DEVELOPMENT OF DROPWISE ADDITIVE MANUFACTURING WITH NON-BROWNIAN SUSPENSIONS: APPLICATIONS OF COMPUTER VISION AND BAYESIAN MODELING TO PROCESS DESIGN, MONITORING AND CONTROL

by

**Andrew Joseph Radcliffe**

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy
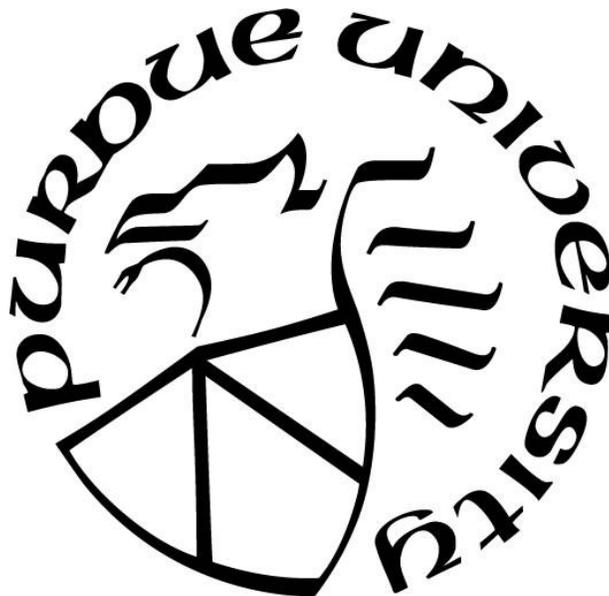
Davidson School of Chemical Engineering

West Lafayette, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Gintaras V. Reklaitis, Co-Chair

Davidson School of Chemical Engineering

Dr. Zoltan K. Nagy, Co-Chair

Davidson School of Chemical Engineering

Dr. Lynne S. Taylor

Department of Industrial and Physical Pharmacy

Dr. Michael T. Harris

Davidson School of Chemical Engineering

Dr. Vivek Narsimhan

Davidson School of Chemical Engineering

Approved by:

Dr. John A. Morgan

for Agnes

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

14

16

NOMENCLATURE, CHAPTER 3

APAP = acetaminophen
API = active pharmaceutical ingredient
$A_{particle}$ = projected area of particle
AR = Aspect ratio
D_{10} = diameter under which 10% (by volume) of particles exist
D_{50} = diameter under which 50% (by volume) of particles exist
D_{90} = diameter under which 90% (by volume) of particles exist
DOD = drop-on-demand
η_0 = carrier fluid viscosity
η_r = suspension viscosity scaled by $(1/\eta_0)$
f_{drop} = number of drops deposited per unit time
γ = surface tension
$\dot{\gamma}$ = shear rate
HMDSO = hexamethyldisiloxane
HLB = hydrophilic-lipophilic balance
k_B = Boltzmann constant
m_{API} = mass of active pharmaceutical ingredient in dose
mg = milligram
mL = milliliter
MA = mefenamic acid
μ-APAP = micronized acetaminophen
μL = microliter
μm = micrometer
N_{drops} = number of drops
nm = nanometer
PB = phenylbutazone
ϕ = particle volume fraction
ϕ_m = maximum particle volume fraction
PSD = particle size distribution
ρ_0 = carrier fluid density
ρ_p = particle density
r_p = particle radius
R 972 = AEROSIL R 972 Pharma
RSD = relative standard deviation
sf-APAP = micronized acetaminophen
V_{drop} = drop volume
vol% = percent by volume
wt% = percent by weight

ABSTRACT

In the past two decades, the pharmaceutical industry has been engaged in modernization of its drug development and manufacturing strategies, spurred onward by changing market pressures, regulatory encouragement, and technological advancement. Concomitant with these changes has been a shift toward new modalities of manufacturing in support of patient-centric medicine and on-demand production. To achieve these objectives requires manufacturing platforms which are both flexible and scalable, hence the interest in development of small-scale, continuous processes for synthesis, purification and drug product production. Traditionally, the downstream steps begin with a crystalline drug powder – the effluent of the final purification steps – and convert this to tablets or capsules through a series of batch unit operations reliant on powder processing. As an alternative, additive manufacturing technologies provide the means to circumvent difficulties associated with dry powder rheology, while being inherently capable of flexible production.

Through the combination of physical knowledge, experimental work, and data-driven methods, a framework was developed for ink formulation and process operation in drop-on-demand manufacturing with non-Brownian suspensions. Motivated by the challenges at hand, application of novel computational image analysis techniques yielded insight into the effects of non-Brownian particles and fluid properties on rheology. Furthermore, the extraction of modal and statistical information provided insight into the stochastic events which appear to play a notable role in drop formation from such suspensions. These computer vision algorithms can readily be applied by other researchers interested in the physics of drop coalescence and breakup in order to further modeling efforts.

Returning to the realm of process development to deal with challenges of monitoring and quality control initiated by suspension-based manufacturing, these machine vision algorithms were combined with Bayesian modeling to enact a probabilistic control strategy at the level of each dosage unit by utilizing the real-time image data acquired by an online process image sensor. Drawing upon a large historical database which spanned a wide range of conditions, a hierarchical modeling approach was used to incorporate the various sources of uncertainty inherent to the manufacturing process and monitoring technology, therefore providing more reliable predictions for future data at in-sample and out-of-sample conditions.

This thesis thus contributes advances in three closely linked areas: additive manufacturing of solid oral drug products, computer vision methods for event recognition in drop formation, and Bayesian hierarchical modeling to predict the probability that each dosage unit produced is within specifications.

1. OVERVIEW

1.1 Preface

With the passage of time humans have incrementally improved their living conditions, if only because it appeared to be to their profit at that moment [1]. The human condition is to struggle, to daily begin anew at a Sisyphean task, as all organisms must, thus, it is not surprising that anything that eases this burden would be immediately recognizable – this may be one way to explain this tendency toward improvement through the millennia [2]. In the present time, the objectives of innovation and improvement are explicitly stated, and many justifications found in economics, ethics, etc. [3]. However, at the individual level, that is, for people actively pursuing such objectives, an ancient need may be satisfied: the exploration of the unknown.

Compared with even one century previous, the work of doctoral students must necessarily be less general as the state of knowledge is considerably advanced [4]. Yet, it is still possible to partake of the same experience of the unknown. This is an opportunity to make a contribution to one or another research field; it is a small beginning which may blossom over a lifetime into realized improvements. To acknowledge the context within the large world seems to this author a necessity for the reader, as it permits one to view the body of the work as it truly is: built upon the work of others, a small step forward. By declaration at the outset, we may dispense with pretension. Instead, we turn our attention to the journey.

1.2 A General Introduction

In light of this, the attempt is made to present the items of this work in as coherent a manner as possible, placing each chapter within its relevant contextual framework. Before embarking on these, it is worthwhile to not only outline the motivations for the research project more generally, but also to provide the background and work of preceding students.

In the past two decades, the pharmaceutical industry has been engaged in modernization of its drug development and manufacturing strategies, spurred onward by changing market pressures, regulatory encouragement, and technological advancement [5, 6, 7]. As a single aspect of this expansive landscape, drug manufacturing processes and strategies are moving toward modalities such as patient-centric medicine and on-demand production, concomitant with improvements to

product quality and manufacturing efficiency which are achieved through introduction of process control strategies and renewed efforts at model-based design [5, 6, 7, 8, 9]. In a certain sense, it has been a substantial effort in all aspects of process and systems engineering which has fertilized this growth for existing technologies, while simultaneously giving rise to new technologies for process sensing, control and even manufacturing.

Of these new process technologies, those based on additive manufacturing have attracted considerable interest for their potential to perform on-demand production of drug products with individually-tailored properties, and furthermore, to serve as robust platforms capable of facilitating scaling from early drug development to commercial production [10]. Commensurate with the level of interest is the variety of possible systems based on two-dimensional or three-dimensional printing techniques, all of which present the same common advantages of flexibility and precision. However, techniques based on the drop-on-demand printing method possess a distinct advantage from the perspective of quality control: the ability to monitor each drop added to each dose in real time enables one to directly monitor unit-specific potency level. This provides the means to develop a robust control strategy through data-driven modeling work, whereby each out-of-specification unit can be detected and summarily excluded from the batch or lot.

The work of prior students on this research project led to the development of a mini-manufacturing platform which utilized the drop-on-demand method to produced doses through sequential addition of drops to one or more substrates [11, 12, 13, 14]. In their work, the students identified the means to overcome a critical limitation of many of the proposed printing processes: production rate, or, manufacturing throughput. They rectified this by exchanging the original inkjet printheads for a mechanically-actuated drop formation process; flexibility and scaling advantages were not sacrificed by this change of technology, it retained the drop-on-demand print mode, and scale-up could still be achieved by adding more print heads in parallel. The distinction between the inkjet and dropwise additive manufacturing system occurs in the size of drops produced and the frequency at which they are generated. Process throughput is proportional to product of volume times frequency, hence, having found a mechanically-actuated drop generator capable of precision deposition with large drops, this provided strong motivation to develop a system which was inherently capable of higher production rates. The choice of the mechanically-actuated drop generator happened to be fortuitous for different reasons as well, which were realized later when printing of concentrated non-Brownian suspensions was considered; this would have been

impossible with inkjet printers simply because many of the particles are larger than the conduit diameters in such devices.

In addition to considering the drop generator, the platform built by these students included process-wide temperature control systems, an online drop monitoring system and precision substrate staging control which were integrated through a custom LabVIEW-based automation program that enabled development of a real-time process management strategy [11, 12]. Their experimental endeavors demonstrated the feasibility of the system for production of drug products utilizing solvent- and melt-based inks [13, 14]. Control of critical product attributes such as crystal phase and particle size was demonstrated using the substrate temperature control system they developed [15, 16]. Monitoring of product content using a custom drop image acquisition system enabled identification of out-of-specification dose units in real time [11, 17]. Taken together, their work demonstrated the capability of a new manufacturing technology to achieve the goals of flexibility in formulation and precision drug products, while ensuring quality by control strategy.

The next phase of the project was envisioned to be its development for production using concentrated non-Brownian suspensions. The potential advantages of producing doses from such inks are multiple: processing such materials as dry powders is often difficult due to the complexities of powder rheology, hence, processing them as a fluid had the potential to offer improved scalability, with less arduous process development, thereby leading to shorter time-to-market and economic gains. This is particularly relevant as an estimated 40% of drug compounds are poorly water-soluble [18], hence the necessity of bioavailability enhancement, which is most readily achieved through particle size reduction. However, micronization of drug particles exacerbates powder flow issues, thus prolonging process development as granulation unit operations must be used to achieve dry flow properties compatible with a tablet press or capsule-filling machine [19].

As an alternative to the traditional sequence of unit operations leading from crystallization to the tablet press, dropwise additive manufacturing with non-Brownian suspensions circumvents the potential problems inherent to micronized powders by using a fluid-based process, and, moreover, it offers scaling advantages which are unavailable to dry powder processes [20]. Furthermore, it has the potential to drastically reduce the number of unit operations needed to transform crystallized drug compounds into finished products, as a connection to upstream processes requires only that the ink be prepared. Additional benefits which should not be

overlooked are that use of concentrated suspensions for dropwise additive manufacturing enables production of units at high dose potency levels (e.g. 100-200mg active) which may be impractical with solvent- or melt-based inks, and that the production rate achievable on a single print head is on the order of 1000 units per hour [20, 21].

The challenges of suspension-based manufacturing are more numerous than for solvent- or melt-based inks due to the complexities of suspension rheology and the effects of appreciable particle size on drop formation. Therefore, the research project must develop a framework for ink formulation and process operation, in hope that some generalization which facilitates operation with new particle types might be possible. Simultaneously, the performance of the process for producing doses must be quantified, and the product properties must be verified. Then, process monitoring and control strategies must be reassessed in light of the non-idealities introduced by suspensions/particles.

In the chapters that follow, the research efforts in pursuit of these objectives are presented along with their accompanying background and motivations. As often seems to occur in research, the results of initial work lead to other areas of investigation which may be needed to proceed further along the originally intended path. Throughout this work the reader will find these branching points, as the work on mapping operating regimes and developing an ink formulation framework leads to automated image analysis and computer vision methods, whereas the process monitoring and quality control challenges posed by suspensions motivate the use of Bayesian modeling. These ideas merge in the final content chapter in which the monitoring and control strategy is re-envisioned so as to combine machine vision algorithms with Bayesian hierarchical modeling in order to make individual predictions of dose unit and batch quality using online measurements and all the historical data gathered on suspension printing.

1.3   On the Arrangement of Chapters

The arrangement of this thesis follows a format which seems to yield a coherent narrative, and thus, requires a slight rearranging of the temporal order of the works in the latter two panels. In Chapter 2, a brief foreword is provided for each panel; these commentaries serve as introductory thoughts to the work and its motivations. Full descriptions of the motivations and background for each work are provided in the chapters themselves, unless noted otherwise in the foreword. In total, there are 3 content panels; this is followed by a chapter which concludes the work and a chapter

that provides perspectives on potential future directions. The content panels are as follows: Panel 1) Dropwise additive manufacturing using non-Brownian suspensions: Chapters 3 and 4; Panel 2) Computer vision for high-speed imaging of drop-formation phenomena: Chapters 5 and 6; Panel 3) Bayesian modeling for process monitoring and quality control: Chapters 7 and 8.

1.4   References

[1] Dostoevsky, F. (1994). *Notes from underground.* (Trans.: Pevear, R.. & Volokhonsky, L.,). Random House, Inc., New York.

[2] Camus, A. (2013). *The myth of Sisyphus*. Penguin UK.

[3] Musil, R. (1995). *The man without qualities.* (Trans.: Wilkins, S. & Pike, B.). Random House, Inc., New York.

[4] Lévi-Strauss, C. (2012). *Tristes tropiques*. Penguin.

[5] Ierapetritou, M., Muzzio, F., & Reklaitis, G. (2016). Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. *AIChE Journal*, *62*(6), 1846-1862.

[6] Nagy, Z. K., & Braatz, R. D. (2012). Advances and new directions in crystallization control. *Annual review of chemical and biomolecular engineering*, *3*, 55-75.

[7] Kayrak-Talay, D., Dale, S., Wassgren, C., & Litster, J. (2013). Quality by design for wet granulation in pharmaceutical processing: assessing models for a priori design and scaling. *Powder technology*, *240*, 7-18.

[8] Lee, S. L., O'Connor, T. F., Yang, X., Cruz, C. N., Chatterjee, S., Madurawe, R. D., ... & Woodcock, J. (2015). Modernizing pharmaceutical manufacturing: from batch to continuous production. *Journal of Pharmaceutical Innovation*, *10*(3), 191-199.

[9] Teżyk, M., Milanowski, B., Ernst, A., & Lulek, J. (2015). Recent progress in continuous and semi-continuous processing of solid oral dosage forms: a review. *Drug development and industrial pharmacy*.

[10] Daly, R., Harrington, T. S., Martin, G. D., & Hutchings, I. M. (2015). Inkjet printing for pharmaceutics–a review of research and manufacturing. *International journal of pharmaceutics*, *494*(2), 554-567.

[11] Hirshfield, L., Içten, E., Giridhar, A., Nagy, Z. K., & Reklaitis, G. V. (2015). Real-time process management strategy for dropwise additive manufacturing of pharmaceutical products. *Journal of Pharmaceutical Innovation*, *10*(2), 140-155.

[12] Içten, E., Joglekar, G., Wallace, C., Loehr, K., Sacksteder, J., Giridhar, A., ... & Reklaitis, G. V. (2016). Knowledge provenance management system for a dropwise additive manufacturing system for pharmaceutical products. *Industrial & Engineering Chemistry Research*, *55*(36), 9676-9686.

[13] Hirshfield, L., Giridhar, A., Taylor, L. S., Harris, M. T., & Reklaitis, G. V. (2014). Dropwise additive manufacturing of pharmaceutical products for solvent-based dosage forms. *Journal of pharmaceutical sciences*, *103*(2), 496-506.

[14] Içten, E., Giridhar, A., Taylor, L. S., Nagy, Z. K., & Reklaitis, G. V. (2015). Dropwise additive manufacturing of pharmaceutical products for melt-based dosage forms. *Journal of pharmaceutical sciences*, *104*(5), 1641-1649.

[15] Icten, E., Nagy, Z. K., & Reklaitis, G. V. (2015). Process control of a dropwise additive manufacturing system for pharmaceuticals using polynomial chaos expansion based surrogate model. *Computers & Chemical Engineering*, *83*, 221-231.

[16] Içten, E., Giridhar, A., Nagy, Z. K., & Reklaitis, G. V. (2016). Drop-on-demand system for manufacturing of melt-based solid oral dosage: effect of critical process parameters on product quality. *AAPS PharmSciTech*, *17*(2), 284-293.

[17] Içten, E., Purohit, H. S., Wallace, C., Giridhar, A., Taylor, L. S., Nagy, Z. K., & Reklaitis, G. V. (2017). Dropwise additive manufacturing of pharmaceutical products for amorphous and self emulsifying drug delivery systems. *International journal of pharmaceutics*, *524*(1-2), 424-432.

[18] T. Wright, 6 March 2015, Solid Dosage Manufacturing Trends, Contract Pharma; March 6 2015. Available at: https://www.contractpharma.com/issues/2015-03-01/view_features/solid-dosage-manufacturing-trends-724749. Accessed October 8, 2018

[19] Williams, H. D., Trevaskis, N. L., Charman, S. A., Shanker, R. M., Charman, W. N., Pouton, C. W., & Porter, C. J. (2013). Strategies to address low drug solubility in discovery and development. *Pharmacological reviews*, *65*(1), 315-499.

[20] Radcliffe, A. J., Hilden, J. L., Nagy, Z. K., & Reklaitis, G. V. (2019). Dropwise Additive Manufacturing of Pharmaceutical Products Using Particle Suspensions. *Journal of Pharmaceutical Sciences*, 108(2), 914-928.

[21] Radcliffe AJ, Reklaitis GV. Dropwise Additive Manufacturing using Particulate Suspensions: Feasible Operating Space and Throughput Rates. In *Computer Aided Chemical Engineering*, Vol. 40, Elsevier; 2017:1207-1212.

2.   FOREWORD TO ARTICLES

2.1   Foreword to Chapter 3, 4

The initial stage of the research project began with a phase of experimental work devoted to collection of data to demonstrate how such suspensions could be printed, to develop multiple platforms for ink formulation, and to determine the bearing of various factors on printing feasibility and performance. The large body of experimental work provided not only the support for the feasibility of this manufacturing strategy, summarized in the first panel of the triptych, but also much data which formed the bases of the second and third panels. Sections 2.1.1 through 2.1.4 provide a cursory outline of the challenges presented by physics; the interested reader is encouraged to consult Appendix A for a thorough review of the literature which supports the narrative given here.

2.1.1   Suspension Rheology

When one is introduced to fluid mechanics, even at the graduate level, assumptions regarding viscosity and material response to stress are necessary to cover the full range of possibilities in purely ideal fluids – that is, constant viscosity, or the a few non-ideal cases, viscosity can be specified by some known functional form. This does not hold for the type of fluid which the process was to be developed for: concentrated suspensions of micrometer-scale particles with non-uniform size and shape. Even in idealized cases in which particles are all smooth spheres and have uniform size, very complex phenomena are observed – shear thinning, shear thickening, discontinuous shear thickening, hysteresis, flow-driven segregation such as migration of particles in tube flow, and particle-pressure driven segregation [1, 2, 3, 4]. Notably, such phenomena are observed for flow in simple geometries (e.g. a tube), and in the absence of a free liquid surface boundary.

These phenomena are shear-rate and shear-stress dependent – the current state of the literature is divided as to the dominant mechanisms that underlie these complex occurrences [1, 4]. This is to say that simulation of the slurry flow, even in a simple enclosed geometry at idealized conditions (i.e. perfect particle-fluid coupling) is challenging if one wishes to recapitulate all possible occurrences. Thus, for non-ideal particles – distributed size, shape parameters; substantial

surface roughness – in a system that has an oscillatory flow regime which switches from shear rates of 0 to 1000 s^{-1} on every pump actuation, representative simulation is likely to be extremely difficult. Realistically, detailed rheology experiments would be necessary to validate simulations in such a non-ideal case – this would have been a research project in itself, outside the realm of my research objectives.

However, an understanding of the phenomenology of such fluids is invaluable for any modeling work, as one is then supplied with the type of approximations possible and the corresponding assumptions, in addition to the relevant variables to flow conditions, ink formulation, and their interactions. Using dimensional analysis, one can reduce the numerous (20+) variables to a smaller set, and hence make inferences to guide development for new particles by using the shear rheology to generalize across particle size and shape.

2.1.2   The Physics of Drop Formation

The manufacturing process relies on moving fluid from the reservoir through tubing to the nozzle, hence the need to consider suspension flow itself, but the critical step of the process is drop formation. Without the ability to reliably and consistently print drops of approximately equal volume, the manufacturing scheme would be untenable, thus, development – operating conditions, ink formulation – must consider both simultaneously [5, 6].

Foremost, drop formation and jet breakup, even in ideal, particle free fluids exhibit remarkably complex physics due to the singular nature of these phenomena and the inherent complexity of free-surface flow in which the liquid surface is not enclosed by walls on one or more sides [7]. Notably, the driving physical phenomena change in phases as the fluid column proceeds toward the singularity – thus, the characteristic length/time-scales are also changing, as are the relevant force balances. Therefore, scaling with fluid properties changes based on the thinning regime – or, in other words, on the filament diameter. Additionally, the added term in the stress tensor – capillary force – is dependent on the curvature of the liquid surface, thus, the whole breakup process is dependent on the initial shape of the filament (as is each and every time step). Various asymptotic scaling laws have been derived which capture the stages of breakoff in ideal fluids, relating dependencies on viscous, inertial and capillary force to filament diameter [5, 6, 7].

Drop breakoff in non-ideal fluids has been studied perhaps most extensively for polymeric or polymer solutes, but this is not the focus here. Colloidal suspensions – where particle size is in

the 10's or 100's of nanometers – are routinely used in inkjet printing [5, 6]. Given the small particle size, it is possible to treat these using continuum assumptions up to a very small filament diameter, thus, literature studies have made using of power-law viscosity models to understand the drop/jet breakup behavior [8].

However, the development at hand was to involve non-colloidal, i.e. non-Brownian, suspensions where particle size is 1 to 500 micrometers. The literature on drop/jet phenomena in such fluids was relatively new and dealt only with limiting conditions: drop breakoff in the pendant drop case [9-15]. This case involves essentially zero flow rate, in which the drop breaks off under gravity, in the absence of inertial effects (i.e. without imparted velocity) which are necessary for drop-on-demand printing. Such studies identified that these large particles had effects at all stages of breakoff, often with conflicting effects [9]. The initial stage appeared to follow the bulk viscosity of the suspension, but at a filament diameter of 2-10 times the particle diameter, particle-driven effects caused more rapid breakup than expected from either the bulk or interstitial fluid viscosity [9-15]. Thus observation was of a new, particle-driven thinning regime, which is induced at a non-deterministic length scale – it varied with the particle size itself, and different studies report different onsets of the effect [9, 12, 14, 15].

Furthermore, the literature reported that the point of breakoff – the spatiotemporal singularity point – became variable (it is very consistent in Newtonian fluids) and appeared to be mediated by single-particle localization in the thinning filament [13]. Essentially, a stochastic process at a particle-dependent length scale. One must note that these observations occurred in experiments conducted at limiting flow conditions, using of smooth, size-uniform spheres in density-matched fluids.

In the literature there were no demonstrated cases of actual drop printing – in which a transient pressure pulse is used to temporarily form a jet which breaks up. The intended manufacturing process would operate at high extension rates, at which the rheology is even less understood/studied [16-19], and equally importantly, with highly non-ideal particles that exhibit a wide size distribution and markedly non-spherical shapes. Whether this was possible was not clear – only the 2004 paper by Furbank and Morris included information on jetting, but this was for continuous jetting [9]. In contrast, our manufacturing process relies on temporary formation of a jet, then allowing breakoff to proceed from the inertial and capillary forces.

2.1.3 A Path Forward

From the literature, it was unclear if concentrated non-Brownian suspensions could be reproducibly and satisfactory printed using a drop-on-demand method. Moreover, the operating conditions, ink formulations and fluid/particle properties which could lead to the desired result were decidedly unknown. Rationalizing non-Brownian suspension rheology in terms of $\left(\frac{\phi}{\phi_{max}}\right)$ provided a method of generalization across a variety of powders, with ϕ_{max} accounting for the effects of particle size distribution and shape; several conditions must be met for this to hold, such that the flow of interest occurs at steady shear wherein particle inertia effects are absent and also that $\frac{D_{conduit}}{d_{particle}} \cong 10$ or greater [20, 21, 22, 23]. Consequently, as $\left(\frac{\phi}{\phi_{max}}\right)$ provides a collapse of low shear rate rheology data, it seems reasonable to posit that a similar collapse may be observed with printing feasibility for suspensions (in fluids of equivalent or near-equivalent viscosity) of particles with nearly equivalent $\frac{D_{conduit}}{d_{p,max}}$ ratio but variable size/shape distribution; the available literature on drop formation in non-Brownian suspensions supports this hypothesis, though the discrepancy in shear rates is of order 10^6 s^{-1} [9, 12, 14, 15]. Use of $\left(\frac{\phi}{\phi_{max}}\right)$ is useful in that the expected non-Newtonian behavior of the suspensions (our DOD shear rates ~10^3 s^{-1}) precludes the direct use of the (Newtonian) *Ohnesorge* number; the dimensionless grouping of fluid inertia and surface tension, We, retains relevance and as the deformation rate ($\dot{\gamma}$) is implicit contained within this quantity, the extent of non-Newtonian flow behavior will be dependent on the magnitude of *Weber* number.

In the preliminary test, inks were made from density-matched, 38-45 micrometer polyethylene spheres suspended in water, with a small amount of polysorbate-80 surfactant to aid in wetting the particles. Testing of suspension-based inks at particle concentrations from 0.01 to 0.5 demonstrated initial feasibility of drop-on-demand printing, with particle fractions above 0.1 resulting in consistent drop volumes, as shown in the Figure 2.1. From the rheology, it was known that the maximum particle fraction for spheres of uniform size is ≈0.63, thus, it was not very surprising to find that printing attempts at volume fraction 0.55 and 0.60 were fraught with inconsistency, given the expected divergence of the viscosity as $\frac{\phi}{\phi_{max}} \to 1$.
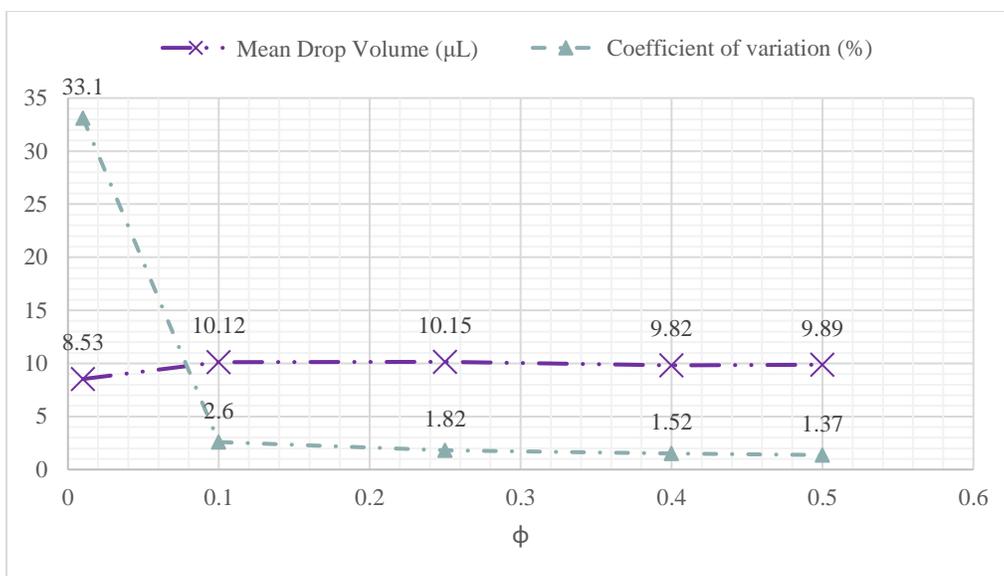
Figure 2.1. Drop printing results for density-matched, 38-45 micrometer polyethylene spheres suspended in water + 0.1-2wt% polysorbate 80.

Such results demonstrated that printing was possible with ideal spheres even at very high dimensionless loadings (up to $0.8 = 0.5/0.63$), and upon presentation of these results to the project managers at Eli Lilly & Co., we were asked to pursue the drop-on-demand printing route with non-Brownian suspensions. The objective was to develop multiple formulation platforms – multiple possible carrier fluids for the particles which were to include both volatile and non-volatile liquids which would be deposited into empty capsules. This would enable formulations to produce doses which are just the API (in which carrier fluid is to be evaporated), in addition to doses that contain the carrier liquid for bioavailability enhancement or controlled-release products. From the perspective of designing a pharmaceutical product, this introduced several constraints: to fill capsules, the choice of fluid must be compatible; the particle properties must be preserved through the process, thus, the fluids must not dissolve the particles.

To this end, two carrier fluids (tricaprylin oil, hexamethyldisiloxane) were selected as suitable based on their physical properties relevant to printing, and as antisolvents in which many pharmaceuticals are insoluble. As the non-volatile carrier, tricaprylin had desirable physical properties based on the dimensional analysis and observations from the initial experiments; it is also an antisolvent for most non-lipid-soluble drug compounds. As the volatile carrier, hexamethyldisiloxane is advantageous in that most drug compounds are not soluble in it; the

physical properties of hexamethyldisiloxane can be adjusted to the needs of printing, as the viscosity of the pure fluid is one half that of water, but can be easily modified by addition of polydimethylsiloxane (i.e. silicone oil), with both the polymer chain length and amount added as means to tune ink properties.

To develop a framework for formulation of future inks, one must consider the effects of particle properties and their interactions with fluid properties and operating conditions, thus, 4 drug powder grades were selected. As discussed in greater detail in Chapter 3, characterization of the powder types, suspensions and printing performance would provide the data which informed the framework. Drops would be printed from suspensions at several particle concentrations in each particle type, with a general goal of using as high a loading as possible. In addition to demonstrating that drops could be printed, doses were to be produced at low (1mg), moderate (10mg) and high (100mg) potency levels, and the critical quality attributes (content & particle properties) in the resultant doses characterized. This was an extensive undertaking which necessitated not only production of hundreds of doses, but also characterization for these. The results of this work are well-summarized by the Journal of Pharmaceutical Sciences article which appears as Chapter 3.

2.1.4   Mapping the Operating Conditions

Coincident with the collection of data on doses, it was also an objective to understand the physics of drop formation in non-Brownian suspensions. This would enable one to provide guidelines for operating conditions and ink formulations for new particle types by combining the rheology and drop physics. Making use of dimensional analysis, the operating conditions could be considered in terms of the *Weber* number and the *Ohnesorge* number [5, 6, 7] – however, as the viscosity of the suspension would require extrapolation from the shear experiments (which, unfortunately, would not even match the shear rates that occur during printing). One considers a scaling based on generalized rheological models, as $f(\phi/\phi_m)$, wherein ϕ_m is dependent on the particle properties, which allows one to both generalize across them, and to explain why it is impossible to print some particle types above certain particle loadings [20-27].

Mapping of the feasible operating conditions requires an in-depth investigation of the drop formation process, which is only possible using high-speed imaging, given that the total elapsed time from initiation of flow to drop breakoff is 50-75 milliseconds. In the actual timeline of the

total research project, the preliminary insights obtained from the high-speed imaging study of drop formation informed dose production efforts leading to the publication in Journal of Pharmaceutical Sciences; these preliminary insights, updated with new information, are presented in Chapter 4, which based on the ESCAPE-27 conference paper.

As an experimental investigation and source of future work, this proved to be a valuable endeavor. In order to conduct the desired high-speed imaging of drop formation, this required design of a suitable setup. Sparing the unnecessary details, this involved locating a suitable high-speed camera (for which Prof. Carl Wassgren must be kindly thanked for lending it), then constructing a physical setup based on the lighting requirements and constraints of the camera. This required learning a small amount about the optical physics related to photography, which later happened to be useful for model purposes (e.g. ESCAPE-28 conference paper).

2.2   Foreword to Chapter 5, 6

Perhaps most substantially, the dataset generated during the high-speed imaging experiments motivated the pursuit of advanced methods for automated image analysis. This would have a profound impact on the remainder of my research work, thus, it is perhaps worthwhile to describe the underlying thought processes. The original intent of the high-speed imaging data collection was for velocimetry – estimating 1-dimensional velocity during drop ejection in order to estimate the Weber number, and hence, to map the feasible operating region. Additionally, temporal evolution of the filament diameter leading to breakoff was to be use to understand the rheology, as the thinning law relates to the fluid (or effective fluid) properties, thus, with the experimental data, one could determine the thinning behavior/regimes, then use this to inform modeling work [5, 6, 7, 9, 14, 15, 28, 29].

From the first, simple analysis of the data, it was apparent that substantially more information was available from the drop formation videos. Much unexpected behavior was immediately observable, in particular, the variability in drop breakup and formation events after detachment, which can have a substantial effect on manufacturing reproducibility. To automate extraction of information on these events and their participant moieties, it was necessary to build upon existing computer vision methods [30-38] in order to develop new algorithms which enabled mapping of object lineages through time and space. This permits the topological representation of the entire video sequences, as shown in Figure 2.2.
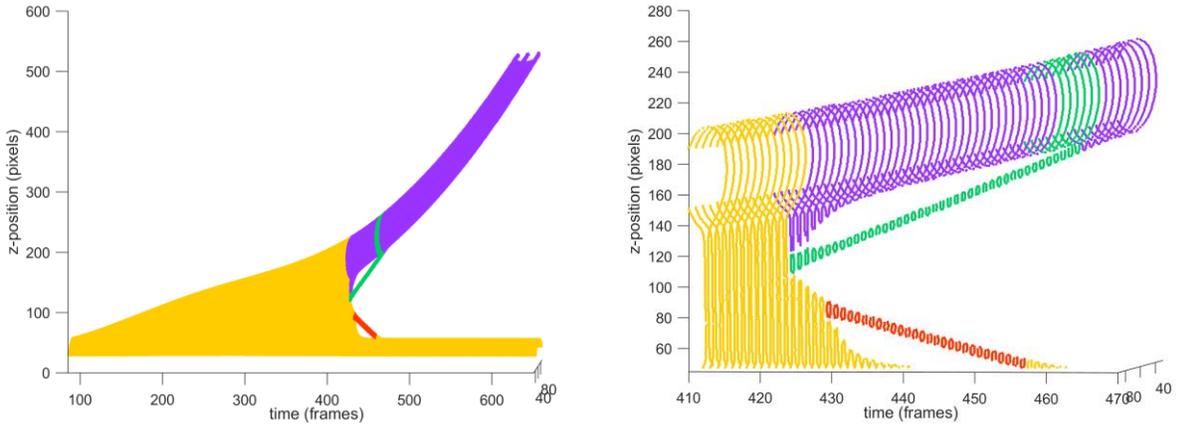
Figure 2.2. Left: Spatiotemporal trajectory of objects through entire video sequence. Right: close-up of space in which: drop break-off events occur, the primary and secondary drop coalescence, and the tertiary drop coalesces with the nozzle-attached fluid meniscus. Coloration: yellow – nozzle-attached fluid filament, purple: primary drop – green, secondary drop, red – tertiary drop. The video of this sequence is found in Figure 5.2, Middle.

While such a representation is interesting in itself, the reader must note that it is automatically constructed from the extracted information, hence, for modeling purposes, the computer vision algorithms provide a detailed history for each object and its interaction partners. The authors envision that these algorithms would be useful to the community studying drop formation phenomena using high-speed imaging techniques, as given such a level of detail one is therefore able to conveniently organize data for simpler procedures such as thinning dynamics. Furthermore, interested researchers can now consider novel methods for validating simulation results against experimental data through comparison of the topological representation of the data and that which is implicitly generated by the simulation. More creative uses are certainly possible.

As a direct use of the object trajectory and event information extracted from the sets of high-speed image sequences, the process monitoring scheme via the online image acquisition system could be improved: the component placement in this online monitoring system could be optimized to accommodate situations in which multiple drops are present, as described in full detail in the ESCAPE-30 chapter. This was motivated by the high frequency of occurrence in printing non-Brownian suspensions, which is clearly observable from the figures in Chapter 5.

2.3   Foreword to Chapter 7, 8

In the development of this additive manufacturing process, a key benefit of the drop-on-demand method is the ability to monitor the total quantity of material deposited to each dose using an online image acquisition system. The work of previous students led to its initial realization, which required selection of the hardware and integration with the LabVIEW-based automation system and real-time process management strategy; this online image acquisition system captures a digital photograph of each drop that is printed by the system. The desired result from the image analysis was the drop volume. To analyze the images, a heuristic thresholding method was developed to binarize the images, and the voxel volume was computed by treating the object as an axisymmetric cylinder. To convert from voxel volume (pixel units) to real volume units (e.g. microliters), a calibration was constructed using a 2-dimensional optical standard presented at a fixed distance from the camera. As demonstrated in their work [39], this approach proved to be very accurate (within 0.33% of actual volume) for fluids without particles.

However, this approach did not work well with drops printed from suspensions. The predicted volumes were sometimes accurate (within 0.5% of actual value), but more often had errors of 5-10%, and sometimes as much as 50% or greater. Understanding the source of variation was not forthcoming, as this could not be dismissed as simple miscalibration. Within a batch of 50 doses, each consisting of a single drop, produced at conditions (μAPAP, Φ=0.2, NEO895 carrier fluid) known to yield stable operation, predictive error was observed to be nearly zero for some dose units, but as much as 10% in others. Interestingly, this varied between over-prediction and under-prediction within the same batch of doses, for which no clear pattern was readily apparent. From other batches of dose units in which each consisted of a single drop, both similar and different observations were made – in the latter, predictive error was sometimes entirely over/under.

Initially, it was hypothesized that these observations might be explained by variability in the drop trajectory, which is known to be affected by the presence of particles – something that was verified by the high-speed imaging experiments. If one considers the optical setup, given the assumption that the image analysis is correct, then this could be explained by the variability in distance from the camera: the apparent size of the object, as recorded in the digital photograph, would be smaller if the object were farther from the camera, and larger if closer to the camera. Essentially, this could be viewed as a source of uncertainty in the measurement, and modeled as such – hence, a Bayesian model based on the historical data is the clear way to model this

uncertainty, and make data-informed predictions for future data. A simple Bayesian model which considers only distance from the camera as a source of uncertainty is presented in the short chapter for the ESCAPE-28 conference paper [40].

However, this approach only seems to work acceptably well for two batches of dose units; a large amount of data was collected to carry out Bayesian modeling on a larger scale involving many batches of dose units, thus, this was initially disappointing. It some cases, it was apparent that the image processing was the source of error, as the predicted volumes were seemingly unrelated to the images and the measured values.

It was not until after completing the computer vision work that it became apparent that it would be possible to enact an automated image analysis method that was robust and capable of consistently extracting information which was sufficiently more detailed. This made it possible to obtain at minimum the most reliable information from the images, but, due to the peculiarities of suspension printing, many sources of uncertainty remain. These include variability in the drop trajectory, variability in drop location when image is captured, and variability in shape and number of drops. Additionally, in a substantial number of images, a portion of a drop is not observable, as it is outside the image boundaries – this can be a major source of uncertainty, as one would like to estimate the missing portion from that which is present.

With the new image analysis method in place, using the online sensor data to make predictions about individual product content and batch content uniformity is once again possible. Use of Bayesian models to achieve this aim provides several advantages: 1) the historical data for all of the doses produced during my research project - 1500 dose units in 41 batches – can be used to inform predictions for future data, thus, one would achieve more robust predictive performance than a fixed calibration, 2) a Bayesian model allows us to quantify our uncertainty, and make risk-based decisions for each end every dose unit; this forms a framework for real-time quality control, and 3) hierarchical modeling enables one to take advantage of the inherent structure in the dataset and therefore make improved predictions for future data obtained at conditions which were not included in the training dataset.

In the context of the dropwise additive manufacturing platform, the Bayesian hierarchical modeling approach provides the means for real-time quality control at any scale of manufacturing – compounding pharmacy, pilot scale, or full-scale.

2.4   References

[1] J.J. Stickel and R.L. Powell, "Fluid Mechanics and Rheology of Dense Suspensions," Annu. Rev. Fluid Mech. **37**, 129 (2005).

[2] M.M. Denn and J.F. Morris, "Rheology of Non-Brownian Suspensions," Annu. Rev. Chem. Biomol. Eng. **5**, 203 (2014).

[3] P. Coussot and C. Ancey, "Rheophysical classification of concentrated suspensions and granular pastes," Phys. Rev. E. **59**, 4445 (1999).

[4] E. Brown and H.M. Jaeger, "Shear thickening in concentrated suspensions: phenomenology, mechanisms and relations to jamming," Rep. Prog. Phys. **77**, 046602 (2014).

[5] O.A. Basaran, H. Gao, and P.P. Bhat, "Nonstandard Inkjets,"Annu. Rev. Fluid Mech. **45**, 85 (2013).

[6] H. Wijshoff, "The dynamics of piezo inkjet printhead operation," Phys. Rep. **491**, 77 (2010).

[7] J. Eggers and E. Villermaux, "Physics of liquid jets," Rep. Prog. Phys. **71**, 036601 (2008).

[8] X. Wang, W.W. Carr, D.G. Bucknall, and J.F. Morris, "Drop-on-demand drop formation of colloidal suspensions," Int. J. Mul. Flow **38**, 17 (2012).

[9] R.J. Furbank and J.F. Morris, "An experimental study of particle effects on drop formation," Phys. Fluid. **16**, 1777 (2004).

[10] R.J. Furbank and J.F. Morris, "Pendant drop thread dynamics of particle-laden liquids," I. J. Mul. Flow **33**, 448 (2007).

[11] T. Bertrand, C. Bonnoit, E. Clement, and A. Lindner, "Dynamics of drop formation in granular suspensions: the role of volume fraction," Gran. Matt. **14**, 169 (2012).

[12] C. Bonnoit, T. Bertrand, E. Clement, and A. Lindner, "Accelerated drop detachment in granular suspesnsions," Phys. Fluid. **24**, 043304 (2012).

[13] A. Lindner, J.E. Fiscina, and C. Wagner, "Single particles accelerate final stages of capillary break-up," Eur. Phys. Lett. **110**, 64002 (2015).

[14] Z. Pan, N. Louvet, Y. Hennequin, H. Kellay, and D. Bonn, "Drop formationin shear-thickeing granular suspensions," Phys. Rev. E. **92**, 052203 (2015).

[15] M.Z. Miskin and H.M. Jaeger, "Droplet formation and scaling in dense suspensions," Proc. Nat. Acad. Sci. **109**, 4389 (2012).

[16] M. Roche, H. Kellay, and H.A. Stone, "Heterogeneity and the Role of Normal Stresses during the Extensional Thinning of Non-Brownian Shear-Thickening Fluids," Phys. Rev. Lett. **107**, 134503 (2011).

[17] P.J. Zimoch, G.H. McKinley, and A.E. Hosoi, "Capillary Breakup of Discontinuously Rate Thickeing Suspensions," Phys. Rev. Lett. **111**, 036001 (2013).

[18] M.I. Smith, R. Besseling, M.E. Cates and V. Bertola, "Dilatancy in the flow and fracture of stretched colloidal suspensions," Nat. Comm. **1**, 114 (2010).

[19] E.E. Bischoff White, M. Chellamuthu, and J.P. Rothstein, "Extensional rheology of a shear-thickening cornstarch and water suspension," Rheol. Acta **49**, 119 (2010).

[20] F. Boyer, E. Guazzelli, and O. Pouliquen, "Unifying Suspension and Granular Rheology," Phys. Rev. Lett. **107**, 188301 (2011).

[21] E. Brown, H. Zhang, N.A. Forman, B.W. Maynor, D.E. Betts, J.M. DeSimone, and H.M. Jaeger, "Shear thickening in densely packed suspensions of spheres and rods confined to few layers," J. Rheo **54**, 1023 (2010).

[22] S. Mueller, E.W. Llewellin, and H.M. Mader, "The rheology of suspensions of solid particles," Proc. R. Soc. A. **466**, 1201 (2010).

[23] P. Moitra, and H.M. Gonnermann, "Effect of crystal shape- and size-modality on magma rheology," Geochem., Geophys., Geosys. **16**, 1 (2015).

[24] J.F. Morris, "A review of microstructure in concentrated suspensions and its implications for rheology and bulk flow," Rheol. Acta **48**, 909 (2009).

[25] E. Bertrand, J. Bibette, and V. Schmitt, "From shear thickening to shear-induced jamming," Phys. Rev. E. **66**, 060401 (2002).

[26] A. Fall, F. Bertrand, G. Ovarlez, and D. Bonn, "Yield Stress and Shear Banding in Granular Suspensions," Phys. Rev. Lett. 178301 (2009).

[27] E. Brown, N.A. Forman, C.S. Orellana, H. Zhang, B.W. Maynor, D.E. Betts, J.M. DeSimone, and H.M. Jaeger, "Generality of shear thickening in dense suspensions," Nat. Mat. **9**, 220 (2010).

[28] C. Clanet and J.C. Lasheras, J. Fluid Mech. **383,** 307-326 (1999).

[29] C. Clasen, P.M. Phillips, L. Palangetic, and Jan Vermant, "Dispensing of Rheologically Complex Fluids: The Map of Misery," AIChE Journal **58**, 3242 (2012).

[30] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679-698.

[31] Elder, J. H., & Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7), 699-716.

[32] Lindeberg, T. (1998). Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 117-156.

[33] Bao, P., Zhang, L., & Wu, X. (2005). Canny edge detection enhancement by scale multiplication. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(9), 1485-1490.

[34] McIlhagga, W. (2011). The Canny edge detector revisited. *International Journal of Computer Vision*, 91(3), 251-261.

[35] Xu, Q., Varadarajan, S., Chakrabarti, C., & Karam, L. J. (2014). A distributed canny edge detector: algorithm and FPGA implementation. *IEEE Transactions on Image Processing*, 23(7), 2944-2960.

[36] Lee, J., Tang, H., & Park, J. (2018). Energy efficient canny edge detector for advanced mobile vision applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(4), 1037-1046.

[37] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.

[38] Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-166.

[39] Hirshfield, L., Içten, E., Giridhar, A., Nagy, Z. K., & Reklaitis, G. V. (2015). Real-time process management strategy for dropwise additive manufacturing of pharmaceutical products. *Journal of Pharmaceutical Innovation*, *10*(2), 140-155.

[40] Radcliffe, A. J., & Reklaitis, G. V. (2018). Bayesian estimation of product attributes from on-line measurements in a dropwise additive manufacturing system. In *Computer Aided Chemical Engineering* (Vol. 43, pp. 1243-1248). Elsevier.

3. DROPWISE ADDITIVE MANUFACTURING OF PHARMACEUTICAL PRODUCTS USING PARTICLE SUSPENSIONS

3.1 Abstract

The principal method of drug delivery is by oral solid doses, the production of which often necessitates multiple post-crystallization unit operations to ensure content uniformity and/or enhance bioavailability. As an alternative to conventional dose production methods, applications of additive manufacturing technologies based on solvent- or melt-based formulations have demonstrated the potential for improvements to process efficiency, flexibility and dosing precision. Here we explore the use of particulate suspensions in a dropwise additive manufacturing process as a method for dosing active ingredients in crystalline form, which may be difficult to achieve via powder processing due to poor flow properties. By employing a fluid-based method, powder flow issues are alleviated and adaptation of the process to new particles/crystals is facilitated by dimensional analysis. In this work, a feasibility study was conducted using four active ingredient powders, each with non-ideal particle properties, and two carrier fluids, in which the active ingredient does not dissolve, to formulate suspensions for dose manufacturing; drug products were analyzed to show reproducibility of dosing and to assess preservation of particle size through the process. Performance across particle types is affected by particle size and shape, and is related through effects on the rheological properties of the formulation.

3.2 Introduction

In the past decade, changes in global/domestic market growth, desire for new drug therapies with faster time to market, pressure to minimize drug cost in healthcare expenditure, and the ever-increasing prevalence of generic medications have prompted revision of operational strategies/technologies in the pharmaceutical industry [1]. Cost reduction and process efficiency improvement in drug manufacturing through implementation of process monitoring technologies,

predictive models and advanced manufacturing technologies have been the focus of much recent work [1-4]. With contributions from academia/industry and regulatory interest, there has also occurred a push to transition from batch to continuous manufacturing to derive the benefits provided by economy of scale and reductions in complexity of process modeling, online measurement and control.[1, 5]

Orally administered solid doses, such as tablets or capsules, constitute approximately 60% of drug formulations [6]. The active component(s) of these drug products are synthesized and purified in fluid-based processes which yield a crystalline product that is solid at ambient temperature. For most active pharmaceutical ingredients (APIs), post-purification steps leading up to a complete tablet or capsule will exclusively involve powder-based technologies. Due in part to the complex nature of powder rheology, API crystals often require additional processing (dry/wet granulation, milling, extrusion) to ensure flow uniformity in feeders, blenders and tableting operations [1, 7]. Thus, for new products with unknown or challenging powder rheology (due to particle shape, size or cohesiveness), process development and formulation design can become time-consuming endeavors.

Due to the prevalence of poorly water-soluble drugs, which constitute an estimated 40% of drugs on the market and in development, particle size reduction is frequently employed to enhance bioavailability[8], though often this necessitates granulation step(s) in order to compensate for poor powder flow characteristics caused by small particle sizes. Thus, a manufacturing process capable of depositing a range of particle sizes (1-500µm) would provide an alternative formulation method for a wide variety of drugs, and could supplement dry powder-based processes for API particles with poor rheological characteristics. Ideally, this would allow for API particles to be drawn directly from crystallization/milling operations, suspended in a fluid in which they are effectively insoluble (to preserve particle size distribution and polymorphic form) and subsequently dispensed to create finished dosage forms after a capsule sealing or tablet coating step.

An alternative to powder processing that has received considerable attention is the adaptation of inkjet printing to pharmaceutical manufacturing [4, 9]. Drop production in modern inkjets is dominated by two methods: continuous inkjet (CIJ), in which a column of liquid is continuously emitted from a nozzle, breaking up into droplets which are then directed onto the substrate, and drop-on-demand (DOD) in which a single drop is produced per trigger of a

piezoelectric, mechanical or thermal actuator [9-11]. The DOD method has been the preferred printing technique for pharmaceutical applications, due to its inherent flexibility, lower raw material requirements and the relative simplicity of such systems compared to CIJ [4, 9]. In addition to being able to handle a variety of fluids, DOD inkjets inherently embody quality-by-design principles and may be outfitted with on-line measurement tools which record the volume of each ejected drop, thereby allowing real-time release of products [12-14]. With the exception of a few works, DOD inkjets have been used to deposit solvent-based API solutions directly onto blank tablets, edible films and other substrates to generate oral dosages; DOD inkjets have also been used in three-dimensional (3D) printing to deposit binders onto which powder is added – both of which have received industrial attention [4, 12-17]. In solvent- or melt-based direct deposition processes, a post-printing heating/cooling cycle is added to crystallize the material; if a control system is implemented this temperature cycle may be used to control the crystallization process on the substrate [13, 15, 17].

With one exception (based solely on a colloidal suspension) [18], suspension-based dropwise pharmaceutical manufacture has yet to be explored, though the benefits appear substantial. Suspension-based dropwise manufacturing directly facilitates control of particle size distribution (PSD) in a drug product by generating the PSD upstream and choosing a carrier fluid that does not solubilize the particles, and permits increases in mass-based production rate (with respect to solvent-based DOD manufacturing). Suspension-based dropwise manufacturing may provide both chemical and physical stability advantages because the API is present in its most stable crystalline form. An additional benefit of employing a fluid-based method is that it may be more easily adapted to a variety of API powders, which can be facilitated through use of dimensional analysis.

In this paper we present the use of the dropwise additive manufacturing method with non-ideal particulate suspensions for production of high dose drug products, and report on the feasibility and performance of the process for operation with several API powders, which exhibit non-ideal particle shapes and size distributions, in select (volatile/non-volatile) carrier fluids. First, a description of the manufacturing process is presented, followed by a discussion of technical and theoretical considerations related to the effects of particles/suspensions on drop formation, drop deposition and formulation selection; insights from the literature are presented in order to provide the reader with a context for the present work. Then, the application of a dropwise additive

manufacturing process using particle suspensions to production of pharmaceutical dosage forms is demonstrated with a variety of particle types/suspension formulations, good reproducibility and preservation of particle size distribution through the process.

3.3 Description of Dropwise Additive Manufacturing Process

The bench-scale dropwise additive manufacturing platform used in this work is shown in Figure 3.1, and the details on the functionality of its process monitoring and control systems are available in references [12, 13, 15, 17]. The drop generator consists of an IVEK Digispense 10 single-channel positive displacement pump with a custom controller (IVEK Corporation, North Springfield, Vermont), which enables operation in the drop-on-demand mode; with appropriate choice of fluid properties, pump/controller parameters and nozzle diameter, this system produces individual fluid drops with volumes in the microliter range, at a rate of one drop per pump actuation event. During operation, the pump draws fluid from a sealed reservoir and subsequently discharges it through a single nozzle, at which point a drop is formed by the breakup of the resultant liquid jet into one (or more) drop(s), as determined by the interaction of capillary, gravitational, inertial and viscous forces. After ejection at the nozzle, the drop accelerates under the effect of gravity, following a trajectory which leads to a prescribed location on the substrate. Pharmaceutical doses are produced by this system through the sequential deposition of fluid drops on appropriate substrates, wherein final product content is determined by the number, volume and composition of the deposited drops.
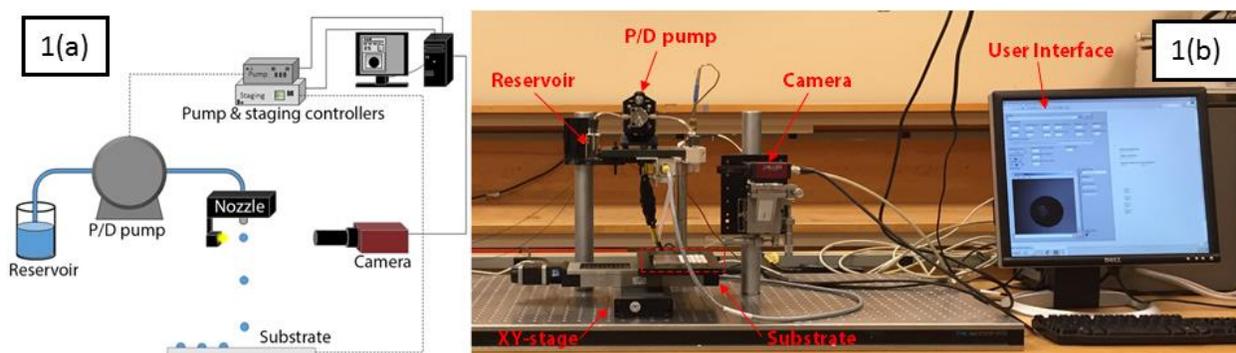
Figure 3.1. System diagram, 1(a), and photograph of system, 1(b); pump/staging controllers not shown in photograph.

The performance metrics of this system are controlled by the jet breakup/drop formation process: uniformity of products is affected by variance in drop volume, content and trajectory; the dynamics of drop formation determine the frequency at which drops can be ejected from the nozzle, and thus, the volumetric throughput, defined as the product of drop volume and drop ejection frequency. In subsequent sections, salient features from the literature on jet breakup/drop formation and suspension rheology are presented in order to identify the fluid properties, particle characteristics and system conditions that are relevant to performance. First, we discuss the dimensionless parameters related to drop formation, and the phase diagrams which may be used to estimate suitable conditions for drop printing; this leads to a discussion of the effect of particles on drop formation, which is preceded by a brief overview of the effect of suspension properties (particle size/shape, particle loading, carrier fluid) on rheological properties.

3.3.1 Drop Formation

Due in part to the many applications of jets/drops in everyday life, the physical mechanisms of drop formation have been extensively investigated through theoretical, numerical and experimental studies.[9-11, 19] These works have identified the regimes (dripping, jetting, spraying/atomization) of drop formation and transitions between them, which, with the use of dimensional analysis, may be used to develop criteria for drop printing applications based on fluid properties and process conditions.[9,10] The dimensionless parameters are used to characterize the contributions of other forces relative to surface tension force, which, through the Laplace pressure, is the primary driver of jet breakup/drop formation; for Newtonian fluids, the relevant dimensionless groupings to consider in drop formation processes are the Ohnesorge ($Oh \equiv \mu /\sqrt{\rho R \gamma}$) number, the gravitational Bond ($Bo \equiv \rho g R^2/\gamma$) number, and the Weber ($We \equiv \rho U^2 R/\gamma$) number, which measure, respectively, the importance of fluid viscosity, gravity, and velocity (of the emerging fluid) at the nozzle [9-11, 19].

Experimental and numerical studies have identified the range of Ohnesorge numbers, $0.1 < Oh < 1$, as suitable for drop-on-demand drop formation processes due to the reduction in undesirable effects related to insufficient or excess fluid viscosity; for $Oh < 0.1$, formation of secondary ("satellite") drops negatively affects product consistency, and for $Oh > 1$, fluid viscosity may result in extended filament break-off lengths, or prevent drop ejection altogether [9, 10, 19, 20]. For most inkjet printers, the Bond number is nearly zero due to the small nozzle sizes

(diameter < 100μm) employed by such processes, thus, effects of gravity on the liquid interface at the nozzle are negligible [8, 19]. However, for *Bo* ≈ *1*, surface tension is unable to maintain the interface, and gravity-induced free dripping occurs; as this would compromise operation in DOD mode, we limit nozzle inner diameters to 1mm (*Bo* < 0.5). The Weber number describes the transition between dripping and jetting, and, when combined with the Ohnesorge and Bond numbers, may be used to estimate operating conditions (e.g. flow velocity) based on available phase diagrams or correlations [9, 10, 11, 21].

The use of particulate suspensions in this study introduces additional factors which must be considered for the drop-on-demand manufacturing process – non-Newtonian flow behavior(s) inherent to suspensions and changes in the drop formation dynamics due to the presence of particles of non-negligible size. In the next section, the rheology of particle suspensions is briefly considered in order to describe the contributions of particle and fluid properties to formulation of the suspensions used in this study; then, the effect of particles on drop formation are discussed.

3.3.2 Suspension Rheology

Experimental, theoretical and numerical analyses indicate that particulate suspensions exhibit several non-Newtonian flow behaviors: [1]shear stress/rate dependence (shear-thinning/thickening, yield stress), [2]segregation due to particle sedimentation/creaming (when particle and fluid not density-matched), [3]multiple types of migration/segregation resulting from fluid deformation (stress or rate induced), and [4]time-dependent material properties (thixotropy). As several aspects of the aforementioned non-Newtonian flow behaviors present problems for the drop-on-demand manufacturing process used in this study, it would be desirable to minimize their effects through adjustment of operating conditions and modification of suspension formulation. To inform such efforts, we briefly review the variables that affect suspension viscosity and flow behavior.

The viscosity, η, of an idealized suspension depends on the particle properties (radius r_p, density ρ_p, number concentration n_p), fluid properties (interstitial fluid viscosity η_0, density ρ_0) and system conditions (thermal energy k_BT, shear rate $\dot{\gamma}$, buoyancy g) such that $\eta = f\left(r_p, \rho_p, n_p, \eta_0, \rho_0, k_BT, \dot{\gamma}, g\right)$ [21-26]. Using dimensional analysis, this variable set may be reduced to $\eta_r = \eta/\eta_0 = f(\phi, Re, St, Pe, Gr)$, in which $\phi = n_p \frac{4\pi}{3} r_p^3$, $Re = \rho_0 r_p^2 \dot{\gamma}/\eta_0$, $St =$

$\rho_p r_p^2 \dot{\gamma}/\eta_0$, $Pe = 6\pi r_p{}^3 \eta_0 \dot{\gamma}/k_b T$, and $Gr = 2r_p^2(\rho_p - \rho_0)g/9\eta_0\langle U\rangle$, identifying the significance of the particle volume fraction (ϕ), particle size (r_p), interstitial fluid viscosity (η_0), deformation rate ($\dot{\gamma}$) and buoyancy effects (ρ_p-ρ_0) [23-25]. An important distinction may be made between Brownian (colloidal) and non-Brownian (non-colloidal) suspensions, which are differentiated by the relative influences of convection (shear) and Brownian (thermal) diffusion on particle motion. This ratio of forces is quantified by the (particle) Péclet number, Pe, large values of which indicate that particles in the suspension are not appreciably affected by Brownian motion, hence rheological effects due to colloidal phenomena may be assumed to be negligible [24, 25]. Typically, suspensions with particle diameters greater than 1μm are taken to be "non-Brownian" (non-colloidal); it is expected that suspensions of micronized API powders will be within this category, based on the range of particle sizes involved.

In real cases, viscosity predictions are difficult for non-Brownian suspensions due to the complex interaction of particle/fluid properties and flow conditions [25, 26]. However, useful insights into the effect of powder properties on viscosity can be obtained through consideration of the models for ideal cases, in which the suspension consists of size-uniform, neutrally buoyant spheres dispersed in a Newtonian fluid and forces acting on the particles are purely hydrodynamic. Under such conditions, in the absence of thermal effects ($Pe \to \infty$), gravitational effects ($Gr \to 0$) and shear rate dependence ($Re, St \to 0$), viscosity depends only on particle volume fraction, which in concentrated suspensions follows an exponential or power law relation.[23-27] One such model is the Krieger-Dougherty equation, given in Eq.(1).:

$$\eta_r = \left(1 - \frac{\phi}{\phi_m}\right)^{-[\eta]\phi_m} \tag{1}$$

in which the empirical constants, $[\eta]$ and ϕ_m, are used to fit a set of rheological data evaluated at a single steady shear rate; $[\eta]$ is the intrinsic viscosity, dependent on particle shape, and ϕ_m is the maximum particle volume fraction, dependent on particle size distribution and shape [28]. More general models (with origins in granular rheology) that permit description of non-ideal, shape-heterogeneous particulate suspensions avoid dependence on the shape-dependent intrinsic viscosity parameter, $[\eta]$, preferring the form, $\eta_r = f(\phi/\phi_m)$, wherein ϕ_m accounts for the effect of both particle shape and size distribution [28-32]. In terms of maximum particle packing fraction, increased particle size polydispersity corresponds to increased ϕ_m (provided that finite-size effects are negligible in the geometry of interest), as the smaller particles can fill the interstitial

voids between large particles, leading to increased solid fraction for a given control volume [31-33]. Deviations from an aspect ratio (AR) of 1 (spheres) correspond to a decrease in ϕ_m in suspensions of both mono-/poly-disperse particles; this results primarily from hydrodynamic effects that arise due to the rotation/re-orientation of the particle during fluid flow [28, 34]. However, geometric constraints are also introduced by existence of an aspect ratio distribution, directly decreasing the maximum solid packing fraction (corresponding to increased void fraction) of the powder in the dry state [35]. The use of a model in the form of $\eta_r = f(\phi/\phi_m)$ permits generalization across particle types, and also indicates the source of limitation on particle loading in a suspension, ϕ_m. Thus, by selective tuning of the particle properties to increase ϕ_m, actual particle loadings (ϕ) may be increased; in practice, this is achieved by forming suspensions from particles with a wide size distribution and approximately spherical shape (low aspect ratio) [28, 31-33].

3.3.3   Effect of Particles on Drop Formation

The use of particle suspensions in drop-on-demand printing has the potential to introduce disturbances to the process due to non-Newtonian flow effects, and through the presence of particles in the liquid filament that forms during drop break-off, which has a substantial effect in non-Brownian suspensions due to the larger particle sizes involved. The literature on drop formation from suspensions may be separated in a manner similar to the rheology: colloidal and non-colloidal. For non-Brownian suspensions, experimental investigation of drop formation dynamics under limiting conditions ($We \rightarrow 0$) has sought to elucidate the underlying physical mechanisms, and effects due to particle size/volume fraction; however, studies of drop-on-demand drop formation have not yet been reported. In comparison, considerably more studies have been published on drop formation and drop-on-demand manufacture employing colloidal suspensions – this is due in part to constraints on particle sizes imposed by conventional piezoelectric/thermal-actuated drop generators, and also to the interests of the ceramics industry, for whom nanometer-sized particles are often desirable [10, 19].

Although the suspensions used in this study are non-colloidal due to their particle size range, insights from the literature on DOD printing of concentrated suspensions of colloidal particles may be useful in interpreting effects related to bulk rheological properties and ink formulation. Notably, the presence of colloidal particles affects the feasibility of inkjet printing

processes, causing disturbances such as clogging and unstable jetting phenomena (inconsistent breakup location, filament swaying/non-straight trajectory), which necessitate reformulation of the ink should a suitable set of process parameters be inaccessible [36-40]. The effects of particle size, particle volume fraction and ink formulation (carrier fluid, surfactants, polymers) on performance have been considered in terms of the rheological behavior, which to is known to be shear rate/stress dependent. Satisfactory performance was observed with ink formulations that were approximately Newtonian, and with adjustment of operating conditions, highly reproducible DOD printing was achieved, though the feasible region of drop ejection frequencies was substantially smaller than for pure fluids [36-38]. For the system used in this study, presence of colloidal particles in an API slurry would not problematic for operation, but are not investigated here due to the interest in delivering larger (non-Brownian) particles which are with the range of typical APIs after crystallization and/or milling.

Regarding the effect of non-Brownian particles on drop formation, experimental studies have been published in the literature on the break-off of pendant drops ($We \rightarrow 0$), and in one case [42], jetting behavior of non-Brownian suspensions with particle sizes between 1-500μm and particle volume fractions up to 0.55 [41-49]. In the pendant drop experiments focused on ideal non-Brownian suspensions, similar phenomenological effects are observed: particles decrease the drop break-off time (compared to pure liquid of equivalent viscosity) [41-49], suspensions undergo break-up in distinct stages (the incipience and duration of which depend on particle size and volume fraction) [41-46, 49] and satellite drop formation may be suppressed by presence of particles [41, 46]. Significantly, Furbank and Morris report that for a given set of experimental conditions, repeatability (liquid intact length, drop volume) is achieved with some difficulty – this is attributed to the random localization of particles within the necking liquid bridge, which varies somewhat for each pendant drop [41]. The phase diagram developed by Pan *et al.*[46] divided the behavior of suspensions into three categories based on the different breakup regimes observed in non-Brownian suspensions: Newtonian fluids, non-Brownian suspensions and shear-thickened fluids; the authors made these classifications by analogy with the radial decay velocities observed in Newtonian fluids [11], dense suspensions [47], and shear-thickening fluids [48, 49]. Beyond pendant drop dynamics, Furbank and Morris made several observations on jetting behavior in non-Brownian suspensions, noting that particles smooth the transition from dripping to jetting (an abrupt phenomenon in Newtonian fluids) to such an extent that a clear transition cannot be

identified and that lower particle concentrations accelerate breakup and higher particle concentrations appear to inhibit radial necking [41].

The observations of phenomena from pendant drop experiments and consideration of the breakup regimes imply that for drop-on-demand drop formation, process operation with non-Brownian suspensions below a critical particle loading/size may yield approximately Newtonian flow behavior. However, as the available literature is based on ideal suspensions (uniform size spheres, neutrally buoyant particles), differences are expected to occur when working with suspensions of real particles which have size/shape distributions and may not be density matched to the fluid. From consideration of the suspension rheology literature [28, 31-34], it is expected that deviation of particle shape and size distribution from monodisperse spheres will result in lower particle loadings (ϕ) than could be achieved in the experiments of Pan $et\ al.$, for example, due to confinement effects; however, when ideal/non-ideal results are compared in terms of ϕ/ϕ_m, similar trends should be observable. The other substantial difference between the pendant drop experiments and the drop-on-demand drop formation experiments presented in this paper is related to the flow conditions: in the pendant drop experiments, flow rates were very low ($We \rightarrow 0$); in drop-on-demand drop formation, flow rate must be sufficiently high as to transiently produce a jet from the nozzle – this is advantageous from the perspectives of production rate and controllability. The increased flow rates may produce undesirable effects in the drop printing process through shear rate/stress-dependent flow behavior in the tubing upstream of/at the nozzle, or through flow-driven migration of particles in the liquid phase during transit through the system; however, from the suspension rheology literature, it can be observed that such non-Newtonian flow effects increase in severity as shear rate ($\dot{\gamma}$) and particle loading (ϕ) increase [24-26] – therefore, these effects can be reduced by controlling the particle loading, ϕ, to be below critical values for a given set of operating conditions.

3.4   Materials and Methods

3.4.1   Materials

Particles

Pharmaceutical powders were selected such that a variety of particle properties (aspect ratio distribution, particle size distribution, crystal habit) were encompassed; to this end, three

APIs were selected, from which four powder grades were realized. Micronized acetaminophen purchased from Mallinckrodt Pharmaceuticals (Item Code 0422: *Acetaminophen USP/Paracetamol Ph Eur Micronized*) and semi-fine acetaminophen purchased from Mallinckrodt Pharmaceuticals (Item Code 0081: *Acetaminophen USP/Paracetamol Ph Eur Semi-Fine Powder*) and were kindly provided by Prof. Carl Wassgren (Purdue University, West Lafayette, IN). Mefenamic acid (≥98%) and phenylbutazone (≥98%) were purchased from Tokyo Chemical Industry America (Philadelphia, PA). Due to the presence of large particles (size > 1mm) in the as-received samples of mefenamic acid and phenylbutazone, secondary samples were produced from the bulk powder by sieving. This separation was achieved using a Ro-Tap® E Sieve Shaker (W.S. Tyler Industrial Group, Mentor, OH) operated in the "Fines Analysis" mode, with a mesh progression (in µm) of: 700, 500, 355, 250, 180; the contents of the pan were kept as the (sieved) samples for mefenamic acid and phenylbutazone.

Liquids

Carrier fluids were selected to preserve particle properties in the final product; that is, liquids in which the APIs were very poorly soluble or insoluble. Further constraints on choice of carriers arose due to the need for the liquid to be compatible with gelatin capsules, and in the case of a nonvolatile carrier, to be an excipient approved for use in oral pharmaceuticals. The volatile solvent, hexamethyldisiloxane (HMDSO) – viscosity = 0.00049 Pa·s, density = 0.760 g/cm^3, surface tension = 15.3 dyne/cm, was selected as the carrier liquid for production of dry dosage forms (which, after evaporation of the solvent, would contain only API powder); pharmaceutical grade hexamethyldisiloxane (99.9%) as *Dow Corning Q7-9180 Silicone Fluid 0.65cSt* was purchased from Univar (Indianapolis, IN). A rheological modifier for the hexamethyldisiloxane, hydrophobic colloidal silicon dioxide (AEROSIL R 972 Pharma, abbreviated as *R 972*), was kindly provided by Evonik Industries (Parsippany, NJ). For production of dosage forms using a non-volatile carrier liquid, pharmaceutical grade triglyceride oil (≥95% tricaprylin, 4% tricaprin) – NEOBEE® 895 (viscosity = 0.021 Pa·s, density = 0.949 g/cm^3, surface tension = 29.9 dyne/cm), was kindly provided by Stepan Company (Millsdale, IL).

3.4.2 Particle Characterization

To quantify effects resulting from particle properties, each powder's particle size distribution, true density and bulk density were determined. To ensure repeatable physical property testing, samples from each bulk powder were obtained following the protocol set forth in *ASTM Standard B215 - 15 (7.1.2 Sampling of a Stationary Powder Using a Scoop Sampler)*; each test was performed in triplicate and mean values are used where appropriate – specifications of each testing procedure are given below.

Particle Size, Aspect Ratio Distribution

Characterization of the powders' particle size was performed using two-dimensional image analysis, as facilitated with a Malvern Morphologi G3-ID (Malvern Instruments Ltd, Worcestershire, UK). High-throughput image analysis, as opposed to laser diffraction, was chosen for this purpose as it generates a particle size distribution that retains information on the two-dimensional shape and aspect ratio distribution of the particles. Particle size analysis was performed in the same manner for each of the four powders, with changes only to the optical magnification and exposure settings to accommodate the apparent range in particle sizes. Initial measurements were performed in triplicate using $9mm^3$ of powder, dispersed on an optical glass stage by means of the Morphologi G3-ID's air-actuated dispersion unit. The injection pressure (4 bar) and time interval (10ms) were fixed for all powders, and chamber dwell time (120s) was sufficiently long to permit sedimentation of all particles greater than 0.5μm. Multiple microscope magnifications were used to determine the appropriate magnification level for analysis of the respective powder samples; the optical magnification for micronized acetaminophen was 20x, and for all other powders was 10x. Analysis at 20x magnification afforded a dynamic range of 0.5-100μm, for which images were comprised of 1936x2584 pixels (0.14 μm/pixel); objects smaller than 1.4μm (10 pixel lengths) were discarded. At 10x magnification the expected range was 2-200μm, and images were comprised of 1936x2584 pixels (0.28 μm/pixel); objects smaller than 2.8μm were discarded from the analysis.

In the initial measurements which used the dry dispersion method, some aggregates were observed in the powder samples, potentially resulting from of the cohesive nature of the materials or limitations of the dry dispersion method. As the inter-particle cohesion would be considerably

reduced when the powders were suspended in a liquid, and the presence of such aggregates following the suspension preparation steps (mixing, sonication) is less likely, a wet dispersion method was used to provide an alternate measurement of particle size/shape for the powders prior to printing. In the wet dispersion method, the powder samples were prepared as described (below) in the suspension preparation section, in the same manner as was used for drop printing experiments, using triglyceride oil (NEOBEE 895) as the suspending fluid and particle volume loadings of $\phi = 0.2$. The powder samples (particles suspended in oil) were presented to the Morphologi G3-ID on a different optical stage, held between a microscope slide and coverslip; respective to particle type, optical magnification and dynamic range were the same as in the dry dispersion method. The microscope slides were prepared by adding one drop of suspension and four drops of oil to the slide, which resulted in a particle volume fraction of approximately 0.04, after which, a coverslip was added and 4 hours was allowed to elapse prior to measurement, in order that the particles settle completely. This procedure enabled measurement of the particle size and shape properties in a state which should have been representative of the suspensions as they entered the process. The same wet dispersion method was applied to the printed products (100mg doses – filled into gelatin capsules), with the only difference being that the source of the drop of suspension was a dose produced by the process.

Particle True Density

The true density of the particles, ρ_{true}, was needed to formulate a suspension at a given calculated particle volume fraction and to estimate the maximum particle volume fraction. Under the assumption that the API powders have negligible inaccessible internal voids, the skeletal density will be equal to the true density of the particles. Using an XS105 DualRange balance (Mettler-Toledo, Switzerland) to measure sample mass, an AccuPyc II 1340 Gas Displacement Pycnometry System (Micromeritics Instrument Corporation, Norcross, GA) was used to determine the skeletal density of each powder sample. Helium gas was used as the measurement gas, and the following instrument specifications were used with each powder sample: 10 purges (purge fill pressure 19.5 psig), 10 cycles (cycle fill pressure 19.5 psig), pressure equilibration rate 0.0050psig/min, 3.5cm^3 cell volume.

Powder Bulk Density

Bulk density of the powder samples is taken to be the mean value of three average bulk density measurements taken using a Micromeritics GeoPyc 1360 (Micromeritics Instrument Corporation, Norcross, GA) with Transverse Axial Pressure (T.A.P.) density option. Analysis was performed with this instrument using two preparation cycles and twenty measurement cycles, with the 50.8mm diameter cylinder at 20N consolidation force (10kPa compaction pressure) – an XS105 DualRange balance was used to determine the sample mass. Sufficient sample quantities were used to fill approximately 25% of the sample chamber volume.

3.4.3   Dropwise Dose Production

Suspensions Preparation, Characterization

To make a suspension of powder in carrier fluid at a particle volume fraction of interest, an appropriate mass of powder, sampled in accordance with *ASTM Standard B215 – 15*, was placed in a 20mL borosilicate glass scintillation vial to which the corresponding mass of liquid was later added – samples were made in such quantities as to not exceed 75% of the vial volume; powder and liquid masses were quantified on an AL-201S (American Weigh Scales, Cumming, GA), which had readability of 0.1mg, minimum weight of 200mg and meets the requirements for repeatability/accuracy. The vial was subsequently sealed, homogenized for 1 minute with an MX-S vortex mixer (Scilogex, Berlin, CT), placed in a sonication bath (P30H, Elmasonic, NJ) at 120W ultrasonic power/37kHz for 5 minutes, and then allowed to rest for 24 hours – this treatment was repeated prior to use in printing or measurements, with homogenization and sonication bath times of 1 minute each.

The effect of suspension settling in the reservoir was investigated using time-lapse photography of prepared suspensions, wherein the rate of sedimentation was estimated from the rate of decrease of bed height in a quiescent suspension [50]. For suspensions of semi-fine acetaminophen ($\phi = 0.2$), and mefenamic acid ($\phi = 0.2$) in triglyceride oil, time points from the sedimentation rate experiments are presented in Figure 3.2. From these images, it can be observed that the sedimentation rate was slow, and thus, settling in the reservoir was not appreciable as drop printing experiments were conducted within 20 minutes of priming the pump.

Suspension Printing and Dose Manufacturing

In production of doses, a freshly mixed reservoir was attached to the pump stage using 15cm of PTFE (1/16th inch inner diameter) tube and a rubber stopper (to prevent solvent evaporation). Once the reservoir was connected, the pump was primed at low speed (RPM), and 3mL of fluid was permitted to flow into a waste jar; flow was then stopped and the pump was switched to drop-on-demand mode. Prior to dosage production, 20 drops were ejected at the process conditions (RPM, stroke length) of interest in order to ensure consistent results. At the end of each experiment, in cases of suspected particle-liquid segregation, or after the allotted time had passed, the reservoir was discarded and the system flushed with: [1] pure carrier liquid (20mL) and then [2] 95vol% ethyl alcohol/5vol% isopropyl alcohol (20mL).
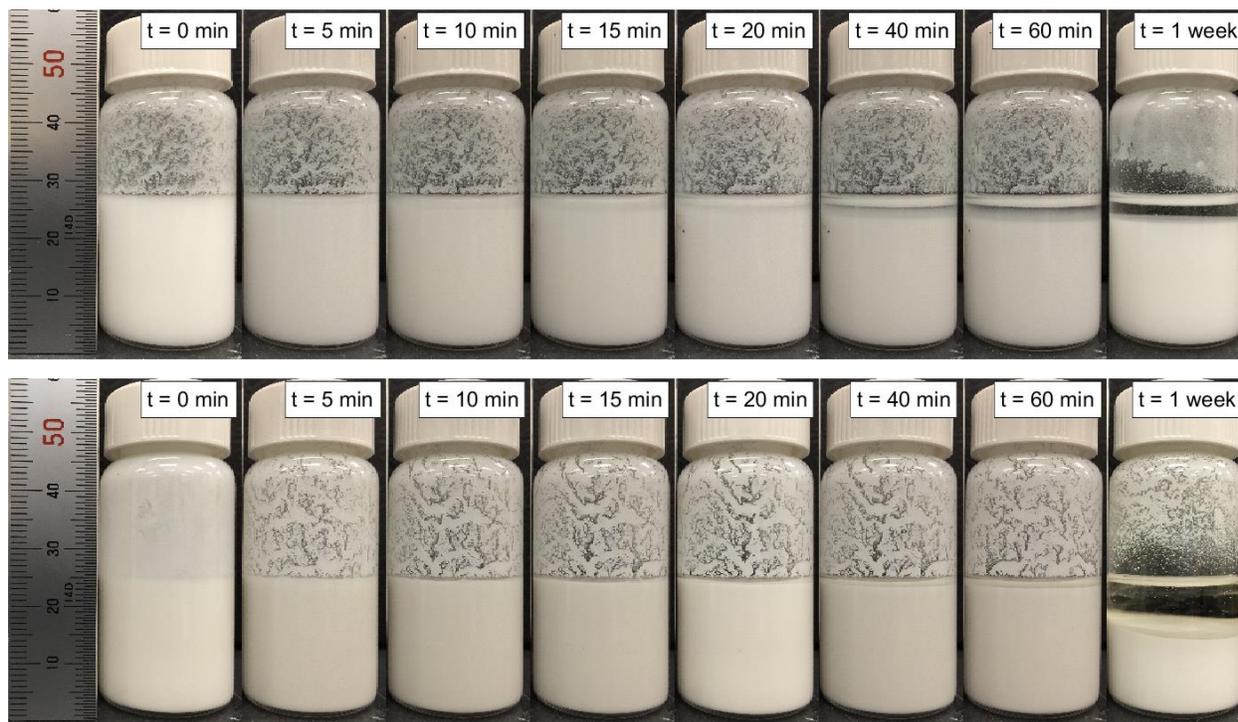
Figure 3.2. Sedimentation of suspended particles in NEOBEE 895: (top) semi-fine acetaminophen, $\Phi = 0.2$; (bottom) mefenamic acid, $\Phi = 0.25$.

3.4.4 Analysis of Capsule Products

Uniformity of Dosage Units by Weight Variation

The slurry mass deposited in each capsule (dose) was quantified gravimetrically using an XS105 DualRange balance in (Fine mode: readability 0.01 mg, minimum sample weight 20 mg; gelatin capsule weight: empty, ≈100 mg; filled dose, 300-600mg). For each dose, the slurry mass was determined by subtracting the mass of the empty gelatin capsule from the total mass of the filled capsule. To preserve the identity of each capsule and its respective contents, the following procedure was used: prior to filling with drops, each empty gelatin capsule (body and cap) was weighed and then stored in its own 1.5mL centrifuge tube which had a unique numerical identifier. In the dropwise addition process, each capsule body was placed on a specified location in a numbered grid and each capsule top was stored in its corresponding 1.5mL centrifuge tube. After addition of drops to each empty gelatin capsule body, its respective cap was snapped on and the total mass of the finished product was quantified on the balance; the slurry-filled capsule was then returned to its respective 1.5mL centrifuge tube for storage.

Dosage (API) mass was determined from the mass of slurry deposited into each capsule and the mass fraction of API in the prepared suspension. For dosage units containing an API amount ≥ 25mg and with API weight fraction ≥25%, the *United States Pharmacopeia* (*USP*) Chapter <905> on Uniformity of Dosage Units states that Weight Variation testing may be applied to hard capsules – specifically, the capsule contents – in order to calculate acceptance values.[51] As the doses produced in this work have a target API amount of 100mg, and most of the suspension formulations have an API weight fraction ≥25%, we employed the Weight Variation method described in *USP* <905> for assessment of dose uniformity. The assay used to determine drug content was the gravimetric: mass fraction of API was determined from the mass of drug substance and mass of carrier liquid, quantified on the AL-201S balance, added to the 20mL vial during suspension preparation.

