the particles used: in air experiment, particles are usually water droplets generated from commercial or dedicated injectors that produce polydispersed particle populations leading to seeding consisting of a mixture of various Stokes number particles so that average Stokes numbers and corresponding uncertainties have to be defined.
Acquisitions of the concentration field requires imaging techniques. If 3D imaging is developing fast, it is still limited to cases when too few particles (thousands) are present in the measurement volume to allow proper study of concentration issues, consequently most of the work devoted to preferential concentration relies on 2D images obtained from cameras aiming at a particular region of the flow illuminated by a Laser sheet whose thickness L_{th} is of order 1 mm. In this case, experimentalists are studying a 3D phenomenon through projections into a 2D space and one may wonder how this projection affects the results and conclusions of their work.
Another bias arising from these imaging techniques is linked to the identification of the particles in the images. Typical water droplets experiments involve mixture of particles whose diameters are in the range $2 - 200$ μm [7, 1, 10, 11, 14]. On a typical image, rough detection of particles is performed through thresholding of the grey level, a more precise position being obtained from sub-pixel accuracy techniques that are relevant when particles span over several pixels in the images. The use of thresholds implies that many particles may not be detected and therefore the particle sets are artificially sub-sampled.
Whenever it is possible to clearly measure the coordinates of the particles, various methods can be implemented to estimate and/or accurately access the concentration field or its global properties: box counting methods, clustering index, correlation dimension, Minkowski functional, Voronoï diagrams... In a recent review, we have tried to present the advantages and drawbacks of these various techniques in the field of particle laden turbulent flows [8]. Here, we would like to focus more particularly on the technical issues raised by the use of Voronoï diagrams when studying concentration fields of particles in turbulent flows. We have introduced the use of this tool in the field in a recent paper to analyse experimental data obtained in a wind tunnel[7]. In the present, we will use various 3D numerical simulations to question the influence of the three biases introduced above. Section 2 presents the Voronoï analysis, the former experimental results and the numerical simulations used here. Section 3 is dedicated to the study of the various biases we consider. We finally discuss our results and draw possible conclusions in section 4.

2. Data, post-processing and former results

As we aim at addressing the possible biases arising in experiments using 2D imaging methods, we present the experimental setup and we summarize the main results presented in [7] that is representative of this family of experiments. We then describe the two numerical data sets we have employed to assess the importance of these biases. All the analysis being carried out using Voronoï diagrams, we first recall the basics of this tool.

2.1. Post-processing and definitions

As explained in [7] and further justified in [8], the use of Voronoï diagrams is particularly relevant for studying preferential concentration of inertial particles in turbulent flows. Given a set of particles, the corresponding Voronoï diagram is the unique decomposition of the nD space into independent cells associated to each particle. One Voronoï cell is defined as the ensemble of points that are closer to a particle than to any other. Use of Voronoï diagrams is very classical to study granular systems and has also been used to identify galaxy clusters. Voronoï diagrams computation is very efficient with the typical number of particles per set usually encountered (up to several hundred thousands). From their definition, it appears that the volume/area of a Voronoï cell is the inverse of the local concentration of particles ; therefore the investigation of Voronoï volumes/areas field is strictly equivalent to that of local concentration field. In the following, we will deal with volume regardless of the space dimension. As the mean value of the Voronoï volumes is nothing but the average concentration, we always normalise these volumes by their mean value to define normalised Voronoï volumes that will be referred to as volumes and denoted $\mathcal{V}$ throughout this article.
In simulations as well as in experiments, several 2D or 3D particle fields are obtained at various instants. We present statistics obtained from ensemble averaging over several time samples. Probability Density Function (PDF) and standard deviation of the normalised Voronoï volumes can be calculated for experimental and numerical datasets and compared to those expected for uniformly distributed particles (see [4] for a presentation of those expectations). Voronoï volumes PDF may be used to identify clusters of particles as follows. Voronoï PDFs for a typical experiment/simulation and for a uniform random process (URP) intersect twice: for low and high values of normalized Voronoï volume, corresponding respectively to high and low values of the local concentration, PDF associated to inertial particles is above that of URP, while the opposite is observed for intermediate volume values. This is consistent with the intuitive image of preferential concentration: inertial particles concentration field is more intermittent than the URP, with more probable *preferred* regions where concentration is higher than the uniform case and subsequently also more probable *depleted* regions where concentration is lower than in the uniform case. We consider the first intersection point $\mathcal{V}_c$ as an intrinsic definition of particle *clusters*: for a given dataset, Voronoï cells whose volume is smaller than $\mathcal{V}_c$ are considered to belong to a cluster. It appears that

cluster cells tend to be connected in groups of various sizes and shapes that we identify as clusters whenever they belong to the same connected component. We then analyze the geometrical structure of the identified clusters. More details are available in [7].

2.2. Original experiment

Here we present our former experimental setup and results that will be referred to as LEGI data.

2.2.1. Experimental setup Experiments are conducted in a large wind tunnel with a 0.75 m × 0.75 m square cross-section where an almost ideal isotropic turbulence is generated behind a grid whose mesh size is 7.5 cm. We can adjust the mean velocity from 3 to 15 m.s^{-1} (the turbulence level remaining relatively low, of the order of 3% at the measurement location and the anisotropy level between the transverse and longitudinal fluctuating velocities is smaller than 10%) and thus the Taylor micro scale Reynolds number R_λ. Inertial particles are water droplets generated by four injectors placed in the convergent part of the wind tunnel, one meter upstream the grid to insure a homogeneous seeding of the flow. According to the injection process, we are able to tune the average Stokes number and the average particle concentration C_0. We always consider regimes of relatively low particles volume loading (volume fraction in our experiments covers the range $2.10^{-6} < \phi_v < 3.10^{-5}$) so that no turbulence modulation by two-way coupling is expected to occur. Acquisitions are performed using a Phantom V12 high speed camera (Vison Research, USA) operated at 10 kHz and acquiring 12 bits images at a resolution of 1280 pixels × 488 pixels corresponding to a 125 mm (along x)× 45 mm (along y) visualization window on the axis of the wind tunnel (covering slightly less than an integral scale in the vertical y direction and almost two integral scales in the streamwise x direction), located 2.95 m downstream the grid. The camera is mounted with a 105 mm macro Nikon lens opening at $f/D = 2.8$. An 8 W pulsed copper Laser synchronized with the camera is used to generate a 2 mm (*i.e.* 3-4η) thick light sheet illuminating the field of view in the stream-wise direction. Particles are identified on the recorded images as local maxima with intensity higher than a prescribed threshold. Sub-pixel accuracy detection is obtained by locating the particles at the center of mass of the pixels surrounding the local maxima. More detailed description of the experimental and acquisition setups are available in [7].

