

## 1 Overview

Mainly we are concerned about what parameters are important in electrostatic transport of relatively large (e.g. millimetric) droplets in low-gravity, and what the values of the respective dimensionless groups, namely  $\mathbb{I}_m$ ,  $\mathbb{E}_u$  and  $\mathbb{E}_{u+}$  might be.

To find typical values of these parameters we used the following approach:

1. We observed spontaneous droplet jumps on charged dielectric super-hydrophobic surfaces under low-gravity conditions in a 2.14 s drop tower, while varying the independent variables  $V_d$ ,  $\sigma$ . However, this raises an interesting challenge; it is impractical to directly measure all of the key physical quantities that appear in the dimensionless groups at once in a drop tower experiment. In particular, determination of net droplet free electric charge  $q$ , is difficult.
2. Using high-speed video and image analysis software we captured the trajectories of the droplets.
3. We solved the inverse problem to find the key parameters by maximizing a statistical likelihood function between an observed trajectory and the trajectory predicted by a dynamical model given that certain set parameters. The best fit parameters obtained by direct-search optimization are those corresponding to the maximum likelihood experimental values. The optimization is constrained by the measurement precision of directly measured independent, and dependent variables.

## 2 Experimental Methods

- A very low-tech superhydrophobic electret was produced, with surface potentials 0.7–4.0 kV and contact angles  $\sim 150^\circ$  with approximately  $20^\circ$  contact angle hysteresis when uncharged. The dielectrics are a lamina of 3–4 corona charged 0.4 mm PMMA sheets. The electric field scales with the

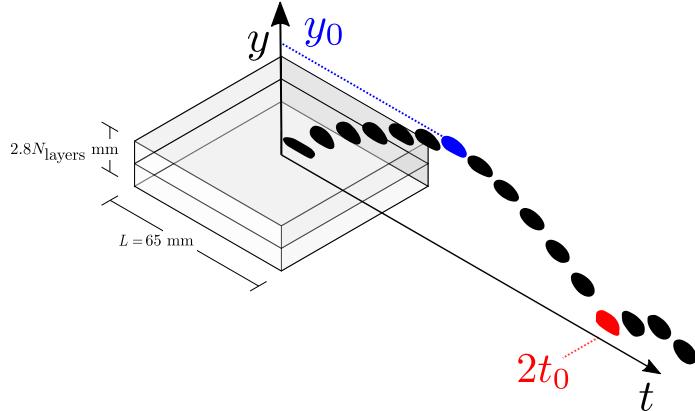


Figure 1: Caption.

number of dielectric lamina. The final, superhydrophobic layer, is produced by laser etching PMMA, and depositing a thin layer of PTFE on the resulting roughness topology to increase the Young's angle. The surface charge density can be modulated during the experiment by means of a 0-2 kV DC-DC converter, which can re-polarize the dielectric substrate by means of an embedded electrode, and the resulting bound charge partially or fully neutralizes the electric field due to the surface ions deposited by corona charging of the electret.

Superhydrophobic substrates were placed on a drop tower rig platform. The rig is released by a solenoid chucked pin at the same moment the exterior drag shield is allowed to fall. This decouples drag acceleration felt by the drag shield from the experiment which experiences approximately  $1 \cdot 10^{-6}g$  of high quality freefall for the 2.1s required for the rig/drag shield assembly to reach the bottom of drop tower 6 floors below.

The superhydrophobic surface was treated with PTFE spray for the first experiment of the day, and approximately every 3<sup>rd</sup> drop following. The general properties and details of the fabrication of these superhydrophobic surfaces are discussed in Section [ref]. Shortly before each drop the surfaces were briefly rinsed with distilled water, and then dried under a fume hood for about 10 minutes. The drop rigs were balanced to prevent unintentional torques from

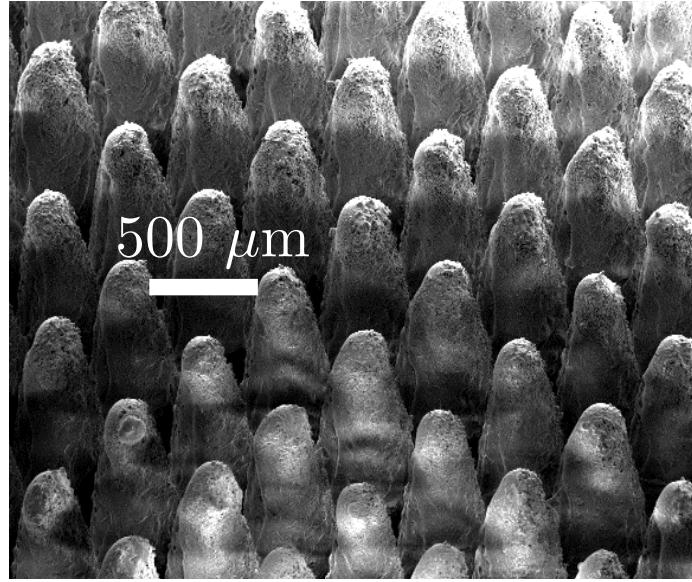


Figure 2: Caption.

contaminating the kinematics data (as angular velocity of the camera-fixed inertial reference with respect to the droplets during freefall would appear as a linear acceleration of the droplet). Surface charge density was measured on the superhydrophobic surface using a *Simco-Ion* FMX-004 electrostatic fieldmeter. This measurement was done with the superhydrophobic surfaces connected by a conductive ground plane by conductive tape, far away from the presence of other conductors. Some additional notes on idiosyncrasies of determining surface voltages (and thence surface charge densities) on insulative dielectric surfaces are provided in Appendix [ref].

Droplets of distilled water, in a range of volumes ( $0 \leq R_d \leq 1\text{mL}$ ), were very carefully deposited on the superhydrophobic surface in the final 1-5 minutes prior to the drop using an ungrounded glass syringe with  $\pm 1\mu\text{L}$  accuracy. Red tracer dye was added to improve trajectory digitization. During the drop, droplet trajectories were observed using a *Panasonic HC-WX970* Camera, shooting in high-speed mode at 120 fps and 1/3000s shutter speed. In a few cases where higher frame rates were needed a *Nikon 1 J1* camera with a 30mm telephoto lens was used, shooting at 400 fps. The experimental test cell

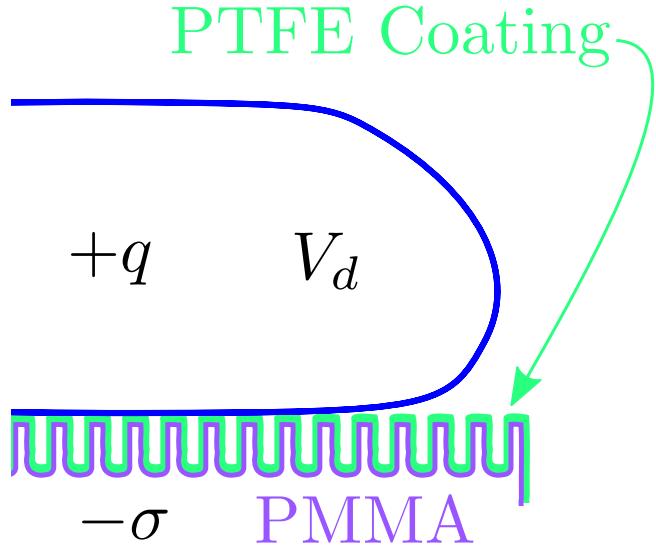


Figure 3: Caption.

was illuminated with a 12V 6000K *SEALIGHT* Chip-On-Board LED strip with a thin semi-opaque plastic film covering to make the light diffuse. The effect of static charges on spontaneously jumped droplet trajectories was studied in terms of the parameters droplet volume  $V_d$ , surface charge density  $\sigma$ , and a dependent variable, initial droplet jump velocity  $U_0$ . Since  $q$  cannot be directly measured insitu during a drop (as high-input resistance electrometers, being notoriously fickle instruments, are not well suited to sudden 15-g decelerations), it was inferred from the statistical distribution determined in corollary experiments. The droplet charge determination is discussed in slightly more detail in Appendix [ref].