Disintegration Testing

Rapid disintegration of capsule products is of substantial importance, and provides a test as to whether the dosage forms produced are capable of delivering API powder, without regard to an individual API's dissolution kinetics. Disintegration testing was performed by introduction of a

capsule into a 500mL glass reactor; disintegration was judged to be complete when the capsule had formed a soft mass that did not visibly contain undissolved gelatin fragments. In the batch reactor, temperature was set at 37±0.5°C by means of a heat transfer jacket supplied with fluid by an external heat exchanger. The hydrodynamic conditions in the vessel were controlled by means of an IKA® Eurostar 60 digital mixer (IKA® Works, Inc., Wilmington, NC) which, mounted overhead, extended an impeller with three paddles into the batch reactor, with 2cm clearance from the vessel floor. Mixing conditions, as determined by the impeller RPM, were approximated using information provided in *USP* <701> on Disintegration, by matching the tip velocity of the impeller to that of a basket assembly as in ascends/descends – this corresponded to a set point of 180RPM. Testing of each capsule was performed separately using 500mL of distilled water, with temperature monitored continuously using a probe mounted through the vessel top; after each test, the vessel was flushed with 3000mL of distilled water.

3.5   Results and Discussion

3.5.1   Particle, Powder Properties

The characterization methods confirmed that the API powders used in this work exhibited markedly different crystal shapes, size distributions and consequently, bulk densities. Particle size, true density, bulk density, and particle shape are given in Table 3.1, with corresponding images of the particles presented in Figure 3.3.

Table 3.1. Particle, powder properties as measured by the procedures described in the methods section; particle size parameters (from both dry and wet dispersion methods) are based on the volume-weighted distributions, and aspect ratio (AR) is based on the number-weighted distribution. Particle diameter is defined by the projected area diameter, $d_{particle} = \sqrt{4A_{particle}/\pi}$.

	Micronized acetaminophen (μ-APAP)		Semi-fine acetaminophen (sf-APAP)		Mefenamic acid (MA)		Phenylbutazone (PB)	
True density (g/cm³)	1.329		1.327		1.271		1.203	
Bulk density (g/cm³)	0.352		0.502		0.696		0.505	
Bulk Solid Fraction	26.7%		37.8%		54.8%		42.4%	
Dispersion Method	Dry	Wet	Dry	Wet	Dry	Wet	Dry	Wet
D[3,2] (μm)	24.2	30.5	61.2	56.7	57.2	81.0	64.2	85.9
D[4,3] (μm)	31.5	40.4	119.6	90.9	112.5	118.0	94.9	103.4
d_{10} (μm)	14.3	18.3	28.64	29.8	27.5	43.6	34.6	59.3
d_{50} (μm)	28.2	37.3	94.0	79.1	108.3	101.8	86.8	97.9
d_{90} (μm)	50.1	63.5	243.5	178.7	196.4	211.5	163.9	161.3
Mean AR	0.632	0.601	0.543	0.494	0.674	0.701	0.498	0.316
Stdev AR	0.168	0.18	0.194	0.211	0.154	0.148	0.211	0.217
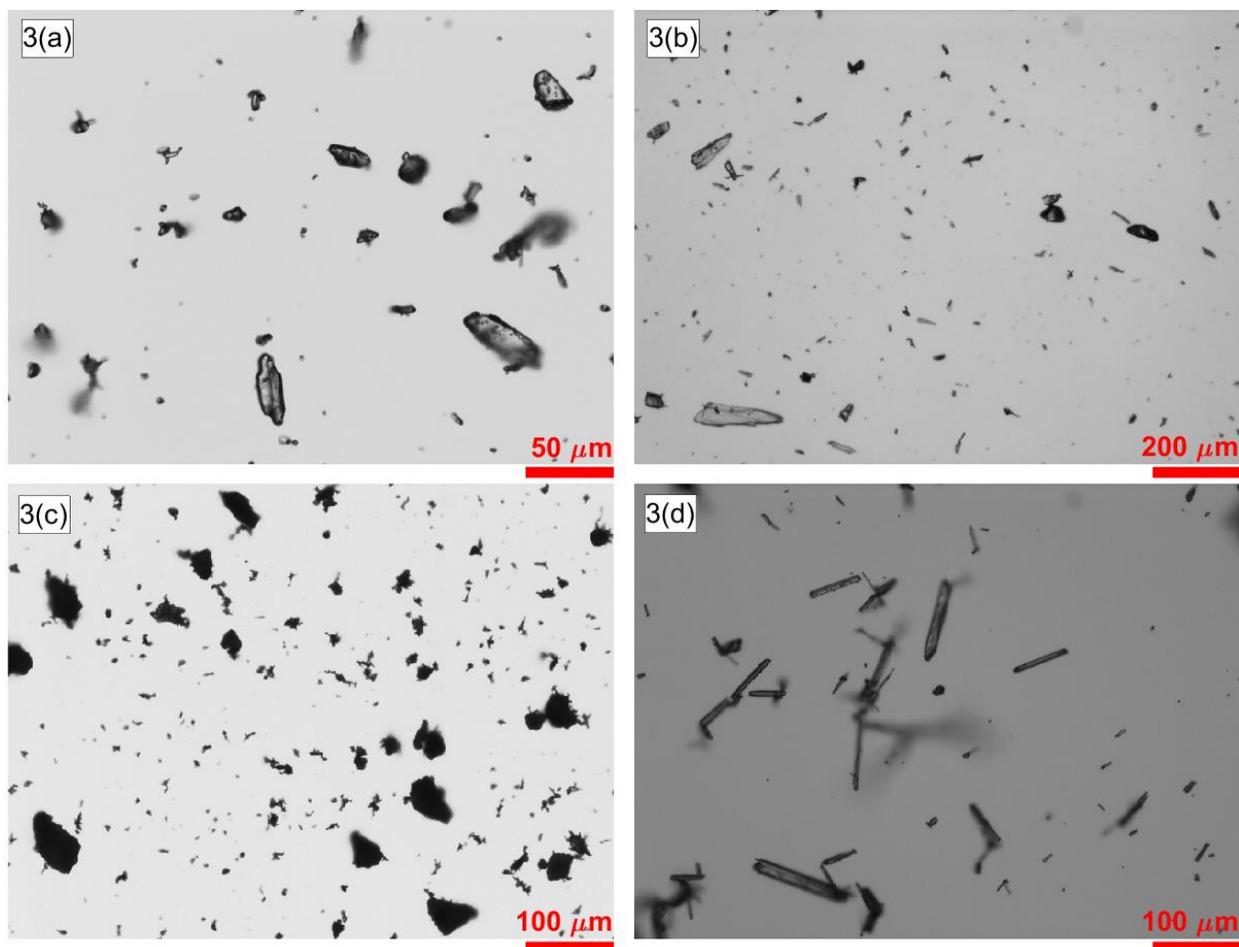
Figure 3.3. Sample images obtained on Malvern Morphologi G3-ID; for each powder, the number of recorded particles in each sample ranged from 250,000-1,500,000 for the dry dispersion method, and 10,000-200,000 for the wet dispersion method. 3(a)-(d) correspond, respectively, to micronized acetaminophen, semi-fine acetaminophen, mefenamic acid and phenylbutazone.

In Table 3.1, the bulk solid fraction, or maximum solid packing fraction, for each API is determined from the ratio of the powder bulk density to the particle true density. For generalized models of shear viscosity in non-Brownian suspensions, bulk solid fraction has been shown to be an accurate predictor of ϕ_m for mixtures of particles with different aspect ratios, provided that the particle size is not appreciably close to the flow dimensions[32, 34]; this exploits the fact that for particles with negligible internal voids, the powder bulk density reflects the effects of both particle size and aspect ratio distributions on maximum achievable solids fraction within a control volume, provided that particle cohesion is negligible. As the alternative to this method is a body of shear rheology experiments, use of bulk solid fraction to estimate ϕ_m may be preferable for

pharmaceutical applications due to the non-destructive nature of bulk/true density tests and the relative ease of performing these measurements.

3.5.2 Drug Product Manufacture Using Suspension Formulations

In this manufacturing process, dosages were produced by depositing drops into Size 0 gelatin capsules; the API mass in each dose (m_{API}) was determined from the product of the deposited slurry mass (m_{slurry}) and the API weight fraction in the suspension, such that $m_{API} = m_{slurry} \times (wt.frac.API)$. The deposited slurry mass is the sum of the individual masses of drops (m_{drop}) directed into a given capsule, which also may be expressed using the drop volume (V_{drop}) and the suspension density ($\rho_{suspension} = \phi\rho_{API} + (1 - \phi)\rho_{carrier\ liquid}$), as in Eq. (2).

$$m_{\text{slurry}} = \sum_i^N m_{drop,i} = \sum_i^N \rho_{suspension}V_{drop,i} \qquad (2)$$

If the drop mass and drop volume are assumed to be constants, then the slurry mass in each dose can be expressed as $m_{slurry} = m_{drop}N_{drop} = \rho_{suspension}V_{drop}N_{drop}$, in which N_{drop} is the number of drops deposited into each capsule. The mass of API in each dose may be expressed in a number of ways, and here we choose to relate it as the product of suspension loading (ϕ, particle volume fraction), drop volume, solids (API) density and number of drops, given in Eq.(3).

$$m_{API} = \phi V_{drop}\rho_{API}N_{drop} \qquad (3)$$

For new formulations which target a given API mass, difficulty arises when determining an appropriate solid loading and number of drops due to the effects of solid loading and particle properties on drop formation, and specifically, drop volume. As noted in the discussion of the literature on drop formation in non-Brownian suspensions, the relation between drop volume and particle fraction is complex, and consequently, an iterative procedure is used to produce doses at the target API amount (100 mg) with a satisfactory RSD. In the first iteration, drop volumes are extrapolated from previously collected experimental data (not shown) on drop-on-demand drop formation from suspensions of 38-45µm spheres (density = 1.25g/cm^3) in triglyceride oil at particle volume fractions of 0.1-0.4. To produce doses with API amounts closer to the targeted mass in the second iteration, the formulator may choose to alter the process conditions, solid loading, or number of drops – given the data on drop volumes respective to each API and particle loading

collected in the first iteration, we opted to adjust the number of drops, as changes to ϕ correspond to different rheological properties, and, thus, different drop volumes.

In the dropwise production experiments presented in Figures 3.4-3.8, the system conditions (pump RPM, stroke length, number of volume strokes, nozzle size) were held constant; the Weber number for this process was approximately 1, and the Ohnesorge number, based on model predictions for shear viscosity, was between 0.1 - 5. For each particle type used in these experiments, particle volume fractions were increased until printing was no longer feasible due to jamming and/or particle-liquid segregation during start-up/operation of the process. The results presented in Figure 3.4 demonstrate that accurate dosing is not necessarily achieved on the first iteration due to error in mean drop volume estimates based on extrapolation of drop formation behavior across different particle types. From Figure 3.5 it can be concluded that further tuning of the API mass in each dose necessitates small changes in API mass per drop or drop volume, which correspond to modified suspension loading and/or process conditions. Thus, for a novel product, three or more iterative cycles may be required to identify optimal solid loadings and process conditions to achieve both accuracy and precision for a given dosage amount.
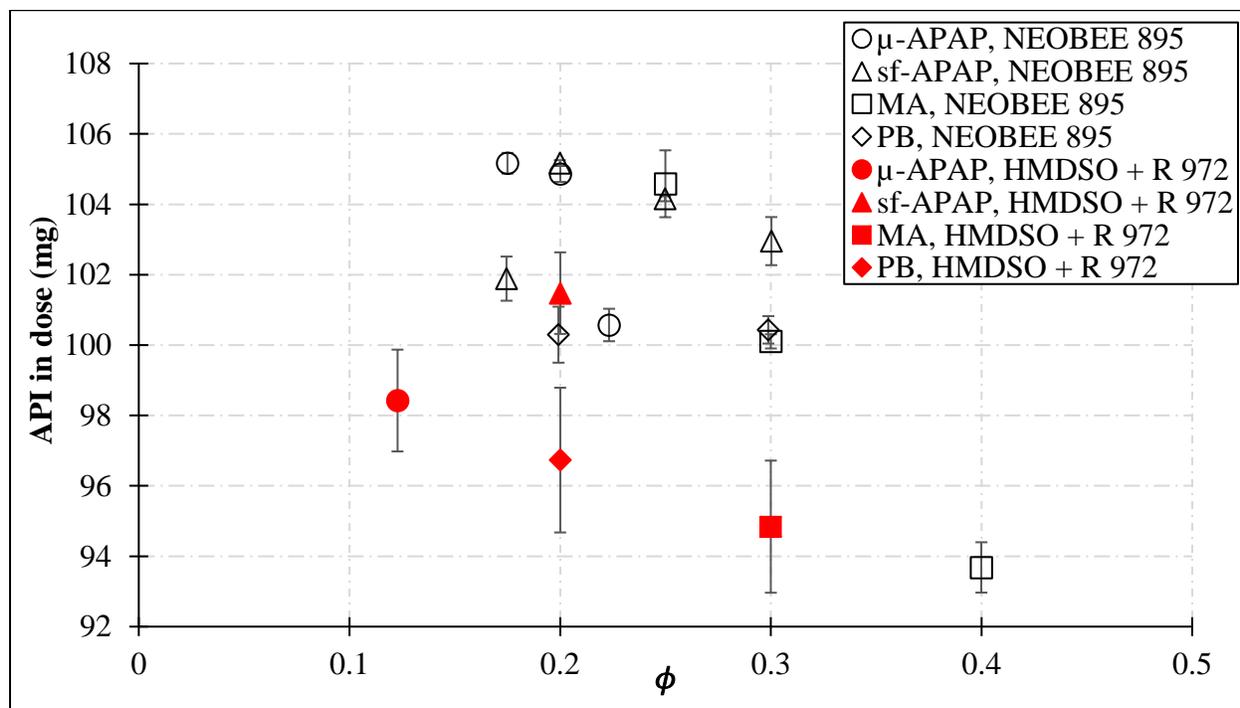
Figure 3.4. Mean API mass (error bars = 1 standard deviation) for API mass target of 100 mg, delivered as a suspension in volatile and non-volatile carrier liquids; n = 20 capsules for each point.
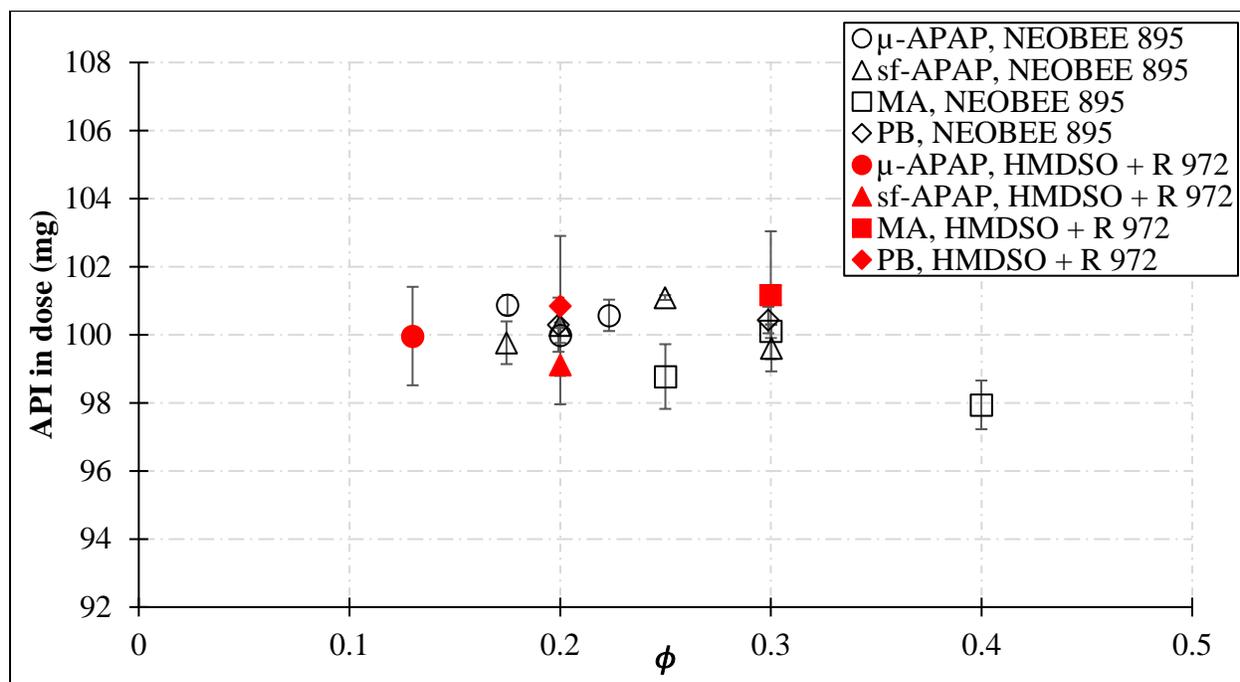
Figure 3.5. Mean API mass (error bars = 1 standard deviation) in second round of manufacturing; n = 20 capsules for each point.
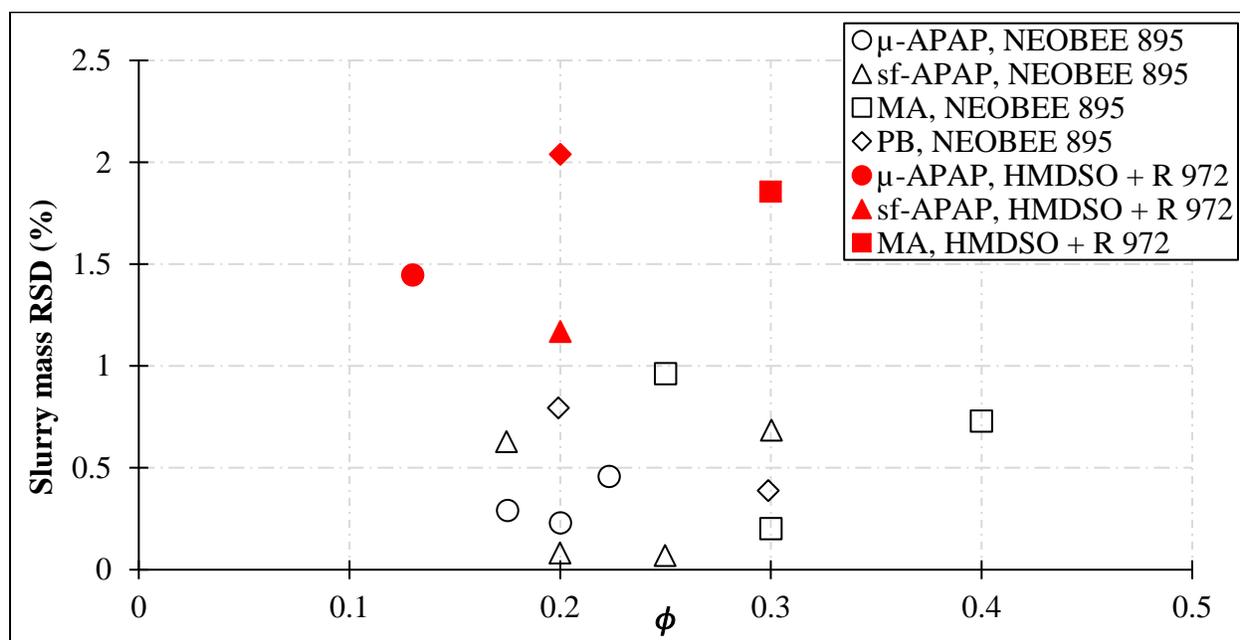
Figure 3.6. Reproducibility, as RSD of slurry mass, for doses (API target – 100mg) from second round of manufacturing; carrier fluids are [1] NEOBEE 895 and [2] 95wt% hexamethyldisiloxane(HMDSO) + 5wt% hydrophobic colloidal silicon dioxide(R 972).

For cases in which the particle volume fraction is assumed to be uniform throughout the process, it can be concluded from Eq. (3) that the reproducibility of the dosing method depends on the stable generation of drops at a chosen volume, given a solid loading and number of drops in the dose. This implies that to achieve the lowest RSD of dosage mass, the particle volume fraction should be tuned so as to minimize the RSD of the drop volume/mass. However, there exist some practical reasons why selecting the particle volume fraction that yields the minimum drop mass RSD may not be the best choice due to its effects on achievable dose amounts (in a given capsule size) and production rate.

The maximum achievable dose in a fixed capsule size is determined by the particle volume fraction of the API in the suspension, as the volume of suspension required to achieve the target API mass must fit within the capsule. This criterion is specified in the left half of Eq. (4), and then combined with Eq. (3) to provide an expression for API mass as a function of capsule volume and particle volume fraction (in the right half of Eq.(4)).

$$\text{in capsule dose,}\ \frac{V_{drop}N_{drop}}{V_{capsule}} \leq 1;\ \Rightarrow m_{API,max} = \phi V_{capsule}\rho_{API}\quad (4)$$

In the scenarios considered in this study, Eq. (4) may be used to specify a lower limit on particle loading that is capable of achieving the target API amount of 100mg; this is the reason for the lowest particle loadings in Figures 4-6 to be $\phi \approx 0.12$. Alternatively, one could extrapolate from the results in Figures 4 and 5 to estimate the maximum achievable dose based on the evidence of feasible operation at the highest values of ϕ for each powder. In this scenario, for Size 0 capsules, dosages up to 250mg should be readily achievable, but were not demonstrated in this study.

3.5.3 Dose Production Rate

For drop ejection/deposition at a constant frequency (f_{drop} = number of drops deposited per unit time), increases in particle loading lead to linear increases in production rate, as evidenced by Eq. (5), which assumes that the contribution of X-Y staging movement (substrate re-positioning between doses) to the total time per dose is negligible.

$$\frac{\text{doses}}{\text{time}} = \frac{\text{rate of API addition}}{\text{target API mass}} = \frac{\phi f_{drop} V_{drop} \rho_{API}}{m_{API}} \qquad (5)$$

In reality, the particle volume fraction itself has an effect on the maximum possible drop ejection frequency due to effects on the drop formation dynamics; the extent of these effects depend on the rheological properties of the carrier fluid and the particle properties, though, for the suspensions in this study, differences in maximum drop ejection frequency are small (≤10%). Based on previous work which estimated the solids mass-based throughput of this system from the drop ejection frequency, drop volume and particle loading [53], the production rate of this system for doses containing 100mg of API is between 1400 and 2000 doses per hour; rates for each API are presented in in Figure B.1, which is included in the supplementary information.

When comparing the production rate of the system used in this study to those achieved by the commercial-scale technologies for capsule filling (liquids: up to 120,000 capsule per h [54]; powders: up to 150,000 per h [55]) or tablet pressing (up to 1,440,000 tablets per h [56]), it is important to note that in the latter two technologies a very good formulation is needed for operation at the highest rates, due to the requirements on powder flowability, compressibility and lubricity. Formulation development may be difficult (or limit production rate) for powders with unfavorable particle properties (e.g. micronized powders, small particle sizes), and consequently it may be a time/resource-intensive endeavor best-suited to commercial products. For cases in which a new dosage form must be rapidly developed and produced (e.g. early clinical trials), commercially available systems such as Xcelodose offer a method by which API powder is directly weighed into capsules. The automated Xcelodose system can fill up to 600 capsules per hour, but encounters similar problems as other dry powder handling operations for particles with poor (dry) flow characteristics.

For micronized powders and/or those with small particle sizes, addition of liquid makes handling considerably easier than it would otherwise be for the dry powder due to the elimination of cohesive and electrostatic effects. Processing of such particles as a slurry takes advantage of the benefits of liquid handling, and the bench-scale dropwise manufacturing platform provides an accurate, flexible dosing method for a range of powders – demonstrated in this study for particle sizes 1-250μm in dose amounts of 100mg. Consequently, the system is capable of functioning as a rapid formulation method for new powders, as measurements of particle true density and powder

bulk density are sufficient to estimate ϕ_m, and thus, feasible values of ϕ for ink formulation, based on prior values of $(\phi/\phi_m, \textit{We, Oh})$. At production rates between 1400-2000 doses per hour, the dropwise system is capable of fulfilling a similar role as the Xcelodose systems, and is unaffected by powders that exhibit poor dry flow behavior. Additionally, linear scale-up of production rate for the dropwise manufacturing process may be achieved by adding more reservoir-pump-nozzles setups – drop generators – in parallel (with corresponding changes to substrate positioning), as the production rates presented in this work are based on a single drop generator.

3.5.4   Effect of Particle Properties on Processing

Based on the RSD values presented in Figure 3.6 and assuming a constant value for f_{drop}, operating at the highest possible particle fraction (ϕ) appears to be the choice that maximizes production rate. Thus, it is worthwhile to consider the factors that limit ϕ. The largest contribution is made by the particle properties themselves (size, shape), which constrain the particle loading to values below that of a "random loose packing," ϕ_m, respective to each particle type. When the data in Figure 3.6 are re-scaled using each powder's bulk solid fraction as in Figure 3.7, it is clear why, for example, DOD printing of suspensions of micronized acetaminophen are limited to values of $\phi \leq 0.225$, whereas for mefenamic acid, values of ϕ up to 0.4 were feasible. For all the particle types used in this study, dimensionless particle loadings, ϕ/ϕ_m, were limited to values below ≈ 0.85; the severity of non-Newtonian flow effects encountered above 0.85 caused such experiments to fail due to jamming in the transit tubing/nozzle and/or flow-induced segregation of liquid from particles.

Figure 3.7 also presents an interesting trend for the upper bound of ϕ/ϕ_m in each particle type: ϕ/ϕ_m is slightly decreased for powders with, on average, larger particles, as indicated by their respective Sauter mean (D[3,2]) and volume-weighted mean (D[4,3]) diameters. This observation may be related to confinement-induced flow disturbances which occur when particles size is comparable to the flow dimensions, as documented in the rheology literature [28, 34]. In this work, the ratio of nozzle diameter to mean particle diameter (D[4,3]) for μ-APAP is ≈ 25, however, for sf-APAP, MA and PB, the ratio is between 8-10; comparison with literature values based on concentrated/densely packed suspensions [28, 34] suggests that for such ratios jamming/shear-thickening related to particle size is less likely for μ-APAP, but may be significant the sf-APAP, MA and PB suspensions.
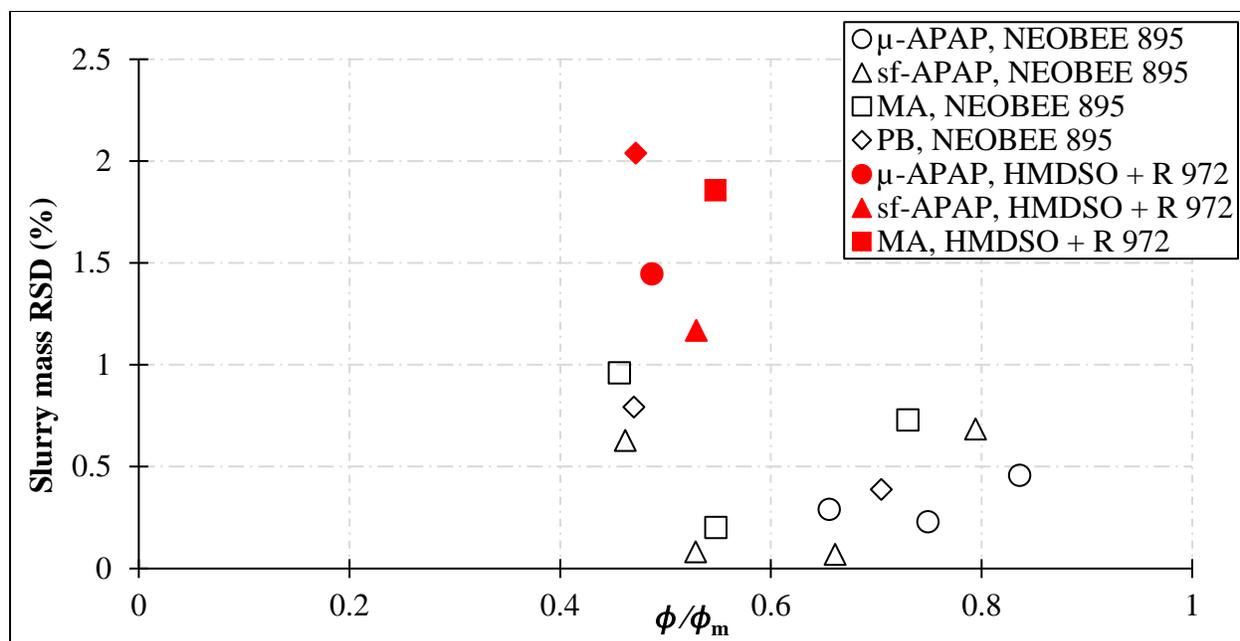
Figure 3.7. Reproducibility, as RSD, with respect to dimensionless particle loading for doses with a target API mass of 100mg.

An additional aspect which is present in Figure 3.6, but becomes apparent in Figure 3.7 is that the RSD of the slurry mass (for 100mg doses) of the oil-based suspensions exhibits a parabolic trend, with the minimum occurring between dimensionless particle loadings (ϕ/ϕ_m) of 0.5-0.75, which correspond to dimensional particle loadings (ϕ) between 0.2-0.3. In light of the literature pertaining to drop formation from non-Brownian suspensions and non-Brownian suspension rheology, this type of trend is sensible – at lower particle volume loadings, satellite drop formation leads to increased variance in the product, then with increasing particle volume loading satellite drops are suppressed, leading to lower variance, and at the highest particle loadings, drop production once again becomes more unstable due to non-Newtonian flow effects (e.g. shear-thinning/thickening). The limitation in the volatile carrier fluid, hexamethyldisiloxane, arose due to evaporation of the solvent at the nozzle: localized loss of liquid caused an increase in the particle volume fraction, which led to jamming of the print head; in order to minimize this effect, values of ϕ/ϕ_m far from the (apparent) maximum were chosen as these suspensions could tolerate some evaporative losses at the nozzle.

As the number of drops deposited into each capsule must be adjusted based on the particle volume fraction and drop volume respective to a set of process conditions and suspension

properties, it is plausible that the number of drops per dose may also affect the slurry mass RSD. This effect is considered in Figure 3.8, which presents the same slurry mass RSD as in Figures 3.5-3.7, but as a function of the number of drops added to each capsule. For comparison with the RSD values of the doses which contained a target API amount of 100mg, additional doses were prepared for API targets of 1 and 10 milligrams using NEOBEE® 895 as the carrier fluid. For low dose products (e.g. 1 mg API), the particle loading in the suspension is determined by the minimum drop volume (~5μL) which can be produced by the system, and in such cases this limit requires that only one drop is deposited in each capsule; consequently, for doses produced from a single drop, the inter-dose variance would be equal to the inter-drop variance. Multiple drops per capsule are required for higher doses (e.g. 10, 100mg), which leads to many feasible combinations of (ϕ, N_{drops}); in such cases, if ϕ does not appreciably affect inter-drop variance, and drop volume is assumed to be a normally distributed random variable, then products comprised of multiple drops should exhibit values of mass RSD which decrease with proportionality $RSD \propto N_{drops}^{-1/2}$. The absence of such a trend in the slurry mass RSD values for doses with API mass target of 100mg suggests that particle loading exerts some effect on the inter-drop variance; this is not unexpected given effect of ϕ on rheology.
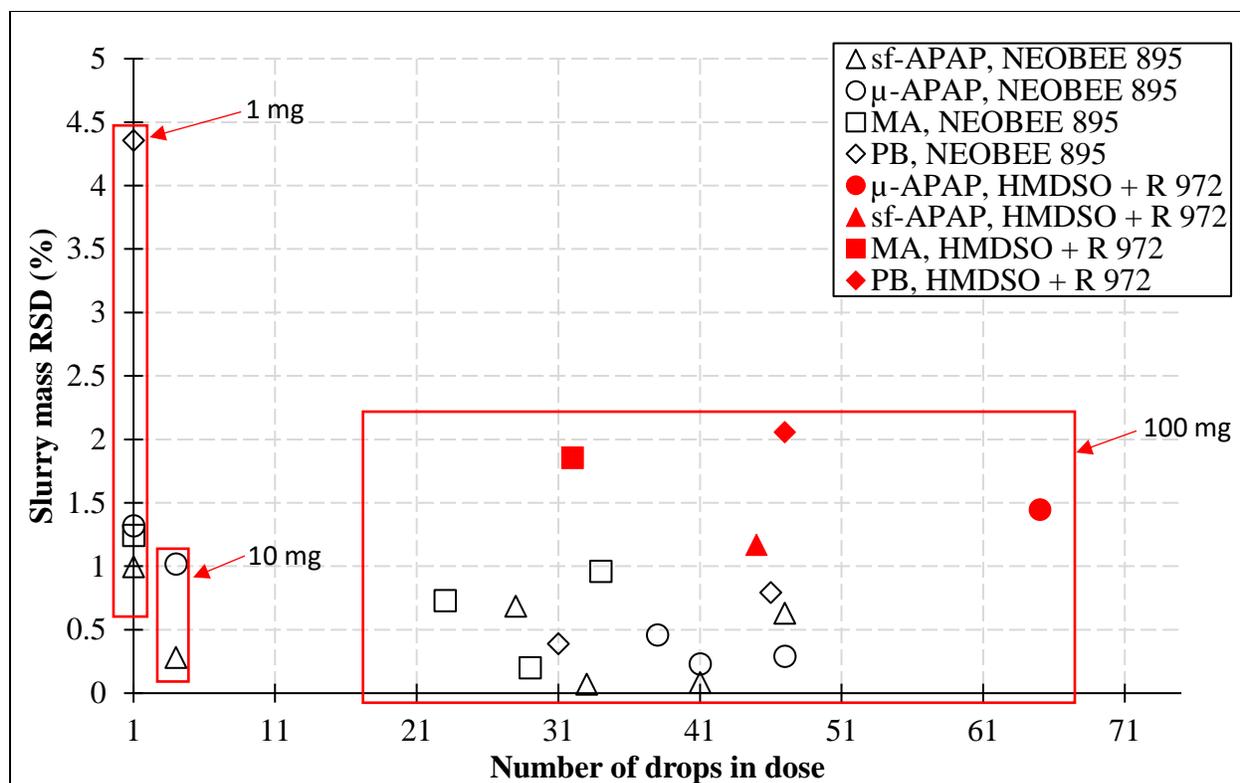
Figure 3.8. Mass RSD as a function of drops per dose for 1, 10 and 100mg API mass targets (in capsules); particle fraction (Φ) adjusted based on drop volume and desired API mass.

3.5.5 Weight Variation of Doses

Using the procedure described in the Methods section to preserve the identity of the capsule contents (slurry mass), uniformity of the dosage units was assessed using the Weight Variation method for hard capsules described in *USP* <905> [51]. As this method is only applicable to doses with ≥25mg of API which contained ≥25% API by weight, the following 100mg doses (present in Figures 3.5-3.8) were excluded as they did not meet the condition on API weight percent (wt%): μ-APAP, $\phi = 0.175$ (22.9 wt%); sf-APAP, $\phi = 0.175$ (22.9 wt%); PB, $\phi = 0.2$ (24.0 wt%). Presented in Table 2 are the acceptance values calculated for the first lots (n=10) from each eligible 100mg dose type (powder type, ϕ), which indicate that the products pass the test for weight uniformity, based on the maximum allowed acceptance value of 15. This result demonstrates the suitability of dropwise manufacturing with API slurries for high dose drug products.

The RSD values in Table 3.2 may be compared with the minimum achievable RSD values as predicted from properties of the particle size distributions [57-59]; the method based on the D[6,3] was selected for its applicability to PSDs with a variety of shapes [59]. From additional

measurements of the particle size properties in the four powders, conducted on a Malvern Mastersizer 3000 (Malvern, Westborough, MA), the D[6,3] values for μ-APAP, sf-APAP, MA and PB are, respectively, 22.9μm, 98.0μm, 213.0μm and 70.7μm, which following the method in Hilden *et al.*, correspond to minimum %RSD values in the 100mg products of 0.029%, 0.256%, 0.802% and 0.149%, respectively. When the predicted values for minimum achievable RSD are compared with the experimental %RSD values in Table 3.2, some discrepancy appears for sf-APAP (0.256%) at $\phi = 0.2$ (0.155%), 0.25 (0.073%) and for MA (0.802%) at $\phi = 0.3$ (0.214%). Assuming that the particle size analysis is sound, potential explanations for this observation may be the occurrence of slight deviations in mass fraction of the products, or, given that the high drug concentration in the suspensions violate a key assumption (i.e. API concentration is dilute) in the model, predictions of minimum %RSD should instead be interpreted as conservative estimates (i.e. upper bounds). To test either supposition, chemical analysis is required for the products, however, it was not performed at the time of writing. The results would be of interest for the three cases mentioned above, and the remaining 100mg doses, given the low RSD values observed in experiments. Similarly, for the low-dose (1mg) products, for which the predicted values for RSD increase by a factor of 10, chemical analysis would be required to assess content uniformity; to relate this to the observed RSD of slurry mass and assess the performance of suspensions for low-dose applications is the subject of a future work.

Table 3.2. Acceptance value (AV) for doses with target API amounts of 100mg, produced from suspensions with ≥25% API by weight. The acceptability constant was k = 2.4, for n = 10 capsules; in this table only, RSD(%) is ($100s/\overline{X}$), in which s is the sample standard deviation.

Particle type	μ-APAP		sf-APAP			MA			PB
AV	0.512	2.065	0.374	0.177	2.263	2.362	0.513	3.289	0.984
$\overline{X}$	99.91	100.54	100.26	101.14	99.30	101.28	100.05	97.52	100.63
s	0.21	0.86	0.15	0.07	0.94	0.98	0.21	0.96	0.41
RSD(%)	0.214	0.856	0.155	0.073	0.95	0.972	0.214	0.985	0.407
ϕ	0.200	0.223	0.200	0.250	0.300	0.250	0.300	0.400	0.300
wt%	25.9	28.7	25.9	31.8	37.5	30.8	36.5	47.2	35.1
N_{drops}	41	38	41	33	28	34	29	23	31

3.5.6 Doses From Solvent-Based Suspensions

For doses produced from suspensions which use a volatile solvent as the carrier fluid, it is possible to quantify the amount of API delivered in each dose by weighing the capsules after evaporation of the liquid. Additionally, by weighing the capsules before drying, it is possible to assess the relation between mass of the wet slurry (immediately after deposition) and dry API mass. This is of interest as it allows for an assessment of API content uniformity in place of chemical analysis and provides support for the assumption of uniform particle concentration in the doses based on the non-volatile carrier fluid (NEOBEE® 895).

In the experiments presented in Table 3.3, doses were produced from suspensions of μ-APAP, MA and PB in a liquid mixture which consisted of hexamethyldisiloxane (90wt%) and polydimethylsiloxane (10wt%, 12,500 centistokes grade); wet slurry mass was measured gravimetrically immediately after printing into the gelatin capsules, and dry slurry mass was measured gravimetrically following 24h of evaporation at 22°C in a fume hood. With respect to the fluid formulation for these experiments, it would be ideal to use only drug particles and volatile solvent; however, in order to achieve suitable rheological properties for drop printing, addition of the polydimethylsiloxane was necessary. The 12,500 cSt grade silicone oil is fully dissolved by the volatile solvent, hexamethyldisiloxane, and consequently, the carrier liquid may be viewed as a homogenous fluid. Thus, a discrepancy between actual drug content (after drying) and the estimate of drug content based on wet slurry mass should be due to deviations in particle concentration during drop printing. Comparison of the mean drug mass estimate with its actual value for the experiments presented in Table 3.3 reveals that for the MA and PB doses, the estimate of drug mass based on the wet slurry is within 1% of the actual value, which suggests that fluctuations in particle concentration are small; however, for the μ-APAP doses the estimate is approximately 4% in error with respect to the dried mass, which indicates non-negligible fluctuations in particle concentration during processing. A potential cause for this observation may be related to the nature of the fluid-particle interactions which are specific to the hexamethyldisiloxane-based suspensions, but absent from the NEOBEE® 895-based suspensions: APAP is very poorly wetted by hexamethyldisiloxane, which results in peculiar (notable yield stress) and undesirable rheological properties. MA and PB, as more hydrophobic compounds, are more easily wetted by hexamethyldisiloxane and, consequently, these suspensions exhibit rheological behavior similar to the NEOBEE® 895-based suspensions.

Table 3.3. Deposits into capsules (n=10), before and after drying, for formulations consisting of μ-APAP, MA and PB particle suspended in the HMDSO(90wt%)/polydimethylsiloxane(10wt%) liquid mixture.

	μ-APAP		MA		PB	
Deposit State	Wet slurry	Dried	Wet slurry	Dried	Wet slurry	Dried
Mean, slurry mass (mg)	81.4	26.9	78.4	32.6	113.4	39.4
RSD (%), slurry mass	2.20	1.70	0.47	0.62	0.39	0.69
Mean, drug mass (mg)	22.2	21.3	27.6	27.5	31.5	31.3
RSD(%), drug mass	2.20	1.70	0.47	0.62	0.39	0.69
ϕ	0.18		0.25		0.20	
wt% drug	27.2	78.9	35.2	84.5	27.8	
wt% polydimethysiloxane	7.2	21.1	6.4	15.5	7.2	
N_{drops}	13		10		15	

3.5.7 Particle Size in Drug Products

In order to further assess the viability of the proposed manufacturing method, it is necessary to consider drug bioavailability, which is determined by the dissolution and absorption of the drug substance, in addition to a host of biological factors which vary between individual persons. *In vitro* dissolution kinetics of a drug product depend on particle size, excipient interactions, and thermodynamic solubility of the active ingredient; thus, determination of the particle size distribution in the final product is of vital interest due to its influence on drug bioavailability. In this manufacturing system, it is plausible that particle size/aspect ratio may change due to shear/flow-induced effects that occur in transit from reservoir to nozzle, or as the result of particle attrition within pump. Consequently, the contents of the 100mg capsule-based doses were analyzed as described in the Methods section for particle size; not all doses were analyzed – one formulation for each powder type was selected for analysis based on presentation with minimum RSD, provided they met the criterion of ≥25% API by weight.

For doses of μ-APAP, sf-APAP, MA and PB produced at ϕ = 0.2, 0.25, 0.3, and 0.3, respectively, the wet dispersion method was used to analyze the particle size properties in the 100 mg doses shown in Figures 3.5-3.8. Comparison of the results for the capsule contents and the prepared suspensions (before printing) revealed only minor differences in the volume-weighted

particle size distributions, as shown in Figure 3.9, and number-weighted aspect ratio distributions, as shown in Figure B.2, for the doses of μ-APAP, sf-APAP, and MA. However, for the doses produced from PB at $\phi = 0.3$, Figure B.2(d) reveals a small shift toward higher aspect ratios, which suggests that for the more needle-like PB crystals, attrition may occur during transit through the system (e.g. in the pump). Based on these results it appears reasonable to conclude that particle size properties are not appreciably affected by transit through the system, thus, for ink formulations (carrier liquid, additives) that do not dissolve the API, dropwise manufacturing with suspensions enables specification of particle size in the final product using upstream processes. For suspensions of low aspect ratio (needle-like) crystals, the effect of processing on particle shape (aspect ratio) should be considered, as suggested by the doses produced from PB in this work.
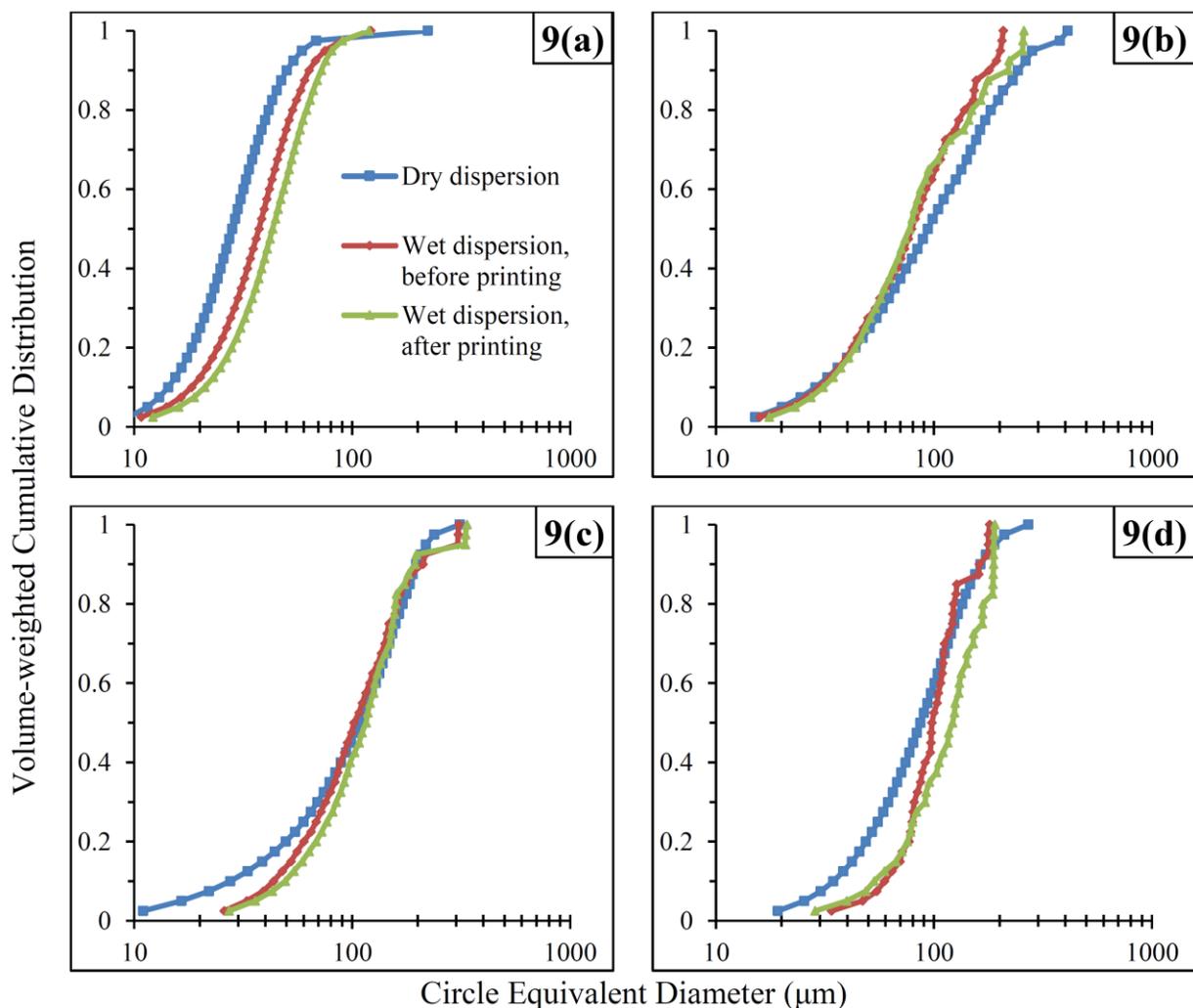
Figure 3.9. Particle size in capsules with target API amount of 100mg: $\phi = 0.2$ micronized acetaminophen, $\phi = 0.25$ semi-fine acetaminophen, $\phi = 0.3$ mefenamic acid, and $\phi = 0.3$ phenylbutazone, denoted as 9(a)-(d), respectively.

3.5.8 Disintegration Testing of Capsule Doses

Capsules from the same lots used in the particle size analysis were tested for their ability to disintegrate under standard conditions using the procedure described in the methods section, in which the filled gelatin capsules were introduced into a stirred, heated (37°C) water bath, with assessment performed visually. Due to the carrier fluid's hydrophobic character (hydrophilic-lipophilic balance (HLB) = 1), dissolution of the gelatin capsule yielded a soft mass (the deposited suspension) that dispersed only when lightly agitated; disintegration times shown in Figure 3.10 correspond to the time at which the drug product formed a soft mass without visible capsule

fragments. The overall average disintegration time of 5.27±0.53 minutes suggests that these products are near the target disintegration time for tablets/capsules, which, should the products meet the dissolution specification, may be a favorable outcome. Despite its utility for preserving particle size of the API, use of NEOBEE® 895 as the carrier fluid may require an additional excipient, such as surfactant, to achieve dissolution specifications for immediate-release products. Alternatively, hydrophilic carriers (e.g. polyethylene glycol) may be substituted for NEOBEE® 895 in order to produce a desired dissolution profile; choice of carrier fluid is restricted only by its rheological properties, and, if preservation of particle size is interest, API solubility in the carrier.
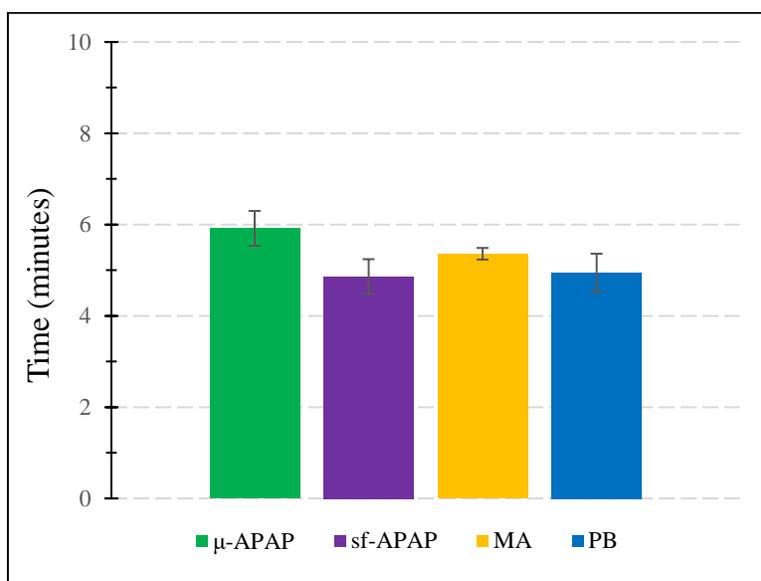
Figure 3.10. Disintegration time of capsules containing 100 mg of API in NEOBEE® 895; particle volume fractions for μ-APAP, sf-APAP, MA, and PB are, respectively, 0.2, 0.25, 0.3, and 0.3.

3.6   Conclusion

The processing of slurries into oral solid doses using a dropwise additive manufacturing platform is investigated in this work for a range of particle size/shape characteristics. Using four active ingredient powders, doses with a target API amount of 100mg were produced by deposition of slurries into gelatin capsules. The effects of particle size/aspect ratio distribution on drop-on-demand dose production were explored at multiple values of particle volume fraction (suspension loading) for each API; doses produced were assayed gravimetrically, which for eligible doses, the Weight Variation method confirmed that drug substance uniformity was within acceptable limits,

and RSD was ≤1%. Based on insights from the literature on rheology of non-Brownian suspensions, the experiments presented in this study suggest that for drop formation from suspensions of non-ideal particles, the effects of particle size distribution and aspect ratio distribution may be modeled using ϕ_m, which may be estimated from two non-destructive measurements: powder bulk density and particle true density. Based on this finding, ink formulation (ϕ, carrier fluid viscosity) for new API powders may be developed based on the feasible range of ϕ/ϕ_m identified in this work; the interstitial fluid of the suspension may be substituted for another carrier liquid and/or excipients (e.g. surfactants) added to increase bioavailability, provided that similar values of *Oh* result. Regarding the effect of particle size on drop formation, it was observed that DOD drop printing was feasible for cases in which the average particle size is not appreciably close to the nozzle size (ratio of particle diameter to nozzle diameter ≥8) for cases in which the Weber number was ≈1. Taken together, these results suggest that dosing of slurries by a dropwise process is a suitable method for manufacturing of high dose drug products from powders with particle sizes of 1-250μm and robust alternative to dry powder-based processes for dosing of powders with poor flow properties.

3.7   Acknowledgements

3.8   References

[1] Ierapetritou M, Muzzio F, Reklaitis G. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AIChE J 2016;62(6):1846-1862.

[2] Nagy ZK, Braatz RD 2012. Advances and new directions in crystallization control. Ann Rev Chem Biomol Engr 2012;3:55-75.

[3] Kayrak-Talay D, Dale S, Wassgren C, Litster J. Quality by design for wet granulation in pharmaceutical processing: assessing models for a priori design and scaling. Powder Tech 2013;240:7-18.

[4] Daly R, Harrington TS, Martin GD, Hutchings IM. Inkjet printing for pharmaceutics–a review of research and manufacturing. Intl J Pharm 2015;494(2):554-67.

[5] Lee SL, O'Connor TF, Yang X, Cruz CN, Chatterjee S, Madurawe RD, Moore CM, Lawrence XY, Woodcock J. Modernizing pharmaceutical manufacturing: from batch to continuous production. J Pharm Innov. 2015;10(3):191-9.

[6] T. Wright, 6 March 2015, Solid Dosage Manufacturing Trends, Contract Pharma; March 6 2015. Available at: https://www.contractpharma.com/issues/2015-03-01/view_features/solid-dosage-manufacturing-trends-724749. Accessed October 8, 2018

[7] Teżyk M, Milanowski B, Ernst A, Lulek J. Recent progress in continuous and semi-continuous processing of solid oral dosage forms: a review. Drug Dev Ind Pharm, 2016;42(8):1195-1214.

[8] Williams HD, Trevaskis NL, Charman SA, Shanker RM, Charman WN, Pouton CW, Porter CJ. Strategies to address low drug solubility in discovery and development. Pharma Rev 2013;65(1):315-499.

[9] Basaran OA, Gao H, Bhat PP. Nonstandard inkjets. Annual Rev Fluid Mech 2013;45:85-113.

[10] Derby B. Inkjet printing of functional and structural materials: fluid property requirements, feature stability, and resolution. Ann Rev of Matl Res 2010;40:395-414.

[11] Eggers J, Villermaux E. Physics of liquid jets. Rep Prog Phys 2008 ;71(3):036601.

[12] Hirshfield L, Içten E, Giridhar A, Nagy ZK, Reklaitis GV. Real-time process management strategy for dropwise additive manufacturing of pharmaceutical products. J Pharm Innov 2015;10(2):140-55.

[13] Içten E, Joglekar G, Wallace C, Loehr K, Sacksteder J, Giridhar A, Nagy ZK, Reklaitis GV. Knowledge provenance management system for a dropwise additive manufacturing system for pharmaceutical products. Ind & Engr Chem Res 2016;55(36):9676-86.

[14] Clarke A, Doughty D, Khinast J, Rantanen J. Development of liquid dispensing technology for the manufacture of low dose drug products. In *Continuous Manufacturing of Pharmaceuticals*; 2016:551-75.

[15] Hirshfield L, Giridhar A, Taylor LS, Harris MT, Reklaitis GV. Dropwise additive manufacturing of pharmaceutical products for solvent-based dosage forms. J Pharm Sci 2014;103(2):496-506.

[16] ZipDose® Formulations, Aprecia Pharmceuticals. Available at: https://www.aprecia.com/technology/zipdose. Accessed October 8, 2018.

[17] Içten E, Giridhar A, Taylor LS, Nagy ZK, Reklaitis GV. Dropwise additive manufacturing of pharmaceutical products for melt-based dosage forms. J Pharm Sci.2015;104(5):1641-9.

[18] Pardeike J, Strohmeier DM, Schrödl N, Voura C, Gruber M, Khinast JG, Zimmer A. Nanosuspensions as advanced printing ink for accurate dosing of poorly soluble drugs in personalized medicines. Intl J Pharm Sci 2011;420(1):93-100.

[19] Wijshoff H. The dynamics of the piezo inkjet printhead operation. Phys Rep 2010;491(4-5):77-177.

[20] Clasen C, Phillips PM, Palangetic L, Vermant AJ. Dispensing of rheologically complex fluids: the map of misery. AIChE J. 2012;58(10):3242-55.

[21] Clanet C, Lasheras JC. Transition from dripping to jetting. J Fluid Mech 1999;383:307-26.

[22] Koch DL, Hill RJ. Inertial effects in suspension and porous-media flows. Ann Rev Fluid Mech 2001;33(1):619-47.

[23] Tropea C, Yarin AL, Foss JF. *Handbook of Experimental Fluid Mechanics*, Springer-Verlag, Würzburg; 2007:680-710.

[24] Stickel JJ, Powell RL. Fluid mechanics and rheology of dense suspensions. Ann. Rev. Fluid Mech. 2005;37:129-49.

[25] Denn MM, Morris JF. Rheology of non-Brownian suspensions. Ann Rev Chem Biomol Engr 2014;5:203-28.

[26] Brown E, Jaeger HM. Shear thickening in concentrated suspensions: phenomenology, mechanisms and relations to jamming. Rep Prog Phys 2014;77(4):046602.

[27] Zarraga IE, Hill DA, Leighton Jr DT. The characterization of the total stress of concentrated suspensions of noncolloidal spheres in Newtonian fluids. J Rheo 2000;44(2):185-220.

[28] Mueller S, Llewellin EW, Mader HM. The rheology of suspensions of solid particles. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, The Royal Society; 2009 Dec 16.

[29] Morris JF, Boulay F. Curvilinear flows of noncolloidal suspensions: The role of normal stresses. J Rheo. 1999;43(5):1213-37.

[30] Cassar C, Nicolas M, Pouliquen O. Submarine granular flows down inclined planes. Phys Fluids 2005;17(10):103301.

[31] Boyer F, Guazzelli É, Pouliquen O. Unifying suspension and granular rheology. Phys Rev Let 2011;107(18):188301.

[32] Moitra P, Gonnermann HM. Effects of crystal shape-and size-modality on magma rheology. Geochemy, Geophys, Geosys 2015;16(1):1-26.

[33] Kalyon DM, Aktaş S. Factors affecting the rheology and processability of highly filled suspensions. Ann Rev Chem Biomol Engr 2014;5:229-54.

[34] Brown E, Zhang H, Forman NA, Maynor BW, Betts DE, DeSimone JM, Jaeger HM. Shear thickening in densely packed suspensions of spheres and rods confined to few layers. J Rheo 2010;54(5):1023-46.

[35] Latham JP, Munjiza A, Lu Y. On the prediction of void porosity and packing of rock particulates. Powder Tech 2002;125(1):10-27.

[36] Reis N, Ainsley C, Derby B. Ink-jet delivery of particle suspensions by piezoelectric droplet ejectors. J Apl Phys 2005;97(9):094903.

[37] Hoath SD, Hsiao WK, Hutchings IM, Tuladhar TR. Jetted mixtures of particle suspensions and resins. Phys Fluid 2014;26(10):101701.

[38] Wang X, Carr WW, Bucknall DG, Morris JF. Drop-on-demand drop formation of colloidal suspensions. Intl J Multiphase Flow 2012;38(1):17-26.

[39] Lee A, Sudau K, Ahn KH, Lee SJ, Willenbacher N. Optimization of experimental parameters to suppress nozzle clogging in inkjet printing. Ind & Engr Chem Res. 2012;51(40):13195-204.

[40] Bienia M, Lejeune M, Chambon M, Baco-Carles V, Dossou-Yovo C, Noguera R, Rossignol F. Inkjet printing of ceramic colloidal suspensions: Filament growth and breakup. Chem Engr Sci 2016;149:1-3.

[41] Furbank RJ, Morris JF. An experimental study of particle effects on drop formation. Phys Fluid. 2004;16(5):1777-90.

[42] Furbank RJ, Morris JF. Pendant drop thread dynamics of particle-laden liquids. Intl J Multiphase Flow 2007;33(4):448-68.

[43] Bertrand T, Bonnoit C, Clément E, Lindner A. Dynamics of drop formation in granular suspensions: the role of volume fraction. Gran Matt 2012;14(2):169-74.

[44] Bertrand T, Bonnoit C, Clement E, Lindner A. Accelerated drop detachment in granular suspensions. Bul Amer Phys Soc. 2011 Nov 22;56.

[45] Lindner A, Fiscina JE, Wagner C. Single particles accelerate final stages of capillary break-up. EPL (Europhysics Letters). 2015;110(6):64002.

[46] Pan Z, Louvet N, Hennequin Y, Kellay H, Bonn D. Drop formation in shear-thickening granular suspensions. Phys Rev E. 2015;92(5):052203.

[47] Miskin MZ, Jaeger HM. Droplet formation and scaling in dense suspensions. Proc Natl Acad Sci. 2012;109(12):4389-94.

[48] Roché M, Kellay H, Stone HA. Heterogeneity and the role of normal stresses during the extensional thinning of non-Brownian shear-thickening fluids. Phys Rev Let 2011;107(13):134503.

[49] Zimoch PJ, McKinley GH, Hosoi AE. Capillary breakup of discontinuously rate thickening suspensions. Phys Rev Let 2013;111(3):036001..

[50] Davis RH, Acrivos A. Sedimentation of noncolloidal particles at low Reynolds numbers. Ann Rev Fluid Mech 1985;17(1):91-118.

[51] *USP* General Chapter <905> "Uniformity of Dosage Units," *USP38*, US Pharmacopeial Convention, December 1, 2015.

[52] *USP* General Chapter <701> "Disintegration," *USP35*, US Pharmacopeial Convention, May 1, 2012.

[53] Radcliffe AJ, Reklaitis GV. Dropwise Additive Manufacturing using Particulate Suspensions: Feasible Operating Space and Throughput Rates. In *Computer Aided Chemical Engineering*, Vol. 40, Elsevier; 2017:1207-1212.

[54] Cole ET, Cadé D, Benameur H. Challenges and opportunities in the encapsulation of liquid and semi-solid formulations into capsules for oral administration. Advanced drug delivery reviews. 2008 Mar 17;60(6):747-56.

[55] Hoag SW. Capsules Dosage Form: Formulation and Manufacturing Considerations. In *Developing Solid Oral Dosage Forms*, 2nd Edition; 2017: 723-747.

[56] Natoli D, Levin M, Tsygan L, Liu L. Development, optimization, and scale-up of process parameters: tablet compression. In *Developing Solid Oral Dosage Forms*, 2nd Edition; 2017:917-951.

[57] Zhang Y, Johnson KC. Effect of drug particle size on content uniformity of low-dose solid dosage forms. Intl J Pharma 1997;154(2):179-83.

[58] Rohrs BR, Amidon GE, Meury RH, Secreast PJ, King HM, Skoug CJ. Particle size limits to meet USP content uniformity criteria for tablets and capsules. J Pharm Sci 2006;95(5):1049-59

[59] Hilden J, Schrad M, Kuehne-Willmore J, Sloan J. A first-principles model for prediction of product dose uniformity based on drug substance particle size distribution. J Pharm Sci 2012;101(7):2364-71.

4. DROPWISE ADDITIVE MANUFACTURING USING PARTICULATE SUSPENSIONS: FEASIBLE OPERATING SPACE AND THROUGHPUT

4.1 Abstract

In dropwise additive manufacturing, particularly in pharmaceutical application, process performance depends critically on stable formation of individual drops of reproducible volume and with a desired trajectory, as determined by the interaction of fluid properties and processing conditions [1]. Current additive manufacturing techniques employ solvent- or melt-based printing inks, with which colloidal particles have been used [2, 3, 4]. However, work with non-Brownian suspensions with particle concentrations relevant to practical application has yet to appear in the literature. Such suspensions involve particle sizes in the 10-200 micron range [5] that are readily produced by conventional crystallization and milling operations. This focuses on elucidating the effects of particle size and aspect ratio distribution as well as particle concentration in suspensions which are used in a dropwise additive manufacturing process for production of pharmaceutical dosage forms. Results with drops produced from non-Brownian particle suspensions consisting of active pharmaceutical ingredient crystals and Newtonian carrier fluids indicate that such suspensions offer substantial advantages in mass-based production rate in comparison to solvent- or melt-based printing formulations. Based on the range of dimensionless parameters accessed in this work, a feasible operating space is developed for drop-on-demand drop formation with non-Brownian suspensions. Process monitoring capabilities sufficient to confirm that the resulting dosages meet quality metrics are also discussed.

4.2 Introduction

The use of dropwise additive manufacturing has become increasingly prevalent in the past two decades, expanding from its origins in traditional printing to the materials (e.g. electronics,

semiconductor, ceramics), biomedical (e.g. pharmaceutical, biomaterials) and precision device (three-dimensional printing) industries [1]. With respect to pharmaceutical applications, particularly individualized medicine, precision drop dispensing on suitable substrates (e.g. inert tablets, thin films) enables flexible, accurate dosing of active ingredients. Recent work has demonstrated the feasibility of producing both solvent- and melt-based pharmaceutical formulations using a dropwise additive manufacturing process (DAM) outfitted with online imaging/sensing, process control (temperature, substrate position, mechanical actuation) and knowledge management systems [2, 3]. Integration of these systems with drop-on-demand (DOD) dispensing technology provides data (e.g. active ingredient mass) for each individual dosage produced.

In pharmaceutics, delivery of precise drug amounts is essential, but equally important is the bioavailability of the active pharmaceutical ingredient (API), which is determined by the crystal size and polymorphic form of the API and potentially the presence of added compounds (excipients). In DAM, crystal size can be modified by controlling the crystallization process at the substrate, though this may give rise to multiple polymorphic forms – in conventional manufacturing desired crystal size/form are achieved through upstream processes. Use of DAM with particles (e.g. API crystals) on this scale necessitates development of the feasible operating regions for this technology with respect to particle size, particle loading and process conditions, due to the complex rheology exhibited by non-Brownian suspensions.

In this work, feasibility of drop-on-demand drop production from non-Brownian suspensions is studied using four types of non-ideal particles, with sizes ranging from 1-200µm. Each particle type exhibits differing distribution of size, and varying degrees of shape anisotropy; suspensions of each are formed at low to high particle loadings and printed across a range of process conditions.

4.3 Drop-on-Demand Drop Production

4.3.1 Process Description

The dropwise additive manufacturing platform consists of a precision positive displacement (P/D) pump, x-y staging, and pump/staging controllers; the setup is a variant of the platform described in [3] less the temperature control system. Drops are produced by the

mechanical action of the P/D pump, which ejects fluid through the nozzle; the conditions (e.g. jetting velocity) at which drops form are determined by the pump displacement, pump rpm and nozzle dimensions.

4.3.2 Image Acquisition, Analysis

As the drop formation process takes place at small timescales (~100 millisecond), a high-speed camera (Photron FASTCAM 100k 1024PCI) is used to record the dynamics of drop ejection; lighting is provided using two 500W tungsten filament lamps. Image sequences are obtained at 6,000 frames per second and at a resolution of 512 x 128 pixels (41µm/pixel) for each set of parameters presented in this work; these image sets are subsequently analysed using a script implemented in MATLAB. Quantitatively, the frequency of secondary ("satellite") drops, jetting velocity and primary drop volume mean/variance are determined.

4.4   Particle, powder characterization

To provide a quantitative description of the particles types and suspensions thereof used in this study, the size distribution and shape characteristics each powder are measured by two-dimensional size analysis using a Malvern Morphologi G3-ID. Particle density and powder bulk density are measured using an AccuPyc II 1340 Gas Displacement Pycnometry System and a GeoPyc 1360 with Transverse Axial Pressure option, respectively. Characterized by the Péclet number, $Pe = 6\pi\mu\dot{\gamma}d^3/kT$, suspensions of the powder grades used in this work may be assumed to be non-Brownian, as evidenced by the particle size ranges presented in Table 4.1.

Table 4.1. Measured particle properties. L_i is the maximum length under which i% of the sample falls (on volume basis); from the particle size cumulative distribution function. Aspect ratio, AR, is defined as ratio of longest to shortest lengths for a given particle. Φ_m is the ratio of particle density to powder bulk density.

	Particle (powder) type			
	micronized paracetamol (μ-APAP)	semi-fine paracetamol (sf-APAP)	mefenamic acid (MA)	Phenyl-butazone (PB)
$\rho_{particle}$ (kg·m^{-3})	1329	1327	1271	1203
ρ_{bulk} (kg·m^{-3})	355	502	696	505
Φ_m	0.267	0.348	0.549	0.424
L_{10} (μm)	15.36	36.7	40.4	50.7
L_{50} (μm)	38.3	118.9	141.4	174
L_{90} (μm)	114.8	351.4	261.4	361
Mean AR	1.57	1.80	1.48	2.01

4.5   Drop Production Results and Discussion

4.5.1   Feasibility

Particulate suspensions were formed from the powder grades and a purified plant-based oil (95% glycerol trioctanoate, 5% glycerol tridecanoate; which is a Newtonian fluid with $\mu = 0.022$ Pa·s, $\gamma = 0.029$ N·m^{-1}, $\rho = 949$ kg·m^{-3}), and in which the APIs are insoluble (or exhibit very low solubility – $\leq$ (10mg API)/(1g oil)); samples were prepared on a mass basis, and particles were dispersed using a vortex mixer and ultra-sonication bath.

For each particle type, drop formation experiments were carried out at multiple process conditions (pump displacement, rpm) and particle loadings, which are defined using the particle volume fraction, given in Eq.(1).

$$\Phi = \frac{V_{particle}}{V_{particle} + V_{carrier\ liquid}} \qquad (1)$$

For drop production from a nozzle, the dimensionless grouping that relates the ratio of fluid inertia to surface tension is the *Weber* number [1], as in Eq.(2).

$$We = \frac{\rho_f U^2 R}{\gamma} \qquad (2)$$

In which the nozzle radius, R, is chosen based on the wetted perimeter – in this case, the outer diameter of the nozzle. The meniscus velocity, U, evaluated here for time points corresponding to

the incipience of flow, is estimated by [1]differentiation (central difference) and [2]linear regression of the *time-position* data obtained by applying an edge tracing algorithm to the image sequences obtained with the high-speed camera – values of U represent averages over 1 millisecond intervals (comprised of 6 individual images).

For each particle/powder type, particle volume loadings were increased until printing the suspension was no longer feasible due to particle-fluid segregation and/or blockage in the suction/discharge lines; these results are presented in Figure 4.1. Limitation on particle loading was found to depend on material type, consistent with the suspension rheology literature, which indicates direct dependence of maximum particle loading (Φ) on the particle properties (particularly, size and shape) [5, 6]. The non-Newtonian properties of such suspensions also exhibits dependence on total stress, deformation rate and flow geometry, details of which will be discussed in a future work. Rheological models (reviewed in.[5, 6]) relate stress to deformation rate in suspensions through a viscosity function which exhibits power-law dependence on Φ, and in particular, the parameter given by (Eq.(3)), for which the viscosity function diverges as Φ/Φ_m $\rightarrow$1.

$$\overline{\Phi} = \frac{\Phi}{\Phi_m} = \frac{particle\ volume\ fraction}{maximum\ particle\ volume\ fraction} \qquad (3)$$

For dry materials, Φ_m is the solid fraction evaluated at given consolidation conditions; regression of data from rheology experiments at several values of Φ can also be used to estimate Φ_m. In both cases Φ_m incorporates the effects of particle size and shape [6, 7]. Solid fraction values from dry powder consolidation experiments have been used provide conservative non-destructive estimates of Φ_m for each respective powder.
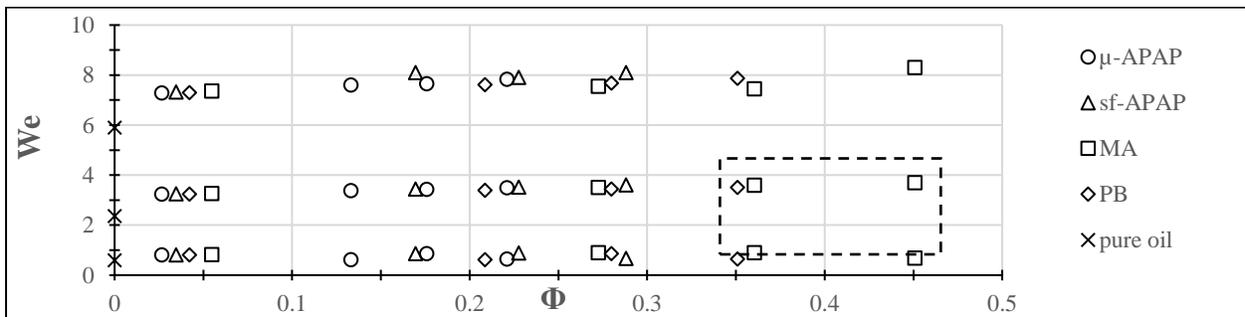
Figure 4.1. Printing of suspensions by *Weber* number and particle volume fraction, Φ; the dashed box contains suspensions that could not be printed reproducibly.
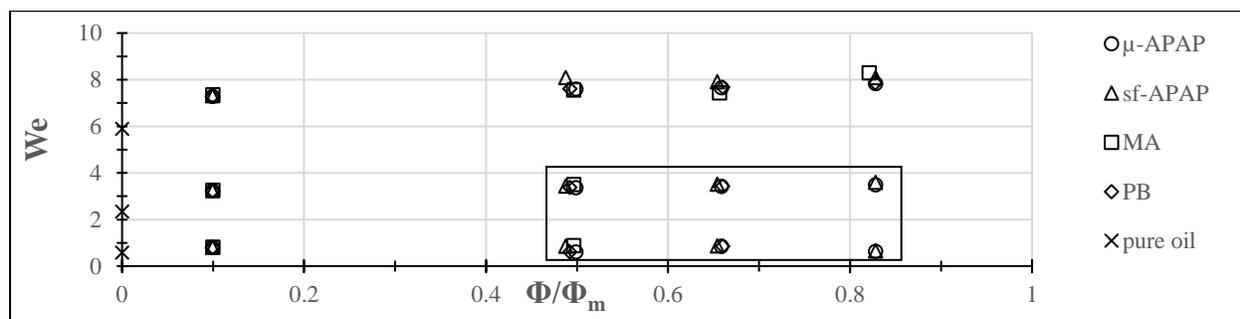
Figure 4.2. Feasible operating region, in solid box, based on the printing metrics for non-Brownian suspensions used in this study.

As shown in Figure 4.2, each particulate suspension type can be brought close to its respective maximum packing limit, which serves to increase mass-based production rates for the process, in addition to suppressing the formation of satellite drops. For the pure oil and suspensions of $\Phi/\Phi_m=0.1$, satellite drops consistently occur with each primary drop produced at low Weber numbers ($We \leq 2$). For the pure oil and all suspensions, higher Weber numbers ($We \sim 10$) produce jets which break up into multiple drops by end-pinching or capillary instability; at higher values of Φ/Φ_m (0.66, 0.83) jetting becomes chaotic, exhibiting behavior similar to that of sprays. However, for moderate values of the Weber number ($We \sim 1\text{-}4$) and $0.5 \leq \Phi/\Phi_m \leq 0.83$, results become more reproducible, with respect to $\Phi/\Phi_m=0.1$. Significantly, primary drop volume variance and satellite drop frequency are minimal for μ-APAP, sf-APAP, MA and PB suspensions printed at $We \sim 1$, $\Phi/\Phi_m=0.83$, 0.66, 0.5, and 0.5, respectively; these results are presented in Table 4.2, accompanied by sample images containing satellite drops in Figure 4.3. The disparity in Φ/Φ_m at which drop printing is most reproducible likely results from larger average/maximum particle sizes (particularly for the latter two particle grades), which in drop formation disrupt the thinning liquid bridge at a length scale proportionate to the particle size [8]; regarding flow prior to exiting the nozzle, it is well-documented in the suspension rheology literature that non-Newtonian effects become more pronounced as the particle size approaches that of the conduit [6].

86

Table 4.2. Primary drop volume coefficient of variation (CV) and satellite frequency (f_{SAT})

	μ-APAP			sf-APAP			MA	PB	
Φ/Φ_m	0.50	0.66	0.83	0.50	0.66	0.83	0.50	0.50	0.66
CV(%)	1.20	0.76	0.72	1.25	0.92	1.45	2.28	1.56	16.2
f_{SAT}(%)	61.5	38.7	21.9	14.8	7.70	14.6	29.1	13.0	19.2
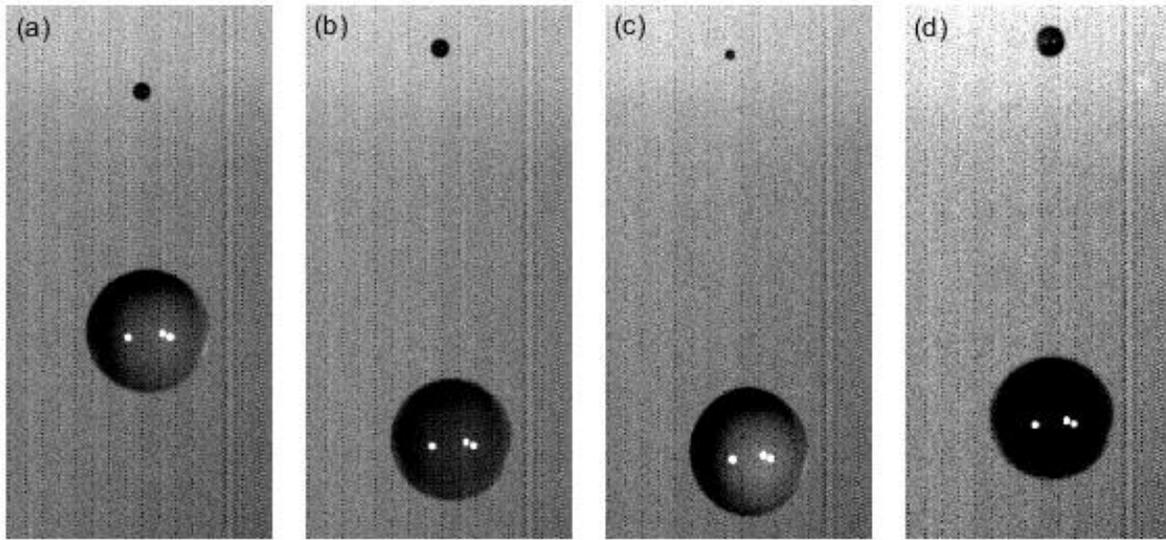
Figure 4.3. Examples of primary drop with satellite; (a)-(d) are, respectively, μ-APAP,Φ=0.22; sf-APAP, Φ=0.23; MA,Φ=0.27; and PB,Φ=0.21.

4.5.2 Throughput

From the image sequences corresponding to drops produced at each set of experimental conditions, individual drop ejection times are computed as the interval from initial meniscus movement through drop break-off at the nozzle, ending when the meniscus is again stable. This time used to estimate the maximum drop generation frequency and production rate for the aforementioned conditions, which are, respectively, Eq.(4) and Eq.(5).

$$f_{drop} = (t_{ejection})^{-1} \tag{4}$$

$$\dot{m} = f_{drop} V_{drop} \rho_{fluid} \tag{5}$$

A comparison of mass-based production rate, on the same dropwise additive manufacturing platform, is presented for solvent-, melt- and suspension-based formulations in Figure 4.4. Noting

that the process is not optimized, and the particles (powders) are used without modification, achievable mass flow rates are substantially higher for suspensions. These rates could be further increased by selecting particles/powders that exhibit high Φ_m values, with smaller average/maximum size.
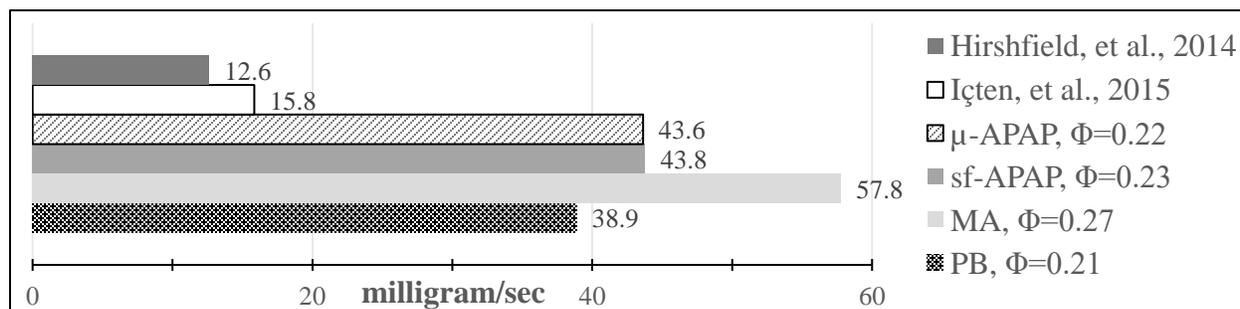
Figure 4.4. Process throughput, based solely on API mass, at maximum drop frequency. Suspensions were selected for minimum drop volume variance; solvent- and melt-based rates were estimated from Hirshfield, et al., 2014 and Içten, et al., 2015, respectively

4.6   Conclusion

This study demonstrates that non-Brownian particulate suspensions can be effectively processed in a dropwise additive manufacturing process. Using dimensional grouping and insights from suspension rheology, a feasible operating region is identified in terms of *We* and Φ/Φ_m. Though segregation due to buoyancy was not observed on the timescale of the experiments, implementation of online and offline Raman spectroscopic technologies are currently underway to confirm solids (API) content of deposits, as segregation effects during suspension transit through the device cannot be ruled out *a priori*. Moreover, offline particle size analysis is being performed to determine whether the particle size distributions and shapes are preserved.

4.7   Acknowledgements

4.8   References

[1] O.A. Basaran, H. Gao, and P.P. Bhat, 2013, Nonstandard Inkjets, Annual Review of Fluid Mechanics, 45, 85-113.

[2] L. Hirshfield, A. Giridhar, L.S. Taylor, M.T. Harris, and G.V. Reklaitis, 2014, Dropwise Additive Manufacturing of Pharmaceutical Products for Solvent-Based Dosage Forms, Journal of Pharmaceutical Sciences, 103, 496-506.

[3] E. Içten, A. Giridhar, L.S. Taylor, Z.K. Nagy, and G.V. Reklaitis, 2015, Dropwise Additive Manufacturing of Pharmaceutical Products for Melt-Based Dosage Forms, Journal of Pharmaceutical Sciences, 104, 1641-1649.

[4] X. Wang, W.W. Carr, D.G. Bucknall, and J.F. Morris, 2012, Drop-on-demand drop formation of colloidal suspensions, International Journal of Multiphase Flow, 38, 17-26.

[5] M.M. Denn and J.F. Morris, Rheology of Non-Brownian Suspensions, 2014, Annual Review of Chemical Biomolecular Engineering, 5, 203-228.

[6] E. Brown and H.M. Jaeger, 2014, Shear thickening in concentrated suspensions: phenomenology, mechanisms and relations to jamming, Reports on Progress in Physics, 77, 046602.

[7] J.P. Latham, A. Munjiza, and Y. Lu, 2002, On the prediction of void porosity and packing of rock particulates, Powder Technology, 125, 1, 10-27.

[8] M.Z. Miskin and H.M. Jaeger, Droplet formation and scaling in dense suspensions, 2012, Proceedings of the National Academy of Sciences, 109, 12, 4389-4394.

5. AUTOMATED OBJECT TRACKING, EVENT DETECTION AND RECOGNITION FOR HIGH-SPEED VIDEO OF DROP FORMATION PHENOMENA

5.1 Abstract

Optical imaging technologies have the potential to provide detailed information which can inform process design decisions via modeling of critical phenomena, or provide innovative process sensors for use in online monitoring and control strategies. In this work, novel algorithms are developed for automated object recognition/tracking and event detection in high-speed video sequences of drop formation phenomena. Using an object-centric structure and generalizations of the physical possibilities, these algorithmic strategies extract and automatically organize the spatiotemporal and modal information for high-level access respective to each object: dynamic motion, deformation; events, event type(s); object lineage through transformations, parent-child relations with other objects. The physics provide the essential patterns, thus, a training dataset is not used or needed. Included results demonstrate the capability of the software for handling complex multi-event scenarios, and with acceptable execution times, processing 2.7GB image data (22000 images/30 videos) in 110-115 seconds in MATLAB on a standard desktop computer.