2.2.2. Experimental results Here, we briefly recall the main results obtained experimentally and presented in [7]. By systematically varying the triplet of parameters (St, R_λ, C_0), we have shown that particle Voronoï volume distributions are always reasonably log-normal, so that preferential concentration can be quantitatively measured by a single scalar (the standard deviation of these distributions $\sigma_\mathcal{V}$). By plotting $\sigma_\mathcal{V}$ versus the Stokes number, the maximum of heterogeneity of the concentration field

for particles with Stokes numbers around unity has been successfully recovered. We have characterized clusters (and voids) geometries and their inner concentration: the cluster areas/volumes ($\mathcal{A}_c$) are algebraically distributed and their structure is fractal (their perimeter $\mathcal{P}_c$ and their area $\mathcal{A}_c$ are not linearly related); in particular clusters do not appear to have any characteristic typical scale. The analysis of particles normalised concentration inside the clusters (C/C_0) have revealed two new and so far unpredicted results: *(i)* average particle concentration inside clusters depends on the global particle loading in a non trivial way and *(ii)* after compensation of this particle loading dependency, average concentration inside clusters exhibits a non monotonic dependency on Stokes number, with a maximum around unity values. Figure 1 gathers all these results.

2.3. Numerical simulations

Our goal in this article is to use 3D Direct Numerical Simulation (DNS) data to produce, as much as possible, sets of data that exhibit biases similar to those naturally present in experiments and to analyse them as experimental data to estimate the consequences of the considered biases on the measured results. We have benefited from two sets of numerical data produced by similar methods. The first set is available on line at `http://mp0806.cineca.it/icfd.php` and has been used and presented in [3], the second set provided by Susumu Goto has been used in [16]. They will be referred to as CINECA and GOTO data respectively. Both of them consist of 3D DNS of isotropic homogeneous turbulence produced in periodic cubic boxes of size $(2\pi)^3$. They employ the simplest model for small heavy particles [5, 6]:

$$\frac{d}{dt}\mathbf{v}_p(t) = \frac{1}{\tau_p}(\mathbf{u}(\mathbf{x}_p(t), t) - \mathbf{v}_p(t)), \tag{1}$$

where $\mathbf{v}_p(t)$ and $\mathbf{x}_p(t)$ are the particle velocity and position at time t and $\mathbf{u}(\mathbf{x}, t)$ is the surrounding fluid velocity field at position $\mathbf{x}$. The latter corresponds to a statistically homogeneous isotropic stationary turbulence field obtained a priori for an incompressible fluid and is used as a frozen forcing to solve equation (1) for various types of particles defined through the Stokes drag coefficient $\tau_p = 2\rho_p a^2/9\mu$ where ρ_p and a are the particle density and radius respectively and μ is the fluid viscosity.
The fluid velocity field $\mathbf{u}(\mathbf{x}, t)$ is obtained by solving the Navier-Stokes equations:

$$\frac{\partial}{\partial t}\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\frac{1}{\rho_f}\nabla p + \nu\nabla^2\mathbf{u} + \mathbf{f}, \tag{2}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{3}$$

where ν and ρ_f are the fluid kinematic viscosity and density respectively, $p(\mathbf{x}, t)$ is the fluid pressure field and $\mathbf{f}$ is an external forcing implemented by fixing the amplitudes of Fourier modes in the low-wavenumber region. The numerical grids used set the associated Taylor micro scale Reynolds numbers R_λ. The relevant parameters defining the simulations are provided in table 1. The number of particles in the simulations

