Supplementary File

Oscillating Dispersed-Phase Co-Flow Microfluidic Droplet Generation: Jet Length Reduction Effect

Amin Shams Khorrami and Pouya Rezai*
Department of Mechanical Engineering, York University, Toronto, ON, Canada

* Corresponding Author: Pouya Rezai, BRG 433B, 4700 Keele St, Toronto, ON, M3J 1P3, Canada; Tel: 416-736-2100 ext. 44703; Email: prezai@yorku.ca

1. Data Analysis and Statistics

1.1. Droplet Size Measurement

In each set of experiments, after length scale calibration, the automatic particle counting and measurement module was used to determine the generated droplet sizes. After subtracting the background values and converting the images into grayscale (8-bit), an empirical threshold value was adopted to perform droplet edge detection and automatic size measurement. The errors between manual and automatic measurements calculated for *We*=13.24 and *Ca*=0.183 at various oscillation frequencies (in manuscript Fig. 2) are summarized in Table S1.

Table S1. Percentage of error difference between the manual and automatic edge detection methods

Oscillation Frequency	**Error %**			
	Cluster #1	Cluster #2	Cluster #3	Cluster #4
0	4.34	5.58	-	-
1	4.75	4.27	4.16	-
5	4.58	5.25	5.68	-
10	4.59	4.81	5.92	-
15	4.54	5.66	6.07	6.32

The statistical methods used in this work imply normal distributions of the analyzed data. For instance, Figure S1 represents the normal distribution of data for various clusters of droplet sizes for *We=13.24* and *Ca=0.183* at 15Hz.

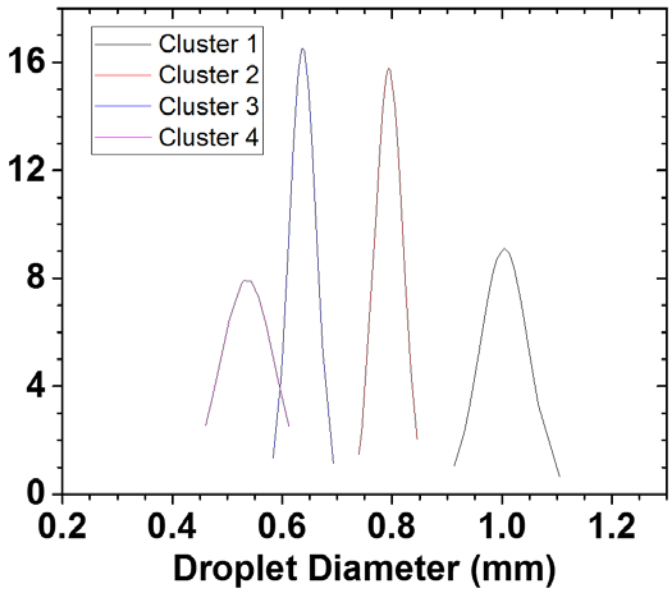

Figure S1. Normal distribution of clustered droplet diameters for We=13.24 and Ca=0.183 at 15Hz

1.2. Data Clustering for Droplet Grouping

Cluster analysis is a tool to detect groups (clusters) that are sharing similar characteristics in a set of data. Similarity can be identified by either internal cohesion or external isolation. Clustering algorithms can be divided into two main categories, i.e. hierarchical and partitional. Hierarchical algorithms benefit predefined nested clusters for categorizing objects while partitional methods simultaneously recognize all clusters as partitions of data points. Due to inherent rigidity of clusters formed via hierarchical methods, the method suffers from a defect that the defined clusters cannot be reformed in further steps. On the other hand, partitional methods offer a dynamic clustering scheme which is not rigid, and reformation is possible at each step.

The most popular and the simplest partitional algorithm is k-means clustering method [1] which automatically partitions a data set into *k* number of clusters. Imagine you have a set of data comprised of *n* numbers $X_1, X_2, \dots, X_n$ which can represent the size of droplets measured in each experiment in our paper. The k-means clustering algorithm [2–4] proceeds as follow:

a) Input: k (number of clusters) and $X_1, X_2, \dots, X_n$ (data set);
b) Place centroids $C_1, C_2, \dots, C_k$ at random data points;

c) For each data point X_i, find the nearest centroid C_j and assign data point X_i to cluster C_j
d) For each cluster j=1, 2, … , k, calculate the average of updated cluster members and assign it to new centroid C_j;
e) Compare calculated C_j with previous value; if they are not equal with a predefined error margin, repeat steps c and d until convergence is achieved;
f) Return clusters;
g) End.

As stated above, alongside the data points, the number of clusters should also be assigned to the algorithm to complete the clustering process. In some cases, the number of clusters is unknown and further investigations are needed to define the optimum number of clusters (ONC). In cases as such, k-means algorithm should be applied for a range of k-values starting from the minimum, k_{min}=2, up to the maximum possible value, k_{max}= population of data points. Simultaneously, a clustering index must also be defined as a function of number of clusters so ONC can be detected by interpretation of behavior of this index.