## 2.1 Electrets

- The method for determination of the evolution of the surface charge density for low conductivity polymers is described in [Davies, 1967]. Surface charge is injected into the surface of the dielectric mounted on top of an grounded electrode. The charge decay is measured with a calibrated probe at periodic intervals.

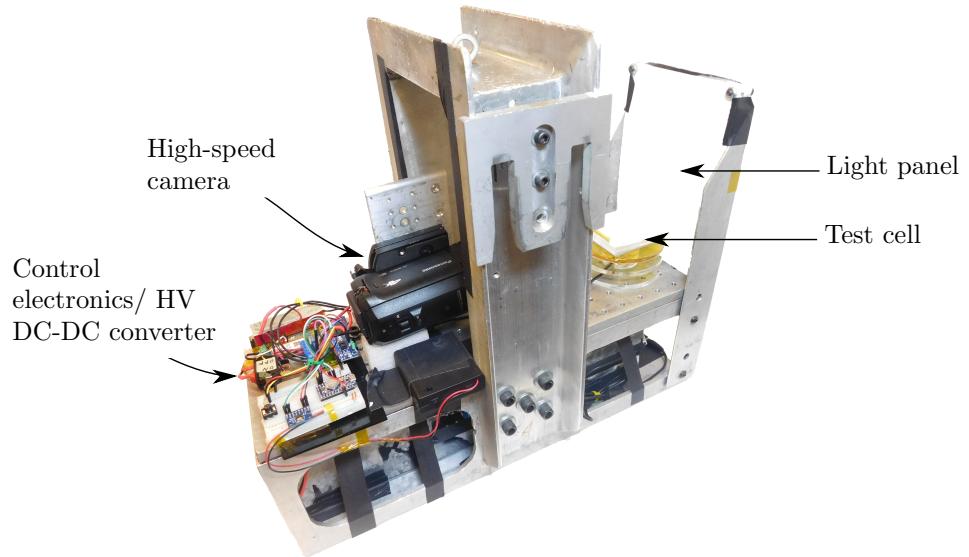


Figure 4: Caption.

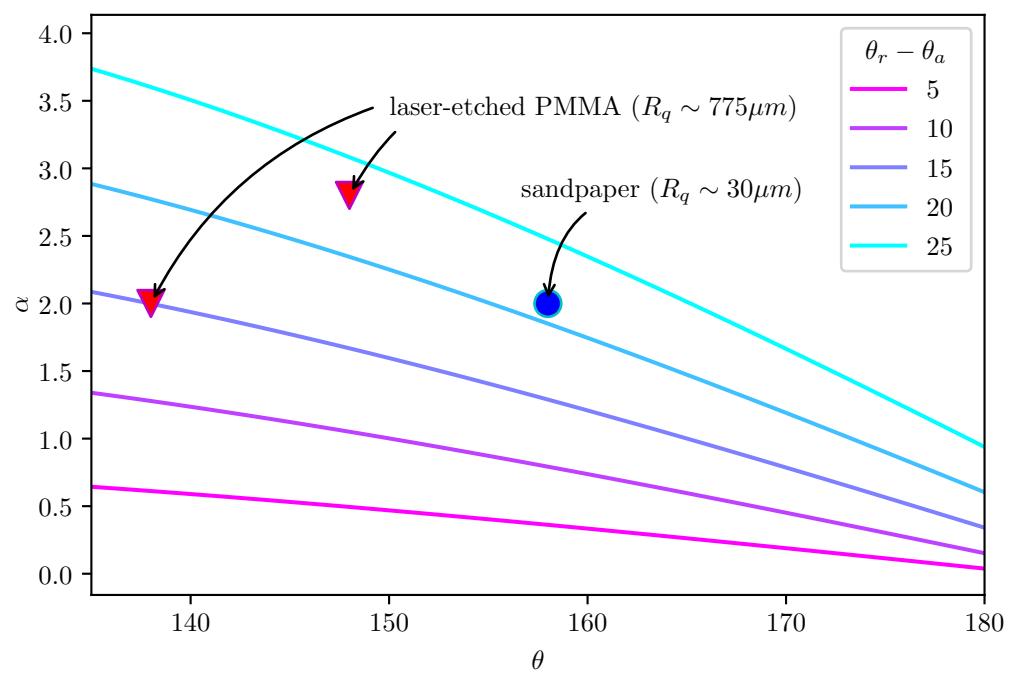


Figure 5: Charge decay.

- Additive stacking of electrets has been used in electret based vibrational energy harvesters [Stacking Electrets for Electrostatic Vibration Energy Harvesters, Wada et al. 2012] and in water desalinization [ref Application of electret technology to low cost desalination, R.C. Amme, 2002]
- There are several charge decay mechanisms: internal ones, such as Ohmic resistance, and external ones such as compensation by environmental ionic species. The relative magnitudes of these charge transport mechanisms, and therefore the stability of the electret varies drastically depending on its initial surface potential, material properties, environment, and charging method. In the case of unshielded electrets compensation by atmospheric ions is significant [E.W. Anderson et al, 1973]. Because environmental convection will tend to maintain a gradient of these ions, sealing an electret in a container from the atmosphere will effectively halt this decay mechanism. Atmospheric humidity and water droplet condensation also significantly increase charge decay (presumably by reducing the surface resistance) [Haenen, 1975].
- According to [Sessler 1987] an electret is a dielectric material with “quasi-permanent” electric charge in the sense that the characteristic decay period of the electret is much greater than a given period of interest.
- Electret charge may be ‘true’ charge in the form of surface or space charges, or polarization charges (such as bound charges). If the electret is not screened by a conductor then it produces an external electric field if the polarization and real charges do not uniformly compensate each other throughout the volume of the electret. For this reason electrets are thought of as electrostatic analogues to permanent magnets (and the name *electret* itself is a portmanteau to the effect conjured by Heaviside in 1892 [ref O. Heaviside, 1982]). Typical commercial electrets are Teflon type polytetrafluoroethylene (PTFE) polymer films on the order of 10-50  $\mu\text{m}$  thick with the charge being primarily real surface charge. Electrets have

a plethora of applications, but most germanely they have been used in Electro-Wetting On Dielectric (EWOD) devices for low-voltage manipulation of small droplets [Wu, 2010].