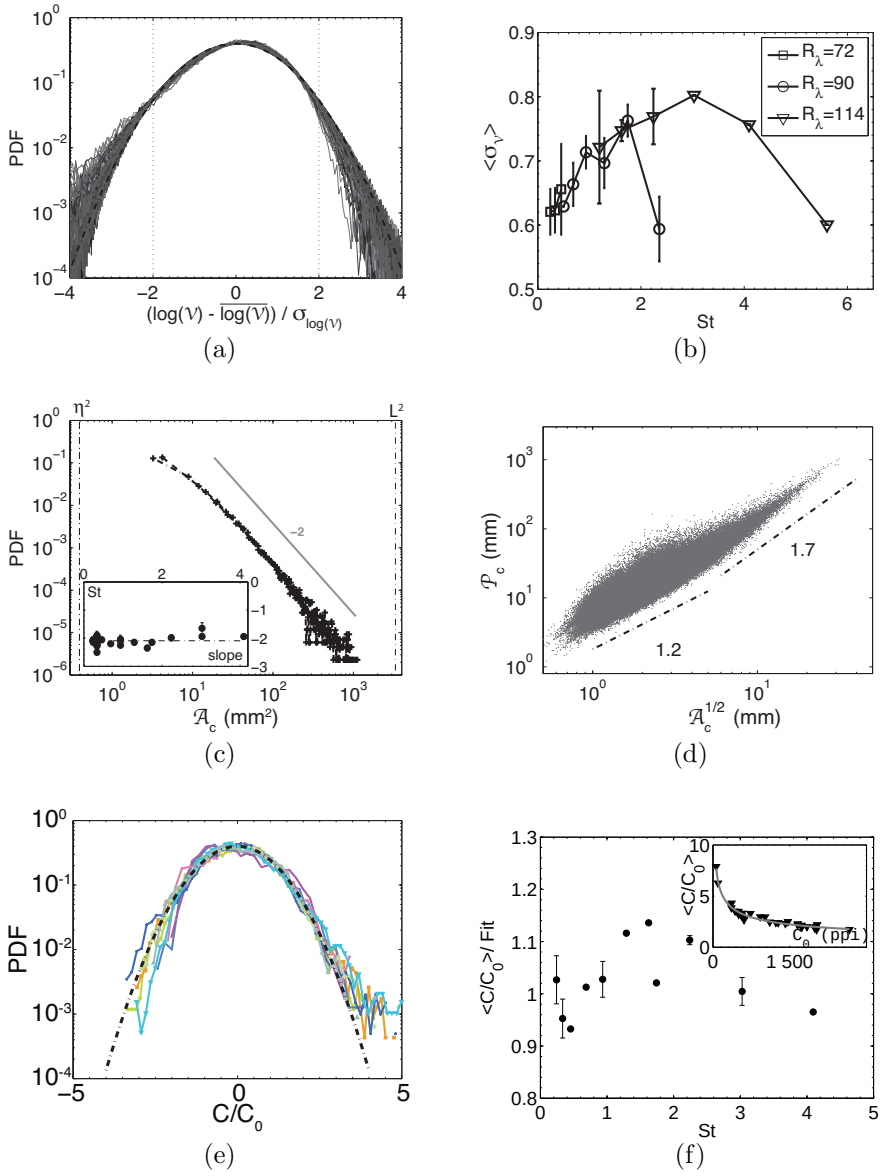

Figure 1: a: centered and normalized PDF of the logarithm of Voronoï area for 40 experiments performed at LEGI; black dashed line represents a Gaussian distribution. b: standard deviation of Voronoï areas as a function of average Stokes number; Reynolds number is constant along lines, each point is estimated as the averaged standard deviation from experiments with same Stokes number (but possibly different average concentration), error-bars represent the dispersion between such experiments. c: PDFs of clusters area, inset shows the evolution of the fitted power law exponent with Stokes number for the 40 experiments performed at LEGI, vertical dash-dotted lines indicates η^2 (left) and L^2 (right). d: geometrical characterization of clusters for the same 40 experiments. e: PDF of normalized-reduced concentration C/C_0 within clusters. f: evolution of the means of PDFs in figure (e) with the Stokes number after compensation by the seeding concentration dependency displayed in the inset.

	LEGI	CINECA	GOTO
grid points	1280×488	512^3	512^3
ϵ	$0.085\text{m}^2.\text{s}^{-3}$	0.877	0.262
R_λ	114	180	190
L	61.4mm	4.6	1.87
λ	$\simeq$ 5mm	0.24	0.15
η	0.45mm	0.01	0.00555
$k_{max}\eta$	NA	3.35	1.34
St	$0.1 - 6$	$0 - 4$	$0 - 10$
L_{th}	$2\eta - 6\eta$	$2\eta - 6\eta$	$2\eta - 6\eta$
N_0	$300 - 6000$	128×10^3	16×10^6
$\overline{\Delta_p}$	$2\eta - 8\eta$	12η	4η

Table 1: Parameters involved in the DNSs (dimension less) and experiments. ϵ: the energy dissipation rate, λ: the Taylor micro scale, R_λ: the associated Reynolds number, L and η the integral and Kolmogorov scales, $k_{max}\eta$ the largest wave number in the simulation in units of η, L_{th}: the 2D slices thickness, N_0: the original number of particles per instant/snapshot and $\overline{\Delta_p}$: the average distance between neighbouring particles

and the considered Stokes numbers vary from one simulation to another, they are also recalled in table 1.

3. Study of bias

3.1. 2D/3D bias

As mentioned in the introduction, most of the current concentration measurements performed in particle laden turbulent flows consist of 2D imaging of particular region of the flow illuminated by a Laser sheet whose thickness is ranging from 0.5 to 2 mm. The consequence of such a technique is that particles living in a 3D space are projected onto a 2D slice. If this may have little implications regarding Particle Image Velocimetry (PIV) since in this case the relevant parameter is the particle velocity that should not much change along the sheet thickness in most of the applications, it turns to be an issue when one deals with the particle concentration since it may artificially increase the concentration locally (think for example about a case where particles would be organised in structures or clusters whose typical relative distance would be smaller than the Laser thickness). In order to estimate the resulting bias, from original DNS boxes of size $(2\pi)^3$, we have defined several slices whose surface along the two first coordinates is $(2\pi)^2$, whose extension along the third coordinate is varied from $L_{th} = 2\eta$ to $L_{th} = 6\eta$ and who are centered at various positions between 0 and 2π. For each slice, we keep only the two first coordinates and we calculate the associated Voronoï tessellations. This exactly mimics experimental 2D imaging of a 3D flow with usual Laser sheets.
Figures 2 and 3 present the resulting normalised Voronoï area distributions and the dependency of the Voronoï area standard deviation with the Stokes number for various slice thickness for GOTO and CINECA data. As for the experimental data, the PDFs of the Voronoï areas are wide, covering more than four decades. The one obtained from

GOTO are close to log-normal (as in the experimental data) while the one obtained from CINECA are not. The shape of the one obtained from GOTO changes when L_{th} is increased which is not the case for the CINECA's ones. Nevertheless, all of them are well described by their standard deviation. When the slice thickness is increased, the standard deviation dependency on the Stokes number is qualitatively preserved over the available range of Stokes numbers. Nevertheless, the exact values of the standard deviation show a systematic increase with the slice thickness. We report a maximal increase of 25% when the slice thickness is increased by 150% for both sets of numerical data. Surprisingly, the standard deviations obtained from GOTO data are twice as big as the one obtained from CINECA data that are closer to the experimental one. The discrepancies between GOTO and CINECA data on the standard deviation values, on the PDFs shapes and on their shape change with L_{th} will be discuss when adressing subsampling issues in 3.2.
A very interesting feature is that for both simulations, 2D and 3D post-processing lead to very similar quantitative results that legitimates the experimental use of 2D slices to investigate a 3D phenomenon.