Clusters should be internally cohesive which requires low-scattered data points in each group. It is obvious that increasing the number of clusters to the maximum possible k-value will continuously reduce the internal dispersion (variance) of data points, but k=n number of clusters for a data set with n members is not desirable. Therefore, another parameter should be considered alongside the variance to ensure the external isolation as well as internal cohesion. In other words, the number of clusters should also be minimized while external isolation ensures the fact that existing clusters are far enough not to merge into a bigger cluster.

By increasing the number of clusters each partition becomes more focused and the variance calculated within the cluster decreases. Unlikely, the sum of these variances does not continuously decrease and changes diminish beyond a specific point. That is where, clusters start to subdivide into daughter clusters while the variance of mother cluster is almost equal to the sum of the variances of daughter clusters. Thus, assuming “sum of variances” as a measure for clustering quality can ensure cluster subdivision effectiveness. That specific point which can also be observed as a knee point in a graph of “sum of variances” vs. “number of clusters” indicates the optimum number of clusters in a data set [5]. In most of the cases with high populations, there is no need to explore the whole possible range for *k* and a rough estimation about maximum number of clusters

can be done based on the number of peaks observed in the graph of data distribution density. Setting the maximum *k* value reasonably above the estimated cluster size can conservatively save a huge portion of computation time. In Figure S2, we apply the described method above on a population of droplet sizes obtained at *We*=13.24 and *Ca*=0.183 at 15 Hz. The maximum number of clusters were manually selected as 10 clusters to reduce the processing time.

