# Data Analysis of Particle Clustering in Turbulence



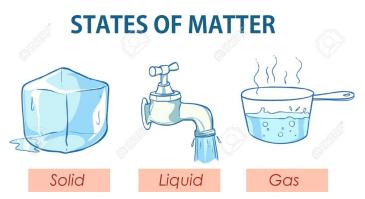
Reza Momenifar & Jonathan Holt

Civil and Environmental Enginnering (CEE), Pratt School of Engineering

Machine Learning and Data Mining (STA 325L-001) course Fall 2019

- Turbulence
- Particle Clustering
- Clustering Analysis (or Unsupervised Learning) via Voronoï Tessellation
- Data Expedition Task

- Fluids, liquids and gases, have the ability to flow.
- Fluid mechanics: the study of how fluids move and the forces on them!



Understanding and predicting how fluids move is of enormous importance to a vast range of problems:



However, in many other problems, the fluid motion is extremely complex, and understanding and predicting the fluid motion and its effect on natural problems is an enormous challenge.

However, in many other problems, the fluid motion is extremely complex, and understanding and predicting the fluid motion and its effect on natural problems is an enormous challenge.

The difference is that in these problems the fluid motion is *turbulent*.

**An extremely brief intro to turbulence:** Depending upon the flow properties (velocity, boundary conditions, viscosity of fluid etc), the motion of a fluid can be either *LAMINAR* or *TURBULENT*, and these are *radically different*!

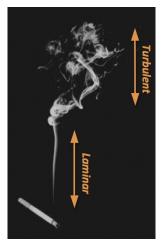
However, in many other problems, the fluid motion is extremely complex, and understanding and predicting the fluid motion and its effect on natural problems is an enormous challenge.

The difference is that in these problems the fluid motion is *turbulent*.

**An extremely brief intro to turbulence:** Depending upon the flow properties (velocity, boundary conditions, viscosity of fluid etc), the motion of a fluid can be either *LAMINAR* or *TURBULENT*, and these are *radically different*!

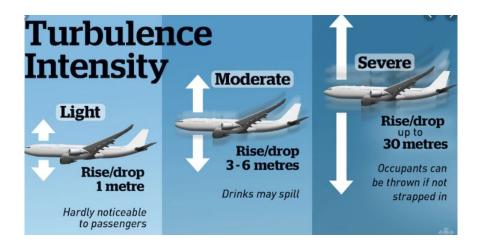
Rising smoke illustrates the difference between laminar and turbulent flows:

Rising smoke illustrates the difference between laminar and turbulent flows:



- Laminar flow is orderly, regular, predictable.
- Turbulent flow looks random, chaotic, irregular, unpredictable (depending on the intensity of turbulence).

Example: Airplane Turbulence.



Turbulence: easy to observe, but extremely difficult to understand.

Turbulence: easy to observe, but extremely difficult to understand.

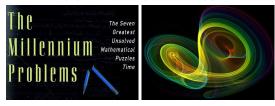


Feynman: "turbulence is the last great unsolved problem in classical physics".

Turbulence: easy to observe, but extremely difficult to understand.



Feynman: "turbulence is the last great unsolved problem in classical physics". It is also connected to deep problems in pure and applied mathematics.



However, turbulence is not only a very intellectually rich and stimulating problem to work on, but it is also extremely important practically.

Example: Turbulence enhances mixing.

Example: Turbulence enhances mixing.

You have probably already observed this:

Example: Turbulence enhances mixing.

You have probably already observed this:



Example: Turbulence enhances mixing.

You have probably already observed this:



This has great implications...for air pollution, combustion, chemical reactions, heat transfer...

Example: Turbulence enhances mixing.

You have probably already observed this:



This has great implications...for air pollution, combustion, chemical reactions, heat transfer...

In fact, turbulence has great implications for understanding and predicting a wide variety of problems...



**Cosmos:** turbulence is important because it distributes magnetism, disperses heat from supernova events and even plays a role in planet and star formation.



**Cosmos:** turbulence is important because it distributes magnetism, disperses heat from supernova events and even plays a role in planet and star formation.