3.2. Sub-sampling bias

Another bias that may come from imaging techniques is an artificial sub-sampling of the data sets. As in any detection problem, two kind of errors are expected: wrong detections and missed detections (false positive and false negative in receiver operating characteristic curve vocabulary). The former overestimates the number of particles while the latter underestimates it. Usually, the signal processing chain implemented in such a case is designed so that there is no wrong detection and so that the number of particles found varies very slowly with the chosen threshold whenever this is possible. As a consequence, there are necessarily several if not many missed detections and experimental data sets are usually sub-sampled.
Sources of missed particles are numerous. One is particles that remain in the shadow of other, another is due to the use of thresholds for particles detection that implies that some particles have to be assimilated to the noisy background. We believe that the latter is the main source of missed detection and this for several reasons: (i) smaller particles appear less luminous than bigger ones, (ii) as the laser sheet intensity is decreasing toward its edges, particles in the center of the sheet appear more luminous than the others, (iii) due to the average seeding and to the Laser sheet thickness, the resulting concentration on 2D images is not very high and the number of shadowed particles might be very small compared to the number of missed particles due to (i) and (ii).
To test the impact of this unavoidable sub-sampling, we have created extra data sets from data sets obtained with a given threshold, by keeping less and less particles from the original data set. Particles were removed by randomly picking them up from the original dataset with a uniform distribution. The same procedure has been applied to 3D DNS data. From an original dataset consisting of N_0 particles, we produce new data

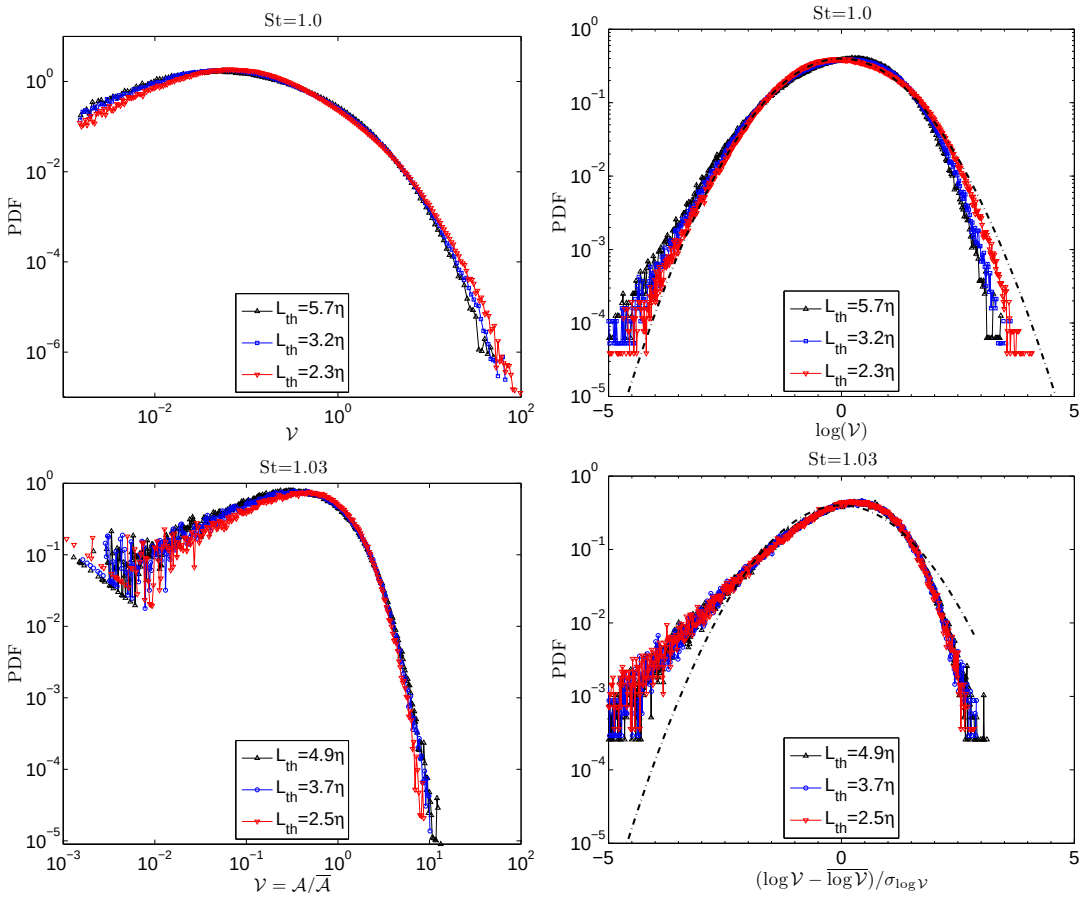

Figure 2: .PDFs of normalised Voronoï areas for 2D slices obtained from GOTO data (top) and CINECA data (bottom). Left: $\mathcal{V} = A/\overline{A}$. Right: centered and normalized PDFs of the logarithm of normalized Voronoï volume.

sets with only αN_0 particles, $\alpha \in [0, 1]$ being the ratio of kept particles.
Results are presented on figure 4. The top left figure shows the changes in the normalised Voronoï area PDF obtained for one experiment at $St = 0.25$ taken from LEGI data when the number of particles is artificially reduced down to 1% of the original number and the top right figure shows the associated standard deviation $\sigma_{\mathcal{V}}$ as a function of the ratio of particles kept. It is seen that the PDFs are not much affected for ratios down to 40% of the original number. This result is confirmed with the standard deviation that shows a change by less than 2% when the number of particles is decreased by more than a factor 2. Note that when only 1% of the particles are kept, the resulting standard deviation corresponds to that of a uniform distribution of particles. The bottom left and right figures on fig. 4 respectively present the dependency of the Voronoï volumes standard deviation on the sub-sampling ratio and the Stokes number for all GOTO data associated to 2D slices of thickness 5.7η. As for the experimental data, the effect of sub-sampling is a monotonous decrease of the estimated Voronoï standard deviation. Whatever the value of the Stokes number, when a few percent of the original particles are kept, the standard deviation is getting very close to the value expected for a uniform random process.