- Real charge electrets can be produced by contact electrification, injection or deposition) of charge carriers by discharge or electron beam, ionizing radiation, or by frictional triboelectrification. Dipolar electrets by contrast are produced by a combination of polarization at elevated temperatures in a strong external electric field, followed by an annealing process.
- Effective surface charge densities are limited to the material dielectric strengths due to internal dielectric breakdown phenomenon (this typically occurs before external breakdown or Paschen breakdown).
- We use an isothermal electret formation process using a variation of the widely applied corona-charging technique. The typical corona-charging technique uses strong inhomogenous DC electric fields to produce discharge in air at ambient conditions; the dielectric substrate is atop a grounded electrode, and there is a screening potential electrode immediately positioned to control (the the surface potential of the dielectric will tend to saturate at this grid potential if the material is not space-charge current limited). The corona field is usually produced by pin-shaped electrode. In air the most common charge carriers thus produced are  $\text{CO}_3^-$  ions. This approach is known to generally produce samples with fairly uniform surface charge densities.
- Some work [ref J. van Turnhout, 1975] showed using a Thermally Stimulated Current (TCS) measurements that in 4.8 mm thick polymethyl methacrylate (PMMA) polarization of the dielectric is non-uniform, due to real space-charge mostly ( $\sim 90\%$ ) residing in a thin (0.1-0.2 mm) layer near the free-surface of the sample.
- Notes on our version of the corona-discharge process, figure, plot of charge

decay.

- We plainly see a cross-over effect in the decay of the surface potential in our electret samples, whereby the samples charged to higher initial surface potentials decayed faster and reached the lower overall final potentials. This is a well known effect in polyethylenes charged by corona [Ferreira et al. 1992]

### 2.1.1 Surface charge density

Since the laser-etched PMMA superhydrophobic surface is a dielectric material there are some additional considerations for determining the surface charge density. Firstly, for an insulated conductor  $A$  with charge  $q$ , the charge distribution in equilibrium will be such that the field on the interior of the conductor is zero, with field lines being normal to the surface and the integral of the field strength  $\mathbf{E}$  from any point  $P$  in or on the conductor to a ground point  $G$  is a constant given by

$$V = \int_P^G \mathbf{E} \cdot d\mathbf{a},$$

where  $V$  is the voltage or potential of the conductor (or alternatively, the surface voltage potential). The voltage  $V$  and the charge  $q$  are proportional and  $q$  is usually given by

$$q = CV,$$

where  $C$  is the capacitance of the insulated conductor and is determined by the conductor's size and shape, and its placement relative to the conductors and ground. However, the case of a charged insulator departs significantly from that of the insulated conductor. Here we have the possibility of the density and net polarity of the charge varying as a function of position about the surface. The field of the interior may not be zero, and the field lines are not necessarily normal to the the surface. The integral of the field strength from a point on or in the insulator to ground is usually different from point to point. In general the surface voltages will vary from point to point on an insulator (including point

on the interior of the insulator. Therefore it is not possible to uniquely define the voltage of a charged insulator, except in certain contrived cases (notably the potential of a uniformly charged sphere positioned infinitely far away from any conductors). It follows that there is no definition for the capacitance of an insulator as well. Though insulative sheets may store charge and discharge in the manner of capacitors, this discharge is never particularly complete.

There are two cases where we can make meaningful quantitative measurements of state variables for charged insulators: we can measure the electric field arising from the charge, and possibly the total charge for uniformly charged free insulative sheets, and uniformly charged insulative sheets backed by a grounded conductor. To do this we can use field meters and non-contacting voltmeters. If the field strength indicated on the meter is  $\mathbf{E}$ , the charge density  $\sigma$  on the part of the insulator in the front should be  $\sigma = \epsilon_0 \mathbf{E}$ . If a non-contacting voltmeter is placed at a distance  $d$  from the sheet, then the surface voltage  $V_s$  indicated on the meter is given by

$$V_s = d\mathbf{E} = d\frac{\sigma}{\epsilon_0}.$$

An ideal approach to determining surface charge on a dielectric surface is to screen perturbing effects of external electric fields. This is partly accomplished by grounding the fieldmeter, and by placing the dielectric sample on a grounded conductive plate backing. In this case the surface charge density is determined from

$$\sigma = \frac{V_s \epsilon}{l},$$

where  $l$  and  $\epsilon$  are the absolute permittivity and thickness of the dielectric surface respectively. The measured surface voltage is a function of position away from the charged dielectric. In most cases this function is relatively constant at a distance about 1-2 cm away from the surface (there is some measurement error in surface voltage due to small mispositioning of the electrostatic fieldmeter, say by  $\pm 1$  mm) and add it to the precision error of measurement in  $V_s$ .

A further consideration is the possibility of the change in total charge during a typical experimental timescale. If we consider the drop rig to be a ground (which seems reasonable given that the rig is isolated from true ground, but is at some reference voltage with respect to the surface charges on the dielectric, it also has an abundance of free charge carriers, that it, it is conductive), then there will be both bulk and surface decay of the charge on the dielectric. The evolution of the charge can be approximated by

$$\sigma = \sigma_0 e^{\frac{-t}{\epsilon\rho}},$$

where  $\sigma_0$  is the initial surface charge density, and  $\rho$  is the bulk resistivity (which can also be reframed in terms of conductivity by  $\rho = 1/\gamma$ , where  $\gamma$  is the conductivity). For an example case of a surface with an initial surface charge density  $\sigma = 2.4 \cdot (10^{-6}) \text{ C/m}^2$ , relative permittivity  $\epsilon = 3.5$  and bulk resistivity  $\rho = 1.6 \cdot (10^{16}) \Omega/\text{cm}$  such as with the case of 1/4" continuous cast acrylic sheet, then the time constant  $\tau = \epsilon\rho$  is approximately 5000 s, which is a great deal longer than the typical time period for of a drop tower experiment. Add specific notes on charge measurement procedure.

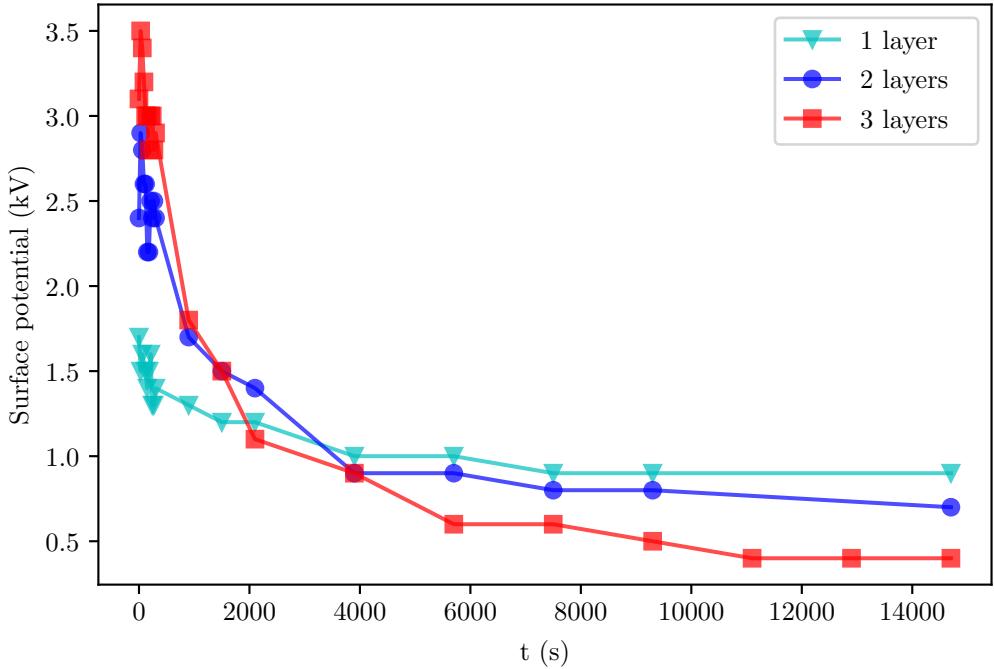


Figure 6: Charge decay.