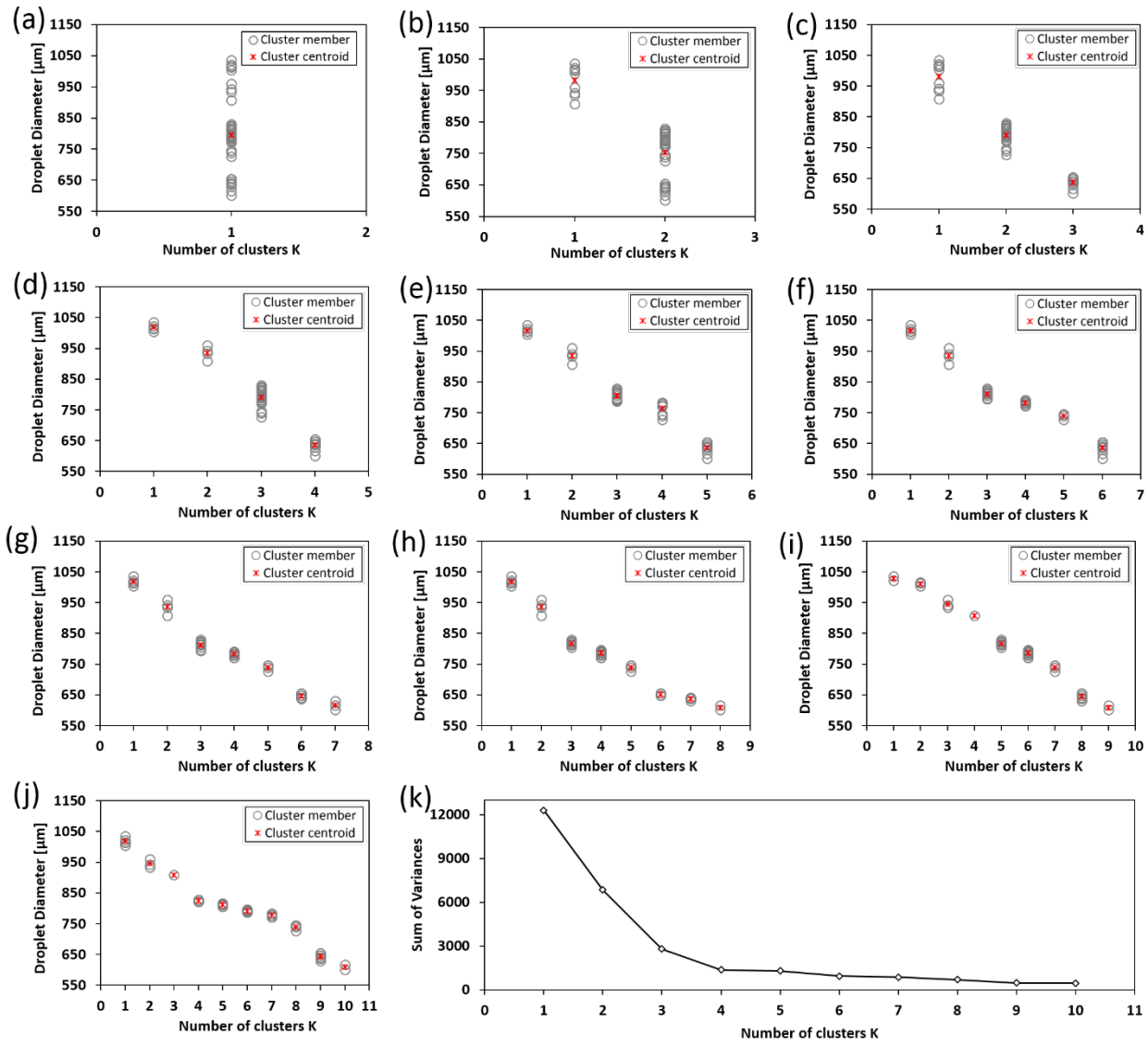

Figure S2. Defining the optimum number of clusters for a set of droplets generated at We=13.24 and Ca=0.183 at 15 Hz using the clustering index curve. Droplet diameters for different cluster numbers, k=2-10, are shown in (a-j). The clustering index curve vs. the number of clusters are plotted in (k) and due to the minimal variations after k=4, this knee point is selected as the number of independent clusters (i.e. 4 droplet size groups). Clusters ineffectively subdivide beyond k=4 and sum of variances of the daughter clusters is almost equal to the mother cluster.

As shown in Figure S2 (a-d and k), by increasing the number of clusters, the clustering index (sum of the variances) decreases effectively till k=4 which is identified as the optimum number of clusters in this data set. Beyond this point in Figure S2 (e-j), increasing the number of clusters resulted in ineffective subdivision of clusters, and sum of variances of daughter clusters being almost equal to the mother cluster. Thus, internal division of clusters beyond this point has no significant effect on the clustering index (Figure S2 k).

2. Effect of d- and c-phase Flow Rates on Stationary Jet Length

The data in the paper were reproduced to show the effect of c-phase (*Ca* number) and d-phase (*We* number) flow rates on the stationary jet length as shown in Figure S3 and Figure S4, respectively.

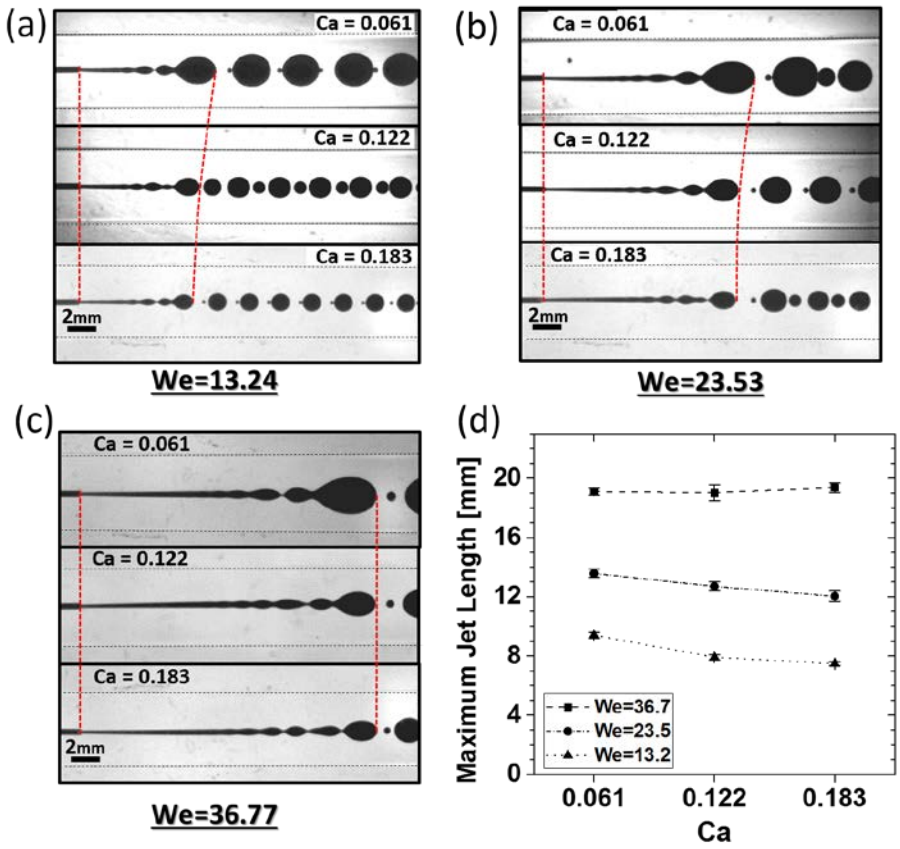

Figure S3. Effect of Q_c (Ca number) on the stationary jet length. Jet images at (a) We = 13.24, (b) We =23.53, and (c) We =36.77 are shown while (d) demonstrates the aggregated effect of Ca number on the jet length at various We numbers.