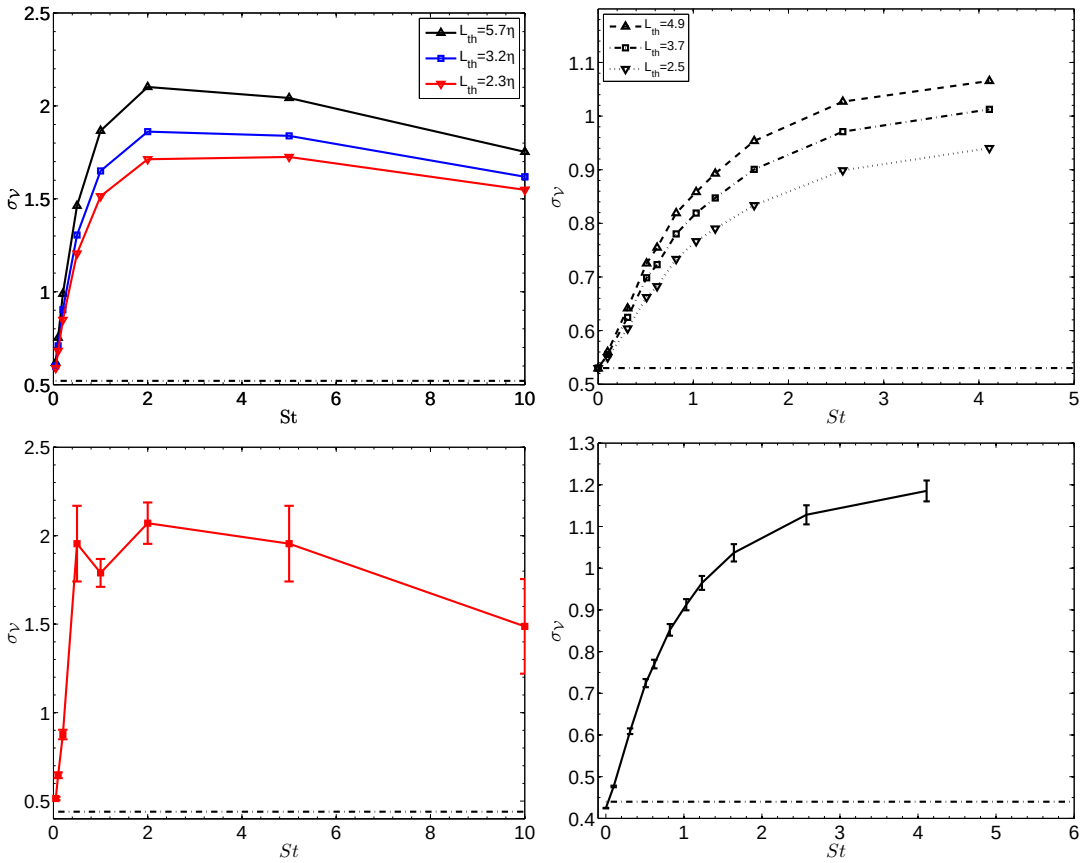

Figure 3: Standard deviation of the normalised Voronoï area distributions as a function of the Stokes number for three different slice thickness. Left: GOTO data, right: CINECA data. Top: 2D data, bottom: 3D data. The dash-dotted lines correspond to the value expected for uniform random distributions of particles.

For small and intermediate values of the sub-sampling (say above 40%) the qualitative dependency upon the Stokes number is preserved while the absolute values are decreased by at most 25%, which is larger that what has been observed from the experimental data sets. It is worth noticing that the plateau exhibited for subsampling values ranging from 40 to 100% in the experimental data is not present anymore in GOTO data. When the subsampling is larger (values below 40%), the qualitative dependency of $\sigma_{\mathcal{V}}$ upon the Stokes number is affected and does not present maximum anymore. The corresponding curves look similar to those obtained from CINECA data that were computed from much smaller particle samples (see table 1). Tagawa and collaborators found consistent values from another set of numerical data[12]. The shift observed on the curve $\sigma_{\mathcal{V}} = f(St)$ when half the particles are removed as to be linked to results presented in 3.1: we noticed that CINECA and GOTO data lead to different values of $\sigma_{\mathcal{V}}$ and to different shapes for the $\mathcal{V}$ PDFs. The only difference between both simulation is the number of particles that is 125 times larger in GOTO data. As a consequence, the use of Voronoï diagrams that does not involve any coarse graining leads to different scales of analysis.