### 3 Data Munging

Digitization of droplet trajectories requires several steps of post-processing. Video is first decomposed into a sequence of still images. Trajectories are captured using the particle tracking module in *Fiji*[ref], a derivative of the popular *ImageJ*[ref] package for scientific image analysis. The series is stabilized to remove the effect of drop transients from the kinematic data[ref]. The series of still images is cropped, and the background (that is, the low-entropy pixels) of the series is removed using a builtin “rolling ball” algorithm. Each still is then split into its constituent RGB maps. In this case the green channel images contained the most information, so these were then globally thresholded using the Triangle algorithm to recover a map of the pixels corresponding to the droplet’s approximate position in the original still. Finally ellipses are fitted to the pixel map stepping through the time series to determine the positions of the centroid, and the semi-major and minor axes of the droplets during the drop. The results

of the particle capture in *Fiji* are shown in Figure [ref]. Out of plane motion for ‘weakly 3D’ trajectories is inferred from the change in tracked area of the droplet. Droplet positional data was smoothed ex post facto using a Savitsky-Golay filter. Additional details regarding the data smoothing methodology are presented in Appendix [ref].

## 4 Parameter Estimation

We find the parameters  $\mathbf{x}$  that solve the inverse problem  $G(\mathbf{x}) = \mathbf{d}$ , using a direct search method (*Nelder-Mead*).

$$\min \chi^2 = \min \sum_{i=1}^n \frac{(y_d(t)_i - y_G(t, \mathbf{x})_i)^2}{y_G(t, \mathbf{x})_i}$$

$$\mathbf{x} = \begin{cases} q \\ V_d \\ \sigma \end{cases} \text{ subject to constraints } g = \begin{cases} V_d \pm u_{exp} \\ \sigma \pm u_{exp} \\ y_0 \pm u_{exp} \\ t_0 \pm u_{exp} \end{cases}$$

where  $y_G(t, \mathbf{x})$  is a numerical solution of the equation of motion

$$my'' = \frac{1}{2}\rho C_D A_d y'^2 + qE(y) + \frac{1}{4} \frac{Kq^2}{y^2} + \frac{1}{2}E(y)^2 \nabla \epsilon$$

In particular, it is impractical to directly measure the droplet free charge  $q$ , during a bounce experiment. Our work flow to identify this parameters is as follows:

1. Experimentally vary  $V_d$ ,  $\sigma$  and capture droplet trajectories using a high-speed camera.
2. Digitize droplet trajectories by using automatic tracking of ellipse-fitted centroids on the thresholded video.

3. Slice droplet trajectories by their bounce minima, and apply a smoothing filter.
4. Extract the droplet charge (and other experimental parameters) by maximizing the log-likelihood of the data given the dynamical model and parameters, by varying the parameter vector using a direct search optimization.

#### 4.1 Inverse Problems

We have a simple model for 1D projectile motion of charged droplets within Coulomb potential wells in low-g. Using various scaling arguments we have gleaned from this model a series of non-dimensional numbers characteristic of droplet bounce apoapses and times of flight, but these dimensionless groups depend on a set of physical parameters. Unfortunately not all of these parameters are physically practical to accurately measure by experiment. Droplet free charge  $q$ , in particular, could in principle be directly measured by collecting the charged drops in a faraday cup under low-g and measuring the change in capacitance of the cup using a very high input-resistance electrometer, but this is a problematic experiment to set up in a drop tower from a practical standpoint. The other state variables we can directly measure by experiment with varying levels of accuracy. To measure the charge,  $q$ , we instead turn to parameter estimation techniques.

Suppose we have a model  $G(\mathbf{x})$ , with a vector of parameters  $\mathbf{x}$ , and set of (noiseless) observations  $\mathbf{d}$ , the we naturally expect there to exist a relationship

$$G(\mathbf{x}) = \mathbf{d},$$

where the operator  $G$  might be an ODE. Suppose the model  $G(\mathbf{x})$  is the ODE

$$\frac{dy}{dt} = f(t, \mathbf{y}; \mathbf{x}), \quad \mathbf{y} \in \mathbb{R}^n,$$

and a collection of  $n$  measurements of experimental data

$$\mathbf{d} = (t_1, \mathbf{y}_1), (t_2, \mathbf{y}_2), \dots, (t_k, \mathbf{y}_k).$$

The process of fitting a function, defined by a collection of parameters, to a data set is called the discrete inverse, or parameter estimation problem (as opposed to the *forward problem* to find  $\mathbf{d}$  given  $\mathbf{x}$  and  $G(\mathbf{x})$ ). This is a familiar procedure when the determination of model parameters is done using linear or polynomial regression. However there are approaches even to fitting an arbitrary function to a noisy and sparse dataset. In this work we use the conventional Maximum Likelihood Estimate (MLE) method to identify the model parameters.

Using MLE we don't ask the question: "what is the probability that my set of model parameters is correct?" (because the probability is very nearly zero!), but rather "given my set of model parameters, what is the probability that this data set occurred (what is the likelihood of the parameters given the data)?". Bayes' Theorem holds that

$$\text{prob}(X|D, I) = \frac{\text{prob}(D|X, I) \times \text{prob}(X|I)}{\text{prob}(D|I)}$$

where  $D$  are our observations (dataset),  $X$  is our vector of parameters, and  $I$  is general background information about the problem including our mathematical model (for instance the ODE above), and

$\text{prob}(X D, I)$	posterior probability density function,
$\text{prob}(D X, I)$	likelihood function,
$\text{prob}(X I)$	prior probability density function,
$\text{prob}(D I)$	evidence.

The posterior probability density function (PDF)  $\text{prob}(X|D, I)$ , is ultimately what we want to estimate, the prior PDF  $\text{prob}(X|I)$ , reflects our knowledge of the system, and the evidence  $\text{prob}(D|I)$ , is the likelihood of the data based

on our knowledge. We also note that since it only makes sense to compare the conditional PDF's for the same data, we can ignore the denominator (that is, the evidence). We further note that the prior  $\text{prob}(X|I)$ , is fixed before our observations and so can be treated as invariant to our problem. We can therefore infer that  $\text{prob}(X|D, I) \propto \text{prob}(D|X, I)$ . The MLE for the model parameters  $\mathbf{x}_0$ , then is given by the maximum of the posterior PDF, which is equivalent to the solution of the ODE given the the parameters  $\mathbf{x}$ , that produces the highest probability of the observed data. Since the likelihood  $\mathcal{L}(\mathbf{x}) = \prod_i^n \mathcal{P}_i$ , and the probability  $\mathcal{P}$ , of any single observation is less than one, then the total likelihood which is the product of a large number of probabilities tends to be vanishingly small. The more well behaved log-likelihood is given by

$$\mathcal{M} = \ln(\mathcal{L}) = \ln(\text{prob}(D|X, I)) = \text{const} - \frac{\chi^2}{2}$$

where

$$\chi^2 = \sum_{i=1}^n \frac{(y_{di} - y_{Gi})^2}{\sigma_{di}^2}$$

is the  $\chi^2$  goodness-of-fit,  $y_d = y_d(t)$  is a observation of droplet position at a point in time, and  $y_G = y_G(t, \mathbf{x})$  is the droplet position predicted by the solution to the equation of motion at time  $t$ , and  $\sigma$  is the standard error of the position measurement. If the number of data points  $n$ , is small was can use the Poisson error  $\sigma_d^2 = y_G$ . The optimal parameter set is the one with the highest probability of observing the data (the maximum of the posterior PDF) and can be determined by maximizing the log-likelihood  $\mathcal{M}$  (or minimizing  $\chi^2$ ) of the data  $\mathbf{d}$  with respect to the parameter set  $\mathbf{x}$ . Thus parameter estimation is a variety of optimization problem.