**Volcano:** the dispersion of ash in the atmosphere is controlled by atmospheric turbulence, with implications for the environment and aviation.



**Cosmos:** turbulence is important because it distributes magnetism, disperses heat from supernova events and even plays a role in planet and star formation.

**Volcano:** the dispersion of ash in the atmosphere is controlled by atmospheric turbulence, with implications for the environment and aviation.

**Plankton:** play a major role in the carbon cycle, and turbulent mixing impacts their population dynamics.



**Cosmos:** turbulence is important because it distributes magnetism, disperses heat from supernova events and even plays a role in planet and star formation.

**Volcano:** the dispersion of ash in the atmosphere is controlled by atmospheric turbulence, with implications for the environment and aviation.

**Plankton:** play a major role in the carbon cycle, and turbulent mixing impacts their population dynamics.

**Clouds:** are important not only for weather, but also for climate, and turbulence controls the thermodynamics of clouds, radiative properties, and the rate at which droplets grow to form rain.

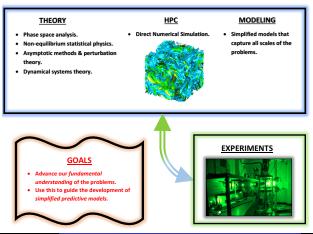
But, we do not really *understand* how turbulence affects these problems, nor are we able to *predict* its effects very well.

But, we do not really *understand* how turbulence affects these problems, nor are we able to *predict* its effects very well.

So there are two aspects to solving these problems: **Understanding** (inference) and **Prediction**.

But, we do not really *understand* how turbulence affects these problems, nor are we able to *predict* its effects very well.

So there are two aspects to solving these problems: **Understanding** (inference) and **Prediction**.



Particle-laden turbulent flows appear in a diverse range of engineering systems and natural phenomena.

In clouds, this is related to how turbulence affects the dynamics of water droplets and ice crystals, how they collide and mix etc.

Particle-laden turbulent flows appear in a diverse range of engineering systems and natural phenomena.

In clouds, this is related to how turbulence affects the dynamics of water droplets and ice crystals, how they collide and mix etc.



In clouds (thunderstorm and non-thunderstorm), particles (droplets or ice crystals) are often charged.

Particle-laden turbulent flows appear in a diverse range of engineering systems and natural phenomena.

In clouds, this is related to how turbulence affects the dynamics of water droplets and ice crystals, how they collide and mix etc.



In clouds (thunderstorm and non-thunderstorm), particles (droplets or ice crystals) are often charged.

How do the electric forces compete with the turbulent air to control the particle collisions and mixing rates?

Particle-laden turbulent flows appear in a diverse range of engineering systems and natural phenomena.

In clouds, this is related to how turbulence affects the dynamics of water droplets and ice crystals, how they collide and mix etc.

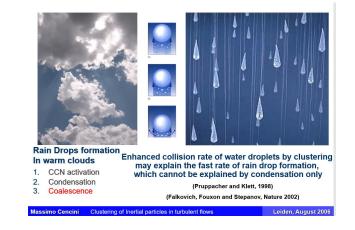


In clouds (thunderstorm and non-thunderstorm), particles (droplets or ice crystals) are often charged.

How do the electric forces compete with the turbulent air to control the particle collisions and mixing rates?

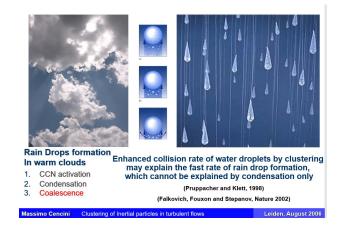
What are the implications for the electrodynamics of clouds?