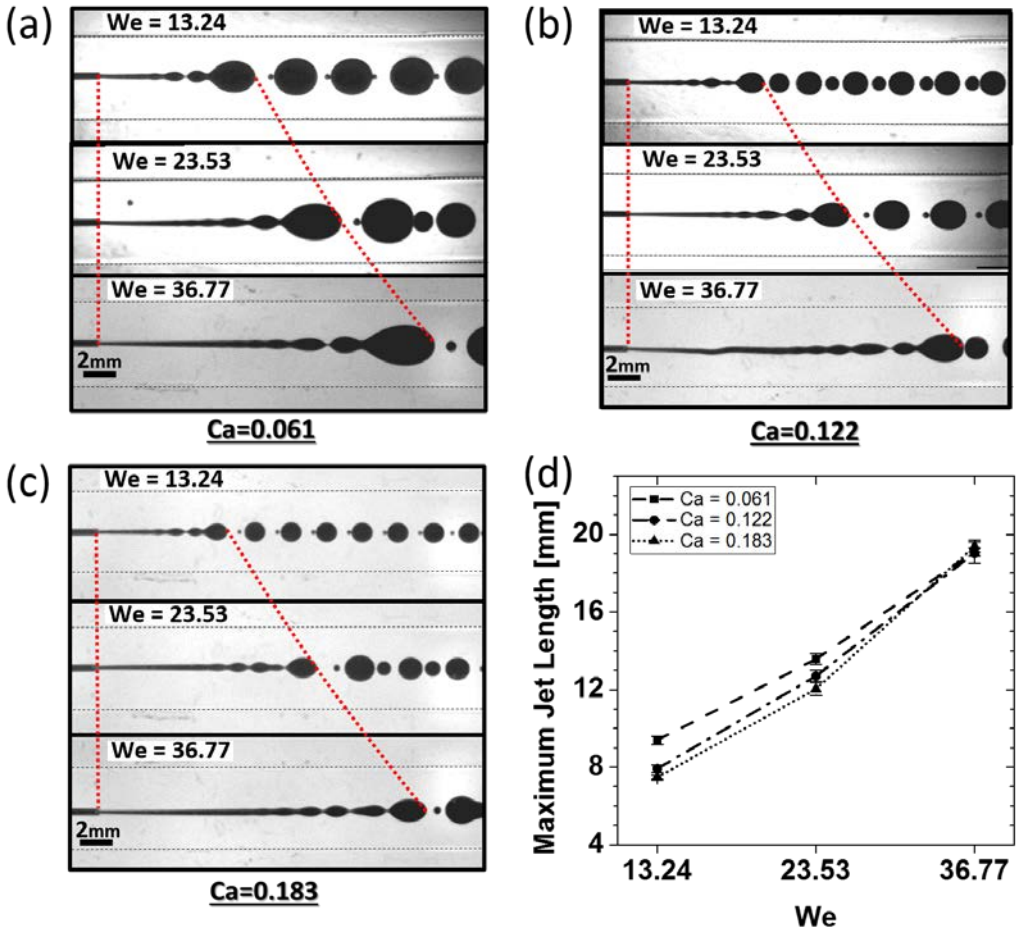

Figure S4. Effect of Q_d (We number) on the stationary jet length. Jet images at (a) Ca = 0.061, (b) Ca =0.122, (c) Ca=0.183 are shown while (d) demonstrates the aggregated effect of We number on the jet length at various Ca numbers.

3. A Semi-Analytical Force Analysis on the Oscillating Jet

In the present study, an oscillatory motion was introduced to a submerged jet (water in oil) to investigate the effect of the oscillation frequency on the jet length. In the stationary mode and because of the increased inertia at the higher flow rates, the d-phase liquid elongated into the c-phase fluid as a jet and formed an interface (Figure S5). As the jet interface grew, the axial drag force imposed on the jet body increased. When the interface reached a critical size, the ellipsoid tail of the jet was sheared off by breaking up the unstable interface and discrete droplets were formed. Upon introducing the oscillatory motion, an additional transverse drag force was imposed onto the d-phase interface in the perpendicular direction to the flow, which yielded into a reduction in the total jet length. Simple droplet generation geometries has been numerically investigated[6], but oscillation-induced droplet generation is very complex and difficult to be analytically studied,

yet we attempted to develop a simplified model solely to examine if our hypothesis of growing axial and lateral drag forces stand valid to describe the results obtained in jet length reduction upon oscillation of the d-phase nuzzle.

As shown in Figure S5, the jet shape was approximated by a cylinder (length: L_1, diameter: nozzle diameter D_N) attached to an ellipsoid (longitudinal diameter: L_2, and transverse diameter: D). The axial and transverse drag forces imposed on each component, respectively invoked by c-phase velocity and nozzle oscillatory motion, can be categorized as:

- F_1: Axial drag on cylinder
- F_2: Axial drag on ellipsoid
- F_3: Transverse drag on cylinder
- F_4: Transverse drag on ellipsoid