5.2 Introduction

Within the process industries, optical imaging can provide detailed information which has the potential to inform design decisions by providing measurements and data – which are otherwise inaccessible – that enable modeling of critical phenomena [1, 2]. Additionally, online image capture systems can provide innovative process sensors, which may be used in real-time process monitoring and control strategies when coupled with automated analyses [3, 4]. Potential applications to physical/chemical processes are numerous, and constitute essentially any scenario in which an indicative visual change occurs – examples are prevalent in the study of fluid and/or particle flows, as such phenomena lend themselves to the use of imaging for quantitative purposes [5, 6, 7]. Motivated by applications in chemical engineering, the work presented here considers the development of algorithms for analysis of high-speed videos of drop formation, breakup and coalescence; this constitutes a complex physical scenario involving motion and free-form

deformation of a random number of objects that are capable of subsuming other objects and/or serving as the source of new objects through non-deterministic events which occur under multiple modes and with variability in temporal ordering.

For free surface flows such as those that occur during jet breakup and drop fragmentation, high-speed imaging is the experimental technique of choice due to the timescales involved, and has been employed over the last three decades to inform the theoretical and computational study of these phenomena [5, 8, 9]. Understanding of the underlying mechanisms is important for inkjet/3D printing, microfluidics, spraying/atomization, liquid/liquid extraction, and the multitude of other processes/applications that involve the formation, fragmentation and/or coalescence of fluid jets, drops and bubbles [5, 10]. In such operations, it is often desirable to establish operating regimes that permit consistent generation of drops in a given size range; this has bearing on design considerations, and on process control strategy, as the moieties produced within such process typically depend directly on the drops formed. Even for ideal Newtonian fluids, this can be a difficult task due to the strongly non-linear physics inherent to liquid breakup and coalescence; with the additional complexities, and perhaps stochastic effects introduced in multi-phase mixtures, experimental investigation with (automated) image processing of results may be the only practical way to acquire data for modeling.

Utilization of high-speed imaging for qualitative or quantitative purposes may involve a considerable time investment by the researcher for analysis due to the size of the data sets, which contain 10^3-10^6 images per experiment [7]. The use of digital image processing to evaluate salient features and facilitate analysis for such applications is certainly not new, however, with respect to drop phenomena, it is often ambiguous as to how the analyses are performed, with few details provided, and it is also unclear to what extent the analyses are automated. In addition to these considerations, it must also be noted that in order to enact more detailed modeling efforts – particularly for applications that involve marked stochastic effects – an advanced approach is necessary to extract and organize the information from such image sequences so that it is accessible via high-level specifications in terms of each object, its spatiotemporal history, and events.

Such a structure is especially relevant to modeling of jet breakup/drop formation in non-Brownian suspensions, which exhibit marked variance in the temporal, spatial and modal transitions, even for 'ideal' suspensions of identically-sized spheres at limiting conditions [11, 12, 13, 14, 15]. Given their dependence on particle size, these stochastic effects would be intensified

for suspensions of non-ideal particles that exhibit a size distribution and shape property (e.g. aspect ratio, convexity) distribution(s). Modeling of these effects is an essential step in optimization of design of drop-on-demand printing processes that utilize such fluids [16], and due to the non-idealities involved, a hybrid approach based on the data is pragmatic.

The objective of this paper is to present a set of systematic methods and algorithmic strategies which enabled fully automated analysis of large image sets for jet formation/breakup, drop fragmentation and drop coalescence, and to demonstrate the implementation speed and results for a sample data set of 22,000 images. Foremost, the set of algorithms is designed to work for both high- and low-quality imaging setups with varying levels of noise and object scales by pre-processing using an adaptive edge detector to reduce the input images/videos to binary edge maps of the relevant features, which are subsequently used for the object and event analyses. Included in this paper are systematic methods for extraction of conventional quantitative information (filament diameter dynamics & velocimetry) which is useful to the fluid mechanician, rheologist, or physicist interested in jet breakup; however, this is not the focus of this work. The novel contribution of our work consists of the development of object recognition/tracking and event detection/classification algorithms for mining statistical and quantitative information in drop formation phenomena. In essence, we develop algorithms to extract and subsequently structure the spatiotemporal and modal information across the total lifetime of each object: translational motion, free-form deformations, breakup/coalesence events, and parent/child relations with other objects.
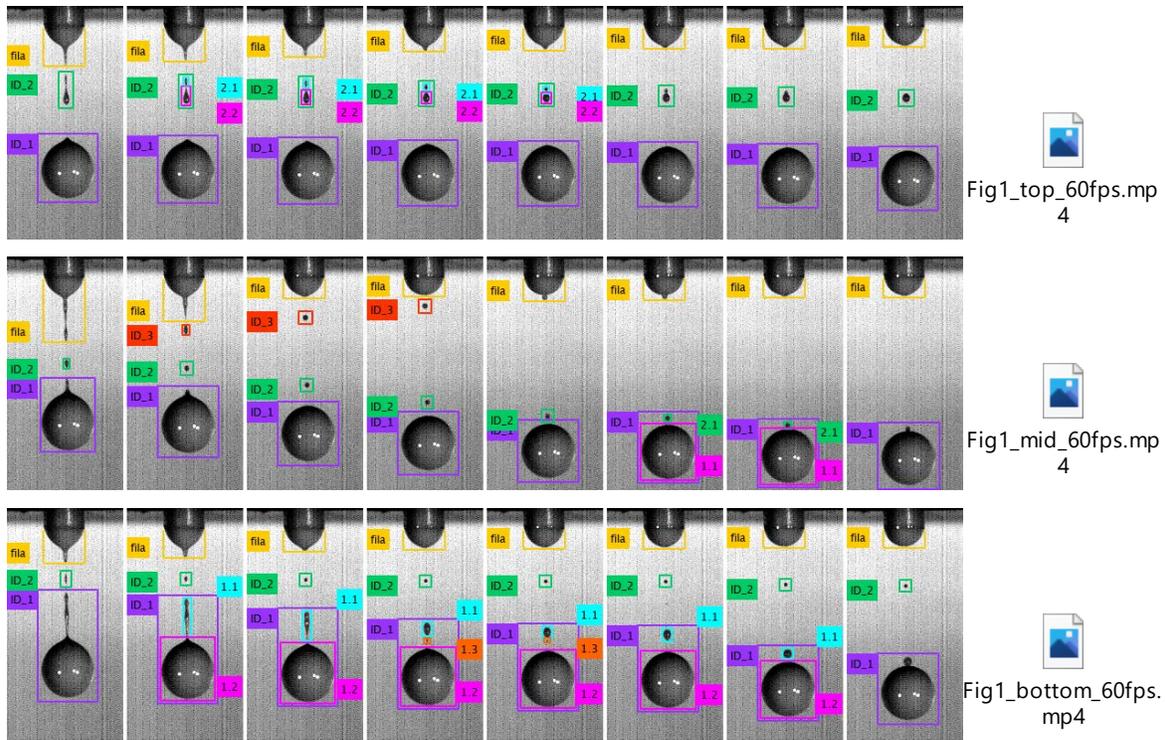
Figure 5.1. Top: Two fragments formed by breakup of the second drop, which then coalesce; 2.1 & 2.2 denote sub-identities of the global ID_2. Middle: Breakoff of three drops, ID_1 & ID_2 coalesce later, and ID_3 merges with the nozzle. Bottom: Two breakup and two coalescence events occur within the primary drop, ID_1; event sequence: breakup-breakup-coalescence-coalescence.

The challenges of developing methods to achieve this arise due to the peculiarities of jet/drop formation, breakup and coalescence, during which persistent objects deform, undergo singularities and give rise to transient objects. These transient objects experience translational motion and free-form deformations, but, perhaps more significantly, starting at their instantiation, several transients enter a realm of possibilities. Each droplet may itself undergo breakup events, thereby forming two or more fragments; additionally, droplets or fragments may coalesce or recombine with both transient and persistent objects, thereby introducing many potential scenarios due to the number of possible events and sequences – combinations – of events in which a given fluid packet participates. The stochastic nature of these phenomena complicates the tracking of objects, tokening and assignment of identities, and association of spatial/temporal/modal information, particularly if one is interested in determining exactly what occurs to each fluid fragment through a series of breakup and recombination events, which is the objective of this work. Consequently, this requires development of algorithms that detect events and then adapt

93

accordingly – dynamically activating or changing methods based on the observed occurrences – by means of generalizations for the possible physical situations and events. This is in contrast to another possible approach, which would be through a deterministic program that exhaustively lists all physical cases. Instead, by using an object-centric structure and generalities of the physical patterns underlying the phenomena, the algorithmic strategies presented in this work can be applied to any image sequence involving the formation/breakup/coalescence of jets, drops or bubbles in order to extract and organize the spatiotemporal and events information. Rather than specifying the type(s) and sequence of events, or number of objects in a given recording, the algorithm automatically recognizes objects from their motion, dynamically instantiates identities, and detects/recognizes events (also participants in events). The range of observations in the data set included in this study demonstrate the necessity of and capability of the methodology in handling a wide variety of scenarios without the need for a training database.

The set of methods is combined into a software package which accepts as its input an image sequence/video, and as output, quantification of the above items; multiple image sequences/videos may be passed to the program, such that by specifying the disk directories which contain the image data of interest, the entire process is automated. The organization of the paper is as follows: in Section 2, details of the high-speed imaging setup are provided, and the edge detector is briefly described; in Sections 3 and 4, the simpler components precede the object tracking and event detection/recognition algorithms which form the novel contribution in this work; then, in Section 5, the results obtained from 30 experiments are briefly presented; thereafter, the work is summarized in light of its potential uses and adaptation to more general cases than those presented by jet breakup, drop fragmentation and drop coalescence.

5.3   Image Acquisition and Edge Detection

5.3.1   Experimental Setup

The image sequences in this work correspond to recordings of drop-wise printing of concentrated non-Brownian (particle size range 1-250 micrometers) suspensions. In the drop-on-demand mode, fluid is ejected from the nozzle by a pressure pulse, supplied by a single actuation of the precision P/D pump (IVEK Digispense 10 single-channel positive displacement pump, IVEK Corporation, North Springfield, Vermont), after which the liquid filament elongates and

ultimately detaches from the nozzle, forming one or more drops. The dynamics of these combined phenomena were recorded using a high-speed camera (FASTCAM 100k 1024 PCI, Photron, San Diego, CA), with front lighting provided by two 250W tungsten filament lamps and one 500W lamp; smooth, non-reflective, white art board (Blick Art Materials, item 13104-1102) was mounted 3cm behind the nozzle to provide a consistent background and improve contrast in the recorded image sequences. This setup is depicted in Figure 5.2
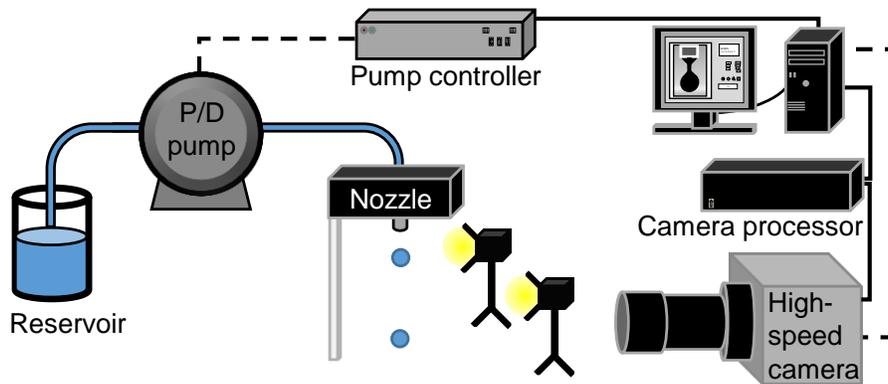
Figure 5.2. Drop generation platform, with high speed camera; combined ejection and detachment timescale is 50 milliseconds
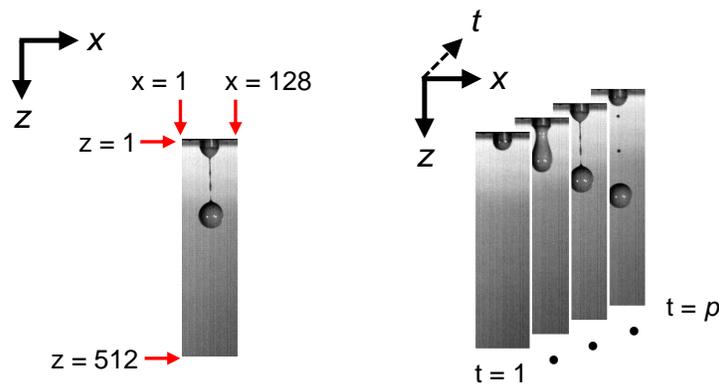
Figure 5.3. Sign convention and coordinate system of images. Left: single image; Right: image sequence, chronologically ordered

Videos were recorded in grayscale at 6,000 frames per second with a resolution of 512 x 128 pixels (scale in images: 45.4 micrometres per pixel, determined from known dimensions of the nozzle); the constituent images of each video were saved as sequences of 16-bit uncompressed TIFF images for subsequent analysis. Each image sequence corresponded to one liquid ejection

event, such that for 30 drop ejections at a given set of experimental conditions, there were 30 respective image sequences.

5.3.2   Image/Sequence Data Conventions

A brief note is warranted on the sign convention and coordinate system used for a single image frame, as well as for an image sequence. A single 16-bit grayscale image may be represented as a matrix whose entries take values between [0, 65535]; the numbers of its rows and columns are equal to the number of pixels in the z and x directions, respectively, thus, $m = 512$ and $n = 128$. The indexing convention used throughout this work, shown in Figure 5.3, is as follows: the upper boundary of the image corresponds to the first row of the matrix, with index $i = 1$, and the left boundary of the image corresponds to the first column of the matrix, with index $j = 1$, such that the upper left corner of the image is entry $(i,j) = (1,1)$. To represent a video – an image sequence that consists of chronologically ordered images – an index, k, is added, $k = 1, \ldots, p$ ($p = 500\text{-}1000$, depending on number of images per sequence), to form an array from the matrices.

5.3.3   Image Processing: Edge Detection, Filtering and Thresholding
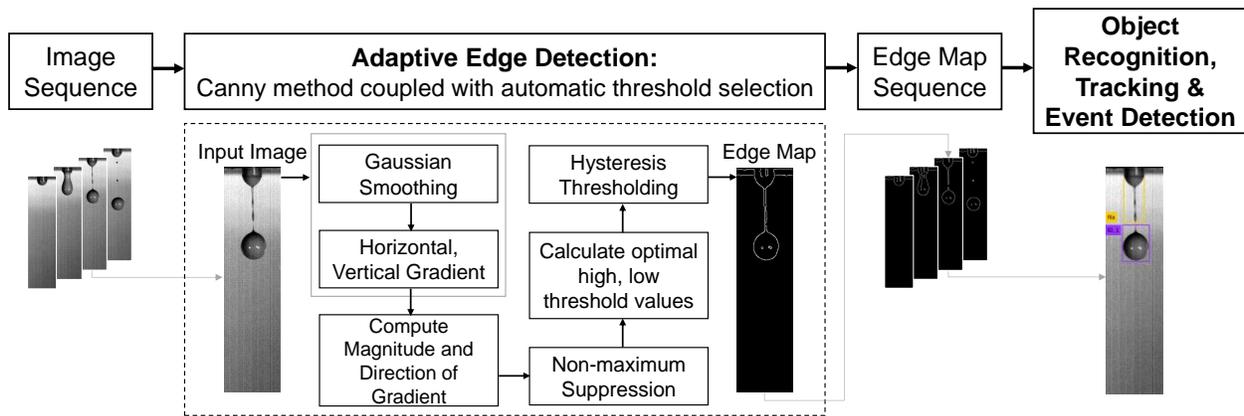
Figure 5.4. Overall information flow; the components in the dashed box are described in Appendix D.

To reduce the data to the salient information, each constituent image respective to a given video sequence was processed by means of an edge detection algorithm in order to produce a binary map of pixels at the feature edges within each image. The resultant time-indexed sequence of edge maps is then the starting point for all subsequent analyses, as shown in Figure 5.4. In this

context, edge detection is a necessary pre-processing method; with the correct design, the edge detector can automatically adapt to the characteristics (e.g. noise) of the imaging setup and therefore can be applied to entire image sequences. For the purposes in this work, a modified Canny detector was used as it proved to be robust in the physical scenarios encountered here; details of the detector, relevant considerations and specific modifications [17, 18, 19, 20, 21, 22, 23, 24, 25] are presented in Online Resource 1. Note that improvements (or substitutions for equivalent methods) can be made to the edge detection algorithm without affecting the methods or algorithms in Appendix E or Section 5.5.

5.4   Transition to Analysis Algorithms

In Sections 5.4.1-5.4.3, the motivation is presented for the systematic methods used to compute physical quantities such as velocity, filament diameter (and dynamics thereof) and drop volume; detailed descriptions are provided in Online Resource 2. Then, Section 4 presents the novel contributions in this work, which are the algorithms developed for extraction of temporal, spatial, modal and statistical information related to free-form deformations, fragmentation (breakoff/breakup), and coalescence (merging) of dynamic objects in motion.

Throughout the following sections, the still image frames and videos were created by writing to the respective matrix of each frame (edge map or real image) at the entries which correspond to the parameters obtained by the algorithms. In the case of horizontal lines that cover the width of the image, the entire row at the parameter of interest, indicated by the color-matched box with its name, is marked by writing a specific red-green-blue (RGB) triplet at all entries of the row. The cases which display the grayscale images with two-dimensional shape are generated by essentially the same method – writing an RGB triplet to specific positions which correspond to parameter values. The additional information required to represent the two-dimensional shape was obtained by applying the bi-directional search (5.5.1.3, 5.5.2.3/5.5.2.5) to the left/right boundaries of each object. In either case, the writing of RGB triplets to mark parameter values is a direct means of visualizing a succinct portion of the information extracted by the algorithms.

5.4.1 Velocity

Velocimetry is useful in understanding the stages of drop ejection through dimensional analysis/scaling, and may provide insight into the effect of particles and suspension on rheology/rheological transitions, which affect filament shape and therefore breakoff through the curvature of the liquid surface. Two different methods, described in the corresponding section of Appendix E, are employed to estimate velocity during the initial stages of liquid ejection: 1) maximum velocity, based on the temporal evolution of the lowest point on the liquid meniscus, and 2) the axial velocity profile and dynamics thereof, constructed from the temporal evolution of the liquid meniscus position at each x-pixel location.

5.4.2 Diameter of Filament as Function of Time, Spatial Variables

The mean curvature of a free liquid surface determines the influence exerted by surface tension on drop dynamics, wherein it is the most important driving force of drop/jet breakup (Eggers 2008). Therefore the diameter of the liquid filament as a function of time and axial position are of significant interest in physical modeling; in particular, the minimum width of the filament in the region which is actively decreasing in diameter is an important parameter in understanding the breakoff dynamics and thinning mode. Consequently, systematic methods for extraction of diameter, $D(t, z)$, and minimum diameter, $D_{\min}(t, z)$, are provided in Appendix E.

5.4.3 Drop Volume, Drop Ejection Time

In drop printing applications (e.g. inkjets), evaluation of process performance in light of particle properties, ink formulation and operating conditions can be considered from a limited perspective, namely, consistency of drop formation (i.e. volume) and drop production frequency. In the respective section of Appendix E, we present systematic methods for computing drop volume, and estimation of drop ejection time; as the latter does not utilize the object tracking/event detection framework from Section 5.5, it suffers some inaccuracy as demonstrated by Figure 5.26.

5.5   Transition to Object Tracking & Event Detection, Recognition

Before proceeding to the next sections, a brief introduction is warranted in order to clarify the capabilities of the developed algorithms which constitute the bulk of the novel contributions

in this work. We begin by clarifying the autonomous nature of the program which requires as its only input from the user an image sequence in order proceed through the analyses. The user does not (and cannot) specify things such as number of detachment events, number of post-detachment events (breakup or coalescence), or the identities of participants in post-/detachment events; similarly, the time step and positions of such events cannot be provided to the program in order to seed it. Instead, the program logic manages all such aspects: (1) it tokens object identities through recognition, and associates features – all positions of the object – to each identity; (2) from their instantiation onward, the features of the existent objects are updated at each time step by matching the motion of each object. In addition to the above feature recognition and modeling which enables object tracking, the program employs a novel construct for the detection and classification of post-detachment events of the free objects, which necessarily involves its own tokening and management of object identities.

In fact, the complexity of post-detachment events necessitates the development of the methods described herein. Since the objective is to apply the analysis without user intervention, then the combined algorithms must be capable of reproducing all the physical possibilities – and correctly identifying them. The difficulty arises not from the phenomena themselves – breakup or coalescence – the physics of which have been widely studied, but from the stochastic nature of free object (post-detachment) behavior; namely, that a single object may exhibit multiple events and combinations of events. From the perspective of developing a computer vision algorithm, this constitutes a decidedly non-deterministic problem: one is unable to specify the type(s), order, or number of events an object experiences, and moreover, the participants are unknown *a priori*. Given the sizeable number of possibilities that can be created starting from just a few objects, the complexity of the problem can be significant.

We approach this problem by developing algorithms to recognize and track object interactions without the need for a training data set or user input; as developed, this enables the program to learn the unique modalities of detachment and post-detachment breakup/coalescence contained in each image sequence. This is made possible by the recognition of singularities (in the spatiotemporal history of the objects), which are inherent to breakup, and which occur in coalescence as well. The only structure supplied to the algorithm is a generalization of the physically possible scenarios, namely, that a transient object, with identity coined at the instant of detachment, may itself break into fragments, and/or undergo coalescence events with any unknown

partner(s). Each instantiated object manages its own motion tracking using the recognition/matching algorithms; simultaneously, it detects object interactions and tokens additional object identities as needed based on the generalized logic structure provided for events. It should be noted here that by generalization we mean that each object has the same motion tracking algorithms, same definition of interactions and equal capability to manage/token sub-identities; therefore, in proceeding from the previous time step to the next, each object – each free liquid volume – is not precluded from breakup or coalescence, just as is physically possible.

In this manner, the following is obtained: the total set of positions respective to each unique object, for each time step at which the object exists, organized with respect to the events the object has participated in. Significantly, this object-centric information can then be used for secondary purposes, i.e. modeling and optimization, and is easily accessible using high-level logic for sets/subsets of objects and/or events of interest. As this information is inevitably high-dimensional due to the motion tubes formed from the dynamics of the total position, succinct visualization of the full information is not possible; however, it is possible to present a single object, its events and its parent-child relation to other objects by reducing the motion information to single (z,t) points, as shown in Figure 5.30.

The next two sections present the development of these algorithms: Section 4.1 considers the dynamics and modalities of detachment of liquid volumes from the nozzle, whereas Section 4.2 is concerned with post-detachment occurrences.

5.5.1 Drop Detachment & Filament Dynamics Before, Through and After Breakoff

Motivation

The reduction in the diameter of the liquid filament after fluid ejection, break-off of a primary drop (and potentially secondary drops) and subsequent retraction of the liquid filament are duly affected by the presence of particles, thus, it is necessary to consider questions such as:

- At what time point and location(s) does the liquid filament break first? What is the shape of the filament and nascent object(s) at this event?
    - o This relates to the effect of particles (particle size, shape; particle volume fraction) on drop formation and extensional rheology.

- Do secondary break-off events occur, and if so, at what time point(s) and location(s)? How many?
  - Same motivations as above, but also important in assessing process performance.
- After break-off of the primary drop, what is the position and shape of the liquid filament as a function of time?
  - This helps to determine the time required for the upper portion of the liquid filament to retract to the initial position.

To answer these questions, this section presents the methods for analysis of drop break-off and filament retraction, namely, recognition of the persistent object – the fluid attached to the nozzle – and association of features, tracking of its motion and deformations, and detection of singularity events.

Overall Algorithm

Consider the temporal sequence of physical situations (Figure 5.5) presented by the first drop to detach from the nozzle, starting from the incipience of liquid meniscus movement in response to the leading edge of the pressure wave generated by the actuator. A jet is temporarily produced, resulting in formation of an unstable liquid filament. Logically, the break-off event must occur at some z-location on the interval occupied by the stretching liquid filament; this interval may be bracketed at each time point with an upper bound equal to the z-location of the nozzle orifice, and the lower bound equal to the location of the attached liquid meniscus farthest from the nozzle (before the first drop breaks off, this is simply the liquid surface location farthest from the nozzle). At the instant of breakoff, the singularity structure is formed: the diameter of the filament goes to zero at some z-location within the bracketed interval, $z(t)$, respective to this time. Therefore, this constitutes the condition which may be used to detect the event – the time at which breakoff occurs is the first time point in the temporal sequence of diameter vectors, $d(t, z(t))$, at which any entry is equal to zero. A note: the terminology of the bound positions used throughout this section is with respect to the spatial orientation, thus, the "upper" bound is actually the left bound of the interval, and the "lower" bound is the right bound of the interval.
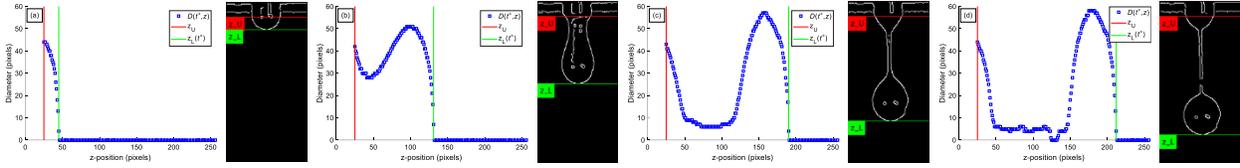
Figure 5.5. Illustrative plots, and associated binary images, with $\boldsymbol{z}(t)$ bounds marked in red/green. Temporal order is (a)-(d), with (a) the start of flow and (d) the time of breakoff.

With the instant of breakoff detected by the recognition of the singularity, the z-positions of the breakoff event are identified as the nascent liquid surface boundaries above and below the z-location at which diameter is equal to zero. The total length of the filament at the instant of breakoff is known from easily identified as the bounds of the bracketed z-interval at the time of breakoff, and the (x,z) positions at each left/right boundary – object shape – are already obtained as they are required to compute $\boldsymbol{d}(t, \boldsymbol{z}(t))$. This allows one to answer the first question above in a systematic manner; the details of the implementation are presented below, with physically possible situations that justify the methods employed.

Before proceeding to these descriptions, it is opportune to discuss how the information required for the second and third questions might be obtained, which relates to the problem of how to track the still-attached liquid thread as it retracts to the nozzle after the singularity event. To follow the temporal evolution of the attached filament, one notes that the nascent liquid surface above the z-location(s) of zero diameter at the time of breakoff must be the starting point of any search algorithm; this provides the basis for reallocation of features to the persistent object when a singularity is detected. What complicates this is that the liquid surface boundary one wishes to track may ascend or descend depending on the particulars of the breakoff, thus, a bi-directional search is needed; the method for this is described later in this section. But first, one must introduce an additional complication related to the second question, which is: secondary, tertiary, quaternary, etc. breakoff events may occur within the liquid thread, coincident with the first breakoff, or during retraction.

Given that one would like to track only the liquid thread attached to the nozzle in order to answer the third question, these extra breakoff events must be recognized and the features belonging to the persistent object updated. The physical reasoning used to detect additional drop detachments during filament retraction is identical to the logic used to identify the first breakoff event, but with one modification: rather than the "first" time point in the temporal sequence of

102

$d(t, z(t))$ at which any entry is equal to zero, one seeks the "next" time step – that is, if one views the temporal sequence in its entirety – but if one restricts the perspective to all times after the most recent breakoff, then "first" is the intuitive objective once again. Viewed as such, the breakoff of extra drops is identical to the breakoff of the primary drop, with the only differences being in the temporal ordering of the events. In the implementation this is enacted by a query at each time step of $d(t, z(t))$; if a breakoff is detected, the relevant z-locations are found, and the positions of the attached thread are shifted to the (uppermost) nascent liquid surface so as to reflect the modal change in features of the persistent object – the attached filament – which also coincides with the tokening of a transient object by means of the singularity.

These observations provide the basis for the general logic for handling any number of detachment events during the tracking of the attached filament which gives rise to the transient objects: free liquid volumes, or droplets. Therefore, when applied to any image sequence, the combined algorithms fully capture the modalities and dynamics of jetting, drop detachment and filament retraction through any number of breakoff events, whether sequential or simultaneous. The extracted information defines the entire shape of the attached liquid filament at not only the time(s) of detachment events, but also for all time points in the recording. Furthermore, a unique (transient object) identity is instantiated for each droplet at the instant of breakoff, and is used to associate respective times, features, locations and quantities which form the starting point of the post-detachment tracking described in Section 5.5.2. Given the potential complexities of object identity/feature association to transients, systematic methods for treatment of simultaneous breakoff events and local/global identity assignments are described in a dedicated section below.

The overall algorithm used to detect the primary breakoff and track the retraction of the filament is shown Figure 5.6. Starting from the first step in the temporal sequence, the z-location of the lower bound is determined (and the bracketed interval is defined: $z(t) = [z_U, \ldots, z_L(t)]$) and a check is performed for the occurrence of any zeros in the diameter on the interval $z(t)$; if true, the algorithm proceeds to find the z-locations associated with breakoff, tokens the transient object identities, updates the persistent object features, and then continues tracking the upper filament; and if false, the next time step in the sequence is subjected to the same procedure.
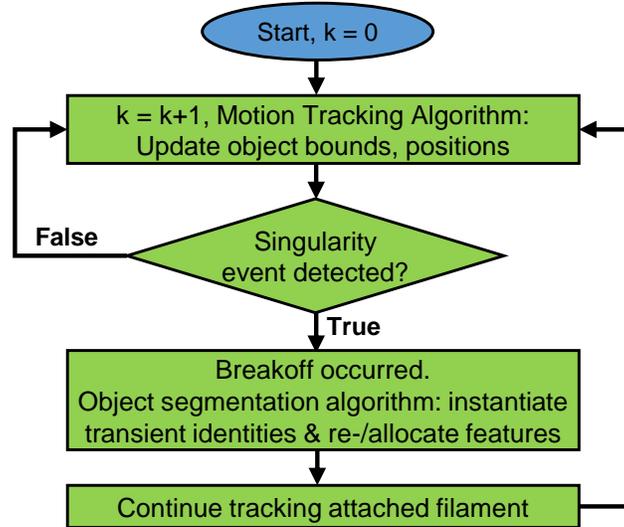
Figure 5.6. Flow chart of overall algorithm for tracking of filament shape and detachment events.

In the following sections, the individual components of the overall algorithm are presented separately to provide a clear description of each; the organization of these sections proceeds in the following sequence: bi-directional liquid surface bound search; detection of breakoff time, search for nascent liquid surface above and below break point; filament retraction (tracking of liquid surface as it ascends toward nozzle), detection of additional breakoff events and re-initialization of the search after each detachment; transient object identity assignment and feature allocation algorithm.

Bi-directional Search for z-Location of Lower Bound

To elucidate the logic applied in the bi-directional search algorithm depicted in Figure 5.7, consider the following physical situation: at a given time step, suppose that the z-location of the meniscus, z_L, and its bound z_{LB}, are known. For the discrete representation afforded by pixels (after edge detection), these are defined, respectively, as the last row occupied by the liquid surface, and the first row not occupied by the liquid surface; this convention is depicted in Figure 5.7 (at t_1). Now, one would like to determine the z-location and bound at the next time step. Since temporal inversion in the axial velocity is possible, the direction of the meniscus movement at the next time step is unknown *a priori*. Therefore, at each new time step, a logical condition must be used to first determine the direction in which to search for the new location of the bound. To this

end, two iterative methods (descending, ascending) are combined and conditional statements are used to select the search direction, which depends on the direction of movement between the current time step and the previous.
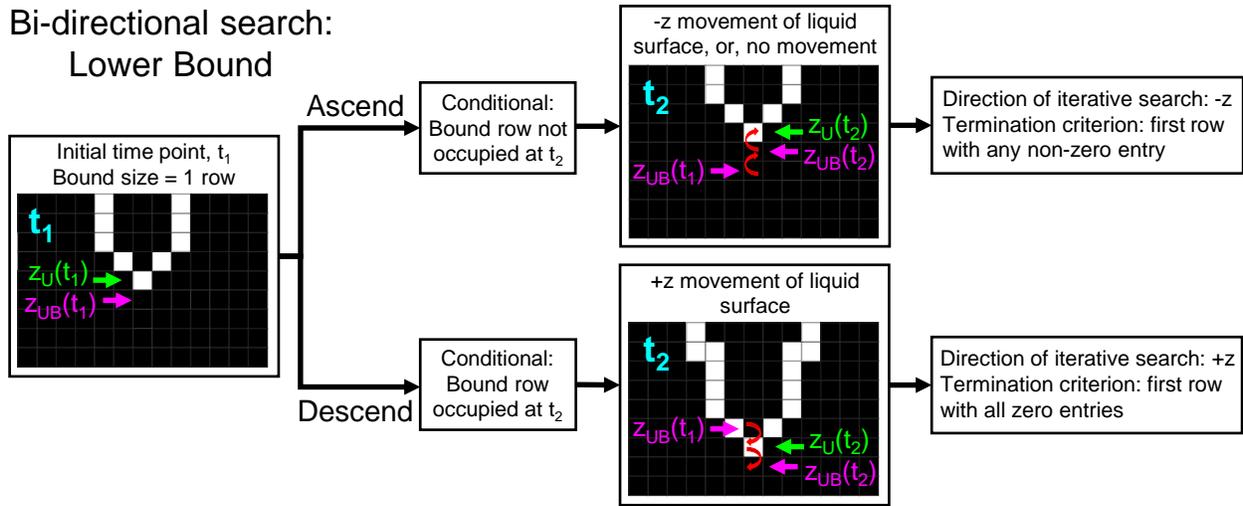
Figure 5.7. Algorithm for updating position, boundary of liquid surface at each time step. Iterations of respective searches indicated by curved red arrows.

In the descending case, one must initiate a search in the +z direction for the new lower bound at t_2, while in the ascending case, the search must be initiated in the -z direction. This necessitates a generalization that admits two mutually exclusive outcomes in order to select the correct search direction, thus, a conditional logic based on the bound, rather than on the liquid surface position, is developed. This type of construct is also useful for later methods (Section 5.5.2) which employ this type of search in identifying breakup/coalescence events.

If one considers the discrete representation of the physical situations, it becomes clear that the new state (at t_2) of the pixels located at $z_{LB}(t_1)$ – the z-location of the lower bound corresponding to t_1 – may be used to formulate the general conditions necessary to select the appropriate search direction at time step t_2. For the descending case: if at t_2 any entry of the row of the binary image at z-index equal to the lower bound at t_1 (i.e. z-index = $z_{LB}(t_1)$) is non-zero, this indicates movement in the +z direction, and an iterative search starting from $z_{LB}(t_1)$ is initiated in that direction; it terminates upon finding the first row which contains all zero entries, the z-index of which is then assigned as $z_{LB}(t_2)$, and $z_L(t_2) = z_{LB}(t_2)$-1. For the ascending case: if at t_2 all entries of the binary image at z-index equal to the lower bound from t_1 are equal to zero, this indicates movement in

the -z direction, and an iterative search starting from $z_{LB}(t_1)$ is initiated in that direction; it terminates upon finding the first row containing any non-zero entries, the z-index of which is then assigned as $z_L(t_2)$, and $z_{LB}(t_2) = z_L(t_2)+1$. The logic for the ascending case also applies when there is no movement of the liquid meniscus – the row located at $z_{LB}(t_1)$ at time t_2 would be all zero entries, therefore the search would initiate in the -z direction, but it would terminate at the first step, and yield $z_{LB}(t_2) = z_{LB}(t_1)$. Though the illustrations in Figure 7 show movement of only a few pixels, the methods are applicable to any amount of movement given the iterative nature of the searches once a direction is determined. This type of search method is in fact highly efficient for finding the liquid surface location at each time step, as the number of iterations required by the algorithm is equal to the movement of the liquid surface – it is exactly equal in the descending case, and only requires 1 extra step in the ascending case – and each iteration searches at most 1 row of the edge map.

Identifying Time Step of Breakoff; Search for Nascent Liquid Boundaries

By applying the bi-directional search at each time step, the respective $z_L(t)$ is determined, which in turn re-defines the vector $\boldsymbol{d}(t, \boldsymbol{z}(t)) = \boldsymbol{d}(t, [z_U, \dots, z_L(t)])$; the next step in the overall algorithm performs a query of this vector for the presence of any zero entries before proceeding to the next edge map in the sequence. When the time step at which diameter goes to zero is identified, a search is performed to determine the z-locations of the breakoff, which are the nascent liquid surface boundaries within the interval $[z_U, \dots, z_L(t)]$. The logic for finding these locations on the still-attached filament and now-freed droplet is developed in this section using the case for breakoff of a single drop.

The physical reasoning is simple given the manner in which one defined the interval of the persistent object, $[z_U, \dots, z_L(t)]$, the time of breakoff, and corresponding $\boldsymbol{d}(t, [z_U, \dots, z_L(t)])$: the new lower bound (which is also the upper breakoff z-location) of the still-attached liquid thread must necessarily be the z-location of the first zero entry of $\boldsymbol{d}(t, \boldsymbol{z}(t))$ found by an iterative search which starts at the top (z_U) and proceeds downward; the converse must be true for the lower breakoff point of the newly detached liquid volume, given an iterative search which starts at the bottom (z_L) and proceeds upward to the first zero entry. The result obtained by applying this method to the edge map identified as the time step of the first breakoff event is shown in Figure 5.8.
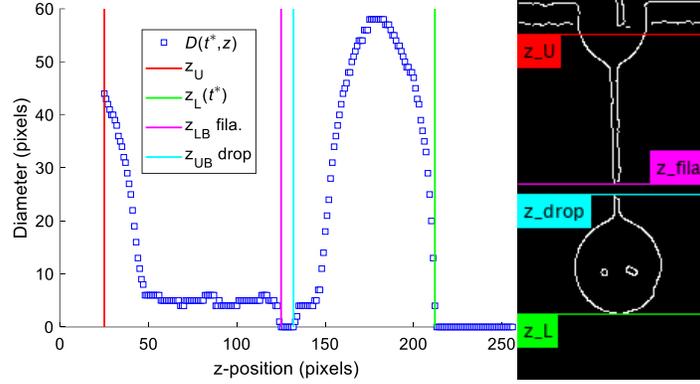
Figure 5.8. Diameter (Left) and corresponding edge map (Right) at time of first breakoff event. Filament interval, $[z_U, \ldots, z_L(t)]$, marked at left/ right bounds with red/green lines, respectively; nascent liquid surface bounds on filament and newly detached droplet marked with magenta/ cyan, respectively.

Re-initialization and Tracking of Liquid Filament

After the breakoff time and associated z-locations are found, one must track the position of the still-attached liquid thread as it retracts toward the nozzle while accounting for the possibility of additional detachment events. The algorithm used to track the lower bound of the liquid thread is the same bi-directional search, which is re-initialized at t* (= time of breakoff) with $z_{LB}(t^*)$ equal to the updated z-position of the still-attached thread (the upper breakoff location) determined from the nascent liquid surface searches – this would be the z-location of the magenta line in Figure 5.8; the new interval $[z_U, \ldots, z_L(t)]$ is depicted in Figure 5.9. The overall algorithm then proceeds through the temporal sequence, updating $z_L(t)$ at each time step. Additional breakoff events are detected in the same manner as for the first breakoff: at each time step, if any entry of the vector $\boldsymbol{d}(t, [z_U, \ldots, z_L(t)])$ is equal to zero, then a singularity event has occurred, therefore the nascent liquid surfaces must be found and the algorithm re-initialized at the boundary of the liquid surface which is attached to the nozzle; in the implementation, this is algorithm is performed as a loop over all edge maps in the sequence.
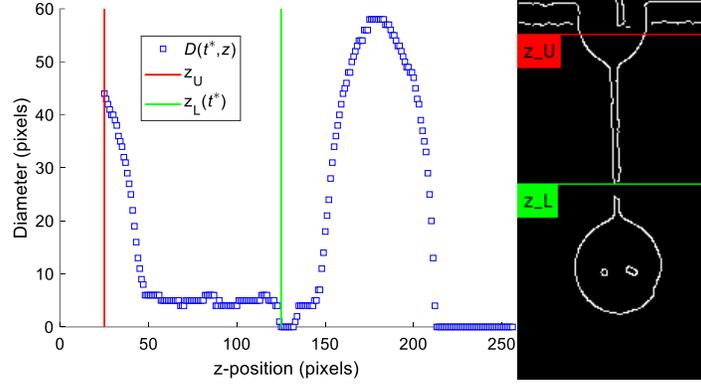
Figure 5.9. Time of first breakoff event. Interval, $[z_U, \ldots, z_L(t)]$, used to re-initiate search through temporal sequence, marked at left/right bounds with red/green lines, respectively.

Transient Object Identity Assignment and Feature Allocation: Object Segmentation Algorithm

Each time a singularity event is detected, a transient object is generated, and the respective (x, z) positions, time step and other quantities at detachment are associated to it. As the coining of these transient identities and systematic association of information to them at their instantiation is crucial to the algorithms in Section 5.5.2, a note on the handling of temporally coincident – simultaneous – breakoff events is warranted by their occurrence in the data set. Observations such as this are caused by the finite frame rate of the imaging system: the attached liquid thread appears to break up at multiple locations in the single frame which is the transition from the penultimate time step to the singularity time. Essentially, the spatially separated breakup events occur sufficiently near to each other in time that the temporal discretization cannot resolve them; artefacts such as this would be common with lower image capture rates, and in fact cannot be categorically prevented even in ultra-high-speed setups. Thus, this type of problem must be addressed in order to attain the objectives in this work, and also to impart a very much necessary flexibility that enables use of the proposed algorithms irrespective of camera frame rate.

Consequently, for application to each detection of any zero entries in the interval of the persistent object, $\boldsymbol{d}(t, [z_U, \ldots, z_L(t)])$, a general method to determine the total number and z-locations of each object, and also to assign identities in a logically consistent manner was developed. This is performed using two separate iterative searches which begin at z_U and $z_L(t)$, respectively, nested within an outer loop, to collapse the interval at each step by a distance equal to the remaining interval length minus the dual regions of occupied/unoccupied entries – the method alternates depending on the parity of the outer iterative index, the value of which is equal

to the local identity number. The termination criteria of the inner loops, based on the reduced interval produced by the previous iteration, $\tilde{\mathbf{z}}(t)$, alternates as follows: for an odd value of the outer index, the first zero entries of $\mathbf{d}\big(t, \tilde{\mathbf{z}}(t)\big)$, and for an even value of the outer index, the first non-zero entries of $\mathbf{d}\big(t, \tilde{\mathbf{z}}(t)\big)$. When the iterations terminate, the respective stored z-position vectors $(\mathbf{z}_{top}, \mathbf{z}_{bot})$ from the two separate searches are concatenated and global identities are assigned by stepping through the respective pairs of the newly formed array which form the local identities. An example of this method is shown in Figure 5.10 for the simultaneous breakoff of three drops; Fig. 5.10(a) demonstrates the interval search, and Fig. 5.10(b) the concatenation scheme and identity assignments.
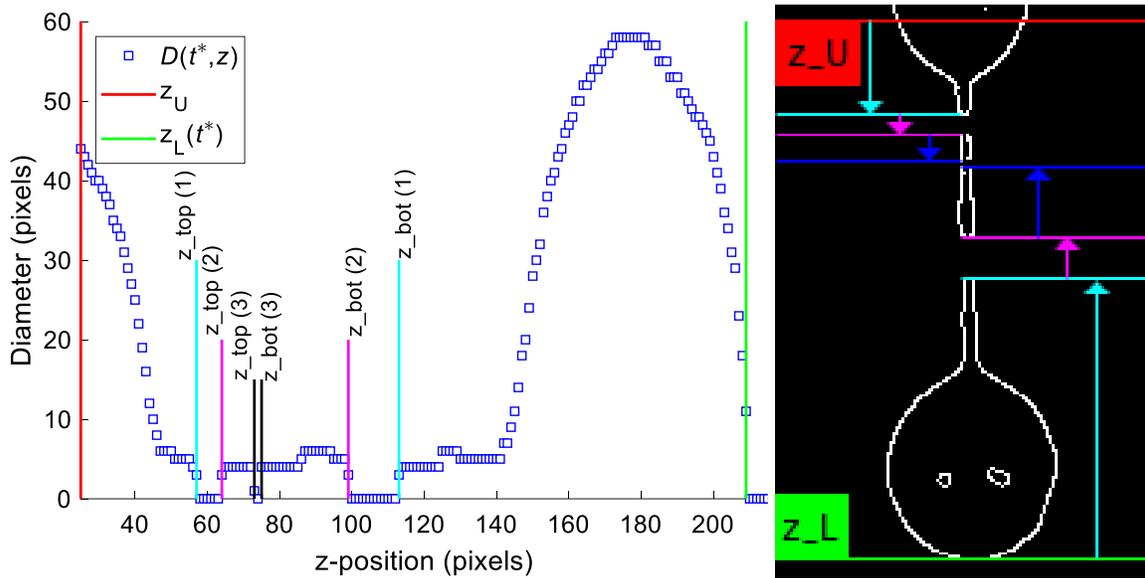
Figure 5.10. (a) Simultaneous breakoff of 3 drops (different repetition than all other figures in Section 4.1/4.2, but same frame capture rate). The cyan, magenta and black vertical lines mark the z-positions determined on the first, second and third iterations which collapse the interval. Numbering upward from $z_L(t)$, the local drop identities are 1, 2, 3; any subsequent breakoff events would start at global identity 4. The same positions are marked on the corresponding edge map; black lines in the plot are recolored blue in the edge map.
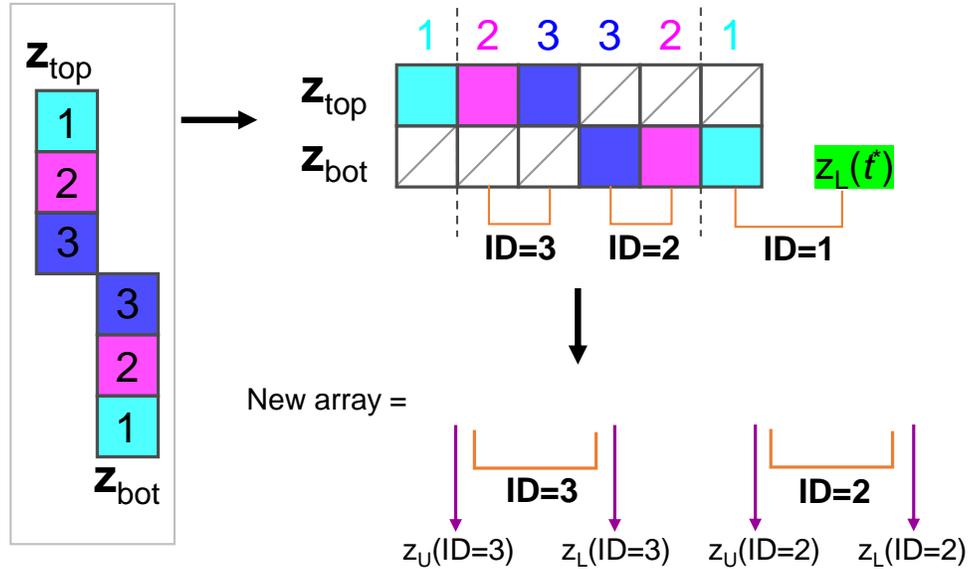
Figure 5.10. (b) The stored vectors, $\mathbf{z}_{top}$ and $\mathbf{z}_{bot}$, each have three entries which correspond to the number of iterations performed by the algorithm; note the correspondence between iteration number and the respective z-positions marked on the edge map. The assignment of identities proceeds by concatenation of the stored vectors as shown above, then looping (decrementing) from the final value of iterative index of the collapsing interval search until the local identity ID=2 is reached. Assignment of local ID=1 is always the same irrespective of the total number of drops: $z_U = z_{bot}(1)$ and $z_L = z_L(t)$.

This scheme provides a coherent method for object recognition and identity assignment in simultaneous breakoff which is completely general in the number of drops. The algorithm continues iterations until all object features have been recognized by determining the structure of the singularity; correspondingly, the assignment of these features to distinct drop identities through concatenation of the arrays generated by the search guarantees that the identity numbers are consistent regardless of the local or global detachment count. To demonstrate this, examples for the simultaneous breakoff of 4, 5,… drops are provided in the Appendix H. It must be noted that the convention enacted here is to designate the drop farthest from the nozzle as local ID=1, though this could just as easily be inverted and the drop closest to the nozzle assigned as local ID=1; use of the former is simply a convenience in this gravity-oriented reference frame.

Results Obtained from Overall Algorithm

The results presented in this section were obtained by applying the combined algorithms to an image sequence in which three drops detach (non-simultaneous) after a single liquid ejection.

In Figure 5.11, the lower bound of the attached liquid filament is shown for the entire time interval associated with drop formation. Note that for clarity of presentation in the plot (Figure 5.11, Left), only the time of the primary breakoff, for which the associated algorithm results were provided in Fig. 5.8 & 5.9, is marked so as to avoid crowding the time interval (frames 384-402) during which the filament retracts and the extra breakoff events occur. and filament retraction (frame 384-402). In the video (Figure 5.11, Right) of the corresponding edge map sequence, $z_L(t)$ is marked by a green horizontal line on each image frame; at each breakoff event, the algorithm correctly determines the appropriate position on the upper filament, as indicated by the shifts in z_L.

The events after the first breakoff are of particular interest for demonstrating the progression of the algorithm, thus, a closer look at the relevant time interval, frames 384-402, is provided in Figure 5.12. The respective green and red lines at frame 385 and frame 389 indicate the detection of two additional detachment events which occurred as the filament retracted after breakoff of the primary drop. These detachment events are shown in the video (Figure 5.12, Right), which facilitates presentation of the shift of z_L at each breakoff event through reduction of the playback speed to 3 frame/sec. In addition to identifying the time steps at which breakoff occurred during retraction, the associated z-locations were automatically determined for each detachment. Analogously to Fig. 5.8 for the breakoff of the primary drop, Figures 5.13 and 5.14 present the corresponding z-intervals and nascent liquid surface bounds for the detachment of the second and third drop at time points (frame) 385 and 389, respectively.
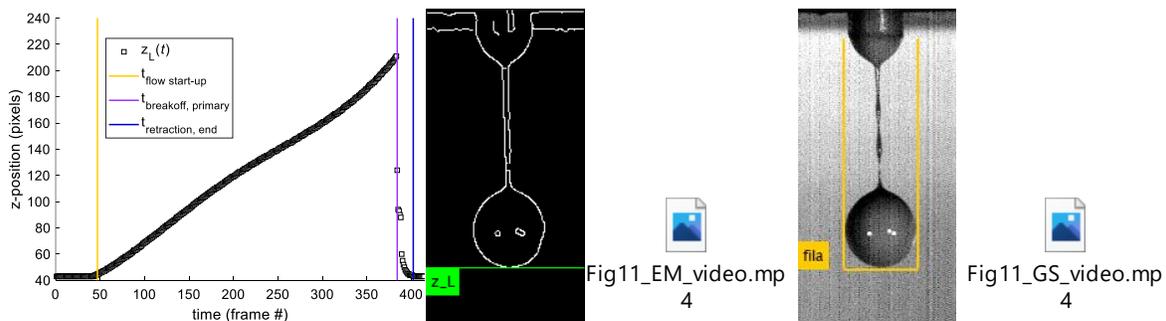
Figure 5.11. Temporal evolution of z-location of liquid filament. Left: scatter plot with times marked. Middle: visualization generated by marking the location of $z_L(t)$ on frames 1-414 of the edge map sequence; playback speed 60 frame/sec. Right: visualization generated from the temporal evolution of the two-dimensional shape of the persistent object (attached filament), rendered on the grayscale image sequence; playback speed 60 frame/sec.

111
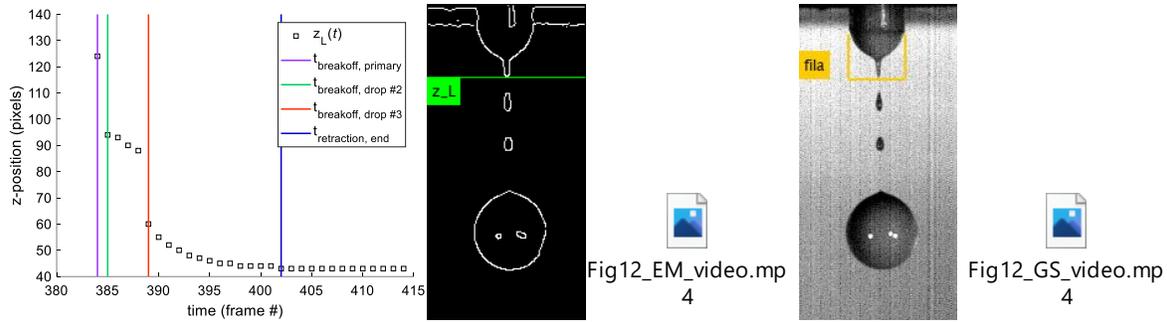
Figure 5.12. Temporal evolution of upper filament position starting from first (primary drop) breakoff. Left: scatter plot with times marked. Middle: visualization of $z_L(t)$ from edge map frames 384-402; playback speed 3 frame/sec. Right: visualization of two-dimensional shape; playback 3 fps.
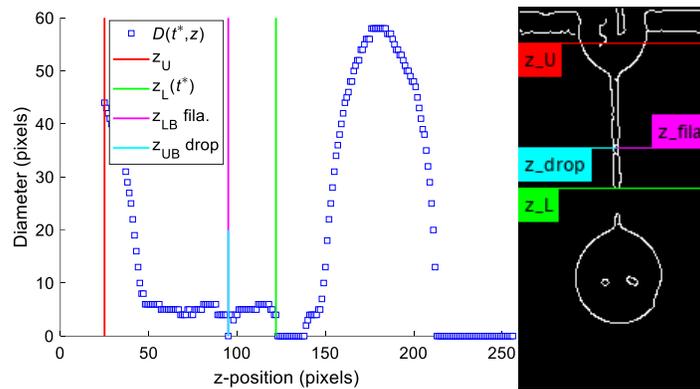
Figure 5.13. Time of second breakoff event, which occurred during retraction of liquid filament to the nozzle. Left: scatter plot of diameter as function of z. Right: corresponding binary image. In this breakoff event, the gap between the filament and the second drop is only 1 pixel, thus, the cyan/magenta markers overlap.
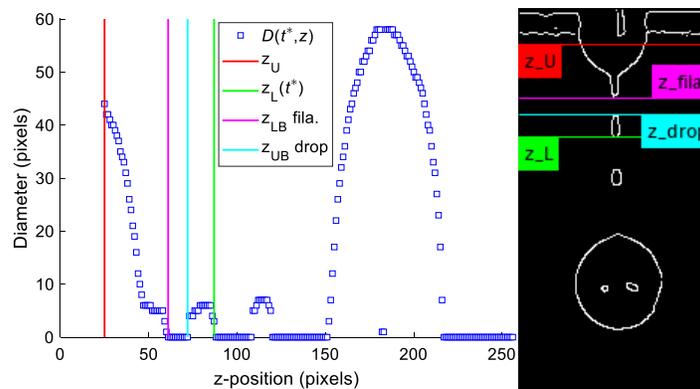
Figure 5.14. Time of third breakoff event which occurred during filament retraction; cyan label in binary image moved to right for aesthetic reasons only.

5.5.2   Object Tracking & Event Detection, Recognition: Drop Breakup & Coalescence

Motivation

         The effect of particle and fluid properties on the process do not abruptly end once the liquid volumes detach from the nozzle. In fact, important events take place once the fluid detaches, particularly if a single liquid volume undergoes fragmentation into two or more droplets, or conversely, two or more droplets merge into one (or more) drop(s). These modal changes in the process output have implications for the reproducibility of drop formation, and also drop stability, the dynamics of which are important if one wishes to use an image sensor to compute the volume of each drop, which can be a major benefit of this type of process in precision applications. The post-detachment motion of the drops is also of significance as this may influence effective drop ejection time (as defined in Appendix E), and also the reproducibility of drop deposition, which is dependent on consistent drop trajectory. Essentially, it is important to understand in quantitative terms how each object moves, deforms and potentially undergoes breakup or coalescence events after detachment from the nozzle, in order to understand particle-driven effects which impact design of ink formulations, system parameters and process sensors.

         This section presents an algorithm for tracking the post-detachment motion of each object by following the upper and lower (and left/right) liquid surface positions as it translates across the image coordinates and deforms in time. Then, the methods to detect and confirm the occurrence of breakup or coalescence events are developed. By applying the combined algorithms to an image sequence, one may systematically:

- Determine the dynamic deformations of each object, and assess when drops have stabilized; e.g. determine spatiotemporal point at which the tail of the drop retracts
    - o This can be used to adaptively specify when/where a drop volume should be evaluated, if stability is relevant to the computation of volume from the image. The dynamic information also provides useful insight into the physics, especially when coupled with event detection, thereby enabling modeling.
- Perform velocimetry for objects
    - o This is useful for assessing effects related to satellite drops, and/or non-straight drop trajectory

- Determine whether the liquid volumes that detach from the nozzle undergo additional breakup and/or coalescence events, and identify the time(s) and location(s) at which such events occur.
  o This can be used to determine the number of drops (primary or satellite) per actuation event, and also relates to drop stability – if breakup or coalescence occurs, then the participant moieties are not stable until after these events occur.

To fully characterize all the occurrences that each object undergoes after detachment, the algorithms developed in this section must extract not only the full set of positions associated with a given object at each time step, but also detect the modal changes inherent to breakup or coalescence. Detection and recognition of these events builds upon motion tracking, and ultimately leads to dynamic creation of multi-layered objects in order to tackle the complex possibilities.

Overall Algorithm

Consider a liquid fragment at the time of detachment from the nozzle: there are many possible events which this object may undergo, such as deformation (drop oscillations, retraction of own filament to the drop body), net upward or downward motion (or any combination of these distributed across some time sequence), self-breakup into two or more smaller fragments, or coalescence with one or more other drops/fragments. Any combination of these occurrences, in any temporal arrangement, is possible for each drop, therefore a general method which permits all of these phenomenological changes, but detects them only if they occur, must be developed in order to obtain the desired quantitative information. One must begin with a feature tracking algorithm that is defined for each unique object. Then, by building on this construct, one may define the circumstances of the more complex interactions, identify candidate time points (and respective z-locations), and subsequently classify the events. This section introduces the algorithms which comprise Section 5.5.2, starting with the physical reasoning associated with motion tracking in order to develop the notion of feature interactions for event detection and the subsequent classification scheme.

As a unique entity at the instant of breakoff from the nozzle, each drop has a time point and several z-locations associated to it by the drop breakoff/filament retraction algorithm from Section 5.5.1. To determine the temporal evolution of the upper/lower liquid surface positions for a reference drop, one must take into account the possible directionalities of both boundaries at each

time step, which depend not only on the net velocity of the object, but on strong dynamic deformations (e.g. retraction of its liquid filament). This can be achieved using the algorithm developed in Section 5.5.1, which performs an iterative search based on the previous time step to update the positions at the current time step; the method for the lower bound of the newly-freed liquid fragment is analogous, and for the upper bound of the fragment, the conditional logic used to initiate (and terminate) ascending/descending iterative searches need only be inverted. Implementation of this algorithm starting from the time of detachment of each drop enabled tracking of the upper/lower liquid surface positions as they evolved with time; the resultant bounds for each object at each point in time are shown in Figure 5.15, which presents the scenario in which 3 drops break off (sequentially) from the nozzle after a single actuator pulse.
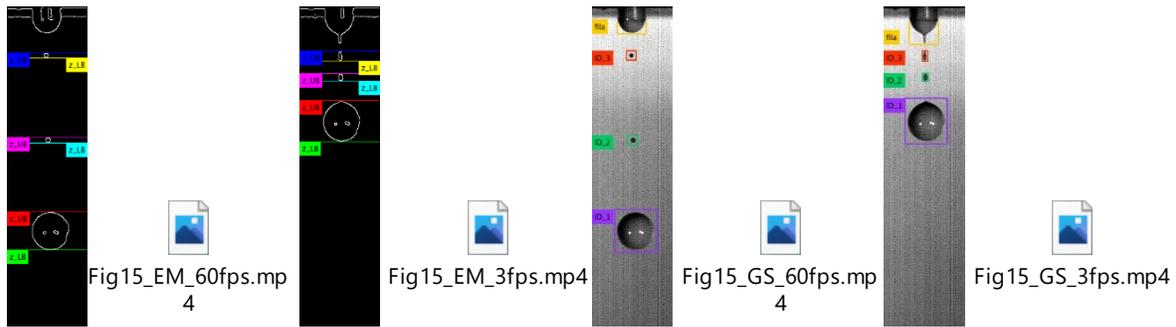
Figure 5.15. Temporal evolution of upper/lower liquid surface bounds, starting at the respective time of detachment for each drop. Left edge map: frames 350-854, 60 frame/sec playback speed. Right edge map: frames 382-395, 3 frame/sec playback rate. Grayscale images: Temporal evolution of two-dimensional shape and bound positions for each object – frames and playback rate identical to respective edge map videos.

From inspection of Figure 5.15, it is clear that fragments might interact with each other after detachment, which could lead to a coalescence event; alternately, each liquid volume could itself break up into two or more fragments. For a droplet before a potential break-up event, the physical situation is as follows: within its respective restricted interval, $\mathbf{z}(t) = [z_U(t), \ldots, z_L(t)]$, the fragment is a single entity with a continuous liquid surface. At the instant of break-up (if it occurs), a singularity forms in the previously-continuous liquid surface – this discontinuity is a location at which the diameter of the liquid volume has gone to zero. Therefore, the event may be detected by finding the time step(s) at which any entry of $\mathbf{d}(t, [z_U(t), \ldots, z_L(t)])$ is equal to zero;

this is analogous to breakoff detection described above, but with the difference that a unique interval – $z(t)$ – is associated (and updated at each time) to each droplet identity.

If a breakup does occur, the nascent liquid surface bounds can be determined by the object segmentation algorithm. Then, given the interval, $z(t)$, and time of breakup, one could either instantiate new (sub)identities and re-start the motion tracking algorithm for each drop, or continue the motion tracking as-is, with the knowledge that each subsequent time step would also be flagged as a "breakup" event due to the (intentional) persistence of the discontinuity within $z(t)$. In the implementation, both are used: sub-identities are tokened for the fragments, but the original object identity is also retained. The reason for this is that the latter option has the advantage of preserving the identity of the original object – or, to be exact, the time-indexed arrays of intervals and parameters – which was tokened at the time of detachment from the nozzle, thereby providing sufficient information to make quantitative statements such as: the 2nd drop to detach from the nozzle broke into two fragments, of volumes a and b, at location c and time d. Additionally, this convention is useful for identifying time sequences, such as that presented in Figure 5.16, during which a liquid volume breaks up, the resultant fragments spend some time apart, and thereafter coalesce to form a single drop; each time step starting at breakup and ending one step before coalescence would be recorded due to the apparent singularity exhibited by the outer object.

While advantageous for the above reasons, identification of multiple time points as "breakup" for a tracked object requires that an algorithm must be developed to classify each time step in the sequence. The physical situations contained in such a sequence supply a simple intuition for classification: enumeration of the liquid surface discontinuities within the respective interval $z(t)$ for each step in a given time sequence, followed by comparison of the total number of elements between each time step and that preceding it – if the total increases, decreases, or stays the same, then breakup, coalescence, or neither occurred, respectively. The pattern which can be used for systematic recognition of the exact time step of breakup events arises from the motion tubes. Within a given object, consider two time steps in forward order, t_k and t_{k+1} : the set of positions in the (x,z) plane that circumscribe the perimeter of the object can be represented alternately using a reduction to a set of diameters, $d(x_{relative}, z)$, in which $x_{relative}$ simply orients the diameter magnitude on the x-axis at each z-position. More details of the event classification method are provided in Section 4.2.6.

Before proceeding to the detailed descriptions of the motion tracking, breakup detection and classification methods, additional consideration must be given to the concept of fragment coalescence, and in particular to the converse of the case depicted in Figure 5.16, which is the coalescence of different droplets that are not already within $z(t)$, (i.e. droplets with distinct global identities conferred at detachment from the nozzle). To develop a criterion that may be used for identification of such coalescence events, it is helpful to examine the physical situations of breakup and of coalescence; their similarity – stylized here, but the essence is the same – becomes apparent if the temporal ordering of images associated with a breakup event is inverted, as illustrated in Figure 5.17. Consider that in droplet breakup, the interval $z(t)$ of the regular (non-sub) identity must encompass the z-location of the event, thus the occurrence of a discontinuity can be detected, and the time step flagged. Conversely, in the case of coalescence shown in Figure 5.17, there is not an apparent change in the number of liquid surface boundaries on the interval $z(t)$, nor would either time step be flagged by the motion tracking algorithm as a potential "breakup" event as none of the entries of $d(t, z(t))$ are equal to zero at any time.
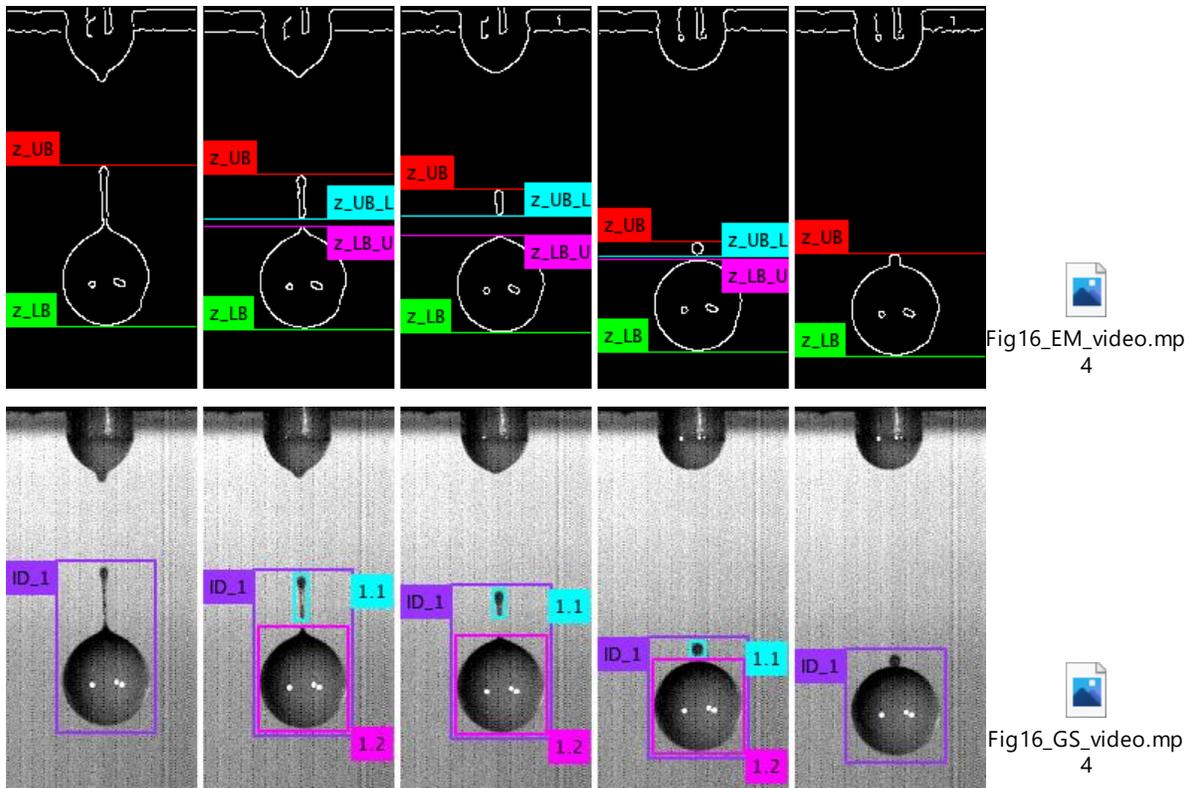
Figure 5.16. Temporal sequence of drop breakup and subsequent coalescence. Top: Edge maps, from left to right: time step before breakup detection, three intermediate flagged times, time step of coalescence. Video: sequence of edge maps, starting before breakup and ending at coalescence; 5 frame/sec playback rate. Bottom: Grayscale version of same image sequence and video, but displayed using identity-associated (main "ID_1", and sub-IDs "1.1", "1.2") information: temporal evolution of positions (in x, z) and two-dimensional shape.
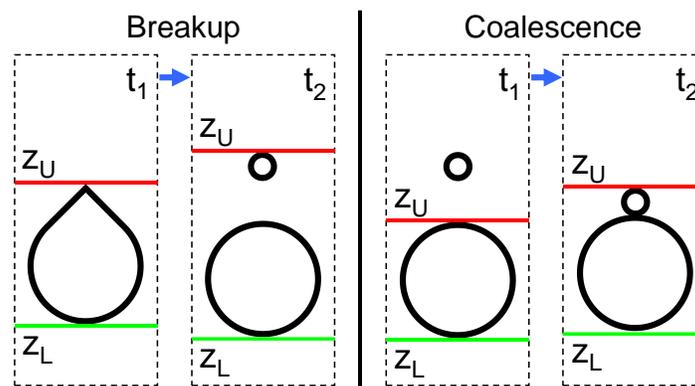
Figure 5.17. Illustration of breakup and coalescence events; note the difference in $z(t)$ between the two cases

From the restricted perspective – $z(t)$ – of the reference object, the lack of apparent changes necessitates a different approach to detection of coalescence events. Consider the time step before merging: if the interval $z(t)$ were to include the non-reference drop, then this time step would be flagged as a breakup event, and the classification algorithm could be used to resolve event type afterward. For this detection to work, the bound of the reference drop must contact the liquid surface of the non-reference drop when at least 1 empty row of pixels remains between the two drops, in other words, when at least one entry of $d(t, z(t))$ is equal to zero. However, the motion tracking based on the bi-directional search (Section 5.5.1) employs search methods for $z_U(t)$ and $z_L(t)$ that ensure only a minimum of 1 unoccupied row above/below the position of the respective liquid surface, and therefore most other objects are excluded from the interval $z(t)$ as the bounds do not contact other liquid surfaces unless they are within 1 pixel of the reference object, and only then if this physical arrangement persists for at least 2 time steps. While possible, this is not a certainty, and, in actuality, is rather unlikely, hence most coalescence events would not be flagged. The solution to this is apparent problem of detection is to expand the size of the bound to more than one row of zero pixels above/below the liquid surface positions, which effectively defines a new object size from the perspective of the motion tracking algorithm. Tuning of effective object size – bound size – is important for obvious reasons, as the bound size should be small in order to preserve object identity, but sufficiently large to make coalescence detection a certainty; a detailed treatment is provided in a subsequent section, and a rigorous expression for a minimum sufficient bound size to guarantee coalescence event detection is developed.

The overall algorithm is presented in Figure 5.18; the boxed portion is applied to each object (transient identity) that detaches from the nozzle, from the instant of detachment until the tracked object exits the image. For each object, if any time steps were flagged as potential breakup or coalescence events, the respective flagged time sequence, $t_{sequence}$, is passed to the classification algorithm.
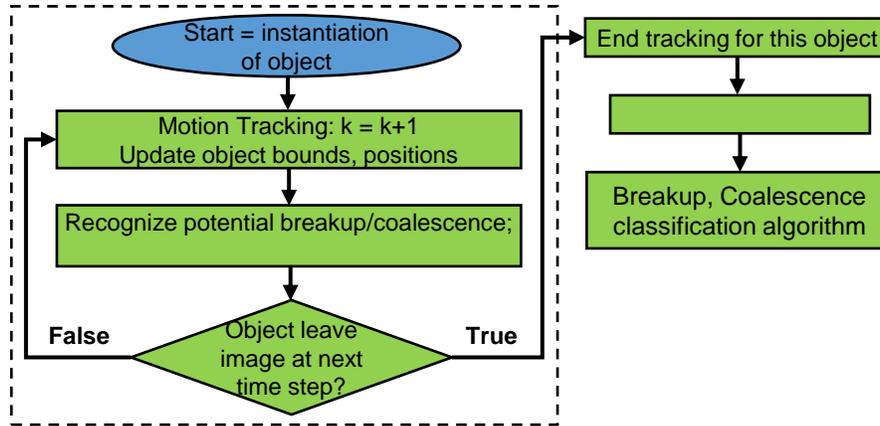
Figure 5.18. Algorithm overview for object tracking and event detection (drop breakup/coalescence).

The subsequent sections describe the components of the overall algorithm, and organization of the material is as follows: object (motion) tracking: 1- and multi-row bound methods, selection of bound region size and relevant factors, detection of breakup/coalescence; algorithm for breakup/coalescence classification. First, the generalization of the motion tracking (MT) algorithm is presented in order to demonstrate the logical structure and results. Next, it is shown that the 1-row bound bi-directional search is in fact a special case of the general multi-row bound method. Then, selection of bound size and relevant factors are considered, and a sufficient minimum bound is presented in terms of an inequality; a comparison of coalescence detection by the 1-row and n-row bound methods is provided. Thereafter, the method for classification of candidate time steps (identified by the MT algorithm) is presented.

Motion Tracking Algorithm & Generalization to Multi-Row Bounds

The object tracking algorithm was developed from a generalization of the bi-directional searches, which is applied to each object along its axes of interest to determine the positions of feature boundaries (e.g. z-axis, at upper/lower boundary). If one considers the liquid surface/bound tracking algorithm presented in Section 5.5.1., the logic applied to both initiate (select direction – ascending/descending) and terminate the searches can be generalized for any case with a bound size greater than 1 pixel = 1 row above or below the liquid surface position. For the sake of brevity and clarity of presentation, the following demonstrates the extension of the algorithm for the upper

bound of an object; the extension is also valid for the lower bound, given the necessary changes due to sign convention.

If one defines an expanded bound of an object with respect to the actual position of the liquid surface, for the upper bound it would be the set of row indices: $\{z_{UB}, z_{UB} + 1, \ldots, z_{UB} + (n-1)\}$, $n =$ (bound size) with the liquid surface position equal to the last entry of the set $+1$; the elements of the set are: $z_{UB} + (j-1)$, $j = 1, 2, \ldots, n$. In other words, the bound is now a matrix with n rows – this is shown in Figure 19 for a bound size of n = 3 – with the "z_{UB}" defined as the first entry of the ordered set of row indices. The n-row bound set in terms of the liquid surface position, z_U, is: $\{z_{UB}, z_{UB} + 1, \ldots, z_{UB} + (n-1)\} = \{z_U - n, z_U - (n-1), \ldots, z_U - 1\}$, and the elements of the set based on z_U are: $z_U - (n - (j-1))$, $j = 1, 2, \ldots, n$; for the 1-row bound this relation simplifies to $\{z_{UB}\} = \{z_U - 1\}$, or equivalently, $\{z_{UB} + 1\} = \{z_U\}$, which should make apparent the relation between the n-row and 1-row bound.

To determine the search direction for the upper bound at the next time step, the n-row segment of the edge map at row indices $\{z_{UB}, z_{UB} + 1, \ldots, z_{UB} + (n-1)\}$ is queried for the presence of any non-zero entries. If any non-zero entries are present, either the liquid surface position (of the reference drop) has ascended, or/and another drop has come sufficiently close to cause interaction by crossing the z-location at z_{UB}, which in both cases corresponds to an iterative search in the upward direction. Iterations terminate at the first occurrence of n all-zero rows, the row indices of which constitute the updated bound set, and which must be situated immediately above the new liquid surface position, therefore $z_{UB} - n = z_U$. This inference is the logical consequence of the following: at the penultimate iteration, only the row corresponding to the index of last element of the bound set could have contained a non-zero entry, thus, at the final iteration, the liquid surface position must be the row located 1 pixel below the row index of the last entry of the (new) bound set. As an aside, note that for cases in which another drop has crossed z_{UB}, the algorithm would iterate upward until the liquid surface position at the top of this object is located; this annexation is the desired effect, as it causes the time step to be flagged as a potential coalescence (breakup) event by virtue of the apparent singularity.
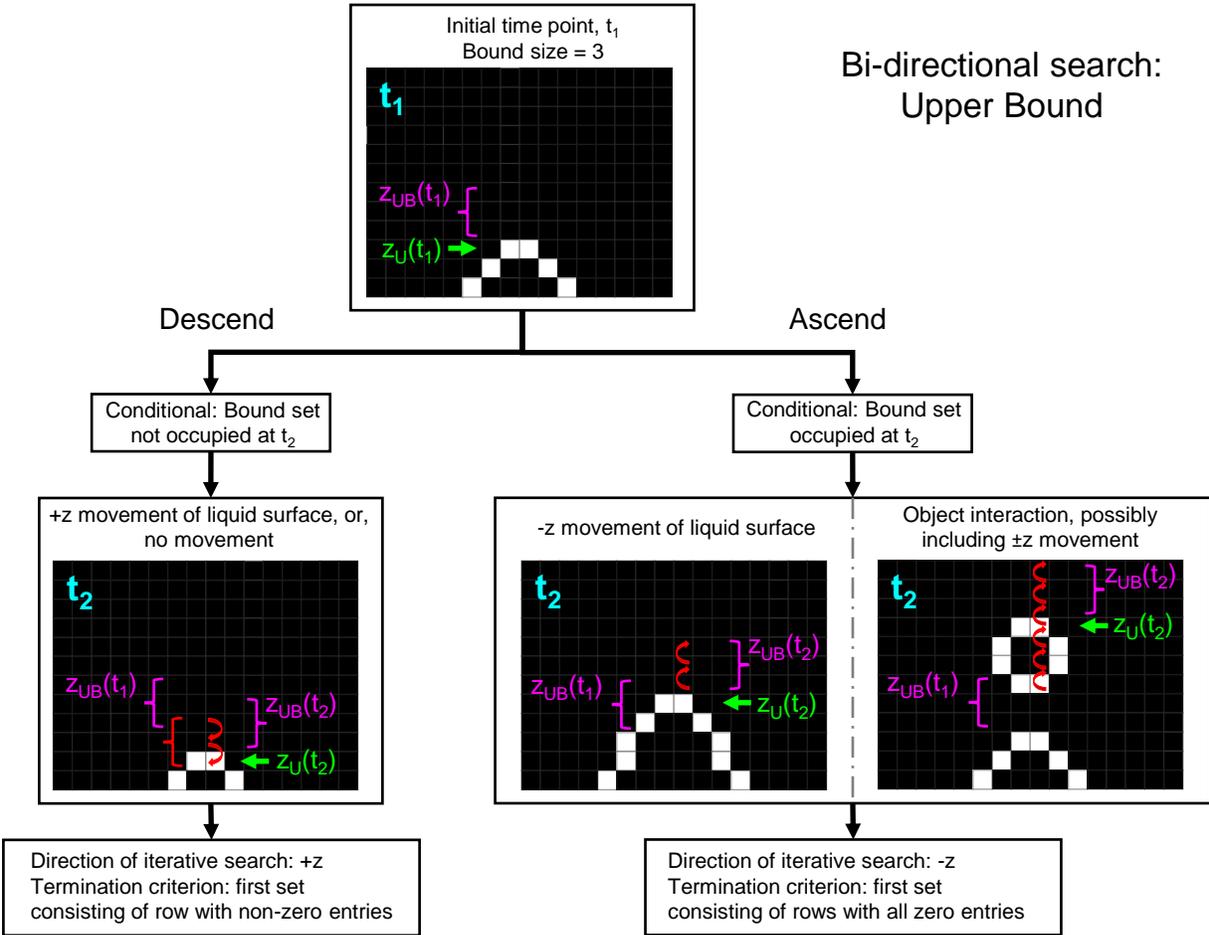
Figure 5.19. Object tracking algorithm with multi-row bound segment, shown for cases of the upper bound. Cases for the lower bound are analogous, given the conditional logic and bound set changes described in this section, with the only noteworthy difference being that the "object interaction" would fall under the respective "descend" branch of a similarly constructed figure

The other possibility admitted by the query of the binary image – all entries of the n rows are zero – indicates that either the liquid surface position (of the reference drop) has descended, or has remained stationary. This corresponds to an iterative search in the downward direction which terminates at the first occurrence of n rows that contain a non-zero entry. At the final iteration, the only row index with a non-zero entry must be the last element of the bound set, which is the new position of the liquid surface; this follows from the nature of the iterative search, as the leading row – the last element of the bound set – would be the first (and only) row to encounter a non-zero entry. The updated bound set at the final iteration is inferred directly from the entries of the bound set at the penultimate iteration, which are the n rows immediately above the updated position of

the liquid surface; these rows are necessarily all-zero as the penultimate iteration would otherwise have been the final iteration.

As mentioned above, the multi-row algorithm for the lower bound is analogous to the upper bound, but necessitates the following changes to account for the difference in sign convention: the set of row indices of the bound is inverted: $\{z_{LB} - (n-1), z_{LB} - (n-2), ..., z_{LB}\} = \{z_L + 1, z_L + 2 ..., z_L + n\}$, such that the first element of the set abuts the liquid surface position, whereas for the upper bound this was the last element. The conditionals used to determine the search direction and respective termination criteria are also inverted: if non-zero entries are present at the row indices of the bound set, the liquid surface position has descended, or another drop has crossed z_{LB}, which in both cases corresponds to an iterative search in the downward direction that terminates at the first occurrence of n all-zero rows (spatially inverted analogy to Figure 19, "Ascend"). If all entries of the n rows are zero, the liquid surface position has ascended, or remained stationary, which corresponds to an iterative search in the upward direction for the first occurrence of n rows with a non-zero entry, and the elements of the updated bound set are inferred from the penultimate iteration (spatially inverted analogy to Figure 19, "Descend").

The algorithm in Section 5.5.1. is a special case of the generalized method presented in this section, and can be recovered by noting that for a bound size = 1, the upper bound set simplifies to a single element: $\{z_{UB}\} = \{z_U - 1\}$. Thus, in the ascending case of the upper bound search, the algorithm finds at the terminal iteration the first all-zero row (updated bound), and directly infers the new position of the liquid surface as 1 row below the updated bound, such that $z_U = z_{UB}+1$; in the descending case the algorithm finds at the terminal iteration the first non-zero row (updated position of liquid surface), and directly infers the location of the new bound as 1 row above the updated liquid surface position, such that $z_{UB} = z_U - 1$. These inferences follow from the same logical consequences as described for the general cases (bound size > 1) by noting that in the special case of bound size = 1, the only element of the bound set is the "last" element, therefore the inferred locations necessarily occur at the iteration immediately preceding termination of the respective iterative search. This demonstrates that in the special case and for the general case of this algorithm, finding either the bound (ascend, upper bound) or actual position of the liquid surface (descend, upper bound) is equivalent to finding both, as the result always depends directly on the actual position of the liquid surface, which is verified as such by the last element of the set

at either the penultimate iteration in the ascending case, or at the terminal iteration in the descending case.