## 4.2 Smoothing

All optimization methods, explicitly or implicitly follow gradients towards an optimum. In a parameter estimation problem, if we approximate these gradi-

ents by finite differences, then the noise manifests itself as amplification of the roughness in the hyper-response surface. Gradient based optimizers do poorly in these situations because they tend to converge to local minima. While so called gradient free algorithms offer an improvement in this regard, speed of convergence and the quality of the MLE is improved by smoothing the objective function. This is equivalent to smoothing the underlying dataset.

Our choice of smoothing approach depends principally on the nature of the errors in the dataset. The sources of error include misalignment of the camera, error in the fiduciary length scale, perspective due to objects (subject or reference scale) being out of the photographic plane, and various errors arising in the digitization process (including the difference between the thresholded ellipse fitted centroid and the true centroid of the non-ellipse drop centroid). Some of these errors are systematic in origin and introduce consistent biases into the data (e.g. coherent spectral sources, rather than truly stochastic noise). Data smoothing does little to help systematic errors in that they are usually of lower frequency than the signal. Random errors, by contrast, are assumed to have a Gaussian distribution (by the central limit theorem), and are independent of the signal (which inherently results from a deterministic process).

We experimented with a variety of filters implemented in the `scipy.signal` *SciPy* module on a representative set of trajectory data; these methods include 1D Gaussian convolution, Wiener, Butterworth, and Savitsky-Golay filters. Qualitatively comparing these smoothing methods (by hand tuning filter orders and window sizes) we find that we loose too many data points in the smoothing process, large amplitudes are overly smoothed by repeated filtering passes, or there are significant end effects for most of these methods. A comparison of these smoothing approaches on a representative trajectory data set are shown in Figure 7.

The Savitsky-Golay, and Butterworth filters both produce fairly smooth derivatives as can be seen in Figure 8; but the small window-size needed for Butterworth filter tends to also produce a noticeable end effect. The Savitsky-Golay

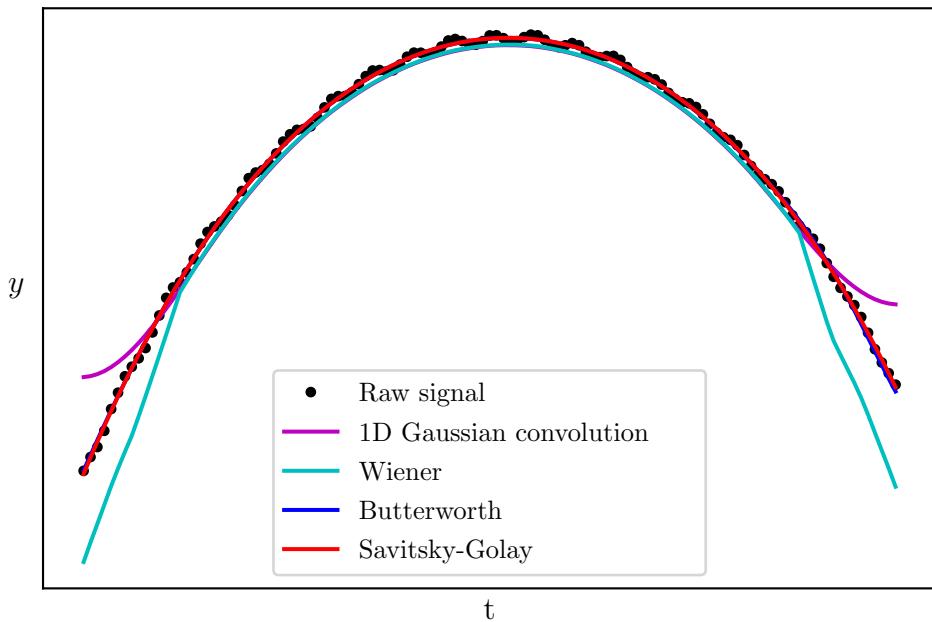


Figure 7: The underlying signal is ‘noisy’, due partially to deterministic errors in determining the centroid position. These deterministic errors are largely due to droplet oscillations, especially the rapidly damped higher harmonics which do not have azimuthal symmetry. There is also Gaussian error in the ellipse fitting due to thresholding and noise in the video itself. We see that 1D Gaussian convolution and Wiener filters suffer from significant end effects. At this scale Butterworth and Savitsky-Golay filters are nearly indistinguishable.

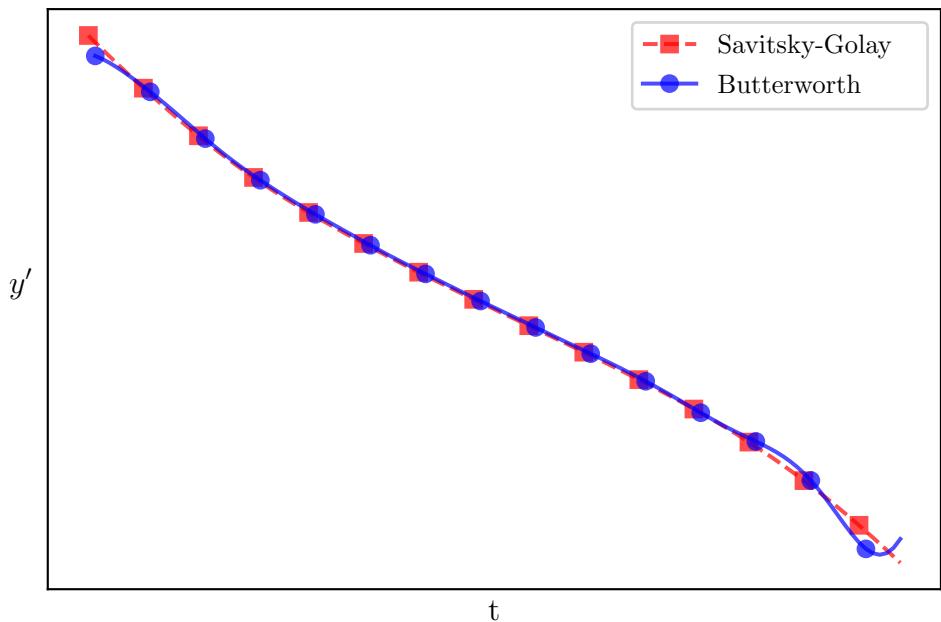


Figure 8: Comparing the first derivatives of the Butterworth and Savitsky-Golay filters we see that the Butterworth filter also suffers from a slight end effect. This implies that the optimization will find a different optima of the likelihood depending on which type of filter is used.