#### Problem of interest: rain formation

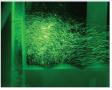


• The particle properties (particularity their inertia and fall speed) can significantly change the interaction of particles with the turbulent flow

Problem of interest: rain formation



 Indeed, inertial particles are distributed non-uniformly in the turbulent flow. In addition, the presence of gravity results in finite settling velocities of inertial particles



Fluid Streamlines (Polystyrene spheres 25 µm in diameter are used as tracers in experiments and illuminated by a laser with)



Inertial particles trajectories

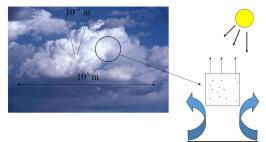


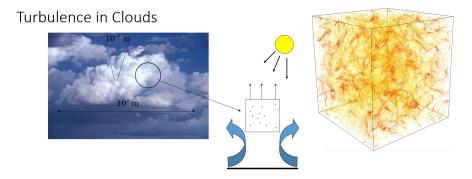
Snapshots of particle distributions in a turbulent flow field



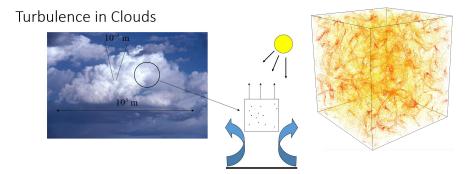
- The non-uniform distribution of particles can lead to the formation of clusters.
- Clusters, the group regions of highly concentrated particles, which can significantly change the structure of the turbulence and produce substantial inhomogeneities in the spatial concentration of particles.

#### Turbulence in Clouds

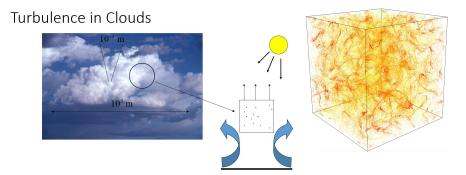




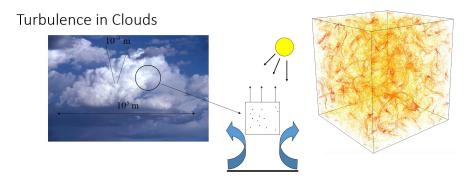
• We employ Direct Numerical Simulations (DNS) to model turbulent flow and track particles lying under an idealized turbulence in a cubic box.



• In these simulations, three independent control parameters representing the properties of turbulent flow and particles are varied and particles' position and other dynamical properties of particles and flow fields (e.g. velocity) are stored.



• More specifically, we want to explore how intensity of turbulence (quantified by Reynolds number or Re), gravitational acceleration (quantified by Froud number or Fr) and particles' characteristics (e.g. size, density which is quantified by Stokes number or St) affect the spatial distribution and clustering of particles in an idealized turbulence (homogeneous isotropic turbulence or HIT).



#### • Numerical Simulation of Particle Distribution in Turbulence\*\*video .

• High-resolution numerical simulation of the turbulent flow field can be achieved by employing direct numerical simulation (DNS) of Navier-Stokes equations (conservation of mass and momentum) which can resolve all scales of motion.

 High-resolution numerical simulation of the turbulent flow field can be achieved by employing direct numerical simulation (DNS) of Navier-Stokes equations (conservation of mass and momentum) which can resolve all scales of motion.

$$\partial_t \boldsymbol{u} + \boldsymbol{\omega} \times \boldsymbol{u} + \boldsymbol{\nabla} \left( \frac{p}{\rho_f} + \frac{\|\boldsymbol{u}\|^2}{2} \right) = \nu \nabla^2 \boldsymbol{u} + \boldsymbol{f}, \, \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

• High-resolution numerical simulation of the turbulent flow field can be achieved by employing direct numerical simulation (DNS) of Navier-Stokes equations (conservation of mass and momentum) which can resolve all scales of motion.

$$\partial_t \boldsymbol{u} + \boldsymbol{\omega} \times \boldsymbol{u} + \boldsymbol{\nabla} \left( \frac{p}{\rho_f} + \frac{\|\boldsymbol{u}\|^2}{2} \right) = \nu \nabla^2 \boldsymbol{u} + \boldsymbol{f}, \, \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

• DNS is computationally prohibitive and cannot be applied to simulate flow at large Reynolds numbers.