Event Detection: Breakup/Coalescence as Interactions

As the motion tracking algorithm identifies the actual liquid surface positions irrespective of bound size, detection of interactions (or breakup) is as follows: at each time step, after the bounds and positions are updated, $\boldsymbol{d}(t, \boldsymbol{z}(t)) = \boldsymbol{d}(t, [z_U(t), \dots, z_L(t)])$ is queried for the presence of zero entries. If the vector contains a zero entry, the time is flagged as a potential interaction or breakup event, and the positions of the nascent liquid surface discontinuities on the interval are found using the iterative methods of the object segmentation algorithm. The motion tracking then continues through the temporal sequence, each step of which would also be flagged unless a merging event occurs, or one of the objects leaves the image (which has a similar appearance in terms of the number of liquid surface boundaries); the algorithm continues until all (sub-)objects in the tracked interval leave the image area. In this context, the motion of an object interval, with a unique identity coined at the instant of detachment, is tracked for its entire lifespan, which does not end until all objects within it exit the image. The times flagged as potential interaction or breakup events are then passed to the classification algorithm.

Tuning of Bound Size for Coalescence Detection

As a parameter in the motion tracking algorithm, the bound size – the unoccupied distance around the liquid surface positions – is critically important for the detection of coalescence events. Recall that utilization of the breakup detection as a means for identifying coalescence relies on the annexation of the new object to the z-interval of the reference object, with at least one time step at which there is still an entry of the vector $\boldsymbol{d}(t, \boldsymbol{z}(t))$ equal to zero. In this context, the time steps preceding the actual merging event may be viewed as an interaction between the two drops, defined for the reference object as the crossing of its bound by the liquid surface of another object. Note that the objects need not undergo a coalescence event if they interact, but if they coalesce, they must interact; this necessary condition is exploited to detect the time steps before coalescence of two drops with distinct identities tokened by separate detachment.

Given an interaction defined in this manner, selection of bound size determines accuracy and sensitivity of detection. Increased bounds improve sensitivity, but if excessively large, diminish the meaning of interaction detection through inappropriate (or substantially premature) annexation of objects which should rightfully have been considered separate entities. Therefore, in order to accurately preserve the identity structure of objects, the bound size should be small, but sufficiently sensitive to detect interactions at least one time step prior to a coalescence event. The nature of the bi-directional bound searches – the object (motion) tracking algorithm – dictates that the relative movement of the two objects in a single time step constitutes the physical basis for selection of a minimum bound size. For the discrete representation of the physical situation, consider Figure 20, which depicts the coalescence of two drops; in one case, the bound is too small relative to the object movement to detect the interaction before it occurs.
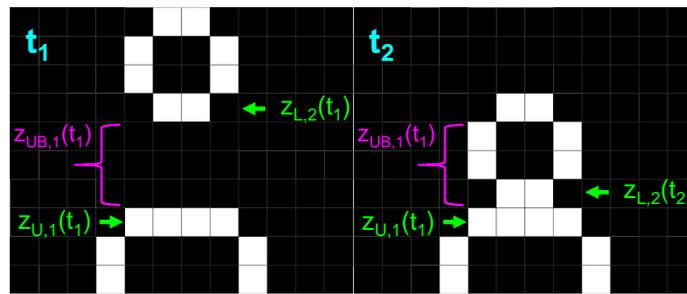
Figure 5.20. (a) Coalescence of drops with bound size (3) too small relative to the object movement – change in $z_{L,2}$, or equivalently, $v_{L,2}\Delta t$ – to detect the interaction before it occurs; the 1, 2 subscripts on z correspond to the reference and non-reference drops, respectively.
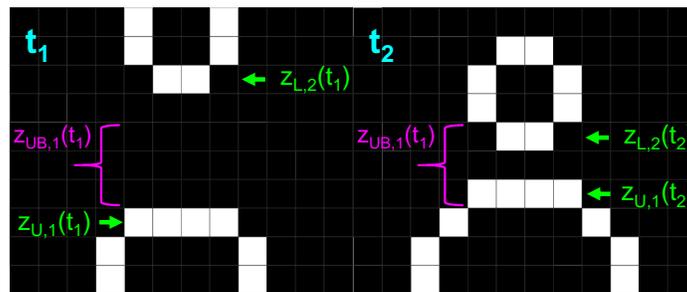
Figure 5.20. (b) Different scenario of coalescence in which the positions of both drops change. The interaction is detected with the same bound size (3) due to the relative positions of the drops, which leaves at least one empty row between the drops.

As demonstrated by Figure 5.20, the necessary bound size to detect the time step before coalescence depends on object velocity, which is dynamically changing in both sign and magnitude. However, in general, object velocity may not be assumed to be *a priori* knowledge at any time point for the following reason: to know (or estimate) the velocity of both drops presupposes that one knows *which* (or *if*) two drops coalesce, and this is simply not possible, given that this is the result to be obtained by the algorithm itself. Therefore, one would like to determine a minimum sufficient bound size using only the (maximum) relative movement and respective positions of the liquid surfaces, the nature of the motion tracking algorithm (the necessary condition for an interaction flag), and the criterion (another necessary condition) that the number of rows of unoccupied pixels between two objects must be at least 1 in order to flag the time step as a potential interaction.

From these considerations, a mathematical expression may be developed for the minimum bound size based on the sufficiency condition which results from the intersection of the two necessary conditions, as demonstrated in Appendix G. The resultant expression (Eq. (5) in Appendix G) relies on velocity and position information which one is not privy to, therefore application of physical reasoning to several limiting cases of coalescence is used to develop an expression, presented in Eq. (10), which eliminates dependence on local position and velocity.

$$BR > 1 + 2\text{ceil}(v_{max}) \tag{10}$$

This expression can always be evaluated prior to when it is needed for analysis of an image sequence, as the maximum velocity estimate (v_{max}) comes from the velocity at ejection of the liquid. Rigorous descriptions of the development of the sufficiency condition, limiting cases, and mathematical expressions are provided in Appendix G, which includes a real example that demonstrates the importance of bound size for detection of coalescence. The minimum sufficient bound size to guarantee interaction detection for the 30 image sequences presented in this work was BR = 4, given a $v_{max} \approx 0.6$.

A final comment on the use of the expanded bound region is warranted by the potential errors it may introduce. Ideally, the expanded bound would be applied at each step of object tracking, however, at the instant of drop detachment, the empty region between the still-attached liquid thread and the newly-freed drop may be fewer rows than the bound, therefore the filament would be inappropriately included in the z-interval. This scenario is easily avoided by use of a reduced bound size for the initial time step(s) of the motion tracking algorithm, and in the

implementation the bound region is set to 1-pixel row for the first three time steps, after which it reverts to a size of 4-pixel rows. Alternatively, bound size at the initial time steps could be a function of time and/or (an approximation to) velocity, scaling with time until the desired bound size is reached. As a related matter, an adjustment to the bound size must be made when the drops leave the image at the bottom (z=512), as it is not meaningful to apply a bound which extends outside the remaining image area; this can be dealt with by reverting to a 1-pixel bound size when the lower liquid surface position exceeds 512-(BR+1).

Uncertainty arising from the discrete representation of object motion is incidentally accounted for by using the floor/ceiling of the maximum velocity estimates. To provide a full accounting of the uncertainty introduced by the imaging setup requires a re-examination of the local noise for each region in the image, and is beyond the scope of this work. However, it should be noted that noise can be adaptively accounted for during feature detection, as shown in the edge detector here; as noted above, modeling of noise and scale-space approaches can be used to further refine such algorithms.

Event Recognition: Drop Breakup and/or Coalescence

Detection of potential breakup/coalescence events by the motion tracking algorithm produces time intervals, specifically associated to each object identity, in which the constituent time steps must be classified. Given the physical situations and the available information, a method based on enumeration of the liquid surface discontinuities and subsequent comparison of the totals between time steps is sufficient for the task, provided that some minor aspects are accounted for.

The z-interval used for enumeration of the discontinuities could be either the entire z-interval below the nozzle, or, the specific $z(t)$ for each drop at each time. While the former is attractive for its generality, it complicates determination of which liquid surface discontinuity has appeared/disappeared and which respective fragment this change belongs to; this is especially problematic for distinct temporal sequences which share as entries a common time step. Therefore, it is preferable to use the identity-associated $z(t)$'s for enumeration, as restriction to distinct spatial intervals removes any possible ambiguity.

The time interval that the classification algorithm acts on is also of importance, given the way in which time steps are added to the respective sequence for each droplet identity (if indeed any are recorded at all) by the object tracking algorithm. As only the times flagged as exhibiting a

singularity – any $\boldsymbol{d}(t, [z_U(t), \ldots, z_L(t)]) = 0$ – are recorded, the resultant time sequence, $[t_m, \ldots, t_n]$, would appear to indicate neither breakup nor coalescence events, as comparisons of the total liquid surface discontinuities would indicate no changes. Consequently, it is necessary to include the time points immediately before and after the interval in order for the classification scheme to work appropriately; thus, the interval passed to the algorithm is always: $[t_{m-1}, [t_m, \ldots, t_n], t_{n+1}]$.

With these aspects determined, enumeration of the liquid surface discontinuities on the interval $\boldsymbol{z}(t) = [z_{UB}(t), \ldots, z_{LB}(t)]$ proceeds by finding all their z-locations in a manner similar to that described in Section 5.5.1. for simultaneous breakoff; this is repeated for all entries in the time interval for a given object identity. Then, the classification structure compares the number of discontinuities at the current step (t_m) to the total from the previous step (t_{m-1}), and binary entries are assigned to a matrix of indicator variables at the respective column index of t_m. The logic used to assign values for time steps $t_m, \ldots, t_{n+1}$ is as follows: if the total at the current time step is greater than at the previous time step, then the binary indicator for breakup is assigned a value of 1; alternately, if the total is less than at the previous time, then the binary indicator for coalescence is assigned a value of 1; if the totals are equal, then zero is assigned to both indicators. Storage of the binary indicator variables in this manner enables convenient retrieval of times for coalescence/breakup within each object identity, and is particularly useful when an object undergoes multiple events within a single time sequence. The result obtained from this method is presented in Figure 5.21 for the same image sequence as shown in Figure 5.16.

Entry of time interval	t_{m-1} $= 459$	t_m $= 460$	t_{m+1} $= 461$	$t_{m+2}, ...,$ t_{n-2}	t_{n-1} $= 478$	t_n $= 479$	t_{n+1} $= 480$
Number of discontinuities	2	4	4	4, …, 4	4	4	2
Breakup indicator	0	1	0	0	0	0	0
Coalescence indicator	0	0	0	0	0	0	1
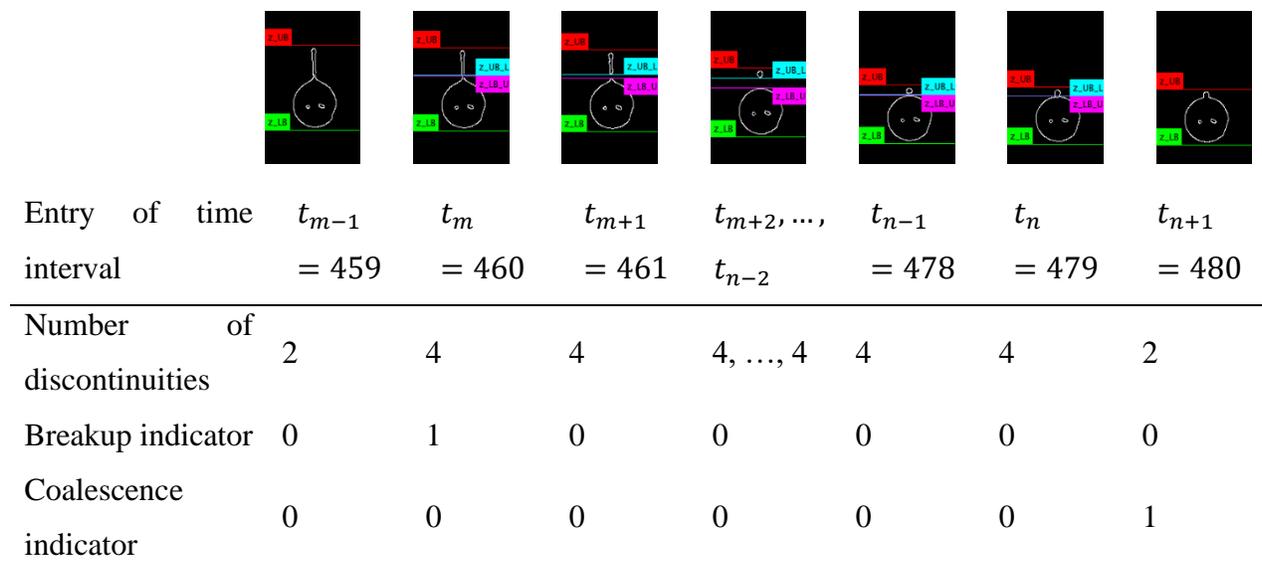
Figure 5.21. Event classification result for time sequence which includes breakup and coalescence

5.6   Results and Discussion

Some results obtained by applying the algorithm to 30 image sequences are presented below. There were 21,504 images in total (2.7GB total), with the number of images per sequence ranging from 600-950. The total time required to process the 30 image sequences, which included loading from hard disk, edge detection, analysis (Section 4.1 & 4.2), and saving of results, was 110-115 seconds in MATLAB 2019a, run on a standard desktop computer.

The intent of this section is to demonstrate the capability of the program to extract information, and only to note the presence of interesting trends and/or observations related to particles, details of which are of interest in examining jet breakup and drop formation/fragmentation. It is important to consider when examining the figures in this section that the plotted parameters are only a small sample, primarily focused on the modalities, of the total extracted information.

The image sets were gathered using the setup described in Section 2.1: repeated printing (in drop-on-demand mode) of drops under the same operating conditions from a concentrated slurry of non-ideal particles. Specifically, the particle concentration was 20 vol%, the particle type was micronized paracetamol, and the carrier liquid was NEOBEE®895; additional details on the particle/suspension properties, preparation and characterization may be found in Radcliffe and

Reklaitis (2019). The dimensions of the nozzle were: outer diameter, 2.05 millimetres, and inner diameter, 1.15 millimetres; the parameters of the IVEK pump were: nominal displacement, 1.5, and cycle speed on actuation, 1440 RPM, which yielded an approximate emission period of 21 milliseconds.

5.6.1   Drop Detachment: Time(s), Location(s), Number of Events

Results pertaining to the breakoff events – drop detachment – from the nozzle, as determined by the algorithms in Section 5.5.1., are presented in this section. Within each image sequence, the number of identities coined by the algorithm is exactly equal to the number of drops that detach, which therefore enables the clear association of time and location(s) of each event to the corresponding drop; this permits the succinct presentation of salient data in the following figures. It must be noted that as presented below, the times are relative to the start-up of flow (defined as the instant at which the meniscus first moves; t_{start} from Appendix D), whereas the spatial locations are in the absolute (non-relative) image coordinates due to the constant location of the nozzle/meniscus starting position across all image sequences; it is straightforward (and perhaps helpful) to change any of these to relative locations. Due to the modalities of the data, the first/second/third/… occurrence of an event is indicated by the presence (and color) of the data entries for that experimental repetition.
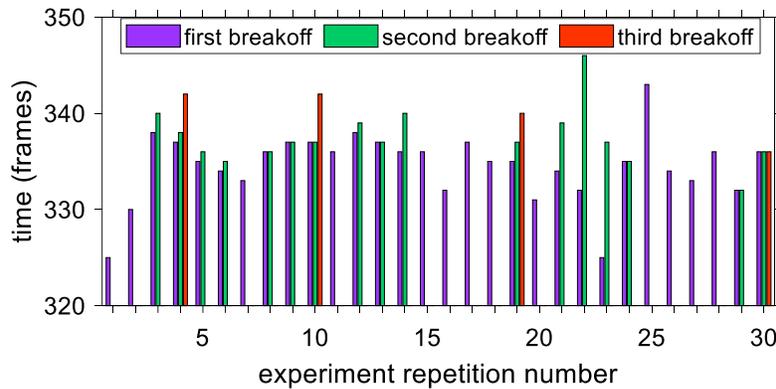
Figure 5.22. Time of drop detachment(s) in each image sequence; note both the variance in time of detachment, in addition to the variability in number of drops formed per liquid ejection.
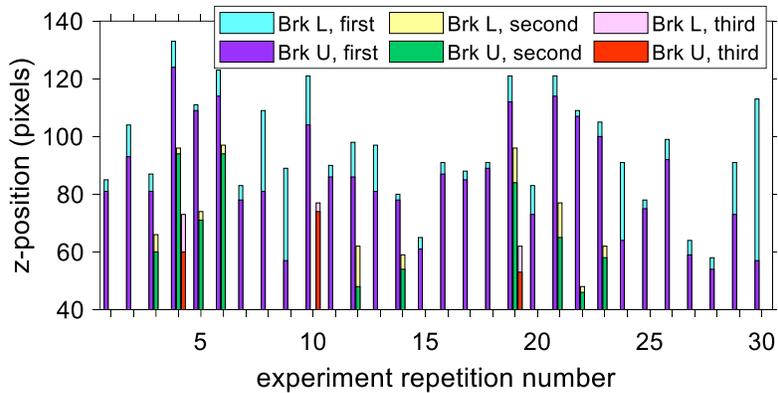
Figure 5.23. BrK U: Upper location of singularity at time of breakoff (physical correspondence: magenta-colored "z_fila." line in Fig. 8), respective to the detachment identity. Brk L: Lower location of singularity at time of breakoff (physical correspondence: cyan-colored "z_drop" line in Fig. 8), respective to detachment identity.

Figure 5.22 explicitly contains the time(s) of breakoff in each repetition; the number of transient object identities coined for each is equal to the number of bars present at each x-axis tick mark (up to 3, which happened to be the largest number of drops to detach), therefore the number of breakoff events is shown implicitly – if the height of the bars are equal for a given experiment, this indicates that a simultaneous breakoff event has occurred at that time. Figure 5.23 demonstrates the significant variance in the locations of the singularity at the instant of breakoff. For the simultaneous breakoff events (repetitions 8, 9, 10, 13, 24, 29, & 30), the absence of bars in Figure 23 for the second and third (30 only) objects is intended as these identities are considered by the algorithm to be part of the same global identity which is coined at each breakoff event. Repetition 10 (corresponding video found in Figure 5.1, Middle) demonstrates the difference between the global breakoff event identity and individual object identities: two drops break off simultaneously (1st global breakoff identity, two object identities), then later a third drop breaks off (2nd global breakoff identity, third object identity).
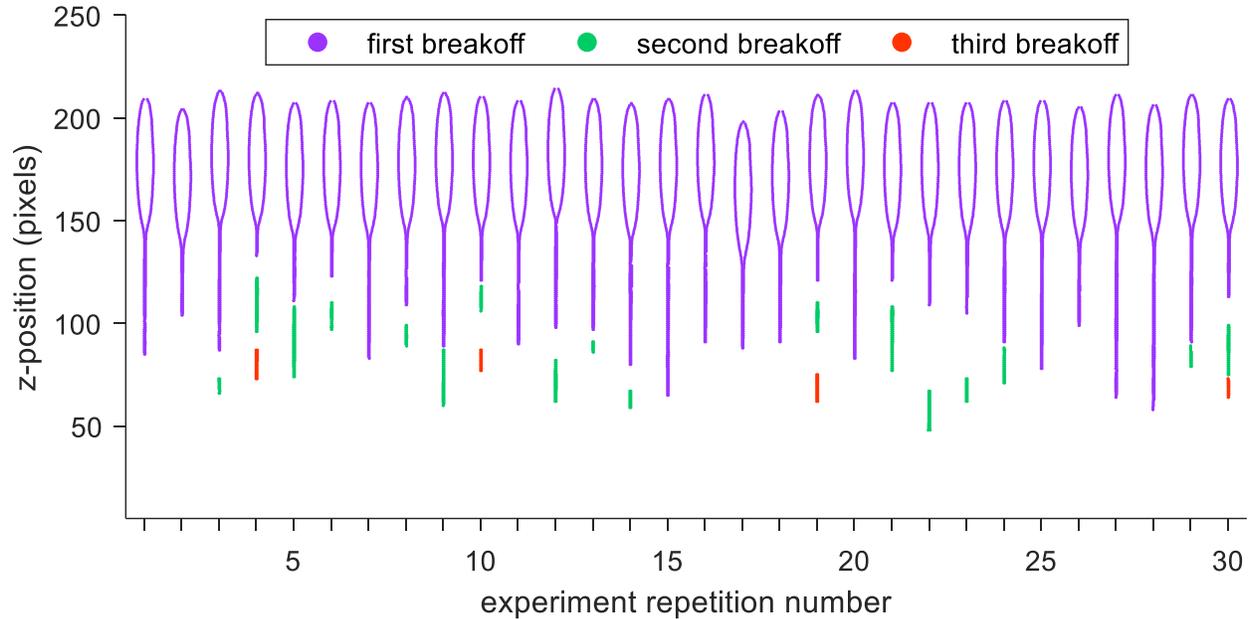
Figure 5.24. Total positions (x, z) of each drop at its respective time of breakoff

In Figure 5.24, for each experimental repetition, all positions (*x*, *z*) of each drop identity (coined) at its instant of detachment are presented by overlaying the respective locations on to a single z-axis. This enables a direct visualization of the upper and lower z-positions (also: proximity) at each detachment event, and the two-dimensional shape of each object at its instantiation. Readily observable from this figure is the apparent consistency in the size of the body of the first drop, which contrasts directly with the variability in length of the thin filament attached to said body. As an accompaniment, the volumes of each drop at the instant of detachment are presented in Appendix F, Figure F.2.

5.6.2   Filament Retraction After Breakoff

In this section, the dynamics of the filament after detachment are considered. Essentially, this information must be used to decide when the next drop could be ejected – in other words, the minimum interval between actuator pressure pulses. An accounting of the total drop ejection dynamics must consider not only the filament through drop detachment and recovery therefrom (Figure 5.25), but also the motion of the drops, which may cause them to remain in the vicinity of the nozzle for considerably longer than average (Figure 5.26), or even result in a coalescence event with the filament itself (Figure 5.27).

132
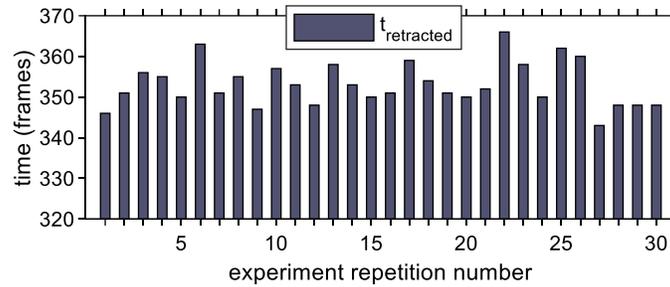
Figure 5.25. Time at which the filament has retracted to its initial position. In Appendix F, retraction time scaled to the respective time of the first breakoff event is presented, which demonstrates that the variance is not solely a reflection of the variance in the time of breakoff.
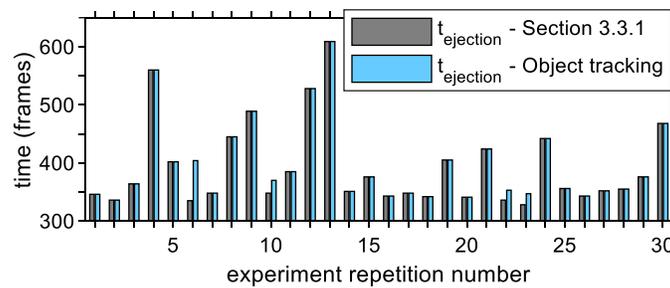
Figure 5.26. Ejection time as defined in Appendix F, obtained using the method of from the drop ejection time sub-section (gray), and obtained using the object tracking (light blue) to determine the final time at which a transient object remained in the vicinity of the nozzle. Discrepancies between the values result from drops that merge with the nozzle.
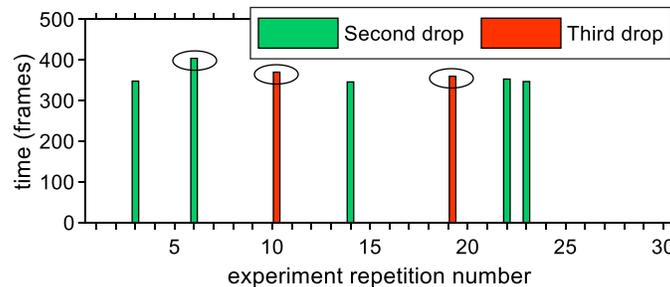
Figure 5.27. Time of event and identity of drop which merged at nozzle-attached meniscus – seemingly rare, but very important to consider. For the cases in which merging at the nozzle occurs after the filament has fully retracted, the bar is circled; the re-stabilization times are presented in Appendix F.

In these figures, take repetition 3 as an example: two drops break off, the second of which merges with the retracting filament prior to completion of the retraction. Alternately, take repetition 6 as an example: two drops break off, the second of which merges with the meniscus

133

after it has fully retracted, therefore the adjusted time must be used. If the nozzle-attached meniscus is considered for determining when the next drop may be ejected, then clearly it is not just the time required for retraction, as merging events may complicate matters. Therefore, the time from flow start-up to returning to a fully stable meniscus is equal to max $(t_{retracted}, t_{merge})$; alternately, if the presence of objects in the vicinity matters, then max $(t_{retracted}, t_{merge}, t_{ejection})$.

5.6.3   Post-Detachment Events: Breakup and Coalescence of Drops

As noted in Section 5.5.2, the events taking place after drop detachment can be quite important. In this section some relevant figures which summarize aspects of these occurrences are presented; the times and positions at which breakup and coalescence events occur are shown in Figure 5.28 for the object identity which corresponds to the "first breakoff" or primary drop. As the repetitions which exhibited a breakup and/or coalescence event for a non-primary drop – that did not involve merging with the nozzle-attached filament – were limited in number (two total: numbers 10 & 21 the corresponding videos for which are shown in Fig. 5.1, Middle & Top, respectively), only the timeline of the post-detachment events of the secondary/tertiary drops is shown in Figure 5.29.
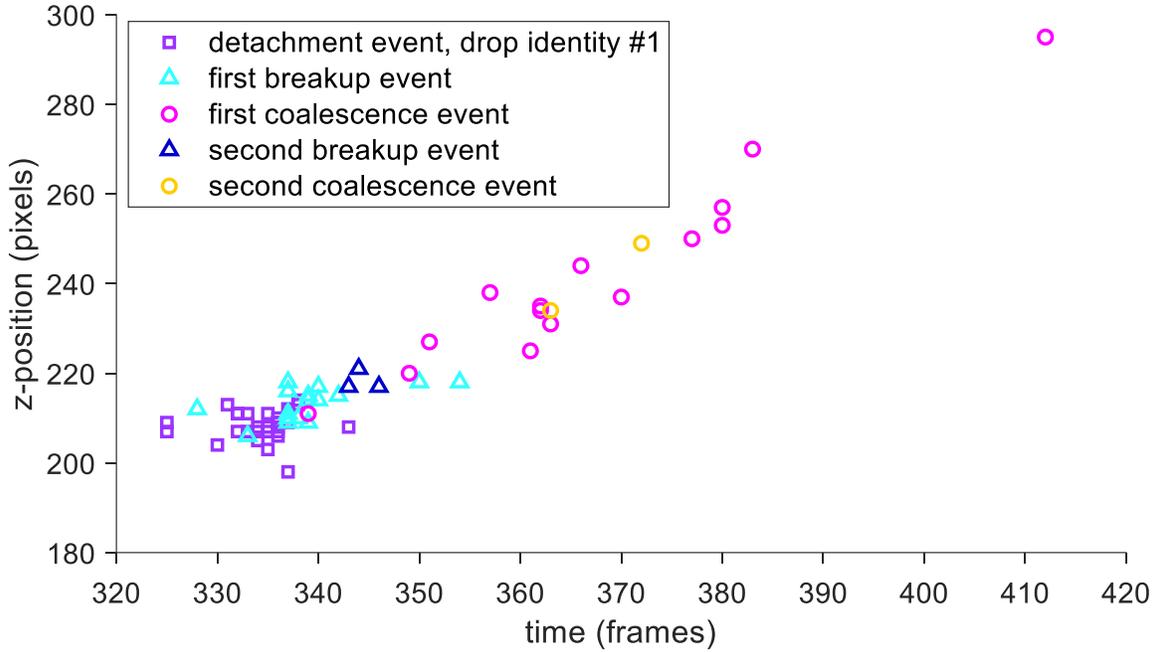
Figure 5.28. Timeline of events for primary drops: at time of event, z-position at bottommost point on respective object.
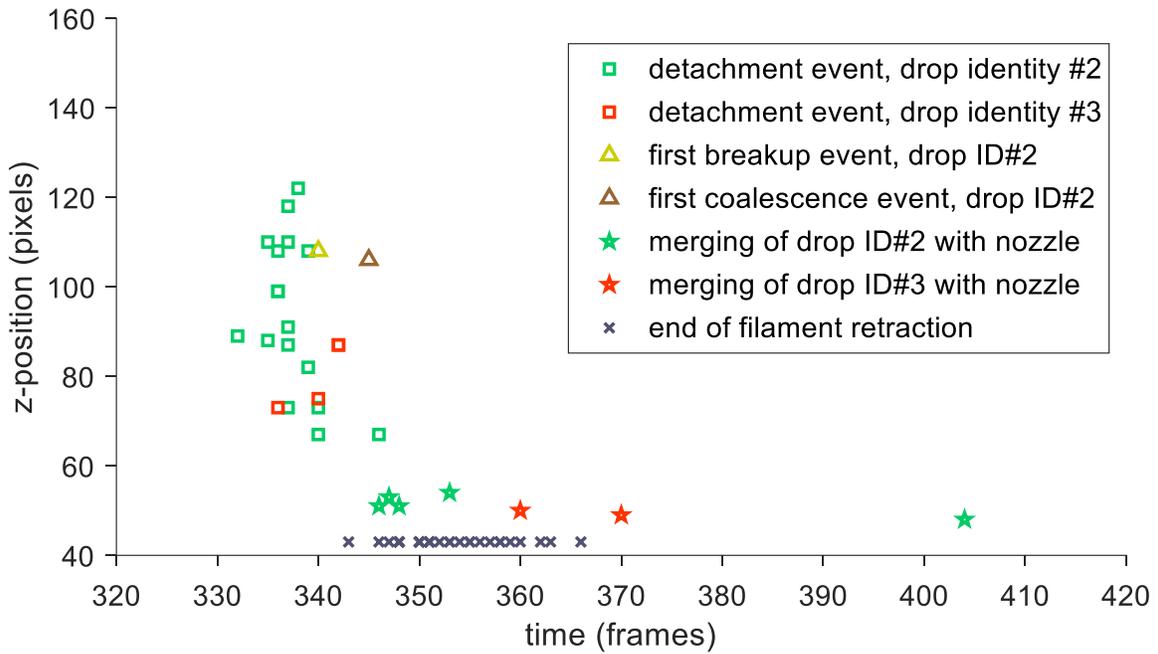
Figure 5.29. Timeline of events for secondary and tertiary drops: detachment, breakup, coalescence, merging with nozzle.

135

Figure 28 and 29 demonstrate the stochastic nature of the post-detachment phenomena, which manifests in the variability of the spatial/temporal points of the breakup/coalescence events, and the multiple modes under which they occur (Figure 32). Within the context of a single sequence, a more intuitive figure can be automatically constructed by linking the spatiotemporal points of the events in order to visualize the total lineage of an object. Constructed from the same image sequence as in Figure 5.1., Bottom, an example of this is presented in Figure 5.30, which, in addition to the plot of linked events, includes the corresponding image frames as a further means of illustration.
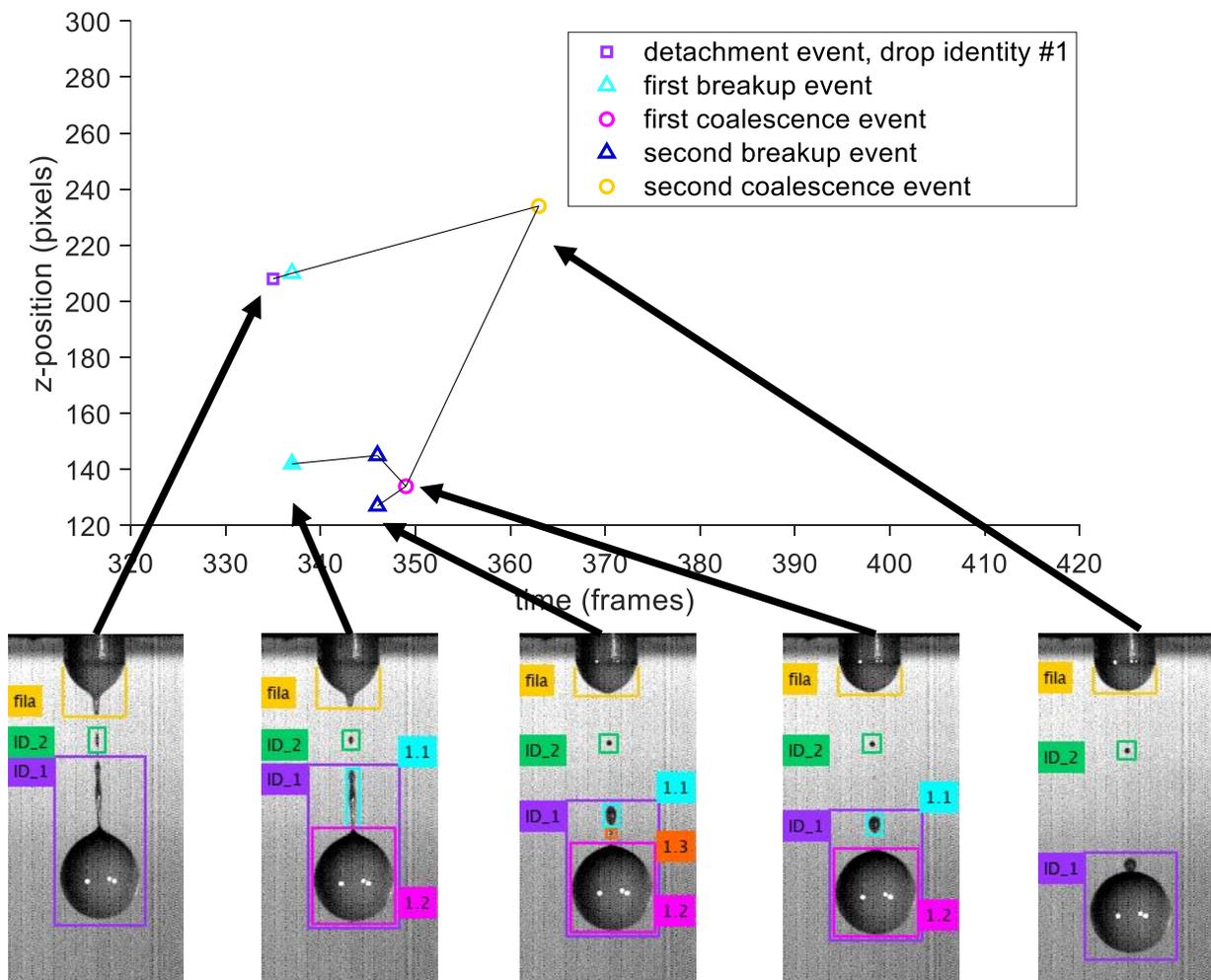
Figure 5.30. Lineage of primary drop in terms of its events. Thick black lines with arrows indicate to which time point the image frame belongs.

Given that the population-level perspective obscures the individual differences within each repetition, the times of the events are presented in Appendix F, and total positions of the objects at the occurrences are shown in Figure 5.31. From these figures, it can be observed that in general the primary drops frequently underwent at least 1 breakup and coalescence event, and less often, 2 events. For the thirteen cases in which a single breakup event occurred, the resultant fragments coalesced in all but two instances, repetitions 11 and 15. Considered together with the less frequently observed outcomes involving the primary drop (e.g. repetition 10 – coalescence of primary and secondary drop, repetition 24 – four total breakup/coalescence events, shown in Figure 1, Bottom), this indicates that in 90% of the repetitions, the primary drop was or eventually became a single entity, rather than comprised of fragments. In fact, if one considers the eventual result of all the post-detachment events – merging at nozzle, breakup, coalescence – for all the drops (primary, secondary, tertiary), and the fragments formed by subsequent breakup(s), it can be observed that in 56% (17 out of 30) of the repetitions the sequence of events ultimately yielded a single primary drop without satellites.
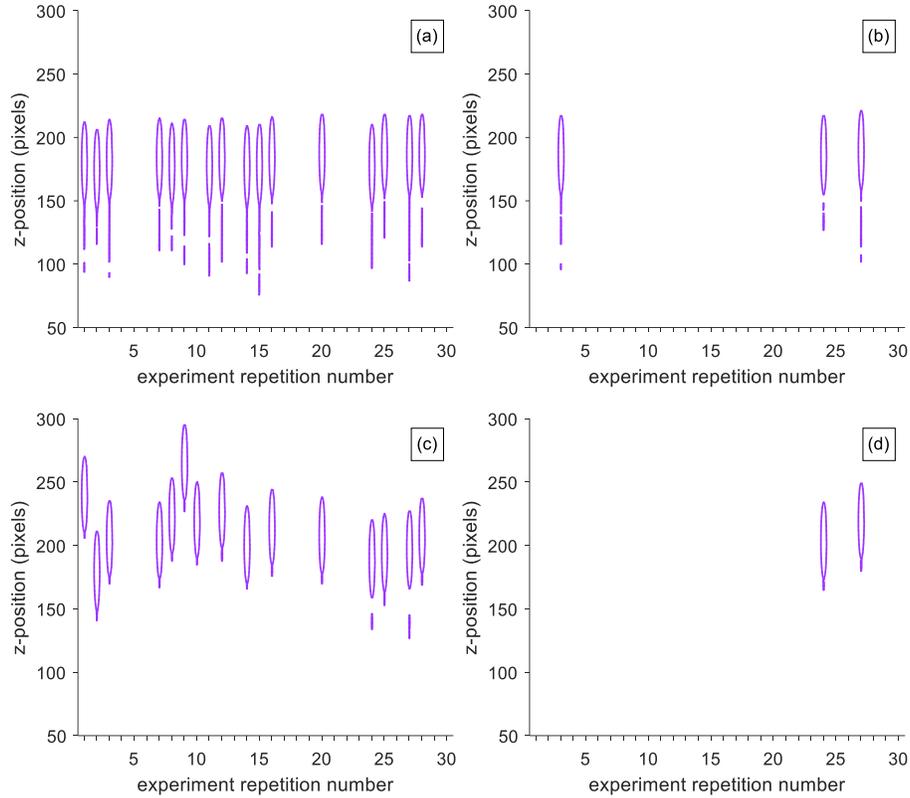
Figure 5.31. Drop shape (all x, z positions) at time of first breakup event, second breakup event, first coalescence event, and second coalescence event, which correspond to (a), (b), (c) & (d), respectively.
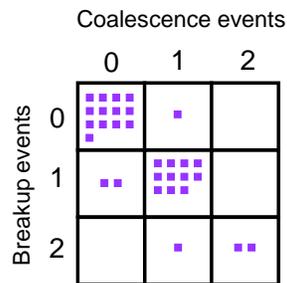
Figure 5.32. Number observed (indicated by purple squares) for each of the 9 possible outcomes (temporal ordering not considered) afforded by 0-2 coalescence and 0-2 breakup events of the primary drop. As mentioned above, the algorithms adapt to any particular modality, combination or sequence of events; it simply happens to be that 2 was the upper limit for coalescence and breakup events in the included data set.

5.7   Conclusion

In this paper, methods were developed for systematic analysis of big data sets which are generated through high-speed imaging of jet breakup and drop formation. The need to extract quantitative information from image data was motivated by the utility this has for modeling of the aforementioned phenomena, which can inform design efforts for the process, as well as sensors for online monitoring. The need to study a large number of experiments becomes especially relevant when particle-containing fluids are considered, due to the modal changes introduced by the mere presence of particles. For real particles, the stochastic effects introduced by non-ideal particles (size, shape distributions) complicate matters further, as demonstrated by the variability evident in the results shown in this work. As the analysis of many experiments, each of which consists of thousands of images, is virtually impossible by hand, the need to automate the image analysis methods is clear. This work reports on a software implementation of a suite of algorithms, and was created with a sufficiently general perspective as to be applicable to any image data which resembles the physical situation described above.

When the moieties are considered as objects which have the potential to create by breakup or destroy by coalescence other objects, one reaches the concept of object interactions with its necessary and sufficiency conditions; this is supported by the underlying patterns which occurs in the image sequences, and which have direct physical meaning here. Combined with logical systems for assignment of identities by object segmentation at the detection of an interaction, and coupled with frame-to-frame matching from motion, one is then able to recapitulate all the possibilities proffered by reality; this does not lead to an infinite regress as an overt or enumerative description would. Instead, the algorithmic strategies and logical systems have the capability to learn any possible scenario, be it a particular sequence and/or combination, of any number of objects. The recapitulated world of objects then contains the motion, deformation, events – all associated to each object – and the parent/child relationship levels, through which each object is related to other objects that appear/disappear in the given temporal (i.e. image) sequence.

The low/noisy quality of the image data generated on our imaging setup meant that the development must be robust in the presence of noise (and/or issues such as simultaneous occurrence of singularities); therefore, for higher quality imaging data which offers improved spatial and temporal resolution, the potential results will correspondingly improve in physical precision. As a related matter, since the pixel-specific search algorithms (i.e. Section 4) generalize

to images of any size and return results which are as precise as the imaging setup permits, the methods developed are fully scalable. With respect to the run time of the program, with minimal code optimization, 2.7GB of image data (22,000 images) could be loaded, analyzed and the results saved in 110-115 seconds – this equates to 190-200 images per second; this implementation was in MATLAB 2019a, run on a standard desktop computer (Intel i7-7700K CPU; without GPU parallelization). We expect that with coding optimization and an implementation in C++, at least an order of magnitude increase in efficiency could be realized; work is currently underway for this.

The information extracted from the high-speed recordings may be used to identify suitable regions of operation within which complications (e.g. satellite drops) are minimal; specifically, this enables one to model the effects of particles on the dynamics and modalities, thereby informing selection of ink formulation properties and operating conditions. Analysis of the image sets may also be used to determine optimal sensor placement, in either a lab or process setting, for a camera which acquires high-resolution images of objects for quantitative purposes [26]. For in-process measurements, low-speed, high-resolution cameras can provide the necessary image data to generate parameters and make control decisions, even if all the phenomena are not overtly observed; in these scenarios, the fully detailed information from high-speed imaging can be used to model the phenomena, and subsequently make predictions of the state or operation parameters from the more limited information afforded by the "snapshots" (captured by the in-process camera). In this sense, the program can be used to automatically annotate a training database to be used for subsequent analysis of image data with low temporal resolution – of course, such a database could be used to analyze other high-speed videos as well.

Beyond applications to other instances of jet breakup and drop breakup/coalescence in drop printing, the program may be directly used to extract information for modeling of droplet size and fragmentation in liquid spraying/atomization; this could inform design of equipment for spray-drying or wet granulation, for example. Adaptation to other similar physical situations is straightforward provided that the observation point is at relatively consistent conditions, which is easy to achieve in controlled setups. Examples to which the program is directly applicable include high-speed imaging of drops/particles/crystals – objects – undergoing growth, breakage and/or agglomeration, such as occurs in wet/dry granulators, fluidized-bed reactors/dryers, and crystallization unit operations.

The program developed in this work is available for sharing with interested members of the community, particularly those studying jet- and drop-related phenomena, in order to advance modeling work and other creative uses.

5.8   Acknowledgements

5.9   References

[1] Sick, V. (2013). High speed imaging in fundamental and applied combustion research. *Proceedings of the Combustion Institute*, 34(2), 3509-3530.

[2] Kamp, J., Villwock, J., & Kraume, M. (2017). Drop coalescence in technical liquid/liquid applications: a review on experimental techniques and modeling approaches. *Reviews in Chemical Engineering*, 33(1), 1-47.

[3] Maaß, S., Rojahn, J., Hänsch, R., & Kraume, M. (2012). Automated drop detection using image analysis for online particle size monitoring in multiphase systems. *Computers & Chemical Engineering*, 45, 27-37.

[4] Qin, S. J., & Chiang, L. H. (2019). Advances and opportunities in machine learning for process data analytics. *Computers & Chemical Engineering*, 126, 465-473.

[5] Eggers, J., & Villermaux, E. (2008). Physics of liquid jets. *Reports on progress in physics*, 71(3), 036601.

[6] Wereley, S. T., & Meinhart, C. D. (2010). Recent advances in micro-particle image velocimetry. *Annual Review of Fluid Mechanics*, 42, 557-576.

[7] Versluis, M. (2013). High-speed imaging in fluids. *Experiments in fluids*, 54(2), 1458.

[8] Thoroddsen, S. T., Etoh, T. G., & Takehara, K. (2008). High-speed imaging of drops and bubbles. *Annu. Rev. Fluid Mech.*, 40, 257-285.

[9] Basaran, O. A., Gao, H., & Bhat, P. P. (2013). Nonstandard inkjets. *Annual Review of Fluid Mechanics*, 45, 85-113.

[10] Basaran, O. A. (2002). Small-scale free surface flows with breakup: Drop formation and emerging applications. *AIChE Journal*, 48(9), 1842-1848.

[11] Furbank, R. J., & Morris, J. F. (2004). An experimental study of particle effects on drop formation. *Physics of Fluids*, 16(5), 1777-1790.

[12] Bonnoit, C., Bertrand, T., Clément, E., & Lindner, A. (2012). Accelerated drop detachment in granular suspensions. *Physics of Fluids*, 24(4), 043304.

[13] Miskin, M. Z., & Jaeger, H. M. (2012). Droplet formation and scaling in dense suspensions. *Proceedings of the National Academy of Sciences*, 109(12), 4389-4394.

[14] Lindner, A., Fiscina, J. E., & Wagner, C. (2015). Single particles accelerate final stages of capillary break-up. *EPL (Europhysics Letters)*, 110(6), 64002.

[15] Pan, Z., Louvet, N., Hennequin, Y., Kellay, H., & Bonn, D. (2015). Drop formation in shear-thickening granular suspensions. *Physical Review E*, 92(5), 052203.

[16] Radcliffe, A. J., Hilden, J. L., Nagy, Z. K., & Reklaitis, G. V. (2019). Dropwise Additive Manufacturing of Pharmaceutical Products Using Particle Suspensions. *Journal of Pharmaceutical Sciences*, 108(2), 914-928.

[17] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679-698.

[18] Elder, J. H., & Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7), 699-716.

[19] Lindeberg, T. (1998). Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 117-156.

[20] Bao, P., Zhang, L., & Wu, X. (2005). Canny edge detection enhancement by scale multiplication. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(9), 1485-1490.

[21] McIlhagga, W. (2011). The Canny edge detector revisited. *International Journal of Computer Vision*, 91(3), 251-261.

[22] Xu, Q., Varadarajan, S., Chakrabarti, C., & Karam, L. J. (2014). A distributed canny edge detector: algorithm and FPGA implementation. *IEEE Transactions on Image Processing*, 23(7), 2944-2960.

[23] Lee, J., Tang, H., & Park, J. (2018). Energy efficient canny edge detector for advanced mobile vision applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(4), 1037-1046.

[24] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.

[25] Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-166.

[26] Radcliffe, A.J., & Reklaitis, G.V. (2020). An application of computer vision for optimal sensor placement in drop printing. *Computer Aided Chemical Engineering*, 30th European Symposium on Computer-Aided Process Engineering (in press)

6. AN APPLICATION OF COMPUTER VISION FOR OPTIMAL SENSOR PLACEMENT IN DROP PRINTING

6.1 Abstract

The use of imaging technologies in the process industries has the potential to provide state, dynamic, statistical or modal information that cannot otherwise be obtained from systems which limit or invalidate the use of physical property and spectroscopic measurement methods for real-time acquisition. To extract the desired information, computer vision techniques can be used to reduce the high-dimensional image data to the salient features, which can then be employed for modelling, optimization, and control purposes. This work applies such an approach to determine the optimal placement of a high-resolution camera for online quantification of drops printed in a slurry-based additive manufacturing process for pharmaceuticals. Systematic selection requires a model of drop instability, which, due to stochastic effects introduced by particles, must be based on real data with sufficient temporal discretization to capture the event dynamics. Thus, high-speed imaging was utilized to generate representative videos, and computer vision algorithms were employed to automate extraction of the coupled temporal evolution of object shape and position, and to detect and recognize breakup and coalescence events. The resultant data were subsequently employed to determine the image sensor and fiber-optic placements that maximize measurement information acquired by online acquisition system. Results are presented for drop printing slurries at 15 distinct ink formulations – spanning several particle size distributions and particle loadings – which are comprised of a set of 250,000 images.

6.2 Introduction

Drop printing serves as an important component of widely varied manufacturing processes, encompassing examples from semi-conductors, pharmaceuticals and biomaterials – essentially, anywhere precision deposition of material is required. In such applications, drops serve as the

144

source of a critical chemical component themselves, or as a liquid binder, the patterning of which on the substrate enables creation of two-/three-dimensional structures. Consequently, in order to control the drop formation/deposition, online monitoring of the generated drops is desirable; this would be best achieved by high-speed imaging due to the small timescales involved, but economic and practical considerations encourage the use of a normal camera to capture one image of each emergent drop. In this context, placement of the online imaging system in relation to the printing nozzle determines what can be extracted and inferred from the resultant images, thus, it is important to optimize the arrangement of the camera, flashbulb & fiber-optic sensor (camera trigger) so as to maximize the information about each drop.

Previously work by colleagues [1] designed a system for online image acquisition that considered the optics, using the system geometry to select lens and camera y-axis placements that maximize the depth of field for objects emerging from the nozzle; this incorporated an actuation system based on a fiber-optic sensor positioned at the same z-height as the center of the imaged area. The position of these two components will be selected by the user; for well-behaved inks (i.e. ideal fluids), a trial and error approach can affect a suitable measurement, with errors as low as 0.33% [1]. However, for cases in which printing exhibits non-ideal effects – multiple satellite droplets, variability in trajectory, non-deterministic events (breakup, coalescence) – the captured images can be quite different (Figure 6.1, Inset #1, 2, 3), despite apparent uniformity of drop volume at the eventual point of deposition. Consequently, it becomes difficult to even arrive at a satisfactory result, or to be convinced that such a placement captures the maximum amount of information.

In the pharmaceutical manufacturing process considered here, the printed drops contain the active drug compound, thus, the images obtained by the online acquisition system enable model-based estimates of the product content; maximizing the information contained in the captured images through optimization of component placement with respect to drop formation phenomena reduces uncertainty in the model structures and in the product content predictions. This work focuses on non-Brownian particle inks, which motivate the use of a data-driven approach due to the substantial non-idealities involved– high particle loading, distributed particle size, non-spherical shape. High-speed imaging of drop printing across a range of operating conditions and ink formulations is used to record the dynamics from the incipience of flow to the eventual stabilization of the drop(s) after detachment. Then, computer vision algorithms are applied to each

145

image sequence to extract the total lifetime of each drop: dynamic evolution of position and shape; non-deterministic events and spatiotemporal location thereof; parent-child relations with other drop(s). With this information, particle-driven effects on detachment, trajectory and events can thus be accounted for.

6.3 Problem Description

6.3.1 Components of Online Image Acquisition System

In order to formulate the problem, one must first define the geometry and adjustable components of the online imaging system. From the previous work, camera and lens are fixed, thus leaving the position of the camera and fiber-optic sensor as the only variables. The camera is to be placed at some z-position below the nozzle orifice so as to capture a single image for each drop; the fiber-optic sensor consists of a light beam, which when interrupted by the drop triggers the acquisition of the image. The high-resolution camera captures an area with dimensions H, W, which is effectively constant due to the fixed distance between the lens and drop paths. The high-speed camera captures a substantially larger actual area; with the actual sizes and known conversion factors (real length per pixel), the potential positions of the area captured by the online camera can be superimposed onto the images from the high-speed videos, as shown in Figure 6.1.

6.3.2 Feasible Camera, Fiber-Optic Sensor Placements

Given the geometry (Figure 6.1) thus defined, and the total imaged area from the high-speed sequences, one can then define the set of feasible camera positions as $\mathbf{Z} = \left\{ \boldsymbol{\zeta} \ \forall \ z_{Box,L} \mid z_{Box,L} = (z_{nozzle} + H), \dots, z_{limit} \right\}$, in which a single camera position encompasses the set of points: $\boldsymbol{\zeta} = \left\{ z_{Box,L} - j \mid j = 1, \dots, H \right\}$. As the area imaged by the online camera covers 90% of the x-axis, the problem can justifiably be simplified by fixing the camera box x-position at the centerline. Feasible positions of the fiber-optic sensor for each $\boldsymbol{\zeta}$ are constrained to be $z_{FbOp} \leq z_{Box,L}$, though, given the configuration in which interruption of the fiber-optic signal is exactly coincident with the camera trigger, one can further limit the feasible points to $z_{FbOp} \in \boldsymbol{\zeta}$, as image acquisition, triggered at some position outside the camera box is readily identifiable as leading to suboptimal results. The former, more lenient constraint could be used if one considers adding a time delay between the fiber-optic signal interrupt and image acquisition.
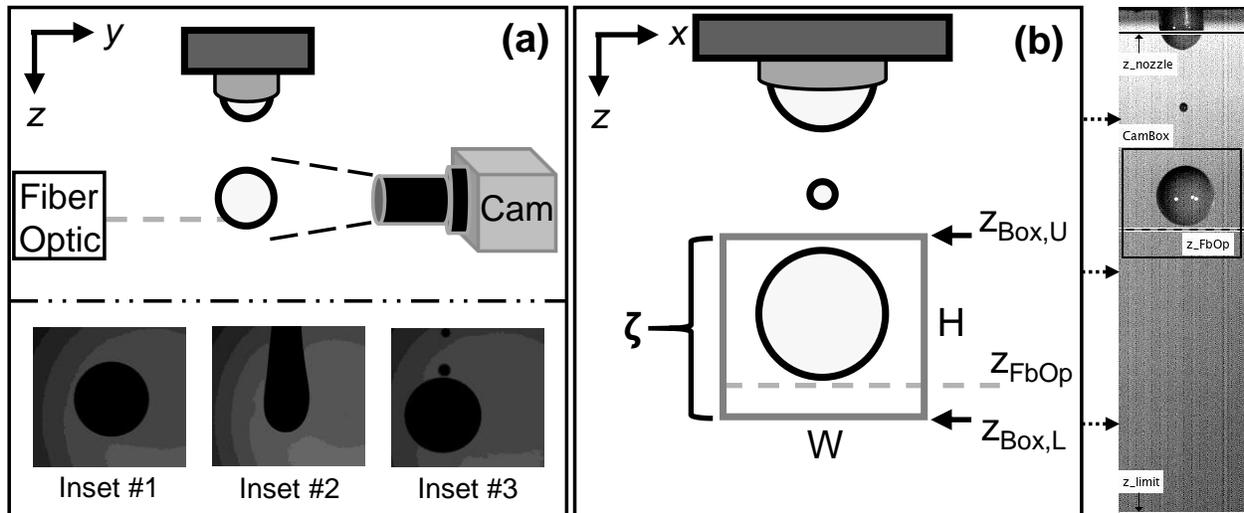
Figure 6.1. (a): Side-view of online image acquisition setup. (b): Scene as observed by camera.

6.3.3 Objective Function

With the feasible sensor placements defined, consider the following scenario in which a drop is printed: mechanical actuator induces pressure wave, fluid is emitted from the nozzle, and after thinning of the liquid thread, one or more drops detach; thereafter, these drops deform, and potentially undergo breakup or coalescence. Let us assume that the operating conditions and ink formulation are such that per actuation event, one large drop ("primary drop") is formed, with the possibility of one or more smaller drops ("satellites"). The system can acquire one image per actuation event, given limitations on the flashbulb frequency and online camera, and the volume of fluid deposited will be estimated using an object detection and segmentation algorithm [2] to extract the diameter for each object, $D(z)$, which, assuming axial symmetry, is used to compute: $V_{Drop} = (\pi/4) \int D(z)^2 dz$.

For the purposes of online process monitoring in the pharmaceutical manufacturing process, as accurate as possible estimate of drop volume is desired; as illustrated by Figure 1, there is considerable uncertainty as to whether the acquired image contains representative information of the deposited volume (e.g. Inset #2, 3), due to differing detachment phenomena, and the potential for satellite drops. Thus, the optimal camera placement and fiber-optic sensor position are the ζ, z_{FbOp} that maximize the observed drop volume – either across a range of conditions (if ζ, z_{FbOp} are to be fixed henceforth), or, for each operating condition/ink formulation, provided the flexibility exists to make automatic adjustments in response to process changes.

147

Noting that the volume computed from the image is proportional to the area, the objective function can be defined based on the area of the object(s) of interest inside the box, normalized to the total area of said object(s), evaluated at the instant of image acquisition – that is, the point at which the lowermost drop interrupts the fiber-optic signal. If only the primary drop is considered, the object function can be defined as the sum across p sequences, as shown in Eq. (1):

$$f(\zeta, z_{FbOp}) = \sum_p \left(\frac{A_{obj,in}}{A_{obj}} \right)_p \tag{1}$$

In which the total area of the object, A_{obj}, is defined as in Eq. (2):

$$A_{obj} = \int_{z_L}^{z_U} D(z)dz = \sum_j D(z_j)\Delta z, \tag{2}$$

$$\{\forall j : z_j \in ID_{Primary}\}$$

wherein the limits on the integral are the respective lower/upper boundaries of the primary object. The area of the object inside the camera box, $A_{obj,in}$, is evaluated analogously, with the limits on the integral such that only the area contained within ζ contributes, illustrated in Figure 6.2; the index set is then: $\{\forall j : z_j \in ID_{Primary} \cap z_j \in \zeta\}$.

This serves as a reasonable approximation for cases in which the primary drop is much larger than the satellite drops. However, given the detailed information, one can more rigorously re-define the objective function for multiple drops (Figure 2) as the sum of areas of all objects inside the camera box, normalized by the total areas of all objects both inside/outside, as shown in Eq (3):

$$F(\zeta, z_{FbOp}) = \sum_p \left(\frac{\sum_{\widetilde{ID}(p)} A_{obj,in,ID}}{\sum_{\widetilde{ID}(p)} A_{obj,ID}} \right)_p \tag{3}$$

With the object-specific event information extracted by the computer vision algorithms, one is able to make an important distinction in terms of the objects and their eventual fates: Eq.(3) should be evaluated for all objects that do not merge with the nozzle, thus, $\widetilde{ID}$ is the subset of all object identities, ID, that satisfy this condition. The reasoning for this is straightforward: through the objects detach, they do not actually contribute to the deposited volume, hence, they should be excluded from the set of objects over which the multi-drop objective function is evaluated.

A further note on the definition of what precisely constitutes an object – $A_{obj,in,ID}$ or $A_{obj,ID}$ – is warranted. Consider a case in which the camera box is positioned as close to the nozzle as possible: when the fluid filament begins to form, it then extends into the box, and may perhaps

trigger image acquisition before the drop detaches. If the area in the camera box were simply the apparent area, this would result in objective functions that favour being close to the nozzle – as perhaps more volume is extended from the nozzle than actually forms a drop. This problem can be elegantly dealt with utilizing the object-specific events construction inherent to the CV algorithms, therefore enabling one to evaluate the objective function for all dynamically instantiated objects active at a given spatiotemporal point, while excluding persistent objects (i.e. fluid attached to nozzle). This enables consideration of the full range of possible $\zeta's$ without the need to artificially set an upper bound so as to exclude contributions from the nozzle-attached filament. The convenience of the object-oriented perspective is that it permits us to specify in the objective function the high-level criterion that "valid" area within the camera box must belong to a free liquid fragment (nozzle-attached filament can be present, but its contribution is always zero), and to easily re-define subsets of objects to evaluate Eq. (3).
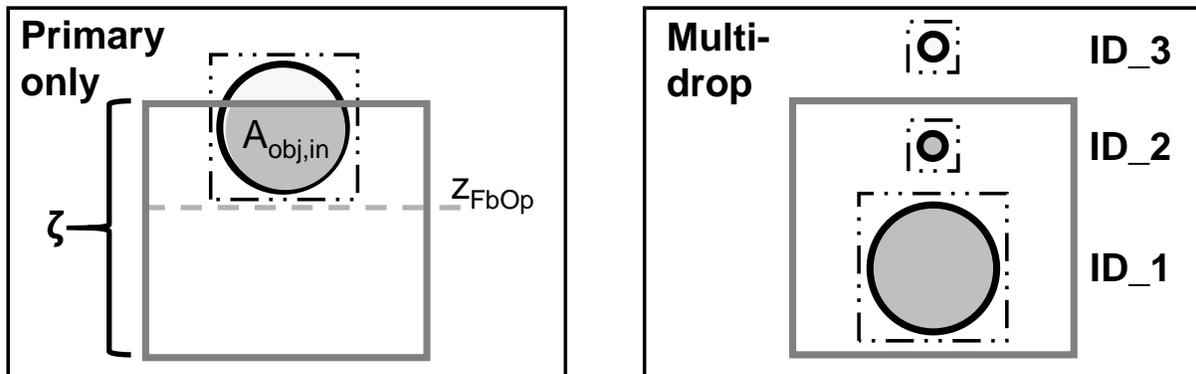
Figure 6.2. Illustration of contributions to objective function: area inside and total areas

6.3.4   Solution Method

From the definitions of the objective functions, it can be readily inferred that a gradient-based method may not be preferable. As one expects that a range of ζ, z_{FbOp} are likely to yield quite similar objective function values, a derivative free method is selected; the classic formulation of the particle swarm optimization algorithm, using 50 particles and 150 generations is used to obtain solutions. An additional advantage for the problems posed here is due to the shape of the objective function (obtained by exhaustive evaluations for p sequences at one set of operating conditions), which resembles a plateau with very steep slopes along the boundary where z_{FbOp} is very near to

149

$z_{Box,L} + H$. Intuitively, this must be the case: as one allows more of the drop area to enter the imaged box, the function changes rapidly.

6.4   Results: Primary Drop, Multi-Drop, Primary Drop + Interaction Partners

The solution to the optimization problem yields a set $\boldsymbol{\zeta}$, parameterized by $z_{Box,L}$, which comprises the camera box positions and a z_{FbOp} which constitutes the location of the fiber-optic sensor. Results are presented in this section for the cases involving only the primary drop, multiple drops, and a third case that takes further advantage of the vent information which pertains to the parent-child relations of the objects. In this third case, the primary object and its interaction partners form a subset of the objects, therefore, by replacing $\widetilde{ID}$, one can evaluate the respective objective function (Eq.(3) directly.

The solutions obtained indicate that when only the primary drop is considered in the objective function, multiple optimal solutions exist, as shown in Figure 3 by the light gray dots, at which points $(\boldsymbol{\zeta}, z_{FbOp})$ the objective function is essentially equal. Including the extra drops in the objective function – multi-drop form, $F(\boldsymbol{\zeta}, z_{FbOp})$ – has a dramatic effect (black dots, Figure 6.3), in that solutions are then essentially limited to one location which has strong dependence on the satellite drop trajectories. In the third case (dark gray dots, Figure 6.3), which considers all the events and drop-drop interactions, but limited to the perspective of the primary drop and its interaction partners; the optimal placements are nearly equal – around $z_{FbOp} \approx 400$, which corresponds to a physical distance approximately 4 drop diameters from the nozzle, essentially in the middle of the potential placements.

Closer inspect of the primary-only objective function evaluated on all points in the feasible region indicates larger gain as one moves away from the nozzle orifice, but, from ≈ 225 onward, the gain levels off rapidly with respect to camera placement, akin to a plateau. With respect to fiber-optic placement, generally the closer it is to the upper limit of the camera box, the lower the objective function value – which fits with intuition.
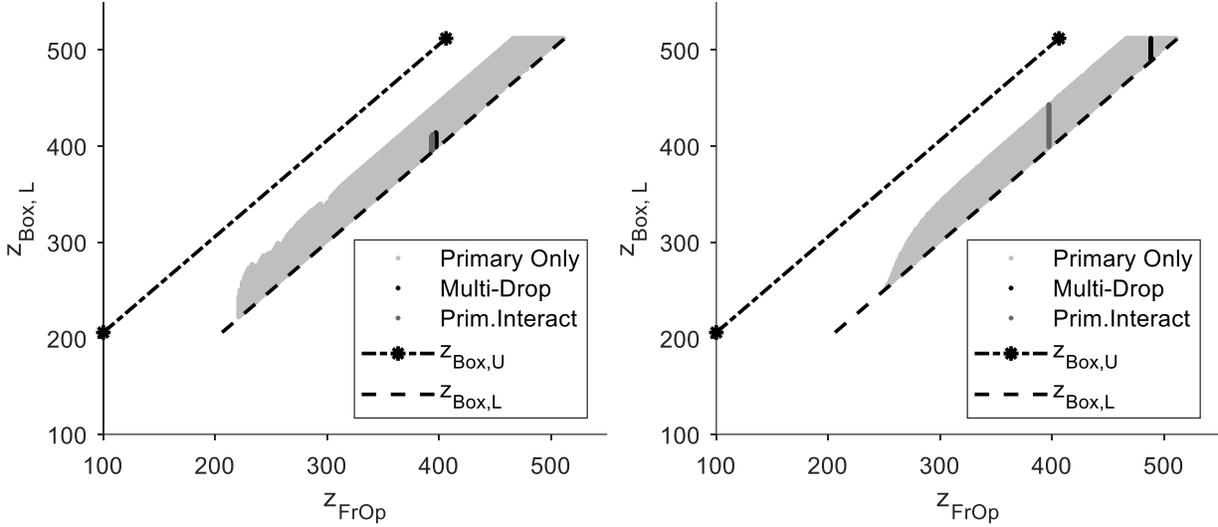
Figure 6.3. Optimal solutions for each objective. Left: μ-APAP, Φ=0.22; Right: sf-APAP, Φ=0.28.

6.5   Discussion

The difference in the placements for the primary-only and multi-drop objective function fits with intuition – for a single object, there should be many camera/sensor placements that enable one to capture the entire volume of the primary drop across all the sequences at a given condition. Coincidentally, this is why it is easy to select by trial and error a camera/sensor position capable of capturing the whole deposited volume when the fluid properties yield a single drop per actuation event. On the other hand, the placement for the primary+interaction and full multi-drop cases yield very specific regions which are respectively far from many of the points that are potential optima for the primary drop. There are only a few possible ζ, z_{FbOp} that maximize the objective function; this results from the variability in the spatiotemporal events (particularly those that involve the primary drop) which lead to satellite drop formation. Comparison of the two conditions shown in Figure 3 indicates a significant difference in the placements between them for the multi-drop case; in the μ-APAP set, the placement is nearly the same as the primary+interaction, whereas in the sf-APAP set, the multi-drop placement is notably farther from the nozzle, and from the primary + interaction. This is due to the difference in the volumes of the satellite drops, as their contribution is proportionate – μ-APAP generally produced smaller satellites, but sf-APAP yielded larger satellites whose volume more significant, hence the shift to a position which was able to achieve approximately 1% greater imaged volume by their inclusion.

6.6   Conclusion

The optimization problem posed here becomes particularly important when one expects to have multiple droplets which undergo breakup or coalescence, as exemplified by the severely restricted regions which contain optima obtained for the primary+ and multi-drop cases. As these images provide a critical process measurement, minimizing uncertainty at the point of acquisition improves predictions from any models that subsequently make use of this information. Though trial and error enables one to guess the conditions when only the primary drop is considered, the treatment presented here demonstrates that some information would inherently be missing. What enables us to approach this problem despite the marked stochastic effects is the use of a data-driven approach which utilizes computer vision to extract object-specific and event type information, from a large number of images – 250,000+ - for real data obtained by high-speed imaging. This provides clear guidance on placements for the complex (multiple drops) scenarios regularly encountered when working with non-ideal inks.

6.7   References

[1] L. Hirshfield, E. Icten, A. Giridhar, Z.K. Nagy, G.V. Reklaitis, 2015, "Real-Time Process Management Strategy for Dropwise Additive Manufacturing of Pharmaceutical Products", J. Pharm. Innov., 10, 2, 140-155

[2] A.J. Radcliffe, G.V. Reklaitis, 2020, "Automated Object Tracking, Event Detection and Recognition for High-Speed Video of Drop Formation Phenomena", (manuscript to be submitted to AIChE Journal)

7. BAYESIAN ESTIMATION OF PRODUCT ATTRIBUTES FROM ON-LINE MEASUREMENTS IN A DROPWISE ADDITIVE MANUFACTURING SYSTEM

7.1 Abstract

In the pharmaceutical industry, application of additive manufacturing technologies presents unique opportunities for monitoring and control of drug product quality through on-line image acquisition systems which estimate process output from images of emergent drops. For doses produced from pure fluids on-line imaging has demonstrated utility [1]; however, processing of particulate suspensions requires consideration of uncertainty regarding gravity- and flow-mediated particle-liquid segregation. Even for well-mixed suspensions, random localization of particles during the necking of the liquid bridge results in variable drop trajectory [2], which must be corrected for in the estimates of drop volume provided by an on-line imaging system. This work explores the use of Bayesian statistical models to resolve potential inaccuracy in drop volume estimates obtained by an on-line image sensor in a dropwise manufacturing process for suspension-based drug products. The framework in the case study presents a method by which uncertainty arising from on-line process measurements may be reconciled for through sufficient sampling and off-line analysis of process outputs.

7.2 Introduction

Concurrent with the expansion of inkjet printing methods into precision materials (e.g. semiconductor, ceramics) industries, considerable research effort has been devoted to development/adaptation of inkjet technology for pharmaceuticals manufacturing. Applications of additive manufacturing technology have received substantial attention for production of oral solid doses (e.g. tablets, capsules, thin films). The mechanics of this fluid-based process facilitate monitoring and control of drug product properties, as the output of the system, individual drops,

may be monitored in real time using an on-line imaging system [1]. Recent work with homogenous fluids (solvents, melts) has demonstrated feasibility of an image capture system for process monitoring in dropwise additive manufacturing system [1, 3]. However, quantification of process output with an image capture system becomes ambiguous for suspensions due the possibility of particle-liquid segregation during transit through the system. Additional uncertainty is presented by drop ejection, due to random localization of particles which may result in irregular drop trajectory; this manifests as uncertainty in the computation of drop volume from the on-line image system. The optical sensor calibration is sensitive to the distance between drop and camera, for which variable drop trajectories would result in inaccurate drop volume estimates from the on-line image system. In this work, a Bayesian statistical model is applied to reconcile errors in computed/actual drop volumes from a set of suspension formulations and process conditions which yielded reproducible drops/products, as determined by gravimetric analysis. Multiple model structures are considered for estimation of actual drop volumes from computed drop volumes; prior information is obtained from separate experiments conducted at the same set of conditions. The estimated parameters from the statistical models are compared with predictions afforded by linear regression models, in both classical and Bayesian formulations.

7.3   Dropwise Additive Manufacturing Platform, On-line Image Acquisition

The dropwise manufacturing system consists of a precision positive displacement pump, x-y staging, pump/staging controllers and on-line image acquisition system; the image acquisition system is comprised of a camera actuated by a motion sensor; the setup is a variant of the process described in [1]. The mechanical action of the positive displacement pump forces fluid through the dispensing nozzle, which, dependent on process conditions and suspension properties, results in formation of one or more drops; the present work considered only the experimental conditions which produced a single drop per actuation event of the pump. A process diagram, including the image acquisition system, is presented in Figure 7.1.
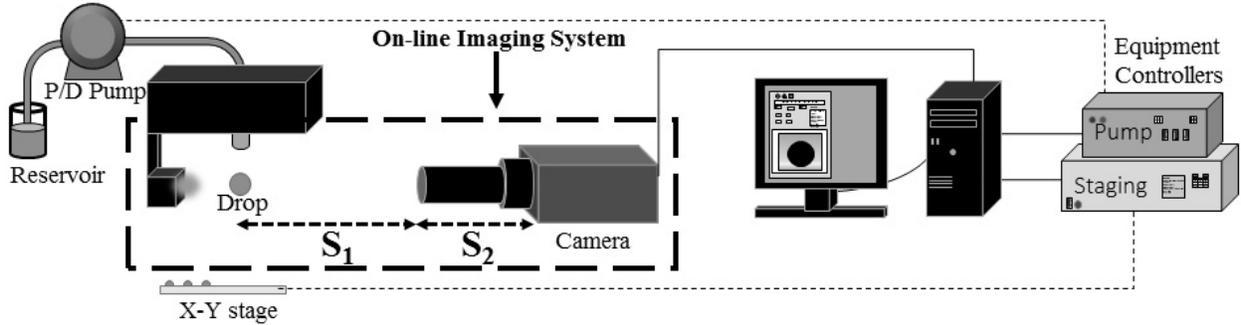
Figure 7.1. Process diagram for dropwise additive manufacturing system

The on-line image acquisition system computes drop volume from the two-dimensional image by estimating the volume of the sphere from a stack of equivalent cylinders, for which diameters are defined using the drop edges; further detail is available in [1]. A calibration procedure which defines a unit pixel in relation to physical dimensions is necessary to complete the volume estimate; this was performed in a configuration analogous to that depicted in Figure 7.2, with the drop replaced by an optical standard.
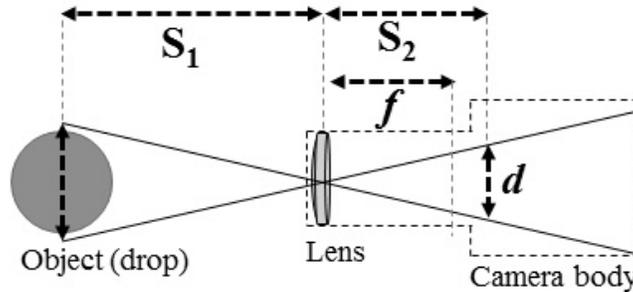
Figure 7.2. Ray tracing diagram for image acquisition system; drop not proportionate to scale.

The calibration procedure assumes constant values for system properties: S_1 (distance of object from lens), S_2 (distance from lens to image), f (focal length), and d (dimension of image sensor). As S_1 is determined by each the drop's trajectory, deviations result in errors for computed drop volume due to a changes in the effective magnification, which for a thin lens the linear magnification would change by a factor proportional to the ratio of ($S_{1,Calibration}$/ $S_{1,NEW}$). For pure fluids, the assumption of constant drop trajectory (fixed S_1) has demonstrated utility [1, 3]; for particle suspensions, drop trajectory appears to be a distributed parameter (variable S_1).

7.4   Bayesian Statistical Modeling for Actual Drop Volume in Suspensions

Drops produced from suspensions exhibit larger variance in volume than drops from pure fluids due to the inherent heterogeneity imparted by particles. The presence of particles disrupt thinning of the fluid filament during drop break off at a length-scale proportional to particle size; for particle samples with a size distribution, this may cause drops to pinch off at different spatial locations, dependent on the presence of (relatively) large particles in the thinning filament [2]. Variability in the drop break off process may lead to altered drop trajectory, which may be one of several causes for discrepancies between computed drop volume and actual drop volumes as determined by gravimetric analysis. The severity of other process disturbances (flow-induced segregation; sedimentation) depends on suspension properties, process conditions and experimental timescale, thus, potential data sets for modelling were selected based on reproducible drop production which appeared to only exhibit error in computed drop volumes. This corresponded to drops produced from suspensions with particle volume fractions of 0.15-0.25 formulated from four pharmaceutical powders in a triglyceride oil carrier fluid; details related to particle characterization and properties may be found in [4].

The use of a statistical model for estimation of actual drop volume incorporates uncertainty into parameter estimates obtained using the computed drop volume as surrogate for actual drop volume. Several Bayesian model formulations with respective parameters sets and structures are developed for estimation of actual drop volume, based on future observation of computed drop volume. The Bayesian model results are compared with predictions from a linear regression model, for which unknown parameters are estimated by (1) ordinary least squares and (2) a Bayesian formulation of ordinary least squares. In the Bayesian models, informative prior distributions were estimated from the data (actual/computed drop volume) at given set of process conditions and suspension properties; the same data was used for the linear regression and Bayesian linear regression.

The first Bayesian model was structured with the likelihood based on the assumption that each computed drop volume (Y) was drawn from a normal distribution with parameters $(\theta_1 - \theta_2, \sigma^2)$, in which $(\theta_1 - \theta_2)$ and σ^2 are the mean and variance of computed drop volume. Actual drop volume (θ_1), and an offset parameter (θ_2), which represented a combined estimate of errors related to drop trajectory and other process disturbances; a multivariate normal distribution was assigned to (θ_1, θ_2) in order to preserve details of the relationship between the two parameters. The variance

parameter (σ^2) of the likelihood was assigned a inverse gamma prior distribution, with parameters estimated from the data set; likewise, the mean vector and covariance matrix of the prior on (θ_1,θ_2) were estimated from the experimental data. Treatment of the future observations, y_i, individually or as a sample result in different interpretations of the available information. Treated individually, the y_i's are used to estimate ($\theta_1,\theta_2)_i$, and when treated together, the identity of y_i's are subsumed under the sample average. Formulations produced by the dropwise platform contain 5-50 drops per dose, which motivates treatment of y_i's using the sample average due to the large number of drops required by a dose regimen of 30 capsules. The joint posterior distribution for this model, conditioned on a future set of observations, Y, is given in Eq. (1).

$$p\left(\theta_1,\theta_2,\sigma^2\mid\underline{Y}\right)=\frac{p\left(\underline{Y}\mid\theta_1,\theta_2,\sigma^2\right)p(\theta_2\mid\theta_1)p(\theta_1)p\left(\sigma^2\right)}{p(\underline{Y})} \qquad (1)$$

The second Bayesian model considers a case in which variable drop trajectory affects computed drop volumes through a magnification re-scaling factor proportional to ($S_{1,C}/ S_{1,N}$), based on the image acquisition configuration and calibration. The likelihood distribution assumes that computed drop volume(s), Y, are drawn from a normal distribution with mean and variance parameters, θ_4/θ_5, and σ^2, respectively, which have known prior distributions. To preserve the relationship between them, the actual drop volume (θ_4) and scale parameter (θ_5) are assigned a multivariate normal distribution, with mean vector and covariance matrix estimates obtained from the experimental data. An inverse gamma distribution is assigned to the prior on σ^2, with parameters estimated from the experiments. The joint posterior distribution for this model is given in Eq. (2), which treats the y_i's using the sample average.

$$p\left(\theta_4,\theta_5,\sigma^2\mid\underline{Y}\right)=\frac{p\left(\underline{Y}\mid\theta_4,\theta_5,\sigma^2\right)p(\theta_5\mid\theta_4)p(\theta_4)p\left(\sigma^2\right)}{p(Y)} \qquad (2)$$

The Bayesian models were compared with a normal linear model given two sets of parameters estimates obtained by: [1]ordinary least squares and [2]Bayesian linear regression. The specifications of the linear model were as follows: actual drop volume, y, was the response variable; computed drop volume, X, was the explanatory variable; with parameter vector $\beta=[\beta_1,\beta_2]$; y, X were the same data as the prior distributions, which contained n=39 samples. In matrix form, the model consisted of the response column vector, y_i, design matrix X_{ij} with column j=1 specified as the control (=1), and column j=2 as the computed drop volumes and parameter vector, β; the Bayesian procedure used the same model, but assigned a prior distribution to the variance of

observations. As these were standard models in both cases, additional details are not included here, but may be found in reference texts [5].

7.5   Sampling From Joint Posterior Distributions: Markov Chain Simulations

The joint posterior distributions presented in Eq. (1) and Eq. (2) are used to estimate actual drop volume for a given set of process conditions and suspension properties. For a given model, the joint posterior density was sampled using the Metropolis-Hastings algorithm in Markov chain simulations, with the density ratio defined using the proportionality of the joint posterior distribution of interest. The prior distribution was selected as the jumping distribution, which is used to generate of parameter proposals, θ_1^*; proposals are evaluated using the ratio of the proposal joint posterior density (step θ^*) to that of the joint posterior density at the previous step (θ^{t-1}). The algorithm proceeds by comparison of a random draw from the standard uniform distribution to the posterior density ratio, r; θ^* is accepted as step t of the Markov chain when the density ratio is greater than the standard uniform draw, otherwise the algorithm proceeds from step θ^{t-1} [6]. Markov chain simulations using the Metropolis-Hastings algorithm were executed and analysed in MATLAB; for a given joint posterior density, 10 chains with unique starting positions were simulated until the vectors of "accepted" posterior samples had lengths of 1,000,000; convergence monitoring and assessment of mixing followed the framework presented by [5].

7.6   Results and Discussion

Using the dropwise additive manufacturing platform, data sets consisting of n drops were generated; the on-line image acquisition system recorded a digital photograph of each drop, which landed in a pre-weighed container for gravimetric analysis. The actual volume of each drop was estimated from its mass, with a working assumption based on constant fluid density. The linked computed image volume and corresponding mass-based volume were used for estimation of prior distribution parameters in models 1, 2 and for the data used for linear regression. The results of one case study, based on drops produced from micronized acetaminophen (20 vol%) suspensions in triglyceride oil, are presented in Figures 7.3, and 7.4.
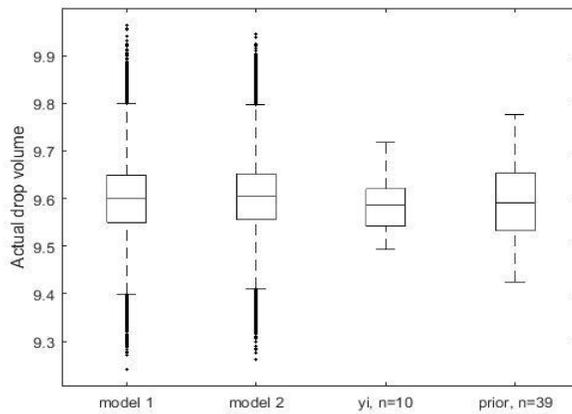
Figure 7.3. Posterior distributions of model 1,2 parameters (θ_1, θ_4), respectively, distribution of true values correspond to y_i's from the likelihood, and distribution of prior data.

The simulated posterior distributions are presented in Figure 7.3 as a box plot, in addition to the set of future observations of actual drop volume estimated from image (y_i in model 1,2 notation); the prior distribution of actual drop volume as the fourth entry. From Figure 3, it can be observed that model 1 (Eq. (1)) and model 2 (Eq. (2)) provide a reasonable estimate for mean value and 25/75th percentile; however, the posterior distributions appear to have much longer tails, as indicated by the whiskers – this is not unexpected as uncertainty incorporated in the model structures may be reflected as wider confidence intervals for the posteriors. In Figure 7.4, the linear regression results are presented; though the difference in mean predicted value is negligible, the Bayesian approach results in a wider confidence interval as obtained from the posterior predictive distribution.