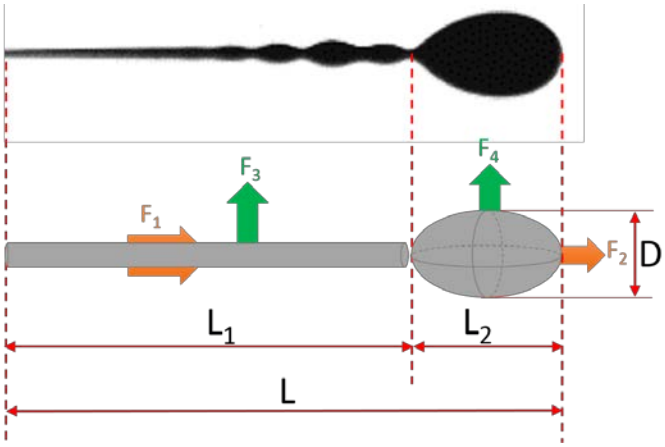

Figure S5. Simplified jet shape and attributed drag forces on each part

The axial shear force on a cylinder (F_1) can be calculated by Equation 1, while the rest of the drag forces above can be calculated by the general form of Equation 2.

$$F_{a,cyl} = \mu \frac{du}{dy} A \tag{1}$$

$$F_D = \frac{1}{2} C_D \rho A V^2 \tag{2}$$

where, μ is viscosity, du/dy is velocity gradient in Y-direction (which was assumed to be linear in this simplified model), A is the effective surface area, C_D is the drag coefficient and V is the

velocity. Thus, axial and transverse drag forces for the simplified jet shape in Figure S5 can be estimated as:

$$F_1 = \mu \frac{V_d - V_c}{w} \pi D_N L_1 \quad (3)$$

$$F_2 = \frac{1}{2} C_D \rho_c \frac{\pi D^2}{4} (V_a)^2 \quad (4)$$

$$F_3 = \frac{1}{2} C_D \rho_c D_N L_1 (V_t)^2 \quad (5)$$

$$F_4 = \frac{1}{2} C_D \rho_c \frac{\pi D L_2}{4} (V_t)^2 \quad (6)$$

where, w is the distance between the needle and the channel side wall and subscripts *d*, *c*, *a* and *t* refer to the dispersed phase, continuous phase, axial direction and transverse direction, respectively.

The axial and transverse drag forces were estimated in terms of their order of magnitude based on the jet lengths obtained in all experimental conditions and shown in Figure S6 at various oscillation frequencies. On the one hand, upon increasing the oscillation frequency and due to a reduction in the jet length in constant flow conditions, the axial drag force magnitude decreased by 50%-100%. On the other hand, upon increasing the oscillation frequency and despite the total length reduction of the jet, due to the increased transverse velocity of the jet, the magnitude of the transverse drag force increased by approximately two orders of magnitude for all experimental conditions. This approximation supports our claim that upon increasing the oscillation frequency, the role of the transverse drag force in jet breakup becomes more dominant.