 High-resolution numerical simulation of the turbulent flow field can be achieved by employing direct numerical simulation (DNS) of Navier-Stokes equations (conservation of mass and momentum) which can resolve all scales of motion.

$$\partial_t \boldsymbol{u} + \boldsymbol{\omega} \times \boldsymbol{u} + \boldsymbol{\nabla} \left( \frac{p}{\rho_f} + \frac{\|\boldsymbol{u}\|^2}{2} \right) = \nu \nabla^2 \boldsymbol{u} + \boldsymbol{f}, \, \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

- DNS is computationally prohibitive and cannot be applied to simulate flow at large Reynolds numbers.
- Reynolds Number (*Re*), quantifies the intensity of a turbulent flow.

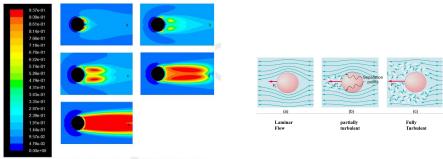


FIG. 11: Contour of void fraction around the cylinder for (a)  $Re_{TP} = 1.000$ , (b)  $Re_{TP} = 3,162$ , (c)  $Re_{TP} = 10,000$ , (d)  $Re_{TP} = 31,623$ , and (e)  $Re_{TP} = 100,000$ .

• In the real problems of interest, oceanic and atmospheric turbulent flow, the Re is very large ( $O(10^7)$  or more).

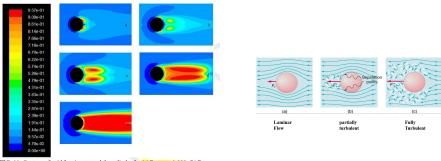
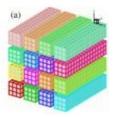


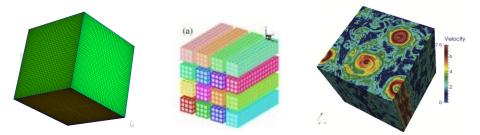
FIG. 11: Contour of void fraction around the cylinder for (a)  $Re_{TP} = 1,000$ , (b)  $Re_{TP} = 3,162$ , (c)  $Re_{TP} = 10,000$ , (d)  $Re_{TP} = 31,623$ , and (e)  $Re_{TP} = 100,000$ .

- In the real problems of interest, oceanic and atmospheric turbulent flow, the Re is very large ( $O(10^7)$  or more).
- We use DNS over a range of Re to look for trends in the behavior. This can provide insight regarding the extent to which results obtained at low/moderate Re might be extrapolated the real problems of interest.

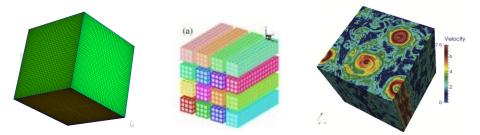








- Turbulent flow property: Our DNS study considers  $R_{\lambda} = 90,224,398$ , where  $R_{\lambda} \equiv u'\lambda/\nu$ .
- The higher the Re, the higher the resolution (more grid points;  $N_g = 128^3$ ,  $N_p = 16$  for  $R_{\lambda} = 90$  versus  $N_g = 1024^3$ ,  $N_p = 1024$  for  $R_{\lambda} = 398$ ; 2 days versus 4 months!)



- Turbulent flow property: Our DNS study considers  $R_{\lambda} = 90,224,398$ , where  $R_{\lambda} \equiv u'\lambda/\nu$ .
- The higher the Re, the higher the resolution (more grid points;  $N_g = 128^3$ ,  $N_p = 16$  for  $R_{\lambda} = 90$  versus  $N_g = 1024^3$ ,  $N_p = 1024$  for  $R_{\lambda} = 398$ ; 2 days versus 4 months!)
- o particle tracking:

$$\frac{d^2}{dt^2} \boldsymbol{x}^p(t) = \frac{d}{dt} \boldsymbol{v}^p(t) = \frac{\boldsymbol{u}(\boldsymbol{x}^p(t), t) - \boldsymbol{v}^p(t)}{\tau_p} + \boldsymbol{\mathfrak{g}}$$



**cumulus clouds:** low-level clouds, generally less than 2 km (6,600 ft)



**Cumulonimbus clouds:** high-level clouds, up to 12 km (40,000 ft)

- Particle property (effect of gravity or fall speed):  $Fr = \infty$ , 0.3 (cumulonimbus clouds), 0.052 (cumulus clouds). The stronger gravity, the larger the domain.
- Particle property (effect of inertia; e.g. size. density): fifteen St in the range of  $0 \le St \le 3$ . The larger the St, the more difficult to solve the particle's equation.