159
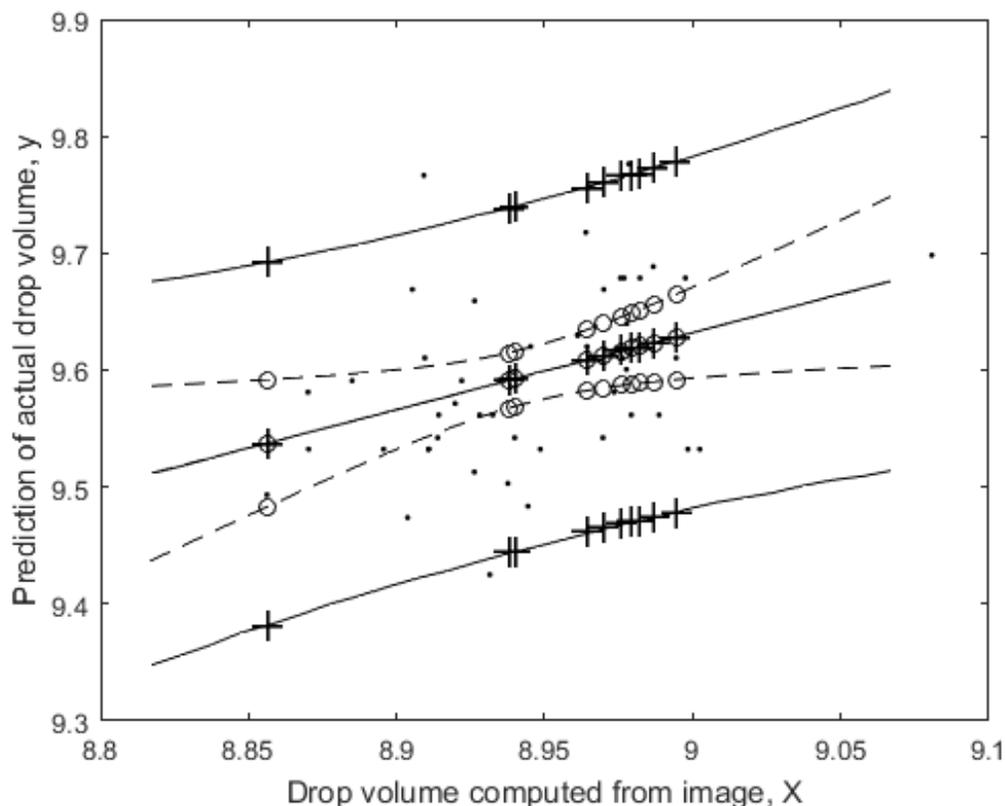
Figure 7.4. Predicted mean value, 95% confidence bounds of ordinary linear regression (dashed line, circle) and Bayesian linear regression (solid line, cross). Dots correspond to points to the associated actual/computed drop volume vectors.

7.7 Conclusions

In this study a Bayesian approach was developed for estimation of product parameters from observations made with offset; prior information from the pharmaceutical manufacturing system was used to update future estimates of drop volume output. The parameter estimates from the Bayesian posteriors provide reasonable estimates for items such as mean, variance, but with a more diffuse density than that observed in the data due incorporated uncertainty, which is also reflected in the respective confidence intervals of the ordinary/Bayesian linear regressions. For implementation in the dropwise manufacturing process, model complexity must at minimum reflect potential disturbances in the process, as demonstrated by peculiarities observed in using the on-line imaging system as the sole process monitor for suspension printing. An improved model

necessitates additional explanatory variables such as concentration, which may assist in associating potential causes with experimental behaviour observed with suspensions.

7.8   References

[1] L. Hirshfield, E. Içten, A. Giridhar, Z.K. Nagy, & G.V. Reklaitis, 2015, Journal of Pharmaceutical Innovation, 10, 140-155, 2015

[2] R.J. Furbank & J.F. Morris, 2004, Physics of Fluids, 16, 1777-1790.

[3] E. Içten, A. Giridhar, L.S. Taylor, Z.K. Nagy, and G.V. Reklaitis, 2015, Journal of Pharmaceutical Sciences, 104, 1641-1649

[4] Radcliffe, A. & G.V. Reklaitis, 27th European Symposium on Computer Aided Process Engineering, Oct. 1-5, 2017, Barcelona, Spain

[5] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, & D.B. Rubin, 2014, Boca Raton, FL: CRC press.

[6] S. Chib & E. Greenberg, 1995, The american statistician, 49, 327-335.

8. BAYESIAN HIERARCHICAL MODELING FOR ONLINE PROCESS MONITORING AND QUALITY CONTROL

8.1 Abstract

With the increasing availability of real-time process data, the Bayesian framework is an adaptable means for dealing with complex models and data structures which can provide readily interpretable outputs for first-principles and hybrid models. Applied to process monitoring and control scenarios, Bayesian models have the capability to provide informed decisions about quantities of interest derived from the online process measurements through the use of probability as a measure of uncertainty. This work presents one such application, wherein real-time image data is used to predict the content of pharmaceutical dose units produced by an additive manufacturing process; a probabilistic quality control strategy is therefore enacted at the level of each dosage unit. Utilizing a hierarchical modeling approach, we are able to draw upon a large historical database to model sources of uncertainty across a wide range of conditions, thus yielding more reliable predictions for future data at in-sample and out-of-sample conditions.

8.2 Introduction

As dictated not only by market pressures, but also by regulatory statute, pharmaceuticals manufacturers must ensure the quality of their products. Conventionally, this has been achieved by drawing a small number of samples from each batch or lot and performing a full assay, the results of which determine whether it can be released to market or must be incinerated. Improvement to this scheme could be realized though implementation of model-based quality control strategies, though it is acknowledged that difficulty arises in mapping available measurements to the critical quality attributes of the resultant product, due to both complex physics and monitoring limitations [1, 2, 3].

With technological advances, various online monitoring tools may be capable of providing measurements which could enable predictive modeling that bridges the gap, albeit not necessarily in the realm of physics-based models with which the community is acquainted. Such data-driven modeling approaches, particularly, neural networks and principal component analysis, are gaining increased acceptance in the process systems community [4, 5]. However, as identified by these

162

authors, these approaches lack ease of interpretation in comparison to first principles models, do not deal with uncertainty in a satisfactory manner, and require a considerable amount of data in order to give reasonable results. In contrast, Bayesian modeling avoids these pitfalls. The scientific insight from the known physics can be directly added to a model through the prior distributions or model structure, and the unknown components determined by the data, with the possibility to yield reasonable first principles interpretations. Moreover, the use of probability as a measure of uncertainty is directly interpretable in the Bayesian paradigm, thereby providing a ready framework for decision analysis. For models with many parameters and multilayered structure, the Bayesian approach provides ready guidelines for dealing with unknown parameters, hence, it provides the freedom to set up complex models which would not otherwise be possible [6, 7].

The application considered in this work involves monitoring and quality control strategy for a pharmaceutical manufacturing process which produces dosage units through drop printing, that is, drug products are formed by the sequential addition of drops to a substrate. Work of prior colleagues created a real-time automation system which manages temperature, substrate movement, and drop generation [8, 9, 10], and an online image acquisition system which captures a digital photograph of each drop added to each dose [11]. As a monitoring tool, online image acquisition has the potential to provide the required information with which to make real-time quality control decisions, provided that automated image analysis algorithms and models are deployed to support it.

In this manufacturing scheme, the volume of material deposited in each capsule directly determines the quantity of active drug in the dosage unit (otherwise referred to as the dosage strength, or dose potency level). Accordingly, to produce dose units at a desired dosage strength, a target specification is set for drop volume (given predetermined ink concentration and number of drops per unit). By monitoring drop volume, it is therefore possible to make a model-based determination as to whether each dose is within the tolerance limits of the nominal dosage strength intended by the operator. Coupling this monitoring scheme with Bayesian models for probabilistic prediction would enable quality assurance in real time, leading to gains in safety and efficacy in the context of flexible, small-scale manufacturing in a compounding pharmacy, whereas in large-scale manufacturing, the ability to discard out-of-specification dosage units would likely lead to economic gains, as batches and lots would be more likely to pass content uniformity testing. One notes that though the work here is concerned with dosage strength (or, for a newly printed batch,

content uniformity) as perhaps the most important critical quality attribute (CQA), the approach could be readily applied to other CQAs.

To phrase the matter succinctly, the Bayesian approach enables one to utilize both the historical data and the real-time data (current observations) in order to make risk-based decisions under uncertainty in a rigorous manner. The strength is that this permits treatment of uncertainty in all aspects of the model, rather than to arbitrarily choose a handful of potential sources while keeping others fixed [2]. Though not discussed in Tabora *et al.*, Bayesian methods are particularly relevant to prediction for future data at in-sample and out-of-sample conditions, wherein hierarchical regression models provide the proper conceptual framework for uncertainty quantification and forecasting [6, 7].

To distinguish between a hierarchical model and non-hierarchical model, consider a simple example in which one claims that the observed data in a batch of experiments is controlled by a parameter related to particle properties, and one has multiple batches of experiments from four distinct particle types. In a non-hierarchical model, there are two possible extremes: claim that either a single parameter describes all batches or each particle type exerts a unique effect, and hence deserves an independent parameter. In a hierarchical model, one provides a unique parameter for each batch and estimates the hyperparameters of the distribution which generated these parameters; this can be represented schematically as presented in Figure 8.1.
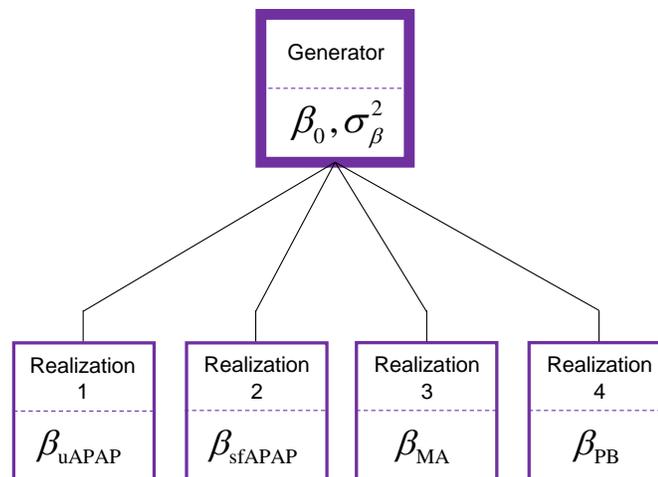
Figure 8.1. Each realization corresponds to a cluster of the data from which the respective parameter is estimated.

Furthermore, a hierarchical modeling approach enables one to treat the data as is naturally structured, to thereby estimate effects across experiments, strata or clusters. The basis for hierarchical modeling can be found in the concept of exchangeability, which is the theoretical justification for treatment of repetitions from a single experiment, $y_1,...,y_n$, as independent and identically distributed. In a hierarchical model, one treats the model parameters themselves as exchangeable across multiple realizations from a group [6]. This is the natural way to construct models when one has multiple realizations of datasets which are stratified at multiple levels and is best implemented in the Bayesian framework in order to sensibly estimate the large number of parameters which result.

8.3   Background

A detailed description of the dropwise additive manufacturing platform used to produce the batches of drug products which form the dataset for this work can be found in [8, 9, 12]. Additional information on the image acquisition system can be found in [13] where component geometry and drop physics are considered in greater detail. For this work, each dose unit was produced by printing of non-Brownian suspensions into empty capsules, with the final quantity of active drug in each capsule determined by the volume(s) of the drop(s) deposited in it and the composition –drug particle concentration – of the ink; detailed descriptions of aspects relevant to suspension printing may be found in [12, 14] .

The training dataset consists of 41 batches of dose units, produced using 4 particle types and 3 fluid types, at a variety of dose potency levels and particle loadings. An exhaustive list of conditions is provided in the supplementary material, which includes the entire dataset; this dataset was gathered by the authors and was unpublished at the time of writing. Of the complete dataset constituted by the 41 batches, one half is used as the training dataset, and the other half is held in reserve to test probabilistic prediction for out-of-sample data from in-sample conditions.

Not included in the training data was the dataset henceforth referred to as out-of-sample data from out-of-sample conditions, which was provided by PinPoint Pharma [15] . This dataset consists of 5 different batches of doses, in which each batch was produced from a new particle type – Hormones (ethinylestradiol and norgestrel), Ibrutinib, Loratidine, Tacrolimus, Warfarin. A new carrier liquid was used in the production of these batches but was the same for all 5. These

batches of doses were produced on the same dropwise additive manufacturing process, hence the data collection procedure was the same.

8.4   Data Structure and Image Processing

8.4.1   Data Structure

The data which fills out the 41 unique batches consists of the real image data acquired during production of ≈1500 total doses at 41 different sets of conditions. At each condition, 25-50 doses were produced by the dropwise additive manufacturing process, and the online imaging system acquired one photo for each drop deposited to each capsule. Given that the data structure has direct bearing on the model structure, we provide a visualization of this in Figure 8.2, along with a standard notation for the batches, doses, images and drops.
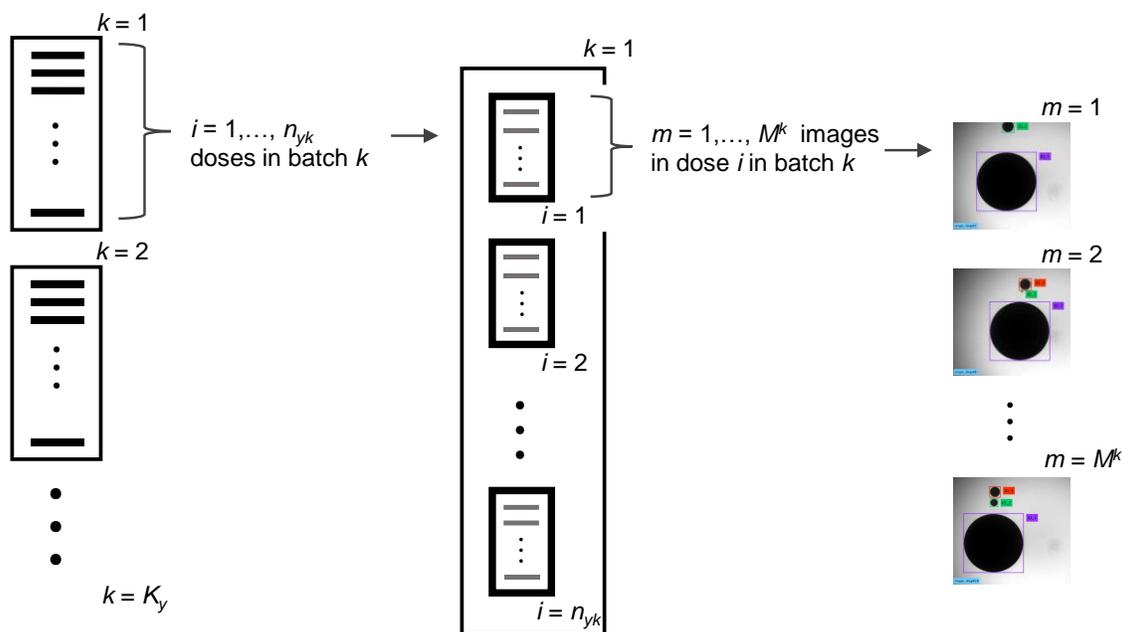
Figure 8.2. Expansion of batch data object (left-most) into dose data objects (middle), each of which expands into image data objects (right-most), represented here by drop images themselves.

Each image is a many-membered data object, each dose is comprised of its respective image data objects, and each batch is comprised of the dose data objects. The total dataset is then comprised of batch data objects. Batch data objects also contain the respective physical

measurements and physical property data, in addition to numerous metainformation members. All together, the members in the batch data objects enable the automatic generation of hierarchical groups, based on the sets desired.

From the long list of quantitative and categorical items present in each image data object, there exist objects – drops – within the image about which the respective information is organized. These objects fall into two categories, essentially: primary drop, i.e. the largest drop and bottommost (as it is this drop that triggers the image acquisition by interrupting the fiber optic signal), and secondary, tertiary, etc. drops, collectively referred to as satellite drops, or satellites, which are by definition all drops that are not the primary drop. This organization enables the construction of explanatory variables as described in Section 8.4.5; a visualization of the contributions made by each drop to the vector of explanatory variables respective to each image data object is provided in Figure J.1.

8.4.2   Image Processing

Automated analysis of the images occurs by a sequence of computational steps which may be divided into two major operations: edge detection and object extraction. The edge detection step takes the original grayscale image and transforms it to a binary edge map; a detailed description of the adaptive edge detection algorithm used here is available in [16]. The binary image, or edge map, that results from applying the edge detector to each image is the starting point for the combined object detection and information extraction step. In the object extraction step, many operations are performed to extract quantitative information from the edge map – this begins with image segmentation as a means of object detection, followed by quantitative computations for numerous "physical" attributes of the drops.

A full description of the information extraction steps is provided in Appendix J. In actuality, each image has much more information extracted from it than the items used to create the its explanatory variable vector, $X_i^{k,m}$; a partial list is given in Figure J.3, with illustrations of the variables provided in Figure J.4 and J.5.

8.4.3 Real Volume by Gravimetric Measurement & Related Modeling Considerations

As determined by the ink formulation and dose potency level, a prespecified number of drops was added to each capsule, after which the dose unit was considered complete; the deposited mass was determined by gravimetric analysis on a balance with a readability of 10 micrograms (the minimum mass deposited in a single capsule was approximately 9 milligrams, and the maximum was approximately 500 milligrams). Thus, regardless of the number of drops deposited in each capsule, the measurement of deposited mass was made for the summed drop volume. Given that the drop generation frequency during normal operation is approximately 10 drops per second, the summed mass is the most practical gravimetric measurement frequency. The empty capsules were individually pre-weighed and each capsule was stored in a uniquely labeled microcentrifuge tube; therefore, the gravimetric measurement of deposited mass (in each capsule) reflects only the mass of deposited liquid, with potential errors related to something such as assumption of a standard empty capsule mass categorically eliminated. In essence, the only potential error is that introduced by the balance itself, but one has little choice other than to use the gravimetric measurements as-is, and allow the model itself to include whatever measurement uncertainty there is.

8.4.4 Model Concept

Due to the method by which deposited volume (of all the drops added to the capsule) was measured, and given that the batches range from 1-49 in the number of drops per capsule, one must therefore aggregate the $m = 1, \ldots, M^k$ images in each dose. For the 21 out of 41 batches in which there is one drop per capsule ($M^k = 1$), one clearly has no need to aggregate the image information, and thus $X_i^k = X_i^{k, M^k}$ directly in both of the following modeling approaches, as the single image maps directly to the single deposited volume. When $M^k > 1$, aggregating the images in each dose is necessary, as only one measurement exists for the sum total of individual volumes; while it is true that the sum of drop volumes in the $m = 1, \ldots, M^k$ images corresponds to the summed deposited mass, this may not be strictly true as a satellite drop could have landed outside of the capsule.

Bearing in mind that the objective of the model is to predict the deposited volume in each capsule from the corresponding $m = 1, \ldots, M^k$ images, there are two possible approaches available:

1. **ModelDrop:** Predict the deposited volume respective to each image, and thereby predict the deposited volume in each capsule. That is, model the images themselves, which in the single drop case has an unambiguous interpretation, as noted above, but in the multi-drop case, requires that a single averaged composite image be formed for each dose.

2. **ModelDose:** Predict the deposited volume in each dose directly, modeling the aggregated information from the images, in order to predict the sum total deposited volume in each dose.

For reasons elaborated below, the distinction between the two approaches is more subtle than whether to simply average or sum the proposed explanatory variable vectors $X_i^{k,m}$, for the $m = 1,\dots,M^k$ images in each dose i. The two modeling concepts are illustrated in Figure 8.3, with specific details presented in Sections 8.4.6. and 8.4.7.

ModelDrop: average $X_i^{k,m}$ over the M^k images in each dose i to obtain X_i^k

$$\frac{y_i^k}{M^k} = X_i^k \beta$$

ModelDose: exclude incompatible explanatory variables, then sum $X_i^{k,m}$ across the M^k images in each dose i to obtain X_i^k

$$y_i^k = X_i^k \beta$$
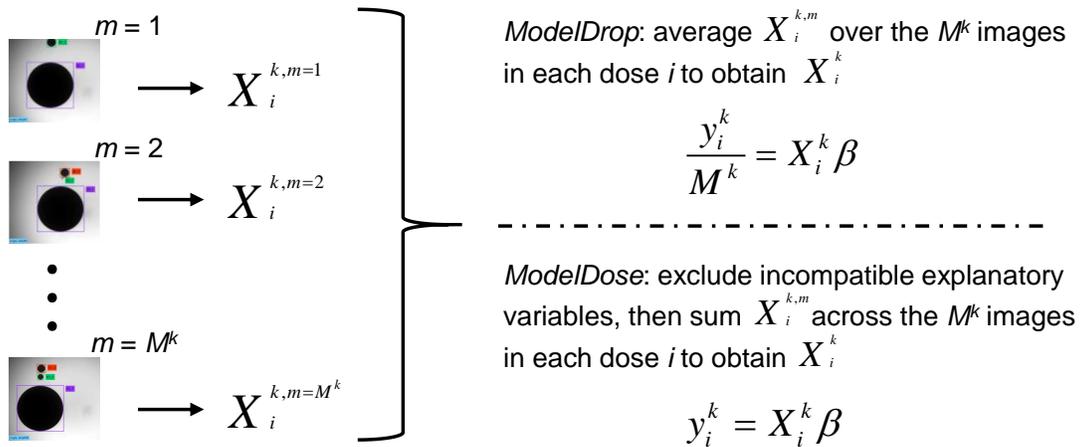
Figure 8.3. Model Concept

8.4.5 Extracted Image Information into Explanatory Variables for Each Model

In the context of modeling the images themselves, (i.e. ModelDrop), one should add explanatory variables for both the primary drops, and for the satellite drops in order to form $X_i^{k,m}$. For each drop, primary or satellite, all the same information is available; for clarity of description below, the primary drop is considered first, then the satellite drops are dealt with.

With respect to the primary drop, a measure of drop volume,

$$V = \tfrac{\pi}{4} \int D(z)^2 \, dz = \tfrac{\pi}{4} \sum_{l \in \{\text{primary drop}\}} D(z_l)^2 \, \Delta z_l$$

169

computed from the image based on the assumption of axial symmetry, is included as an explanatory variable. As the optics suggest that a small correction may be needed due to depth of field effects as objects move increasingly far from the center of the image, the Euclidean distance from the centroid of the drop to the image center is also included as an explanatory variable.

For the satellite drops, one is faced with a choice: add separate sets of explanatory variables for each, or reduce all of the satellites in a given image to a single set. If the former option were enacted, this would present some disagreeable qualities: a unique parameter is created for each explanatory variable for each satellite, which seems to be lacking in justification given that there does not appear to be much to categorically distinguish amongst satellites (cf. primary drop and satellites). One could link the respective parameters in a hierarchy in order to allow the data to determine the extent of pooling, but since individual measurements of the satellites are unavailable and, in fact, one cannot even be certain that the satellites land inside the capsule, this seems unfavorable. Furthermore, one notes that the number of satellites per image varies throughout the dataset (the 12,000 images), with the total satellites very rarely exceeding 2, whereas 1 satellite is common, and 2 less so.

Therefore, the satellite drop explanatory variables are constructed as a composite from the respective information of the satellites present in each image; in essence, a composite satellite object is formed by summing the volumes, and setting the Euclidean distance as the arithmetic average of the satellites present.

Formal definitions for the primary and satellite drop explanatory variables are given below, first denoting the drop-specific information items as:

$$V_i^{k,m,o} = \text{volume of object } o \text{ in image } m \text{ of dose } i \text{ of batch } k$$

$$E_i^{k,m,o} = \text{Euclidean distance from image center to centroid of object } o \text{ in image } m \text{ of dose } i \text{ of batch } k$$

The explanatory variable vector, including a constant intercept, $x_{ij=1}^{k,m} = 1$, for each image is then:

$$x_{ij=2}^{k,m} = V_i^{k,m,o=primary}$$

$$x_{ij=3}^{k,m} = E_i^{k,m,o=primary}$$

for the primary drop, and for the satellite drops:

$$x_{ij=4}^{k,m} = \sum_{o \in \{\text{satellite}\}_i^{k,m}} V_i^{k,m,o}$$

$$x_{ij=5}^{k,m} = \frac{1}{\left|\{\text{satellite}\}_i^{k,m}\right|} \sum_{o \in \{\text{satellite}\}_i^{k,m}} E_i^{k,m,o}$$

The peculiarities of partially observed satellite drops necessitate additional explanatory variables. Specifically, a categorical variable could be included to indicate whether one or more of the drops cross the image border,

$$x_{ij}^{k,m} = \begin{cases} 1 & \text{if any o} \in \{\text{satellite}\}_i^{k,m} : z_l = z_{border} \\ 0 & \text{otherwise} \end{cases}$$

though it is preferable to instead use explanatory variables extrapolated from the available information, as described below, based on the apparent shape, position and volume of the portion in the image.
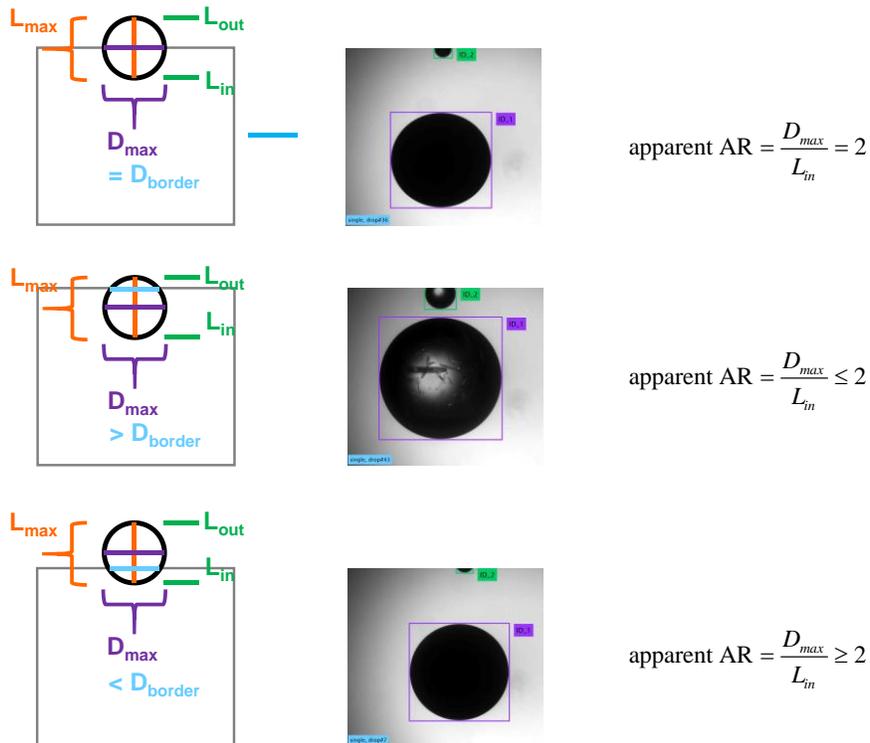
Figure 8.4. Idealized representations (Left), with real examples (Right) of images with satellites only partially observed.

As a first guess, the missing satellite volume can be approximated by considering an ideal case for the partially observed spherical entity (Figure 8.4, Left, Top), in which the real volume is

$V = \frac{\pi}{4} \int_{L_{out}}^{L_{in}} D(z)^2 \, dz$, and the apparent volume is equal to $V = \frac{\pi}{4} \int_0^{L_{in}} D(z)^2 \, dz$. As the real aspect ratio is equal to 1, thus, $D_{max} = L_{out} + L_{in}$, the apparent aspect ratio can be rewritten as:

$$AR = \frac{L_{out} + L_{in}}{L_{in}}$$

Therefore, $\frac{L_{out}}{L_{in}} = AR - 1$, and if apparent aspect ratio is equal to 2, then the missing volume is exactly:

$$CylMsV = \left(\frac{L_{out}}{L_{in}}\right) \frac{\pi}{4} \int_0^{L_{in}} D(z)^2 \, dz = (AR - 1) * \text{apparent Volume}$$

Of course, this is only true in this limited case, in which all the conditions above are met. If $1 \le AR \le 2$ and the object is indeed a sphere (as in Figure 8.4, Top, Middle), then the missing volume would be better approximated using the available dimensions to compute the volume of the missing/unobserved spherical cap,

$$SphMsV = \frac{\pi (L_{out})^2}{3} \left(3 \left(\frac{D_{max}}{2} \right) - L_{out} \right)$$

In which $L_{out} = (AR - 1) L_{in} = D_{max} - L_{in}$.

However, there is uncertainty as to which approximation should be used, in particular due to uncertainty in the real shape of the satellites. If the shape is non-spherical, or the apparent aspect ratio is ≥ 2, as occurs in cases such as Figure 8.4, Bottom, then the simpler approximation would likely yield a better approximation to the unobserved volume than the spherical cap method, though it is acknowledged that it will be an under/over-estimate in most instances.

From the available information it is possible to distinguish between the possible cases illustrated in Figure 8.4, and therefore to apply the what seems to be the better approximation to each satellite object with a border crossing. As shown in the line drawings of Figure 8.4, the cases are easily distinguished if the maximum diameter of the object is known – unfortunately, this level of certainty is not available in most real cases. However, an equivalent conditional can still be formulated based on the z-position of the apparent maximum diameter: if the apparent maximum diameter occurs at the border, then the case is always as in Figure 8.4, Bottom, and the simpler cylindrical approximation should be used,

$$CylMsV_i^{k,m,o} = \begin{cases} \left(AR_i^{k,m,o} - 1\right) V_i^{k,m,o} & \text{if } o \in \{\text{satellite}\}_i^{k,m} : z_l = z_{border} \wedge z_{D_{max}} = z_{border} \\ 0 & \text{otherwise} \end{cases}$$

Whereas if the apparent maximum diameter occurs below the border, then the spherical cap approximation should be applied,

$$SphMsV_i^{k,m,o} = \begin{cases} \dfrac{\pi(L_{out})^2}{3}\left(3\left(\dfrac{D_{max}}{2}\right) - L_{out}\right) & \text{if } o \in \{\text{satellite}\}_i^{k,m} : z_l = z_{border} \wedge z_{D_{max}} > z_{border} \\ 0 & \text{otherwise} \end{cases}$$

As these are mutually exclusive for each object, we are then able to add the two explanatory variables:

$$x_{ij=6}^{k,m} = \sum_{o \in \{\text{satellite}\}_i^{k,m}} CylMsV_i^{k,m,o}$$

$$x_{ij=7}^{k,m} = \sum_{o \in \{\text{satellite}\}_i^{k,m}} SphMsV_i^{k,m,o}$$

to account for the missing satellite volumes.

Additional explanatory variables are included in order to account for the crossing of the top and bottom border by the primary drop, which, though rare, does occur in a few of the batches; various instances of this are presented in Figure 8.5.
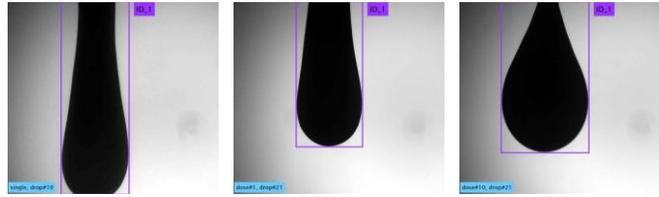
Figure 8.5. Examples of primary drop crossing border. Left: MA, NEO895, particle fraction 0.4, single drop; Middle: MA, NEO895, particle fraction 0.4, 100mg dose; Right: PB, NEO895, particle fraction 0.3, 100mg dose.

As it is the leading edge of the primary drop that triggers the image acquisition, the correction for the border crossing at the bottom should almost always be approximated as a spherical cap. However, given that the drop can be far from spherical, the arguments based on aspect ratio do not apply. Instead, we use the maximum diameter and the distance between the z-position of the maximum diameter and the border to coarsely approximate the length extended outside the image:

$$L_{out} = \frac{D_{max}}{2} - \left(z_{D_{border}} - z_{D_{max}}\right)$$

and compute the missing volume as:

$$BotMsV_i^{k,m,o=primary} = \begin{cases} \dfrac{\pi L_{out}}{3}\left(3\left(\dfrac{D_{border}}{2}\right)^2 + \left(L_{out}\right)^2\right) & \text{if } o = \text{primary drop} \wedge z_l = z_{border,Bottom} \\ 0 & \text{otherwise} \end{cases}$$

in which we use the absolute value of L_{out} to avoid possible negative values. Thus, we add the explanatory variable $x_{ij=8}^{k,m} = BotMsV_i^{k,m,o=primary}$.

It is less clear as to how one should approximate the missing volume when the primary drop crosses the top border – the only reasonable conclusion seems to be that considerable volume may be unobserved. Without additional information, not much can be done other than to propose an approximation that is linearly proportionate in the hopes that the Bayesian procedure can determine the appropriate coefficient. To this end, we employ a simple proposal: the length of the missing cylindrical volume is equal to the diameter at the border, and the fluid volume outside the image mirrors that inside the image, thus, the explanatory variable for the missing volume at the top border is:

$$x_{ij=9}^{k,m} = TopMsV_i^{k,m,o=primary} = \begin{cases} \dfrac{\pi}{4} \displaystyle\int\limits_{z_{border,Top}}^{L=D_{border}} D(z)^2\, dz & \text{if } o = \text{primary drop} \wedge z_l = z_{border,Top} \\ 0 & \text{otherwise} \end{cases}$$

A final explanatory variable must be added in order to account for 2 out of the 41 batches which were produced at notably different conditions that resulted in two drops of approximately equal size, one of which was captured in the images, as shown in Figure 8.6.
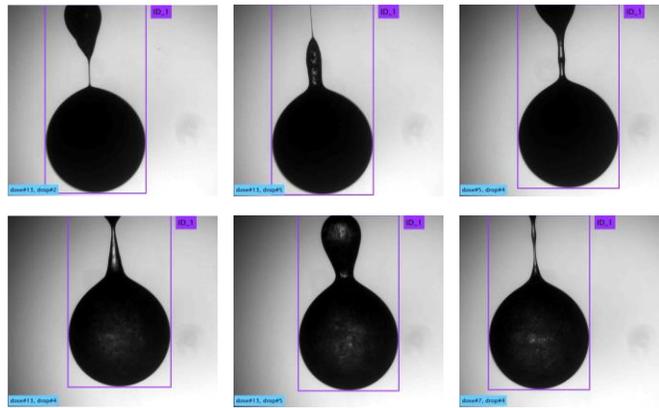
Figure 8.6. Example of drops from batches. Top: MA, PEG300, particle fraction 0.087, 10mg dose; Top: PB, PEG300, particle fraction 0.083, 10mg dose.

8.4.6 ModelDrop

In the context of modeling the image information in order to predict drop volume, the vector of explanatory variables respective to each image is:

$$X_i^{k,m} = \begin{bmatrix} x_{ij=1}^{k,m} & x_{ij=2}^{k,m} & x_{ij=3}^{k,m} & x_{ij=4}^{k,m} & x_{ij=5}^{k,m} & x_{ij=6}^{k,m} & x_{ij=7}^{k,m} & x_{ij=8}^{k,m} & x_{ij=9}^{k,m} \end{bmatrix}$$

with $m = 1,...,M^k$ vectors respective to each dose i in batch k, as illustrated in Figure 8.3. The average composite, $\bar{X}_i^k$, is constructed from the individual image explanatory variable vectors in the following manner:

$$\bar{x}_{ij}^k = \frac{1}{M^k} \sum_{m=1}^{M^k} x_{ij}^{k,m} \text{ , for } j = 1,2,3,4,6,7,8,9$$

$$\bar{x}_{ij}^k = \frac{1}{\left|\left\{\forall m : \{\text{satellite}\}_i^{k,m} \neq \varnothing\right\}\right|} \sum_{\left\{\forall m:\{\text{satellite}\}_i^{k,m} \neq \varnothing\right\}} x_{ij}^{k,m}, \text{ for } j = 5$$

For $j = 5$, the latter definition is necessary in order to preserve the interpretation as an average measure across the images, since not all images contain a satellite drop. If one were to include simple indicator categorical variables for the various border crossings, then one has the choice of interpreting these in the multi-drop case as a frequency of occurrence in terms of the composite, if the first definition above is applied, or as simply the categorical indicator if the second definition is used.

8.4.7 ModelDose

In the context of modeling the aggregate information from the $m = 1,...,M^k$ images in each dose, a vector of explanatory variables for dose i in batch k is created as:

$$X_i^k = \sum_{m=1}^{M^k} X_i^{k,m}$$

However, some of the explanatory variables in $X_i^{k,m}$ must be excluded or altered. Specifically, the Euclidean distance measures, $x_{ij}^{k,m}, j = 3,5$, lose their meaning upon summation as a physical interpretation is not valid, hence, they are excluded. Additionally, the constant at $j = 1$ must be set to 1 for all X_i^k; to sum the constant would be to use the number of drops per dose (M^k) instead, which may be undesirable. Use of categorical variables such as the border indicators is still possible, but due to the difference in scales between a single and multi-drop dose, the only

reasonable choice would be to sum the categorical variables respective to each image; whether this adequately takes into account differences in scale in a linear model is questionable.

8.4.8   Batch-wide Variables

To improve the modeling of variation between batches and between doses, it is possible to introduce batch-wide variables for $\bar{x}_{ij}^k$, $j = 2, 4, 6, 7, 8, 9$. To illustrate, take $\bar{x}_{ij=2}^k$ as an example for which two variables might be constructed from all the $\bar{x}_{ij}^k$'s in the batch, that is, for the $i = 1, \ldots, n_{yk}$ doses in the batch. First, the batch-wide mean of the chosen variable is computed as:

$$\overline{\overline{x_{ij=2}^k}} = \frac{1}{n_{yk}} \sum_{i=1}^{n_{yk}} \bar{x}_{ij=2}^k$$

Then, the deviation from the batch-wide mean,

$$X_{ij=10}^k = \bar{x}_{ij=2}^k - \overline{\overline{x_{ij=2}^k}}$$

is added to the vector of explanatory variables, and the original variable located at $\bar{x}_{ij=2}^k$ is replaced by $\overline{\overline{x_{ij=2}^k}}$ for all doses in the batch. It should be noted that introduction of a batch-wide variable is suitable to $j = 2$, whereas for the other variables which do not take non-zero values consistently, this is likely to be detrimental.

8.5   Bayesian Modeling Approach; Model Structure, Setup and Solution Method

8.5.1   Modeling Approach: Hierarchical Linear Model

From the information extracted from the images, one expects that the deposited volume would be a linear function of the computed image volumes for the drops present. The inherent uncertainty arises in conversion from computed pixel volume to real volume; this is affected by numerous factors, amongst them object blur, position within image, distance of object from camera, validity (or lack thereof) of axial symmetry assumption. Some of these variables can be included in the model, when an appropriate explanatory variable can be added, but others must be accounted for through other means, namely, corrections determined by the data, which are best implemented in a hierarchical model.

The hierarchical model takes the general form given below:

$$y \sim \mathrm{N}\left(X\beta, \Sigma_y\right)$$

$$\beta \sim \mathrm{N}\left(\beta_0, \Sigma_\beta\right)$$

In which y is the vector of observed data, X, the explanatory variable matrix, β, the predictors. The likelihood is assumed to have a normal distribution, with mean $X\beta$ and covariance Σ_y; the prior distribution of β has prior mean β_0 and covariance Σ_β. The structure of the data covariance matrix is assumed to be diagonal,

$$\Sigma_y = Diag\left(W_y \sigma_y^2\right)$$

in which $W_y \in \mathbb{R}^{n_y \times K_y}$ is an indicator matrix of 0's and 1's denoting the batching of the data under a variance parameters in the model, and the variance parameter vector is $\sigma_y^2 : \sigma_{yk}^2, k = 1, \ldots, K_y$. As the batches of doses were produced separately, with full cleaning of the equipment between production runs as described in Radcliffe *et al.,* 2019, there is little reason to believe that correlations exist between the batches, hence, good theoretical justification for use of a diagonal data covariance matrix; we do however allow for unequal variances, as this is well-supported by the physics. The structure of the parameter covariance matrix is also assumed to be diagonal,

$$\Sigma_\beta = Diag\left(W_\beta \sigma_\beta^2\right)$$

in which $W_\beta \in \mathbb{R}^{J \times K_\beta}$ is an indicator matrix denoting the batching of the parameters into hierarchical K_β groups, with $\sigma_\beta^2 : \sigma_{\beta k}^2, k = 1, \ldots, K_\beta$ the vector of hierarchical variance parameters. The data and hierarchical variance parameters are assigned inverse-χ^2 prior distributions,

$$\sigma_{yk}^2 \sim \mathrm{Inv}\text{-}\chi^2\left(v_{yk}, \sigma_{0yk}^2\right)$$

$$\sigma_{\beta k}^2 \sim \mathrm{Inv}\text{-}\chi^2\left(v_{\beta k}, \sigma_{0\beta k}^2\right)$$

in which we set the prior degrees of freedom and prior variance as: $v_{yk} = 0$, $\sigma_{0yk} = 0$ and $v_{\beta k} = -1$, $\sigma_{0\beta k} = 0$, which corresponds to a noninformative uniform prior, $p\left(\sigma_y, \sigma_\beta\right) \propto \prod_k^{K_y} \sigma_y^{-1}$, and for components of β with noninformative flat uniform prior distributions, $\sigma_{0\beta k} = \infty$ [17].

In a non-hierarchical model, the prior distributions on β's are such that for each β_j there is a corresponding prior, and the parameters are not batched whatsoever – each is distinct and unique; by convention, the prior distributions on the β_j's are typically assigned noninformative flat uniform distributions. This constitutes a special case of the above [6, 18, 19].

Motivation for Hierarchical Models

In a hierarchical model, we seek to take advantage of inherent structure in a problem in order to improve the fit to the training data and also to improve the predictive accuracy. Structure in a problem is commonly encountered as the natural categories into which the data falls [6, 7]. This categorical information can be treated in a non-hierarchical fashion, adding explanatory variables to the explanatory variable matrix, but then one must be more certain as to the type of structure each problem should take. Often, this is difficult as we know the data bears some relationship to the categorical structure, but the precise relation is not definitely known [20]. Of course, one can supply a structure and proceed with the non-hierarchical model, but then the model is constrained by beliefs about which one is uncertain. A better approach seems to be to let the data itself determine the precise structure, rather than to impose our beliefs – such an approach is achieved through hierarchical models. [21, 22]

To make a concrete example from the abstract perspective presented above, consider in this model the example of adding explanatory variables (and predictors β_j) for a known categorical partition of the data based on particle type. Noting that there exist multiple y^k, X^k batches for each particle type, each batch of which has some distinct covariates, how should the categorical distinction which arises due to particle type be treated? The following paragraphs provide four potential answers to this question in order to make the distinction between modeling approaches, and to motivate the need for a hierarchical model.

1. The first option is to simply ignore the categorical distinction in the data, but this seems to directly conflict with what we know from the physics – that is, that particle properties affect drop formation.

2. The second option is to use a single explanatory variable distinguished by levels to take effects into account. This could be useful if the underlying physics were amenable to it, for example if for 4 particle types, 4 categories existed, the basis for which was something such as an average particle size/shape measure. Furthermore, suppose the magnitude of the effect increased linearly with the average size/shape measure. Then, we could assign to each X_i^k vector an explanatory variable which would indicate its particle effect category, e.g. 1,2,3,4 respective to particle type. Unfortunately, this relies on several assumptions: 1) that the appropriate distinguishing characteristic between the particle types is readily identifiable (e.g. average particle size), 2) that the effect has a linear relationship with the

particle measure, and 3) that the proposed categorical structure is appropriate for the effect (and ideally applicable outside the data being fit). Often, one does not have sufficient confidence to make assumptions such as these. Proceeding with the arbitrary imposition, the model would become infused by our beliefs, but perhaps the data may tell a different story – other factors may explain the dependence on particle type, the levels selected may not be appropriate to the problem, etc.

3. The third option is to add multiple explanatory variables which serve as indicators for the categorical structure respective to particle type. In this example, four particle types would correspond to 4 additional columns in the explanatory variable matrix. For each data point, one of the 4 columns would be set to 1, and the rest to 0; the index of the column set to 1 corresponds to the same particle type across all data points. That is, particle type 1 corresponds to column index 1, particle type 2 to column index 2, and so on.

   This allows for a unique-to-particle type correction to be made to all the observations respective to a given particle type. However, one must consider: is this the correct approach? – Does each particle type merit its own parameter? We not that this also makes the model more specific to the data it is fit to, thus, we lose some ability to generalize, specifically for future data (X^k's) that are from particle types not included in the training data. Certainly, one could assume that there is sufficient similar to one of the training data particle types and simply re-use the parameter, but such an approach is not rigorous.

4. The fourth option is the hierarchical model. One adds multiple explanatory variables which serve as categorical indicators of particle type as described in the third option. However, rather than assigning each explanatory variable its own unique parameter β_j, with a unique prior distribution, the group of β_j's is assigned a common prior distribution, thereby linking them in a hierarchy. In doing so, one allows the data to determine to what extent the β_j's are actually unique parameters, as the hierarchical linkage can allow for more or less pooling of the data. In contrast, the second option represents complete pooling of the information, i.e. all effects can be described by a single parameter. The third option represents no pooling of the information, in which each effect is unique, and thus, unique parameters are assigned to each particle type. In the fourth option, the common prior which creates the hierarchical group allows the data to determine the extent to which pooling of information should be performed – if the posterior variance of the hierarchical group is

179

small, this approaches the complete pooling limit; if the posterior variance is large, then this approaches the no pooling limit.

The use of hierarchical groupings for the categorical structures also provides another important advantage: for future data which is not within the particle types used as training data, there are no conceptual difficulties in generating new parameters, $\tilde{\beta}$, which correspond to a new group from the hierarchy. Thus, we avoid the inherent problems of the second/third option in making predictions for unseen particle types. In addition to allowing the data to determine the extent of pooling, thereby avoiding explicit bias introduced by one's (uncertain) assumptions, addition of structure to the problem through hierarchical grouping enables the flexibility we seek from the model. Specifically, the generation of new $\tilde{\beta}$'s from the hierarchical groups enables one to make predictions for unseen data in a rigorous manner. This is the expected use of the model, where we will predict deposited volumes from images acquired at conditions which were not included in the training data. Of course, the model can also be used for prediction at conditions that were included in the training data, but we would like to have a more robust model, if possible.

The essence of this modeling approach is to let the data determine the effective number of parameters (and, of course, the parameter values themselves), rather than to impose our beliefs arbitrarily to cases in which we are uncertain.

Hierarchical Groupings: Explanation and Examples

The example provided above contains the essential concept for the hierarchical groupings used in the models presented in this work. Thus, as for the categorical groupings based on particle type, we also consider groupings based on fluid type, rheological regime, and dose potency level – these are the basic categorical distinctions which arise from the physics.

In addition to these basic categorical groupings, interactions between the variables are considered in the model, as expected from the physics. An example of this is a (particle type $\times$ fluid type) specific correction, physically justifiable by the interaction of particle properties with fluid properties – again, one need not be specific as to whether to account for this with a single parameter (complete pooling), or with a unique parameter for each (no pooling). Analogously to the fourth option above, an indicator column is provided for each (particle type $\times$ fluid type) combination, and the respective β_j's are linked in a hierarchy with common mean and variance.

This can be taken one level further, considering interactions such as (rheological regime $\times$ particle type $\times$ fluid type), in which it is posited that there is a Φ-dependent correction specific to each (particle type $\times$ fluid type) combination. The physical justification is clear: rheological regime has a strong effect on drop formation, and is (in simplification) dependent on particle volume fraction (amongst other things, particle and fluid properties in particular), thus, we might account for these effects through the addition of corrections for each instantiation of the categorical (rheology regime $\times$ particle type $\times$ fluid type), denoted ($\Phi \times p \times f$). Once again, since one is uncertain as to the extent of parameterization required, an indicator column is supplied for each realization of ($\Phi \times p \times f$), and the respective β_j's are linked in a hierarchy with common mean and variance. To this end, either 2 or 3 rheological regimes are considered, in accordance with the observed physics for drop formation: low/high (2) or low/medium/high (3).

Taken one level further still, in which interactions of dose potency level with each ($\Phi \times p \times f$) combination are considered in the model, for which the interaction term is (dose potency level $\times$ rheological regime $\times$ particle type $\times$ fluid type), denoted ($\delta \times \Phi \times p \times f$), with division of dose potency level into 3 regions: doses consisting of a single drop each, doses consisting of a few (2-5) drops each, and doses consisting of many (20-49) drops each. The physical basis for distinction based on potency level arises from the consideration that while the images for each dose should be completely representative, there may be effects unaccounted for by the rheological regime, particle property- and/or fluid property-based interactions. Upon consideration, one may expect that there is the possibility for a δ-level specific interaction due to dynamic fluctuation that my occur during production of multiple (10-20) doses, each of which contains many (20-49) drops. This may manifest as small changes in ink composition throughout the production run, during which large quantities of the ink reservoir are consumed. In contrast, production of multiple single-drop doses consumes only a small fraction of the reservoir, thus, it is unlikely that the same type of ($\delta \times \Phi \times p \times f$) interaction as above would be observed.

To account for this in the model, indicator columns are added respective to each ($\delta \times \Phi \times p \times f$) combination, analogously to above. The respective β_j's are linked in hierarchical groupings, but rather than link all the β_j's into a single hierarchy with one common mean and variance, it is preferable to keep with the physical intuition given above. Thus, the appropriate subsets of the β_j's are linked in 3 separate hierarchical groupings, one for each δ-level – that is, the β_j's respective to (δ=*single drop* $\times \Phi \times p \times f$) are linked with common mean and variance, and so on

for δ=*few drops* and δ=*many drops*. In structuring the model, a distinction is made based on the underlying physical basis which is reflected in the mathematical structure – we claim that the β_j's are exchangeable across $(\delta \times \Phi \times p \times f)$ combinations, but only within a given δ-level. In other words, exchangeability holds only within a phenomenological region, thus, each region (each δ-level) necessitates its own hierarchical link.

The aforementioned hierarchical groupings are certainly not the only possible categorical interactions which could be considered, any valid interaction can be added to the model. The physical reasoning underlying each of these can be inferred from the comments on each variable δ, Φ, p, f, as they were added to the expansion of interactions. Essentially, each would posit some unique interaction effect which could be accounted for by addition of group indicator columns, the β_j's of which are linked in a hierarchical grouping. For interactions with sufficient β_j's and the physical reasoning to support it, it is possible to link the subsets of the β_j's in separate hierarchies, as described for the $(\delta \times \Phi \times p \times f)$ example.

8.5.2   Model Structure

In each model considered for this work, one or more of the hierarchical groupings can be added at will – the judgment is up the modeler, who is limited only by what is viewed as reasonable to account for the interactions. The data itself does impose some limitation, as there must be sufficient information to afford at minimum 3 β_j's (3 indicator columns) in order to form a hierarchical grouping, though it is advisable that groupings be constructed with more parameters per batch in order to avoid heavy-tailed distributions when using noninformative priors for the hierarchical variances [17].

The general form of the model consists of batches of observations, $y^k, k=1,...K_y$, with respective non-grouped explanatory variable matrices, X^k, derived from the images. Each batch y^k has $i=1,...n_{yk}$ elements, and each element, $y_i^k \in \mathbb{R}^{1 \times 1}$ has a corresponding vector of non-group explanatory variables, $X_i^k \in \mathbb{R}^{1 \times J_{nongroup}}$. Augmenting the batches together, this yields:

$$y = \begin{bmatrix} y^{k=1} \\ \vdots \\ y^{k=K_y} \end{bmatrix}, \quad X = \begin{bmatrix} X_{nongroup}^{k=1} \\ \vdots \\ X_{nongroup}^{k=K_y} \end{bmatrix}$$

To add a hierarchical grouping, indicator columns will be added to the explanatory variable matrix, X^k, of each batch. The appropriate column will then be assigned values of 1, and the remaining columns values of 0. To illustrate the addition of group-level predictors linked in a hierarchy, an example is presented using the simple particle type grouping; the procedure is analogous for other hierarchical groups.

In the particle type grouping, the data presents a set of 4 particle types: micronized grade APAP (μAPAP), semi-fine grade APAP (sfAPAP), mefenamic acid (MA), and phenylbutazone (PB), thus, 4 indicator columns would be added to the existing columns of X^k. For each batch, k, the respective column of the particle type used to produce that dose would be set to 1, and the other 3 columns to zero. For example, if $p = \text{sfAPAP} \equiv \text{column index } 2$, and the batch was produced from sfAPAP particles, then column 2 of $X_{HG,\,particle\ type}$ at the row indices respective to batch k would be set to 1, and the other 3 columns set to zero (at the same row indices). For all batches $k = 1,...,K_y$, the same 4 columns are added and the appropriate row, columns set respective to the particle type in the batch. Thus, $X^k_{nongroup}$ becomes

$$X^k = \begin{bmatrix} X^k_{nongroup} & X_{HG,\,particle\ type} \end{bmatrix} \in \mathbb{R}^{n_{yk} \times \left(J_{nongroup} + 4 \right)}.$$

A more formal treatment of this is helpful in understanding how more elaborate groupings are added with ease. In general, each batch y^k, X^k has 4 pieces of information which distinguish it:

$$\{\delta - dose\ potency\ level \quad \phi - rheological\ regime \quad p - particle\ type \quad f - fluid\ type\}$$

Thus, a preferable way to write down and conceptualize a hierarchy is using sets. In this context, the particle type hierarchical grouping could be viewed as assigning a unique parameter to the 4 subsets of the index set of all batches ($k = 1,...,K_y$) which can be formed based on particle type distinction:

$$S_{uAPAP} = \{k : p(k) = uAPAP\}$$

$$S_{sfAPAP} = \{k : p(k) = sfAPAP\}$$

$$S_{MA} = \{k : p(k) = MA\}$$

$$S_{PB} = \{k : p(k) = PB\}$$

For each index subset that is non-empty, a corresponding column is provided for indicator variables. Then, proceeding through the batch indices in each subset, e.g. $k \in S_{sfAPAP}$, 1's are

assigned at the respective row indices (i.e. at the respective dose indices i) in said column. Performing this for each subset, and augmenting the columns together yields:

$$\begin{bmatrix} I_{S_{uAPAP}} & I_{S_{sfAPAP}} & I_{S_{MA}} & I_{S_{PB}} \end{bmatrix} = X_{HG, particle\ type}$$

which is the matrix of indicator variables respective to the group-level predictors in the particle-type categorization.

Notice that the number of subsets resulting from the categorization corresponds to the number of columns, and each column corresponds to a particular index subset, e.g. S_{sfAPAP}, of the batches. Thus, the set-based construction can be carried out for the group predictors and hierarchies that consider interactions, for example ($p \times f$):

$$S_{p \times f} = \{ k : p(k) = \text{particle type} \land f(k) = \text{fluid type} \}$$

which for this dataset yields:

$$S_{uAPAP \times NEO895} = \{ k : p(k) = uAPAP \land f(k) = NEO895 \}$$

$$S_{uAPAP \times LPG} = \{ k : p(k) = uAPAP \land f(k) = LPG \}$$

$$S_{sfAPAP \times NEO895} = \{ k : p(k) = sfAPAP \land f(k) = NEO895 \}$$

$$S_{sfAPAP \times LPG} = \{ k : p(k) = sfAPAP \land f(k) = LPG \}$$

$$S_{MA \times NEO895} = \{ k : p(k) = MA \land f(k) = NEO895 \}$$

$$S_{MA \times PEG300} = \{ k : p(k) = MA \land f(k) = PEG300 \}$$

$$S_{PB \times NEO895} = \{ k : p(k) = PB \land f(k) = NEO895 \}$$

$$S_{PB \times PEG300} = \{ k : p(k) = PB \land f(k) = PEG300 \}$$

which are non-empty index sets, along with 4 empty sets:

$$S_{uAPAP \times PEG300} = \{ k : p(k) = uAPAP \land f(k) = PEG300 \}$$

$$S_{sfAPAP \times PEG300} = \{ k : p(k) = sfAPAP \land f(k) = PEG300 \}$$

$$S_{MA \times LPG} = \{ k : p(k) = MA \land f(k) = LPG \}$$

$$S_{PB \times LPG} = \{ k : p(k) = PB \land f(k) = LPG \}$$

Hence, only 8 indicator columns are created for the ($p \times f$) interaction group-level predictors, rather than the 12 possible, since 4 ($p \times f$) combinations do not have any corresponding batches (no data was collected at these conditions). A corresponding β_j is assigned for to each of the 8 columns, and the β_j's are linked together by selecting a set of the index subsets which will share a common

mean and variance. In the particle type grouping, all $4\,\beta_j$'s are grouped into one hierarchy; in the $(p\times f)$ interaction, all $8\,\beta_j$'s are linked in one hierarchy. In contrast, for the $(\delta \times \Phi \times p \times f)$ interaction, the β_j's are split and linked into separate hierarchies in keeping with the physically-justifiable exchangeability.

To illustrate, consider the construction of the model with the (p), $(p \times f)$, $(\Phi \times p \times f)$, and $(\delta \times \Phi \times p \times f)$ group-level predictors:

$$X = \begin{bmatrix} X_{nongroup} & X_{HG,p} & X_{HG,p\times f} & X_{HG,\phi\times p\times f} & X_{HG,\delta\times\phi\times p\times f}^{\delta=\text{single drop}} & X_{HG,\delta\times\phi\times p\times f}^{\delta=\text{few drop}} & X_{HG,\delta\times\phi\times p\times f}^{\delta=\text{many drop}} \end{bmatrix}$$

Which has a corresponding β-vector:

$$\beta = \begin{bmatrix} \beta^{k=1} & \beta^{k=2} & \beta^{k=3} & \beta^{k=4} & \beta^{k=5} & \beta^{k=6} & \beta^{k=7} \end{bmatrix}^T$$

The batch of parameters, $\beta_j^{k=1}, j=1,\ldots,J_{nongroup}$, corresponds to the non-group explanatory variables, for which the β_j's are not linked in a hierarchy. Each β_j is assigned its own prior distribution – by default, the noninformative uniform prior distribution is used for all the non-group β_j's.

For each batch of parameters that corresponds to group-level predictors, we assign as the common prior a normal distribution with mean of zero and unknown variance (to be determined from data). The reasoning behind assigning a mean of zero is that the group-level predictors are meant to serve as a correction to the quantity $X_{nongroup}\beta_{nongroup}$; if more variability is needed in the model, a hyperprior distribution can be added for the mean, though often the amount of information available to determine this from the data may be limited.

As noted above, noninformative uniform prior distributions are assigned to the hierarchical variance parameters, $\sigma_{\beta k}$, which corresponds to the conjugate inverse-χ^2 prior with $\nu_{\beta k} = -1$ prior degrees of freedom and $\sigma_{0\beta k}^2 = 0$ prior scale parameter. This results in a vector of $\sigma_{\beta k}^2$, as there are K_β total batches of β's. The matrix-vector notation for $\Sigma_\beta = Diag\left(W_\beta \sigma_\beta^2\right)$ provides a convenient and general method for construction of the covariance matrix for models with any hierarchical groups desired. W_β is formed form K_β columns of 1's and 0's, in which the row elements of column k are set equal to 1 if the corresponding β_j is a member of the kth batch of parameters.

The remaining item which has not yet been discussed in the original model presented above is Σ_y. In this matter, there are a few natural options: one can assign to each batch a σ_{yk}, resulting in a unique variance for each batch (which intuitively seems reasonable given that one would

expect different conditions to yield different variances). However, this may not be desirable depending on how the model will be used, for example, if the intended use was to make predictions for unseen data, then the data-level variance becomes a problem. Consequently, the data can be re-batched solely for the data variances – in essence, one would map the batches to a new data variance, $\sigma_{yk,model}$ [23]. This procedure leads to consideration of data variance structures, which can be succinctly presented using the same notation as for the parameters, in which S denotes the index sets of relevant batches which constitute a mapping to a data variance. As reasonable examples, we consider re-batching of the data based on δ-level:

$$S_{\delta=single\,drop} = \{k : \delta(k) = single\,drop\}$$

$$S_{\delta=few\,drop} = \{k : \delta(k) = few\,drop\}$$

$$S_{\delta=many\,drop} = \{k : \delta(k) = many\,drop\}$$

Or perhaps even more simply, assuming that all observations have the same data variance:

$$S_{\forall\delta} = \{k : \delta(k) = single\,drop \wedge \delta(k) = few\,drop \wedge \delta(k) = many\,drop\}$$

$W_{y,model}$ and $\sigma^2_{y,model}$ are assembled analogously to W_{β} and σ^2_{β}, hence the convenience of representation in matrix-vector notation, $\Sigma_y = Diag(W_{y,model}\sigma^2_{y,model})$.

8.5.3   Solution Method

Markov chain Monte Carlo Simulation

At this point a discussion of the method for such models is appropriate; the data structure and detail on how y^k, X^k are derived from the images will be discussed afterward. For the model form given in the equations at the beginning of Section 8.5.1, both 2-level and multi-level hierarchical models can be represented in this canonical form through appropriate augmentation of the data matrix and prior parameter vector [6, 7, 24, 25, 26]. For such multi-level regression models with conjugate prior distributions, Markov chain Monte Carlo simulation using the Gibbs sampler is a possible method to obtain samples from the posterior distributions of the parameters. In constructing a Gibbs sampler for these problems, one considers the augmented equations:

$$y_* = \begin{bmatrix} y \\ \beta_0 \end{bmatrix}, \qquad X_* = \begin{bmatrix} X & 0 \\ 0 & I_J \end{bmatrix}, \qquad \Sigma_* = \begin{bmatrix} \Sigma_y & 0 \\ 0 & \Sigma_\beta \end{bmatrix}$$

$$y_* \sim N(X_*\beta, \Sigma_*)$$

which enable one to sample successively from the joint posterior distribution of β's, then sample from the posterior distributions of the data and hierarchical variances, with one cycle of sampling corresponding to one full iteration.

The conditional posterior distribution of β is $p(\beta|\sigma_y, \sigma_\beta, y) \sim N(\hat{\beta}, V_\beta)$, which involves the regression computation of y_* on X_* given the variance matrix Σ_*. Rather than perform the regression explicitly, that is,

$$\hat{\beta} = (X_*^T \Sigma_*^{-1} X_*)^{-1} X_*^T \Sigma_*^{-1} y_* \quad , \quad V_\beta = (X_*^T \Sigma_*^{-1} X_*)^{-1}$$

substantial computational savings (and improvements to numerical stability) can be achieved by taking advantage of the QR decomposition, $QR = \Sigma_*^{-1/2} X_*$, computing $\hat{\beta}$ by back-substitution on the system, $R\hat{\beta} = Q^T \Sigma_*^{-1/2} y_*$ and drawing the update of $\beta = \hat{\beta} + R^{-1}z$, in which z is a $J \times 1$ column vector of draws from the standard univariate normal distribution (in fact, one need not compute R^{-1} explicitly, but instead use back-substitution on the system $Rv = z$, and update $\beta = \hat{\beta} + v$). A further improvement to computational efficiency is enacted by noting that since Σ_* is diagonal, the Cholesky factor of its inverse is readily computed without actually inverting the

matrix as: $\Sigma_*^{-1/2} = \begin{bmatrix} \Sigma_y^{-1/2} & 0 \\ 0 & \Sigma_\beta^{-1/2} \end{bmatrix}$, in which $\Sigma_y^{-1/2} = Diag(W_y(1 \oslash \sigma_y))$, $\Sigma_\beta^{-1/2} =$

$Diag(W_\beta(1 \oslash \sigma_\beta))$, where 1 is an appropriately sized vector of 1's, and $\oslash$ denotes component-wise division.

The conditional posterior distribution of the variance components are independent, given β,

$$\sigma_{yk}^2 \sim \text{Inv } \chi^2 \left(v_{yk} + n_{yk}, \frac{\sigma_{0yk}^2 + SS_{yk}}{v_{yk} + n_{yk}} \right)$$

$$\sigma_{\beta k}^2 \sim \text{Inv } \chi^2 \left(v_{\beta k} + n_{\beta k}, \frac{\sigma_{0\beta k}^2 + SS_{\beta k}}{v_{\beta k} + n_{\beta k}} \right)$$

in which n_{yk} , $n_{\beta k}$ denote the number of data or parameter components in the batch, and SS_{yk} , $SS_{\beta k}$ denote the residual sum of squares for the batch.

This is the all-at-once Gibbs sampler for such hierarchical models, in which each update of β consists of a vector updating step, with the new vector β drawn from its posterior distribution in such a manner that its components are independent. This methodology can be advantageous

when the components of β are highly correlated in their joint posterior distribution, which is a common occurrence when working with hierarchical models with many parameters.

An inherent problem of the Gibbs samplers for hierarchical regression models is potentially slow convergence to stationarity can occur when hierarchical variance parameters are close to zero, which is a common occurrence in such models [6, 19]. This results from the effect of near-zero variance parameters on the β-update which causes the possible range of β draws to shrink, which then in turn causes the σ_β updates to result in smaller values as the result of lower variability in the just-updated β's. this can cause the algorithm spending inordinate amounts of time in one location, though the nature of the MCMC algorithm will allow it to escape eventually.

To preempt such problems, we utilize redundant parameterization, also known as parameter expansion, as described in Gelman *et al*., 2008. In this computational method, the models remain the same, but working parameters are introduced which index a transformation of variables which in the joint posterior of the model and working parameters, leads to greater variability, and therefore avoids the pitfalls of the regular Gibbs sampler [19]. Succinctly, a transformation of variables is introduced through addition of a vector of working parameters α, such that

$$A = \left[Diag(W_\beta \alpha \right]^{-1}$$

yielding a transformed canonical form:

$$y | \xi, \sigma_y, \alpha \sim N\left(XA^{-1}\xi, \Sigma_y \right)$$

$$\xi | \xi_0, \Sigma_\xi \sim N\left(\xi_0, \Sigma_\xi \right)$$

$$\alpha | \alpha_0 \sim N(\alpha_0, \Sigma_\alpha)$$

in which $\xi = A\beta$, $\xi_0 = A\beta_0$, and $\Sigma_\xi = A\Sigma_\beta A$. If α is given an informative prior distribution, then α is identifiable in the joint posterior distribution and the result is a multiplicative parameter model [27]. Alternately, if α is given a noninformative prior distribution, then it is not identifiable in the posterior, and upon marginalization of alpha out of the joint posterior, the original model is recovered, in which case α meets the definition of a working parameter [19, 26]. To recover the original model parameters at each step, the transformed parameters are multiplied by the appropriate α^k from the same iteration of the Gibbs sampler:

$$\beta_j^k = \alpha^k \xi_j^k$$

$$\sigma_{\beta k} = |\alpha^k| \sigma_{\xi k}$$

When alpha is a working parameter, only the original model parameters from each iteration are stored; convergence is determined from the original model parameters, rather than the transformed model parameters.

Convergence Monitoring

For all of the models considered, the parameter-expanded Gibbs sampler was used to obtain samples from the posterior distribution of $\beta, \sigma_y, \sigma_\beta$. Multiple chains were simulated in parallel, each of which was provided with different over-dispersed starting points for the data and hierarchical variance parameters. The starting points are generated by performing one step of a "one-batch-at-a-time" Gibbs sampler, then drawing σ_y, σ_β from inverse-χ^2 distributions with low degrees of freedom (2) and scale parameter equal to the sum of squares respective to the kth batch of y or β. Each simulation consisted of at least 10 chains, each with length of at least 1000; the first half of each chain is discarded as the warm-up period, and all subsequent computations are made using the latter halves.

Using the multiple chains, convergence is assessed by computing the potential scale reduction achieved by the simulation, $\hat{R}$, and the effective number of independent simulation draws $\hat{n}_{eff}$ [6, 28, 29]. The simulation was considered to have converged when the data and hierarchical variance parameters reached $\hat{R} < 1.1$, computed using σ_y, σ_β, rather than $\sigma_y^2, \sigma_\beta^2$, and the effective number of simulation draws was sufficient to yield low Monte Carlo error. For the results presented in Section 6, 20 chains of length 2000 were simulated, resulting in $\hat{R} < 1.05$ and values of $\hat{n}_{eff}$ greater than 250; with the first halves discarded, this yielded a total of 20000 posterior samples.

8.6   Model Checking; Probabilistic Prediction

8.6.1   Posterior Predictive Distribution Under the Models

Using the training data y, X, one obtains the samples from the posterior distribution of $\beta, \sigma_y, \sigma_\beta$. The key use of the model is in prediction for future/new data, that is, probabilistic forecasting for future observables $\tilde{y}$ given a new matrix of explanatory variables, $\tilde{X}$, as would become available during the online monitoring of production. Thus, we seek to obtain the posterior

predictive distribution of future observations, conditional on the posterior distribution of the model parameters (which is itself conditional on the past data and priors). Therefore, we use

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta) p(\theta|y) d\theta$$

In which θ represents all the parameters of the model, i.e. $\beta, \sigma_y, \sigma_\beta$. More specifically in the case of a regression model, one could write:

$$p(\tilde{y}|\tilde{X}, y, X) = \int p(\tilde{y}|\tilde{X}, \theta) p(\theta|y, X) d\theta$$

The posterior predictive distribution can also be used to predict the other possible realizations of y under the model given the original matrix of explanatory variables. When used for this purpose, the alternate possible realizations, or replicates, are denoted as y^{rep}, and their posterior predictive distribution is

$$p(y^{rep}|y, X) = \int p(y^{rep}|X, \theta) p(\theta|y, X) d\theta$$

8.6.2 Model Checking, Comparison: Test Variables for Training and Replicated Data; Information Criteria

To assess the fit of the model to the data, test variables which are considered to be representative of model performance are constructed, and their posterior (predictive) distribution is computed under the training data and the replicated data. The replicated data, y^{rep}, is generated from the posterior predictive distribution of the observables, using the posterior samples of $\beta, \sigma_y, \sigma_\beta$. Using the realized test quantity, $T(y, \theta)$, and the predictive test quantity, $T(y^{rep}, \theta)$, we compute the posterior predictive p-value as $\Pr(T(y^{rep}, \theta) \geq T(y, \theta)|y)$; displaying the results graphically can also provide insight [30, 31]. The posterior predictive p-value represents the probability that the replicated data could be more extreme than the observed data, with test variables constructed to assess particular aspects of the interest relevant to use of the model for predictive purposes [30, 32].

The p-values provide insight into potential failures of the model, which can help to guide practical use of the model; however, it should not be interpreted as a probability of the model being true given the data. Instead, extreme p-values near 0 or 1 imply that the model is inconsistent with the data respective to that test quantity. For a model that is consistent with the data, one might obtain a p-value near 0.5 for a test quantity of interest, which is interpretable as a 50% chance that

replicated data will exceed the observed data. The range of p-values which may be deemed acceptable are determined by the practical significance of the test quantity in the context of the intended use of the model [30, 32, 33].

Test Variables

We consider two test variables based on the mean squared error, with one computed from the within-batch residuals, and the other computed for multiple batches based on the batch-wide mean residual. The residual for each dose is always:

$$r_i^k = y_i^k - X_i^k \beta, \quad i = 1, \ldots, n_{yk}, \quad k = 1, \ldots, K_y$$

The within-batch mean squared error then computed as:

$$MSE^k = \frac{1}{n_{yk}} \sum_{i=1}^{n_{yk}} \left(r_i^k \right)^2$$

And thus provides a measure of the model fit to the data for each batch, though, for 41 batches, this can be overwhelming to consider graphically. However, it is useful in identifying the conditions at which the model may poorly fit the data, thereby providing a path to potential improvements.

The batch-wide mean residual is computed as:

$$\bar{r}^k = \frac{1}{n_{yk}} \sum_{i=1}^{n_{yk}} r_i^k, \quad k = 1, \ldots, K_y$$

with the cross-batch mean squared error:

$$MSE = \frac{1}{\left| \{ Desired\ Index\ Set \} \right|} \sum_{k \in \{ Desired\ Index\ Set \}} \left(\bar{r}^k \right)^2$$

In which $\{ Desired\ Index\ Set \}$ denotes a generic index set of interest in assessing model performance, perhaps one such as $S_{\delta = single\ drop} = \{ k : \delta(k) = single\ drop \}$.

We compute both of these using the $s = 1, \ldots S$ draws from the posterior distribution of the model parameters given the training data. For the replicated data generated from the posterior predictive distribution, $y_i^{k^{rep}}$, the residual is computed as:

$$r_i^{k^{rep}} = y_i^{k^{rep}} - X_i^k \beta$$

and the MSE^k, MSE are computed as described above.

191

Information Criteria for Model Comparison

To further evaluate the accuracy of each model in the context of probabilistic forecasting, one can consider scoring rules based on the log predictive density [34, 35]. To this end, we consider information criteria in order to measure model performance and enable comparison of models of different size on the same scale – accounting for the ability of a larger model to fit data better, if only by chance. As out-of-sample data is not available, the selected methods estimate the expected predictive accuracy from the existing data and make adjustments for the bias introduced by evaluating the log predictive density on the model was fit to.

The Watanabe-Akaike information criterion (WAIC) was selected for this purpose as it is a fully Bayesian measure which accounts for posterior uncertainty in the model parameters [36] and it is asymptotically equal to Bayesian cross-validation [37]. The advantage is that by using the entire posterior distribution, WAIC would resemble how the posterior predictive distribution would be used in forecasting for new data; this is in contrast to the Akaike information criterion and the deviance information criteria (DIC), which use point estimates of the parameters, i.e., the maximum likelihood estimate and the posterior mean, respectively [38, 39].
The WAIC is defined using the log pointwise predictive density,

$$\text{lppd} = \sum_{i=1}^{n} \log \int p\left(y_i \,|\, \theta\right) p\left(\theta \,|\, y\right) d\theta$$

and a correction for the effective number of parameters,

$$p_{\text{WAIC 1}} = 2\sum_{i=1}^{n} \left(\log\left(\text{E}_{p(\theta|y)} \, p\left(y_i \,|\, \theta\right) \right) - \text{E}_{p(\theta|y)} \left(\log p\left(y_i \,|\, \theta\right) \right) \right)$$

$$p_{\text{WAIC 2}} = \sum_{i=1}^{n} \text{var}_{p(\theta|y)} \left(\log p\left(y_i \,|\, \theta\right) \right)$$

In which $\text{E}_{p(\theta|y)}$ and $\text{var}_{p(\theta|y)}$ denote the respective quantities taken over the posterior distribution of θ. Using the notation and re-scaling which makes it comparable to other information criteria [6], this yields:

$$\text{WAIC} = -2\text{lppd} + 2p_{\text{WAIC 2}}$$

The intermediate quantities $p_{\text{WAIC 1}}$, $p_{\text{WAIC 2}}$, provide useful insight into the model, as they are approximate estimates of the number of unconstrained parameters in the model, in which each parameter contributes a value from 0 (fully constrained) to 1 (unconstrained), based on the data and informativeness of the prior distribution [39]. Such interpretations are especially useful in

assessing hierarchical models, where the effective number of parameters can be used to understand the extent of pooling performed by the model.

8.6.3 Prediction for Out-of-Sample Data from In-Sample Conditions

Of the total data available, one half is used as the training data, and the remaining half is reserved for assessing the model performance in probabilistic forecasting. In this context, the new explanatory variable matrix respective to each batch, $\tilde{X}^k : \tilde{X}_i^k, i = 1,...,\tilde{n}_{yk}$, is used with the posterior predictive distribution to forecast $\tilde{y}^k : \tilde{y}_i^k, i = 1,...,\tilde{n}_{yk}$. Additionally, the actual ("real") measured value, $y_i^{k^{new}}$, respective to each $\tilde{X}_i^k$ are available to assess the predictive error with respect to the forecasted distribution.

Point Estimate of Predictive Error

In assessing the model, a number of items based on the posterior predictive distribution of $\tilde{y}$'s are considered. Most directly, the predictive error for each dose i in each batch k is computed as: $e_i^k = y_i^{k^{new}} - \tilde{y}_i^k$, averaging over the posterior distribution of the parameters $\beta, \sigma_y, \sigma_\beta$, as $s = 1,...S$ simulated values of $\tilde{y}_i^k$ from its posterior predictive distribution. Thus, we compute $e_i^{k^{(s)}} = y_i^{k^{new}} - \tilde{y}_i^{k^{(s)}}$, and average over the posterior predictive samples to obtain the mean predictive error,

$$e_i^k = \frac{1}{S}\sum_{s=1}^{S} e_i^{k^{(s)}} = y_i^{k^{new}} - \frac{1}{S}\sum_{s=1}^{S} \tilde{y}_i^{k^{(s)}}$$

From these, the batch-wide mean predictive error can be computed as:

$$\bar{e}^k = \frac{1}{\tilde{n}_{yk}}\sum_{i=1}^{\tilde{n}_{yk}} e_i^k$$

Such measures of predictive error can be informative in the context of model validation, in particular for identifying conditions which may have dubious predictions for out-of-sample data.

However, one would prefer to use the model for probabilistic forecasting, rather than for the point estimates described above. Thus, by revisiting the context in which the model will actually be employed, several useful items appear for which probabilistic forecasting is relevant.

Probabilistic Prediction

As noted in the Introduction, probabilistic prediction using this model is focused on dosage strength. In production of a batch of doses, nominal dosage strength is set at a design value, that is, a target specification for active drug in each dosage unit, denoted here as Θ^k, respective to each batch k. To guarantee efficacy of the product in the patient, the actual amount of active drug in each unit must be within some tolerance limits with respect to the nominal amount, $\left[\Theta^k (1-z), \Theta^k (1+z)\right]$. By FDA regulation, the resultant batch of doses must yield at maximum a 5% RSD, calculated with respect to the nominal amount, as defined by the USP 905 for content uniformity [40].

Such regulations are developed for large batches, from which a small sample is drawn and destructively tested. Thus, in the context of a monitoring system which enables individual prediction for each dose, the 5% RSD regulation provides a suggestion for an upper bound of tolerance limits – that is, we may wish to accept/reject doses at the point of production based on the probability that the dose is within z% of the nominal strength. To compute the probability for each dose, one uses the posterior predictive distribution of $\tilde{y}_i^k$, as defined below:

$$\Pr\left(\Theta^k (1-z) \leq \tilde{y}_i^k \leq \Theta^k (1+z)\right) = \int_{\Theta^k (1-z)}^{\Theta^k (1+z)} p\left(\tilde{y}_i^k \mid y\right) d\tilde{y}_i^k$$

In which Θ^k is the batch target specification and z is the tolerance limit, e.g. 1%, 2.5%, 5% (as percentages).

To enact a quality control setup, a probability threshold could be set for acceptance/rejection of each dose given a tolerance limit. Or perhaps something more creative such as a probability threshold for each tolerance limit, e.g. Pr = 0.5 for z = 1%, Pr = 0.75 for z = 2.5%, Pr = 0.95 for z = 5%. This is a decision that depends on the manufacturer and specification considerations (e.g. patient safety, economics, etc.) for a given product.

Perhaps a more traditional calculation, which also happens to be less demanding of individual doses, is to compute the probability of the entire batch being within specification,

$$\Pr\left(\Theta^k (1-z) \leq \tilde{y}_{i=1}^k \leq \Theta^k (1+z) \cap \cdots \cap \Theta^k (1-z) \leq \tilde{y}_{i=\tilde{n}_{yk}}^k \leq \Theta^k (1+z)\right) = \int_{\Theta^k (1-z)}^{\Theta^k (1+z)} p\left(\tilde{y}^k \mid y\right) d\tilde{y}^k$$

Though it may be more useful to compute the number of doses in the batch which are within specification; normalized to the total number of doses in the batch, $\tilde{n}_{yk}$, that is, the fraction of the

batch within specification, denoted $F_{\tilde{y}^k}$. This allows one to consider not only the probability, but also the size of the credible region through the percentiles of $F_{\tilde{y}^k}$, which gives a means of visualizing the uncertainty in content uniformity. We compute the fraction of the batch within specification, starting from the posterior predictive distribution of $\tilde{y}^k$.

An indicator function is applied to each dose, transforming $\tilde{y}_i^k$ to $I_{\tilde{y}_i^k}$:

$$\text{Indicator Function} = \begin{cases} 1 & \Theta^k\left(1-z\right) \leq \tilde{y}_i^k \leq \Theta^k\left(1+z\right) \\ 0 & \text{otherwise} \end{cases}$$

Then, the fraction of doses in batch k within specification is:

$$F_{\tilde{y}^k} = \frac{1}{\tilde{n}_{yk}} \sum_{i=1}^{\tilde{n}_{yk}} I_{\tilde{y}_i^k}$$

From the posterior predictive simulations, the distribution of $F_{\tilde{y}^k}$ quantity is directly available, hence, it is straightforward to answer questions such as: what is the probability that all doses in the batch are within specification? – this is simply $\Pr\left(F_{\tilde{y}^k} = 1\right)$.

As a means of visualizing uncertainty in content uniformity for all the batches, a line plot can be constructed for a given z from the respective $F_{\tilde{y}^k}$, with the line spanning from the 2.5th to the 97.5th percentile. For comparison with the forecasted result, the fraction of the batch which would pass based on the actual measured values of $y^{k^{new}}$ is marked on each row of the line plot.

It is difficult to provide a rigorous interpretation of the meaning of the fraction based on $y^{k^{new}}$ in relation to the predicted distribution $F_{\tilde{y}^k}$. But, if we take the $y^{k^{new}}$ to be the "true" values, then qualitatively, one would expect the credible region of $F_{\tilde{y}^k}$ to include the fraction computed based on $y^{k^{new}}$ if the model is reasonably accurate. Thus, this provides another means of assessing the model for its intended use. If there is substantial mismatch between the forecast and the actual measurement for many batches, then the model may need rethinking; if the mismatch is present in only a few batches, then this may identify weak points in the model which might be improved through additional explanatory variables.

8.6.4 Prediction for Out-of-Sample Data from Out-of-Sample Conditions

In addition to the probabilistic forecasting described above, hierarchical models enable prediction for out-of-sample data from new groups, that is, for new $\tilde{X}$'s which are obtained at conditions which differ from any of those included in the training data. In predicting observables $\tilde{y}$ for new data from new groups, one notes that the posterior distribution of the β's in the hierarchical groupings is with respect to the conditions in the training data, thus, at out-of-sample conditions, these parameters would not be appropriate (e.g. how can one justify using parameters for sfAPAP for a new particle type?). We posit that the parameters at the new conditions are exchangeable with those at the old, and therefore come from the same parent distribution (which has been estimated from the training data). Hence, we use the posterior predictive distribution of the hierarchical regression parameters,

$$p\left(\tilde{\beta}|\beta,\sigma_{\beta},y,X\right) = \int p\left(\tilde{\beta}|\beta_0,\sigma_{\beta}\right)p\left(\beta,\sigma_{\beta}|y,X\right)d\beta d\sigma_{\beta}$$

to predict the regression parameters for data from new groups, and then obtain the desired result, the posterior predictive distribution of the observables,

$$p\left(\tilde{y}|\tilde{X},y,X\right) = \int p\left(\tilde{y}|\tilde{X},\tilde{\beta},\sigma_y\right)p\left(\tilde{\beta}|\beta,\sigma_{\beta},y,X\right)d\tilde{\beta}$$

For the non-group regression parameters, it is not necessary to generate from their posterior predictive distribution – by virtue of not being in a hierarchical group, it is implicitly assumed that they are equally applicable to all data. However, for the group-level predictors, the above procedure must be performed when encountering data from new groups which were not included in the training data. This is one of the strengths of hierarchical models, that the predictions for new groups have more components of variability, and hence, should be more reliable for new particle types, fluid types, combinations thereof, etc.

To test the predictive performance for out-of-sample data from out-of-sample conditions, we use the data on the 5 batches (Hormones (ethinylestradiol + norgestrel), Ibrutinib, Loratidine, Tacrolimus, Warfarin) of doses provided by PinPoint Pharma (Giridhar 2020).