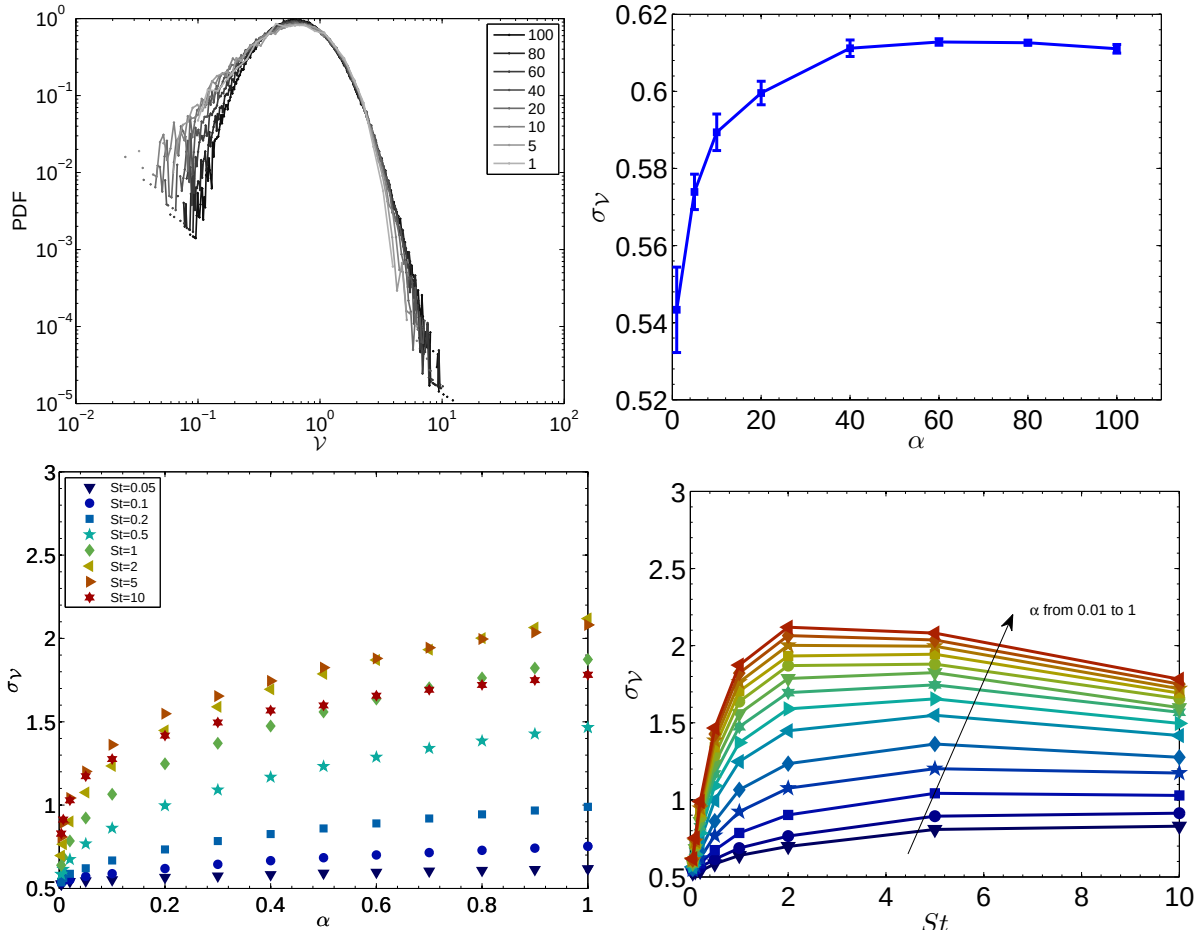

Figure 4: Top left: PDFs obtained using various sub-sampling ratio from one original LEGI data experiment at $St = 0.25$. Top right: dependency of $\sigma_\mathcal{V}$ upon this ratio α for the same data. Bottom: dependency upon the sub-sampling ratio α (left) and St (right) of $\sigma_\mathcal{V}$ for data created from original 2D slices of thickness $L_{th} = 5.7\eta$ from GOTO data.

GOTO data allow probing at dissipative scales while CINECA data are fully in the inertial range and as a consequence both datasets lead to different results.
All the results presented in 2.2.2 have been recovered from the sub-sampled experimental datasets. We illustrate this on figure 5 which presents the PDFs of the identified cluster areas for various sub-sampling ratios. PDFs are compensated by $\mathcal{A}_c^2$ to evidence that the algebraic behaviour with a -2 exponent is always roughly preserved even if more than 50% of the particles have to be kept to observe the scaling over more than one decade. Direct visualisations of the non compensated PDFs reveal that the departure from the -2 scaling law appear mostly in the tails that shrink when the number of particles kept is decreased (see inset).