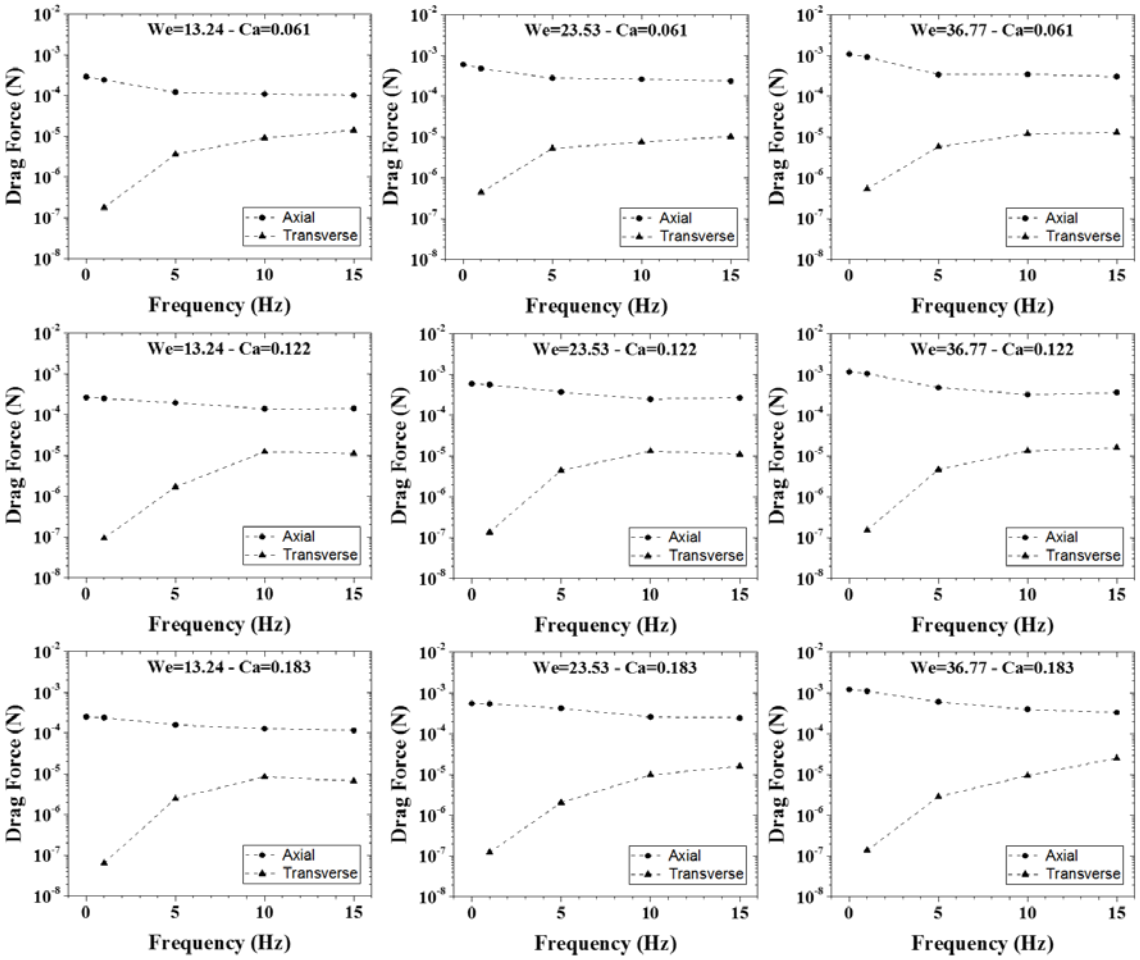

Figure S6. Axial and transverse drag forces vs. oscillation frequencies at different experimental conditions (We-Ca)

Since the flow rates and physical properties of the d- and c-phase fluids remained constant in each experimental condition (We-Ca), one can safely assume that the total magnitude of the drag force should remain constant before and after introducing the oscillatory motion. Accordingly, we assumed that the total drag force in the stationary mode $\left(\vec{F}_{static} = \vec{F}_1 + \vec{F}_2\right)$ should be similar in magnitude (Equations 7-8) to the sum of all four drag forces in the oscillation mode $\left(\vec{F}_{dynamic} = \vec{F}_1 + \vec{F}_2 + \vec{F}_3 + \vec{F}_4\right)$, with the equality unknown being the jet length at the cylindrical part, L_1. One requirement for this approximation to stay valid is that the jet remains as a straight strand and not bend significantly due to the oscillatory motion, an effect observed mostly at lower frequencies of 1 Hz and 5 Hz.

$$|F_{static}| = F_1 + F_2 \tag{7}$$

$$\left|F_{dynamic}\right| = \sqrt{(F_1 + F_2)^2 + (F_3 + F_4)^2} \tag{8}$$

By substituting the drag forces from Equations 3-6 in Equations 7 and 8, the cylinder length L_1 for all oscillation frequencies can be calculated based on Equation 9.

$$(F_{static})^2 = \left(\mu \frac{V_d - V_c}{w} \pi D_N \underline{\boldsymbol{L_1}} + \frac{1}{2} C_D \rho_c \frac{\pi D^2}{4} (V_a)^2\right)^2 + \left(\frac{1}{2} C_D \rho_c D_N \underline{\boldsymbol{L_1}} (V_t)^2 + \frac{1}{2} C_D \rho_c \frac{\pi D L_2}{4} (V_t)^2\right)^2 \quad (9)$$

Dimensions of ellipsoids, D and L_2, were extracted from experiments using image analysis. The total jet length would be estimated by summing the calculated L_1 and the longitudinal diameter of the ellipsoid L2. Figure S7-S10 illustrate the calculated and actual jet lengths at various *We* and *Ca* numbers at different oscillation frequencies.

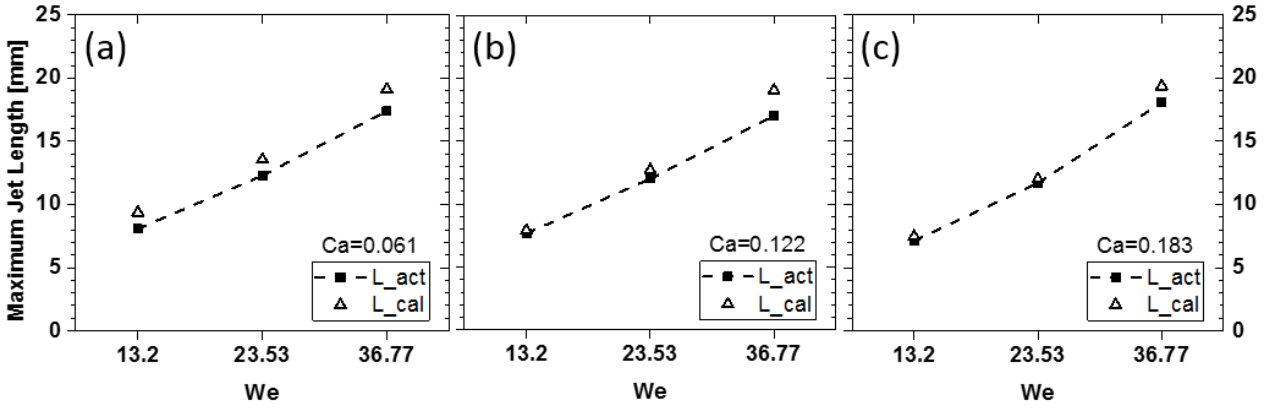

Figure S7. Calculated and actual jet lengths vs We number for f=1Hz at various Ca numbers: (a) Ca=0.061, (b) Ca=0.122 and (c) Ca=0.183.