8.7 Results

For both the ModelDrop and ModelDose conceptual approaches, many models were evaluated; this is inevitable given the many possible combinations of $X_{nongroup}$, the hierarchical

groupings, and the data variance choices. Throughout the evaluations, model performance was first assessed using the information criteria to discard models which fit the data poorly, as indicated by larger values of WAIC. As might be expected, the WAIC's for the remaining models were close; thus, while a useful means of eliminating poor models, the choice of a preferred model must be supplemented by 1) comparison of probabilistic predictions for out-of-sample data from in-sample conditions and 2) considerations for future uses of the model. As the latter has a substantial influence in selection, the relevant aspects of the model are discussed separately in Section 8.8.

8.7.1 ModelDose

For reasons discussed in Section 8.8, we restrict ourselves to the 3 data variances, and for the hierarchical groupings: $(\delta \times \Phi \times p \times f)$ with 4 separate linkages, $(\delta \times p \times f)$ with 1 linkage, and $(\Phi \times p \times f)$ with 1 linkage. This was selected as the preferred model from the many essentially equal WAIC scores which were obtained by trying different hierarchical groupings, as it is sufficiently general to be applicable to most future data, in particular out-of-sample data from out-of-sample groups. The non-group explanatory variables were as described in Section 3.5, with the Euclidean distance measures excluded. Combined with the group-level predictors, this yielded a total of 82 regression parameters, 6 hierarchical variance parameters, and 3 data variance parameters. However, as computed as part of the WAIC, the effective number of parameters (Table 8.1, column 1) in the model is ≈ 60 based on $p_{\mathrm{WAIC\,2}}$, which indicates that some pooling of the data occurs.

Posterior Predictive Model Check

Inspection of the test variable scatterplots and posterior predictive p-values (Figure 8.7) for the batch-wide RMSE's based on the $S_{\delta=single\,drop}$, $S_{\delta=few\,drop}$, $S_{\delta=many\,drop}$ index sets confirms that the model is capable of reproducing the data to which it was fit. However, a more detailed examination of the batch-specific RMSE[k]'s (Figure 8.8) reveals that for some batches the model struggles to replicate the data, indicating potentially poor predictive accuracy at such conditions. This is due primarily to the data variance structure used in the model – in actuality, the within-batch sample variance appears to be lower than the data variance, particularly for the single-drop batches. This is a consequence of the re-batching of the data into just 3 σ_y's, the motivation for which was a

more general model; it is straightforward to allow a unique σ_{yk} for each batch to eliminate this obstacle to improved predictions. On the whole, however, this is a sacrifice we are willing to make given that we are modeling 41 unique batches of data.
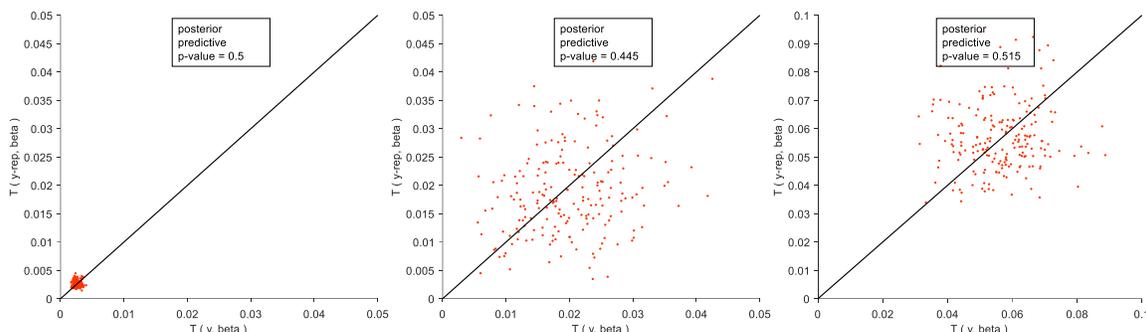
Figure 8.7. Square root of batch-wide MSE, evaluated on batch index sets: $S_{\delta=single\,drop}$, $S_{\delta=few\,drop}$, and $S_{\delta=many\,drop}$, from left to right, respectively.
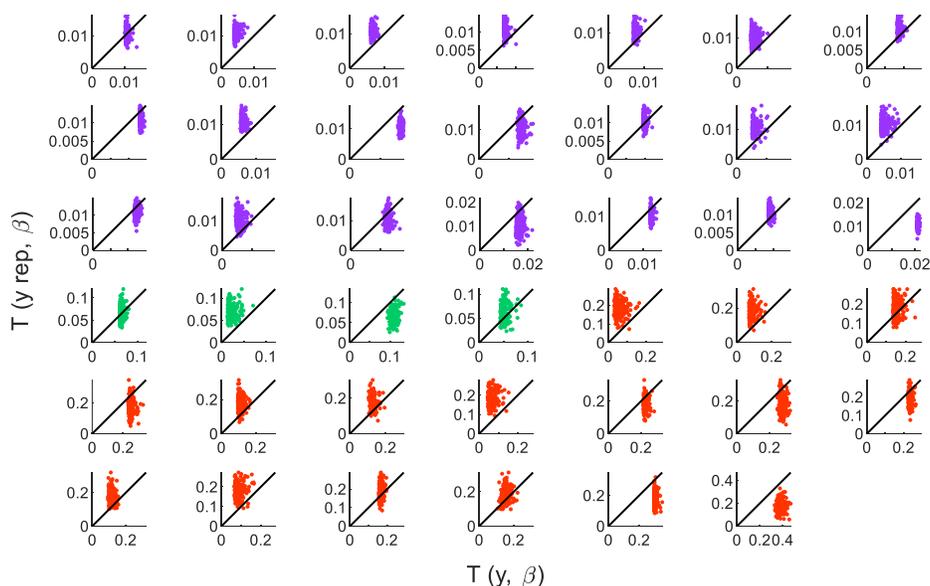
Figure 8.8. Square root of batch-specific MSE^k, evaluated for all 41 batches; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

Prediction for Out-of-Sample Data from In-Sample Conditions

The predictive performance of the ModelDose approach for point estimation and probabilistic forecasting, as described in Section 8.6.3, are presented in Figure 8.9 and Figure 8.10, respectively. As shown in Figure 8.9, in all but 2 batches out of 41 the normalized mean predictive

error is below ±1%, which indicates that the model performs well for most of the out-of-sample data. The notable exception is batch number 18 (condition details: MA, NEO895, particle fraction 0.4, single drop) at 4%, though upon consideration of the presence of multiple missing volume approximations (example in Figure 8.5) in the constituent images, this result is not surprising.

The probabilistic predictions for the fraction of the batch within specification, presented in Figure 8.10, have good accuracy at the $z = 2.5\%$ and 5% of target values, with the 50% interval of the posterior predictive density encompassing the "true" fraction in most cases, and often with the predicted posterior median matching the "true" fraction. This occurs less frequently at $z = 1\%$, but in most of the batches, the 95% interval does encompass the "true" fraction; the breakdown in accuracy would be expected given that the model has non-zero predictive error and variance is likely to shift some simulated $\tilde{y}_i^k$'s outside the $z = 1\%$ specification.
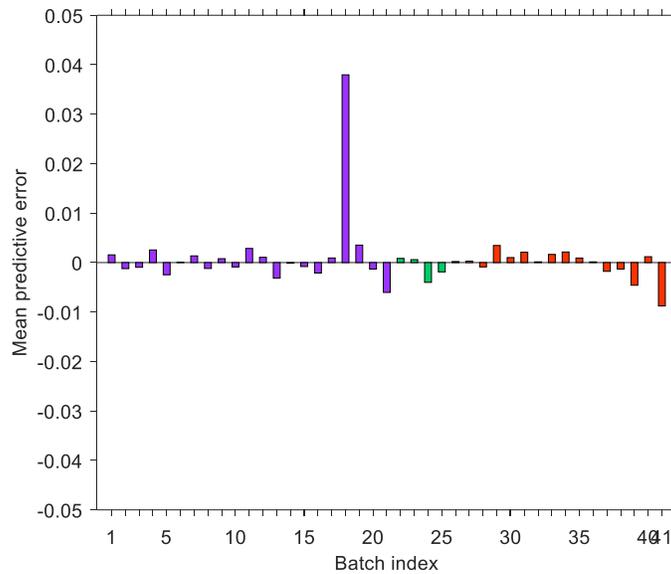
Figure 8.9. Batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k / \bar{y}^{k^{new}}$, for all 41 batches of new data from existing groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

Figure 8.10. $F_{\tilde{y}^k}$, for all 41 batches of new data, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

Prediction for Out-of-Sample Data from Out-of-Sample Conditions

Figures 8.11 and 8.12 present the same point error estimate and probabilistic prediction items, but with the difference that the hierarchical regression parameters are first generated from their posterior predictive distribution, and then the $\tilde{y}_i^k$'s are generated from their posterior predictive distribution as described in Section 8.6.4. Though the mean predictive error is only slightly higher than for out-of-sample data from in-sample conditions, the posterior intervals are considerably wider, as would be expected when generating from the posterior predictive distribution of the hierarchical regression parameters. The posterior predictive 50% intervals seem to be reasonably accurate at $z = 1$, 2.5, and 5% only for the Hormones batch, while it is arguable as to whether the Ibrutinib and Loratidine batches can be considered accurate at $z = 5\%$. In contrast, for Tacrolimus and Warfarin batches the posterior predictive 50% intervals contain the "true" fraction at $z = 5\%$ but the medians suggest that the predictions are off; a potential explanation may be that these batches were produced from suspensions with very low particle loading, hence the predicted regression parameters from the $S_{\delta=many\,drop\,\times\,\phi=high\,\times\,p\,\times\,f}$ hierarchical group are not appropriate.

200
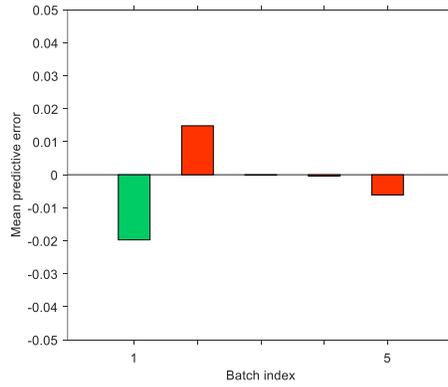
Figure 8.11. Batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k / \bar{y}^{k^{new}}$, for the 5 batches of new data from new groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses. Batch Indices 1, 2, 3, 4 and 5 correspond to Hormones, Ibrutinib, Loratidine, Tacrolimus, and Warfarin, respectively.
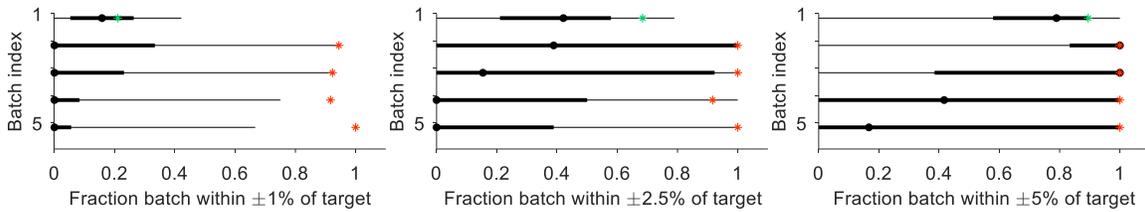
Figure 8.12. $F_{\tilde{y}^k}$, for the 5 batches of new data from new groups, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

8.7.2 ModelDrop

The same data variance structure and hierarchical groupings which yielded the preferred model in the ModelDrop approach are applicable in the ModelDrop approach. The difference between the two approaches is the possibility of including extra non-group explanatory variables, such as: the Euclidean distance measures, (separate) categorical indicators for border crossings by the primary and satellite drops, and total number of satellites. In total, up to 6 extra explanatory variables can be added to the model; however, the result is a marginal improvement in WAIC, as can be seen from comparison of columns 2, 3, and 4 of Table 8.1. Consequently, we prefer the model which uses only the original 8 non-group variables, for reasons discussed below.

201

For the model which includes the Euclidean distance measures, graphical comparison of the predictive performance for out-of-sample data from in-sample conditions and out-of-sample data from out-of-sample conditions reveals few differences from the original model, as would be expected based on the information criteria. Examination of the posterior distributions of the non-group parameters for these two models reveals the only major change to be some shifting of information away from the missing primary drop volume at the top border, with the destination seeming to be the Euclidean distance measure for the primary drop – the other non-group predictors do not change substantially. Essentially, the parameter is the contribution of not only the actual distance rescaled to the maximum, but also the drop shape. The shape of the drop and location in the image may bear some relation to the satellite drops, or trends overall (e.g. unobserved fluid outside image), but it is difficult to decipher without additional data. In summary, this parameter cannot be interpreted as easily as the β 's for the volume-based explanatory variables, for which the intuition is clear.

If all 6 parameters are added, the posterior distribution of the non-group parameters changes in various ways, for which an interpretation is not easily obtainable – it seems that the parameters related to the primary drop now take precedence, with the posterior modes of the satellite parameters being negative, which is contrary to physical intuition. Consequently, we disregard this model as being potentially unrealistic for use on out-of-sample conditions.

As the models have essentially the same predictive performance, we consider additional arguments to make the case for selecting the model without the Euclidean distance measures. The advantage of the 8-parameter model is that has easily interpretable non-group explanatory variables, which seems to provide greater assurance that future predictions would be reasonable. It is possible that in this dataset, the Euclidean distance measures correspond to the missing primary volume (as seems to be the case from the comparison of the posterior distributions of the parameters), but for future data this may not be case. Thus, one is less confident of this model compared to a model which is based solely on the "physical" volume explanatory variables.

Table 8.1. Information Criteria for the models. Note: comparison of scoring functions is limited to the same modeling approach, i.e. when the data is identical. More negative values of DIC and WAIC imply higher predictive accuracy.

	MODELDOSE	MODELDROP, 8 PARAMS	MODELDROP, INCL. 2 EXTRA PARAMS	MODELDROP, INCL. 6 EXTRA PARAMS
-2lpd	-2817.1	-4158.7	-4181.8	-4213.3
p_{DIC}	49.64	49.68	51.73	55.76
DIC	-2718.4	-4059.3	-4078.4	-4101.8
-2lppd	-2819.9	-4161.1	-4184.9	-4214.4
$p_{\text{WAIC 1}}$	51.87	52.10	54.82	56.80
$p_{\text{WAIC 2}}$	59.96	60.23	63.58	66.30
WAIC	-2700.0	-4040.6	-4057.8	-4081.7

Posterior Predictive Model Check

The test variable scatterplots and posterior predictive p-values (Figure 8.13) for the batch-wide RMSE's confirms that the model is capable of reproducing the data to which it was fit, though it does appear that $S_{\delta=few\,drop}$ actually exhibits higher variance. Similarly to Figure 8.8, the batch-specific RMSE^k's in Figure 8.14 suggest that one should be wary when interpreting results for batches for which the posterior predictive p-value is extreme (i.e. majority of scatter points above or below 45° line). Interestingly, the comparatively lower RMSE^k's in Figure 8.14 for the many-drop doses provide further justification for pursuing a model with at least 3 σ_y's, which, given the manner in which the explanatory variable vector is constructed, does seem to be sensible.
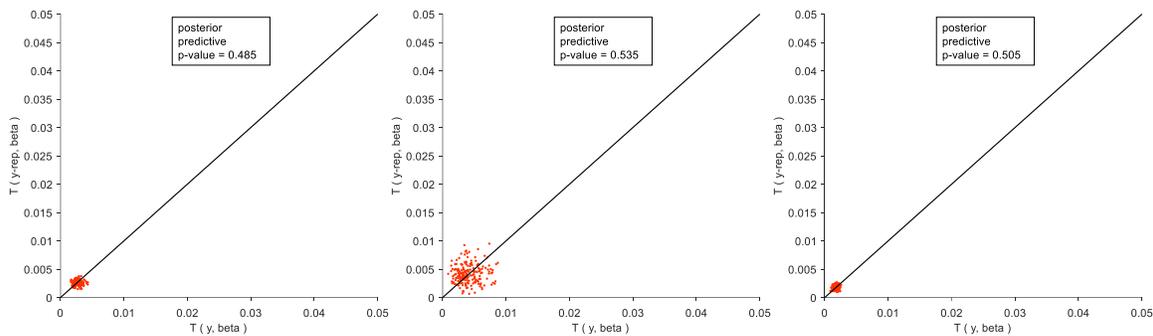
Figure 8.13. Square root of batch-wide MSE, evaluated on batch index sets: $S_{\delta=single\,drop}$, $S_{\delta=few\,drop}$, and $S_{\delta=many\,drop}$, from left to right, respectively.
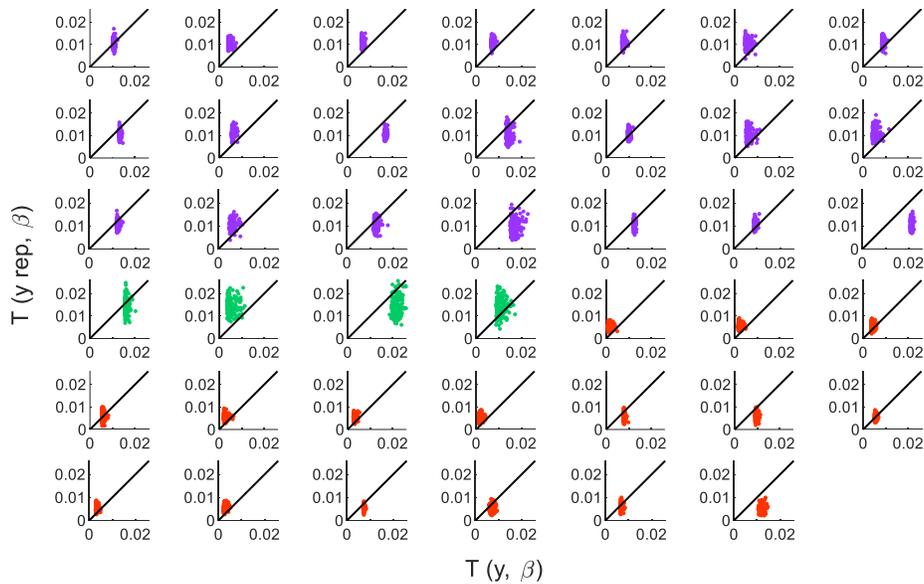
Figure 8.14. Square root of batch-specific MSE^k, evaluated for all 41 batches; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

Prediction for Out-of-Sample Data from In-Sample Conditions

Despite the differences in the ability of the models to replicate the data, the result for the point estimate of predictive error (Figure 8.15) is largely the same between the ModelDose and ModelDrop approaches, with batch number 18 yielding the only predictive error greater than ±1%. Probabilistic predictions for the fraction of the batch within specification, presented in Figure 8.16, exhibit good accuracy at the $z = 2.5\%$ and 5% of target values, 50% interval and median of the posterior predictive density situated about the "true" fraction in most cases; in comparison analogous prediction from ModelDose approach, the only noticeable difference occurs for batch number 39 at $z = 2.5\%$, which the ModelDrop approach appears to better predict. Differences between the ModelDrop and ModelDose results are small but noticeable at $z = 1\%$ for batch numbers 25-32, in which the ModelDrop posterior predictive density is more diffuse and less accurately situated about the "true" fraction.
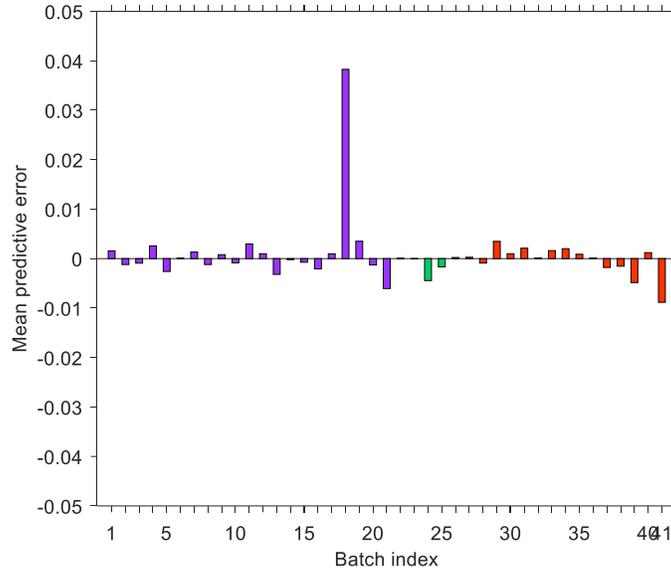
Figure 8.15. Batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k/\bar{y}^{k^{new}}$, for all 41 batches of new data from existing groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.
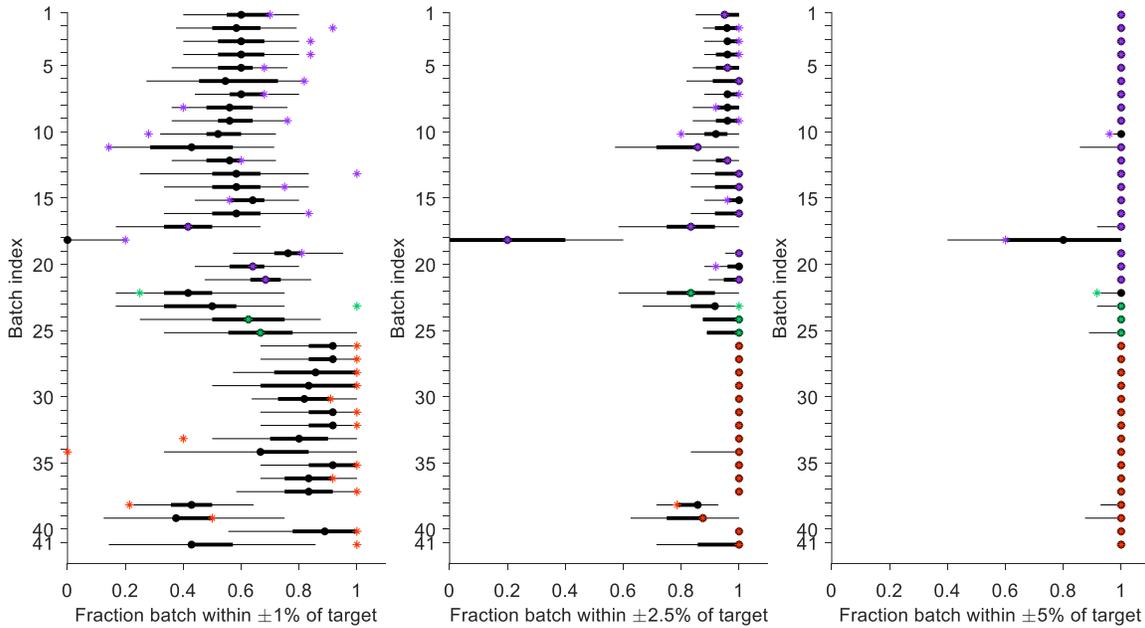
Figure 8.16. $F_{\bar{y}^k}$, for all 41 batches of new data, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\bar{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

Prediction for Out-of-Sample Data from Out-of-Sample Groups
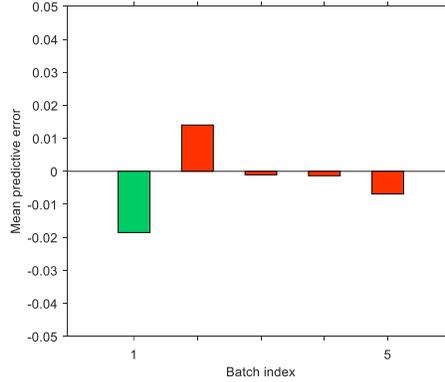
Figure 8.17. Batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k/\overline{y^{k^{new}}}$, for the 5 batches of new data from new groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses. Batch Indices 1, 2, 3, 4 and 5 correspond to Hormones, Ibrutinib, Loratidine, Tacrolimus, and Warfarin, respectively.
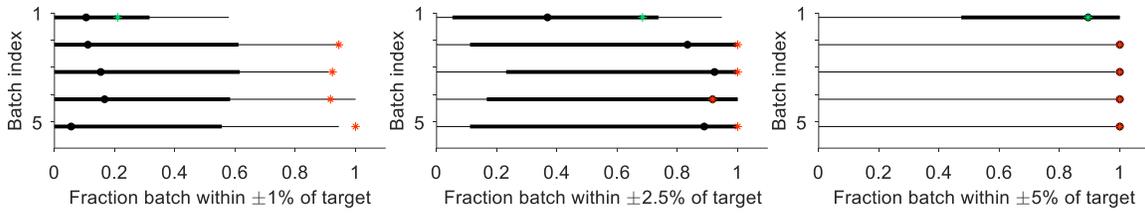
Figure 8.18. $F_{\hat{y}^k}$, for the 5 batches of new data from new groups, computed at $z = 1, 2.5,$ and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\hat{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

8.8   Discussion

8.8.1   Data Variances

In the ModelDose approach, it is apparent that the vector y is heteroscedastic, with each batch potentially exhibiting its own unique variability, hence it is necessary to allow data variances for each approximate scale. Given the system and manner in which doses are produced, the number of drops per dose provides the best approximation for mapping of batches y^k to data variances, hence, we use $s_{\delta=single\ drop}$, $s_{\delta=few\ drop}$, $s_{\delta=many\ drop}$, as described in Section 8.5.2, which yields 3 data variances. Compared to a model with a single data variance, use of the above immediately realizes

a large gain in predictive accuracy, regardless of the hierarchical groupings, coincident with a decrease in WAIC, indicating a better fit to the data; correspondingly, the accuracy of the probabilistic predictions is markedly increased. Predictive accuracy can be further improved by allowing a separate variance for each batch, but to do so reduces the generality of the model since the data variances were not modeled hierarchically; this is unfavorable for practical reasons as the model will be applied to not only future data from in-sample conditions, but also to future data from out-of-sample conditions.

In the ModelDrop approach, the y^k's are all of the same approximate scale, thus there is not as clear a motivation to use more than 1 data variance, in contrast to the ModelDose case where heteroscedasticity is obvious. Comparison of the WAIC for two models which are identical in all other respects than the data variance structure reveals a small improvement to WAIC for the 3 data variance version (compare Table 8.1, column 2 with Table K.1). However, it is most helpful to examine the probabilistic prediction results: the 3 data variance model exhibits visibly better performance in this respect for the credible regions of $F_{\tilde{y}^k}$ at $z = 1\%$ in batches of many-drop doses (compare Figure 8.16 with Figure K.2). Consequently, the 3 data variance structure seems to be preferable in the ModelDrop approach as well. An explanation for this might be based on differences in production conditions (and the treatment of the data) which are reflected in the variance of the many-drop doses, thus, one may apply the three data variance structure as described above.

8.8.2   Hierarchical Groupings

At the lowest level, it is necessary to model the interaction of dose level, rheological regime, particle type, and fluid type, hence, the use of $(\delta \times \Phi \times p \times f)$ interaction terms via hierarchical groupings. Each of the 41 batches is a unique realization of $(\delta \times \Phi \times p \times f)$, as each particle fraction is different, thus, we allow the possibility of a unique interaction – a unique correction parameter – for each batch. Exchangeability of the parameters is logical within the context of a given δ-level and Φ-regime, across $p \times f$ combinations, but not across all $(\delta \times \Phi \times p \times f)$ parameters. Thus, this yields $5 = (6 - 1)$ index sets:

$$S_{\delta \times \phi \times p \times f} = \left\{ k : \delta(k) = \textit{dose level} \wedge \phi(k) = \textit{rheological regime} \right\}$$

as the $(\delta = many\ drop \times \phi = low)$ combination was not in the training data set; the index sets for $(\delta = few\ drop \times \phi = low)$ and $(\delta = few\ drop \times \phi = high)$ were combined in order to obtain a sufficient number of parameters to create a hierarchy. Thus, the final index sets were: $S_{\delta = single\ drop \times \phi = low \times p \times f}$, $S_{\delta = single\ drop \times \phi = high \times p \times f}$, $S_{\delta = few\ drop \times \phi = low,high \times p \times f}$, and $S_{\delta = many\ drop \times \phi = high \times p \times f}$. In contrast to all the other hierarchical groupings used, each member of the index set is assigned a unique parameter, for which a corresponding column is allocated. The batch of parameters corresponding to each index set is then linked in a hierarchical group with common mean (of zero) and unknown variance, thus, this contributes 4 separate hierarchical groups.

The tradeoff between separated or single hierarchies is in the specificity of the model with respect to a given interaction, or region within an interaction; separated linkages within a given interaction allow one to model specific regions of an interaction, treating only those parameters as exchangeable, whereas a single linkage treats all parameters in the interaction as exchangeable. The choice must be determined by the need to explain effects, as guided by available scientific insights. A general rule is not possible, other than to examine measures of predictive accuracy (Section 8.6.2) and assess probabilistic predictions relevant to the intended application of the model.

In addition to including the possibility of unique effects for each realization of ($\delta \times \Phi \times p \times f$) at a given particle fraction, we would like to model trends using the set-specific parameters as discussed in Section 8.5.2. In doing so, we note that it is preferable to construct hierarchical groupings that contain more than the minimum needed (3 parameters) for the variance to be determined from the data (i.e. hierarchical variance assigned a noninformative prior), so that the posterior distribution is not heavy-tailed. In essence, we would like to model effects for which ample realizations exist in the data. This predisposes one against applying hierarchical groupings based only on δ, p, or f, as, respectively, only 3, 4, and 3 realizations exist; Φ-regime only has 2 realizations, thus it is not applicable due to insufficient parameters.

Furthermore, we note that of the 72 possible combinations of (δ, p, f, Φ), only 23 unique combinations were visited, though 41 unique batches were realized as the particle fractions are distinctly different in the repeated visits of ($\delta \times \Phi \times p \times f$) – e.g. single drop dose, high particle loading rheological regime, sfAPAP, NEO895, particle fractions 0.175, 0.2, 0.25, 0.3. Therefore, one must be judicious in selecting which effects to model, taking into consideration not only the

number of parameters available for estimating the effect, but also whether two proposed hierarchies are sufficiently different due to the availability of data (or lack thereof). Additionally, the number of realizations in the dataset is not necessarily balanced in this scenario, thus, the hierarchical groupings should ideally be constructed in a manner that makes the contributions to each parameter reasonably balanced.

To this end, two types of possible trends are used in the selected models: those over ($\Phi \times p \times f$) and ($\delta \times p \times f$), as they are sufficiently different in the available data and provide ample parameters to make estimates of hierarchical variances. For the ($\Phi \times p \times f$) there are 14 realizations out of a possible 24, and for the ($\delta \times p \times f$) grouping there are 18 realizations out of a possible 36; within these groupings, the data is reasonably balanced. In setting up each grouping, we choose to link the each set of β's in a single hierarchy rather than linking with respect to potential sub-divisions such as rheological regime or dose potency level; this decision was based on comparison of the model predictions under single/separated linkages.

It is worth commenting on the various other hierarchical groupings which can be used, namely the ones which seem to offer a sufficient number of parameters and have data to support them: ($p \times f$), ($\Phi \times p$), ($\Phi \times f$), ($\delta \times p$), ($\delta \times \Phi \times p$), ($\delta \times \Phi \times f$), etc. In evaluating models, many of the possible combinations were observed to give essentially equal predictive performance in terms of test variables, information criteria and forecasting for out-of-sample data at in-sample conditions. This may be due to there being little remaining information which was not already accounted for by the other groupings. The main effect of adding further groupings was to increase the posterior uncertainty in predictions for out-of-sample data at out-of-sample conditions, which is what is expected under a hierarchical model with non-informative uniform priors assigned to the hierarchical variance parameters; while uncertainty is desirable in a model, addition of too many hierarchical groupings of this sort can result in the predicted intervals that too diffuse to be useful for inference. A potential solution to resolve this is to use weakly informative prior distributions such as the half-Cauchy distribution.

8.8.3 Utility of both ModelDrop & ModelDose

Given the similar predictive performance of both model concepts, it is worthwhile to elaborate on why both remain useful, depending on the intended use. The ModelDrop approach has advantages in certain situations – one can more confidently employ the model to forecast

individual drop volumes, as this finds use for a scenario in which one is producing a multi-drop dose, and perhaps for economic reasons, it is desired to pause in the middle of deposition, e.g. after half of the total drops have been deposited, and forecast what the final amount would be given the remaining ($M^k/2$) drops and the $\tilde{X}$ for the first half. The forecasted final amount would then be used to make a control decision, for which many possibilities exist – perhaps most simply by stopping deposition in that dose unit if predicted to be outside specification, thereby eliminating unnecessary expenditure of (expensive) ink on what would ultimately become a rejected dose unit.

For this scenario, the ModelDrop approach does not encounter any conceptual difficulty in forecasting the "missing" data for the as-to-yet unprinted drops. The ModelDose approach presents a dilemma: to forecast the missing data is slightly inappropriate, given that normally one waits for all drops to be printed to form the vector of explanatory variables. One could treat the drops individually as single drop doses, but then there is the problem of which data variance is appropriate – clearly, the heteroscedastic structure necessitates use of the data variance for $s_{\delta=single\,drop}$, and perhaps this may work. Similarly, one encounters problems with the group-level predictors – it seems one is forced to use those relevant to single drop doses, rather than multi-drop doses – that only get worse if one is working with new data from new groups (where $\tilde{\beta}$'s must be generated from their appropriate posterior predictive distributions).

The ModelDose approach is generally applicable to any completed dose and/or batch of doses, for which all M^k drops have been deposited. Hence it remains relevant, and, indeed, may be better suited to forecasting whether a partially completed batch will be within specification – one could perform an analogous forecast of "missing" data to that described above, in which the remainder of the batch is predicted using the $\tilde{X}$ for the preceding completed doses. Though, it is difficult to imagine why one might reject the entirety of a (partial) batch which was predicted to be outside specifications, rather than simply to reject the offending units and print replacements.

Thus, both modeling approaches have potential uses. Namely, if the online control scheme includes a mid-deposition pause and forecast step as described above, then ModelDrop is preferably used. If the goal in the online control scheme is only to forecast for the completed doses (i.e. units to which all drops have been printed), then either the ModelDose or ModelDrop can work equally well. There are no conceptual problems with using the ModelDrop in a "pause and forecast" step, and subsequently using the ModelDose to forecast and accept/reject for the completed doses. Ideally, for the completed dose forecast, whichever of the two models has better

predictive accuracy would be selected, perhaps taking into consideration the predictive accuracy for new data from new groups. In summary, depending on the intended use, it may be preferable to use a different model than the one which best fits the training data.

8.8.4 Final Remark

Comparison of the predictions by the two model concepts reveals only one notable difference: the ModelDrop approach makes better predictions for new data from out-of-sample conditions in that more of the density lands near the "actual" values, though both models appear to have approximately equal performance in terms of predicting the mean value. It should be noted that in generating from the posterior predictive distribution of the group-level predictors for the ($\delta \times \Phi \times p \times f$) hierarchies, batches 2, 3, 4, and 5 are not naturally exchangeable with any of the groupings, as they belong in the ($\delta = many\ drop \times \phi = low$) combination which was not included in the training data. One might attempt to treat them as exchangeable with $s_{\delta = many\ drop \times \phi = high \times p \times f}$, which is what we opt to do here; the usefulness of such an assumption depends on the model, as can be seen from comparison of Figure 8.12 and Figure 8.18.

Both model concepts yield the wide confidence regions one would expect under a hierarchical model for prediction at out-of-sample conditions, but the ModelDose approach appears to suffers more from this assumption. A potential solution to this problem is to use a 3-level model for the ($\delta \times \Phi \times p \times f$) hierarchies, linking the means of these distributions in another hierarchical layer. Additionally, one could add more explanatory variables at the second or third layer. A simple but effective solution for prediction at out-of-sample conditions under the ModelDose approach is to replace the 1's in the indicator columns with the respective M^k values, so that the regression on the second level is inherently scaled to the number of drops per dose. The result is a model which is essentially identical for prediction at out-of-sample data from in-sample conditions, but arguably performs better at out-of-sample conditions.

Careful inspection of the $F_{\hat{y}^k}$ for the new data from existing groups (Figure 8.10 vs. Figure 8.16) reveals what appears to be slightly better predictive performance by the ModelDose approach in the sense that the 50% interval of the predictive density encompasses the "actual" (i.e. measured) values for a few more batches than the ModelDrop approach. This suggests that for new data from in-sample conditions, either model approach can work equally well (with the dose model perhaps

slightly better), but for new data from out-of-sample conditions, the drop model is the preferred model.

8.9   Conclusions

In this work a probabilistic modeling approach is proposed as means of utilizing non-traditional online sensor data to predict critical quality attributes. The Bayesian method enables the use of the historical data from a substantial variety of conditions to quantify the uncertainty inherent to both the sensor and process in the model, and therefore employ probabilistic forecasting for newly acquired data as a means of quality control. Through hierarchical structuring of the data to represent the conditions at which it was gathered, we are able to capture the effects of variable interactions without needing to provide detailed (and currently unavailable in the literature) knowledge of the underlying physics of the process. Moreover, this enables us to make better predictions for new data collected at conditions which were not included in the training data, while continuing to quantify our uncertainty about these predictions; this can only be accomplished in a hierarchical setting, where we utilize the posterior predictive distribution of the group-level regression parameters to forecast new parameters, then forecast new observables.

The methodology and framework presented here are applicable to linear and generalized linear models. It is expected that with this extent of flexibility, Bayesian models such as these might become more widespread in the process and systems community, where decisions must often be made under uncertainty. Given the increasing amounts of available data, Bayesian approaches seem to be preferred way to make use of what is known from the physics, and let the remainder be informed by the data; coincidentally, such methods are sufficiently flexible as to allow the model structure to be largely determined by the data, thus, there is no better way to quantify uncertainty.

8.10 Acknowledgements

8.11 References

[1] Ierapetritou, M., Muzzio, F., Reklaitis, G., 2016. Perspectives on the continuous manufacturing of powder-based pharmaceutical processes. AIChE J., 62(6), 1846-1862.

[2] Tabora, J. E., Lora Gonzalez, F., Tom, J. W., 2019. Bayesian probabilistic modeling in pharmaceutical process development. AIChE J., 65(11), e16744.

[3] Mockus, L., Reklaitis, G., Morris, K., LeBlond, D., 2020. Risk-Based Approach to Lot Release. J. Pharm Sci., 109(2), 1035-1042.

[4] Qin, S. J., Chiang, L. H., 2019. Advances and opportunities in machine learning for process data analytics. Comput. Chem. Eng., 126, 465-473.

[5] Venkatasubramanian, V., 2019. The promise of artificial intelligence in chemical engineering: Is it here, finally. AIChE J.,65(2), 466-78.

[6] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B., 2013. Bayesian Data Analysis, third ed. CRC press.

[7] Gelman, A., Hill, J., 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models, first ed. Cambridge University Press.

[8] Hirshfield, L., Giridhar, A., Taylor, L. S., Harris, M. T., Reklaitis, G. V., 2014. Dropwise additive manufacturing of pharmaceutical products for solvent-based dosage forms. J. Pharm. Sci., 103(2), 496-506.

[9] Içten, E., Giridhar, A., Taylor, L. S., Nagy, Z. K., Reklaitis, G. V., 2015a. Dropwise additive manufacturing of pharmaceutical products for melt-based dosage forms. J. Pharm. Sci., 104(5), 1641-1649.

[10] Icten, E., Nagy, Z. K., Reklaitis, G. V., 2015b. Process control of a dropwise additive manufacturing system for pharmaceuticals using polynomial chaos expansion based surrogate model. Comput. Chem. Eng., 83, 221-231.

[11] Hirshfield, L., Içten, E., Giridhar, A., Nagy, Z. K., Reklaitis, G. V., 2015. Real-time process management strategy for dropwise additive manufacturing of pharmaceutical products. J. Pharm. Innov., 10(2), 140-155.

[12] Radcliffe, A. J., Hilden, J. L., Nagy, Z. K., Reklaitis, G. V., 2019. Dropwise Additive Manufacturing of Pharmaceutical Products Using Particle Suspensions. J. Pharm. Sci., 108(2), 914-928.

[13] Radcliffe, A. J., Reklaitis, G. V., 2020a. An application of computer vision for optimal sensor placement in drop printing. In Comput. Aided Chem. Eng. (in press). Elsevier.

[14] Radcliffe, A. J., Reklaitis, G. V., 2017. Dropwise additive manufacturing using particulate suspensions: feasible operating space and throughput rates. In Comput. Aided Chem. Eng. (Vol. 40, pp. 1207-1212). Elsevier.

[15] Girdhar, A. PinPoint Pharma, personal communication, April 10, 2020

[16] Radcliffe, A. J., Reklaitis, G. V., 2020b. Automated Object Tracking, Event Detection and Recognition for High-Speed Video of Drop Formation Phenomena. In submission.

[17] Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Analysis, 1(3), 515-534.

[18] Gelman, A., 2004a. Parameterization and Bayesian modeling. J. Am. Stat. Assoc., 99(466), 537-545.

[19] Gelman, A., Van Dyk, D. A., Huang, Z., Boscardin, J. W., 2008. Using redundant parameterizations to fit hierarchical models. J. Comput. Gr. Stat., 17(1), 95-122.

[20] Robinson, G. K., 1991. That BLUP is a good thing: the estimation of random effects. Stat. Sci, 6(1), 15-32.

[21] Rubin, D. B., 1980. Using empirical Bayes techniques in the law school validity studies. J. Am. Stat. Assoc., 75(372), 801-816.

[22] Deely, J. J., Lindley, D. V., 1981. Bayes empirical bayes. J. Am. Stat. Assoc., 76(376), 833-841.

[23] Boscardin, W. J., Gelman, A., 1994. Bayesian computation for parametric models of heteroscedasticity in the linear model. Univ. Calif. Berkeley.

[24] Dempster, A. P., Rubin, D. B., Tsutakawa, R. K., 1981. Estimation in covariance components models. J. Am. Stat. Assoc., 76(374), 341-353.

[25] Hodges, J. S., 1998. Some algebra and geometry for hierarchical models, applied to diagnostics. J. R. Stat. Soc.: Ser. B (Stat. Methodol.), 60(3), 497-536.

[26] Van Dyk, D. A., Meng, X. L., 2001. The art of data augmentation. J. Comput. Gr. Stat., 10(1), 1-50.

[27] Gelman, A., 2005. Analysis of variance—why it is more important than ever. Ann. Stat., 33(1), 1-53.

[28] Gelman, A., Rubin, D. B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci., 7(4), 457-472.

[29] Brooks, S. P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Gr. Stat., 7(4), 434-455.

[30] Gelman, A., 2003. A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. Int. Stat. Rev., 71(2), 369-382.

[31] Shirley, K. E., Gelman, A., 2015. Hierarchical models for estimating state and demographic trends in US death penalty public opinion. J. R. Stat. Soc.: Ser. A (Stat. Soc.), 178(1), 1-28.

[32] Gelman, A., 2004b. Exploratory data analysis for complex models. J. Comput. Gr. Stat., 13(4), 755-779.

[33] Gelman, A., 2013. Two simple examples for understanding posterior p-values whose distributions are far from uniform. Electron. J. Stat., 7, 2595-2602.

[34] Gneiting, T., Balabdaoui, F., Raftery, A. E., 2007. Probabilistic forecasts, calibration and sharpness. J. R. Stat. Soc.: Ser. B (Stat. Methodol.), 69(2), 243-268.

[35] Vehtari, A., Ojanen, J., 2012. A survey of Bayesian predictive methods for model assessment, selection and comparison. Stat. Surv., 6, 142-228.

[36] Watanabe, S., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res., 11(Dec), 3571-3594.

[37] Vehtari, A., Gelman, A., Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. and Comput., 27(5), 1413-1432.

[38] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. R. Stat. Soc.: Ser. B (Stat. Methodol.), 64(4), 583-639.

[39] Gelman, A., Hwang, J., Vehtari, A., 2014. Understanding predictive information criteria for Bayesian models. Stat. Comput., 24(6), 997-1016.

[40] USP General Chapter <905> "Uniformity of Dosage Units," USP38, US Pharmacopeial Convention, December 1, 2015.

9.   CONCLUSION

Through the combination of physical knowledge and data-driven methods, a framework was developed for ink formulation and process operation in drop-on-demand manufacturing with non-Brownian suspensions. With a sound basis provided by the rheology and drop physics, the experimental work presented in Chapter 3 and Chapter 4 provides support for the general applicability of drop printing from non-Brownian suspensions as a manufacturing method for not only pharmaceuticals, but for any application in which the deposition of micrometer-scale particles is desired.

The importance of particle properties for ink formulation have been explored in terms of particle size and particle shape, however, the investigation of the effect of carrier fluid properties was limited. This work identified the importance of maintaining particle-fluid coupling during printing, that is, the importance of selecting a carrier fluid with sufficient viscosity to keep the *Stokes* number below 1. Underlying reasons for this were eminently practical, in that it was observed that for *Stokes* numbers greater than 1, the extent of particle-liquid segregation was catastrophic. Failure would occur during attempted start-up of the process, or shortly thereafter; as an example: the liquid would visibly be filtered from the particles, yielding optically translucent fluid in the transfer tubing, though the prepared ink was optically opaque. As it relates to ink formulation, these observations suggest an interesting area of future work in order to elucidate the interaction of particles, fluid properties and operating conditions on flow-driven segregation.

Furthermore, the effect of non-Newtonian carrier liquids (polymeric melts, or carrier fluid containing dissolved polymer) on drop formation in non-Brownian suspensions is a fruitful area of investigation given the prevalence of bioavailability modifiers which are polymeric in nature. In this work, a limited number of such suspensions were tested (MA, PB in PEG300), but with surprising results: drop formation was considerably different from that in NEO895, despite a difference in carrier viscosity of a factor of 2.5.

As the underlying physical mechanisms at such extreme conditions were a new area of study, investigation combined experimental work with theory and simulation. Collection of experimental data on drop formation necessitated the use of high-speed imaging, which generated a massive image dataset. Motivated by the need to automate extraction of quantitative information from these high-speed videos, new computational image analysis techniques were developed by

combining first-order logic (as informed by the physics) with a set-based approach. From the perspective of the flow phenomena, this yielded insight into the effects and interactions caused by non-Brownian particles in drop formation. In a simpler sense, the dynamic velocity profiles provide insight into the effects on flow which have implications for ink formulation and process operation. However, the more interesting – modal and statistical – information combined with the spatiotemporal trajectories presents and opportunity for advancing the fundamental understanding of drop formation in non-Brownian suspensions; time constraints abrogated the possibility in this work. On the other hand, an eminently practical use of the extracted information was achieved in this work: optimizing the components of the online image sensor with respect to the drop breakup and coalescence phenomena will lead to improved monitoring capabilities, thus, perhaps some of the observations documented in Chapter 8 (e.g. Figure 8.5) could be preempted.

More generally, it is expected that these computer vision algorithms can readily be applied by other researchers interested in the physics of drop coalescence and breakup in order to further modeling efforts, which in particle-containing fluids exhibits stochastic effects that necessitate an increasingly data-driven approach. To the author, this seems to fill an unmet need, as there are no publicly available tools for analyzing high-speed video of drop formation, despite the large number of researchers working on the topic. Though a personal observation, one notes that in many of the publications which contained experimental work involving drop formation, the number of repetitions used to draw a conclusion was usually small, $n = 10$ or so, for each condition; given the ease of collecting the high-speed videos, one suspects this is due to the burden of analyzing the image sequences, which if done by hand, would be extremely time-consuming. If these researchers availed themselves of the computer vision algorithms developed here, analysis would be equally easy as collection; moreover, there would be far more detailed information available to inform the modeling work.

Returning to the realm of process development to deal with challenges of monitoring and quality control initiated by suspension-based manufacturing, these new machine vision algorithms enabled development of robust models from an existing online process image sensor. As what seems to become a recurring theme in this research, image data held valuable information that can enable modeling, but to make use of it required automated image analysis. For this process, and potentially for many others (as noted in Chapter 5, Introduction), image sensors can provide

complex measurements of quantities of interest; the difficulty lies in extracting the information of interest in an automated manner – for this, one must turn to computer vision methods.

An image sensor used as a means of online process monitoring is subject to various sources of uncertainty, just as any other process monitoring tool would be. Consequently, to make the best possible use of the information, additional modeling work is necessary, and it is the viewpoint of this author that these models should be Bayesian. To this end, implementation of a real-time quality control strategy using the Bayesian approach incorporated the various sources of uncertainty inherent in the data collection process; the use of hierarchical models permitted inclusion of inherent structures in the large historical dataset of ≈1500 dose units produced at 41 unique conditions, therefore providing improved probabilistic prediction for new data collected at out-of-sample conditions.

Taken together, these contributions constitute a small advance in the large picture of this research. The insightful reader may be inclined to find the general structure of particular concepts and adapt them to their purposes – to the author, this appears to be the use of hierarchical Bayesian modeling for monitoring and quality control under uncertainty, perhaps in concert with image-based process data.

10. FUTURE DIRECTIONS

10.1 Temperature Control

Much work remains to be done in the development of this dropwise additive manufacturing system for production of pharmaceuticals. As a small-scale manufacturing platform, the need to accommodate more potential inks and dosage forms motivates a re-envisioning of the temperature control systems on the process. Specifically, the opportunities to improve upon existing systems at the substrate appear to be fruitful; in this respect, the combination of data-driven control strategies with imaging sensors is promising as a means of spatial temperature control. This would enable customization of critical quality attributes related to crystal structure and morphology, hence, the ability to produce a wide variety of novel dosage forms with tailored bioavailability.

Furthermore, improvements to control of temperature in the printing system itself would enable control of fluid rheology via precision temperature adjustment. To enact this greatly increases the range of ink properties that can be accommodated by the system, and consequently, the capability of the system to produce specialized products – from new inks – with greater ease. Immediately available examples would be non-Brownian suspension inks in which the carrier fluid is a melted lipid, a natural extension of the NEOBEE 895 formulations already demonstrated; the advantage is that the printed product would solidify upon cooling to ambient temperature, thereby forming a more stable dosage form. Perhaps more importantly, precise temperature control provides the means to modify the rheology of complex polymeric fluids, potentially making otherwise unprintable inks printable, or to optimize rheological properties so as to achieve the desired printing performance.

10.2 Online Concentration Monitoring

In addition to temperature control, a robust means of online concentration monitoring is necessary when working with particle suspensions. This is a challenge due to not only the small scale of the process, but also the presence of particles themselves, which preclude the use of many inline measurement technologies. A cost-effective measurement technology is under investigation in the form of an inline ultrasound sensor, with inline Raman spectroscopy as an expensive, albeit quite capable alternative. With the future uses of the process in mind, online concentration

monitoring is a necessity for integration with upstream continuous processes, hence the motivation to find one or more measurement technologies capable of inline measurement.

Integration of the measurements from such technologies with the Bayesian models for the online image sensor is readily achieved: in the simplest case by including these in the explanatory variable matrix, or perhaps more elaborately, expanding the models to include a multivariate output. From the concentration sensor(s) and online image sensor, more complex models including time-series data are eminently feasible. Seemingly, the appropriate modeling methods are Bayesian, as only the Bayesian approach provides the means to estimate the many observable and latent parameters necessary to construct such models. Moreover, the Bayesian modeling inherently enables one to quantify uncertainty so as to make risk-based decisions about product quality attributes – a case for which has been demonstrated here.

10.3 Integration of Dropwise Additive Manufacturing System with Upstream Processes

The above are improvements to various aspects of the DAMP itself, but in an expanded future perspective, this process is envisioned as a dose production platform downstream from a small-scale synthesis and purification platform which has bene developed by colleagues. By combining these two platforms, the overall process would convert raw materials into finished drug products within a physical footprint not much larger than a lab bench. The suggested improvements to the DAMP anticipate some of the issues which must be addressed to achieve this integration, but other challenges are expected.

The foremost of these is the real-time mixing of the ink which must occur so as to achieve printable fluid properties. As the solution or slurry leaves the upstream processes, the fluid properties of nearly all synthesis and crystallization solvents are not suited to drop-on-demand printing due to insufficient viscosity. For slurries, this is even more problematic as while the bulk viscosity may be in the correct range, the high shear rates (10^3-10^4 s^{-1}) will cause substantial particle-liquid segregation if the carrier fluid viscosity is too low; this relationship is described by the Stokes number, $St = \rho_p R^2 \dot{\gamma} / \mu_f$, which as a measure of particle-fluid coupling, must be < 1, or preferably < 0.1, based on experimental evidence.

There exist several options for modifying viscosity, however, as this simultaneously affects the drug product properties, they must be considered together. If a solvent is to be the carrier fluid, a

soluble polymer (e.g. hydroxypropyl methylcellulose) or surfactant (e.g. polysorbate 80) may be added to adjust its viscosity so that the Stokes number is acceptably low and the bulk viscosity is not too high; this soluble polymer (and/or surfactant) may be chosen so as to enhance bioavailability, create a controlled release product, or otherwise modify the drug product properties. The alternative is to perform a swap of the select carrier liquid for an inherently more viscous fluid such as an edible oil (e.g. NEOBEE 895), oil-surfactant blend, or PEG – any of these may be solid at ambient temperature, and thus supplied for process use in melted form. Analogously to the above, the alternative carrier fluid should be selected so as to achieve the desired formulation properties. In either option, temperature control can be used to modify (i.e. decrease) viscosity of the bulk fluid in order to make it printable.

10.4 Modeling Support

This implicitly envisions drop-on-demand printing from a wide range of unexplored inks. The experiments conducted here provide a general guideline as to the operating conditions and ink formulation properties for printing of non-Brownian suspensions, but it is apparent that more is needed to support rapid adjustments as may be demanded in the future. Given the complexity of the underlying physics, a pragmatic path is envisioned by combination of high-speed imaging with the computer vision algorithms to extract the quantitative information, thereby enabling data-driven modeling of drop-on-demand drop formation in non-Brownian suspensions.

The machine vision algorithms enable automated analysis of drop formation videos, thus, there is little more burden to obtaining spatiotemporal trajectories and events information from any number of high-speed videos than the collection of the videos themselves. Hence, the collection of a large number (e.g. 100-1000) of videos at each condition under investigation could be used to form a database, the quantitative information from which would enable a variety of modeling approaches; in this sense, each new printing condition attempted would contribute to an increasingly vast historical dataset.

It seems to this author that the models should be Bayesian and seek to take advantage of hierarchical structuring so as to model trends across exchangeable conditions. First-principles knowledge could be incorporated in a variety of ways depending on the level of granularity attempted by the modeler. This undertaking represents a considerable challenge due to the complex

physics of drop formation and also due to the computational difficulties anticipated for the spatial statistical models under the Bayesian paradigm.

Consequently, it may be more tractable to employ advanced statistical modeling techniques and use the known physics to add structure to the models. While a departure from the traditional realm of chemical engineering, it seem to this author to be an acceptable compromise in the face of complex physics marred by various (and seemingly unknown) stochastic elements. This type of scenario has frequently been encountered by chemical engineers; whereas in the past data may not have been available, it is increasingly possible to acquire data to supplant the known physics and utilize a hybrid modeling approach. In this respect, the author would like reader to note that replacing physics-based modeling entirely is an untenable proposition. Rather, the use of data to supplement the known physics in a hybrid approach can serve as a pragmatic means of grasping the unknown. To this end, the author must point out that the Bayesian approach is the only data-driven modeling technique that permits incorporate of the unknown physics in order to impart inherent structure to the data at hand, and moreover, permits quantification of uncertainty in both parameters and predictions.

APPENDIX A. CHAPTER 2 LITERATURE REVIEW

DROP FORMATION

In a drop-on-demand (DOD) manufacturing processes, productivity maximization and final product consistency are the most important requirements. Productivity may be taken to be the product of drop volume and drop ejection frequency, thus, the maximum value of this variable will be constrained by the fluid and system properties that permit consistent drop formation at some frequency. When considering a DOD process, the relevant performance criteria are drop volume, drop speed, drop shape, jet straightness (drop trajectory) and drop content uniformity (particle mass) – all measured at some drop ejection frequency; to generate a consistent final product, these criteria must be satisfied to within deviation levels the user permits. To identify the fluid/system properties that affect our performance criteria, it is necessary to consider the physics governing drop formation.

Drops are created from the fragments formed by the breakup of a column of liquid, which may be produced during a flow or exist statically. When a perturbation of sufficient wavelength is introduced along the axial dimension of a liquid column, the radius must vary such that volume is conserved as the wave traverses the column; the decrease in radius creates an instability due to the scaling of capillary pressure (surface tension $\equiv \gamma$, $P_{capillary} = \gamma H \sim \gamma R^{-1}$ where R is the mean curvature of an infinite cylinder and H is the mean curvature for an arbitrary surface) [1]. The imbalance in pressure results in a flow which initiates the eventual collapse of the liquid filament into drop(let)s. Examples are found in the jet of water one observes upon opening a faucet (a differential fluid element here would be convected at approximately the nozzle velocity), or a liquid bridge which forms between two solid surfaces (in static equilibrium). In both cases, the column of liquid is stable to small wavelength perturbations, but unstable for sufficiently large perturbations. As we are interested in industrial applications, we shall restrict our scope to phenomena which result from flows; consequently, the driving force for initiation of breakup is most frequently surface tension (also possibly gravity), which is resisted by viscosity, inertia and elasticity (if polymers are present) [2, 3, 4, 5].

The radial decay (collapse) of Newtonian liquid filaments has been observed to be a universal phenomenon, where the rate of radial decay (in absence of gravitational force) is

determined by the balance between fluid viscosity (μ), density (ρ) and surface tension (γ), as modified by the radius (R) of the liquid filament; this force balance is most frequently quantified using the *Ohnesorge* number,

$$Oh = \frac{\mu}{\sqrt{\rho R \gamma}} = \frac{viscous}{inertial \cdot surface},$$

noting that when $Oh \leq 1$, resultant drops are expected to form with little influence of fluid viscosity [4]. For sufficiently wide columns of liquid, the effect of gravity will cause the filament to sag, thus accelerating the break-up process moreso than surface tension alone; the relative influence of surface tension and gravity are quantified using a *Bond* number,

$$Bo = \frac{\rho g R^2}{\gamma} = \frac{gravitational}{surface},$$

where values of $Bo \geq 0.1$ constitute cases in which gravitational force plays a significant role [4, 6]. These two dimensionless groups provide information on the form assumed by the filament and its radial decay rate, but they do not tell us under what mode the drop is formed; to determine this, the effect of the flow velocity at the nozzle (jet velocity, U_{jet}) must be considered [4]. Dependent upon U_{jet}, liquid exiting a nozzle will experience qualitatively different phenomena: [1]pendant drops of consistent size where detachment is initiated by the fluid's weight (stable dripping), [2]pendant drops of variable size (chaotic dripping), or [3]a column of liquid (jetting) which eventually breaks up due to perturbation-initiated instability [6, 7]. For a fixed set of fluid properties (ρ,μ,γ) and nozzle size (R), stable dripping is observed at the lowest flow rates, chaotic dripping at modest flow rates (only occurs at low Oh), and jetting at a critical value of U_{jet}; the dimensionless grouping used to describe the relative influence of forces (inertial, surface) governing this transition is the *Weber* number,

$$We = \frac{\rho U_{jet}^2 R}{\gamma} = \frac{inertial}{surface},$$

for which values of $We \geq 1$ form jets in inviscid and viscous Newtonian fluids [4].

Thus, for Newtonian fluids, the relevant dimensionless groupings to consider in drop-on-demand drop formation are the *Bond, Ohnesorge* and *Weber* numbers [2, 3, 4, 5, 8]. Studies of drop-on-demand printing (using Newtonian and non-Newtonian fluids) make use of these dimensionless groups to quantify the effect of ink properties on the performance criteria (typically defined in the same manner as in the first paragraph of this section) [9, 10, 11]. Studies of non-Newtonian fluids introduce additional dimensionless groups which relate the relative timescales of viscoelastic to surface forces (*Deborah* number), and to the relative timescales of the process

and that of viscoelasticity (*Weissenberg* number) [12, 13]. As the fluids used in this work will exhibit non-Newtonian rheological properties, and due to the expansive nature of non-Newtonian fluids in DOD inkjet applications, further discussion will focus on the subset of non-Newtonian fluids comprised of hard-particle suspensions.

SHEAR RHEOLOGY OF NON-BROWNIAN SUSPENSIONS

As drop formation is heavily influenced by the viscosity and elasticity of a given fluid, which in dimensionless form enter the stress balance as the *Ohnesorge*, *Deborah* and *Weissenberg* numbers, understanding of the rheological properties is imperative to the development of a reproducible process. A suspension is broadly defined as a dispersion of particles in a liquid [14], though this provides no information about the particles, interstitial fluid or rheological properties of the composite. Two categories of non-Newtonian fluids may be distinguished from this general definition, with the distinction arising due to the (particle size-mediated) relative influences of convection and Brownian diffusion on particle motion. A (particle) Péclet number,

$$Pe = \frac{r_p{}^2 \dot{\gamma}}{\left\{\frac{k_B T}{6\pi\eta_{liq} r_p}\right\}} = \frac{Convection\ due\ to\ shear}{Brownian\ diffusion} = \frac{6\pi r_p{}^3 \eta_{liq} \dot{\gamma}}{k_B T},$$

is often used as the defining characteristic; a large Péclet indicates that the particles in the suspension are not appreciably affected by Brownian motion and their motion is governed only by the fluid flow, gravity and interactions with other particles/surfaces [15, 16]. Typically, for a particle diameter greater than 1μm the suspension is taken to be non-Brownian; note that non-Brownian suspension are often referred to as non-colloidal suspensions [16]. The rheological properties of the interstitial fluid can introduce a further distinction, should the particles be suspended in a non-Newtonian (e.g. polymer solution) carrier fluid, leading to a possibility of four distinct types of suspensions [17, 18, 19]. The particle sizes presented by micronized API powders causes suspensions thereof, in Newtonian or non-Newtonian fluids, to be non-Brownian (non-colloidal), thus, the remaining influences on particle motion in the fluid will result from shearing, gravity, and potentially viscoelastic effects (non-Newtonian interstitial fluids only).

The viscosity of a non-Brownian suspension (with monodisperse spheres) depends on a set of variables related to the particle/fluid properties and system conditions [15, 16, 20, 21]:

$$\eta = f\left(r_p, \rho_p, N_p, \eta_0, \rho_0, k_B T, \dot{\gamma}, g\right);$$

By dimensionless grouping, this variable set may be reduced to:

$$\eta_r = \frac{\eta}{\eta_0} = f(\phi, Re, St, Pe, Ar),$$

in which: $\phi = N_p \frac{4}{3}\pi r_p^3 = \frac{\frac{m_p}{\rho_p}}{\frac{m_p}{\rho_p}+\frac{m_0}{\rho_0}}$, $Re = \frac{\rho_0 r_p^2 \dot\gamma}{\eta_0}$, $St = \frac{\rho_p r_p^2 \dot\gamma}{\eta_0}$, $Pe = \frac{6\pi r_p^3 \eta_0 \dot\gamma}{k_B T}$, $Ar = \frac{2 r_p^2 (\rho_p - \rho_0) g}{9\eta_0 \langle U \rangle}$

identifying the significance of the particle volume fraction (ϕ), particle size (r_p = *particle radius*), interstitial fluid viscosity (η_0) and deformation rate ($\dot\gamma$).

In absence of thermal effects ($Pe \rightarrow \infty$), density mismatching/gravitational effects ($Ar \rightarrow 0$) and shear rate dependence ($Re, St \rightarrow 0$), viscosity dependence on volume fraction in concentrated suspensions follows an exponential or power law relation [15, 16, 21, 22, 23]. One example of such a model, derived based on hydrodynamic considerations is:

$$\eta_r = \left(1 - \frac{\phi}{\phi_{max}}\right)^{-B\phi_{max}} \text{[Krieger-Dougherty model]},$$

in which the empirical constants, B and ϕ_{max}, are used in fitting a set of rheological data evaluated at a single steady shear rate; B is the Einstein coefficient, dependent on particle shape, and ϕ_{max} is the maximum particle volume fraction, dependent on particle size distribution and shape [24]. More general models (with origins in granular rheology) that permit description of shape-heterogeneous particulate suspensions eliminate dependence on the shape factor, B, preferring a model written as $\eta_r = f\left(\frac{\phi}{\phi_{max}}\right)$, wherein ϕ_{max} accounts for the effect of both particle shape and size distribution [25, 26, 27].

Regardless of whether the given rheological model specifically takes into account a particle shape factor, the effects of particle size polydispersity and shape deviation from that of a spheres are much the same. In terms of maximum particle packing fraction, increased particle size polydispersity corresponds to increased ϕ_{max} (provided that finite-size effects are negligible in the geometry of interest), as the smaller particles can fill the interstitial voids between large particles, leading to increased solid fraction for a given control volume [18, 24, 28, 29, 30, 31, 32]. Defining particle aspect ratio (*AR*) as the ratio of longest to shortest characteristic length, the effect of deviations from *AR*=1 (spheres) is to decrease ϕ_{max} (maximum value occurring at *AR*=1); this results primarily from hydrodynamic effects that arise due to the rotation of the particle during fluid flow [28, 31] – however, geometric constraints are also introduced by existence of an aspect ratio distribution, directly decreasing the maximum solid packing fraction (corresponding to increased void fraction) of the powder in the dry state [33].

In addition to particle volume fraction dependent effects, non-Brownian suspensions exhibit: [1]shear-stress/rate dependence (shear-thinning/thickening, yield stress), [2]segregation due to particle sedimentation (when particle and fluid not density-matched), and [3]multiple types of migration/segregation resulting from shear. Due to the enormity of the body of works on these topics a select few results will be presented, but in the future a more comprehensive review will be provided. Non-Brownian suspensions' viscosity dependence on shear-rate (or shear stress) is a well-known phenomenon resulting from a number of complex (and presently disputed) causes; predictive models do not yet exist for the shear-rate(stress) dependence for a given combination of particles and suspending fluid [23] though with details pertaining to particle volume fraction, particle properties, deformation rate and interstitial fluid properties, reasonable predictions may be made as to the expected categories of flow phenomena.

The occurrence of shear-thinning behavior in a fluid is generally taken indicate the existence of microstructure within the fluid that develops as $\dot{\gamma} \to 0$, and under increasing shear rate this microstructure undergoes rearrangement to a configuration that is more favorable to flow (exerts lower resistance to flow), which appears macroscopically as a decrease in viscosity across some shear rate interval [34]. In non-Brownian suspensions, shear-thinning behavior results from individual particle rearrangement in response to an imposed flow field: non-spherical (e.g. high aspect ratio) particles exhibit preferred flow orientations, and during flow start-up such particles take time to reorient, resulting in a viscosity decrease [31]. The existence of shear-thinning behavior have also been observed to result from particle-fluid density mismatching; in this case shear-thinning results from shear-rate dependent resuspension of particles through flow motion, causing an observed decrease in viscosity [35, 36].

The incipience of a yield stress in a non-Brownian suspension depends upon particle size polydispersity, density (mis)matching, particle roughness/frictional interactions, attractive interactions (e.g. electrostatic) and particle volume fraction [23]. The magnitude of a yield stress in a non-Brownian suspension, which results from the existence of microstructure at zero shear rate, has been successfully related to particle-fluid density mismatching for the case in which particles do not exhibit appreciable surface interactions. [36, 37].

Identifying the causes of shear-thickening in non-Brownian suspensions is presently the topic of much experimental and computational work, recent reviews of which are found in *Brown/Jaeger* and *Denn/Morris* [23, 16]; to keep this section sufficiently brief, the potential causes

for this phenomenon will be given without further detail. Shear-thickening results from a normal stress difference arising during deformation and, wherein underlying the physical mechanisms for a normal stress difference are thought to result from hydroclustering (particle migration under shear resulting in increased effective viscosity), order/disorder transition in microstructure, and/or dilatancy effects. Though each of these mechanisms appears to be sufficient (but not necessary) for producing a shear-thickening response, the one that has received much attention recently is dilatancy, or the tendency of dense granular material to increase in volume when sheared [38]. Following the dilatancy route of analysis, a minimum stress scale for onset of shear-thickening effects has been proposed by Brown/Jaeger; this analysis takes into account the effect of particle size, electrostatic/gravitational forces, surface tension, and osmotic pressure [23, 39] – the result is best viewed graphically, hence, the interested reader is directed to Figure 8 in their article.

Non-Brownian suspensions exhibit several other relevant phenomena during shear flow: particle migration in Poiseuille flow, particle pressure and particle-pressure-induced self-filtration. The work of Matas *et al*. studied the effect of increasing *Re* on particle migration in Poiseuille flow (a phenomenon first noted by *Segre & Silberberg* in 1962); in this work the authors show that the radial equilibrium position of a particle identified by Segre & Silberberg, $r_{equil} = 0.6R$, is influenced by fluid inertia for spherical particles of size $\frac{D_{pipe}}{d_{particle}} = 8 - 42$. The work of Matas *et al*. examined particle distribution along the radial coordinate as a function of distance from entrance length for cases of $Re_p \geq 1$; their findings indicate that as *Re* increases, the particles equilibrium position shifts toward the wall ($r_{equil} \rightarrow R$) for small particles, but for large particles the equilibrium position shifts toward the center of the pipe – comparison of experimental results in a similar system at higher particle volume fractions corroborates the relevance of this phenomenon to flow of more concentrated suspensions in pipes [40, 41, 42].

Deformation of non-Brownian suspensions has been predicted to result in normal stress differences, and experimental investigation thereof has confirmed a manifestation of this normal stress difference: particle pressure [34, 43, 44]. Defined as the mean normal stress exerted by the particle phase in a shear suspension, the particle pressure quantifies the tendency of the particulate phase to spread under shear; it may be expressed as:

$$\Pi = \eta_n^* \eta_0 \dot{\gamma}, \ \eta_n^* \sim \left(\frac{\phi}{\phi_{max}}\right)^2 \left(1 - \frac{\phi}{\phi_{max}}\right)^{-2}$$

The shear-induced particle pressure is of significant relevance to the present work in that it may be used to describe self-filtration effects in concentrated suspensions undergoing a converging flow (at reservoir-suction pipe or at discharge pipe-nozzle) [45]. For suspensions of non-Brownian spheres at $Re_p \rightarrow 0$, $\frac{D_{contraction}}{d_{particle}} = 8 - 11$, it was found that for values of $\left(\frac{\phi}{\phi_{max}}\right)$ greater than ≈ 0.85, $\left(\frac{\phi_{in}}{\phi_{out}}\right)$ decreased steadily to ≈ 0.9 in a $\left(\frac{\phi}{\phi_{max}}\right)$-dependent manner; this self-filtration effect led to flow oscillations at the higher values of $\left(\frac{\phi}{\phi_{max}}\right)$, and could be triggered at lower $\left(\frac{\phi}{\phi_{max}}\right)$ by increasing the shear rate. Performing additional experiments with particles such that $\frac{D_{contraction}}{d_{particle}} = 16 - 24$ and using of the definition of particle pressure, Π, the authors obtain an expression for ϕ_{out} in terms of ϕ_{in}, $\frac{\Pi}{P_{applied}}$, $A = f(\frac{D_{contraction}}{d_{particle}})$ [45]. Moreover, the concept of particle pressure has proven to be instrumental in describing concentrated suspensions in terms of granular rheology [15, 27, 46].

EXTENSIONAL RHEOLOGY OF NON-BROWNIAN SUSPENSIONS

Extension-rate dependence of viscosity has been observed in concentrated Brownian/non-Brownian suspensions, with the incipience of strain-hardening (increased extensional viscosity) and/or viscoelastic effects arising at a critical extension rate that is analogous to the critical shear rate for incipience of shear-thickening [47, 48, 49, 50]. As is observed in shear, the incipience of strain-hardening depends on both the particle volume fraction and the deformation rate [47, 49]; as mentioned in the review of drop formation, viscoelastic effects in filament thinning are quantified using a *Weissenberg* number,

$$Wi = \frac{\dot{\gamma}_{flow}}{\dot{\gamma}_{crit}},$$

which Zimoch *et al.* use to develop a model for the extensional viscosity of a shear-thickening suspension; the aforementioned authors use shear rheology to identify a critical deformation rate corresponding to the incipience of viscosity increase in shear, and analyze the radial decay of suspension filaments to identify the point of viscosity increase in extension [48]. The mode of radial decay described in Zimoch *et al.* corresponds to suspensions that thin as $R_{min} \sim \exp(-t/Wi)$, though this only occurs for suspensions in which $\left(\frac{\phi}{\phi_{max}}\right)$ exceeds a critical value, as identified by

the works of Roche *et al.* and Smith *et al.* This exponential thinning in time is analogous to the qualitative change that occurs in the pendant drop studies [51, 52, 53, 54, 55, 56, 57], thus we observe that the suspension's extensional viscosity and hence radial decay rate are functions of $\left(\frac{\phi}{\phi_{max}}\right)$ and deformation rate ($\dot{\gamma}$).

The magnitude of deformation rate in our drop-on-demand manufacturing process is on the order of 10^3 s^{-1}, thus, shear-thickening effects appear unavoidable at high values of $\left(\frac{\phi}{\phi_{max}}\right)$. Thus, in order to have predictive power for suspensions of unknown powders, it is necessary to characterize a set of suspensions in shear (identifying the critical deformation rate as function of $\left(\frac{\phi}{\phi_{max}}\right)$) and subsequently analyze the radial decay of liquid filaments at the process conditions; this analysis should remain valid so long as $\frac{D_{conduit}}{d_{p,max}}$ is nearly equivalent for a pair of suspension.

The extensional viscosity of a suspension is expected to be:

$$\eta_{E,red} = \frac{\eta_E}{\eta_0} = f\left(\frac{\phi}{\phi_m}, \dot{\epsilon}, \frac{D_{conduit}}{d_{p,max}}\right),$$

in which $\dot{\epsilon}(t) = \left(\frac{\gamma We}{R\rho_{suspension}}\right)^{0.5} \left(\frac{1}{z_{meniscus}}\right) = \frac{U_z(t)}{z}$

In the cylindrical necking region that occurs during radial thinning of a viscoelastic fluid, wherein $R_{min} \sim exp(-t)$, the extension rate is:

$$\dot{\epsilon}(t) = 2\frac{d}{dt}\ln[R(t)],$$

and in this limiting case the extensional viscosity can be estimated as:

$$\eta_E(t) = \frac{\gamma}{\dot{\epsilon}(t)R(t)}.$$

For cases in which the necking region is poorly approximated by a cylinder, knowledge of the evolution minimum radius in time and surface curvature may be used to calculate the extensional viscosity by adding a correction factor to the above equation for η_E [48, 58, 59].

EFFECT OF RHEOLOGICAL PROPERTIES ON DROP VOLUME

For Newtonian fluids, estimates of maximum drop volume may be made by equating the timescales for axial and radial decay of a free fluid filament [60, 61] – filaments below a critical aspect ratio, $\Lambda_c = \frac{L}{2R}$, may be expected to coalesce into a single drop, and filaments exhibiting

$\Lambda_c = \frac{L}{2R}$ may be expected to break up into multiple drops; the maximum drop volume is estimated as $V_{drop} = V_{cyl,max} = 2\pi R^3 \Lambda_c$ For a cylindrical fluid filament of length L and radius R, the radial velocity is estimated from scaling arguments, which discretize modes of decay based the *Ohnesorge* number [17]; the axial velocity is estimated from the velocity of a collapsing fluid sheet [74]. Based on the theoretical prediction and confirmed in experiment: for viscous Newtonian fluids (*Oh>1*), $\Lambda_c = 33\sqrt{2}Oh$, and for inviscid Newtonian fluids, $\Lambda_c \approx 6$ [62]. It would appear that a similar prediction can be made for concentrated suspensions using the radial necking timescale given by Miskin and Jaeger [57]; the prediction contains a fitting parameter (ψ) (from their analysis) that lumps particle volume fraction, particle wetting and contact angle effects – as such, $\Lambda_c = \psi 2\sqrt{2}\sqrt{\frac{R_{nozzle}}{d_{particle}}}$ Thus, searching experimentally for the maximum drop volume (for a given set of particle wetting/contact angle effects – which for hydrophobic carrier fluid and API particles are likely to be very similar) will allow regression of $\psi = f\left(\frac{\phi}{\phi_{max}}\right)$, such that $\Lambda_c = f\left(\frac{\phi}{\phi_{max}}\right)$, for a chosen carrier fluid platform (e.g. triglyceride oils).

DROP FORMATION IN NON-BROWNIAN SUSPENSIONS

Particulate suspensions of micronized API powders are expected to be non-Brownian (non-colloidal), due to minimum particle sizes greater than 1 μm. The non-ideal nature of this real system introduces phenomena, such as sedimentation, shear-driven segregation and shear-stress/rate dependent rheology, which obfuscate *a prior* prediction of flow behavior for a given solid/liquid mixture – this is especially true for the drop formation process in such suspensions, though some experimental investigation of drop formation in ideal, non-Brownian suspensions under limiting conditions has been performed. In addition to the work on non-Brownian suspensions, considerably more studies have been published on drop formation and drop-on-demand manufacture employing colloidal suspensions – this is due in part to constraints on particle sizes imposed by conventional piezoelectric/thermal-actuated drop generators, and also to the interests of the ceramics industry, for whom nanometer-sized particles were most desirable [3, 63]

Regarding drop formation in non-Brownian suspensions, experimental studies have been published in the literature on the break-off of pendant drops and in one case, jetting behavior

(*We→0*) [51-57]. Furbank and Morris (51, 52) studied the effects of non-Brownian particles on break-off of pendant drops and jetting behavior using suspensions of ideal, density-matched spheres (106, 212-250μm diameter) at particle volume fractions up to 0.4. The authors observe that for drops formed in the pendant drop mode, particles lead to differences (c.f. pure Newtonian fluid) in pinch-off structure, drop thread dynamics, drop size and number of satellite drops – the nature and extent of changes are dependent on particle volume fraction. At low particle volume fractions ($\phi \leq 0.1$), pinch-off structure and drop thread dynamics resemble those of the pure fluid; however, capture of individual or groups of particles in the liquid neck leads to formation of larger satellite drops than are observed for the pure fluid. At higher particle fractions ($\phi \geq 0.2$), stages are observed during thinning of the liquid thread, coinciding with a different pinch-off structure; the frequency and volume of satellite drops is substantially altered: satellite drops are generally suppressed, but, when formed, are large. In addition to pendant drop dynamics, Furbank and Morris studied the behavior of suspensions under jetting conditions [51], observing that particles smooth the abrupt nature of this transition and generally lead to jetting at lower *Weber* numbers than for Newtonian fluids.

Following the work of Furbank and Morris, Bonnoit *et al*., Bertrand *et al*., and Lindner *et al*. focused on the influence of volume fraction and particle diameter on pendant drop formation using high speed photography of ideal, density-matched suspensions pinching off from a 4mm nozzle; their works encompassed volume fractions ranging from 0.01-0.55 and monodisperse particle diameters of 20-140μm[53-55]. In Bonniot *et al*., the authors identify that drop formation in non-Brownian suspensions takes place via three regimes: [1]effective fluid, [2]interstitial fluid, and [3]final detachment; the regimes are differentiated via radial thinning dynamics, which correspond to the rheological properties of the bulk suspension, interstitial fluid, and dense suspension (defined in 57). The authors determine that the transition points between regions depend on the particle size and volume fraction, with lower volume fractions and smaller particle sizes more closely resembling the dynamics of a pure fluid of equal viscosity [53, 54]. The work of Lindner *et al.* clarifies the influence of individual particles on drop formation, observing that even single particles result in substantially different thinning dynamics – this causes significant variance in drop formation experiments using suspensions with very low particle concentrations [55].

The dynamics of the final detachment regime were studied by Miskin and Jaeger, and a general scaling law presented for the radial thinning of dense suspensions – the result of their work

is a predictive tool that takes into account the relevance of particle size and nozzle diameter, as modified by an adjustable parameter accounting for contact angle, wetting and volume fraction effects [57]. Most recently, Pan *et al*. performed experiments spanning the entire range of volume fraction and particle diameters (1-500μm) using density-matched monodisperse spheres in order to generate a "regime map" of drop formation from pendant suspension drops [56].

In the pendant drop experiments focused on ideal non-Brownian suspensions, the authors observe similar phenomenological effects: particles decrease the drop break-off time (compared to pure liquid), suspensions undergo break-up in distinct stages (the incipience and duration of which depend on particle size and volume fraction), and satellite drop formation may be suppressed by presence of particles [51-55]; significantly, the authors report that for a given set of experimental conditions, repeatability (liquid intact length, drop volume) is achieved with some difficulty – this is attributed to the random localization of particles within the necking liquid bridge, which varies somewhat for each pendant drop. Additional work by Roche *et al.* identified particle volume fractions at which suspensions behave as: low/high-viscosity Newtonian fluids or with marks of viscoelasticity [47]. The regime map of Pan *et al.* places the behavior of suspensions into three categories: [1]Newtonian fluids, [2]non-Brownian suspensions and [3]shear-thickened fluids. The authors make these classifications by analogy with the radial decay velocity observed in Newtonian fluids[4], and by Pan *et al.* with the radial decay velocities observed in dense suspensions [57], and shear-thickening fluids [47, 48].

In addition to pendant drop dynamics, Furbank and Morris [51] studied the behavior of suspensions under jetting conditions, in which the flow rate from the nozzle is increased beyond creeping flow (*We~0.1*). The authors observe that particles smoothen the transition from dripping to jetting (an abrupt phenomenon in Newtonian fluids) to such an extent that a clear transition cannot be identified [51]. Additionally, the authors also observe that particles accelerate or inhibit breakup as in the pendant drop experiments, but higher particle concentrations appear to inhibit radial necking, thus, particles exert volume fraction-dependent, opposing effects on liquid jets [51]. As jetting behavior significantly alters DOD drop formation, these aberrant phenomena may be problematic for designing a predictable and reproducible jetting-based manufacturing process.

DROP FORMATION IN BROWNIAN (COLLOIDAL) SUSPENSIONS

Though the suspensions here differ somewhat from colloidal suspensions as a result of the disparity in particle size, their similarity merits consideration of colloidal suspension use in DOD inkjets for ceramics manufacture, 3D printing and other applications. An extensive review of DOD inkjet processes by Derby summarized a number of investigations into the effect of colloidal particles (particle size, volume fraction) in situ and compared results to those of Newtonian fluids using dimensionless quantities, which permit evaluation of relevant physical influences across a range of fluid properties and process conditions [3, 5].

The presence of colloidal particles affects the feasibility of DOD drop formation processes, causing disturbances such as clogging and unstable jetting phenomena (inconsistent breakup location, filament swaying/non-straight trajectory), which necessitate reformulation of the ink should a suitable set of process parameters be inaccessible [10, 64, 65, 66, 67]. Reis et al. [67] investigated printing feasibility of fine alumina suspensions (average particle size 0.3µm) with particle volume fractions up to 0.4 and found drop volumes to be well-described by the dimensionless quantities used for purely Newtonian fluids, attributing this to the suspensions' asymptotic flow behavior at the high shear rates present in their DOD setup. Lee *et al*. [64] sought to elucidate the influence of dispersant/particle mass ratio on the clogging behavior of aggregation-prone ZnO nanoparticle (20nm average particle size) suspensions at volume fractions up to 0.2. The authors identified flow-induced aggregation as one source of process disruption and note the necessity of introducing additional criteria in order to describe colloidal suspensions. Hoath et al. [65] explored DOD feasibility for a range of particle sizes (0.8-3.6µm) at volume fraction 0.25 using suspensions of inorganic pigment, polymer resin and solvent. Across the range of particle sizes and resin concentration investigated in the work, jetting behavior was found to be very similar to that of a Newtonian analogue of equivalent shear viscosity. Bienia et al. [66] used image analysis to study the breakup behavior of several colloidal (mean particle size 10nm, 80, 170nm) suspension filaments ejected from a DOD inkjet. In their work, the formulation properties such as particle fraction, surfactant/polymer concentration were found to have a substantial effect on jetting behavior. Non-axisymmetric filaments appear to result from the presence of surfactant/polymer additives, components which are often necessary to produce stable suspensions.