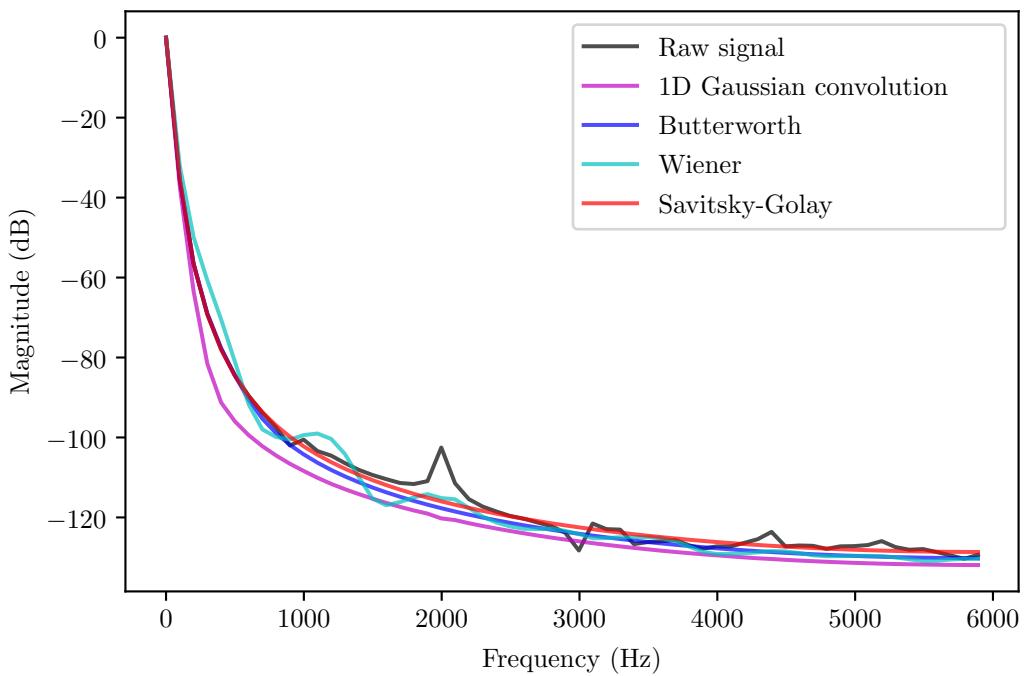


Figure 9: The power spectra has a peak at 1 Hz, which is the droplet trajectory parabola itself, and smaller peaks in the kHz range corresponding to various noise frequencies. The Savitsky-Golay and Butterworth filters seem to have the least distortion of the power spectra of the true signal. Both also do a good job of attenuating the noise at the 2 kHz peak.

filter essentially uses a moving-window based on local least-squares polynomial approximations. It was shown that fitting a polynomial to a set of input samples and then evaluating the resulting polynomial at a single point within the approximation interval is equivalent to discrete convolution with a fixed impulse response [Savitsky, 1964]. A beneficial property of this kind of low-pass filter is their tendency to maintain waveform amplitudes, and so they are attractive in applications having noisy signals with sharply pointed waveforms such as ultrasound or synthetic aperture radar [Schafer, 2011]. Because Savitsky-Golay is a Finite Impulse Response (FIR) filter it requires data points to be equally spaced; to accommodate this we interpolate points between the small gaps which sometimes occur in the tracking results from image analysis. We use a moving window size slightly smaller than the length of the current bounce in a drop jump data set. The windows are piecewise defined by partitioning the data set into a series of individual bounces (the dataset is sliced at minima identified after an initial rough smoothing pass, using the `scipy.signal.argrelextrema()` function). The Savitsky-Golay polynomial order is 4. To understand how these filters differ it is useful to look at their frequency response. In Fourier space, convolution becomes a multiplication, and we can understand what a filter does by looking at which frequencies it lets pass through. We can do this using a Discrete Fourier Transform (though it is worth noting that our signal is not truly periodic). The power spectra for these same data are shown in Figure 9.

### 4.3 Optimization

Most generally a constrained optimization problem is stated as

$$\begin{aligned} \text{minimize: } & f(\mathbf{x}) && \text{objective function} \\ \text{subject to: } & g_j(\mathbf{x}) \leq 0 && \text{inequality constraints} \\ & h_k(\mathbf{x}) = 0 && \text{equality constraints} \end{aligned}$$

$$\text{where } \mathbf{x} = \begin{cases} x_1 \\ x_2 \\ \vdots \\ x_n \end{cases} \quad \text{design variables}$$

Mathematical optimization is the problem of finding minima of a function  $f$ . In this context the function is called the cost, or objective function. The field of mathematical optimization is as old as calculus itself, and the number of particular optimization techniques is correspondingly myriad; particular techniques lend themselves well to particular types of optimization problems. The minima of the objective function  $f$  is sought on a domain  $A$  specified by the constraints of the problem; this domain is usually called the feasible region. Minima of objective function  $f : A \rightarrow \mathbb{R}^m$  are called feasible solutions. If the function  $f$  is convex the feasible solution is the global minimum, otherwise additional local minima exist. The scale of the optimization problem is set ultimately by dimensionality of the objective function. Functions may not always be smooth in the sense of having continuous derivatives, and this is problematic in that optimization methods fundamentally rely on gradients of the objective function. Problems with anisotropic objective functions where there is strong covariance between the parameters, and the gradient vector generally to differ significantly from the Newton direction  $(-\mathbf{H}^{-1}f'(\mathbf{x})^T$ , where  $\mathbf{H}$  is the Hessian matrix) are considered ill-conditioned. Ill-conditioned problems gradient based deterministic search tend to converge slowly as they take a zigzagging path determined by the local value of the gradient rather than following the Newton-direction vector which towards the minimum. Numerical optimization may deal with black box functions (where we do not have an explicit mathematical expression of the function we are optimizing). Black box problems are challenging because we do not have access to analytic gradients of the objective function, and approximating them by finite-differences is slow and noisy. In general, noisy, black box, non-linear, non-quadratic, non-convex, constrained, ill-conditioned, high-

dimensional objective functions are problematic to optimize. Unfortunately, problems of this type are the essence of the parameter estimation, which often leads to its characterization as an ‘art’ rather than a precise science (though we submit that it is a dark art).

The equation of motion behaves stiffly due to the large disparity in Coulombic, image charge, and dielectrophoretic length scales. We integrate it numerically using the `odeint` *Scipy* module. This is a shake-and-bake Python wrapper for the venerable 1982 *netlib ODEPACK* library double-precision `lsoda` (Livermore Solver for Ordinary Differential equations with Automatic method switching for stiff and nonstiff problems) integrator [ref]. The function switches between Adams (nonstiff) and Backwards Differentiation Formulas (BDF, stiff) according to the dynamic value of a set of stiffness eigenvalues.

Our specific optimization problem is non-convex, mixed discrete-continuous black-box (noisy), and highly ill-conditioned which is essentially the worst the worst case scenario for an optimization problem. The ill-conditioning arises due to the strong covariance between several of the model parameters (particularly  $q = q(V_d, E_0)$ ). The non-convexity of the problem implies that there are many local minima of the objective function. While in principle a gradient-based optimizer (for instance using the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) [Nocedal, J, and S J Wright. 2006]) could be used by using finite-differences to obtain approximate gradients of the  $\chi^2$  objective function, in practice doing so is extremely problematic because the noise-to-signal ratio of the objective function scales like  $\mathcal{O}(f)$  for  $\frac{df}{dt}$  and  $\mathcal{O}(f^2)$  for  $\frac{d^2f}{dt^2}$  which will tend to cause convergence to a local minima with is an artifact of the likelihood response surface [Wood, 1982: Refs 5, 14, 49, 95.]. As a further practical matter, given the relatively expensive function-calls (which requires solving a stiff, non-linear ODE) gradient-free approaches tend to offer better performance regardless[Singer and Mead, 2009].