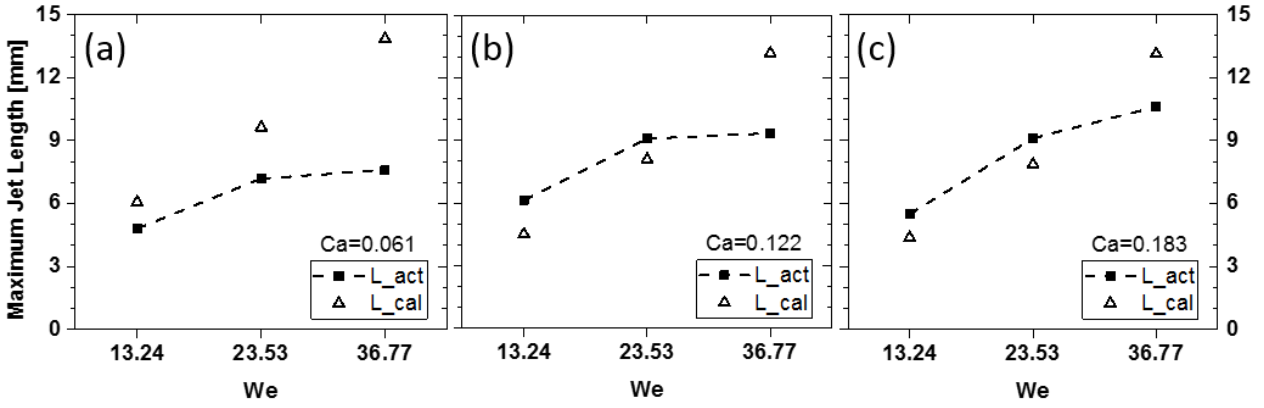

Figure S8. Calculated and actual jet lengths vs We number for f=5Hz at various Ca numbers: (a) Ca=0.061, (b) Ca=0.122 and (c) Ca=0.183.

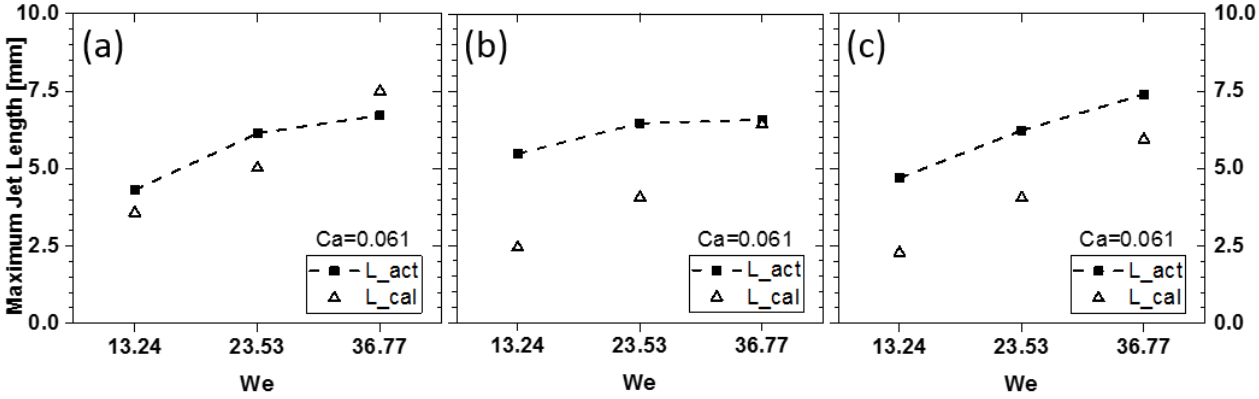

Figure S9. Calculated and actual jet lengths vs We number for f=10Hz at various Ca numbers: (a) Ca=0.061, (b) Ca=0.122 and (c) Ca=0.183.