In addition to work focused on colloidal ink formulation involving surface-active dispersants/polymers, investigation of liquid thread pinch-off and drop formation in colloidal suspensions by Wang et al. [10] provided insight into the influence of non-Newtonian rheological effects on DOD drop formation. Wang et al. worked with suspensions of 130nm pigment particles (stabilized with sulfonate groups) in water/glycerol mixtures, up to a particle volume loading of 15%. The rheological response of the inks was quantified in a capillary viscometer allowing measurement at shear rates up to 10^5 s^{-1}, a value which is representative of the shear rates present in conventional DOD inkjets. Wang et al. conclude that while their colloidal inks exhibit jetting behavior similar to that of Newtonian fluids under optimized conditions, the feasible region of drop ejection frequencies is substantially smaller – outside this, jetted filaments exhibiting non-straight trajectories and non-axisymmetric ligaments compromise reproducibility. The rheological response of the colloidal ink was described best by the low shear rate viscosity, which the authors attribute to the effect of total strain on suspension microstructure (total strains in the DOD process were significantly smaller than the capillary viscometer).

REFERENCES

[1] J. Eggers, "Drop formation – an overview," Z. Angew. Math. Mech. **85**(6), 400 (2005).

[2] O.A. Basaran, H. Gao, and P.P. Bhat, "Nonstandard Inkjets,"Annu. Rev. Fluid Mech. **45**, 85 (2013).

[3] B. Derby, "Inkjet Printing of Functional and Structural Materials: Fluid Property, Requirements, Feature Stability, and Resolution," Annu. Rev. Mater. Res. **40**, 395-414 (2010).

[4] J. Eggers and E. Villermaux, "Physics of liquid jets," Rep. Prog. Phys. **71**, 036601 (2008).

[5] H. Wijshoff, "The dynamics of piezo inkjet printhead operation," Phys. Rep. **491**, 77 (2010).

[6] C. Clanet and J.C. Lasheras, J. Fluid Mech. **383,** 307-326 (1999).

[7] B. Ambravaneswaran, H.J. Subramani, S.D. Phillips, and O.A. Basaran, "Dipping-Jetting Transitions in a Dripping Faucet," Phys. Rev. Lett. **93**, 034501

[8] G.H. McKinley and M. Renardy, "Wolfgang von Ohnesorge," Phys. Fluid. **23**, 127101 (2011).

[9] H. Dong, W.W. Carr, and J.F. Morris, "An experimental study of drop-on-demand drop formation," Phys. Fluid **18**, 072102 (2006).

[10] X. Wang, W.W. Carr, D.G. Bucknall, and J.F. Morris, "Drop-on-demand drop formation of colloidal suspensions," Int. J. Mul. Flow **38**, 17 (2012).

[11] B. Derby, "Inkjet printing ceramics: From drops to solid J. Eur. Ceram. Soc. **31**, 2543-2550 (2011).

[12] G.H. McKinley, "Visco-Elasto-Capillary Thinning and Break-Up of Complex Fluids," HML Report Number 05-P-04 (2005).

[13] C. Clasen, P.M. Phillips, L. Palangetic, and Jan Vermant, "Dispensing of Rheologically Complex Fluids: The Map of Misery," AIChE Journal **58**, 3242 (2012).

[14] IUPAC, "Suspension," *Compendium of Chemical Terminology*, 2nd ed. (2014). At goldbook.iupac.org/S06198.html

[15] J.J. Stickel and R.L. Powell, "Fluid Mechanics and Rheology of Dense Suspensions," Annu. Rev. Fluid Mech. **37**, 129 (2005).

[16] M.M. Denn and J.F. Morris, "Rheology of Non-Brownian Suspensions," Annu. Rev. Chem. Biomol. Eng. **5**, 203 (2014).

[17] P. Coussot and C. Ancey, "Rheophysical classification of concentrated suspensions and granular pastes," Phys. Rev. E. **59**, 4445 (1999).

[18] D.M. Kalyon and S. Aktaş, "Factors Affecting the Rheology and Processability of Highly Filled Suspensions," Annu. Rev. Chem. Biomol. Eng. **5**, 229–54 (2014).

[19] Bonn, D., Denn, M. M., Berthier, L., Divoux, T., & Manneville, S. (2017). Yield stress materials in soft condensed matter. *Reviews of Modern Physics*, *89*(3), 035005.

[20] D.L. Koch and R.J. Hill, "Inertial Effects in Suspension and Porous-Media Flows," Annu. Rev. Fluid Mech. 2001 **33**, 619 (2001).

[21] C. Tropea, A.L. Yarin, and J.F. Foss, *Handbook of Experimental Fluid Mechanics* (Springer-Verlag, Würzburg 2007), p. 680-710.

[22] I.E. Zarraga, D.A. Hill, and D.T. Leighton, "The characterization of the total stress of concentrated suspensions of noncolloidal spheres in Newtonian fluids," J. Rheol. **44**, 185 (2000).

[23] E. Brown and H.M. Jaeger, "Shear thickening in concentrated suspensions: phenomenology, mechanisms and relations to jamming," Rep. Prog. Phys. **77**, 046602 (2014).

[24] A.B. Metzner, "Rheology of Suspensions in Polymeric Liquids," J. Rheo **29**, 739 (1985).

[25] J.F. Morris and F. Boulay, "Curvilinear flows of noncolloidal suspensions: The role of normal stresses," J. Rheo. **43** 1213 (1999).

[26] C. Cassar, M. Nicolas, and O. Pouliquen, "Submarine granular flows down inclined planes," Phys. Fluid. **17**, 103301 (2005).

[27] F. Boyer, E. Guazzelli, and O. Pouliquen, "Unifying Suspension and Granular Rheology," Phys. Rev. Lett. **107**, 188301 (2011).

[28] S. Mueller, E.W. Llewellin, and H.M. Mader, "The rheology of suspensions of solid particles," Proc. R. Soc. A. **466**, 1201 (2010).

[29] S. Mueller, E.W. Llewellin, and H.M. Mader "The effect of particle shape on suspension viscosity and implications for magmatic flows," Geophys. Res. Lett. **38**, L13316 (2011).

[30] C. Cimarelli, A. Costa, S. Mueller and H.M. Mader, "Rheology of magmas with bimodal crystal size and shape distributions: Insights from analog experiments," Geochem., Geophys., Geosys. **12**, 7 (2011).

[31] E. Brown, H. Zhang, N.A. Forman, B.W. Maynor, D.E. Betts, J.M. DeSimone, and H.M. Jaeger, "Shear thickening in densely packed suspensions of spheres and rods confined to few layers," J. Rheo **54**, 1023 (2010).

[32] P. Moitra, and H.M. Gonnermann, "Effect of crystal shape- and size-modality on magma rheology," Geochem., Geophys., Geosys. **16**, 1 (2015).

[33] J.P. Latham, A. Munjiza, and Y. Lu, "On the prediction of void porosity and packing of rock particulates," Powder Tech. **125**, 10 (2002).

[34] J.F. Morris, "A review of microstructure in concentrated suspensions and its implications for rheology and bulk flow," Rheol. Acta **48**, 909 (2009).

[35] E. Bertrand, J. Bibette, and V. Schmitt, "From shear thickening to shear-induced jamming," Phys. Rev. E. **66**, 060401 (2002).

[36] A. Fall, F. Bertrand, G. Ovarlez, and D. Bonn, "Yield Stress and Shear Banding in Granular Suspensions," Phys. Rev. Lett. 178301 (2009).

[37] E. Brown, N.A. Forman, C.S. Orellana, H. Zhang, B.W. Maynor, D.E. Betts, J.M. DeSimone, and H.M. Jaeger, "Generality of shear thickening in dense suspensions," Nat. Mat. **9**, 220 (2010).

[38] O. Reynolds, "On the dilatancy of rigid particles in contact. With experimental illustrations," Phil. Mag. Ser. **5**, 469 (1885).

[39] E. Brown and H.M. Jaeger, "The role of dilation and confining stresses in shear thickening of dense suspensions," J. Rheo. **56**, 875 (2012).

[40] J.P. Matas, J.F. Morris, and E. Guazzelli, "Inertial migration of rigid spherical particles in Poiseuille flow," J. Fluid Mech. **515**, 171 (2004).

[41] M. Kim and S. Lee, "Particle migration in tube flow of suspensions," J. Rheo. **43**, 1157 (1999).

[42] Segre, G., & Silberberg, A. (1962). Behaviour of macroscopic rigid spheres in Poiseuille flow Part 2. Experimental results and interpretation. *Journal of fluid mechanics*, *14*(1), 136-157.

[43] Y. Yurkovetsky and J.F. Morris, "Particle pressure in sheared Brownian suspensions," J. Rheo. **52**, 141 (2008).

[44] A. Deboeuf, G. Gauthier, J. Martin, Y. Yurkovetsky, and J.F. Morris, "Particle Pressure in a Sheared Suspension: A Bridge from Osmosis to Granular Dilatancy," Phys. Rev. Lett. **102**, 108301 (2009).

[45] S.D. Kulkarni, B. Metzger, and J.F. Morris, "Particle-pressure-induced self-filtration in concentrated suspensions," Phys. Rev. E. **82**, 010402 (2010).

[46] P. Mills and P. Snabre, "Apparent viscosity and particle pressure of a concentrated suspension of non-Brownian hard spheres near the jamming transition," Eur. Phys. J. E. **30**, 309 (2009).

[47] M. Roche, H. Kellay, and H.A. Stone, "Heterogeneity and the Role of Normal Stresses during the Extensional Thinning of Non-Brownian Shear-Thickening Fluids," Phys. Rev. Lett. **107**, 134503 (2011).

[48] P.J. Zimoch, G.H. McKinley, and A.E. Hosoi, "Capillary Breakup of Discontinuously Rate Thickeing Suspensions," Phys. Rev. Lett. **111**, 036001 (2013).

[49] M.I. Smith, R. Besseling, M.E. Cates and V. Bertola, "Dilatancy in the flow and fracture of stretched colloidal suspensions," Nat. Comm. **1**, 114 (2010).

[50] E.E. Bischoff White, M. Chellamuthu, and J.P. Rothstein, "Extensional rheology of a shear-thickening cornstarch and water suspension," Rheol. Acta **49**, 119 (2010).

[51] R.J. Furbank and J.F. Morris, "An experimental study of particle effects on drop formation," Phys. Fluid. **16**, 1777 (2004).

[52] R.J. Furbank and J.F. Morris, "Pendant drop thread dynamics of particle-laden liquids," I. J. Mul. Flow **33**, 448 (2007).

[53] T. Bertrand, C. Bonnoit, E. Clement, and A. Lindner, "Dynamics of drop formation in granular suspensions: the role of volume fraction," Gran. Matt. **14**, 169 (2012).

[54] C. Bonnoit, T. Bertrand, E. Clement, and A. Lindner, "Accelerated drop detachment in granular suspesnsions," Phys. Fluid. **24**, 043304 (2012).

[55] A. Lindner, J.E. Fiscina, and C. Wagner, "Single particles accelerate final stages of capillary break-up," Eur. Phys. Lett. **110**, 64002 (2015).

[56] Z. Pan, N. Louvet, Y. Hennequin, H. Kellay, and D. Bonn, "Drop formationin shear-thickeing granular suspensions," Phys. Rev. E. **92**, 052203 (2015).

[57] M.Z. Miskin and H.M. Jaeger, "Droplet formation and scaling in dense suspensions," Proc. Nat. Acad. Sci. **109**, 4389 (2012).

[58] A.M. Ardekani, V. Sharma, and G.H. McKinley, "Dynamics of bead formation, filament thinning and breakup in weakly viscoelastic jets," J. Fluid Mech. **665**, 46 (2010).

[59] P.P. Bhat, S. Appathurai, M.T. Harris, and O.A. Basaran, "On self-similarity in the drop-filament corner region formed during pinch-off of viscoelastic fluid threads," Phys. Fluid. **24**, 083101 (2012).

[60] A.A. Castrejon-Pita, J.R. Castrejon-Pita, and I.M. Hutchings, "Breakup of Liquid Filaments," Phys. Rev. Lett. **108**, 074506 (2012).

[61] S.D. Hoath, S. Jung and I.M. Hutchings, "A simple criterion for filament break-up in drop-on-demand inkjet printing," Phys. Fluid. **25**, 021701 (2013).

[62] G. Taylor, "The dynamics of thin sheets of fluid III. Disintegration of fluid sheets," Proc. R. Soc. A. **253**, 313 (1959).

[63] B. Derby, "Inkjet printing ceramics: From drops to solid J. Eur. Ceram. Soc. **31**, 2543-2550 (2011).

[64] A. Lee, K. Sudau, K. Hyun Ahn, S. Jong Lee, and N. Willenbacher, "Optimization of Experimental Parameters to Suppress Nozzle Clogging in Inkjet Printing," Ind. Eng. Chem. Res. **51**, 13195-13204 (2012).

[65] S.D. Hoath, W.-K. Hsiao, I.M. Hutchings and T.R. Tuladhar, "Jetted mixtures of particle suspensions and resins," Phys. Fluids **26**, 101701 (2014).

[66] M. Bienia, M. Lejeune, M. Chambon, V. Baco-Carles, C. Dossou-Yovo, R. Noguera, and F. Rossignol, "Inkjet printing of ceramic colloidal suspensions: Filament growth and breakup," Chem. Eng. Sci. **149**, 1-13 (2016).

[67] Reis, N., Ainsley, C., & Derby, B. 2005. Ink-jet delivery of particle suspensions by piezoelectric droplet ejectors. Journal of Applied Physics, 97(9), 094903.

APPENDIX B. CHAPTER 3 SUPPLEMENTARY MATERIAL

**SUPPLEMENTARY MATERIAL INCLUDED WITH PUBLICATION**
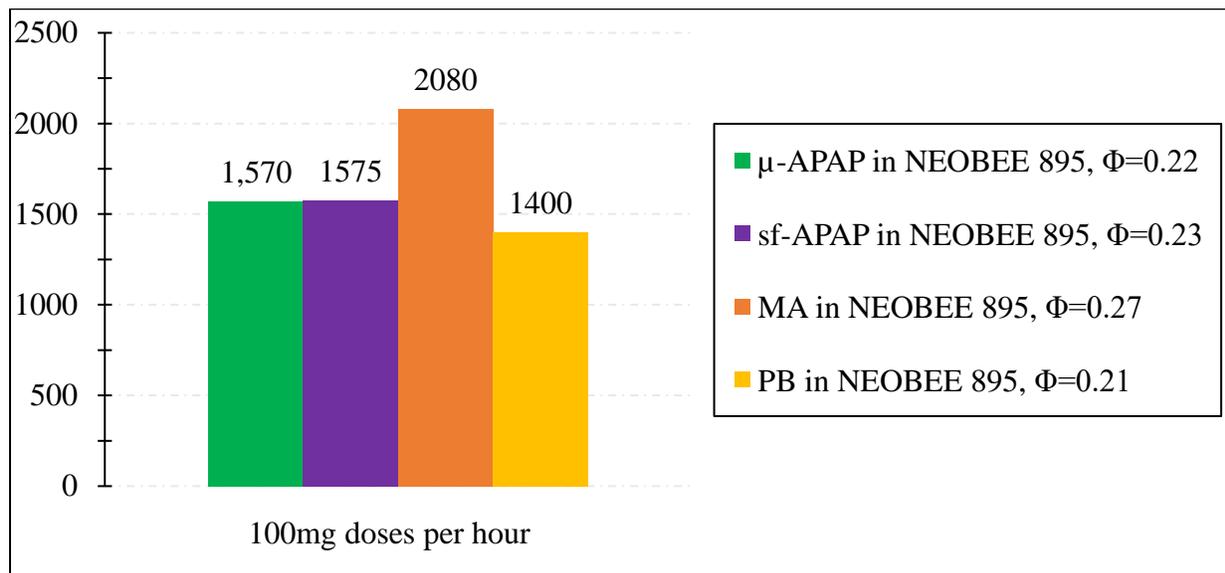
Figure B.1: Production (capsule-filling) rate of system for 100mg doses. Estimate is based on mg API/sec, given operation at the maximum drop frequency (~15 drops per sec) and drop volumes of ~9μL [Radcliffe & Reklaitis 2017].

Table B.1: Particle size parameters, from volume-weighted data, for (printed) capsule contents prepared with same treatment as in Table 3.1. Particle aspect ratio parameters (Mean, Stdev) are based on number-weighted data.

	μ-APAP	sf-APAP	Mefenamic acid	Phenylbutazone
Φ (in dose)	0.2	0.25	0.3	0.3
D[3,2] (μm)	34.9	58.7	86.9	89.9
D[4,3] (μm)	45.8	97.3	123.8	120.8
d_{10} (μm)	21.1	31.0	49.2	53.22
d_{50} (μm)	43.2	77.2	115.7	121.8
d_{90} (μm)	72.2	219.9	194.1	187.5
Mean AR	0.551	0.475	0.681	0.363
Stdev AR	0.198	0.218	0.156	0.240
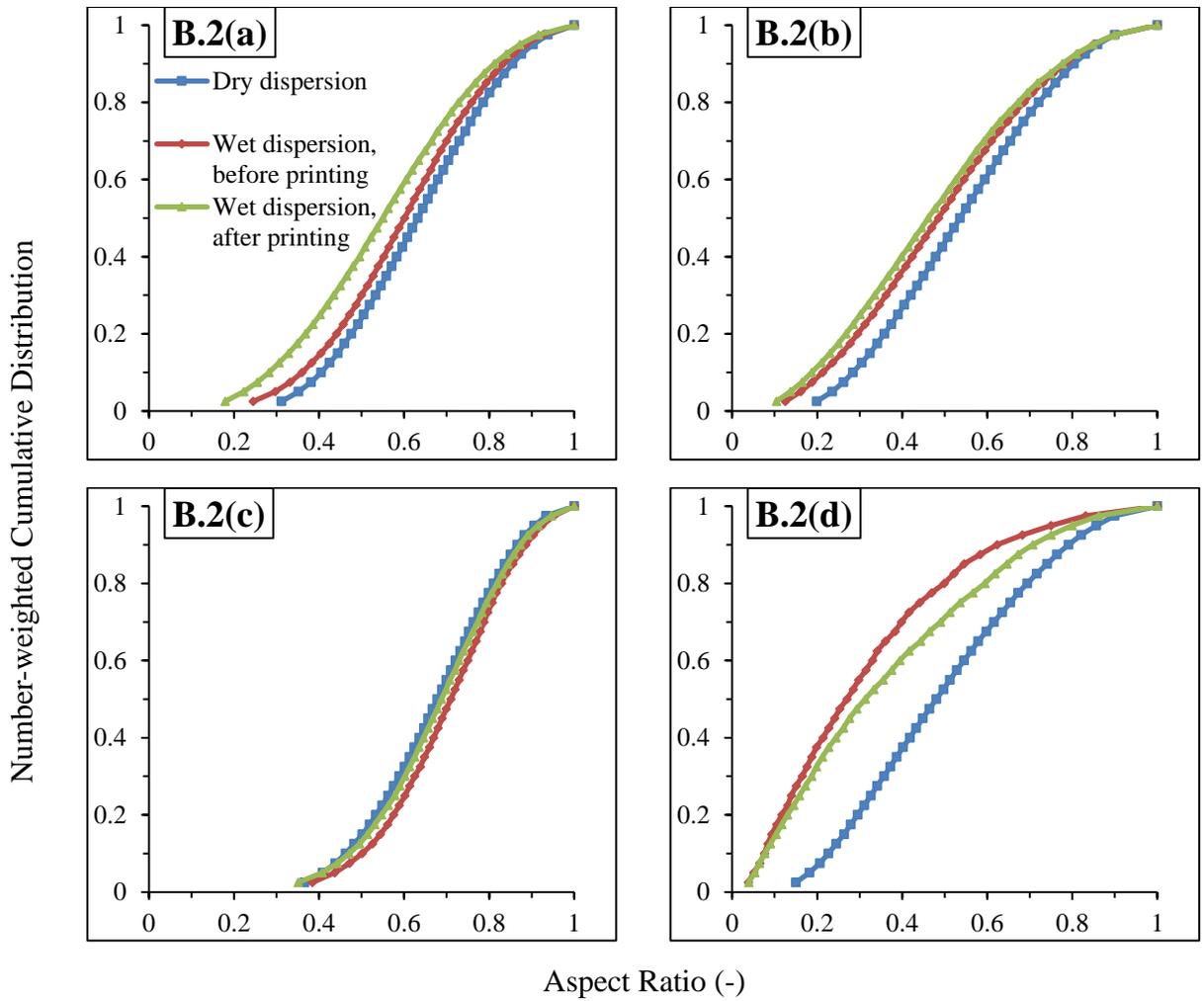
Figure B.2: Aspect ratio in capsules containing: $\phi = 0.2$ micronized acetaminophen, $\phi = 0.25$ semi-fine acetaminophen, $\phi = 0.3$ mefenamic acid, and $\phi = 0.3$ phenylbutazone, denoted as B.2(a)-(d), respectively.

ADDITIONAL FIGURES UPDATED WITH NEW INFORMATION

■ µ-APAP, NEOBEE 895	● sf-APAP, NEOBEE 895
◆ MA, NEOBEE 895	▲ PB, NEOBEE 895
□ µ-APAP, HMDSO + R 972	○ sf-APAP, HMDSO + R 972
◇ MA, HMDSO + R 972	△ PB, HMDSO + R 972
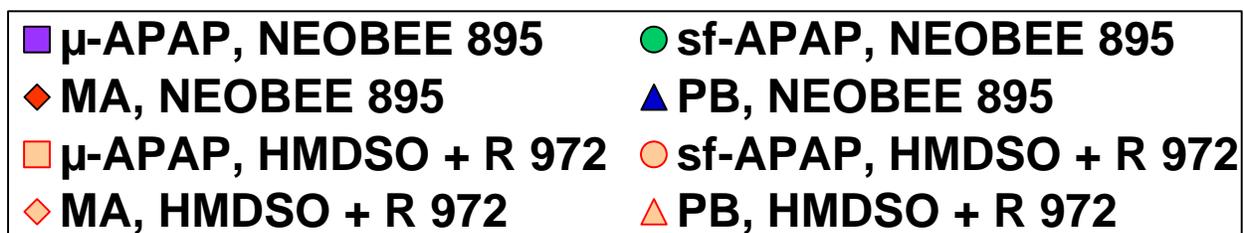
Legend for Figures B.3 & B.4
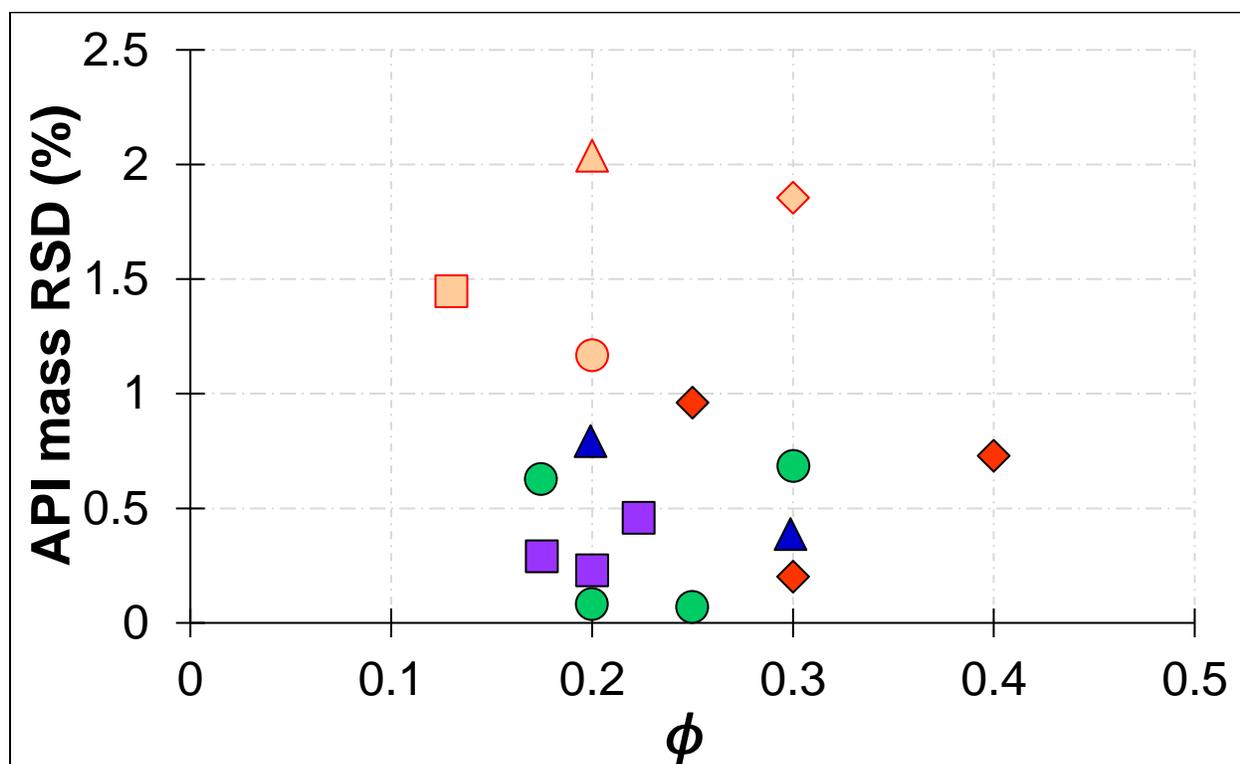
Figure B.3. Identical information to Figure 3.7., but with re-colored markers
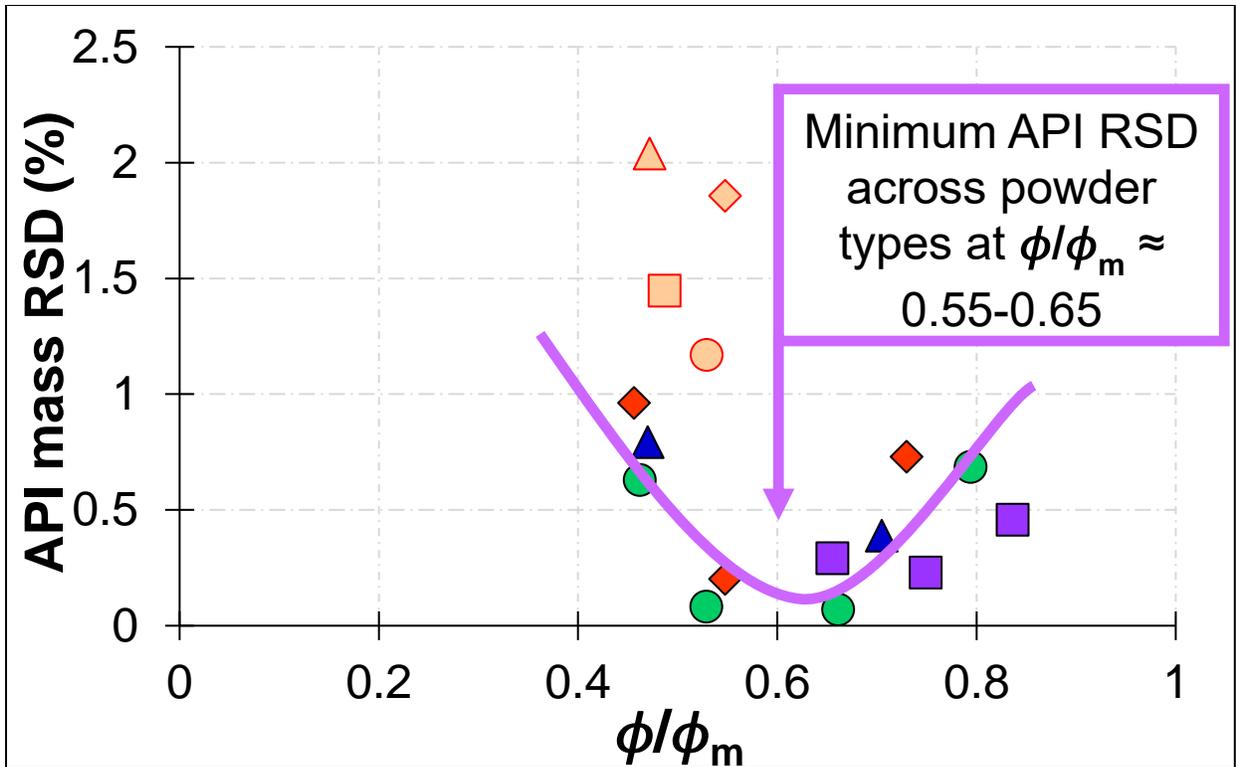
Figure B.4. Identical information to Figure 3.8. Re-colored markers, line to guide eye and comment used to clarify trend discussed in paper.
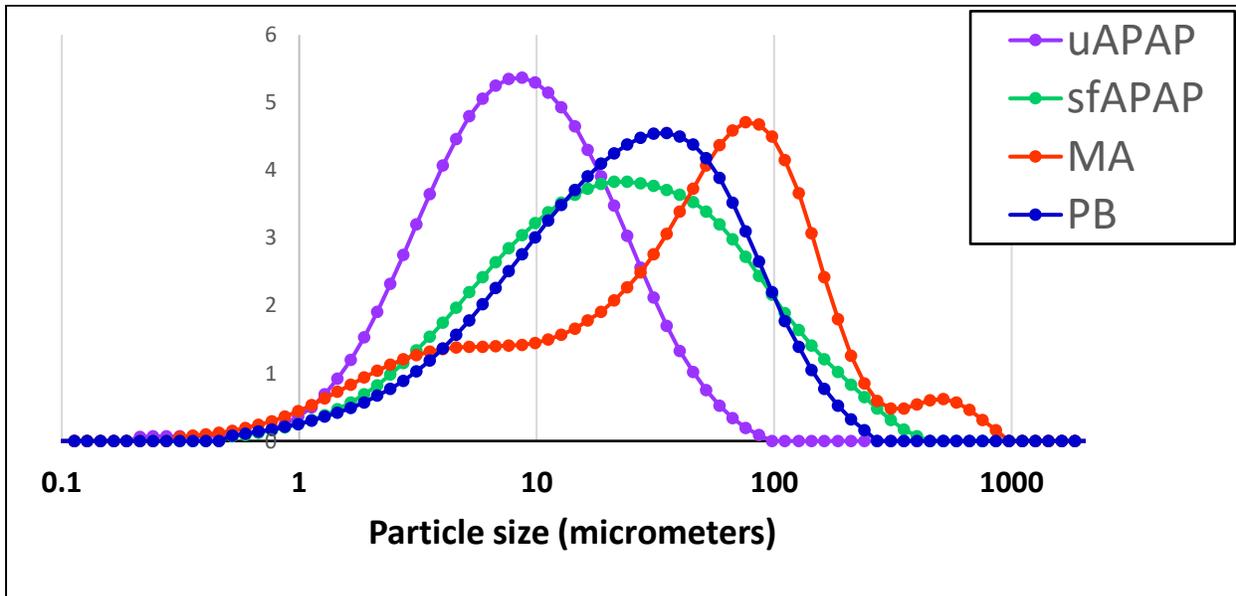
Figure B.5. Particle size measured by laser diffraction method (Malvern Mastersizer 3000)

APPENDIX C. CHAPTER 4 NEW MATERIAL

The following results were obtained from the high-speed imaging data upon analysis with the computer vision codes described in Chapter 5. The results are included in the appendix of Chapter 4 as they relate more closely to its themes and material. The updated regime maps are the result of applying an automated algorithm to extract the one-dimensional velocity information rather than the previous method, which relied on edge detection, but with manually determined time intervals.

OPERATING CONDITIONS: REGIME MAP
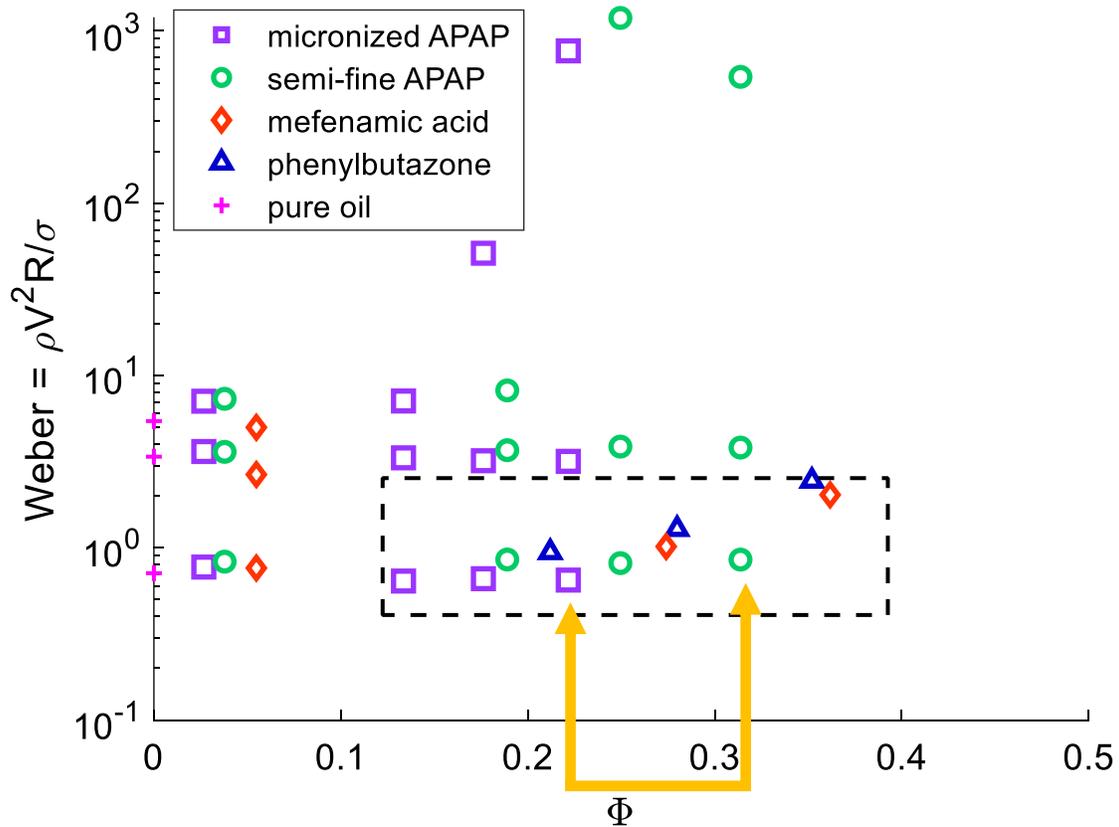
Figure C.1. New version which is equivalent in concept to Figure 4.1.
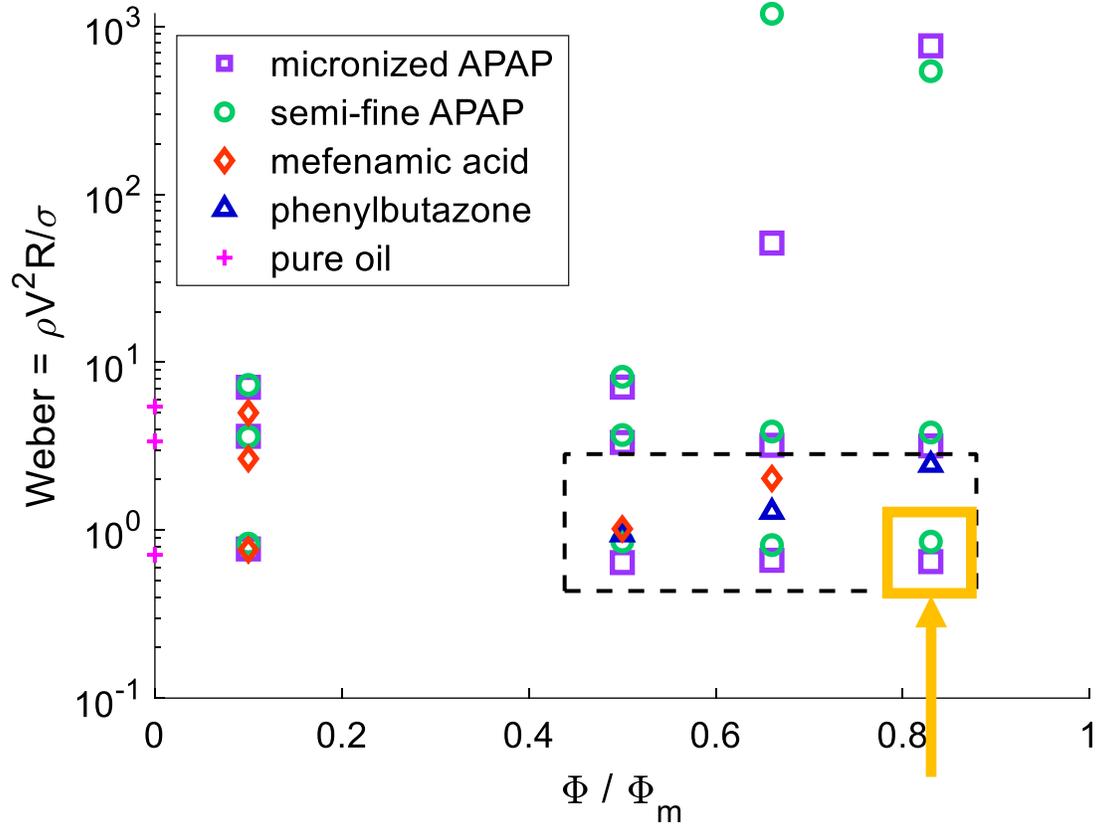
Figure C.2. New version which is equivalent in concept to Figure 4.2.

It is apparent that the old method was inaccurate, as one notices some interesting new trends; note that the placement of the yellow arrows and boxes denote the case of interest examined in the figures below.
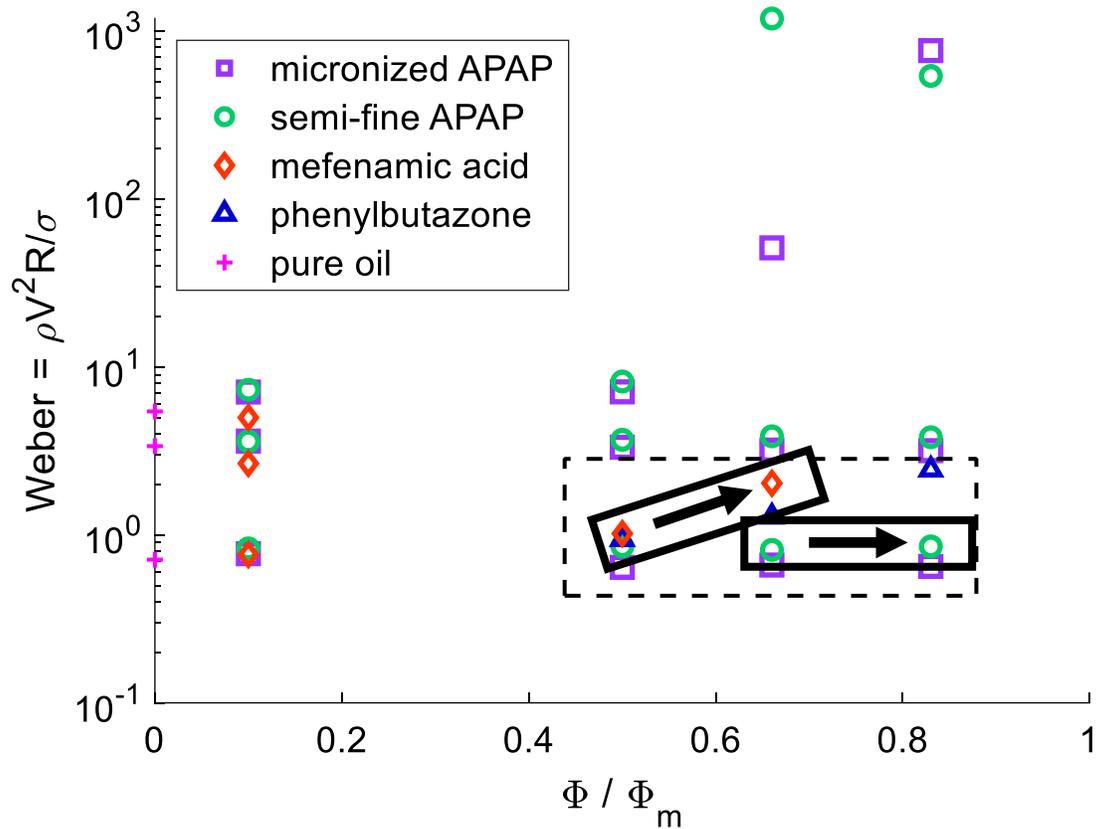
Figure C.3. Identical to Figure C.2., but with boxed arrows to draw attention to the specific cases.
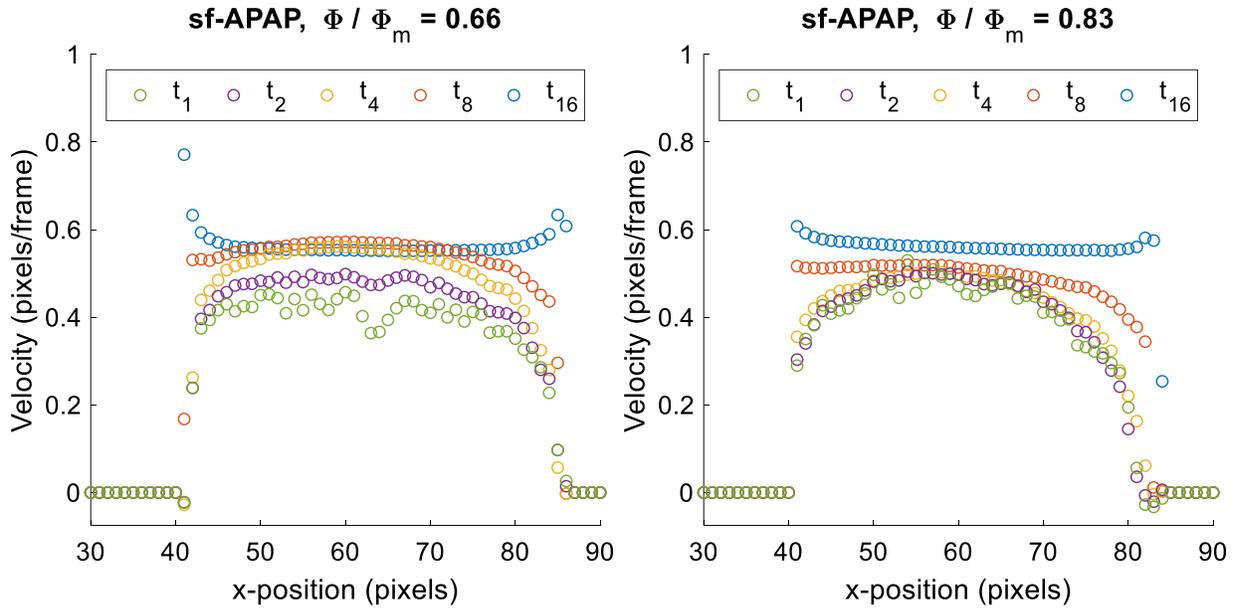
Figure C.4. Dynamic velocity profile during printing of semi-fine acetaminophen suspensions at the boxed conditions in Figure C.3.; increasing values of the subscript on *t* denote later time points; obtained by linear regression of the object trajectories, as described in detail in the computer vision sections.
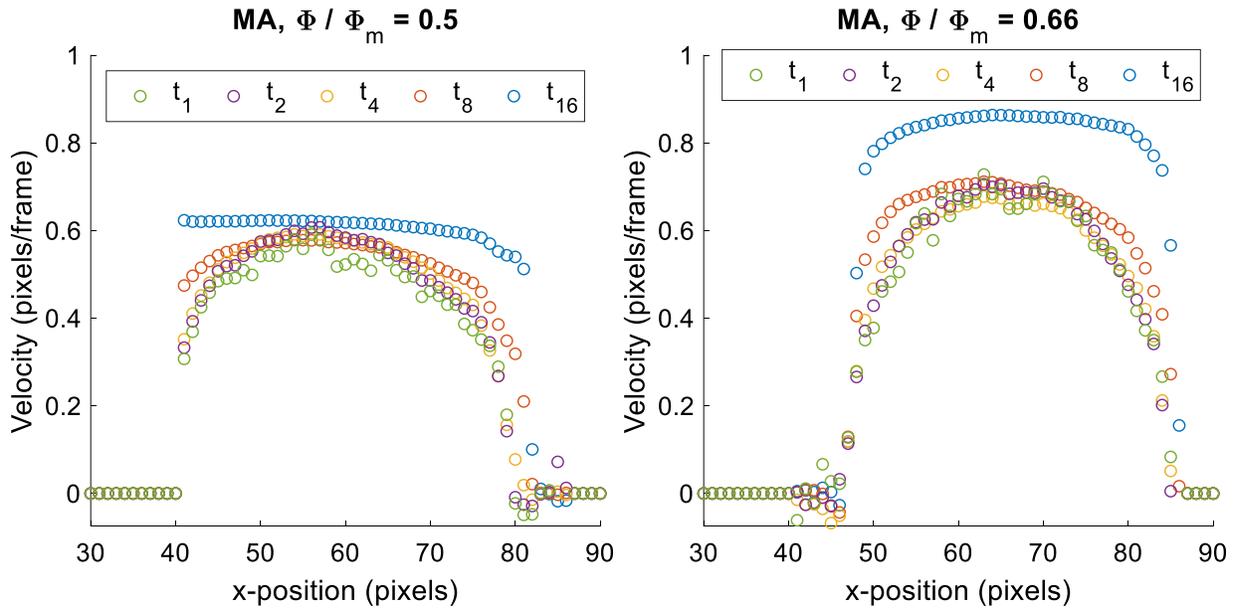
Figure C.5. Dynamic velocity profile during printing of mefenamic acid suspensions at the boxed conditions in Figure C.3.; increasing values of the subscript on *t* denote later time points; obtained by linear regression of the object trajectories, as described in detail in the computer vision sections.

DISCUSSION

If we take a look at the points denoted by the green marker at 0.66 and 0.83, drop ejection is similar, despite approaching the critical particle loading. Upon examination of the dynamic velocity profiles at the respective conditions, one observes that they are approximately Newtonian at earlier time points just after the fluid leaves the nozzle, then later transitioning to plug-flow, as is expected when shear forces are absent on the outer surface.

We can also take a look at a more interesting case: the points denoted by the red markers at 0.5 and 0.66. Note that in fact, the lack of a point at 0.83 is intended – it is not possible to print drops under such conditions. What is the reason for this? The onset of shear-thickening effects. One can infer this by examining the velocity profiles. At the lower loading, the profile is still approximately Newtonian. However, at the higher loading, the curves show a marked increase in velocity near the centerline; this transition is indicative of a power-law rheology associated with shear-thickening.

VELOCITY PROFILES

Further examination of the velocity profiles for printing of drops at *Weber* number between ≈ 0.6-2 provides additional evidence for the rheological transitions. In the figures below, the progression for each particle type is presented.
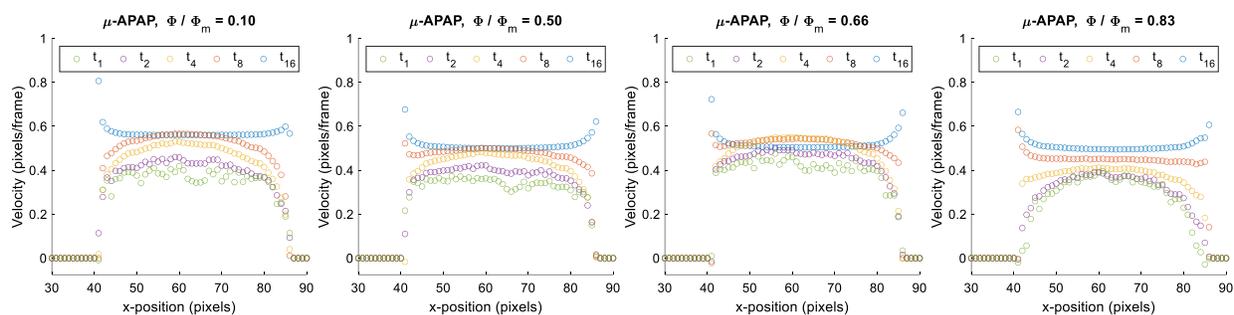

Figure C.6. Dynamic velocity profiles during printing of suspensions of micronized acetaminophen in NEOBEE 895
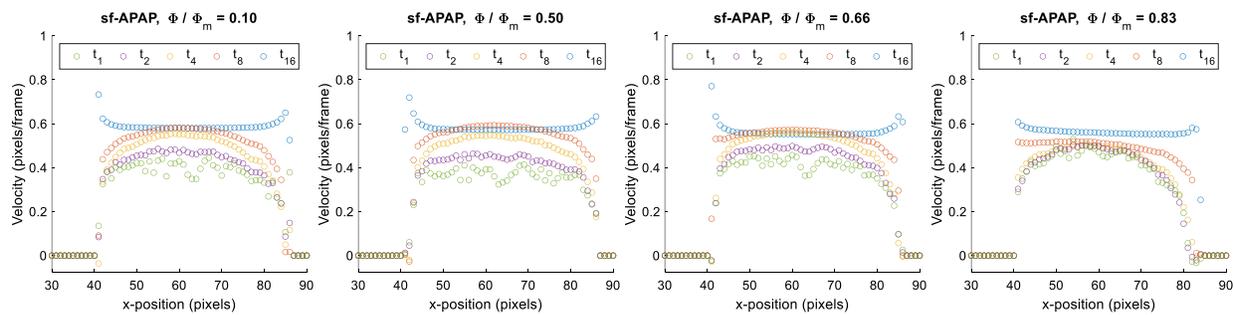

Figure C.7. Dynamic velocity profiles during printing of suspensions of semi-fine acetaminophen in NEOBEE 895
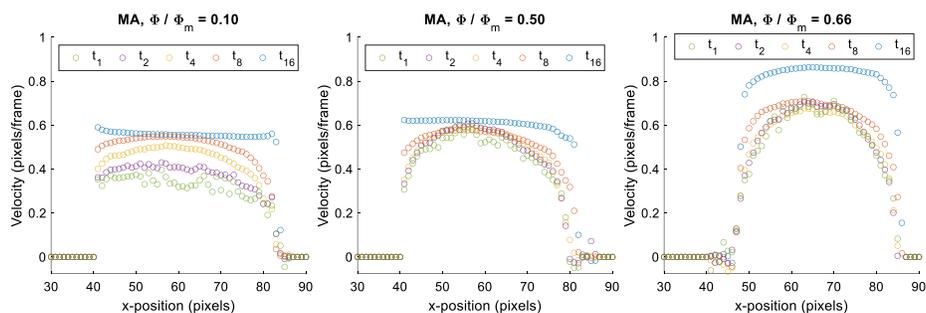
Figure C.8. Dynamic velocity profiles during printing of suspensions of mefenamic acid in NEOBEE 895
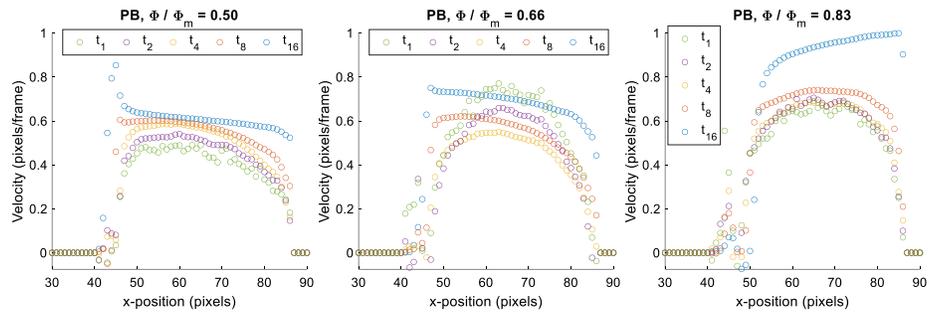
Figure C.9. Dynamic velocity profiles during printing of suspensions of phenylbutazone in NEOBEE 895
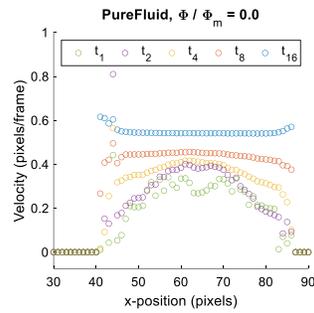

Figure C.10. Dynamic velocity profiles during printing of the pure carrier fluid, NEOBEE 895

APPENDIX D. CHAPTER 5 RESOURCE 1

IMAGE ANALYSIS: EDGE DETECTION, FILTERING, AND THRESHOLDING

To reduce the data to the salient information, and to facilitate the subsequent analysis algorithms (Sections 3 and 4), each constituent image of each video sequence was processed by means of an edge detection algorithm; the overall information flow is shown in Figure 4 in the manuscript. In this context, edge detection is a pre-processing method which, with the correct design, automatically adapts to the characteristics (e.g. noise) of the imaging setup and therefore can be applied to entire image sequences; it produces a binary map of pixels at the feature edges within the image. An edge in an image may be visualized as the locations on the three-dimensional surface with steep slopes, and can be defined more rigorously using the magnitude of the gradient, noting that the edge orientation at a given point is orthogonal to the direction of the gradient. From this consideration and others, many edge detectors have been developed; the interested reader is referred to the references for details (Canny 1986; Elder and Zucker 1998; Lindeberg 1998; McIlhagga 2011). For the purpose in this work, the widely used Canny edge detector was selected as it was designed as an optimal detector which maximizes signal-to-noise ratio and accurate localization of edges (Canny 1986).

Within the edge detection algorithm there are several steps, with specific modifications for this context that enable application to any image set (temporal sequence or single image) which resembles the physical situation here: objects – in this case, fluid volumes – entering (and leaving) a region of interest in which they undergo deformation, fragmentation and coalescence. The modest quality of the imaging setup used in this work demonstrates that the method is robust even in environments with substantial noise. A description of the Canny method and necessary components for the implementation is presented below. The steps of the detector are presented in Figure D.1.

EDGE DETECTION

For each input image, the Canny edge detection algorithm consists of the following steps: 1) Gaussian filtering and calculation of horizontal and vertical gradients for each pixel by

convolving with Gaussian and derivative of Gaussian filters. 2) Computation of gradient magnitude and direction at each pixel. 3) Non-maximum suppression (NMS) to thin the edges of the gradient magnitude image. In this step the direction of the gradient is used to identify with which two immediate neighbors the magnitude of the gradient should be compared: if the magnitude at the location is greater than its two neighbors along the direction of the gradient, the value is retained, otherwise it is suppressed (set equal to zero). 4) High/low thresholds are calculated from the histogram of the gradient magnitude. Here, rather than set the high threshold as a fixed percentage of the histogram as in the classic Canny method, an adaptive threshold is computed for each image using the Otsu method, which determines the value that maximizes the interclass variance. 5) Hysteresis thresholding to determine the edge map. For each pixel, if the gradient magnitude is greater than the high threshold, it is labeled as a strong edge; if the gradient magnitude is greater than the low threshold but less than the high threshold, it is labeled as a weak edge. Strong edges are automatically included in the edge map, but weak edges are included if and only if they are connected to a strong edge. Additional details on each of the steps are given below in order to provide a brief overview.
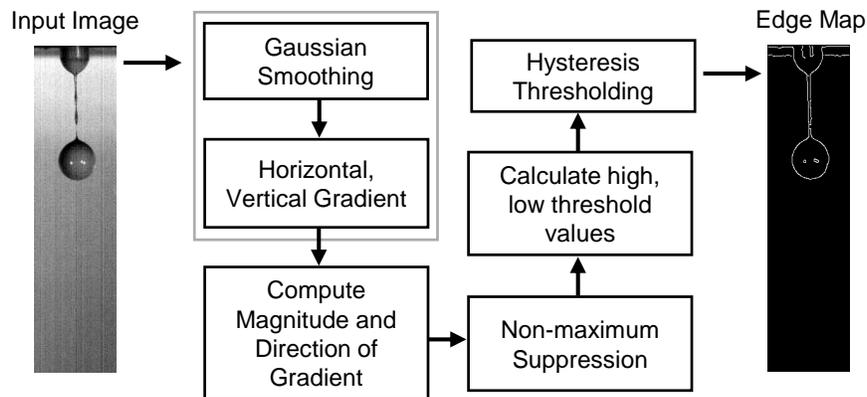

Figure D.1. Flow chart of Canny edge detector

GAUSSING SMOOTHING AND IMAGE GRADIENT

As the Canny detector is based on the first derivatives of the image, the initial step consists of smoothing (to reduce noise induced by differentiation) by convolution with a separable two-dimensional Gaussian filter, $G(x, y) = g^{(1)}(x) * g^{(2)}(y)$, in which * denotes the convolution

operation, followed by convolution with the first derivatives of the Gaussian to yield the horizontal and vertical image gradients. As extensively documented in the literature (Canny 1986; Elder and Zucker 1998; Xu et al. 2014), a significant improvement in computational efficiency can be realized by separating the two-dimensional Gaussian into two one-dimensional Gaussians and using the properties of convolution to compute the image gradient as follows:

$$\nabla(I * G) = \begin{bmatrix} \dfrac{\partial(I * G)}{\partial x} & \dfrac{\partial(I * G)}{\partial y} \end{bmatrix}^T = \begin{bmatrix} I * \dfrac{\partial G}{\partial x} \\ I * \dfrac{\partial G}{\partial y} \end{bmatrix} = \begin{bmatrix} \left(I * g^{(2)}(y)\right) * \dfrac{dg^{(1)}(x)}{dx} \\ \left(I * g^{(1)}(x)\right) * \dfrac{dg^{(2)}(y)}{dy} \end{bmatrix} = \begin{bmatrix} M_x(i,j) \\ M_y(i,j) \end{bmatrix}$$

For a discrete Gaussian kernel of width K, this reduces the number of multiply-add operations per pixel in each gradient mask to 2K, rather than K^2.

The size (K x K) of the Gaussian kernel and its horizontal/vertical derivatives are determined by the spatial scale parameter, σ, which may be selected by the user based on the level of noise in the image, size of objects, and/or the desired features (Canny 1986; Elder and Zucker 1998), as demonstrated in the original publication on the Canny detector which used multiple scales to synthesize a composite edge map from different image features. Alternately, a scale-space approach may be necessary to systematically detect edges at different resolutions (Elder and Zucker 1998; Lindeberg 1998). Although methods may be used to select a minimum reliable scale based on models of sensor noise, or for a local region, this leads to different edge detectors, and therefore may be considered for a future implementation (Elder and Zucker 1998; Lindeberg 1998; Bao et al. 2005; McIlhagga 2011).

As the size and location of objects in the image are expected to vary, a global scale parameter was selected heuristically based on comparison of the edge features obtained by the Canny detector for values of the scale parameter, σ = 0.4-1.4; the corresponding discrete one-dimensional kernels were obtained by sampling from the continuous Gaussian at the points [-n, …, 0, …, n], in which (2n)+1 = K, and K = ceil(6σ). In general, smaller filters are more sensitive to fine edges, whereas larger (σ) filters increased detector robustness in the presence of noise, which was a necessary consideration due to the quality of the images; consequently, a value of σ = 1.4 (kernel size of K = 9) was used for all images as this preserved the salient features. The edge maps obtained with the Canny detector using different values of σ are presented in Figure D.2; the high/low hysteresis thresholds obtained by the Otsu method for the σ = 1.4 (image 4) gradient

mask were applied in all four edge – if the respective Otsu thresholds for each gradient mask in images 1, 2, & 3 (of Figure D.2) are used instead, the results are more or less the same.
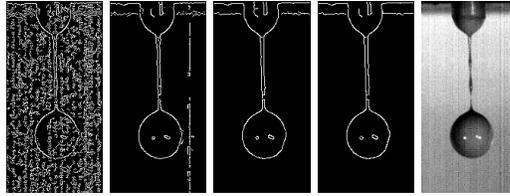

Figure D.2. Effect of scale parameter (kernel width) on edge map obtained by Canny detector. Left to right: scale parameter σ = 0.35, 0.71, 1.2, 1.4; input image.

GRADIENT MAGNITUDE AND DIRECTION

Convolution of the gradient masks with the image results in two matrices that form the components of the gradient vector in the x and y directions for each pixel location, (i,j). The next steps require the magnitude of the gradient at each pixel, which is calculated as the Euclidean norm:

$$M(i,j) = \sqrt{\left(\nabla_x (I * G)(i,j)\right)^2 + \left(\nabla_y (I * G)(i,j)\right)^2}$$

As suggested in the literature, it is not necessary to explicitly compute the angle of the gradient for each pixel, as the direction may be inferred from the sign/value of the components of the gradient vector at the point of interest (Xu et al. 2014; Lee et al. 2018).

NON-MAXIMUM SUPPRESSION (NMS)

In this step, an edge map is constructed from the gradient magnitude image by suppressing all but the local maxima in edge strength along the direction orthogonal to the edge. At each pixel location the direction of the gradient is used to identify with which of its eight neighbors the current magnitude should be compared; if the magnitude at the point (i,j) is greater than the magnitudes of the two neighbors in the positive/negative gradient directions, then the location is retained as a local maximum, otherwise it is suppressed (set to zero) in the NMS edge map.

The gradient may take one of 8 main directions (angles 0°, 45°, …, 315°), though in virtually all cases, it will not correspond exactly to a main direction, therefore an interpolation is required. Alternatively, the direction may be rounded to the nearest main direction in order to

255

simplify the comparisons at a small (5-15%) loss in accuracy (Lee et al. 2018). To preserve accuracy, the magnitudes of the neighboring pixels along the gradient direction are interpolated by weighting with the ratio of components of the gradient vector respective to the current pixel, $r = M_y(i,j)/M_x(i,j)$. As example, consider the gradient in Figure D.3, which has a direction between 0-45° at the current pixel, therefore the interpolated magnitude in the positive direction of the comparison would be $r\big(M(i+1, j+1)\big) + (1-r)\big(M(i+1,j)\big)$. By performing this for all pixels in the gradient magnitude image, a binary map which consists of only the local maxima is produced; it must be thinned further as it contains non-edges in addition to weak and strong edges.
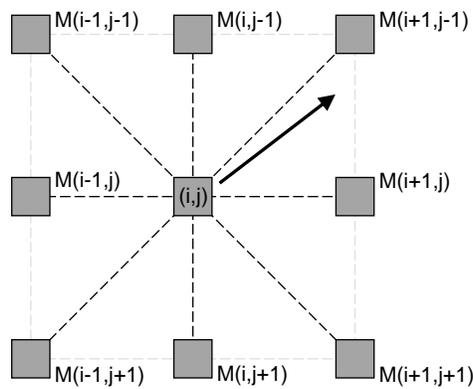

Figure D.3. Each pixel location (i,j) is evaluated using the magnitudes of its eight neighbors, with the comparison determined by the gradient direction.

HIGH/LOW THRESHOLDS AND HYSTERESIS THRESHOLDING

The next step is to determine the threshold values which will be used to thin the edge map obtained by NMS. The low and high thresholds govern the classification of edges as "weak" or "strong", respectively, and are applied to eliminate all local maxima below the low threshold, then to eliminate all weak edges which lack a strong edge neighbor based on 8-connectivity. The choice of threshold values is critically important, as demonstrated in Figure D.4 which compares the resultant edge maps for several possible thresholds.
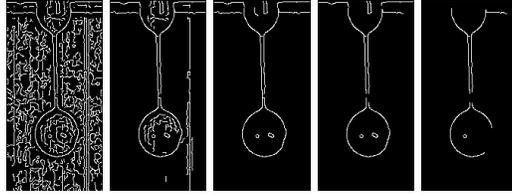
Figure D.4. Edge maps obtained by Canny detector with different threshold values. From left to right, [low, high] threshold values: [0.025, 0.05], [0.05, 0.1], [0.1, 0.2] – Otsu method, [0.2, 0.4], [0.25, 0.5]

Threshold values can be set by the user based on the desired features from an image, but this is unworkable for our purposes as each image demands its own threshold value due to the dynamically changing noise, and the inherent differences between image sets (e.g. clear liquid, or optically opaque). A number of systematic methods have been proposed in the literature for adaptive selection of threshold values based on local or global image statistics (Sezgin and Sankur 2004). One of the most widely-used techniques for automatic selection of threshold values is the Otsu method, which can be paired with the Canny detector. The Otsu algorithm considers the candidate image as a gray-level histogram, and seeks to divide pixels into two classes by selection of a single threshold value that maximizes the interclass variance; the optimal threshold is obtained by an exhaustive search across the bins of the histogram (Otsu 1979). Though simpler than more advanced methods, this algorithm is particularly well-suited to our application as each input image may be expected to contain only two valid classes: background and foreground objects.

For use within the Canny detector, the Otsu algorithm was applied to the gradient magnitude image; the value of the gray-level intensity obtained by solving the optimization problem was selected as the high threshold value, and the value for the low threshold was set to one half the high threshold. For the same input image as in Figure D.2, the gradient magnitude histogram and interclass variance objective function evaluations are presented in Figure D.5, with respective results marked by vertical lines. Use of the Otsu method to automatically determine the unique hysteresis threshold values for each input image enabled implementation of the Canny edge detector for any number of images and/or video sequences without the need for individual tuning.
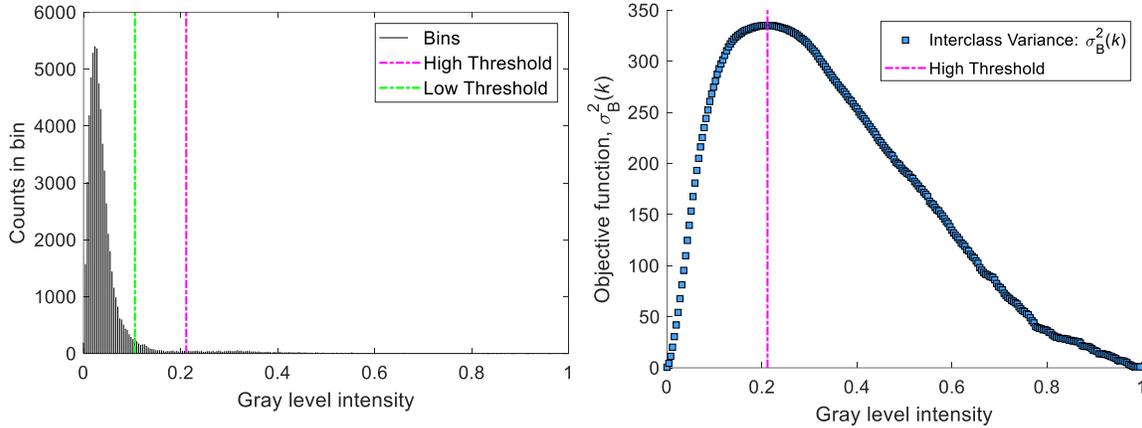
Figure D.5. Otsu method applied to gradient magnitude image. Left: histogram with low/high thresholds marked; Right: objective function evaluated for each bin; the optimal solution (high threshold) is apparent

In the final step of the Canny detector, hysteresis thresholding is used to thin the local maxima edge map (obtained from NMS step) using the high/low threshold values determined by the Otsu method. Starting from the NMS binary edge map (Figure D.6(a)), the procedure is as follows: find the locations in the NMS map which have a gradient magnitude less than the lower threshold, and create a new edge map by setting these locations in the NMS map equal to zero – the result (Figure D.6(b)) contains all weak and strong edges. Next, create another edge map which contains only the strong edges by finding all locations within the NMS (combined weak/strong) map which have a gradient magnitude greater than the high threshold (Figure D.6(c)). Then, determine the locations in the combined weak/strong edge map that are linked by 8-connectivity with the locations in the strong edge map; eliminate weak edges by setting their value to zero (in the final edge map) if none of their 8 neighbors in the 3 x 3 grid centered at (i,j) are a strong edge. The resultant edge map is the output of the Canny detector (Figure D.6(d)).
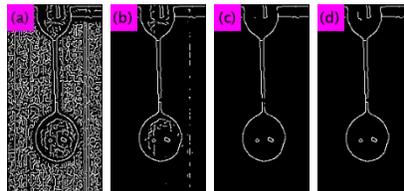

Figure D.6. Hysteresis Thinning. (a): local maxima edge map obtained from NMS, (b): weak/strong combined edge map, (c): strong edge map, (d): strong edge map with connected weak edges.

REFERENCES: Repeated from Chapter 5

[17] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679-698.

[18] Elder, J. H., & Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7), 699-716.

[19] Lindeberg, T. (1998). Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 117-156.

[20] Bao, P., Zhang, L., & Wu, X. (2005). Canny edge detection enhancement by scale multiplication. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(9), 1485-1490.

[21] McIlhagga, W. (2011). The Canny edge detector revisited. *International Journal of Computer Vision*, 91(3), 251-261.

[22] Xu, Q., Varadarajan, S., Chakrabarti, C., & Karam, L. J. (2014). A distributed canny edge detector: algorithm and FPGA implementation. *IEEE Transactions on Image Processing*, 23(7), 2944-2960.

[23] Lee, J., Tang, H., & Park, J. (2018). Energy efficient canny edge detector for advanced mobile vision applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(4), 1037-1046.

[24] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.

[25] Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-166.

APPENDIX E. CHAPTER 5 RESOURCE 2

VELOCIMETRY: Corresponds to 5.4.1

As drop formation depends on the system conditions, fluid properties and particle properties, it is useful to understand their interactions even at the earliest stages of liquid ejection. Consequently, axial velocity (maximum, average) is a useful parameter in understanding the initial stages of drop ejection through dimensional analysis/scaling, and may provide insight into the effect of particles and suspension on rheology/rheological transitions. Furthermore, the axial velocity profile (with respect to the x-axis) can provide insight into particle-driven effects on flow during drop ejection, which affect filament (and pendant drop) shape and therefore breakoff through the curvature of the liquid surface.

Quantitative estimates of velocity can be obtained by regression of position-time data for the liquid meniscus. Two different methods, described in the next sections, are employed to estimate velocity during the initial stages of liquid ejection: 1) maximum velocity, based on the temporal evolution of the liquid meniscus position which is farthest from the nozzle, i.e., the lowest point on the liquid meniscus, and 2) the velocity profile as a function of the x-axis, constructed from the temporal evolution of the liquid meniscus position at each *x*-pixel location.

MENISCUS POSITION AS $f(t,x)$

To perform the regressions, the meniscus z-position as a function of time and x-position must be extracted from the temporal sequence of edge maps respective to a given image sequence. Consideration of the physical situation represented in a single edge map indicates that the z-location of the meniscus at each x-position must be the last row index in each column that contains a non-zero entry. For each x-position (index j), this can be obtained by an iterative search through the row indices (i) which stores the row index in a separate matrix (p rows, n columns) at the $(k^*,j^*)^{\text{th}}$ position if the $(i,j^*,k^*)^{\text{th}}$ entry in the edge map sequence is non-zero. By applying this to all columns (index j) at a given time, k^*, the entire meniscus profile in a single edge map can be obtained. By performing this for all edge maps (index k) in the sequence, the profile at all times, denoted as $Z(t,x)$, is generated. Then, the meniscus z-position farthest from the nozzle at each

time can be recognized as the maximum value of $\mathbf{Z}(t^*, x)$ for all x: that is, the maximum value of all entries of a given row of $\mathbf{Z}$, $z_{max}(t^*) = \max\limits_{x} \mathbf{Z}(t = t^*, x)$. The results obtained for a real image are shown in Figure E.1.
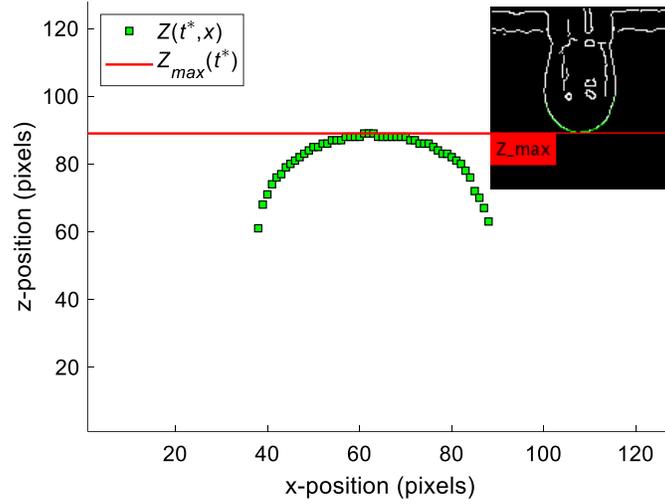

Figure E.1. Scatter plot of meniscus z-position, $\mathbf{Z}(t^*, x)$, with $z_{max}(t^*)$ marked by the red line. Inset: Corresponding edge map with $\mathbf{Z}(t^*, x)$ marked by green pixels and $\mathbf{z}_{max}(t^*)$ by the horizontal red line.

VELOCITY ESTIMATION: POINT

The maximum velocity at flow start-up can be estimated by applying a linear model, z = m·t + c, to $z_{max}(t)$, which yields the following least squares problem:

$$\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_k & 1 \end{bmatrix} [m \quad c]^T = \begin{bmatrix} \max\limits_{x} \mathbf{Z}(t = t_1, x) \\ \max\limits_{x} \mathbf{Z}(t = t_2, x) \\ \vdots \\ \max\limits_{x} \mathbf{Z}(t = t_k, x) \end{bmatrix}$$

in which the slope *m* is the quantity of interest. Representing this system as Ax = b, the familiar solution is x = (A^{T}A)$^{-1}$A^{T}b; though this problem is trivial in itself, some consideration must be given to automating this, as the first and last points of the time interval [t_1, t_k] must be systematically selected for each image sequence.

The starting point of the interval, t_1, should be the approximate time at which flow begins. This can be determined systematically by defining an initial meniscus position,

$z_{start} = \min z_{max}(t), k = 1, \ldots, 20$, and then defining the instant at which the bottom of the meniscus experiences displacement greater than or equal to a threshold value, q_1, here set equal to 2 pixels ($\approx$90 micrometers), from this position as the time point of flow start-up. This time step can be found by performing an iterative search (initial step, $k = 1$) of $\mathbf{z}_{max}$ for the first occurrence of $z_{max}(t) \geq z_{start} + q_1$, the k-index of which is assigned as t_{start}.

An end point of the interval, t_k, may be defined based on: 1) the liquid meniscus reaching a predetermined z-location (e.g. z_{start} + const.), or 2) a fixed time interval (e.g. t_{start} + const). To maintain a more general definition, a location basis, rather than a time basis, was selected as an accurate tuning of the latter presupposes some knowledge of the approximate velocity at a given set of conditions. The displacement from the initial position, q_2, was set at 45 pixels (1 nozzle diameter, $\approx$2 millimeter), and the time step, denoted t_{VP}, at which the meniscus reached this location was determined by applying the same iterative search procedure used for t_{start}, but for the first occurrence of $z_{max}(t) \geq z_{start} + q_1 + q_2$. Now that a systematic method of obtaining [t_{start}, t_{VP}] has been defined, one can compute the least squares estimate of velocity; an example is shown in Figure E.2, as are the images that correspond to the start/end points of the interval.
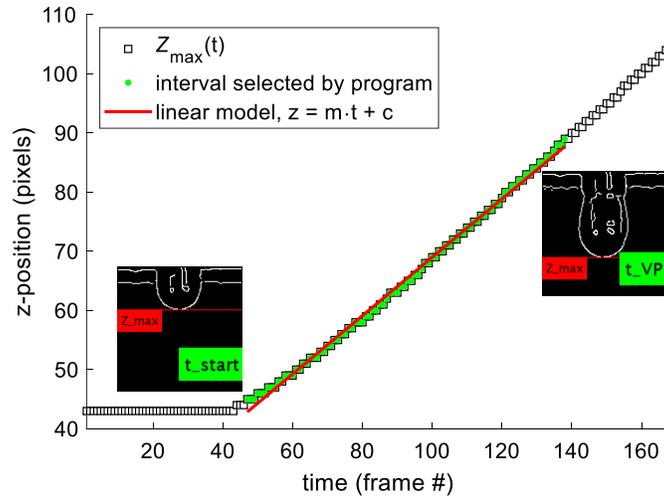

Figure E.2. Plot of maximum meniscus position ($z_{max}(t)$), interval for velocity estimation, and result of least squares model. Inset: binary edge maps at t_{start} (frame 47) and t_{VP} (frame 138) with respective $z_{max}(t)$ marked.

VELOCITY PROFILE ESTIMATION

To estimate the velocity profile, one must solve the least-squares problem for each x-position (j) given the respective time interval $[t_{1,j}, \ldots, t_{k,j}]$ and z-positions $[Z(t = t_{1,j}, x_j), \ldots, Z(t = t_{k,j}, x_j)]$. Given that the shape of the meniscus is changing with time, a common time based on movement of the meniscus at its lowest point – the same time interval as in the preceding section – must be used to ensure a sensible construction of $v(x)$. For each x-position j, the least squares problem becomes:

$$\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_k & 1 \end{bmatrix} [m_j \quad c_j]^T = \begin{bmatrix} Z(t = t_1, x_j) \\ Z(t = t_2, x_j) \\ \vdots \\ Z(t = t_k, x_j) \end{bmatrix}$$

which yields a total of n problems ($j = 1, \ldots, n$) if performed for each element on the horizontal axis; an application specific parameter, n is equal to the number of pixels on the x-axis of the image, thus, in the example here, $n = 128$.

The solutions are obtained in the same manner as before, $x = (A^TA)^{-1}A^Tb$; a simplification can be made by noting that for a given image sequence, the matrix A is the same in each of the least-squares problems above, therefore the pseudoinverse, $A^\dagger = (A^TA)^{-1}A^T$, need be computed only once. Then, the solutions are computed as $x_j = A^\dagger b_j$, which realizes a significant time-savings as each successive x_j requires just one matrix-vector multiplication. The velocity profile, $v(x)$, is constructed from the least-squares solutions, such that $v(x_j) = m_j$; this is shown in Figure 3 (Left) by the blue circles.

In order to gain insight into the dynamics, it may be of interest to estimate the velocity profile at different times; several possible definitions exist for subdividing time intervals, of which we consider two. The first option resembles the standard time interval in that the left bound is fixed at t_{start}, but varies the right bound, such that the interval, $[t_1, \ldots, t_k]$, used in the respective least-squares problems is of the form: $[t_{start}, \ldots, [t_{start} + (t_{VP} - t_{start})/N]]$; the second possibility would be to allow the left bound to also vary with a formula similar to the right bound, but in which the denominator is greater than N. Examples of the results from these two definitions are shown in Figure E.3, for the averaged velocity profile obtained from the 30 image sequences.
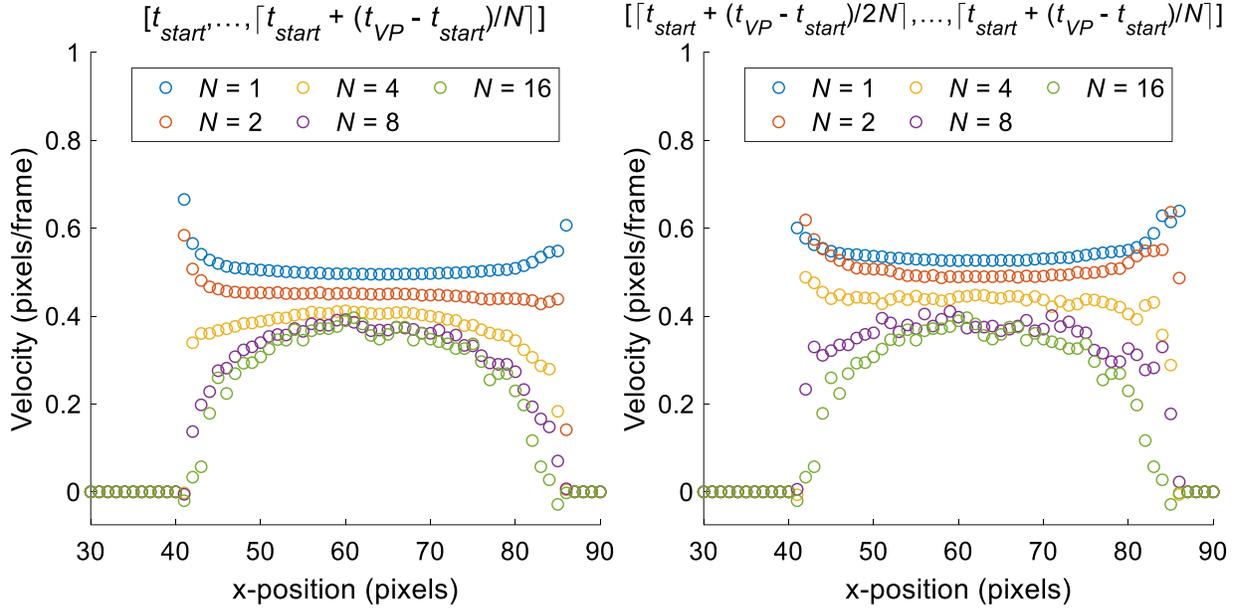
Figure E.3. Axial velocity at each x-position, computed using both definitions of the time intervals (given in figure title) at different time points. Left: larger N corresponds to shorter time intervals, and the legend notes the right bound denominator. Right: larger N corresponds to time intervals temporally closer to the incipience of flow; for $N = 16$, the left bound is simply t_{start}.

DIAMETER AS FUNCTION OF TIME, SPATIAL VARIABLES: Corresponds to 5.4.2

DIAMETER AS $f(t,z)$

An important aspect with significance to the rheology (and other aspects of drop formation) is the diameter of the liquid filament as a function of time and axial position. To obtain this from an edge map, one must determine the left and right boundaries of the liquid surface on the x-axis, and then compute the diameter as the difference. In the discrete representation (edge map), the boundaries of the liquid surface at a given z-position can be determined by finding the x-positions of the first occurrences of non-zero pixels on the left/right sides of the filament using iterative searches. For each z-position, the left boundary may be found by searching across the elements (index j) of the row (index i) of the edge map at time k, incrementing from the left-most image limit, $j = 1, 2, \ldots$, for the first occurrence of a non-zero entry, at which point the search terminates and assigns $X_L(k, i) = j$; the method for the right boundary is analogous, except that the search decrements from the right-most image limit, $j = (n{=}128), 127, \ldots$, and assigns $X_R(k, i) = j$ at the

first occurrence of a non-zero entry, then terminates. By repeating this for all z-positions (row indices, $i = 1, \ldots, (m=512)$) in a given edge map, and for all times ($k = 1, \ldots, p$) in the sequence, the matrices $\boldsymbol{X}_L(t, z)$ and $\boldsymbol{X}_R(t, z)$ are generated.