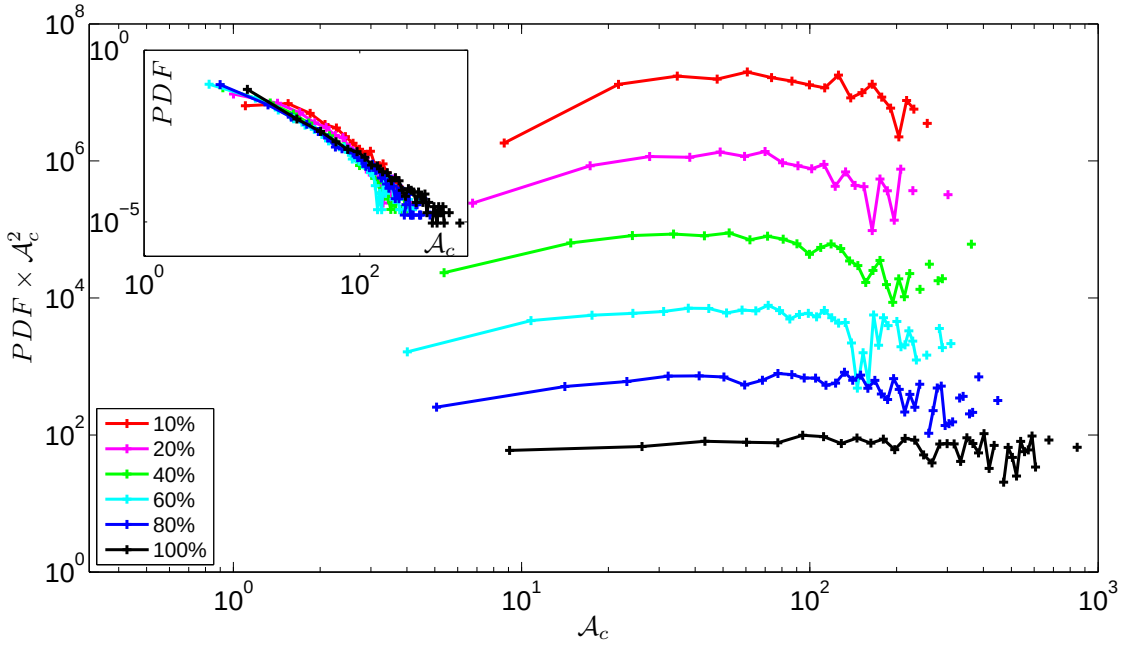

Figure 5: PDFs of clusters areas estimated from the sub-sampled data sets used in fig. 4. PDFs are compensated by a $\mathcal{A}_c^2$ to evidence the -2 power-law scaling and are arbitrarily shifted for clarity. Inset: same PDFs without the compensation. For smaller values of α, clusters can not be identified anymore.

3.3. Mixing Stokes numbers

Figure 6 left presents the distribution of the particles diameters obtained in LEGI experiment for four values of the air pressure varying from 2 to 5 bars at fixed water flow rate (1.2 l/mn for each injector) and fixed wind velocity ($V_0 = 4.5$ m.s^{-1}). Each of them is widely distributed over more than one decade and estimation of a Stokes number implies to use an average or a most probable Stokes number defined from the average or the mode of these PDFs (we actually used the mode in [7]). In both cases, because of the high polydispersity, the standard deviation σ_{St} of Stokes number (based on measured diameters distribution), which could be interpreted as an error-bar for the Stokes number estimation, is large (σ_{St}/St easily exceeds 50%). The same happens in any wind tunnel experiment seeded with liquid dropplets [1, 10] and one may wander how this polydispersity impacts the dependency upon the Stokes number of the various quantities measured in experiments. To tackle this issue, we have built polydisperse data set from GOTO data projected on 2D slices.

In Goto's work, each Stokes number data set was computed using the same fluid DNS (see 2.3). As a result, gathering particle fields of various Stokes number makes sense and properly mimics polydispersity. For each experimental condition, we have estimated a particle diameter distribution similar to those shown on figure 6 left. Using the definition of the Stokes number used in this article, these PDFs can be expressed in

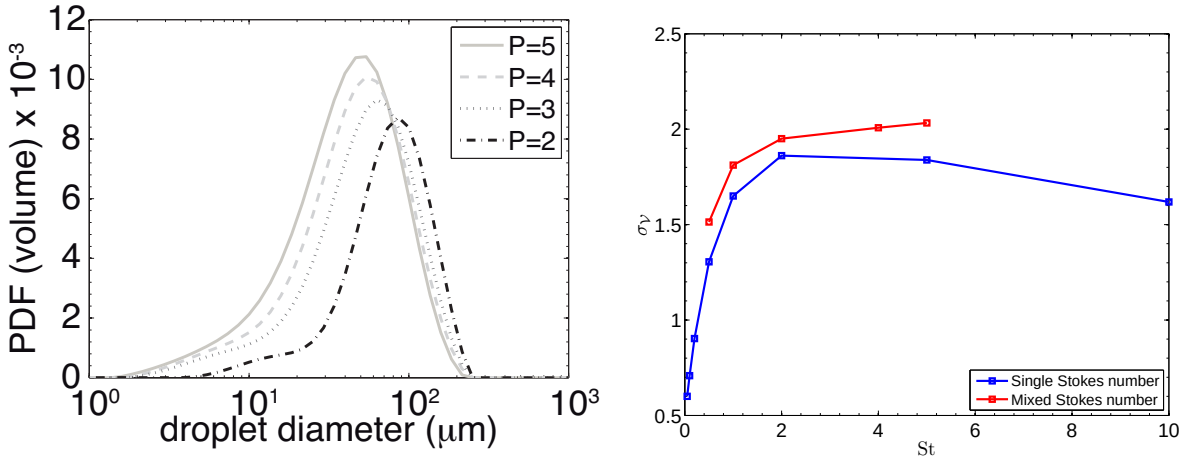

Figure 6: Left: particles diameter Probability Density Function evolution with air pressure (varying from 2 to 5 bars) at fixed water flow rate (1.2 l/mn for each injector) and fixed wind velocity ($V_0 = 4.5$ m.s^{-1}); note that PDFs are calculated from particles volume (and not particles number).

terms of the Stokes number rather than in the particle diameter. In Goto's work, 8 sets of particles associated to 8 values of the Stokes number have been simulated. To mimic one particular experiment, we pick randomly a certain amount of particles from each of these 8 data sets. The relative proportion taken from each of them is given by the presented experimental particle diameter PDF. Figure 6 right presents $\sigma_\mathcal{V}$ as a function of St for single Stokes data (see figure 3) and for mixed data built as described above. In spite of the short range of Stokes numbers covered by the polydispersed data, we can see that quantitative values are changed by less than 15%. The qualitative behaviours of these two curves are close, even if the one associated to single Stokes number data presents a maximum while the other does not. Nevertheless, this maximum relies only on the point at $St = 10$ and on the range where both curves are present, the two behaviours are similar.

4. Discussion and conclusions

2D/3D biases studied in section 3.1 and summarized on figure 3 have been shown to affect less the obtained results. Qualitative results are always recovered when considering slices of thickness $L_{th} \simeq 2\eta - 6\eta$. Quantitatively, the estimated values of $\sigma_\mathcal{V}$ are also not much affected in this range of slices thickness. This is encouraging regarding the reliability of experimental results obtained from 2D acquisitions performed in particle laden turbulent flows.