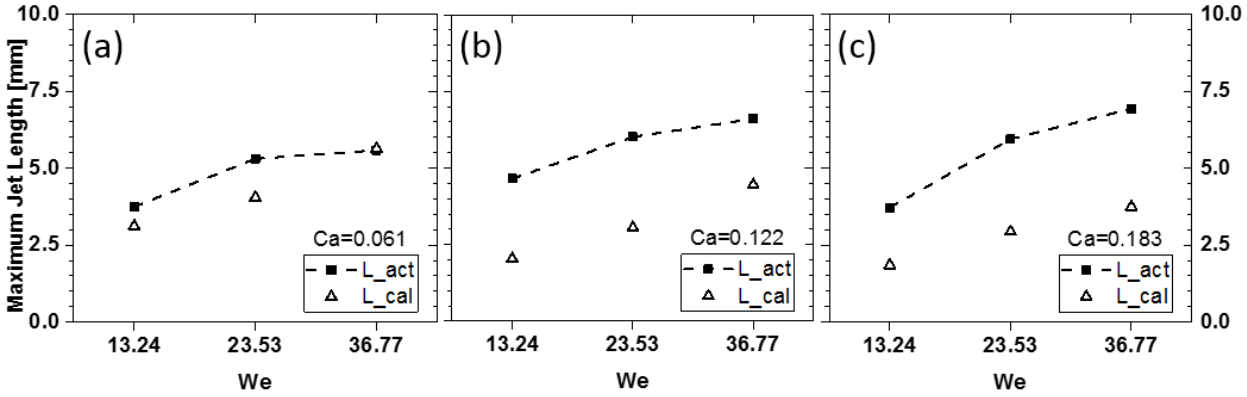

Figure S10. Calculated and actual jet lengths vs We number for f=15Hz at various Ca numbers: (a) Ca=0.061, (b) Ca=0.122 and (c) Ca=0.183.

The model predicted the jet length with acceptable accuracy at the lower frequencies of 1 Hz and 5 Hz because the jet maintained its straight shape during oscillation similar to our simplified model; However, at the higher frequencies of 10 Hz and 15 Hz, jet bending prevented us from achieving a reasonable length prediction, although the trend in jet length increase with *We* number was captured by the model. We will continue improving this semi-analytical model for better prediction of the jet length in the future.

4. Effect of d-phase Flow Rate on the Oscillating Jet Length

The data in the paper was reproduced to show the effect of d-phase flow rate on the oscillating jet length as shown in Figure S11. At constant frequency and *Ca* levels, the jet length was observed to proportionally increase by increasing the d-phase flow rate. However, the effect of d-phase flow

rate was more significant at slower oscillation rates. This observation is in accordance with the fact that jets were longer at increased d-phase flow rates and thus were inherently more unstable at initial oscillation frequency levels.

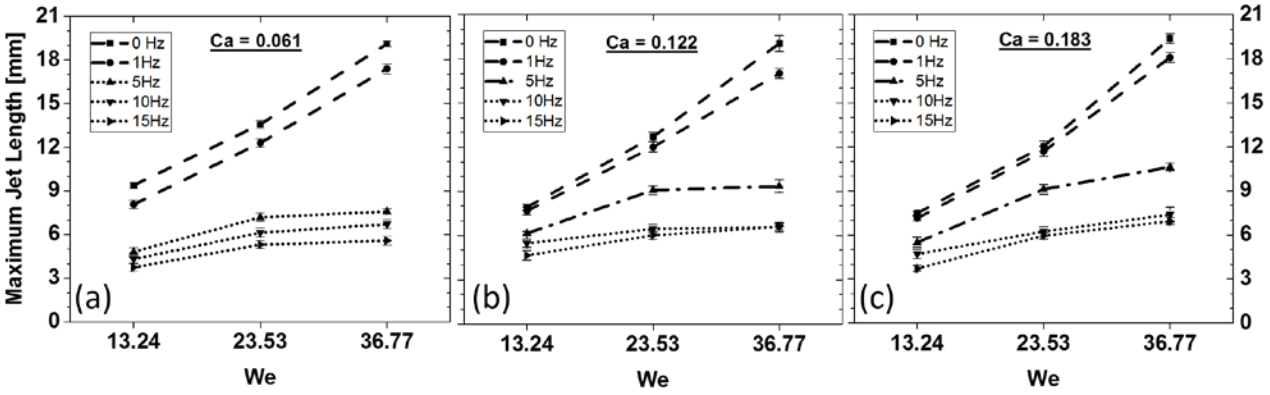

Figure S11. Effect of d-phase flow rate on jet length at different oscillation frequencies for (a) Ca = 0.061, (b) Ca=0.122 and (c) Ca =0.183

References:

1 J. Macqueen, *Proc. Fifth Berkeley Symp. Math. Stat. Probab.*, 1967, **1**, 281–297.

2 P.-N. Tan, M. Steinbach and V. Kumar, *Introd. to Data Min.*, 2005, Chapter 8.

3 J. A. Hartigan and M. A. Wong, *J. R. Stat. Soc. Ser. C (Applied Stat.*, 1979, **28**, 100–108.

4 X. Jin and J. Han, in *Encyclopedia of Machine Learning*, 2011, pp. 563–564.

5 R. L. Thorndike, *Psychometrika*, 1953, **18**, 267–276.

6 V. Cristini and Y. C. Tan, *Lab Chip*, 2004, **4**, 257–264.