We use a gradient-free, direct-search approach: Nelder-Mead [Nelder, 1965] implemented in `scipy.optimize` [Jones et al. 2001 –]. Nelder-Mead is robust to

noise (relatively speaking), and is relatively thrifty with our extremely expensive function-calls. *Nelder-Mead*, sometimes called simplex-search or downhill-simplex, is a heuristic search method, with no guarantee of optimal solutions, but is well-established and widely used despite that. *Nelder-Mead* is based on the concept of a  $N$ -simplex, which generalizes a triangle into higher dimensions as a polytrope of  $N + 1$  vertices in  $N$  dimensions. It uses only-function calls and expands or contracts the simplex according to the function values at its vertices in a way visually reminiscent (in  $\mathbb{R}^2$ ) of the oscillations of the jumping droplets themselves (in fact *Nelder-Mead* is sometimes also called the “amoeba method”). Very little is known about the convergence properties of the *Nelder-Mead* algorithm in its classical form for non smooth objective functions ([Price and Coope, 2003]), except that in general it doesn’t satisfy the properties required for convergence by other direct search algorithms: that the simplex remains uniformly non-degenerate, and that some form of “sufficient” descent condition for function values at the vertices is required at each iteration. Scaling can help solve convergence problems and improve numerical stability. We precondition the optimization problem by minimizing  $\ln(\chi^2)$ , and using a naive scaling (scaling variables such that their magnitudes  $\sim 1$ ) of our constraints by their initial guesses. Here is goal is to make the problem equally sensitive to steps in any direction. *Nelder-Mead* is not a global optimizer, though there are variants which use sequential local searches with probabilistic restarts to achieve globality. However global optimization usually comes at a tremendous computational cost. However, *Nelder-Mead* behaves less locally than many gradient-based approaches. The convergence history of the parameter MLE using *Nelder-Mead* for a single drop jump experiment is shown in Figure 10.

#### 4.4 Identifiability

That we are capable of fitting any arbitrary model to a dataset given sufficient degrees of freedom in our parameters is admittedly a disconcerting issue, begging the question “given the structure of the model is it possible to uniquely estimate

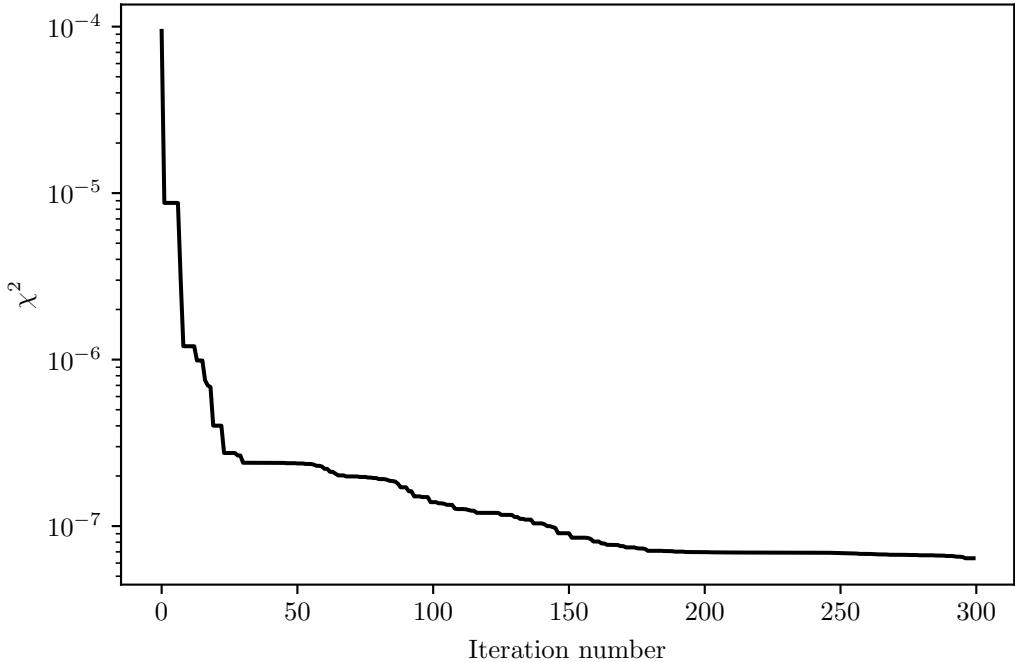


Figure 10: As is typical with *Nelder-Mead* much of the improvement in  $\chi^2$  is realized in the first few iterations. Overall the rate of convergence is sub-linear, which is to be expected for non-linear constrained problems using an heuristic algorithm.

the unknown parameters?” This question is called the problem of identifiability. However, some of the inverse model parameters are constrained by our experimental observations of them and their associated measurement uncertainties. This, we hope, makes the specter of an over fitted model less frightening, but does convert our unconstrained optimization problem to an constrained one which raises special difficulties of its own, which we discuss below.

We’re interested in the variance and co-variance as a means to determine the quality if the parameter estimate. The  $(i,j)$ -th element of the matrix  $\sigma(\mathbf{x}, \mathbf{y})$  is equal to the covariance  $\text{cov}(X_i, Y_j)$  between the  $i$ -th scalar component of  $\mathbf{x}$  and the  $j$ -th scalar component of  $\mathbf{Y}$ . Here the concept of error bars in linear correlation associated with a covariance matrix are not suitable. We might try to generalize the idea of confidence intervals to a multidimensional space, but usually it will be hard to describe the surface of the (smallest) hyper-volume

containing 90% of the probability in just a few numbers. The situation is worse if the probability density function has several maxima. However we notice that

$$[\sigma^2]_{ij} = - \left[ (\nabla \nabla \mathcal{L})^{-1} \right]_{ij} = 2 [\nabla \nabla (\chi^2)]_{ij}^{-1} = - [H^{-1}]_{ij}$$

where  $H$  refers to the Hessian matrix and  $[\sigma^2]_{ij}$  is the covariance matrix  $C$ . The issue of identifiability is especially fraught for non-linear, black box type problems were it is difficult to explicitly evaluate the Hessian. The likelihood function (and thus the posterior probability density function) are defined completely by the optimal solution  $\mathbf{x}$  and the second derivative of  $\mathcal{L}$  at the maximum, which corresponds to the covariance matrix  $C$ . The standard errors (marginal variances) are the square roots of the diagonal of the covariance matrix. Our relative errors thus produced are extremely small ( $> 1\%$ ). The Hessian matrix must be negative definite for  $\mathcal{L}$  to have a maximum at  $\mathbf{x}_0$ . We can use the condition number

$$\text{cond}(A) = \frac{\max[\text{eig}(A)]}{\min[\text{eig}(A)]}$$

as a means to evaluate to stability of our problem. We find typical condition numbers  $\sim \mathcal{O}(10^{27})$  which indicate the problem is strongly ill-conditioned near the mimima  $\mathbf{x}_0$ .

The *Nelder-Mead* direct search method cannot be used with explicitly constrained problems. However, there are various implicit approaches to (approximately) solving general constrained problems using unconstrained algorithms. Generally, this is achieved by domain transformations or the use of penalty functions. By the addition of a penalty function which depends in some way on the values of the constraints to the objective function, we minimize a pseudo-objective function where the in feasibility of the constraints is minimized simultaneously to the objective function.