To find the filament diameter, $\boldsymbol{D}(t, z)$, one cannot simply compute the difference between the two matrices above, as the discrete nature of the problem necessitates an adjustment for the pixel which would be lost; Figure E.4, row 3 makes a poignant example, but the loss – 1 pixel unit – is the same for all rows. Consequently, the diameter must be computed as the difference between the $(k, i)^{\text{th}}$ entries of $\boldsymbol{X}_R$ and $\boldsymbol{X}_L$ to which +1 is added to account for the discrete coordinates, but only if either entry is non-zero. In our implementation this is performed by means of a loop over i nested within a loop over k, which applies the conditional logic to check for presence of a non-zero entry, then assigns $D(k, i) = X_R(k, i) - X_L(k, i) + 1$ if true, and if false, $D(k, i) = 0$.
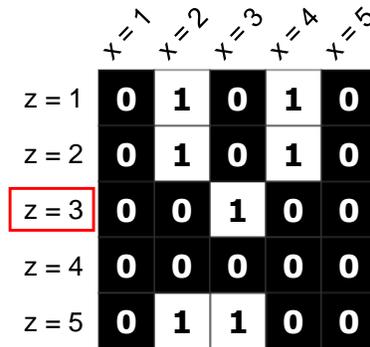
Figure E.4. Sketch for a binary representation of a liquid filament. $X_L(z = 3) = 3$, $X_R(z = 3) = 3$.

FILAMENT THINNING: MINIMUM *D(t,z)*

The minimum width of the filament in the region over which it is actively decreasing in diameter prior to the first breakoff event is an important parameter in understanding the breakoff dynamics and thinning mode. Given that the macroscopic shape of the emergent liquid surface is approximately that of an hourglass with one rounded end, and the other attached at the nozzle, a simple search of each row of $\boldsymbol{D}(t, z)$ for the minimum non-zero entry will inevitably lead to incorrect values as the diameter at the farthest point from the nozzle – the bottom of the rounded end – will likely be smaller than the body of the hourglass at various times. An example of this type of error is marked in Figure E.5 by the arrow pointing to "incorrect D_{min}".

265

However, if one is able to bracket the z-interval of interest, then a search for the minimum on this interval will yield the desired result: $D_{\min}(t, z)$ in the thinning region. Since, logically, it must reside between the nozzle and maximum diameter of the liquid surface for thinning regions that resemble a stretched/deformed hourglass, the function $- D(t^*, z) -$ is convex on this interval. Therefore, one finds $D_{\max}(t^*, z) = \max D(t^*, z)$, and defining $z_{\text{nozzle}} = \text{const.}$, one finds $D_{\min}(t^*, z) = \min D\left(t^*, [z_{\text{nozzle}}, \ldots, z_{D_{\max}(t^*)}]\right)$. An example of this method applied to a single time step is provided in Figure E.5, with the z-interval bounds marked by vertical lines, between which the z-position of the minimum diameter may be found; the edge map is included with corresponding positions marked by horizontal lines. In Figure E.6, the results generated by applying this method to each time step of an image sequence (up to time of breakoff) are presented.
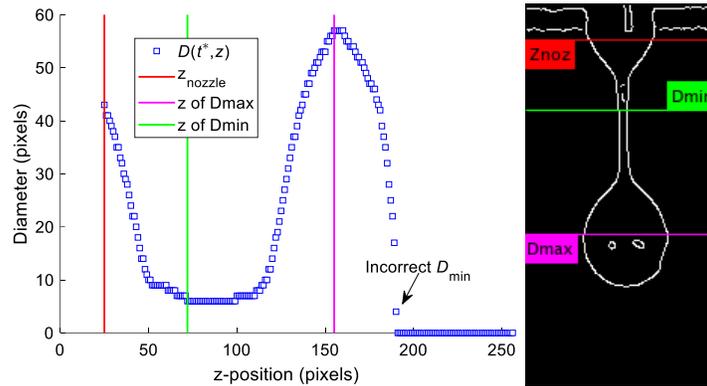

Figure E.5. Results obtained by method above: $D_{\min}(t^*, z)$, $D_{\max}(t^*, z)$ Left: Scatter plot of $D(t^*, z)$, with z-positions marked by vertical lines. Right: edge map with positions marked by horizontal lines.
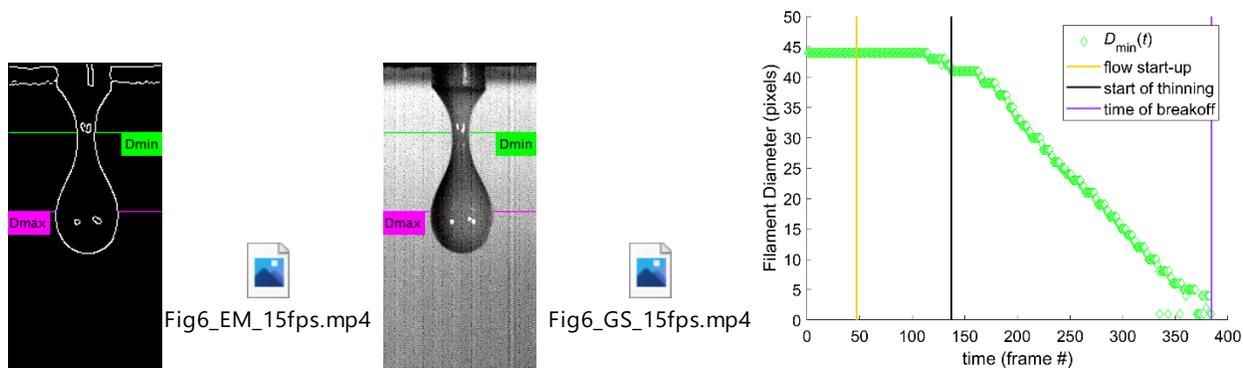
Figure E.6. Left: Minimum filament diameter as function of time and z-position; the video is generated by insertion of horizontal lines at the z-positions of the minimum and maximum filament diameter on each edge map in the temporal sequence. Middle: Same as Left, but for grayscale image sequence. Right: Scatter plot of minimum filament diameter as a function of time. Note: the marked edge map & greyscale image shown here are representative images from the videos.

DROP VOLUME, DROP EJECTION TIME: Corresponds to 5.4.3

To quantify the effects of particle properties, ink properties and process conditions on process performance, one would like to evaluate: 1) the number of drops that are formed per ejection event and their volumes and 2) the time required to eject each drop. These quantities depend somewhat on the perspective of the observer, as one must define both modal, temporal and spatial conditions in order to choose from several possible contingencies. We present specific choices of these in the following section: others can equally be made without affecting the algorithm. Specifically, we select a position ≈6mm from the top of the image, $z = 131$, as a workable boundary for the example below.

DROP EJECTION TIME

In the context of drop printing, one may define the time interval which begins at the time of flow start-up, t_{start}, and ends when all objects leave the vicinity of the nozzle as the ejection time – in other words, the time at which all detached drops have completed their crossings of a z-location which is specified by the interests of the observer. Consequently, the method to determine ejection time is in essence a method which relies on detection of object crossings at a user-selected z-position (e.g. $z_E = 131$). The logic applied in this algorithm relies on knowing the z-position of

the liquid meniscus (attached to the nozzle, or not) farthest from the nozzle, but above z_E, which can be determined by the same search procedures as described for object segmentation in Section 5.5.1., but with iterations across the z-rows restricted to $i < (z_E+1)$; the resultant matrix and vector are denoted as $\boldsymbol{E}(t, z)$ and $\boldsymbol{E}_{max}$, respectively. The algorithm consists of two inner search methods, described below, nested within an outer iterative search which steps forward through the temporal sequence, $E_{max}(t)$, to systematically identify the time points at the start (t_{SC}) and end (t_{EC}) of each crossing event; after the second inner search ends, a check for additional crossing events is made on the remainder of the temporal sequence – $[t_{EC} + 1, \dots, p]$ – and the search continues until all crossings have been enumerated.

If one considers the crossing of a discrete z-position by an object, it can be observed that some part of the object must necessarily occupy at least one x-pixel at this z-position for the duration of the crossing event, which corresponds to a time interval composed of a sequence of discrete time points. Therefore, the start of an object crossing (t_{SC}) can be detected as the first time step at which the row at the selected boundary contains a non-zero entry, i.e. $E_{max}(t) = z_E$. Then, the point (t_{EC}) at which the same object completes its crossing must be the next time step in the sequence at which the row at the boundary contains all zero entries, i.e. $E_{max}(t) \neq z_E$; (consider that for all time points between, the value of $E_{max}(t)$ must be equal to z_E). Specifically, two separate iterative searches are used: for the start time, iterations terminate at the time point k, if $E_{max}(k) = z_E$; the search for the end time is phrased as search for the last time step at which the selected boundary contains a non-zero entry, thus, iterations terminate at k if both $E_{max}(k) = z_E$ and $E_{max}(k + 1) \neq z_E$. The result of applying the algorithm to an image sequence is presented in Figure E.7.
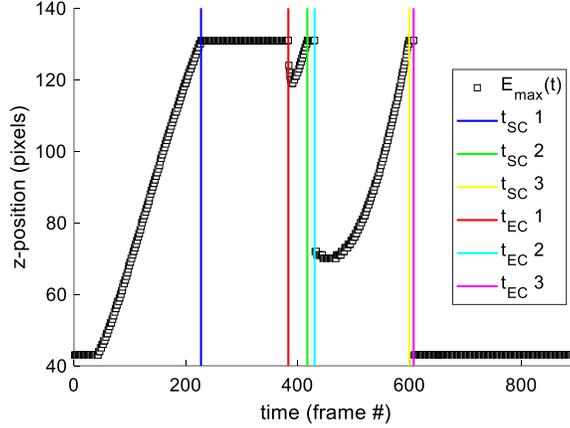
Figure E.7. Plot of $E_{\max}(t)$ with start (t_{SC} #) and end (t_{EC} #) time of crossing events marked by vertical lines; three total crossing events occurred.

The time points selected by the algorithm provide the means for evaluating drop volume at a consistent location, that is, the time points at which the drop has just finished crossing the selected z-location. Depending on whether the drop has stabilized (tail retracted, drop oscillations ended), this may or may not be the preferred location to compute the drop volume. However, if one wishes to enforce consistency of other relevant aspects in computation of drop volume (e.g. lighting conditions, proximity to characteristic length scale), then one must heuristically select a z-position in order to employ a method such as this to identify the specific time points which are passed to the drop volume method. Additionally, the number of crossing events may be used to infer the number of drops formed per drop ejection event, though caution must also be exercised here as breakup/coalescence of drops can occur below the user-selected boundary.

DROP VOLUME

The method for computing the volume of a drop from its image is straightforward, provided that one can assume the drop (or liquid fragment) exhibits axial symmetry. For a cylindrical object of non-constant diameter, the volume can be computed as: $V = (\pi/4) \int_{z_{\text{bottom}}}^{z_{\text{top}}} D(z)^2 dz$, for which the discrete form necessitated by the pixel representation is: $V = (\pi/4) \sum_i D(z_i)^2 \Delta z_i$; Figure E.8 depicts the correspondence between the continuous and discrete representations. In essence, the volume is the sum of the cylindrical slices with cross-sectional area $(\pi/4) \times D(z)^2$ and height Δz = 1 = const, which is computed by looping over the i entries of the interval $[z_{\text{bottom}}, z_{\text{top}}]$ and

269

summing the terms. Therefore, one must obtain the z-locations which are the top and bottom boundaries of the drop, given the times passed from the ejection time algorithm.
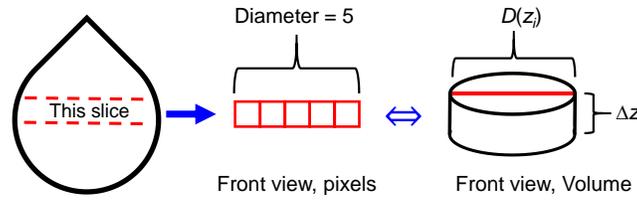

Figure E.8. Pixel representation of one cylindrical slice.

With the object-specific information extracted as described in Section 5.5, evaluation of volume is straightforward, since the respective z-interval of each object is known at all time points. A substantial advantage of using the object-specific reference frame is that it enables creative evaluation of volumes in the various ways seen throughout this this thesis, for example, Chapter 6, and also in Chapter 8.

DROP VOLUME: Linked to Ejection Time

From a practical perspective, evaluation by object reference frame is far easier to manipulate and utilize than the method described below, which is included only for the sake of completion. This method was originally developed for use with the ejection time information determined by the non-object method described above; it is inferior for various reasons.

If one considers the situation depicted in Figure E.9, some insight is provided into how one might structure an algorithm whose objective is to find the bounds of the object which has just completed its crossing. The upper limit of the liquid surface can be identified using an iterative search (incrementing) for the first non-zero entry of the set $\{D(t^*, z): z \geq z_E\}$; the loop terminates at the index of the desired entry, which is the top of the drop, denoted z_{top}. Then, to find the lower limit, a second iterative search is initiated starting at index $i = z_{top}$, and continues until the last non-zero entry which is part of the same drop is found – in other words, the search finds the first zero entry of the restricted set $\{D(t^*, z): z \geq z_{top}\}$, the index of which is denoted z_{bottom}.
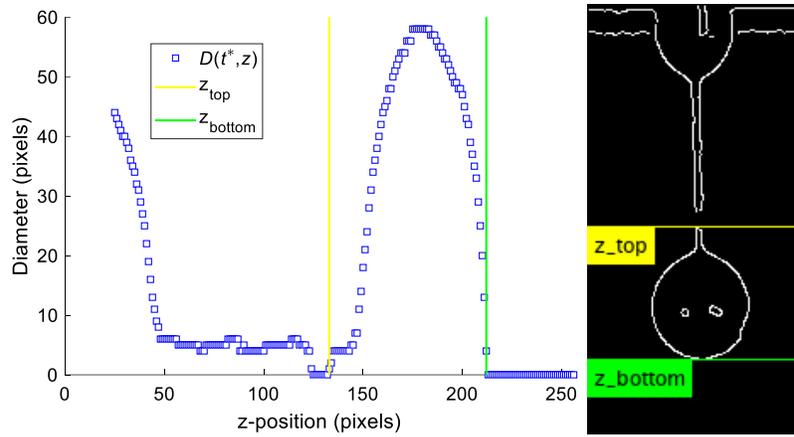
Figure E.9. Time point at end of crossing event (t_{EC} 1 above). Left: Diameter as function of z. Right: corresponding edge map.

With z_{top} and z_{bottom} determined, the volume of the object is computed, after which the outermost layer of the algorithm proceeds to the next time point passed by the ejection time algorithm. The drop volumes computed at the time points (t_{EC}) identified in Figure E.7 are shown below in Figure E.10 with their corresponding edge maps; the volume reported for each image corresponds to the drop which has just finished its crossing of z_E .
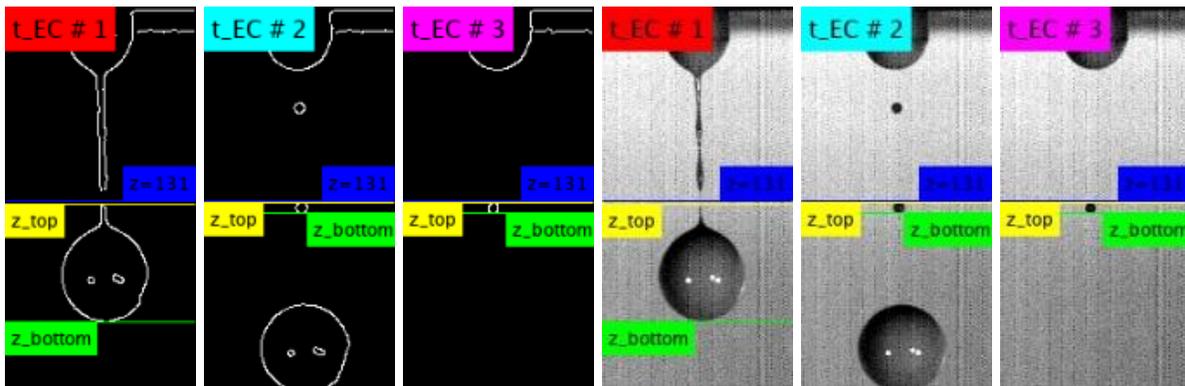
Figure E.10. Edge maps (Left) & actual images (Right) at t_{EC}. Volumes of drops for t_{EC} # 1, 2 & 3 are, respectively, 110328, 341 & 240 cubic pixels.

APPENDIX F. CHAPTER 5 RESOURCE 3

ADDITIONAL FIGURES: SECTION 5.6.1

As shown in Figure F.1, the breakoff lengths for secondary/tertiary detachments would naturally be smaller due to the retraction of the filament, however, across the repetitions there are notable differences in the locations at which the second and third drops detach – presumably this is related to the effect of particles (and distributed particle size). The volumes of each drop at the instant of detachment are presented in, Figure F.2; note that a volume of 1.1×105 pixel3 corresponds to 10.33µL, and a volume of 250 pixel3 to 23.47nL.
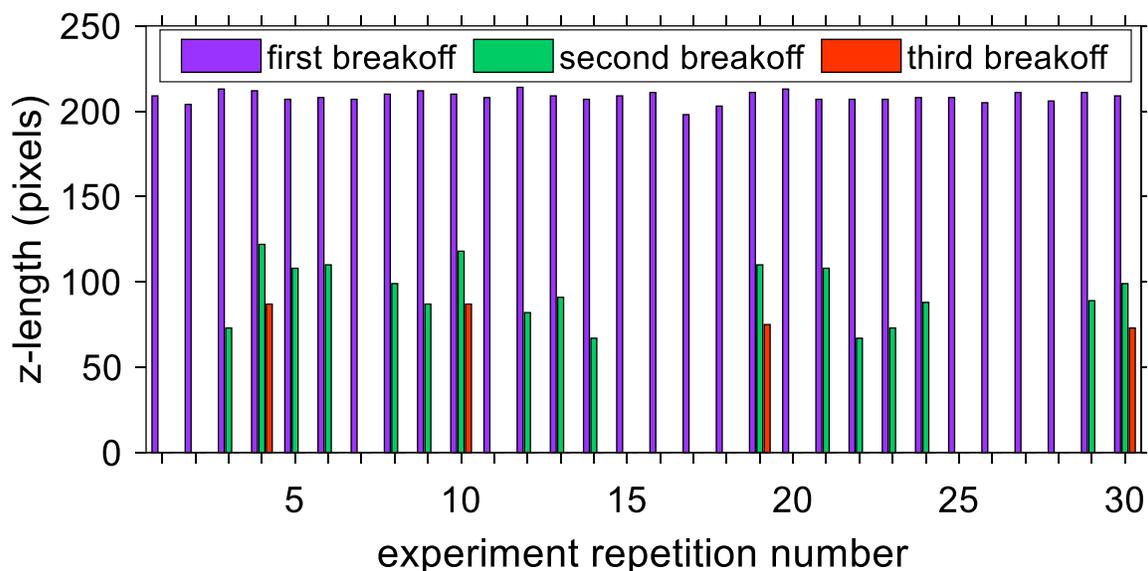
Figure F.1. Length of filament at time of breakoff; this length is physically equivalent to the green line denoted "z_L" in Fig. 24. Note the variance (~11) in the filament length for the first breakoff.
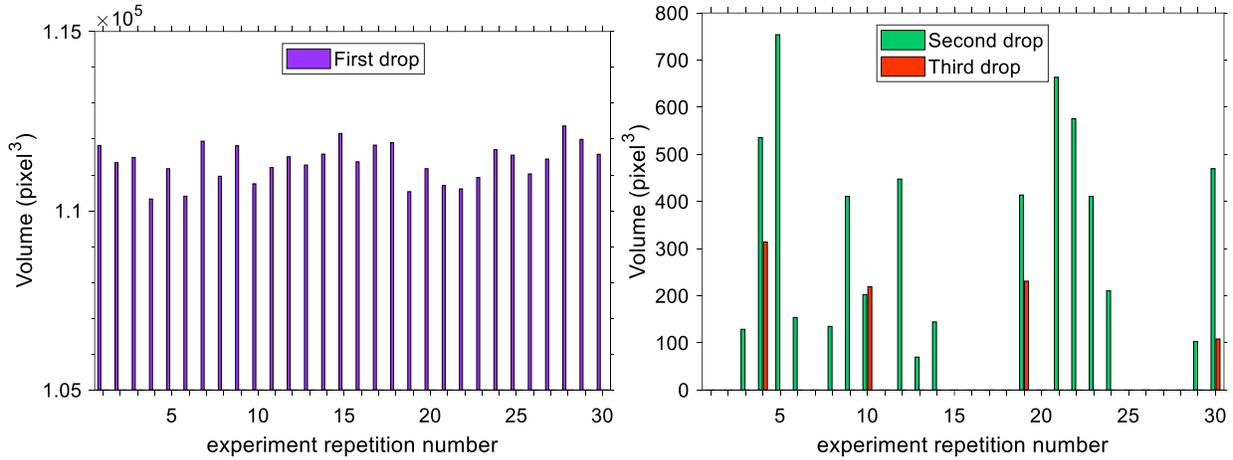
Figure F.2. Volumes of primary drops at detachment; Right: volumes of secondary and tertiary drop(s) at time of breakoff.


ADDITIONAL FIGURES: SECTION 5.6.2
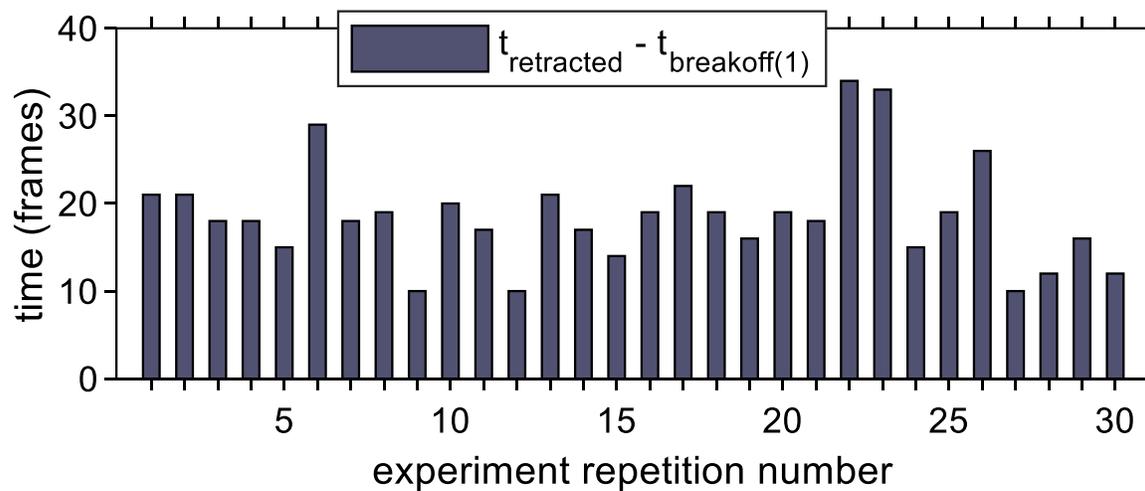
Figure F.3. Time at which the filament has retracted to its initial position. On the left, time is scaled to the usual start-up of flow; on the right, the time is additionally scaled to the respective time of the first breakoff event.
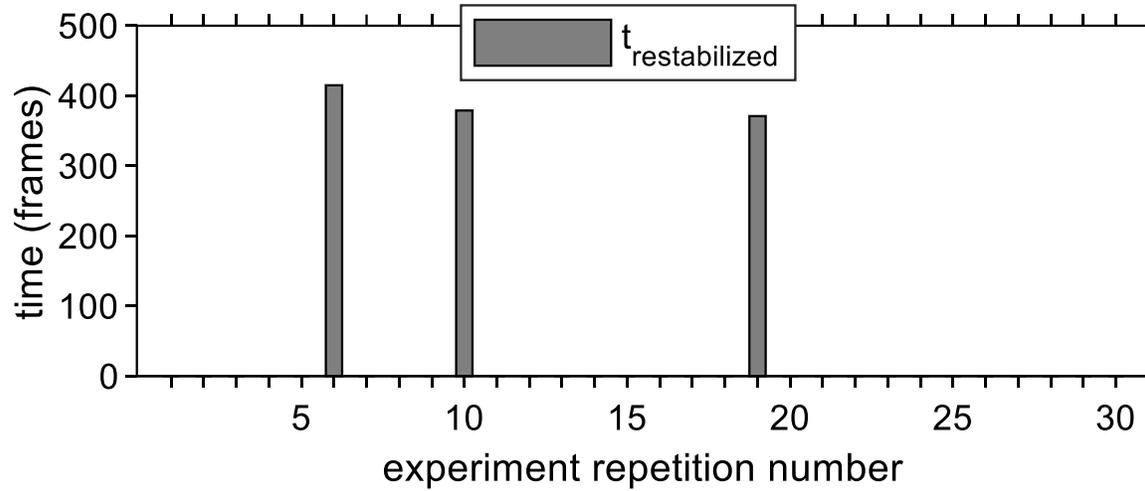
Figure F.4. If merging at the nozzle occurs after the broken filament has fully retracted, then one must instead use the adjusted retraction ("re-stabilized") time.

Rather than use t_{merge}, it would be more appropriate to use $t_{restabilized}$; however, the differences happen to be quite small in this data set, thus, the manuscript uses t_{merge}. More rigorously: the time from flow start-up to returning to a fully stable meniscus is equal to $max\,(t_{retracted}, t_{restabilized})$; alternately, if the presence of objects in the vicinity matters, then $max\,(t_{retracted}, t_{restabilized}, t_{ejection})$.
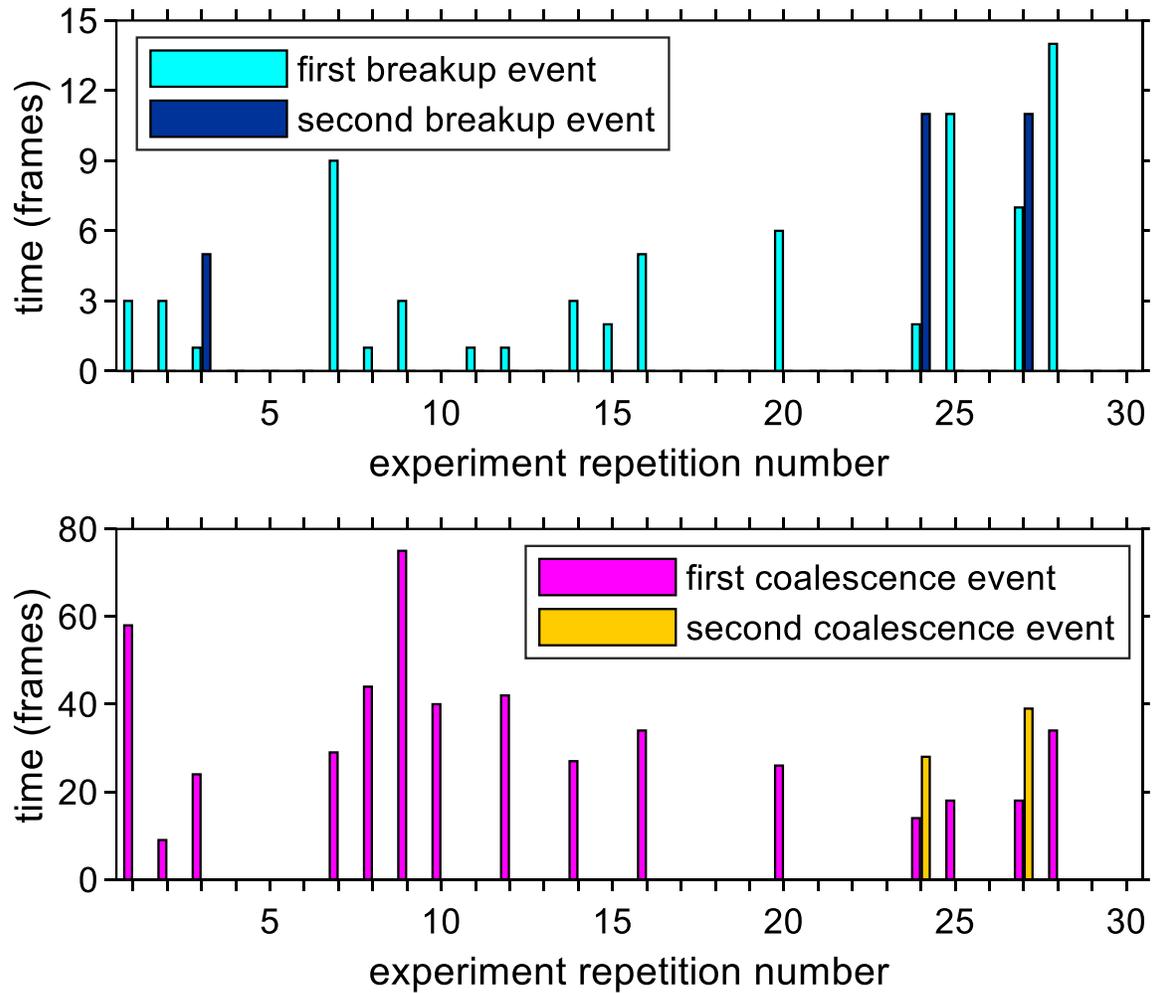
Figure F.5. Top: Number (implicitly indicated by presence/absence of bars) and times (scaled to time of breakoff) of breakup events for primary drop. Bottom: Number (implicitly indicated) and times (scaled to time of detachment) of coalescence events for primary drop

APPENDIX G. CHAPTER 5: MOTION TRACKING PROOF

PROOF OF MINIMUM BOUND FOR COALESCENCE DETECTION

If one considers the motion tracking searches in the context of two drops interacting at the upper bound of the reference drop, as shown in Figure 5.20, the necessary condition (note: not sufficient) for an interaction flag is: the liquid surface position of the non-reference drop at the next time step must be greater than or equal to the bound position of the reference drop, which is expressed in Eq.(1):

$$z_{L,2}(t) \geq z_{UB,1}(t-1) \tag{1}$$

Accounting for the relative motion symbolically, at the time of detection the number of unoccupied rows, UR, between the objects may be expressed as in Eq.(2):

$$\left[\left(z_{UB,1}(t-1) + BR + v_{U,1}\Delta t\right) - \left(z_{L,2}(t-1) + v_{L,2}\Delta t\right)\right] - 1 = UR \tag{2}$$

which is simply the difference between the liquid surface positions at time t_2, extrapolated from t_1, adjusted for the pixel coordinate representation (origin of the "-1" outside the bracket). It follows that the liquid surface position of the non-reference drop is $z_{L,2}(t-1) = z_{L,2}(t) - v_{L,2}\Delta t$, hence, the number of unoccupied rows may be expressed as in Eq.(3):

$$z_{UB,1}(t-1) - z_{L,2}(t) + v_{U,1}\Delta t + BR - 1 = UR \tag{3}$$

Therefore, in order to satisfy the criterion for the minimum bound, the left side of Eq.(3) must be greater than 0, with the bound region size, BR, as the only adjustable parameter. While maintaining that the positions, bounds and velocity are assumed to be unknown for any particular time step, it remains possible to use physical reasoning to gain some insight into Eq.(3).

The first two terms in Eq.(3) represent the difference between the upper bound of the reference drop at the previous time and the position of the non-reference drop at the current time, and may be conveniently expressed as a single quantity: $z_{UB,1}(t-1) - z_{L,2}(t) = p$. Intuitively, this is the distance by which the entrant drop overshoots the bound, and if the necessary condition for an interaction flag (Eq.(1)) is satisfied, then $p \leq 0$. Now, let us express Eq.(3) in terms of p, and apply the criterion for the number of unoccupied rows that must necessarily be present to flag the time step as a potential interaction:

$$p + v_{U,1}\Delta t - 1 + BR > 0 \qquad\qquad (4)$$

next, apply the necessary condition for the interaction flag provided in Eq.(1), which requires: $p \le 0$. Therefore, when both of the two necessary conditions are satisfied, this constitutes the sufficiency condition for the minimum bound, given by Eq.(5):

$$p + v_{U,1}\Delta t - 1 + BR > 0, \quad p \le 0 \qquad\qquad (5)$$

However, as the overshoot, p, and the adjustment for the movement of the upper surface of the reference drop, $v_{U,1}\Delta t$, depend on the relative velocities and positions (at previous time step) of the drops, it seems that this may be a fruitless endeavor. However, physical reasoning can provide a means for elucidating an extremum (minimum) value for p through consideration of limiting cases.

Before proceeding, consider that a coalescence event may occur under two types of conditions, as illustrated in Figure G.1.: 1) two drops, equal or unequal magnitude of velocity with opposite direction, and 2) two drops, unequal velocity magnitude but with same direction. With respect to these two cases, the physical reasoning for a limiting case is most clearly presented through an examination of the second type of coalescence event, for which a detailed diagram is provided in Figure G.2.
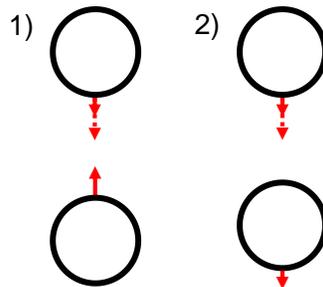

Figure G.1. Possible configurations of a coalescence event.
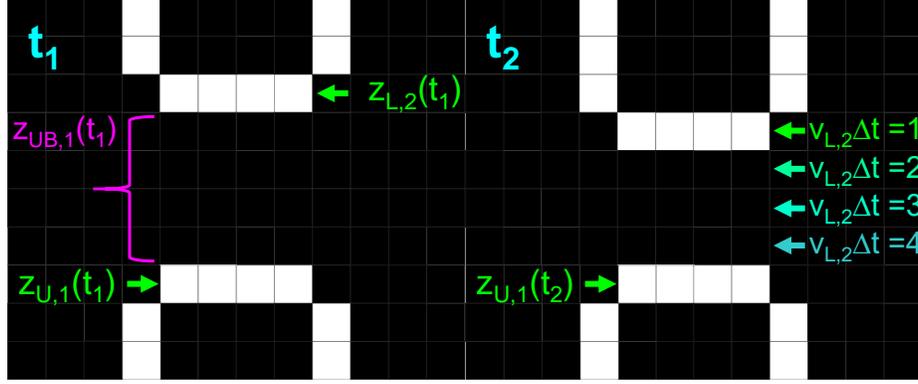
Figure G.2. This physical arrangement represents the limiting conditions for reference drop movement, $v_{U,1}\Delta t = 0$, and also overshoot, as $c = 1$. Consider that for $v_{L,2}\Delta t = 1$ (depicted above), 2, or 3, Eq.(6) is satisfied by BR = 4; however, for $v_{L,2}\Delta t = 4$ there is no longer an unoccupied row separating the two drops

A limiting case occurs for the second type of coalescence event when the velocity of the reference drop, $v_{U,1}$, is such that at the time depicted in Figure G2, $v_{U,1}\Delta t$ is equal to zero – this corresponds to no net movement of the upper surface of the reference drop between the time steps, and may occur for reasons related to the discrete representation of the objects (e.g. a velocity of 0.5 pixel/frame implies movement every other frame). Essentially, this constitutes the worst-case scenario for a reference drop in this context, as any downward movement of the upper surface would otherwise increase the number of unoccupied rows available, and therefore permit larger magnitudes for the overshoot, p. This reduces Eq.(5) to:

$$p - 1 + BR > 0 , \quad p \leq 0 \hspace{4cm} \text{Eq.(6)}$$

A further limiting case for the overshoot, p, occurs when the liquid surface of the non-reference drop is exactly 1 row above the bound of the reference drop at the prior time step; naturally, this need not occur at the time step prior to detection (the liquid surface could be farther from the bound). However, this is the case which minimizes the overshoot, p; to see this, consider the alternate expression for p, $p = z_{UB,1}(t-1) - z_{L,2}(t-1) - v_{L,2}\Delta t = c - v_{L,2}\Delta t$, noting that $c \geq 1$, and in the limiting case, $c = z_{UB,1}(t-1) - z_{L,2}(t-1) = 1$, therefore, $p = 1 - v_{L,2}\Delta t$. As a useful simplifying assumption, it is preferable to use $p \propto -v_{L,2}\Delta t$, or in words, overshoot is proportional to the negative of the velocity of the liquid surface of the non-reference drop. An upper limit for $v_{L,2}\Delta t$ may be estimated from the maximum velocity magnitude, such that $p \propto$

$-\text{ceil}\big(\text{max}\big(v_{L,2}\Delta t\big)\big)$. Therefore, the minimum bound region size for the second type of coalescence event may be expressed as in Eq.(7):

$$BR > 1 + \text{ceil}\big(\text{max}\big(v_{L,2}\Delta t\big)\big) \tag{7}$$

For the first type of coalescence event, depicted in Figure G.1., the proportionality of the overshoot and definition of the corresponding limiting case remain the same. However, the assumption that the upper surface of the reference drop is stationary, i.e. $v_{U,1}\Delta t = 0$, is no longer appropriate; the worst-case scenario is for $v_{U,1}\Delta t$ to be negative, in which the movement of the reference drop actively decreases the number of unoccupied rows available. To account for this, an upper limit for the magnitude of $v_{U,1}\Delta t$ should be restored to the expression for the minimum sufficient bound region, which yields the following expression (Eq.(8)):

$$BR > 1 + \text{ceil}\big(\text{max}\big(v_{L,2}\Delta t\big)\big) - \text{floor}\big(\text{min}\big(v_{U,1}\Delta t\big)\big) \tag{8}$$

Noting that $v_{U,1}\Delta t$ is assumed to be negative in Eq.(8), this may be equivalently expressed by Eq.(9):

$$BR > 1 + \text{ceil}\big(\text{max}\big(v_{L,2}\Delta t\big)\big) + \text{ceil}\big(\text{max}\big(|v_{U,1}\Delta t|\big)\big) \tag{9}$$

To select between the two expressions for the minimum sufficient bound (Eq.(7) and Eq.(9)), reflection upon the intrinsic uncertainty in type of coalescence event dictates that the more difficult detection scenario, which occurs in the first type of coalescence event, should be used to determine the bound size, therefore a value of BR that satisfies Eq.(9) is the minimum sufficient bound for both types, as a value of BR that satisfies Eq.(9) also satisfies Eq.(7). In order to avoid assumptions with respect to the velocities of any two specific drops, the maximum velocity magnitude (v_{max}) during drop ejection (which is presumably near the global maximum) may be substituted for both $v_{L,2}$ and $v_{U,1}$. Thus, the minimum sufficient bound region may be expressed as Eq.(10):

$$BR > 1 + 2\text{ceil}(v_{max}) \tag{10}$$

In the image sequences in this paper, the frame rate of the camera resulted in effective velocities of 0.5-0.8 pixel/frame during drop ejection, therefore application of Eq.(10) would yield a minimum bound region size of 4. An illustrative example is presented below.
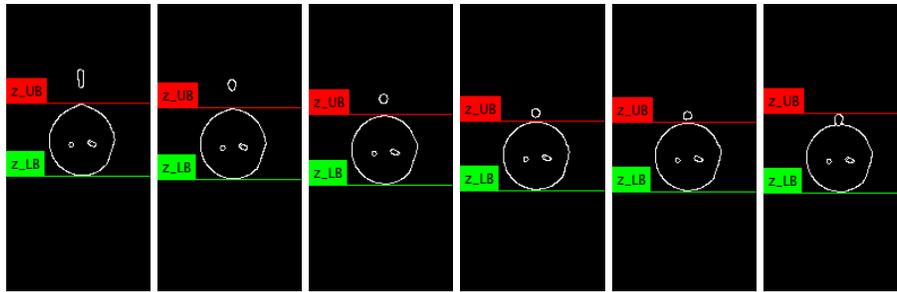
Figure G.3. Failure of detection with bound size = 1. Edge maps, from left to right: t = 1, 4, 9, 15, 16, & 17.
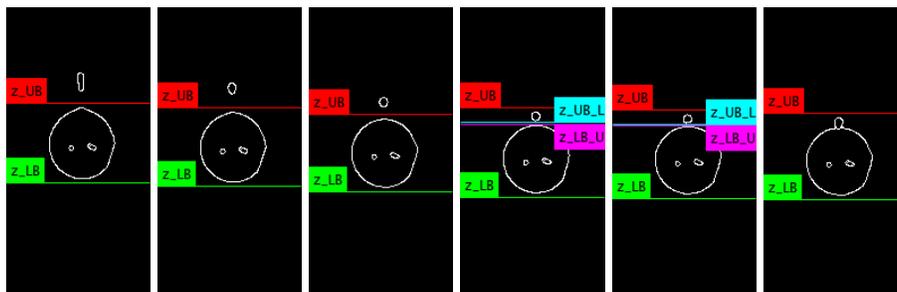
Figure G.4. Successful detection with bound size = 4. Smaller drop first included in z-interval at t = 15 (3rd image from left), which is the first time step flagged as a potential breakup. Edge maps, from left to right: t = 1, 4, 9, 15, 16, 17.

The use of a larger bound – the minimum sufficient bound – leads to detection of a coalescence event which is not detected by a bound size of 1. This is demonstrated by a simple comparison of the coalescence event from Figure 5.16, with the algorithm started at t = 464 and initially only the larger drop encompassed by the z-interval. Application of the 1-pixel bound size to this example is presented in Figure G.3., which demonstrates that the coalescence event is not detected, for even at the time step prior to coalescence, the smaller drop is not included in the z-interval. Conversely, for the 4-pixel bound size presented in Figure G.4, the merging event is flagged at 2 time steps prior to osculation of the respective upper ($z_{U,1}$) / lower ($z_{L,2}$) liquid surfaces of the large/small drops. A real example of coalescence detection outside drop identities is provided by Figure 5.1, Middle.

APPENDIX H. CHAPTER 5: OBJECT SEGMENTATION

The following illustration, Figure H.1., demonstrates the steps of the object recognition/identity assignment algorithm for the simultaneous breakoff of 4 drops. Initialized upon detection of a singularity, the algorithm that recognizes the objects performs 4 outer loop iterations in total; each iteration consists of two separate searches, which start from "z_U" and "z_L" and that converge toward each other. The termination condition for the algorithm is dependent on the parity of the iterative index: if the iterative index is even, the termination condition requires that all remaining entries (on $\tilde{\mathbf{z}}(t)$) of the diameter vector be non-zero; if the iterative index is odd, the termination requires that all remaining entries of the diameter vector be equal to zero. This is demonstrated in both the example presented in Figure H.1. for the even termination point, and in Figure 5.10 for the odd termination point, (though it is perhaps difficult to see as the gap is only 1 pixel).

To better illustrate the odd termination point, Figure H.2. demonstrates the case for the simultaneous breakoff of 5 drops. The general applicability of this algorithm to any number of objects, odd or even, should become clearer when viewed in light of this figure. The codes written to implement the interval search, concatenation and subsequent identity assignments benefit from the general nature of the logic in that switch-case (or analogous) structures, specific to each object total, may be completely avoided, thus, a compact code can implement this scheme for truly any number of objects.
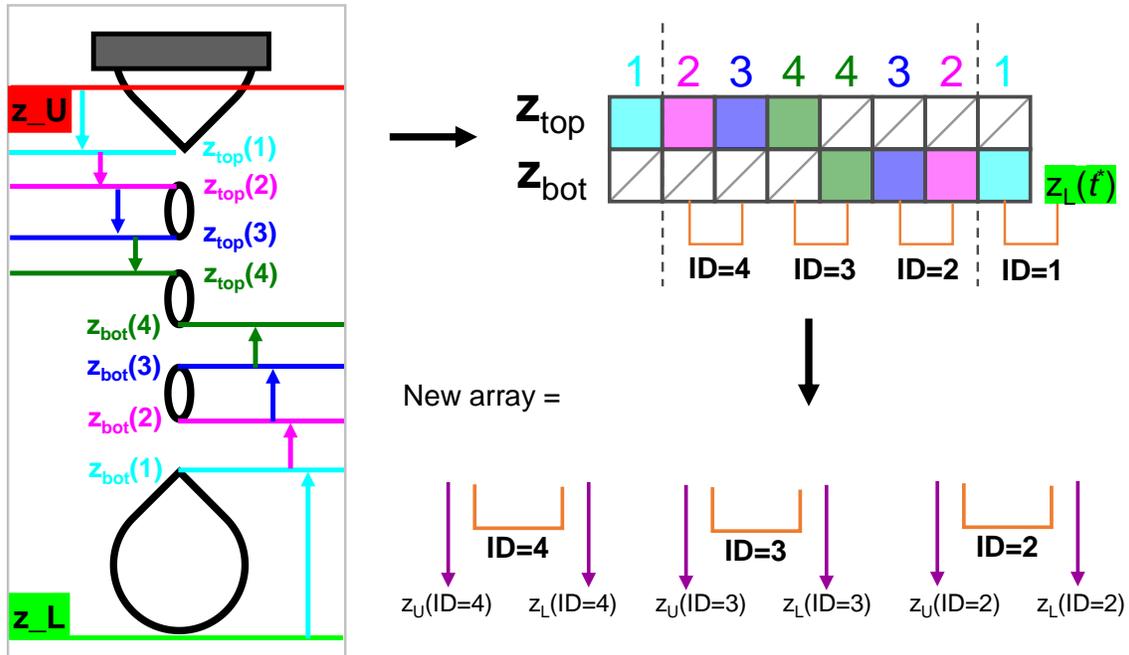
Figure H.1. Illustration of simultaneous detachment of 4 drops. The sequence of iterations that determine the structure of the singularity are marked in cyan, magenta, blue and dark green.
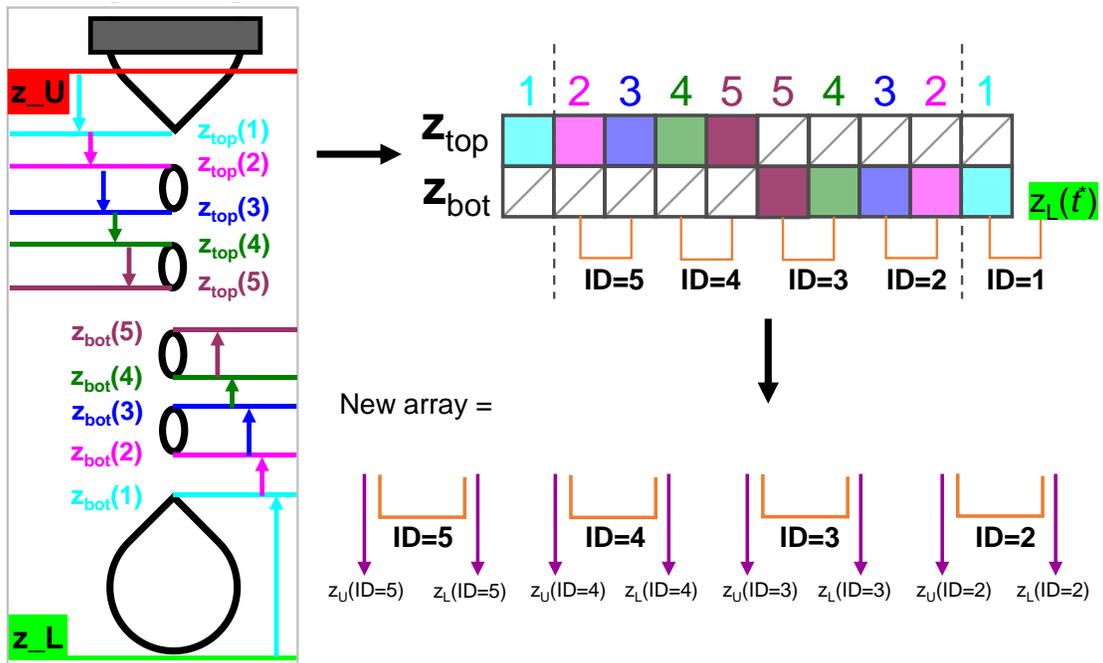

Figure H.2. Illustration of simultaneous detachment of 5 drops. The sequence of iterations that collapse the interval are marked in cyan, magenta, blue, dark green and mauve

APPENDIX I. CHAPTER 8: EXHAUSTIVE LIST OF BATCH CONDITIONS

Table I.1. The format of the string is: **DosePotencyLevel_ParticleType_FluidType_PhiValue**

k = 1	'SingleDrop_uAPAP_NEO895_0.085'
k = 2	'SingleDrop_uAPAP_NEO895_0.175'
k = 3	'SingleDrop_uAPAP_NEO895_0.2'
k = 4	'SingleDrop_uAPAP_NEO895_0.225'
k = 5	'SingleDrop_uAPAP_LPG_0.079'
k = 6	'SingleDrop_uAPAP_LPG_0.175'
k = 7	'SingleDrop_sfAPAP_NEO895_0.085'
k = 8	'SingleDrop_sfAPAP_NEO895_0.175'
k = 9	'SingleDrop_sfAPAP_NEO895_0.2'
k = 10	'SingleDrop_sfAPAP_NEO895_0.25'
k = 11	'SingleDrop_sfAPAP_NEO895_0.3'
k = 12	'SingleDrop_sfAPAP_LPG_0.078'
k = 13	'SingleDrop_sfAPAP_LPG_0.175'
k = 14	'SingleDrop_sfAPAP_LPG_0.25'
k = 15	'SingleDrop_MA_NEO895_0.089'
k = 16	'SingleDrop_MA_NEO895_0.2'
k = 17	'SingleDrop_MA_NEO895_0.3'
k = 18	'SingleDrop_MA_NEO895_0.4'
k = 19	'SingleDrop_MA_PEG300_0.041'
k = 20	'SingleDrop_PB_NEO895_0.084'
k = 21	'SingleDrop_PB_PEG300_0.043'
k = 22	'Dose10mg_uAPAP_NEO895_0.196'
k = 23	'Dose10mg_sfAPAP_NEO895_0.202'
k = 24	'Dose10mg_MA_PEG300_0.082'
k = 25	'Dose10mg_PB_PEG300_0.087'
k = 26	'Dose100mg_uAPAP_NEO895_0.175'
k = 27	'Dose100mg_uAPAP_NEO895_0.2'
k = 28	'Dose100mg_uAPAP_NEO895_0.223'
k = 29	'Dose100mg_uAPAP_LPG_0.175'
k = 30	'Dose100mg_sfAPAP_NEO895_0.175'
k = 31	'Dose100mg_sfAPAP_NEO895_0.2'
k = 32	'Dose100mg_sfAPAP_NEO895_0.25'
k = 33	'Dose100mg_sfAPAP_NEO895_0.294'
k = 34	'Dose100mg_sfAPAP_NEO895_0.3'
k = 35	'Dose100mg_sfAPAP_LPG_0.177'
k = 36	'Dose100mg_MA_NEO895_0.25'
k = 37	'Dose100mg_MA_NEO895_0.3'
k = 38	'Dose100mg_MA_NEO895_0.4'
k = 39	'Dose100mg_MA_NEO895_0.449'
k = 40	'Dose100mg_PB_NEO895_0.199'
k = 41	'Dose100mg_PB_NEO895_0.299'

APPENDIX J. CHAPTER 8: IMAGE PROCESSING

The vector of explanatory variables contributed by each image data object is therefore assembled as illustrated in Figure J.1.
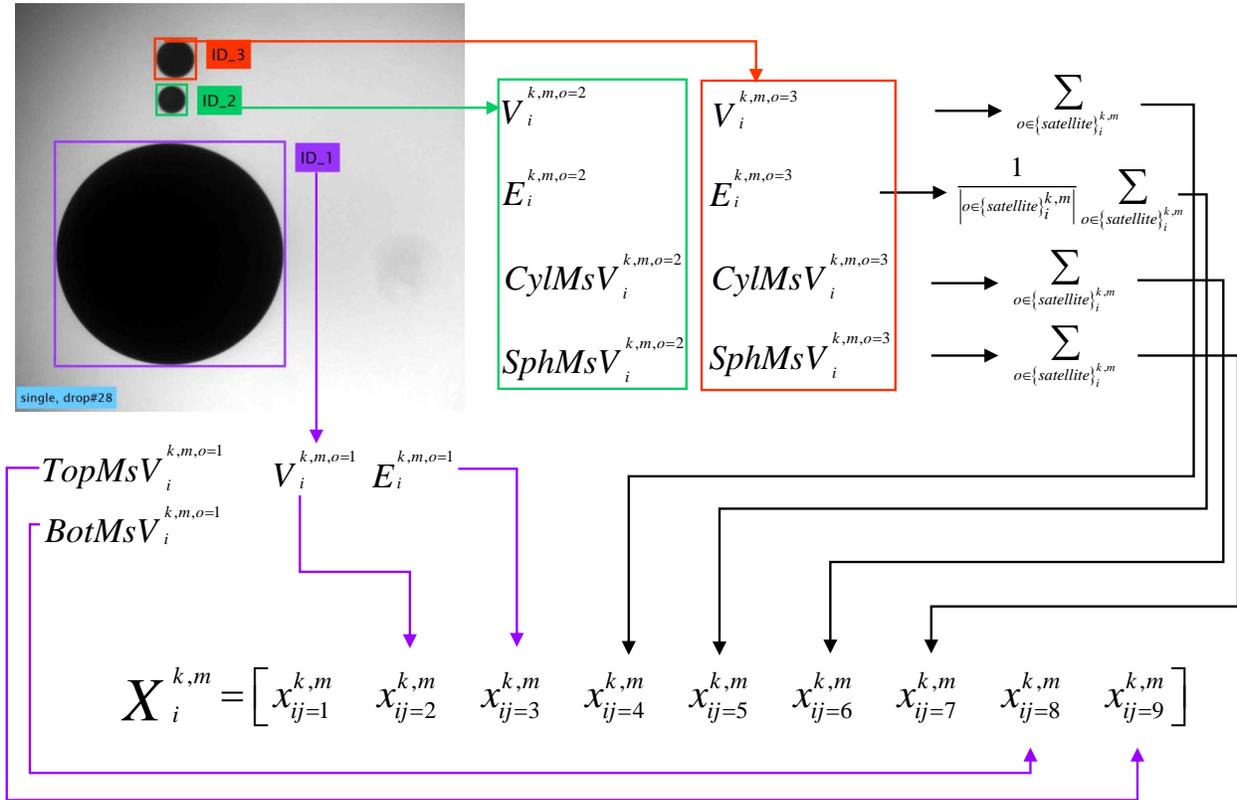

Figure J.1. Vector of explanatory variables for a single image

AUTOMATED IMAGE ANALYSIS

The image segmentation algorithm is described in detail in [16]; in the form used here, it detects objects by searching from the image boundaries for the minimum bounding rectangle around a continuous object, which is defined as an interval of contiguous rows or columns with non-zero elements in the direction being searched. The algorithm searches in the z-direction, finding the z-bounds of the minimum bounding rectangle, then searched in the x-direction on this interval for the x-bounds. The z-interval is then checked for the presence of discontinuities – any row in the edge map that which does not contain at least one non-zero element. If a discontinuity

is detected, the sub-image enclosed by the original bounding rectangle is subjected to the same image segmentation algorithm, but initiated separately from the two known starting points – the z-position of the original bounding rectangle. The searches proceeds to find the intervals of uninterrupted edge-containing rows by finding the first zero elements, the positions of which define the bounding rectangles for the two outermost objects in the original bounding rectangle.

The original z-interval may contain more than two objects, however. Therefore, the sub-image defined by the original x-bounds and the z-interval between the two outermost objects is subjected to the original algorithm – treating the sub-image in exactly the same manner as the original image itself was treated. This alternating procedure is repeated until each object is detected and its bounding rectangle is determined; the algorithm terminates for each object when both the z-direction and x-direction searches enacted on the proposed bounding rectangle both return true values, which correspond respectively to contiguous rows or columns occupied by edges.

Each bounding rectangle therefore corresponds to an object – a drop – from which quantitative information of interest may then be extracted, as described below. With respect to the z-axis of the object, the diameter, $D(z)$, was computed as the distance between the x-position of the right-most non-zero element and the x-position of the left-most non-zero element at each z-position; the positions are determined by iterative searches for the first non-zero element, initiated at the right/left limits of the respective bounding rectangle. These sets of positions are stored as $X_R(z)$, $X_L(z)$. Analogously to above, but on the x-axis, the length $L(x)$ was computed as the distance between the bottom-most and top-most non-zero elements at each x-position, determined by the analogous iterative searches for the non-zero elements initiated at the bottom/top limits of the bounding rectangle. This yields the set of points which is stored as $Z_B(x)$, $Z_T(x)$.

The centroid of the object is computed using the set of all points, $q_t = [z_t \quad x_t], \{q_t : t = 1, \ldots, T\}$, that constitute the perimeter of the object. This set is comprised of all the points $X_R(z)$, $X_L(z)$, and also the additional points which are present on the perimeter, but not included in these. The missing points are located by considering the geometry of the object in discrete coordinates as depicted in Figure J.2.
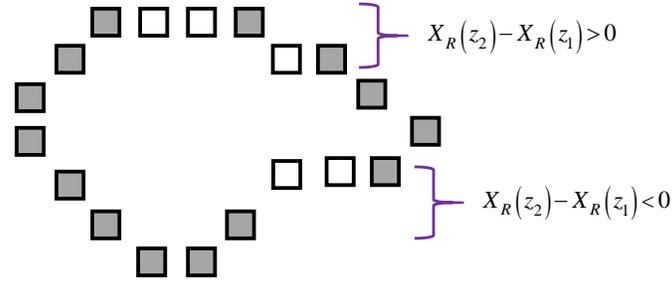
Figure J.2. Depiction of an edge map: each box represents an 'on' pixel in the edge map of the object. Gray-filled boxes correspond to the points captured in $X_R(z)$, $X_L(z)$.

At the top row, the need for additional points is determined by the checking the conditional: if $D(z_{top}) > 2$, then add points to the set using the following formula: $q_{Extra} = \begin{bmatrix} z_{top} & X_L(z_{top}) + j \end{bmatrix}, j = 1,...,(D(z_{top}) - 2)$; additional points for the bottom row are determined analogously. Elsewhere on the perimeter, the points are readily obtained by making use of the existing sets $X_R(z)$, $X_L(z)$ to check if the neighboring point in the perimeter is adjacent. This can be phrased as an iterative search through the set of points defined by $X_R(z)$ by computing $n = |X_R(z_1) - X_R(z_2)|$, and checking if the quantity is greater than 1 – if equal to zero, then the next point is directly below at 270°, and if equal to 1, then the next point is at 225° or 315°.

Thus, when greater than 1, this indicates that $n-1$ points must be added, as the next point in the set is not directly adjacent. The positions of these points depend on the sign of the quantity $X_R(z_1) - X_R(z_2)$: if positive, then the set is: $q_{Extra} = \begin{bmatrix} z_2 & X_R(z_2) - j \end{bmatrix}, j = 1,...,n-1$, and if negative, the set is $q_{Extra} = \begin{bmatrix} z_1 & X_R(z_2) + j \end{bmatrix}, j = 1,...,n-1$. For the left side, the analogous procedure is: $X_L(z_1) - X_L(z_2)$, which if positive, $q_{Extra} = \begin{bmatrix} z_2 & X_L(z_2) + j \end{bmatrix}, j = 1,...,n-1$, and if negative, $q_{Extra} = \begin{bmatrix} z_1 & X_L(z_2) - j \end{bmatrix}, j = 1,...,n-1$. Performing this for all z's on the interval respective to the object yields the perimeter points missing from $X_R(z)$, $X_L(z)$, and therefore, one is now able to compute the centroid, $\underline{c}$, of the object by solving the optimization problem:

$$\arg\min_{\underline{c}} \sum_{\forall t} \left(\left\| \underline{c} - \underline{q}_t \right\| \right)^2$$

Subsequently, the resultant vector for each object can be used to compute the Euclidean distance from the object centroid to the camera center-point.

From these more basic quantities, numerous derived quantities are obtained, representing categorical and quantitative information, as described in subsequent Section 8.4.5. In Figure J.3 a partial list of the extracted information is presented. Illustrations of the physical quantities not discussed in the preceding text are provided in Figure J.4 and Figure J.5.
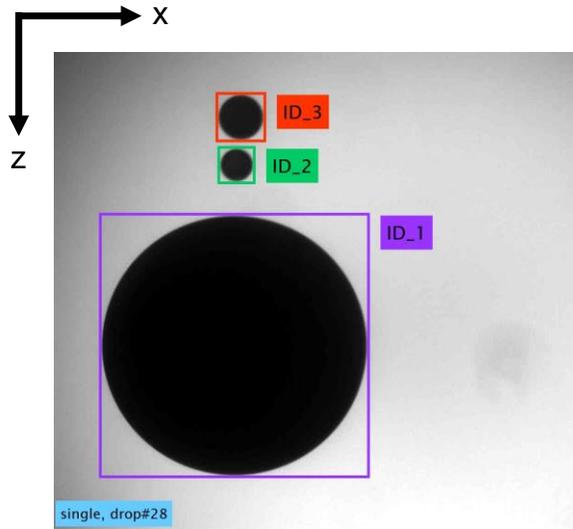

For each object, o, in each image, m

- $D(z)$: Diameter as function of z
  - X_L, X_R
  - D_{max} : maximum diameter of object along z-axis
  - $X_{L,D_{max}}$ & $X_{R,D_{max}}$: left and right x-positions at which D_{max} occurs
- $L(x)$: Diameter as function of x
  - Z_T, Z_B
  - L_{max} : maximum diameter of object along x-axis
  - $Z_{T,L_{max}}$ & $Z_{B,L_{max}}$: top and bottom z-positions at which L_{max} occurs
- minimum bounding rectangle
  - z_{top}, z_{bottom} : z-position of top & bottom of object
  - $X_{L,boundmin}, X_{R,boundmax}$: left & right bounds on object
- $q_t = [z_t \quad x_t], \{q_t : t = 1, \ldots, T\}$: the set of all points on the perimeter of the object
  - $[z_c, x_c]$: centroid of object
  - $\left\| [z_c, x_c] - [z_{\text{image center}}, x_{\text{image center}}] \right\|$: Euclidean distance from object centroid to image center

Figure J.3. Example image with minimum bounding rectangles displayed
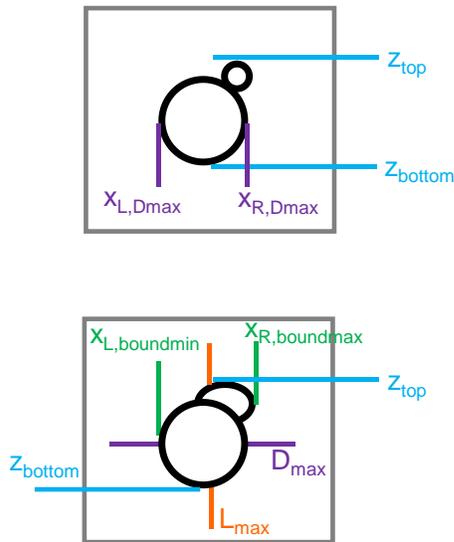
$$AR_{Natural_{DL}} = \frac{D_{max}}{\left(z_{bottom} - z_{top}\right)}$$

$$AR_{Inherent_{DL}} = D_{max} / L_{max}$$

$$AR_{BoundingRectangle_{DL}} = \frac{\left(X_{R,boundmax} - X_{L,boundmin}\right)}{\left(z_{bottom} - z_{top}\right)}$$

Figure J.4. Illustration of extracted information for variable names in Figure 2.
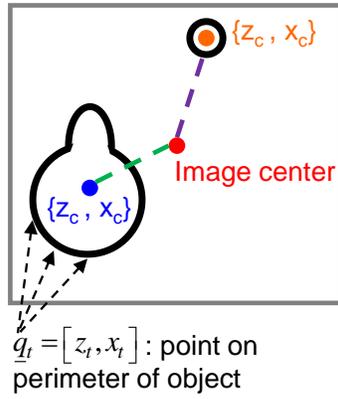
Figure J.5. Illustration of distance between object centroid and image center.

APPENDIX K. CHAPTER 8. MODELDROP VARIANTS

MODELDROP WITH SINGLE DATA VARIANCE

Posterior Predictive Model Check

Table K.1. Information Criteria for the models. Note: comparison of scoring functions is limited to the same modeling approach, i.e. when the data is identical.

	MODELDROP, SINGLE DATA VARIANCE
-2lpd	-4056.6
p_{DIC}	48.01
DIC	-3960.5
-2lppd	-4054.7
$p_{\text{WAIC 1}}$	45.15
$p_{\text{WAIC 2}}$	52.14
WAIC	-3950.4
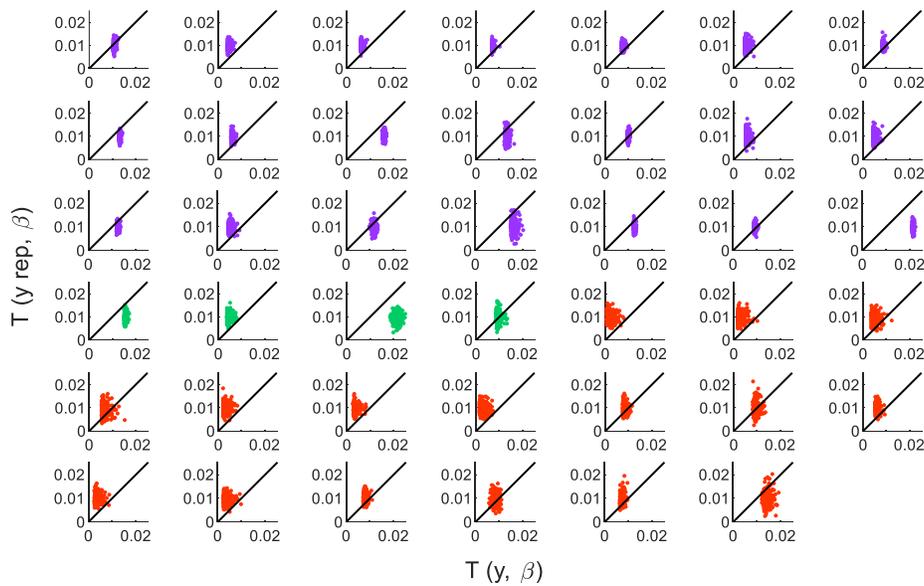
Figure K.1.: square root of batch-specific MSE^k, evaluated for all 41 batches; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

**Prediction for Out-of-Sample Data from In-Sample Conditions**
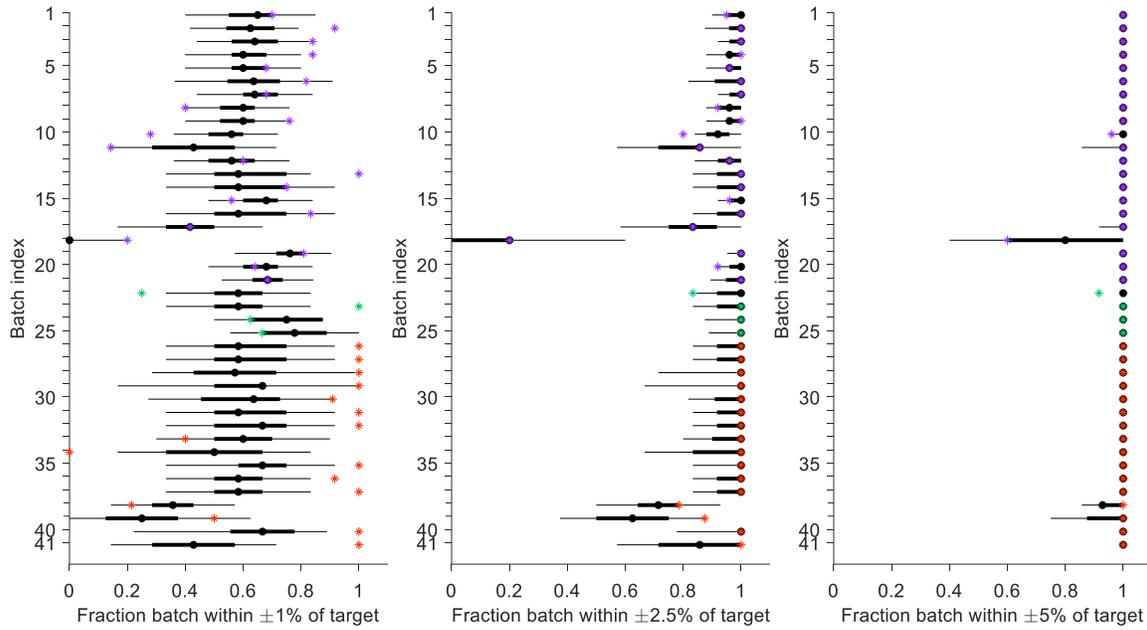
Figure K.2.: $F_{\tilde{y}^k}$, for all 41 batches of new data, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50$^{\text{th}}$ percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

Prediction for Out-of-Sample Data from Out-of-Sample Conditions
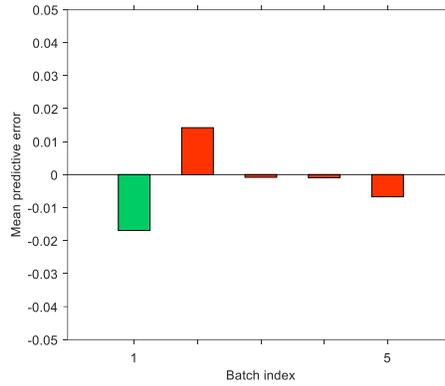
Figure K.3.: batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k / \bar{y}^{k^{new}}$, for the 5 batches of new data from new groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses. Batch Indices 1, 2, 3, 4 and 5 correspond to Hormones, Ibrutinib, Loratidine, Tacrolimus, and Warfarin, respectively.
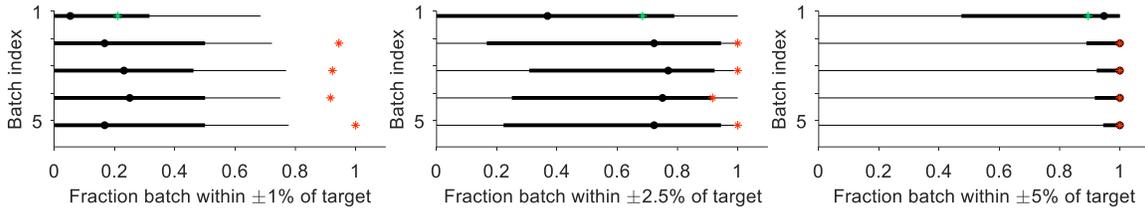
Figure K.4.: $F_{\bar{y}^k}$, for the 5 batches of new data from new groups, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\bar{y}^k}$, the black dot indicates the 50ᵗʰ percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

MODELDROP WITH 2 EXTRA EXPLANATORY VARIABLES

Posterior Predictive Model Check
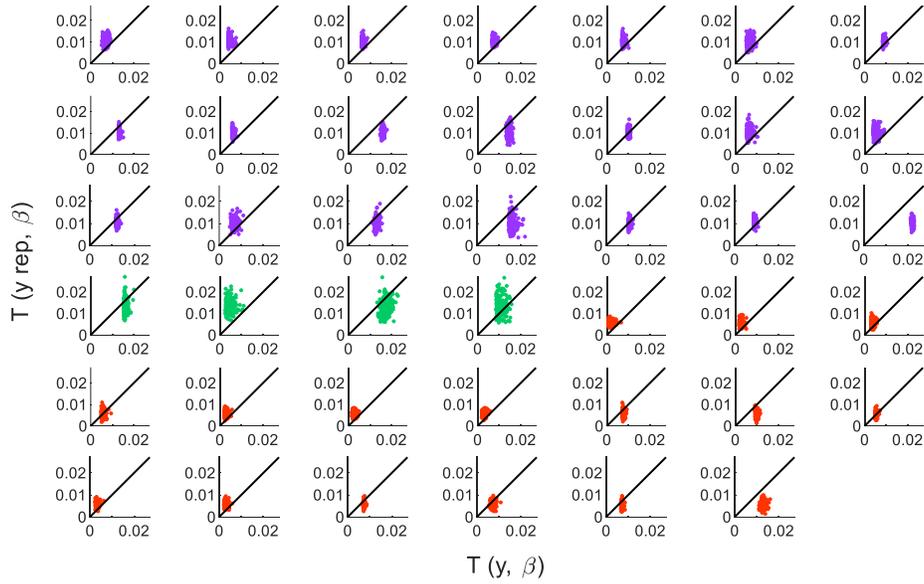
Figure K.5.: square root of batch-specific MSE^k, evaluated for all 41 batches; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

**Prediction for Out-of-Sample Data from In-Sample Conditions**
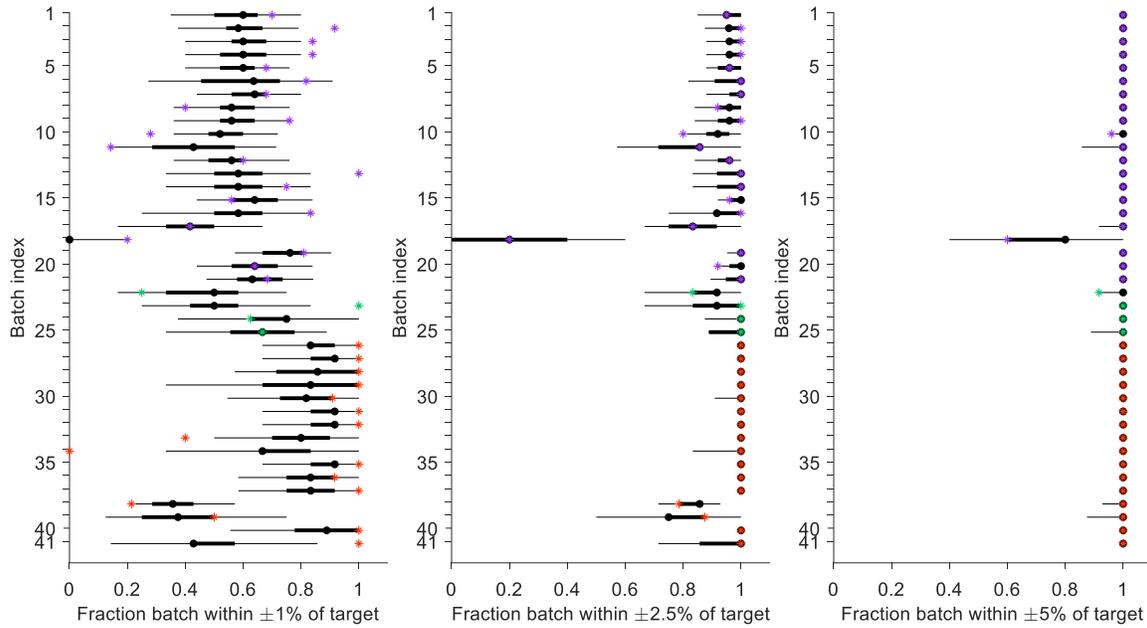
Figure K.6.: $F_{\tilde{y}^k}$, for all 41 batches of new data, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

Prediction for Out-of-Sample Data from Out-of-Sample Conditions
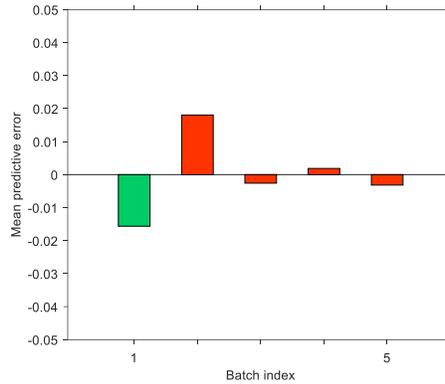
Figure K.7.: batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k / \bar{y}^{k^{new}}$, for the 5 batches of new data from new groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses. Batch Indices 1, 2, 3, 4 and 5 correspond to Hormones, Ibrutinib, Loratidine, Tacrolimus, and Warfarin, respectively.
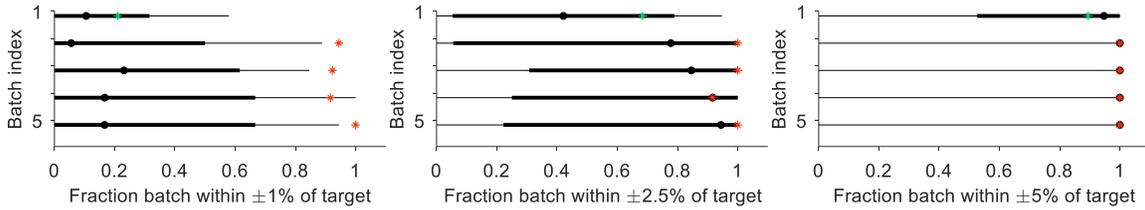

Figure K.8.: $F_{\tilde{y}^k}$, for the 5 batches of new data from new groups, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.

MODELDROP WITH 2 EXTRA EXPLANATORY VARIABLES

Posterior Predictive Model Check
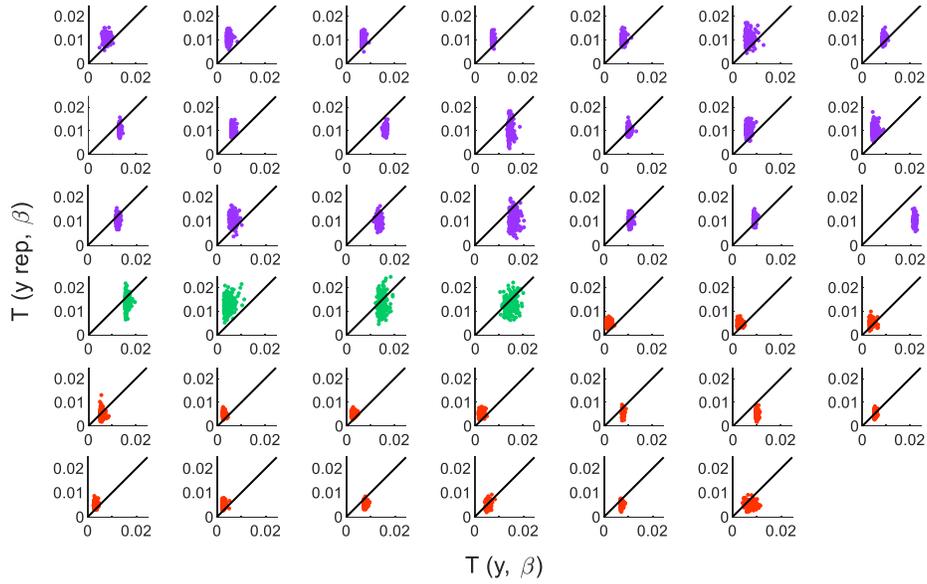
Figure K.9.: square root of batch-specific MSE^k, evaluated for all 41 batches; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses.

**Prediction for Out-of-Sample Data from In-Sample Conditions**
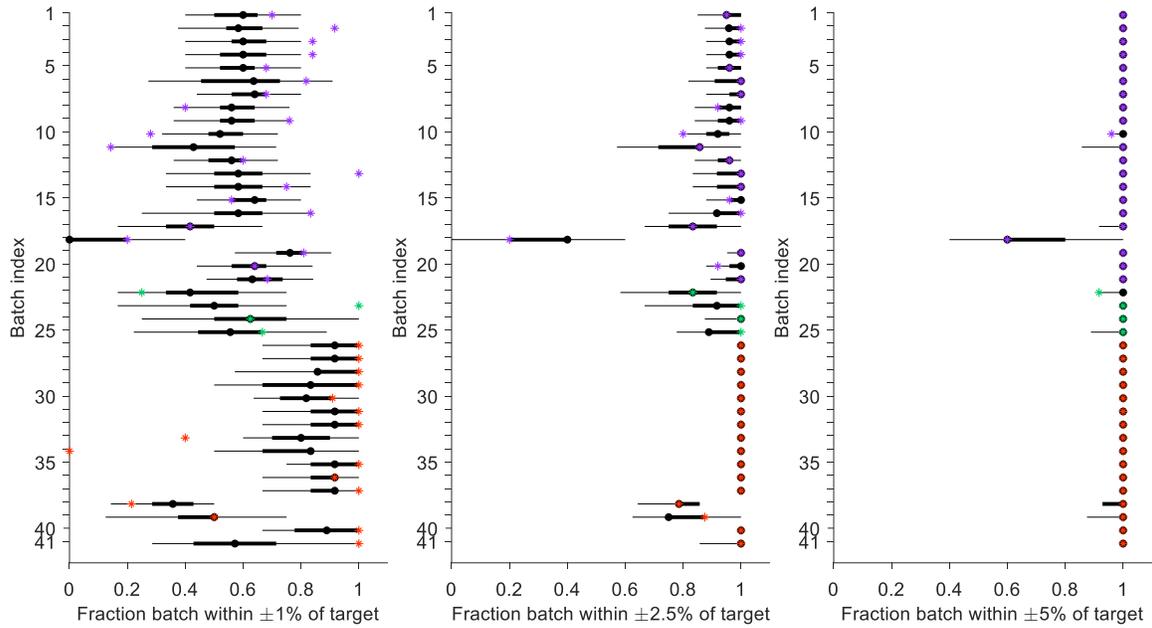
Figure K.10.: $F_{\tilde{y}^k}$, for all 41 batches of new data, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50$^{\text{th}}$ percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements y_k^{new}.

**Prediction for Out-of-Sample Data from Out-of-Sample Conditions**
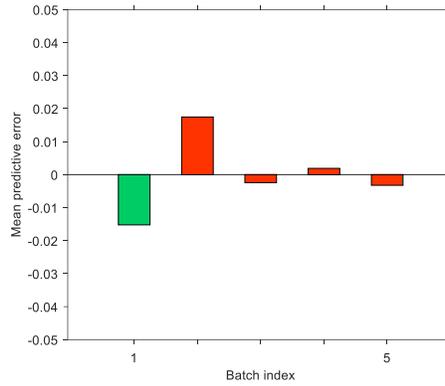
Figure K.11.: batch-wide mean predictive error normalized to the mean respective batch mean, $\bar{e}^k / \bar{y}^{k^{new}}$, for the 5 batches of new data from new groups; in this figure, color corresponds to δ-level: purple, green, and red to single-, few- and many-drop doses. Batch Indices 1, 2, 3, 4 and 5 correspond to Hormones, Ibrutinib, Loratidine, Tacrolimus, and Warfarin, respectively.
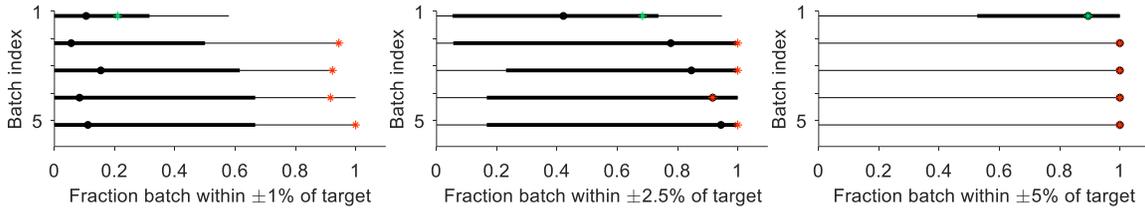

Figure K.12.: $F_{\tilde{y}^k}$, for the 5 batches of new data from new groups, computed at $z = 1$, 2.5, and 5%, from left to right. The horizontal lines indicate the 95% (thin) and 50% (thick) intervals for $F_{\tilde{y}^k}$, the black dot indicates the 50th percentile, and the color-coded asterisks indicate the "true" fraction of doses in the batch that pass based on the measurements $y^{k^{new}}$.