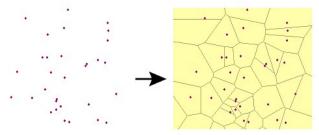
• To obtain local insight into the clustering of particles, the Voronoi diagrams (tessellation) approach is employed as the data processing technique to identify regions of high (clusters) and low (voids) concentration of particles.

- To obtain local insight into the clustering of particles, the Voronoi diagrams (tessellation) approach is employed as the data processing technique to identify regions of high (clusters) and low (voids) concentration of particles.
- Using this technique enables characterization of the general features of individual clusters such as topology and spatial orientation (and the kinematics and dynamics of clusters) in the underlying turbulence.

- To obtain local insight into the clustering of particles, the Voronoi diagrams (tessellation) approach is employed as the data processing technique to identify regions of high (clusters) and low (voids) concentration of particles.
- Using this technique enables characterization of the general features of individual clusters such as topology and spatial orientation (and the kinematics and dynamics of clusters) in the underlying turbulence.
- The appearance of clustering and its strength is diagnosed by exploring the distribution of Voronoï volumes over a significant range of the three parameter space  $Fr = \infty, 0.3, 0.052, R_{\lambda} = 90, 224, 398$  and  $0 \le St \le 3$  which are varied independently.

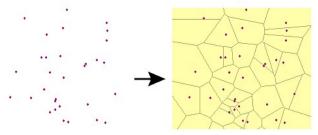
- To obtain local insight into the clustering of particles, the Voronoi diagrams (tessellation) approach is employed as the data processing technique to identify regions of high (clusters) and low (voids) concentration of particles.
- Using this technique enables characterization of the general features of individual clusters such as topology and spatial orientation (and the kinematics and dynamics of clusters) in the underlying turbulence.
- The appearance of clustering and its strength is diagnosed by exploring the distribution of Voronoï volumes over a significant range of the three parameter space  $Fr = \infty, 0.3, 0.052, R_{\lambda} = 90, 224, 398$  and  $0 \le St \le 3$  which are varied independently.

• Voronoi diagrams are used to understand patterns over an area of interest.



(https://www.safe.com/transformers/voronoi-diagrammer/)

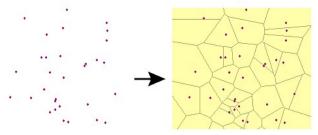
• Voronoi diagrams are used to understand patterns over an area of interest.



(https://www.safe.com/transformers/voronoi-diagrammer/)

- After collecting data points and recording their location, the plane can be divided into sections, or Voronoi cells, that are representative of each data point.
- A Voronoi diagram describes the spatial relationship between points that are near each other, or their nearest neighbours. It is a set of connection polygons derived from points or locations.

• Voronoi diagrams are used to understand patterns over an area of interest.

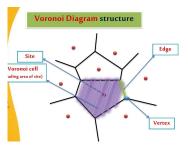


(https://www.safe.com/transformers/voronoi-diagrammer/)

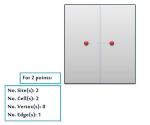
- After collecting data points and recording their location, the plane can be divided into sections, or Voronoi cells, that are representative of each data point.
- A Voronoi diagram describes the spatial relationship between points that are near each other, or their nearest neighbours. It is a set of connection polygons derived from points or locations.

• Voronoï diagram definition: The partitioning of a plane with *n* points into convex polygons such that each polygon contains exactly one generating point and every point in a given polygon is closer to its generating point than to any other.

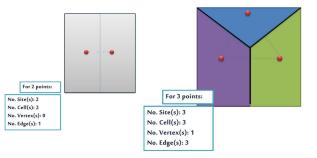
• Voronoï diagram definition: The partitioning of a plane with *n* points into convex polygons such that each polygon contains exactly one generating point and every point in a given polygon is closer to its generating point than to any other.