There are various penalty function schemes. We use an Exterior Penalty Function as a simple way of converting converting the constrained problem into an unconstrained one. These are are especially useful in cases where the con-

straints are not “hard” in the sense that they need to be satisfied precisely. General penalty functions, which are sequential unconstrained minimization techniques, reformulate the general constrained problem as the pseudo-objective function given by

$$\phi(\mathbf{x}, r_p) = F(\mathbf{x}) + r_p P(\mathbf{x})$$

where  $P(\mathbf{x})$  the penalty function, is given by

$$P(\mathbf{x}) = \sum_{j=1}^m \{\max[0, g_j(\mathbf{x})]\}^2 + \sum_{k=1}^l [h_k(\mathbf{x})]^2. \quad (1)$$

We see from Equation 1 that there is no penalty if the constraints  $g_j(\mathbf{x})$ , and  $h_k(\mathbf{x})$  are satisfied.

The Exterior Penalty Function specifically (and all Penalty Function approaches in general) do have several drawbacks. Namely these include the possibility of the objective function being undefined outside of the set of feasible solutions. Additionally, by naively “encouraging” feasibility of the solution using large values of the penalty parameter,  $r_p$ , we will tend to ill-condition the unconstrained formulation of the problem (though in our implementation the preconditioning tends to make the pseudo-objective function less and less sensitive to the constraints as the likelihood is approaches a maximum). We use the measured values of  $u_0$ ,  $V_d$ , and  $E_0$ , and the informed guess  $q \approx kV_d^{2/3}E_0$  where  $k$  is a constant  $k \approx 1^{-11}$  as an initial guess for the parameter vector  $\mathbf{x}$ . We stop the optimization after 300 iterations rather than waiting for convergence.

Despite their wide use in parameter estimation, there is very little known about convergence properties of *Nelder-Mead*. The iterative procedure may thus converge very slowly near the optimum, unless the Hessian is ”well-conditioned”, i.e., the condition number is near one. However, if parameter estimates are highly correlated among themselves, the Hessian matrix is near-singular and its inversion is either impossible or involves significant numerical error. Posteriori verification of the results is crucial to bound the identifiability of the parameter

estimation problem, yet it has not been done yet. [Ref] suggests verification by generation of Monte Carlo data sets.

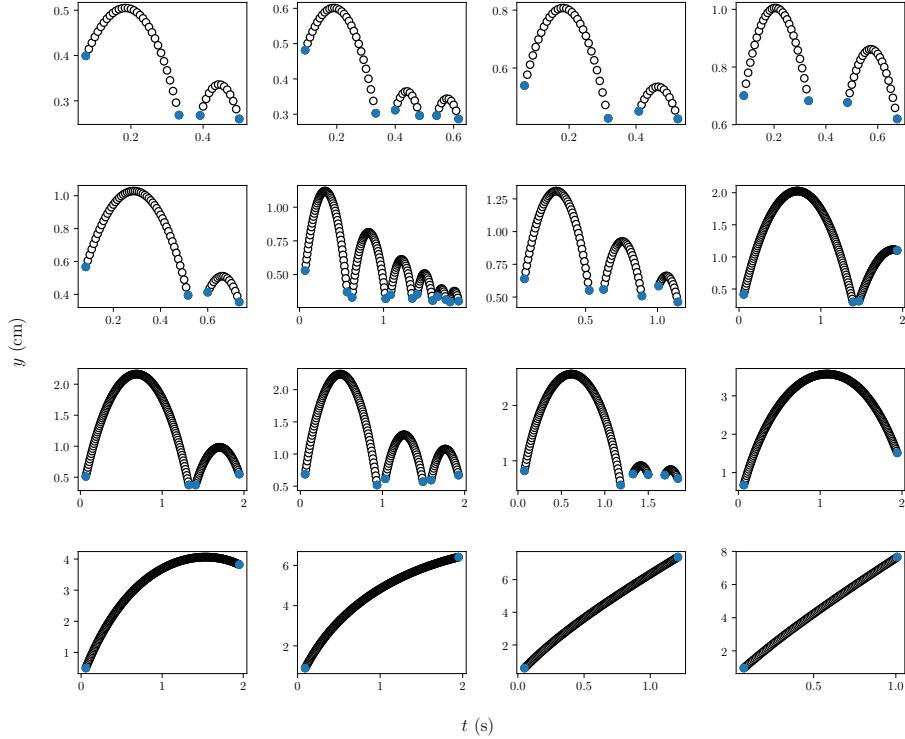


Figure 11: A series of filtered droplet trajectories arranged by increasing apoapse. The blue dots represent either the beginning and end of the experiment, or points at which the droplet is either coming into, or leaving contact with the surface.

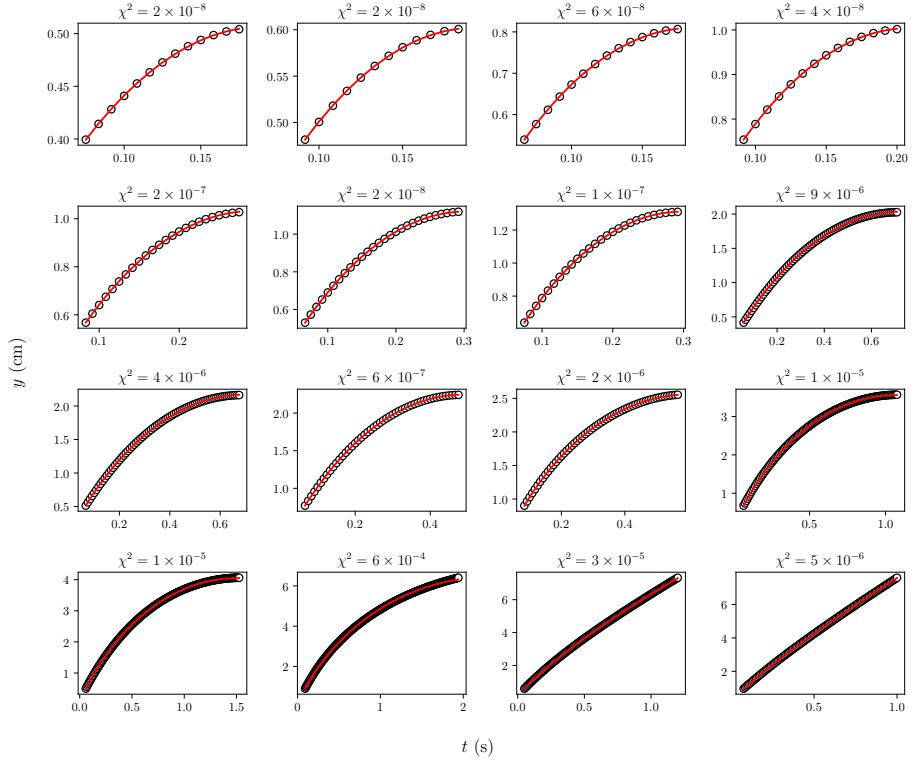


Figure 12: A series of droplet trajectories showing the results of the parameter estimation. The trajectories are shown only up to the apoapse of the first bounce. The red lines show the ODE solution with given the MLE parameter vector.  $\chi^2$  goodness-of-fit varies between  $1 \times 10^{-5}$  and  $1 \times 10^{-8}$  with the better fit occurring typically for the droplets with the lowest apoapses.