The study of sub-sampling experimental and numerical data shows that none of the so far obtained results are affected qualitatively nor quantitatively when half of the detected particles are removed from the original data. This demonstrates the robustness of the Voronoï analysis with respect to the almost unavoidable missed detections occurring

when processing experimental data. On the other hand, the impact of sub-sampling seems slightly more important for GOTO data than for LEGI data and we noticed several discrepancies between GOTO and CINECA data. We may wander about the reasons explaining these important differences. Considering values reported in table 1, it is seen that the ratio between λ, the Taylor micro-scale and η, the Kolmogorov scale is larger for GOTO data. As a consequence, when sub-sampling the data from the experiment, we actually probe scales fully included within the inertial range while when sub-sampling in GOTO data we probe scales ranging from the dissipative to the deep inertial range. Similar reasoning can be done on the global seeding (N_0) differences between GOTO and CINECA data. Yoshimoto and Goto [16] have shown that preferential concentration is self similar within the inertial range while Bec and coworkers [2] insisted on the crucial differences between the mechanisms involved to achieve preferential concentration in the dissipative and inertial range. The differences observed here between the numerical and the experimental data on the one hand and between GOTO and CINECA data on the other hand could thus be explained from this considerations about the involved scales and the global seeding of the flow. This is also consistent with the self similarity observed in our experimental results [7]. Similarly, differences observed between GOTO and CINECA data can also be understood from this scale argument and help us formulating the obvious warning that one should be very careful about the scales involved in an experiment or in a simulation according to the particle seeding density: changing the seeding density (or the number of particles used for Voronoï calculations) has an impact on the scales that can be probed.
Regarding the polydispersity problem, we have also shown that qualitative behaviour of $\sigma_{\mathcal{V}}$ as a function of St is preserved and that associated absolute values are slightly affected by polydispersity. This shows that the maximum enhancement of preferential concentration observed around $St \simeq 1-2$ is a very robust effect that is even recovered when various types of particles are mixed together.
Even if the present study does not address all the possible biases arising from imaging measurement of concentration in particle laden flows, we have considered the more relevant in the context of Voronoï analysis. The same work could be undertaken for the other classical preferential concentration estimators such as pair correlation or box counting methods for example.

Acknowledgments

This work has been initiated during my stay in Grenoble to answer the various questions emanating from my colleagues Mickaël Bourgoin, Alain Cartellier, Romain Volk and Jean-François Pinton who are warmly thanked. I also acknowledge the important effort done by the various ceramicists who feed the CINECA data base and more particularly Enrico Calzavarini and Federico Toschi who helped me in reading the data used in this article. I am also indebted to Susumu Goto who provided me half the numerical data processed to write this article. I am also grateful to the NJP editors who invited me to

write this article.

[1] A. Aliseda, A. Cartellier, F. Hainaux, and J. C. Lasheras. Effect of preferential concentration on the settling velocity of heavy particles in homogeneous isotropic turbulence. *Journal of Fluid Mechanics*, 468:77–105, October 2002.

[2] J. Bec, M. Cencini, and R. Hillerbrand. Clustering of heavy particles in random self-similar flow. *Phys. Rev. E*, 75(2):025301–+, February 2007.

[3] E. Calzavarini, M. Kerscher, D. Lohse, and F. Toschi. Dimensionality and morphology of particle and bubble clusters in turbulent flow. *Journal of Fluid Mechanics*, 607:13–24, 2008.

[4] J.-S. Ferenc and Z. Néda. On the size distribution of Poisson Voronoi cells. *Physica A Statistical Mechanics and its Applications*, 385:518–526, November 2007.

[5] R. Gatignol. The Faxen formulae for a rigid particle in an unsteady non-uniform Stokes flow. *Journal de Mecanique theorique et appliquee*, 2:143–160, 1983.

[6] M. R. Maxey and J. J. Riley. Equation of motion for a small rigid sphere in a nonuniform flow. *Physics of Fluids*, 26:883–889, April 1983.

[7] R. Monchaux, M. Bourgoin, and A. Cartellier. Preferential concentration of heavy particles: A Voronoi analysis. *Phys. Fluids*, 22:103304, 2010.

[8] R. Monchaux, M. Bourgoin, and A. Cartellier. Analyzing preferential concentration and clustering of inertial particles in turbulence. *International Journal of Multiphase Flow*, 40(0):1–18, 2012.

[9] W. C. Reade and L. R. Collins. Effect of preferential concentration on turbulent collision rates. *Physics of Fluids*, 12:2530–2540, October 2000.

[10] J. P. L. C. Salazar, J. De Jong, L. J. Cao, S. H. Woodward, H. Meng, and L. R. Collins. Experimental and numerical investigation of inertial particle clustering in isotropic turbulence. *Journal Of Fluid Mechanics*, 600:245–256, April 2008.

[11] E. W. Saw, R. A. Shaw, S. Ayyalasomayajula, P. Y. Chuang, and Á. Gylfason. Inertial Clustering of Particles in High-Reynolds-Number Turbulence. *Physical Review Letters*, 100(21):214501–+, May 2008.

[12] Y. Tagawa. private communication. 2011.

[13] L. P. Wang and M. R. Maxey. Settling velocity and concentration distribution of heavy particles in homogeneous isotropic turbulence. *Journal of Fluid Mech.*, 256:27–68, 1993.

[14] Z. Warhaft. INVITED REVIEW: Laboratory studies of droplets in turbulence: towards understanding the formation of clouds. *Fluid Dynamics Research*, 41(1):011201–+, February 2009.

[15] C. Y. Yang and U. Lei. The role of the turbulent scales in the settling velocity of heavy particles in homogeneous isotropic turbulence. *Journal of Fluid Mechanics*, 371:179–205, September 1998.

[16] H. Yoshimoto and S. Goto. Self-similar clustering of inertial particles in homogeneous turbulence. *Journal of Fluid Mechanics*, 577:275–+, April 2007.

[17] L. I. Zaichik, O. Simonin, and V. M. Alipchenkov. Two statistical models for predicting collision rates of inertial particles in homogeneous isotropic turbulence. *Physics of Fluids*, 15:2995–3005, October 2003.