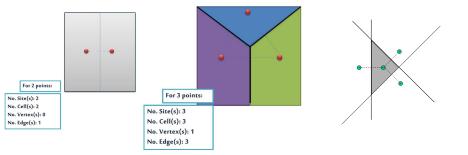
- First step is to draw a line connecting adjacent points.
- Second step is to draw a perpendicular line to the one you just drew in the midpoint of it.
- Last step is to connect lines, drawn in the second step, in to an network.



- First step is to draw a line connecting adjacent points.
- Second step is to draw a perpendicular line to the one you just drew in the midpoint of it.
- Last step is to connect lines, drawn in the second step, in to an network.



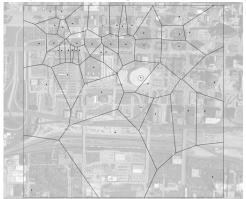
- First step is to draw a line connecting adjacent points.
- Second step is to draw a perpendicular line to the one you just drew in the midpoint of it.
- Last step is to connect lines, drawn in the second step, in to an network.



• Application: They find widespread applications in areas such as nearest neighbor queries for data structure problems in computational geometry, business applications such as determining where to locate a store so it is no closer to any existing store of its kind, epidemiology, geophysics, and meteorology.

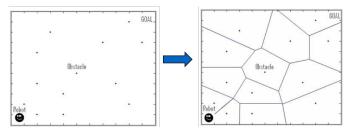
• Application: They find widespread applications in areas such as nearest neighbor queries for data structure problems in computational geometry, business applications such as determining where to locate a store so it is no closer to any existing store of its kind, epidemiology, geophysics, and meteorology.

**City Planning**: one can easily determine where is the nearest shop or hospital, and urban planners can study if certain area need a new hospital

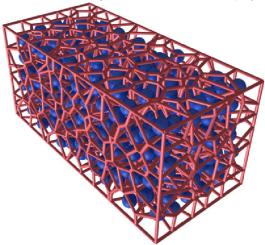


• Application: They find widespread applications in areas such as nearest neighbor queries for data structure problems in computational geometry, business applications such as determining where to locate a store so it is no closer to any existing store of its kind, epidemiology, geophysics, and meteorology.

**Robot Path Planning**: restricting a robot to traverse the edges created by the voronoi diagram

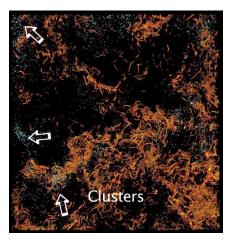


- A voronoi diagram records information about the distances between sets of points in any dimensional space.
- Higher Dimensions Voronoi Diagrams: Cells are convex polytopes



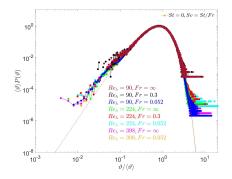
• Now that we are familiar with Voronoï Tessellation concept and particle clustering in turbulence, we are ready to analyze clusters.



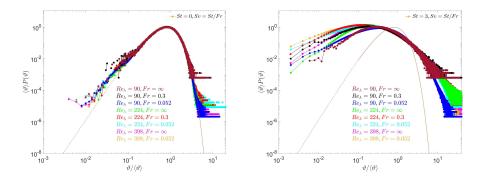


- Recall that for each simulation we write out the particles' position and this data serve as input for the 3D Voronoï analysis.
- We employ the 3D Voronoï diagrams on DNS simulations over a wide range of three control parameters  $R_{\lambda}$ , Fr and St. For each tuple (Re, Fr, St), we analyze multiple (~25) snapshots (time step) of simulations to collect sufficient data for statistical convergence.
- Given the particles' position and box volume, an id (index) is assigned to each particle. The output of 3D Voronoï analysis is the cell volume as well as neighbors list for each particle.
- Considering the Voronoï volume as a random variable, we compute its probability distribution function (PDF) for different combinations of (*Re*, *Fr*, *St*).

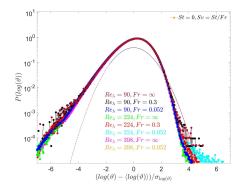
 Indeed, the local particle concentration field is represented by the inverse of the volume of Voronoï cells. The small/large Voronoï volumes represent the high/low concentration regions of the flow.



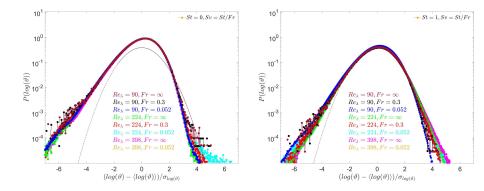
 Indeed, the local particle concentration field is represented by the inverse of the volume of Voronoï cells. The small/large Voronoï volumes represent the high/low concentration regions of the flow.



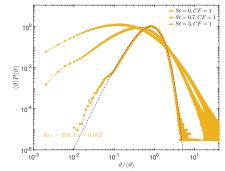
In certain regimes, the PDF of Voronoï volumes follows a log-normal distribution!



In certain regimes, the PDF of Voronoï volumes follows a log-normal distribution!

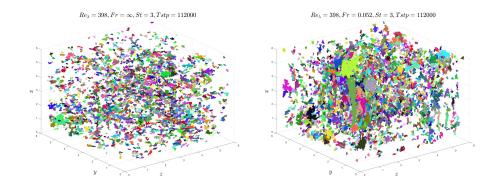


• The PDF of inertial particles intersect the RPP at two points, one is smaller and the other one is greater than the mode of RPP, which are denoted by  $\vartheta_C$  and  $\vartheta_v$ , respectively. These points serve as thresholds to detect clusters and voids.

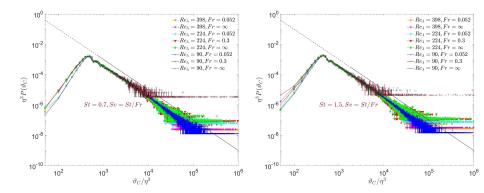




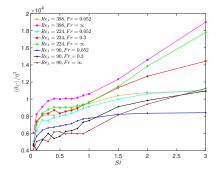
- We then consider the properties of particles in clusters, which are regions of connected Voronoï cells whose volume is less than a certain threshold.
- To detect the clusters in each simulation of (Re, Fr, St), we first choose the particle ids that their cell volumes is smaller than the threshold ( $\vartheta_C$ ). These particles may form a cluster, if they are connected together.
- Then we run a searching algorithm to go through the neighbors of those particles and pick neighbors that also their cell volumes are smaller than the threshold.
- We continue this procedure recursively for the neighbors of neighbors of neighbors of ... until we find all the connected cells of clusters. For some cases it may take 12 days to find all the clusters of one snapshot of flow!



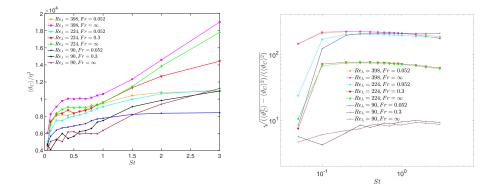
 Volume of each cluster is sum of the Voronoï volumes of its constituent particles. Let's see how the PDF of cluster volumes behaves:



• How about the average (first moment) size and the standard deviation (second moment) of clusters?



• How about the average (first moment) size and the standard deviation (second moment) of clusters?



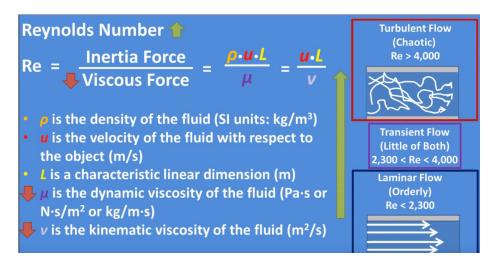
• Given the first four moments of cluster volumes for 112 tuple of (Re, Fr, st), build a supervised machine learning model that take a tuple and outputs the first four moments.

1	Α	В	С	D	Е	F	G	Н	1	J		
1	St	Re	Fr	Sv	eta	u_eta	R_moment_1	R_moment_2	R_moment_3	R_moment_4		
2	0	398	0	1	0	0.09	0.00022182	0.0010347	0.011354	0.12509		
3	0	398	0	2	0	0.09	0.00027479	0.0032549	0.058006	1.0344		
4	0	398	0	4	0	0.09	0.0002954	0.0041561	0.079922	1.5377		
5	0	398	0	6	0	0.09	0.00030066	0.0043488	0.083446	1.6023		
6	0	398	0	8	0	0.09	0.00029691	0.0041375	0.076124	1.4014		
7	1	398	0	10	0	0.09	0.00030716	0.0043494	0.080143	1.477		
8	1	398	0	12	0	0.09	0.00031431	0.0044672	0.08206	1.5077		

# **Thank You**

#### Backup slides: DNS table

Parameter	DNS 1	DNS 2	DNS 3	DNS 4							
N	128	128	1024	512							
$R_{\lambda}$	93	94	90	224							
Fr	~	0.3	0.052	$\infty$	Reynolds-	number e	ffects on	inertial part	icle dynamic	s. Part 1	621
L	$2\pi$	$2\pi$	$16\pi$	$2\pi$							
ν	0.005	0.005	0.005	0.0008289	Simulation	I	П	III	IV	V	
e	0.324	0.332	0.257	0.253		1			-		
l	1.48	1.49	1.47	1.40	$R_{\lambda}$	88	140	224	398	597	
$l/\eta$	59.6	60.4	55.6	204	v	0.005	0.002	0.0008289	0.0003	0.00013	
u'	0.984	0.996	0.912	0.915	e	0.270	0.267	0.253	0.223	0.228	
$u'/u_{\eta}$	4.91	4.92	4.82	7.60	l	1.46	1.41	1.40	1.45	1.43	
$T_L$	1.51	1.50	1.61	1.53	$\ell/\eta$	55.8	107	204	436	812	
$T_L/\tau_\eta$	12.14	12.24	11.52	26.8	<i>u</i>	0.914	0.914	0.915	0.915	0.915	
$\kappa_{\max}\eta$	1.5	1.48	1.61	1.66	$u'/u_{\eta}$	4.77	6.01	7.60	10.1	12.4	
Np	262,144	262,144	16,777,216	2,097,152	$T_L$	1.60	1.54	1.53	1.58	1.57	
Parameter	DNS 5	DNS 6	DNS 7	DNS 8	$T_L/\tau_\eta$	11.7	17.7	26.8	43.0	65.4	
N	512	1024	1024	1024	$T/T_L$	15.0	10.4	11.4	11.1	5.75	
$R_{\lambda}$	237	230	398	398	$k_{max}\eta$	1.59	1.59	1.66	1.60	1.70	
Fr	0.3	0.052	~	0.052	N	128	256	512	1024	2048	
L	27	47	27	27	$N_p$	262 144	262 144	2097152	16777216	134 217 728	
v	0.0008289	0.0008289	0.0003	0.0003	Niracked	32768	32768	262 144	2 097 152	16777216	
	0.2842	0.239	0.223	0.223	Nproc	16	16	64	1024	16384	
i i	1.43	1.49	1.45	1.45	TABLE 1. Flow parar	neters for	the DNS	study. All di	mensional pa	rameters are in	arbitrary
$l/\eta$	214	213	436	436	units and all statistic						
u'	0.966	0.914	0.915	0.915	§§ 2.1 and 2.2.		1		1		
$u'/u_n$	7.82	7.7	10.1	10.1	33						
$T_L$	1.48	1.63	1.58	1.58							
$T_L / \tau_\eta$	27.36	27.66	43.0	43.0							
	1.62	1.68	43.0	43.0							
$\kappa_{\max}\eta$ $N_p$	2,097,152	16,777,216		2.097.152							
Np	2,097,152	16,777,216	2,097,152	2,097,152							



- Given the Voronoï volumes, the clustering of inertial particles is explored by comparing the PDF of inertial particles (St > 0) with the fluid particles (St = 0).
- The fluid particles always follow the streamlines, have a uniform random distribution and their Voronoï tessellation PDF is unique (Random Poisson process or RPP).
- We then consider the properties of particles in clusters, which are regions of connected Voronoï cells whose volume is less than a certain threshold.